

HYBRID BEAMFORMING DESIGN FOR MILLIMETER WAVE  
MASSIVE MIMO COMMUNICATIONS

GUILHERME MARTIGNAGO ZILLI

A THESIS  
IN  
THE DEPARTMENT  
OF  
ELECTRICAL AND COMPUTER ENGINEERING

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY (PH.D. IN ELECTRICAL AND COMPUTER ENGINEERING) AT  
CONCORDIA UNIVERSITY  
MONTRÉAL, QUÉBEC, CANADA

April 2021

© Guilherme Martignago Zilli, 2021

CONCORDIA UNIVERSITY  
School of Graduate Studies

This is to certify that the thesis prepared

By: **Guilherme Martignago Zilli**  
Entitled: **Hybrid Beamforming Design for Millimeter Wave Massive MIMO  
Communications**

and submitted in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy (Ph.D. in Electrical and Computer Engineering)**

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

\_\_\_\_\_ Chair  
*Dr. Chun Wang*

\_\_\_\_\_ External Examiner  
*Dr. Yu-Dong Yao*

\_\_\_\_\_ External to Program  
*Dr. Arash Mohammadi*

\_\_\_\_\_ Examiner  
*Dr. Dongyu Qiu*

\_\_\_\_\_ Examiner  
*Dr. Yousef Shayan*

\_\_\_\_\_ Thesis Supervisor  
*Dr. Wei-Ping Zhu*

Approved by \_\_\_\_\_  
Dr. Wei-Ping Zhu  
Graduate Program Director

April 22, 2021 \_\_\_\_\_  
Dr. Mourad Debbabi, Dean  
Gina Cody School of Engineering and Computer Science

# Abstract

Hybrid Beamforming Design for Millimeter Wave Massive MIMO Communications

**Guilherme Martignago Zilli, Ph.D.**

**Concordia University, 2021**

Wireless connectivity is a key driver for the digital transformation that is changing how people communicate, do business, consume entertainment and search for information. As the world advances into the fifth generation (5G) and beyond-5G (B5G) of wireless mobile technology, new services and use cases are emerging every day, bringing the demand to expand the broadband capability of mobile networks and to provide ubiquitous access and specific capabilities for any device or application. To fulfill these demands, 5G and B5G systems will rely on innovative technologies, such as the ultradensification, the mmWave, and the massive MIMO. To bring together these technologies, 5G and B5G systems will employ hybrid analog-digital beamforming, which separates the signal processing into the baseband (digital) and the radio-frequency (analog) domains. Unlike conventional beamforming, where every antenna is connected to an RF chain, and the signal is entirely processed in the digital domain, hybrid beamforming uses fewer RF chains than the total number of antennas, resulting in a less expensive and less energy-consuming design. The analog beamforming is usually implemented using switching networks or phase-shifting networks, which impose severe hardware constraints making the hybrid beamforming design very challenging. This thesis addresses the hybrid analog-digital beamforming design and is organized into three parts.

In the first part, two adaptive algorithms for solving the switching-network-based hybrid beamforming design problem, also known as the joint antenna selection and beamforming (JASB) problem, are proposed. The adaptive algorithms are based on the minimum mean square error (MMSE) and minimum-variance distortionless response (MVDR) criteria and employ an alternating optimization strategy, in which the beamforming and the antenna selection are designed iteratively. The proposed algorithms can attain high levels of SINR while strictly complying with the hardware limitations. Moreover, the proposed algorithms have very low computational complexity and can track channel variations, making them suitable for non-stationary environments. Numerical simulations have validated the effectiveness of the algorithms in different operation scenarios.

The second part addresses the phase-shifting-network-based hybrid beamforming design for narrowband mmWave massive MIMO systems. A novel joint hybrid precoder and combiner design is proposed. The analog precoder and combiner design is formulated as constrained low-rank channel decomposition, which can simultaneously harvest the array gain provided by the massive MIMO

system and suppress intra-user and inter-user interferences. The constrained low-rank channel decomposition is solved as a series of successive rank-1 channel decomposition, using the projected block coordinate descent method. The digital precoder and combiner are obtained from the optimal SVD-based solution for the single-user case and the regularized channel diagonalization method for the multi-user case. Simulation results have demonstrated that the proposed design can consistently attain near-optimal performance and provided important insights into the method's convergence and its performance under practical phase-shifter quantization constraints.

Finally, the phase-shifting-network-based hybrid beamforming design for frequency-selective mmWave massive MIMO-OFDM systems is considered in the third part. The hybrid beamforming design for MIMO-OFDM systems is significantly more challenging than for narrowband MIMO systems since, in these systems, the analog precoder and combiner are shared among all subcarriers and must be jointly optimized. Thus, by leveraging the OFDM systems' multidimensional structure, the analog precoder and combiner design is formulated as constrained low-rank Tucker2 tensor decomposition and solved by a successive rank-(1,1) tensor decomposition using the projected alternate least square (ALS) method. The digital precoder and combiner are obtained on a per-subcarrier basis using the techniques presented in the second part. Numerical simulations have confirmed the design effectiveness, demonstrating its ability to consistently attain near-optimal performance and outperform other existing design in nearly all scenarios. They also provided insights into the convergence of the proposed method and its performance under practical phase-shifter quantization constraints and highlighted the differences between this design and that for the narrowband massive MIMO systems.

# Acknowledgments

First and foremost, I thank the Fonds de Recherche du Québec – Nature et Technologies (FRQNT), the Gina Cody School of Engineering and Computer Science and Faculty Research Support (FRS) Fund, and the Concordia Accelerator Award donors for their generous financial support.

I would like to express my genuine appreciation and gratitude to my supervisor, Dr. Wei-Ping Zhu, who has provided valuable support, guidance, and encouragement throughout my Ph.D. at Concordia University. His great advice, constructive feedback, and careful editing contributed enormously to the writing of this thesis and the publications we completed together.

I also would like to express my appreciation to my thesis supervisory committee members, Dr. Yousef R. Shayan, Dr. Dongyu Qiu, and Dr. Arash Mohammadi, for taking their time to carefully read my comprehensive examination, doctoral research proposal, and Ph.D. seminar reports and this thesis, and for providing valuable comments on my research. I also thank the external examiner, Dr. Yu-Dong Yao, from Stevens Institute of Technology, for his careful review of my thesis and his participation in my defence.

I thank Dr. Ciro André Pitz, Dr. Eduardo L. O. Batista, and Dr. Rui Seara for collaborating with me in the work developed in the first chapter of this thesis. I also thank Dr. Seara for providing access to the high-performance computing facility at his laboratory. Thanks are also due to Dr. Pitz and Dr. Eduardo Kuhn for giving relevant feedback to some of the manuscripts written as a result of this research.

I gratefully acknowledge my colleagues at Concordia University for the constant exchange of ideas and support. In particular, I would like to thank Mojtaba Hasannezhad, Ali Mohebbi, Kalpesh Ranipa, Hongjiang Yu, and Zhiheng Ouyang. I extend my acknowledgment to all Concordia University staff members, especially to Ms. Sheryl Tablan, the Electrical and Computer Engineering Department staff, the GradProSkills staff, and the librarians.

I would like to express my special gratitude to Bruno Alves, for his unconditional support and for providing valuable help in designing some of the figures present in this thesis.

Finally, I would like to extend my sincere thanks to my parents for their unwavering support and encouragement throughout my studies.

# Contents

<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xiv</b>
<b>List of Acronyms</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Looking Forward to 5G and Beyond . . . . .	1
1.2 Hybrid Beamforming: putting mmWave and Massive MIMO together . . . . .	5
1.3 Objectives and Contributions . . . . .	8
1.4 Related Publications . . . . .	11
1.5 Organization . . . . .	12
<b>2 Adaptive Joint Antenna Selection and Beamforming</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.2 System Model and Problem Statement . . . . .	15
2.2.1 Scenario Definition . . . . .	15
2.2.2 Joint Antenna Selection and Beamforming Problem . . . . .	18
2.3 MMSE-based Adaptive Joint Antenna Selection and Beamforming . . . . .	19
2.3.1 MMSE Beamforming . . . . .	19
2.3.2 MMSE Antenna Selection . . . . .	20
2.3.3 MMSE-based AJASB Algorithm . . . . .	23
2.3.4 Simulation Results . . . . .	24
2.4 MVDR-based Adaptive Joint Antenna Selection and Beamforming . . . . .	30
2.4.1 MVDR Beamforming . . . . .	30
2.4.2 MVDR Antenna Selection . . . . .	31
2.4.3 MVDR-based AJASB Algorithm . . . . .	34
2.4.4 Simulation Results . . . . .	35
2.5 Concluding Remarks . . . . .	40

<b>3</b>	<b>Hybrid Beamforming for Narrowband mmWave Massive MIMO Systems</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	System Model and Problem Formulation . . . . .	44
3.2.1	System Model . . . . .	44
3.2.2	Channel Model . . . . .	47
3.2.3	Problem Formulation . . . . .	48
3.3	Constrained Channel Decomposition-based Hybrid Beamforming Design for MIMO Systems . . . . .	49
3.3.1	Analog Precoder and Combiner Design . . . . .	50
3.3.2	Digital Precoder and Combiner Design for Single-User Systems . . . . .	54
3.3.3	Digital Precoder and Combiner Design for Multi-User Systems . . . . .	55
3.3.4	Computational complexity . . . . .	58
3.4	Simulation Results . . . . .	59
3.4.1	Single-User Systems . . . . .	60
3.4.2	Multi-User Systems . . . . .	68
3.4.3	Convergence Analysis . . . . .	80
3.4.4	Further Discussions . . . . .	82
3.5	Concluding Remarks . . . . .	83
<b>4</b>	<b>Hybrid Beamforming for OFDM mmWave Massive MIMO Systems</b>	<b>84</b>
4.1	Introduction . . . . .	84
4.2	System Model and Problem Formulation . . . . .	87
4.2.1	System Model . . . . .	87
4.2.2	OFDM Channel Model . . . . .	89
4.2.3	Problem Formulation . . . . .	90
4.3	Constrained Channel Tensor Decomposition-based Hybrid Beamforming Design for MIMO-OFDM Systems . . . . .	91
4.3.1	Analog Precoder and Combiner Design . . . . .	91
4.3.2	Digital Precoder and Combiner Design for Single-User Systems . . . . .	97
4.3.3	Digital Precoder and Combiner Design for Multi-User Systems . . . . .	97
4.3.4	Computational complexity . . . . .	98
4.4	Simulation Results . . . . .	99
4.4.1	Single-User Systems . . . . .	100
4.4.2	Multi-User Systems . . . . .	104
4.4.3	Convergence Analysis . . . . .	110
4.4.4	Effects of Power Method Approximation . . . . .	111
4.4.5	Comparison with the Narrowband Design in Chapter 3 . . . . .	113
4.4.6	Further Discussions . . . . .	115

4.5	Concluding Remarks . . . . .	116
<b>5</b>	<b>Conclusions</b>	<b>117</b>
5.1	Summary of the Contributions . . . . .	117
5.2	Future Work . . . . .	119
	<b>Bibliography</b>	<b>120</b>
	<b>Appendices</b>	<b>131</b>
<b>A</b>	<b>Obtaining (3.13) from (3.12)</b>	<b>132</b>
<b>B</b>	<b>Proof of the optimality of (3.20)</b>	<b>134</b>
<b>C</b>	<b>Tensors Preliminaries</b>	<b>137</b>
C.1	Tensor Notations and Operations . . . . .	137
C.1.1	Fibers . . . . .	137
C.1.2	Slices . . . . .	138
C.1.3	Inner Product and Norm . . . . .	138
C.1.4	Matricization . . . . .	139
C.1.5	Tensor Multiplication . . . . .	140
C.1.6	Other Useful Operations . . . . .	140
C.2	The Tucker Decomposition . . . . .	141
C.2.1	The $n$ -Rank and the Truncated Tucker Decomposition . . . . .	142
C.2.2	Tucker Decomposition Computation . . . . .	143

# List of Figures

1.1	Conventional (fully-digital) MIMO system. . . . .	6
1.2	Switching-network-based hybrid combiner. . . . .	7
1.3	Phase-shifter-network-based hybrid precoder (left) and combiner (right). . . . .	8
2.1	Multi-cell system configuration with the signal of interest (blue line) and the interfering signals (red lines) arriving at the considered BS [62]. . . . .	16
2.2	Example 1. SINR curves of the MMSE-based algorithms for $M = 64$ and $L = 4$ . . .	26
2.3	Example 1. EPDF curves for the MMSE-based algorithms with $M = 64$ and $L = 4$ . . .	26
2.4	Example 1. SINR curves of the MMSE-based algorithms for $M = 64$ and $L = 8$ . . .	27
2.5	Example 2. SINR curves of the MMSE-based algorithms for $M = 64$ and $L = 8$ . . .	28
2.6	Example 3. Steady-state SINR curves for different numbers of RF chains using the MMSE-based algorithms. . . . .	29
2.7	Example 4. SINR curves for the MMSE-based algorithms in the non-stationary scenario. . .	29
2.8	Example 1. SINR curves of the MVDR-based algorithms for $M = 64$ and $L = 4$ . . .	36
2.9	Example 1. EPDF curves for the MVDR-based algorithms with $M = 64$ and $L = 4$ . . .	36
2.10	Example 1. SINR curves of the MVDR-based algorithms for $M = 64$ and $L = 8$ . . .	37
2.11	Example 2. SINR curves of the MVDR-based algorithms for $M = 64$ and $L = 8$ . . .	38
2.12	Example 3. Steady-state SINR curves for different numbers of RF chains using the MVDR-based algorithms. . . . .	39
2.13	Example 4. SINR curves for the MVDR-based algorithms in the non-stationary scenario. . .	39
3.1	System diagram for the downlink of a multiuser massive MIMO system with hybrid beamforming. . . . .	45
3.2	Example 1. SU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with $N_t = N_r = 64$ antennas and $N_s = 4$ data streams. . . . .	61
3.3	Example 2. SU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = N_r = 64$ antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	62

3.4	Example 3. SU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas $N_t = N_r$ for a system communicating through $N_s = 4$ data streams at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	63
3.5	Example 4. SU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ , and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	64
3.6	Example 1. SU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with $N_t = N_r = 64$ antennas and $N_s = 4$ data streams. . . . .	65
3.7	Example 2. SU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = N_r = 64$ antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	66
3.8	Example 3. SU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas $N_t = N_r$ for a system communicating through $N_s = 4$ data streams at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	67
3.9	Example 4. SU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	68
3.10	Example 1. MU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with $N_t = 144$ transmit antennas and $K = 8$ users, each having $N_r = 16$ receive antennas and communicating through $N_s = 4$ data streams. . . . .	69
3.11	Example 2. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = 144$ transmit antennas and $K = 4$ users having $N_r = 16$ receive antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	70
3.12	Example 3. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas $N_t$ for a system with $K = 4$ users having $N_r = 16$ receive antennas and communicating through $N_s = 4$ data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	72
3.13	Example 4. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of users $K$ for a system with $N_t = 144$ transmit antennas and users having $N_r = 16$ and communicating through $N_s = 4$ data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	73

3.14	Example 5. MU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	74
3.15	Example 1. MU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with $N_t = 144$ transmit antennas and $K = 8$ users, each having $N_r = 16$ receive antennas and communicating through $N_s = 4$ data streams. . . . .	74
3.16	Example 2. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = 144$ transmit antennas and $K = 4$ users having $N_r = 16$ receive antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	76
3.17	Example 3. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas $N_t$ for a system with $K = 4$ users having $N_r = 16$ receive antennas and communicating through $N_s = 4$ data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	77
3.18	Example 4. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of users $K$ for a system with $N_t = 144$ transmit antennas and users having $N_r = 16$ and communicating through $N_s = 4$ data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	78
3.19	Example 5. MU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	79
3.20	Convergence analysis for SU-MIMO systems. Empirical probability mass function (PMF) and empirical cumulative density function (CDF) of the number of iterations required for convergence of the proposed method considering a $64 \times 64$ MIMO systems with $N_s = 8$ in Rayleigh (a) and mmWave (b) channels. . . . .	80
3.21	Convergence analysis for MU-MIMO systems. Empirical probability mass function (PMF) and empirical cumulative density function (CDF) of the number of iterations required for convergence of the proposed method considering a system with $N_t = 144$ transmit antennas and $K = 8$ users, each having $N_r = 16$ receive antennas and communicating through $N_s = 8$ data streams in Rayleigh (a) and mmWave (b) channels. . . . .	81
4.1	Downlink of an MU-MIMO-OFDM system with hybrid beamforming architecture. . . . .	87
4.2	Example 1. SU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with $N_t = N_r = 64$ antennas, $N_s = 4$ data streams, and $M = 512$ subcarriers. . . . .	100

4.3	Example 2. SU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = N_r = 64$ antennas and $M = 512$ subcarriers at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . .	102
4.4	Example 3. SU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas $N_t = N_r$ for a system communicating through $N_s = 4$ data streams and $M = 512$ subcarriers at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	103
4.5	Example 4. SU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	104
4.6	Example 1. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. SNR for a system $N_t = 144$ transmit antennas and $K = 4$ users, each having $N_r = 16$ receive antennas, communicating through $N_s = 4$ data streams and $M = 32$ subcarriers. . . . .	105
4.7	Example 2. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams $N_s$ for a system with $N_t = 144$ transmit antennas and $K = 4$ users, each having $N_r = 16$ receive antennas, communicating through $M = 32$ subcarriers at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	107
4.8	Example 3. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas $N_t$ for a system with $K = 4$ users having $N_r = 9$ receive antennas and communicating through $N_s = 4$ data streams and $M = 32$ subcarriers at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). . . . .	108
4.9	Example 4. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of users $K$ for a systems with $N_t = 144$ transmit antennas and users having $N_r = 16$ receive antennas, communicating through $N_s = 4$ data streams and $M = 32$ subcarriers at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). . . .	109
4.10	Example 5. MU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits $N_b$ and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d). . . . .	110
4.11	Empirical PMF and empirical CDF of the number of iterations needed for computing each analog beamforming vector pair, considering a SU-MIMO-OFDM system with $N_t = N_r = 64$ , $N_s = 8$ , and $M = 512$ . . . . .	111
4.12	Empirical PMF and empirical CDF of the number of iterations needed for computing each analog beamforming vector pair, considering an MU-MIMO-OFDM system with $N_t = 144$ , $N_r = 16$ , $K = 8$ , and $N_s = 4$ , and $M = 32$ . . . . .	112

4.13	Effects of the power method approximation. Spectral efficiency (a) and sum-rate (b) vs. the SNR for an MU-MIMO-OFDM system with $N_t = 144$ transmit antennas and $K = 4$ users, each having $N_r = 16$ receive antennas, communicating through $N_s = 4$ data streams in $M = 32$ subcarriers. . . . .	113
4.14	Effects of the power method approximation. Empirical CDF of the number of iterations required for computing each analog beamforming vector pair, considering an MU-MIMO system with $N_t = 144$ , $N_r = 16$ , $K = 4$ , and $N_s = 4$ . . . . .	113
4.15	Comparison with the narrowband design in Chapter 3. Spectral efficiency (a) and sum-rate (b) vs. the SNR for an MU-MIMO system with $N_t = 144$ transmit antennas and $K = 4$ users, with each user having $N_r = 16$ receive antennas and communicating through $N_s = 4$ data streams. . . . .	114
4.16	Comparison with the narrowband design in Chapter 3. Empirical CDF of the number of iterations required for computing each analog beamforming vector pair, considering an MU-MIMO system with $N_t = 144$ , $N_r = 16$ , $K = 4$ , and $N_s = 4$ . . . . .	114
C.1	Fibers of a 3rd-order tensor [135]. Reproduced with permission from SIAM (Society for Industrial and Applied Mathematics). . . . .	138
C.2	Slices of a 3rd-order tensor [135]. Reproduced with permission from SIAM. . . . .	138
C.3	Tucker decomposition of a 3rd-order tensor [135]. Reproduced with permission from SIAM. . . . .	142
C.4	Truncated Tucker decomposition of a 3rd-order tensor [135]. Reproduced with permission from SIAM. . . . .	143

# List of Tables

2.1	Example 2. Gain (in dBi) at the direction of the SOI and the interferers for the MMSE-based algorithms. . . . .	28
2.2	Example 2. Gain (in dBi) at the direction of the SOI and the interferers for the MVDR-based algorithms. . . . .	38
3.1	Computational Complexity of SU-MIMO Hybrid Beamforming Design Methods . . .	59
3.2	Computational Complexity of MU-MIMO Hybrid Beamforming Design Methods . .	59
4.1	Computational Complexity Comparison for Single-User OFDM Hybrid Beamforming Design Methods . . . . .	99
4.2	Computational Complexity Comparison for Multi-User OFDM Hybrid Beamforming Design Methods . . . . .	99

# List of Acronyms

**4G** fourth-generation

**5G** fifth-generation

**ADC** analog-to-digital converter

**AJASB** adaptive joint antenna selection and beamforming

**ALS** alternate least square

**AltMin** alternating-minimization

**AR** augmented reality

**AS** antenna selection

**B5G** beyond 5G

**BD** block-diagonalization

**BS** base station

**CNLMS** constrained normalized least mean square

**CPA** cyclic prefix addition

**CPR** cyclic prefix removal

**DAC** digital-to-analog converter

**DFT** discrete Fourier transform

**DL** downlink

**EGT** equal gain transmission

**eMBB** enhanced mobile broadband

**FFT** fast Fourier transform

**FWA** fixed wireless network

**HetNet** heterogeneous network

**HRCDD** hybrid regularized channel diagonalization

**i.i.d.** independent and identically distributed

**ICI** inter-carrier interference

**IFFT** inverse fast Fourier transform

**IoT** internet of things

**JASB** joint antenna selection and beamforming

**LNA** low-noise amplifier

**MIMO** multiple-input-multiple-output

**MMSE** minimum mean square error

**mMTC** massive machine type communications

**mmWave** millimeter wave

**MO** manifold optimization

**MRC** maximum ratio combining

**MRT** maximum ratio transmission

**MSE** mean square error

**MU** multi-user

**MUI** multi-user interference

**MVDR** minimum-variance distortionless response

**NFV** network function virtualization

**NLMS** normalized least mean square

**OFDM** orthogonal frequency-division multiplexing

**OMP** orthogonal matching pursuit

**PA** power amplifier

**PE** phase-extraction

**PS** phase-shifter

**QoS** quality-of-service

**RCD** regularized channel diagonalization

**RF** radio frequency

**RZF** regularized zero-forcing

**SDN** software defined network

**SDP** semidefinite programming

**SINR** signal-to-interference-plus-noise ratio

**SOI** signal of interest

**SU** single-user

**SVD** singular value decomposition

**UL** uplink

**ULA** uniform linear array

**URLLC** ultra-reliable low latency communications

**VR** virtual reality

**WMMSE** weighted minimum mean square error

**ZF** zero-forcing

# Chapter 1

## Introduction

### 1.1 Looking Forward to 5G and Beyond

Over the past decade, the world has witnessed an accelerated digital transformation that has impacted how people communicate, do business, consume entertainment and search for information. A crucial component of this digital transformation is wireless connectivity. As societies advance into the fifth generation of wireless mobile technology (5G) and beyond-5G (B5G), industry and researchers need to evolve the concept of wireless connectivity to expand the broadband capability of mobile networks and to provide ubiquitous access and specific capabilities for any device or application, not only for consumers but also for various industries and societies as a whole [1], [2].

The potential use cases for 5G and B5G systems spread across different industries. By compiling the prospects from recent reports from Ericsson, Nokia, McKinsey, and STL Partners [1]–[7], some examples of use cases, classified according to industry sectors, include:

- **Agriculture:** low-power sensor networks and IoT devices will enable smart crop monitoring to integrate irrigation and nutrient-distribution for real-time optimized crop growth and resource usage, and smart-livestock monitoring to provide individualized feeding-and-care plans aimed at detecting illnesses and providing optimal feed and medicine mix to maximize each animal's growth; low-latency and reliable connectivity will enable autonomous-farming machinery to perform interventions that optimize the resource usage and reduce the labour requirements.
- **Consumer:** fixed wireless networks (FWA) will provide a faster and more reliable alternative

to wired broadband in households; virtual and augmented reality (VR/AR) will enable immersive experiences for education and gaming; enhanced video streaming will allow interruption-free video calls and media streaming anytime and anywhere, with the ability to stream from drones, bikes and other non-wired locations; smart venues will ensure connectivity and reliability in crowded events, such as sports tournaments, concerts, festivals, and allowing high-quality video uploading and a blend of physical and virtual experiences.

- **Healthcare:** remote patient monitoring will enable 24/7 patient monitoring and support through e-health and wearable devices; virtual consultations and remote surgeries will increase the access of patients to healthcare; connected ambulances will stream patient data and information from sensors and high-definition cameras between paramedics and in-hospital emergency departments.
- **Industry:** fast and reliable communication between sensors and machinery will allow manufacturing industries to monitor the status of machines in real-time and to perform predictive and preventative maintenance; wireless connectivity will reduce the requirement for wired systems, allowing a more reconfigurable and flexible production lines; VR/AR will speed up the worker's training and guide worker or remote expert when carrying out maintenance.
- **Mobility:** high-speed, low-latency, and reliable connectivity will enable autonomous vehicle control to increase safety and allow drivers to use their commuting time profitably for leisure or business activities; intelligent transportation systems will facilitate traffic management and dynamic traffic rerouting; vehicle-to-vehicle and vehicle-to-infrastructure communications will increase driving efficiency and safety; other in-vehicle features will enable precise navigation, entertainment streaming, and alike.

Nevertheless, there will be other use cases in mining, energy, smart cities, public safety and surveillance sectors. Despite the plethora of use cases, at this stage, it is impossible to provide an exhaustive list of applications since many others will likely emerge. Furthermore, 5G and B5G use cases have utterly different requirements regarding availability, connection density, energy efficiency, latency, reliability, and throughput. The use cases are, thus, classified according to these requirements in three essential types of communication [8]: enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive machine type communication (mMTC).

The eMBB provides extremely high data-rates and low latency with high coverage and addresses human-centric use cases, such as multimedia content delivery and mobile telephony [1]. It will also address use cases that require a large coverage area and hotspots. Large coverage area cases

require seamless coverage and high mobility, while hotspots (i.e., areas with high user density) need very high traffic capacity and low-mobility [8]. Some of the requirements for eMBB communication include peak data-rates of 20 Gbit/s and 10 Gbit/s for downlink (DL) and uplink (UL), respectively; peak spectral efficiency of 30 bit/s/Hz (DL) and 15 bit/s/Hz (UL); user experienced data rate of 100 Mbit/s (DL) and 50 Mbit/s (UL) in a dense urban environment; area traffic capacity of 10 Mbit/s/m<sup>2</sup> in indoor hotspots; latency of 4ms; and mobility support to vehicular (up to 30 km/h) and high speed vehicular (above 120 km/h) in dense urban and rural environments, respectively [9].

The URLLC and mMTC, on the other hand, address machine-centric use cases. The URLLC focuses on use cases with stringent requirements on throughput, latency and reliability, such as autonomous vehicles, remote surgery, and real-time control of industrial manufacturing and processes. Its requirements include a 0.5 ms latency (for both UL and DL) and reliability of  $1 - 10^{-5}$  for a 32-bytes packet transmission with a latency of 1 ms [10]. The mMTC provides connectivity to a massive number of low-complexity narrowband devices with infrequent and small data transmission, such as low-cost sensors, meters, trackers, and wearables [2]. The main requirements for mMTC are to support 1 million connected devices per square kilometre and devices with battery life beyond ten years [9], [10].

These requirements focus on four dimensions, which will drive innovations in 5G and B5G system [11]: (i) Spectrum allocation: to make more spectrum available, 5G and B5G systems will adopt frequencies above 3 GHz – one can expect  $2\times$  more spectrum at below-6 GHz and more than  $200\times$  at above-6 GHz when compared with the current 4G systems [12]; (ii) Spectral efficiency: an increase in spectral efficiency will be achieved by using large-scale antenna arrays, resulting in improved link budget through beamforming, and by developing new waveforms and protocols; (iii) Spatial efficiency: improvements in the order of 5 to 10 times are expected for spatial efficiency, which will be achieved by supporting reuse over smaller cells/geographic areas through denser networks; (iv) System efficiency: additional efficiency will be achieved by leveraging software defined network (SDN) and network function virtualization (NFV), which enable the network to reconfigure itself dynamically to meet instantaneous requirements; To carry out advances in these dimensions, three big technologies will take part in the 5G and B5G deployment [13], [14]: ultra-densification, millimeter wave (mmWave), and massive MIMO.

In the concept of cellular communication, as initially conceived in the early 80s, the cells were intended to provide coverage for a large area, on the order of hundreds of square kilometres [13]. The idea of densifying the wireless network consists of deploying a diverse set of base-stations (BS), including low power, low cost, and smaller (e.g., picocells and femtocells) BSs in the same area

covered by the traditional macrocell or in locations where the macrocell cannot provide enough signal strength [11], [15]. This concept, known as heterogeneous network (HetNet) [15], has already been applied and has provided gains over a thousandfold in the spectral efficiency [11], [16]. The ultra-densification comes from deploying highly dense HetNet, with tens or hundreds of small cells under one macrocell's coverage area, providing spectral reuse across smaller geographical areas and reducing the number of users sharing the same resources [13]. In dense networks, however, the small cells are closely located, which comes at the cost of increasing the inter-cell interference [16], [17]. In this context, the ultra-densification becomes attractive when utilizing higher frequencies such as mmWave, as the significant atmospheric attenuation, characteristic in those frequencies, imposes a natural cell radius on the order of tens to hundreds of meters [11].

The use of mmWave is also motivated by the spectral scarcity at below-3 GHz bands. As 5G and B5G systems will require more bandwidth, a natural tendency is to move to the 3-300 GHz<sup>1</sup> [12], [13], [18]. One of the most significant advantages of using mmWaves is the availability of large continuous unused or underused bandwidths [19]. While the total bandwidth allocated for today's mobile communication system does not exceed 780 MHz [20], the mmWave has up to 252 GHz of potentially available bandwidth [12], which will allow much higher data rates. On the other hand, one of the main characteristics of the mmWave frequencies is its adverse propagation characteristics, which is severely affected by path loss, penetration loss, atmospheric and rain attenuation, among other issues [12], [20]. Nevertheless, the high-frequency waves of the mmWave do not penetrate well in certain materials, especially in building materials such as concrete and bricks [21]. Although these effects may significantly degrade the mmWave propagation and make mmWave mobile communication infeasible for large areas, they are much less present in the short-range mobile communications, such as in the picocell and femtocell scenarios envisaged for the dense HetNets. In reality, for small-cells, these propagation effects can be beneficial as they will help promote larger densification by reducing the frequency reuse distance and limiting the inter-cell interference [14], [22], [23]. Furthermore, these propagation effects can be mitigated by using MIMO techniques. In particular, the shorter wavelength of mmWaves enables reducing the size of the antenna elements and thus, packing a much large number of antennas into small form factors [24], and even embed antenna arrays into portable devices [25], which favours the massive MIMO.

MIMO systems have long been known for providing diversity, power, and multiplexing gains [26]. The concept of massive MIMO was introduced by the breakthrough works of Marzetta [27], [28], which have shown that, in the single-cell scenarios, increasing the number of antennas is always

---

<sup>1</sup>Despite technically only the 30-300 GHz band corresponds to the mmWave band, as the 3-30 GHz band share similar propagation characteristics, it is common to refer to the entire 3-300 GHz band as mmWaves [12].

beneficial and, the effects of noise and small-scale fading tend to vanish as the number of antennas grows to infinity [27]. Furthermore, in massive MIMO, simple linear processing techniques, such as maximum ratio transmission (MRT) and maximum ratio combining (MRC), or zero-forcing (ZF) and minimum mean square error (MMSE) detector, can asymptotically achieve near-optimal capacity [27]–[30]. The above conclusions typically rely on the assumption of an ideal rich scattering propagation scenario, which is usually not the case for point-to-point and mmWave systems, where line-of-sight or poor scattering conditions cannot leverage the multiplexing gain of massive MIMO [24], [31]. However, mmWave systems can still benefit from the high directivity and high array gain provided by the large-scale arrays, which are expected to provide massive MIMO with a  $10\times$  increase in capacity and  $100\times$  improvement in the radiated energy efficiency [31]–[33]. Nevertheless, in terms of energy, massive MIMO systems are expected to use low power, less bulky components and benefit from additional robustness against malfunctioning or failure of few antenna elements [29], [33].

These three technologies are complementary in the sense that one will depend on or benefit from the others. For instance, the short wavelengths of mmWave will lead to smaller antennas, thus allowing massive MIMO arrays; massive arrays will help reduce the effects of poor propagation in mmWave; finally, this poor propagation will enable larger densification. However, putting these technologies together is very challenging. The next section discusses these challenges and a promising solution to overcome them.

## 1.2 Hybrid Beamforming: putting mmWave and Massive MIMO together

Conventional (fully-digital) MIMO arrays at sub-6 GHz frequencies usually handle less than ten antennas and have all the signal processing performed in the digital domain. Such architecture, depicted in Fig. 1, is known as the fully-digital MIMO. The fully-digital MIMO architecture requires each antenna to be connected to its own radio frequency (RF) chain devices [e.g. power amplifier (PA), low-noise amplifier (LNA), analog-to-digital converter (ADC), digital-to-analog converter (DAC), and frequency converters] [14], [24]. For massive MIMO, however, the RF chain’s cost and power consumption, especially in high frequencies, and the space occupied by all these devices, despite the small size of the antennas, will likely prevent the systems from using a complete RF chain per antenna [34].

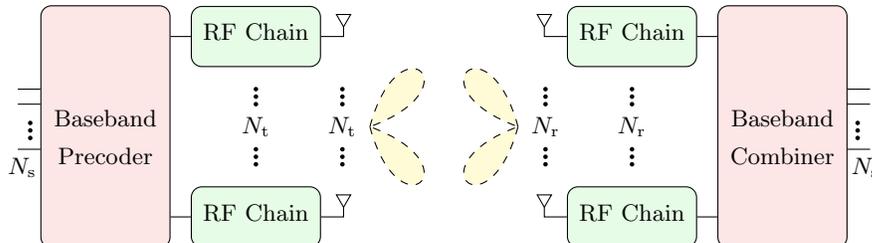
A promising solution to overcome these issues is using hybrid analog-digital beamforming, which

splits the signal processing between analog and digital domains. The hybrid beamforming uses much fewer RF chains than the number of antennas in the array. Here, the term hybrid beamforming is adopted to refer to both the hybrid precoding and the hybrid combining used, respectively, at the transmitter and receiver. At the hybrid precoding, the low-dimensional transmitted symbols are digitally precoded in the baseband domain, upconverted, and processed in the analog domain to produce the high-dimensional transmitted waveform. At the hybrid combining, the high-dimensional received waveform is processed in the analog domain and downconverted into a low-dimensional signal, which is further digitally combined.

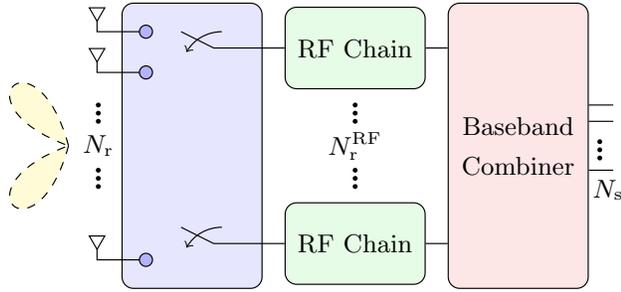
Hybrid beamforming combines the high array gain, provided by the use of many antenna elements, with the multiplexing gain, provided by the use of multiple RF chains. It also yields a very appealing trade-off between performance and complexity by supporting the high array and multiplexing gains of massive MIMO systems while preserving the reduced hardware complexity due to the lower number of RF chains. Besides, digital processing allows reducing intra- and inter-user interference [34], to support multi-stream and multi-user transmissions, and more robust operation in broadband channels [35].

The digital precoder and combiner are implemented in software (e.g., using digital signal processors). In contrast, the analog precoder and combiner stages are implemented in hardware, using either switches or phase-shifters (PS). These two different analog components establish two classes of hybrid beamforming [34]: switching-network and phase-shifting-network-based hybrid beamforming.

In the switching-network-based hybrid beamforming, the antennas are connected to the RF chains through switches as illustrated in Fig. 1.2. In this architecture, the analog beamforming design alone corresponds to the classical antenna selection (AS) problem [36]–[47], while the hybrid analog-digital beamforming design is often referred to as the joint antenna selection and beamforming (JASB) problem [48]–[55]. Although the joint antenna selection and beamforming problem has been widely studied for conventional MIMO [36]–[55], it has been brought to mmWave massive MIMO



**Figure 1.1:** Conventional (fully-digital) MIMO system.

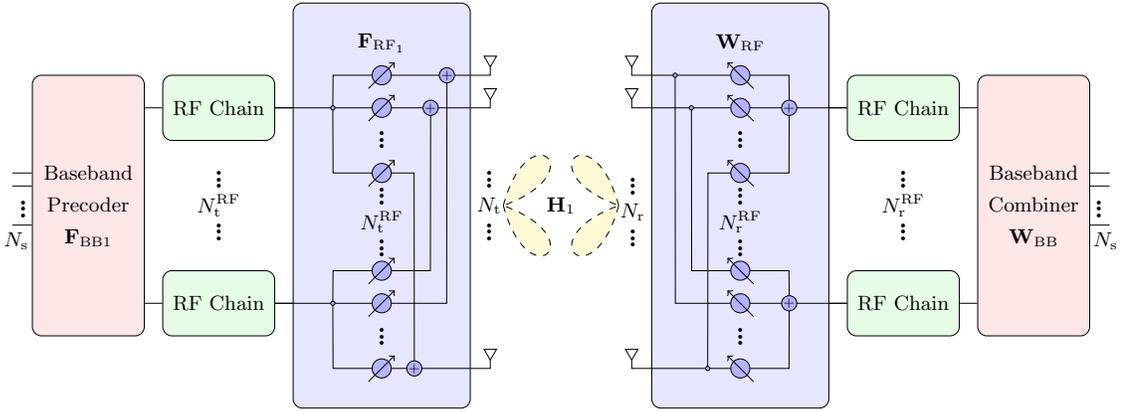


**Figure 1.2:** Switching-network-based hybrid combiner.

systems by [56], [57]. Switching-network-based hybrid beamforming uses very simple, low-power, low-cost switches, resulting in increased energy efficiency [56]. However, this architecture suffers from two critical drawbacks. First, an optimal solution for the antenna selection problem requires an exhaustive search over all possible antenna selection combinations. Since the number of combinations grows very rapidly with the increase in the number of antennas, finding the optimal solution becomes impractical even for small arrays [38]. Secondly, this architecture can not achieve large array gains as only a tiny fraction of the antennas are used [58].

In the phase-shifting-network-based hybrid beamforming, the RF chains are connected to all antennas or a subset of antennas (respectively, in the fully- and partially-connected architectures) through a phase-shifter network. At the precoder, the RF chain output signal is fed into all phase-shifters connected to that RF chain and the phase-shifted signals from different RF chains are combined before being transmitted by the antenna element. Conversely, at the combiner, the signal received by each antenna is fed into the phase-shifters connected to different RF chains, and the phase-shifted signals from different antennas are then combined before being processed by the RF chain. A phase-shifting-network-based hybrid beamforming scheme is presented in Fig. 1.3. Phase-shifter-based hybrid beamforming had been initially explored for conventional MIMO in [59] and others but has gained attention after Ayach’s work [24], which broke the ground in leveraging the mmWave channel characteristics for designing hybrid beamforming.

Compared to switching-network-based hybrid beamforming, the hybrid beamforming based on phase-shifting-network has access to all antennas, translating into a significant performance improvement in array gain and diversity gain [59]. Furthermore, with proper design, phase-shifting-network-based hybrid beamforming can provide near-optimal performance compared to the fully-digital beamforming, with a much lower number of RF chains, and consequently, lower complexity and power consumption. Despite these advantages, the phase-shifting-network-based hybrid beamforming design is also challenging: the PS imposes a constant-modulus and phase quantization



**Figure 1.3:** Phase-shifter-network-based hybrid precoder (left) and combiner (right).

constraints, leading to non-convex, combinatorial, NP-hard design problems [60]; and the coupling between analog and digital beamformers adds non-linearity to the problem and makes the channel estimation more difficult [35].

For conciseness, the following sections and chapters will refer to the switching-network-based hybrid beamforming as joint antenna selection and beamforming and will use hybrid beamforming to refer to the phase-shifting-network-based hybrid beamforming.

### 1.3 Objectives and Contributions

The objective of this thesis is to develop novel switching-network-based and phase-shifting-network-based hybrid beamforming designs for massive MIMO wireless communication systems. The contributions are presented in three chapters: the first chapter addresses the joint antenna selection and beamforming for multi-cell massive MIMO systems, and the second and third chapters address the hybrid beamforming design, respectively, for narrowband and OFDM mmWave massive MIMO systems. The objectives and contributions of each chapter are described below.

Chapter 2 aims to fill an important gap in JASB problems: the lack of adaptive algorithms capable of dealing with fixed limits on the number of RF chains. The chapter mainly focuses on developing two new adaptive algorithms to solve the receive JASB problem in cellular systems. These algorithms are based on an alternating optimization strategy that searches for the optimal solution by performing two steps at every algorithm iteration: first, the beamforming vector is obtained by considering an *a priori* antenna selection, then the antenna selection is updated using the *a posteriori* beamforming vector. The proposed algorithms can attain high levels of SINR while strictly complying with limitations on the number of RF chains. Moreover, the proposed algorithms

can track channel changes, making them suitable for operation in non-stationary environments, requiring only information about the desired signal (i.e., the corresponding direction of arrival or a reference signal). The main contributions of Chapter 2 are:

- First, a novel adaptive JASB (AJASB) algorithm based on the minimum MSE (MMSE) criterion is proposed. The idea behind such an algorithm is to compute the beamforming vector and perform the antenna selection aiming to reduce the MSE. In this context, the beamforming problem is formulated as an unconstrained MSE minimization problem, and the beamforming vector is updated using the normalized least mean square (NLMS) algorithm. Additionally, the antenna selection problem is formulated as a constrained MSE minimization problem (taking into account the hardware limitations), and the antenna selection is updated by using a constrained NLMS (CNLMS)-based algorithm. Simulation results are provided to assess the performance of the proposed algorithm under different scenarios.
- A second AJASB algorithm based on the minimum-variance distortionless response (MVDR) approach is also proposed. In this algorithm, the beamforming vector is obtained from the MVDR problem, i.e., by minimizing the variance at the beamformer's output while maintaining a constant gain toward the desired user. Similarly, the MVDR approach is used along with an additional constraint on the number of available RF chains to formulate the antenna selection problem. Both the beamforming vector and the antenna selection are updated by using CNLMS-based algorithms. The performance of the proposed algorithm under different scenarios are assessed through numerical simulation.

Chapter 3 aims to jointly design hybrid precoders and combiners for single-user and multi-user narrowband mmWave massive MIMO systems. The proposed designs are based on an iterative optimization framework that performs a constrained channel matrix decomposition. Additionally, the proposed design does not require *a priori* computation of the optimal unconstrained beamforming. The main contributions of Chapter 3 are summarized as follows:

- A two-stage approach for designing hybrid beamforming for narrowband mmWave massive MIMO systems is proposed. At the first stage, the analog precoder and combiner are designed to harvest the array gain provided by the massive MIMO system while simultaneously reducing the intra-user interference and multi-user interference (MUI). Such a design is formulated as a constrained low-rank channel decomposition problem to account for the PS's constant modulus constraint, resulting in an intractable problem. To make the problem tractable, a successive rank-1 channel decomposition approach is proposed. The rank-1 decomposition

problem computes one analog precoder and combiner vector pair and is solved by using the projected block coordinate descent method. This approach is very general and supports both single-user and multi-user systems. At the second stage, the digital precoder and combiner are computed. For single-user systems, the digital precoder and combiner is obtained from the optimal singular value decomposition (SVD)-based solution, considering the effective baseband channel. For multi-user systems, the digital precoder and combiner are computed using the regularized channel diagonalization approach [61], aiming to minimize the MUI plus noise.

- Extensive numerical simulation results are provided to confirm the effectiveness of the proposed design. The examples consider a wide range of scenarios for both Rayleigh and mmWave channel models and under practical constraints, such as quantized PS. The simulation results show that the proposed method consistently attains near-optimal performance and outperforms other hybrid beamforming designs in the literature in nearly all scenarios. In addition, a proof of optimality for the projected block coordinate descent solutions in the analog beamforming design is provided. This proof allows establishing important insights on the proposed algorithm’s guaranteed convergence, further supported by numerical simulation results.

In Chapter 4, hybrid beamforming designs for single-user and multi-user frequency-selective mmWave massive MIMO systems are proposed. The systems are assumed to use orthogonal frequency-division multiplexing (OFDM) and the designs are formulated by leveraging the OFDM channels’ multidimensional structure. The main contributions of Chapter 4 are:

- A two-stage hybrid beamforming design method for OFDM mmWave massive MIMO communication systems is proposed. The analog precoder and combiner are design in the first stage. Using a tensor representation of the OFDM channel (i.e., by stacking together the channel matrices of all subcarriers into a tensor), the analog beamforming design problem is formulated as a constrained Tucker2 channel tensor decomposition, which allows maximizing the sum of the effective baseband gains over every subcarrier while suppressing the interference among data streams within the same subcarrier. The problem is solved through successive rank-(1,1) tensor decomposition by using the projected alternate least square (ALS) method. The proposed analog beamforming design supports both single-user and multi-user systems. The digital precoder and combiner are designed in the second stage. The designs are obtained on a per-subcarrier basis, considering the effective baseband channel at each subcarrier. The optimal SVD-based beamforming and the regularized channel diagonalization are considered, respectively, for single-user and multi-user systems.

- The performance of the proposed design is assessed through extensive simulation. The results show that the proposed design achieves near-optimal sum-rates and outperforms other hybrid beamforming designs in nearly all scenarios. Examples also assess the convergence of the proposed algorithm and the impact of the power method approximation (adopted to reduce the computational complexity of the proposed design). Finally, an additional example illustrates the differences between this chapter’s design and that in Chapter 3, in particular, demonstrating that the design in Chapter 3 is not simply an special case of this chapter’s design.

## 1.4 Related Publications

Throughout the research work that has originated this thesis, the following peer-reviewed conference and journal papers have been published/submitted:

- G. M. Zilli, C. A. Pitz, E. L. O. Batista, R. Seara, and W.-P. Zhu, “An adaptive approach for the joint antenna selection and beamforming optimization,” *IEEE Access*, vol. 7, pp. 99 017–99 030, 2019
- G. M. Zilli and W.-P. Zhu, “Constrained-SVD based hybrid beamforming design for millimeter-wave communications,” in *Proc. IEEE 92nd Vehicular Technology Conference: VTC2020-Fall*, Victoria, BC: IEEE, Nov. 2020, pp. 1–5
- G. M. Zilli and W.-P. Zhu, “Constrained channel decomposition-based hybrid beamforming for mmWave massive MIMO systems,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1707–1720, 2020
- G. M. Zilli and W.-P. Zhu, “TUCKER2-based hybrid beamforming design for mmWave OFDM massive MIMO communications,” in *Proc. IEEE 93rd Vehicular Technology Conference: VTC2021-Spring*, Helsinki, Finland: IEEE, Apr. 2021, pp. 1–5
- G. M. Zilli and W.-P. Zhu, “Constrained Tensor Decomposition-Based Hybrid Beamforming for mmWave Massive MIMO-OFDM Communication Systems”, under the second round of review.

## 1.5 Organization

This thesis is organized as follows. Chapter 2 presents an overview of the literature on JASB, along with the system model and problem formulation, the derivation of the two adaptive algorithms for solving the adaptive JASB problem proposed herein, and numerical simulation results. Chapter 3 presents a literature review on current hybrid beamforming design approaches for narrowband mmWave massive MIMO systems; the system model, the narrowband channel model, and the problem formulation; the proposed narrowband hybrid beamforming design based on the constrained channel matrix decomposition; and the numerical simulation results. Chapter 4 presents a literature review on existing hybrid beamforming designs for wideband OFDM mmWave massive MIMO systems and an overview of tensor operations; the system model, the OFDM channel model, and the problem formulation; the proposed single-user and multi-user hybrid beamforming design for OFDM systems based on the constrained channel tensor decomposition; and the numerical simulation results. Finally, Chapter 5 summarizes this thesis's contributions and points out future extensions for the research problems discussed herein.

Throughout this thesis, the following notations are adopted: non-bold, boldfaced lower-case, and boldfaced upper-case letters are used, respectively, for scalars, vectors, and matrices; operators  $(\cdot)^*$ ,  $(\cdot)^T$ , and  $(\cdot)^H$  denote, respectively, the complex conjugate, the transpose, and the Hermitian (conjugate transpose);  $|\cdot|$ ,  $(\cdot)^{-1}$ , and  $(\cdot)^\dagger$  denote, respectively, matrix determinant, inverse, and Moore-Penrose pseudo-inverse;  $|\cdot|$  and  $\|\cdot\|$  denote, respectively, the absolute value of a scalar and the norm of a vector, while  $\|\cdot\|_F$  denote the Frobenius norm of a matrix;  $\|\cdot\|_0$  denotes the zero-norm (i.e., the number of non-zero elements) of a vector;  $\mathbf{I}_M$  denotes an  $M \times M$  identity matrix and  $\mathbf{0}_{M \times N}$  denotes an  $M \times N$  zero matrix, while  $\mathbf{0}$  and  $\mathbf{1}$  represent, respectively, vectors of zeros and ones;  $[\mathbf{v}]_m$  represents the  $m$ -th element of a vector  $\mathbf{v}$  and  $[\mathbf{A}]_{m,n}$  denotes the element in the  $m^{\text{th}}$  row and the  $n^{\text{th}}$  column of matrix  $\mathbf{A}$ , while  $\mathbf{A}(n, :)$  and  $\mathbf{A}(:, n)$  represent, respectively, the  $n^{\text{th}}$  row and the  $n^{\text{th}}$  column of matrix  $\mathbf{A}$ ; The matrix trace, the expectation, and the real-part operator are denoted by  $\text{Tr}(\cdot)$ ,  $\text{E}(\cdot)$ , and  $\text{Re}(\cdot)$ , respectively;  $\mathbb{C}$  denotes the set of complex numbers and  $\mathcal{CN}(\mathbf{a}, \mathbf{A})$  represents the complex Gaussian distribution with mean  $\mathbf{a}$  and covariance matrix  $\mathbf{A}$ ;  $\angle\{\cdot\}$  denotes the phase of a complex number.

## Chapter 2

# Adaptive Joint Antenna Selection and Beamforming

### 2.1 Introduction

Antenna arrays have been considered a promising solution to several problems of cellular communication systems, such as limited spectrum availability, reduced battery life of mobile devices, interference due to frequency reuse, and time and frequency selective fading [66], [67]. The performance gain obtained by using antenna arrays gets larger as the number of array elements increases [66], which has motivated the use of large-scale arrays, the so-called massive MIMO [28]. However, implementation costs often limit the expansion of the number of antennas in massive MIMO arrays. This problem is especially relevant in mmWave communication systems, in which the very high carrier frequencies and signal bandwidths result in severe hardware constraints [34]. Although the cost of additional antenna elements and the associated signal processing hardware is relatively small, the high cost of RF chains restricts the implementation of mmWave communication systems with a higher number of antennas [38], [68], [69]. Furthermore, it is well known that, in a real-world propagation environment, the elements of an antenna array do not contribute equally to the overall system performance [70]. In this context, several strategies have been devised to efficiently allocate the available RF chains, such as hybrid beamforming [24], [34] and antenna selection techniques [38], [68]–[70]. In particular, antenna selection has gained considerable attention in mmWave massive MIMO, as it allows reducing both hardware complexity and power consumption [57], [71]–[73].

Antenna selection techniques account for finding the optimal subset of antennas within the available set. Solving this problem is not a trivial task since, in general, it involves an extensive search within all possible combinations of antennas, making the problem computationally prohibitive even for arrays with a few dozens of elements [38], [69]. As a result, several suboptimal strategies have been proposed in the literature aiming to solve antenna-selection problems. For instance, in [41], a decremental antenna-selection scheme is introduced by using a capacity-maximization criterion for iteratively removing antenna elements until a desired number of antennas is reached. By exploiting a similar concept, an incremental antenna-selection technique is proposed in [42], where the antenna elements are iteratively included. Despite their conceptual simplicity, these techniques assume time-invariant channel state information and require evaluating capacity loss/gain for every removed/included antenna. The resulting computational costs limit the application of these techniques, especially in large-scale arrays and non-stationary environments. In [43], receive antenna selection is cast as a capacity-maximization-based convex problem that can be solved in polynomial time. However, the obtained solution is channel-dependent, which restrains the application of this method to stationary environments. Convex optimization has also been applied to antenna selection, aiming to minimize the mean-square error (MSE) [44] and to maximize the channel capacity in massively distributed antenna systems [45]. The main drawback of these approaches is their reliance on computationally costly optimization solvers, usually with a computational complexity of order  $\mathcal{O}(M^3)$  (where  $M$  is the number of antennas) [74], making a real-time implementation very hard.

In [46] and [47], adaptive antenna selection algorithms based on stochastic optimization are proposed to maximize both the channel capacity and the signal-to-noise ratio (SNR) and minimize the bit error rate. These algorithms focused on the real-time operation and can track a suboptimal antenna subset in time-varying channels. However, they require a dictionary containing all possible combinations of antenna selection and their corresponding probabilities and require the computation of the cost function twice in each iteration [46], [47], thus incurring a large amount of memory and have high computational complexity.

The approaches proposed in [41]–[47] consider the antenna selection problem apart from the beamforming design. In contrast, the joint antenna selection and beamforming problem has been studied in several other research works [48]–[55]. Transmit JASB problems can be formulated under different criteria, such as total power minimization under quality-of-service (QoS) constraints, energy-efficiency maximization, and decremental selection based on power minimization. These problems involve, in general, solving one or multiple semidefinite programming (SDP) problems, and therefore, also rely on computationally costly optimization solvers.

On the other hand, several low-complexity adaptive techniques have been proposed for solving receive JASB problems. For instance, the adaptive algorithms proposed in [52]–[55] indirectly promote antenna selection by inserting sparsity-inducing  $\ell_1$ -norm or  $\ell_0$ -norm constraints into the derivation of constrained least-mean-square (CLMS) algorithms [75], [76]. In doing so, antenna selection can be performed by thinning the array, i.e., by switching off the antenna elements for which the beamforming coefficients are very small [54]. This an approach is useful when the aim is to reduce power consumption and heat dissipation. However, one common assumption for the development of the algorithms in [52]–[55] is that the signal received by each antenna element is available at every iteration of the adaptive process, which does not comply with physical limitations on the number of RF chains.

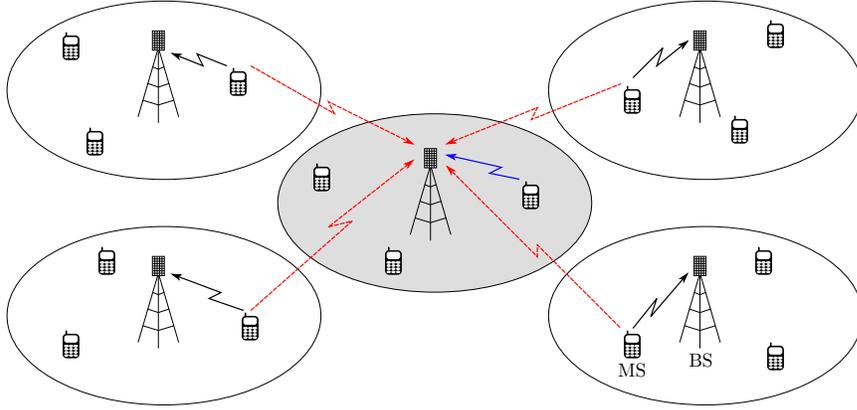
This chapter proposes two new adaptive approaches for solving the JASB problem. These algorithms are based on the MMSE and MVDR criteria and are designed to attain the hardware-imposed constraint on the number of RF chains at every iteration. This chapter is organized as follow: Section 2.2 describes the system model and presents the problem formulation; Sections 2.3 proposes the new adaptive algorithm for solving the MMSE-based JASB problem and presents simulation results to asses its effectiveness; Section 2.4 proposes the new adaptive algorithm for the MVDR-based JASB problem and also presents simulation results to asses its effectiveness; and finally, Section 2.5 presents the concluding remarks.

## 2.2 System Model and Problem Statement

In this section, the communication scenario considered in this chapter is firstly depicted, followed by the description of the adopted performance metric, and the optimization problem that lays the groundwork for developing the proposed algorithms.

### 2.2.1 Scenario Definition

The scenario considered in this work involves a multi-cell cellular system in which each BS serves multiple single-antenna users. For a particular BS, it is assumed that intra-cell users are allocated to different channel resources and do not interfere with each other. On the other hand,  $K - 1$  inter-cell users located at nearby cells share the same channel resources with a given intra-cell user, as illustrated in Fig. 2.1. Moreover, it is also assumed that the considered BS is equipped with a massive antenna array. The array is partitioned into subarrays formed by  $M$ -element antenna and  $L(\leq M)$  RF chains, as shown in Fig. 1.2, where the subarrays are individually assigned to each



**Figure 2.1:** Multi-cell system configuration with the signal of interest (blue line) and the interfering signals (red lines) arriving at the considered BS [62].

intra-cell user [57], [72], [73].

The equivalent baseband representation of uplink signals arriving at the subarray allocated to a given intra-cell user are arranged in the input vector  $\mathbf{x}(n) \in \mathbb{C}^{M \times 1}$ , such that

$$\mathbf{x}(n) = \sum_{i=0}^{K-1} \mathbf{a}(\theta_i) \beta_i(n) + \mathbf{z}(n) \quad (2.1)$$

where  $\mathbf{a}(\theta_i) \in \mathbb{C}^{M \times 1}$  represents the steering vector related to the angle of arrival  $\theta_i$  of the  $i$ th user; the complex scalar  $\beta_i(n)$  denotes the zero-mean baseband symbol with power  $p_i = \mathbb{E}[|\beta_i(n)|^2]$  corresponding to the  $i$ th user; and  $\mathbf{z}(n) \in \mathbb{C}^{M \times 1}$  models the complex additive white Gaussian noise (AWGN) with power  $\sigma_z^2$  at each antenna element. The signals transmitted by the remaining intra-cell users are not considered in (2.1), since they are allocated to different channel resources. The input vector after the antenna selection is written as

$$\mathbf{x}_s(n) = \mathbf{S}\mathbf{x}(n) \quad (2.2)$$

which contains the signals processed by the RF chains, with  $\mathbf{S}$  representing an  $M \times M$  diagonal antenna-selection matrix whose  $m$ th diagonal element is given by

$$[\mathbf{S}]_{m,m} = \begin{cases} 1, & \text{if the } m\text{th antenna is selected} \\ 0, & \text{if the } m\text{th antenna is not selected.} \end{cases}$$

Note that, in practice, only  $\mathbf{x}_s(n)$  is available for the system. Considering the number of selected

antennas equal to the number of available RF chains,  $\text{Tr}(\mathbf{S}) = L$ . In particular, when all antennas are selected (i.e.,  $L = M$ ),  $\mathbf{S}$  becomes the  $M \times M$  identity matrix  $\mathbf{I}_M$ , resulting in  $\mathbf{x}_s(n) = \mathbf{x}(n)$ .

Denoting  $\mathbf{w} \in \mathbb{C}^{M \times 1}$  as the beamforming vector, the beamforming output is given by

$$y(n) = \mathbf{w}^H \mathbf{x}_s(n). \quad (2.3)$$

Furthermore, defining the first inter-cell user ( $i = 0$ ) as the signal of interest (SOI), the reference signal becomes  $d(n) = \beta_0(n)$ , and the error signal between the reference signal and the beamformer output can be defined as

$$e(n) = d(n) - y(n) = \beta_0(n) - \mathbf{w}^H \mathbf{x}_s(n). \quad (2.4)$$

Next, using (2.2), (2.3), and (2.4), the MSE is given as

$$\begin{aligned} \mathbb{E}[|e(n)|^2] &= \mathbb{E}[(\beta_0(n) - y(n))(\beta_0^*(n) - y^*(n))] \\ &= \mathbb{E}[\mathbf{w}^H \mathbf{S} \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{S} \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{S} \mathbf{x}(n) \beta_0^*(n)\} + |\beta_0(n)|^2] \\ &= \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{S} \mathbf{p}\} + p_0^2 \end{aligned} \quad (2.5)$$

where  $\mathbf{R}_x = \mathbb{E}[\mathbf{x}(n) \mathbf{x}^H(n)]$  denotes the autocorrelation matrix of the input vector, and  $\mathbf{p} = \mathbb{E}[\mathbf{x}(n) \beta_0^*(n)]$  is the cross-correlation vector between the input vector and the desired signal.

To obtain an expression for the signal-to-interference-plus-noise ratio (SINR) at the beamformer output, the mean power (variance) of  $y(n)$  is derived from (2.3) as follows:

$$\mathbb{E}[|y(n)|^2] = \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w}. \quad (2.6)$$

If it is further assumed that the baseband symbols are independent, i.e.,  $\mathbb{E}[\beta_i(n) \beta_j^*(n)] = 0$  for  $i \neq j$ ,  $\mathbf{R}_x$  can be expressed as

$$\mathbf{R}_x = \sum_{i=0}^{K-1} p_i \mathbf{a}(\theta_i) \mathbf{a}^H(\theta_i) + \sigma_z^2 \mathbf{I}_M. \quad (2.7)$$

Then, rearranging (2.7) into

$$\mathbf{R}_x = \mathbf{R}_{\text{soi}} + \mathbf{R}_{\text{in}} \quad (2.8)$$

with

$$\mathbf{R}_{\text{soi}} = p_0 \mathbf{a}(\theta_0) \mathbf{a}^H(\theta_0) \quad (2.9)$$

representing the autocorrelation matrix of the SOI and

$$\mathbf{R}_{\text{in}} = \sum_{i=1}^{K-1} p_i \mathbf{a}(\theta_i) \mathbf{a}^H(\theta_i) + \sigma_z^2 \mathbf{I}_M \quad (2.10)$$

representing the autocorrelation matrix of the interference plus noise, (2.6) can be rewritten as

$$\text{E} [|y(n)|^2] = \mathbf{w}^H \mathbf{S} \mathbf{R}_{\text{soi}} \mathbf{S} \mathbf{w} + \mathbf{w}^H \mathbf{S} \mathbf{R}_{\text{in}} \mathbf{S} \mathbf{w}. \quad (2.11)$$

Thus, the SINR at the beamformer output is given by

$$\gamma = \frac{\mathbf{w}^H \mathbf{S} \mathbf{R}_{\text{soi}} \mathbf{S} \mathbf{w}}{\mathbf{w}^H \mathbf{S} \mathbf{R}_{\text{in}} \mathbf{S} \mathbf{w}}. \quad (2.12)$$

### 2.2.2 Joint Antenna Selection and Beamforming Problem

To address the joint antenna selection and beamforming problem with a limited number of RF chains, a performance metric  $\Phi(\mathbf{w}, \mathbf{S})$  is optimized with respect to the beamforming vector and the antenna selection matrix, subject to a constraint on the number of selected antennas. As a result, the following constrained optimization problem is formulated:

$$\begin{aligned} & \underset{\mathbf{w}, \mathbf{S}}{\text{minimize}} && \Phi(\mathbf{w}, \mathbf{S}) \\ & \text{subject to} && [\mathbf{S}]_{m,m} \in \{0, 1\} \\ & && \text{Tr}(\mathbf{S}) = L. \end{aligned} \quad (2.13)$$

Here, the performance metric  $\Phi(\mathbf{w}, \mathbf{S})$  focuses on two criteria: MMSE and MVDR. It is noteworthy that regardless of the choice of  $\Phi(\mathbf{w}, \mathbf{S})$ , the problem in (2.13) is NP-hard due to the binary constraints involved [43], [69]. Therefore, finding an optimal solution to (2.13) is very computationally expensive, especially for large scale arrays [69].

It is important to notice that a single SOI is considered in (2.13) due to the assumption that an individual subarray is allocated for each intra-cell user. In cases where the intra-cell users share the same channel and RF hardware resources, the objective function of (2.13) must be reformulated to include all intra-cell users.

## 2.3 MMSE-based Adaptive Joint Antenna Selection and Beamforming

This section is focused on the development of a new MMSE-based adaptive algorithm for solving (2.13). Assuming  $\Phi(\mathbf{w}, \mathbf{S}) = \text{E}[|e(n)|^2]$ , (2.13) is partitioned such that  $\mathbf{w}$  and  $\mathbf{S}$  can be obtained by using an alternating optimization approach. Therefore, the proposed algorithm consists of alternately optimizing the beamforming vector for a fixed antenna selection matrix and the antenna selection matrix for a fixed beamforming vector. The following subsections describe the detailed procedure and present simulation results to validate the proposed approach.

### 2.3.1 MMSE Beamforming

The beamforming procedure considered in this section is based on the minimization of  $\text{E}[|e(n)|^2]$  as defined in (2.5). Thus, the beamforming vector is obtained from the following optimization problem:

$$\underset{\mathbf{w}}{\text{minimize}} \quad \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{S} \mathbf{p}\} + p_0^2. \quad (2.14)$$

The optimum solution for (2.14) is the well-known Wiener-Hopf solution [77], i.e.,

$$\mathbf{w}_o = (\mathbf{S} \mathbf{R}_x \mathbf{S})^\dagger \mathbf{S} \mathbf{p}, \quad (2.15)$$

where  $\mathbf{w}_o$  represents the optimum beamforming vector and  $(\cdot)^\dagger$  denotes the Moore-Penrose inverse, which is required since  $\mathbf{S} \mathbf{R}_x \mathbf{S}$  results in a rank-deficient matrix [i.e.,  $\text{rank}(\mathbf{S} \mathbf{R}_x \mathbf{S}) \leq L \leq M$ ]. One can notice that the evaluation of the Wiener-Hopf solution requires not only *a priori* knowledge of  $\mathbf{R}_x$  and  $\mathbf{p}$ , but also the computation of  $(\mathbf{S} \mathbf{R}_x \mathbf{S})^\dagger$ , which results in a high computational burden, especially in cases involving real-time processing, massive arrays and/or non-stationary scenarios. Therefore, to overcome these difficulties, the steepest descent method [77] is used for iteratively obtaining the beamforming vector, which leads to the following update rule:

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu_w \nabla_{\mathbf{w}^*} \text{E}[|e(n)|^2] \\ &= \mathbf{w}(n) - \mu_w [\mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w}(n) - \mathbf{S} \mathbf{p}]. \end{aligned} \quad (2.16)$$

with  $\nabla_{\mathbf{w}^*} \text{E}[|e(n)|^2] = [\mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w}(n) - \mathbf{S} \mathbf{p}]$  representing the gradient of  $\text{E}[|e(n)|^2]$  with respect to the complex conjugate of  $\mathbf{w}$ . Then, considering (2.2)-(2.4) and replacing  $\mathbf{R}_x$  and  $\mathbf{p}$  by their corresponding

instantaneous estimates given, respectively, by  $\mathbf{x}(n)\mathbf{x}^H(n)$  and  $\mathbf{x}(n)\beta_0^*(n)$ , one obtain an LMS-type update expression, given as

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu_w [\mathbf{x}_s(n)\mathbf{x}_s^H(n)\mathbf{w}(n) - \mathbf{x}_s(n)\beta_0^*(n)] \\ &= \mathbf{w}(n) + \mu_w \mathbf{x}_s(n)e^*(n)\end{aligned}\tag{2.17}$$

where  $\mu_w$  is the step-size parameter.

It is important to highlight that only the input signals present at the selected antennas are used in (2.17), and therefore, only the coefficients corresponding to these antennas are updated at a given iteration. Therefore, the algorithm can be targeted to practical applications, where only a limited number of RF chains is available. Moreover, it is assumed that the error signal  $e(n)$  required by (2.17) is obtained during a training period, wherein the intra-cell users transmit a predefined training sequence (reference signal) known to the BS. The design of this training sequence, including its length and repetition rate, directly impact the ability of the system to track channel variations, thereby influencing the performance of the AJASB algorithm. However, such a design is out of the scope of this work, being a recurrent research topic available in the open literature (see [78]–[81]).

Finally, to achieve a faster convergence and easier parameter adjustment, a normalized version of (2.17) is also formulated here following the same idea used for developing the NLMS algorithm [77]. Thus, the following NLMS update rule is obtained:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_w \frac{\mathbf{x}_s(n)e^*(n)}{\mathbf{x}_s^H(n)\mathbf{x}_s(n) + \varepsilon}\tag{2.18}$$

where  $\varepsilon$  is a regularization term, included to prevent divisions by very small values [77].

### 2.3.2 MMSE Antenna Selection

The MMSE antenna selection problem considered here consists of finding the antenna subset that minimizes (2.5). This problem can be mathematically described as

$$\begin{aligned}\underset{\mathbf{S}}{\text{minimize}} \quad & \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{S} \mathbf{p}\} + p_0^2 \\ \text{subject to} \quad & [\mathbf{S}]_{m,m} \in \{0, 1\} \\ & \text{Tr}(\mathbf{S}) = L.\end{aligned}\tag{2.19}$$

In spite of being easier to solve it in comparison to (2.13) (since only  $\mathbf{S}$  needs to be optimized), (2.19) is still an NP-hard problem due to the binary constraints involved. To circumvent this drawback,

the binary constraints are relaxed, allowing  $[\mathbf{S}]_{m,m}$  to be a real number. Then, defining the auxiliary vector  $\mathbf{s}$  as the diagonal vector of  $\mathbf{S}$ , defining also  $\mathbf{W}$  as a diagonal matrix whose entries are given by  $\mathbf{w}$  [i.e.,  $\mathbf{W} = \text{diag}(\mathbf{w})$ ], and considering that  $p_0^2$  does not depend on the optimization variable, the relaxed version of (2.19) becomes

$$\begin{aligned} & \underset{\mathbf{s} \in \mathbb{R}^M}{\text{minimize}} && \mathbf{s}^T \mathbf{W}^H \mathbf{R}_x \mathbf{W} \mathbf{s} - 2 \text{Re}\{\mathbf{s}^T \mathbf{W}^H \mathbf{p}\} \\ & \text{subject to} && \mathbf{1}^T \mathbf{s} = L. \end{aligned} \quad (2.20)$$

Note that  $\mathbf{W}^H \mathbf{R}_x \mathbf{W}$  is a positive semidefinite matrix and, therefore, (2.20) is a convex problem [82].

To find the optimal solution for (2.20), one can use the Lagrangian form

$$\mathcal{L}(\mathbf{s}, \lambda) = \mathbf{s}^T \mathbf{W}^H \mathbf{R}_x \mathbf{W} \mathbf{s} - 2 \mathbf{s}^T \text{Re}\{\mathbf{W}^H \mathbf{p}\} + \lambda(\mathbf{1}^T \mathbf{s} - L) \quad (2.21)$$

where  $\lambda$  is the Lagrange multiplier. Next, by taking the gradient of (2.21) with respect to  $\mathbf{s}$ , one obtain

$$\nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda) = 2 \mathbf{R}_{rw} \mathbf{s} - 2 \mathbf{r}_p + \lambda \mathbf{1} \quad (2.22)$$

with

$$\mathbf{R}_{rw} = \text{Re}\{\mathbf{W}^H \mathbf{R}_x \mathbf{W}\} \quad (2.23)$$

$$\mathbf{r}_p = \text{Re}\{\mathbf{W}^H \mathbf{p}\}. \quad (2.24)$$

The optimum selection vector  $\mathbf{s}_o$  is obtained by solving  $\nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda) = \mathbf{0}$ , which leads to

$$\mathbf{s}_o = \frac{1}{2} \mathbf{R}_{rw}^{-1} (2 \mathbf{r}_p - \lambda \mathbf{1}). \quad (2.25)$$

Regarding  $\lambda$ , its value is obtained by reinforcing the constraint  $\mathbf{1}^T \mathbf{s}_o = L$ . Then, pre-multiplying both sides of (2.25) by  $\mathbf{1}^T$  and manipulating the resulting expression yields

$$\lambda = 2 \left[ \frac{\mathbf{1}^T \mathbf{R}_{rw}^{-1} \mathbf{r}_p - L}{\mathbf{1}^T \mathbf{R}_{rw}^{-1} \mathbf{1}} \right]. \quad (2.26)$$

Next, by substituting (2.26) into (2.25), the optimal solution to the relaxed problem in (2.20) is given as

$$\mathbf{s}_o = \mathbf{R}_{rw}^{-1} \mathbf{r}_p - \frac{\mathbf{R}_{rw}^{-1} (\mathbf{1}^T \mathbf{R}_{rw}^{-1} \mathbf{r}_p) \mathbf{1}}{\mathbf{1}^T \mathbf{R}_{rw}^{-1} \mathbf{1}} + L \frac{\mathbf{R}_{rw}^{-1} \mathbf{1}}{\mathbf{1}^T \mathbf{R}_{rw}^{-1} \mathbf{1}}. \quad (2.27)$$

Note that  $\mathbf{s}_o$  is a vector of real numbers and, thus, the diagonal elements of the antenna-selection

matrix  $\mathbf{S}$  cannot be obtained directly from such a vector. To circumvent this problem, a binary vector  $\mathbf{s}_b$  is obtained considering the minimum disturbance with respect to  $\mathbf{s}_o$  (i.e.,  $\mathbf{s}_b$  is obtained minimizing its Euclidean distance to  $\mathbf{s}_o$ ). This strategy can be mathematically expressed as follows:

$$\begin{aligned} & \underset{\mathbf{s}_b}{\text{minimize}} && \|\mathbf{s}_b - \mathbf{s}_o\|^2 \\ & \text{subject to} && [\mathbf{s}_b]_m \in \{0, 1\} \\ & && \mathbf{1}^T \mathbf{s}_b = L, \end{aligned} \tag{2.28}$$

By rewriting the cost function in (2.28) as  $\|\mathbf{s}_o\|^2 + \|\mathbf{s}_b\|^2 - 2\mathbf{s}_b^T \mathbf{s}_o$  and having the constraints imposing  $\|\mathbf{s}_b\|^2 = L$ , one can verify that the minimum is attained when the inner product  $\mathbf{s}_b^T \mathbf{s}_o$  is maximized. The a maximization, in turn, can be reached by setting to one the  $L$  components of the binary vector  $\mathbf{s}_b$  corresponding to the  $L$  largest components of  $\mathbf{s}_o$ . The antenna selection matrix is then determined by making the diagonal of  $\mathbf{S}$  equal to  $\mathbf{s}_b$  [i.e.,  $\mathbf{S} = \text{diag}(\mathbf{s}_b)$ ], which is a feasible solution to the original problem (2.19). It is important to note that the matrix  $\mathbf{S}$  obtained from  $\mathbf{s}_b$  will not usually be an optimal solution for (2.20). Despite this fact, the use of the minimum disturbance as criterion for obtaining  $\mathbf{s}_b$  ensures that such a vector will generally be a solution close to the optimal one.

The closed-form solution given in (2.27) requires *a priori* knowledge of  $\mathbf{R}_x$  and  $\mathbf{p}$ , as well as the inversion of  $\mathbf{R}_{rw}$ , which incurs high computational costs. Thus, one can use the steepest descent method to iteratively obtain  $\mathbf{s}$  as follows:

$$\begin{aligned} \mathbf{s}(n+1) &= \mathbf{s}(n) - \mu_s \nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda) \\ &= \mathbf{s}(n) - 2\mu_s \mathbf{R}_{rw} \mathbf{s} + 2\mu \mathbf{r}_p - \mu \lambda \mathbf{1}. \end{aligned} \tag{2.29}$$

where  $\mu_s$  is the step-size parameter and the gradient  $\nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda)$  is defined in (2.22). The Lagrangian multiplier  $\lambda$  can be found in (2.29) by reinforcing the *a posteriori* constraint  $\mathbf{1}^T \mathbf{s}(n+1) = L$ , which, along with  $\mathbf{1}^T \mathbf{1} = M$ , leads to

$$\lambda = \frac{\mathbf{1}^T \mathbf{s}(n) - 2\mu \mathbf{1}^T \mathbf{R}_{rw} \mathbf{s}(n) + 2\mu \mathbf{1}^T \mathbf{r}_p - L}{\mu M}. \tag{2.30}$$

Then, substituting (2.30) into (2.29) yields

$$\mathbf{s}(n+1) = \mathbf{P}_s [(\mathbf{I} - 2\mu \mathbf{R}_{rw}) \mathbf{s}(n) + 2\mu \mathbf{r}_p] + \mathbf{f}_s \tag{2.31}$$

with

$$\mathbf{P}_s = \left[ \mathbf{I}_M - \frac{\mathbf{1}\mathbf{1}^T}{M} \right] \quad (2.32)$$

and

$$\mathbf{f}_s = \frac{L}{M} \mathbf{1}. \quad (2.33)$$

Finally, using the instantaneous estimates of  $\mathbf{r}_p$  and  $\mathbf{R}_{\text{rw}}$  given by  $\text{Re}\{\mathbf{W}^H(n)\mathbf{x}_s(n)\beta_0^*(n)\}$  and  $\text{Re}\{\mathbf{W}^H(n)\mathbf{x}_s(n)\mathbf{x}_s^H(n)\mathbf{W}(n)\}$ , respectively, leads to

$$\mathbf{s}(n+1) = \mathbf{P}_s \left[ \left( \mathbf{I}_M - 2\mu \text{Re}\{\mathbf{W}^H(n)\mathbf{x}_s(n)\mathbf{x}_s^H(n)\mathbf{W}(n)\} \right) \mathbf{s}(n) + 2\mu \text{Re}\{\mathbf{W}^H(n)\mathbf{x}_s(n)d^*(n)\} \right] + \mathbf{f}_s. \quad (2.34)$$

### 2.3.3 MMSE-based AJASB Algorithm

The proposed MMSE-based AJASB (MMSE-AJASB) algorithm, summarized in Algorithm 2.1, is based on an iterative scheme that concurrently uses the MMSE-based beamforming strategy from Section 2.3.1 and the MMSE-based antenna-selection strategy from Section 2.3.2.

From Algorithm 2.1, one can notice that, in the first steps, the received signal  $\mathbf{x}_s(n)$  is used for computing the beamforming output  $y(n)$  and the error signal  $e(n)$ . Next, the beamforming vector  $\mathbf{w}(n)$  is updated employing the NLMS algorithm presented in Section 2.3.1. Once the *a posteriori* beamforming vector  $\mathbf{w}(n+1)$  is calculated, it is further used to compute  $\mathbf{s}(n+1)$  considering the algorithm proposed in Section 2.3.2. Finally, the antenna-selection matrix is evaluated by selecting the antennas corresponding to the  $L$  largest components of  $\mathbf{s}(n+1)$ . This operation is denoted by  $\text{select}[\mathbf{s}(n+1)]$ , which results in a diagonal matrix with ones at the diagonal elements corresponding to the  $L$  largest components of  $\mathbf{s}(n+1)$  and zeros otherwise.

In terms of real operations per iteration, the MMSE-based AJASB algorithm requires  $17L + M$  sums,  $18L + 5$  multiplications, and one division for the beamforming and antenna-selection-vector updates, whereas the remaining  $\text{select}[\cdot]$  operation requires at most  $(M - L)L$  real comparisons. In contrast, the standard MMSE-based NLMS beamforming requires  $10M$  sums,  $10M + 2$  multiplications, and one division. Since, in general,  $M \gg L$ , the proposed MMSE-based AJASB algorithm usually has lower computational costs than its NLMS counterpart. The proposed algorithm's computational cost is also much smaller than those of the algorithms based on convex optimization solvers, for which the complexity is on the  $\mathcal{O}(M^3)$  order.

An accurate convergence analysis of the proposed algorithm can only be achieved from the corresponding stochastic modeling. However, this type of modeling is complicated for the proposed

---

**Algorithm 2.1:** Summary of the MMSE-AJASB Algorithm

---

```

1 Initialize:
2    $\mu_w, \mu_s, \varepsilon$ 
3    $\mathbf{P}_s = \left[ \mathbf{I}_M - \frac{\mathbf{1}\mathbf{1}^T}{M} \right]$ 
4    $\mathbf{w}(0) = [1 \ 0 \ \dots \ 0]^T$ 
5    $\mathbf{s}(0) = \frac{L}{M} \mathbf{1}$ 
6    $\mathbf{S}(0) = \begin{bmatrix} \mathbf{I}_L & \mathbf{0}_{L \times (M-L)} \\ \mathbf{0}_{(M-L) \times L} & \mathbf{0}_{(M-L) \times (M-L)} \end{bmatrix}$ 
7 for  $n = 0, 1, 2, \dots$  do
8    $y(n) = \mathbf{w}^H(n) \mathbf{x}_s(n)$ 
9    $e(n) = d(n) - y(n)$ 
10   $\mathbf{w}(n) = \mathbf{w}(n) + \mu_w \frac{e^*(n) \mathbf{x}_s(n)}{\mathbf{x}_s(n)^H \mathbf{x}_s(n) + \varepsilon}$ 
11   $\mathbf{W}(n) = \text{diag}[\mathbf{w}(n+1)]$ 
12   $\mathbf{s}(n+1) = \mathbf{P}_s \left[ \mathbf{I}_M - 2\mu_s \text{Re}\{\mathbf{W}^H(n) \mathbf{x}_s(n) \mathbf{x}_s^H(n) \mathbf{W}(n)\} \right] \mathbf{s}(n)$ 
13   $\quad + 2\mu_s \text{Re}\{\mathbf{W}^H \mathbf{x}_s(n) d^*(n)\} + \frac{L}{M} \mathbf{1}$ 
14   $\mathbf{S}(n+1) = \text{select}[\mathbf{s}(n+1)]$ 
15 end

```

---

algorithm due to the coupling between the antenna selection matrix and beamforming coefficients. Despite this, both the MMSE-based beamforming problem given in (2.14) and the relaxed antenna selection problem in (2.20) are convex, and their individual adaptive solutions have guaranteed convergence for small values of the step-size parameters [77], even in non-stationary scenarios. Consequently, the convergence of the concurrent update involved in the proposed method is secured for small step size values.

### 2.3.4 Simulation Results

In this subsection, the performance of the proposed MMSE-AJASB algorithm is assessed through numerical simulations and compared with other MMSE-based adaptive beamforming algorithms from the literature. Two versions of the MMSE-based NLMS beamforming algorithms are considered in the simulations: the first using all  $M$  antennas of the subarray (denoted NLMS-M) and the second using only the first  $L$  antennas (termed NLMS-L). It is important to highlight that, as mentioned at this chapter's Introduction (Section 2.1), other AJASB algorithms from the literature do not abide by the limited number of RF chains, and thus, they cannot be fairly compared with the proposed algorithms.

Four examples are presented. In all examples, a uniform linear array (ULA) with half-wavelength spacing between consecutive elements is considered, resulting in a steering vector given by

$$\mathbf{a}(\theta_i) = [1 \ e^{-j\pi \sin \theta_i} \ e^{-j2\pi \sin \theta_i} \ \dots \ e^{-j(M-1)\pi \sin \theta_i}]^T. \quad (2.35)$$

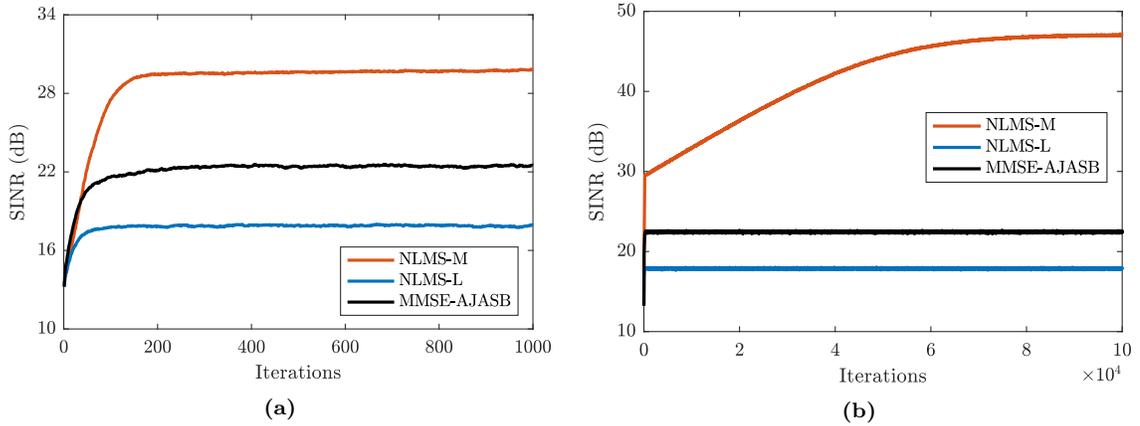
This type of array is chosen without loss of generality since the proposed algorithms can also be effectively applied to other array geometries, such as planar, circular, or non-uniform. For all examples, the algorithms have their beamforming vectors initialized with  $\mathbf{w}(0) = [1 \ 0 \ \dots \ 0]^T$ , which corresponds to an omnidirectional radiation pattern over the azimuth plane. The auxiliary antenna-selection vector is initialized as  $\mathbf{s}(0) = (L/M)\mathbf{1}$  and the antenna-selection matrix  $\mathbf{S}$  is initialized in such a way that the  $L$  leftmost antenna elements are selected. The symbols are generated from a zero-mean unit-variance Gaussian distribution, and the normalized power (relative to the noise variance) of the transmitted signals is 30 dB. Since the interfering signals come from co-channel cells, they are attenuated by 20 dB due to free-space path loss.

The algorithms' performance is assessed in terms of transient and steady-state SINR, which are obtained through Monte Carlo simulations with 200 independent runs. The steady-state values used to obtain the empirical probability density function (EPDF) curves are those from the last iteration of each Monte Carlo run.

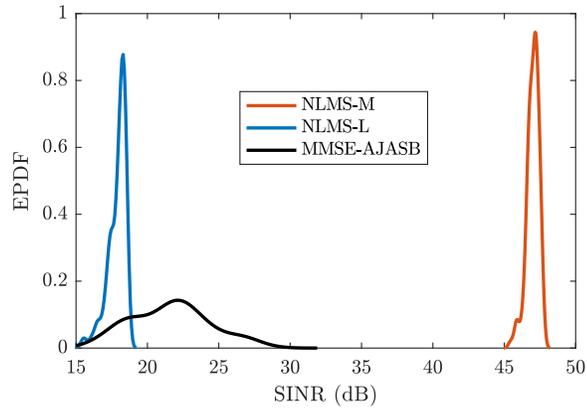
### Example 1

The first example considers a stationary scenario where the subarray allocated to a given intra-cell user (SOI) is equipped with  $M = 64$  antennas and initially with  $L = 4$  RF chains. It is further assumed  $K = 6$  inter-cell users, with the SOI positioned at  $0^\circ$ , and five interferers located at  $-35^\circ$ ,  $-10^\circ$ ,  $12^\circ$ ,  $25^\circ$ , and  $40^\circ$ . The step-size parameters were adjusted to provide similar initial convergence rates, resulting in the following values:  $\mu = 1$  for both the NLMS-M and NLMS-L;  $\mu_w = 1$  and  $\mu_s = 0.01$  for the MMSE-AJASB.

The transient SINR curves are shown in Fig. 2.2. Since the NLMS-M algorithm has a very slow convergence component due to the high number of beamforming coefficients, two figures are shown: Fig. 2.2a shows the results for the first 1000 iterations and Fig. 2.2b shows the results for  $10^5$  iterations. From this figure, one can notice that the proposed algorithm outperforms the NLMS-L by more than 4 dB in terms of average SINR performance. The NLMS-M algorithm obtained the best performance, which is an expected result since the number of RF chains used by this algorithm is 16 times larger than that of the others (64 RF chains versus 4 for both the NLMS-L and proposed



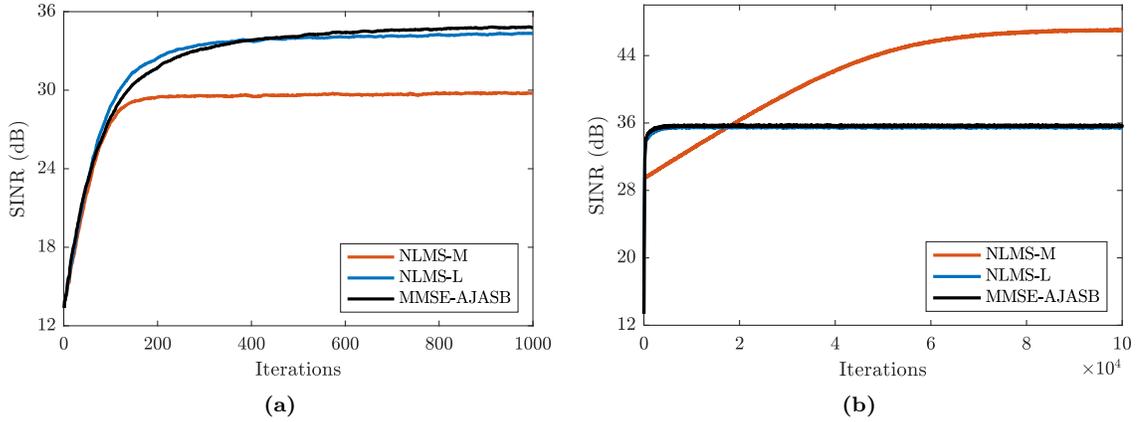
**Figure 2.2:** Example 1. SINR curves of the MMSE-based algorithms for  $M = 64$  and  $L = 4$ .



**Figure 2.3:** Example 1. EPDF curves for the MMSE-based algorithms with  $M = 64$  and  $L = 4$ .

one). Fig. 2.3 shows the EPDF of the steady-state SINR after  $10^5$  iterations. Note that the proposed algorithm achieves higher steady-state SINR values in most of the runs as compared to the NLMS-L. In fact, the best run of the proposed algorithm provided an SINR that is more than 10 dB higher than that of the best run of the NLMS-L.

A second scenario, where the number of RF chains is changed from  $L = 4$  to  $L = 8$ , is also considered. This case is of particular interest since it consists of a reasonably favorable scenario in which the SOI is positioned afar from the interfering signals and the number of RF chains is larger than the number of users. The SINR curves obtained for the MMSE-based algorithms, depicted in Fig. 2.4a, which shows the transient SINR for the first 1000 iterations, and Fig. 2.4b, which shows the transient SINR for  $10^5$  iterations. From these figures, note that the proposed algorithm and the NLMS-L have nearly the same performance. This result is expected since the additional RF chains allow higher levels of interference suppression and the same fading is assumed across all antennas.



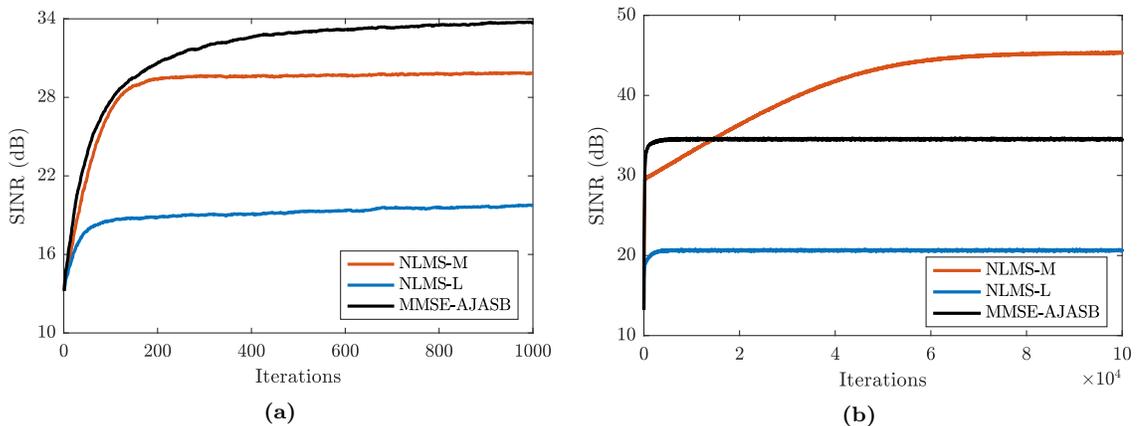
**Figure 2.4:** Example 1. SINR curves of the MMSE-based algorithms for  $M = 64$  and  $L = 8$ .

## Example 2

In this example, the proposed algorithms' performance is assessed in a scenario where the SOI is positioned nearly in the same direction as one interferer. The system is assumed to operate with  $L = 8$  RF chains, and the SOI is located at  $11^\circ$  degrees, while the interferer positions and all other parameters are the same as in Example 1. Here, the transient SINR and the radiation pattern obtained at the last iteration of a Monte Carlo run are assessed.

The SINR curves obtained for the MMSE-based algorithms are shown in Fig. 2.5a and Fig. 2.5b. Compared with the results from Example 1 with  $L = 8$  RF chains (shown in Fig. 2.4), note that the SINR obtained here for both the NLMS-M and the proposed algorithm remain nearly the same, whereas a reduction of nearly 14.5 dB is observed for the NLMS-L. These results demonstrate the robustness of the proposed algorithm to adverse scenarios compared to the beamforming algorithm without the use of antenna selection with the same number of RF chains.

The gains at the SOI's and the interferers' directions, obtained from the radiation patterns at the last iteration of a given Monte Carlo run of each algorithm, are summarized in Table 2.1. These results show that both the NLMS-M and the proposed MMSE-AJASB algorithms are capable of significantly attenuating the interferer at  $12^\circ$  (the one close to the SOI), achieving attenuations of 10.02 dB and 17.64 dB, respectively. In contrast, the NLMS-L achieves an attenuation level of only 1.42 dB.



**Figure 2.5:** Example 2. SINR curves of the MMSE-based algorithms for  $M = 64$  and  $L = 8$ .

**Table 2.1**

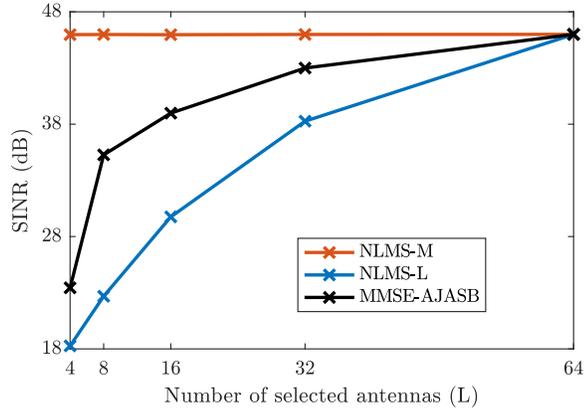
Example 2. Gain (in dBi) at the direction of the SOI and the interferers for the MMSE-based algorithms.

$\theta$	NLMS-M	NLMS-L	MMSE-AJASB
$11^\circ$	-0.01	0.10	0.01
$-35^\circ$	-37.80	-6.89	-12.87
$-10^\circ$	-20.02	-6.74	-13.96
$12^\circ$	-17.64	-1.42	-10.02
$25^\circ$	-20.86	-5.41	-12.79
$40^\circ$	-19.27	-6.53	-6.70

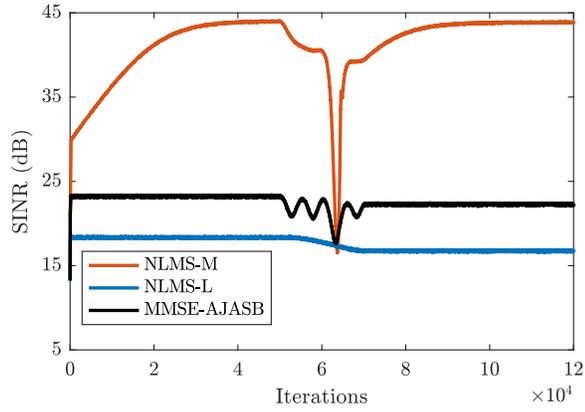
### Example 3

In the third example, the impact of the number of RF chains on the proposed algorithms' steady-state performance is assessed. It is assumed that the BS is equipped with  $M = 64$  antenna elements and  $L = 4, 8, 16, 32,$  and  $64$  RF chains. The steady-state SINR is obtained by averaging the SINR values achieved after  $2 \times 10^6$  iterations of each Monte Carlo run. Note that such a high number of iterations is required since the NLMS-M has a very slow-rate convergence component. It is further assumed  $K = 6$  users, with the SOI positioned at  $5^\circ$ , and five interferers located at  $-42^\circ, -18^\circ, 25^\circ, 44^\circ,$  and  $6^\circ$  (the last one is very close to the SOI). The step sizes for the considered algorithms are:  $\mu = 0.5$  for both the NLMS-M and NLMS-L; and  $\mu_w = 0.5$  and  $\mu_s = 0.01$  for the MMSE-AJASB.

The results are depicted in Fig. 2.6. These results show that the array gain provided by the proposed algorithms increases much more quickly as the number of RF chains increases compared to the beamforming algorithms without the use of antenna selection (i.e. the NLMS-L). Moreover, it is important to highlight that, except for the cases in which  $L < K$ , the SINR levels obtained by the proposed algorithms are closer to those obtained by using the complete array (i.e., NLMS-M) than to those obtained by the arrays with a reduced number of RF chains (NLMS-L).



**Figure 2.6:** Example 3. Steady-state SINR curves for different numbers of RF chains using the MMSE-based algorithms.



**Figure 2.7:** Example 4. SINR curves for the MMSE-based algorithms in the non-stationary scenario.

#### Example 4

In this last example, a non-stationary scenario is investigated. In such a scenario, the SOI, initially positioned at  $15^\circ$ , moves to  $0^\circ$  between iterations 50,000 and 70,000, at a constant angular speed. The interferers are positioned at  $-45^\circ$ ,  $-10^\circ$ ,  $5^\circ$ ,  $25^\circ$ , and  $40^\circ$ , and the corresponding subarray has  $M = 32$  antennas and  $L = 4$  RF chains. The step sizes used are the same as in Exp. 1.

The SINR curves are shown in Fig. 2.7. This figure shows a clear advantage of the proposed algorithm over its counterpart that uses the same number of RF chains. Moreover, by comparing the steady-state SINR before iteration 50,000 with that at the end of the process, it is noted that: i) the SINR obtained for the NLMS-M has no significant difference; ii) the SINR falls by 8.8% for the NLMS-L; and iii) the SINR falls by nearly 4.1% for the proposed MMSE-AJASB algorithm. These results confirm the proposed algorithm's ability to track the SOI and attenuate the interfering signals in non-stationary scenarios.

## 2.4 MVDR-based Adaptive Joint Antenna Selection and Beamforming

In this section, a second alternating-optimization-based adaptive algorithm for solving the JASB problem described in (2.13) is proposed. The algorithm is formulated considering the MVDR criterion [75], which consists of minimizing  $\Phi(\mathbf{w}, \mathbf{S}) = \text{E} [|y(n)|^2]$  along with an additional constraint to ensure unit gain towards the SOI direction. Similar to the MMSE-based AJASB algorithm, an alternate optimization approach is employed. Thus, the solution to the beamforming vector and the antenna selection matrix are alternately obtained while considering the other fixed. The following subsections present the development of the optimization algorithm and simulation results to verify the proposed algorithm's performance.

### 2.4.1 MVDR Beamforming

The linearly constrained minimum variance (LCMV) criterion has been widely used in beamforming optimization due to its convexity property [82]. This criterion is focused on minimizing the variance of the beamforming output  $y(n)$ , subject to a set of linear constraints. Particularly, using a single linear constraint that ensures unit gain towards the SOI direction, the criterion is known as the MVDR problem, which leads to the SINR maximization [66].

For the scenario considered here, the MVDR beamforming problem is formulated as follows:

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w} \\ & \text{subject to} && \mathbf{w}^H \mathbf{S} \mathbf{a}(\theta_0) = 1. \end{aligned} \tag{2.36}$$

Note that, if all antennas are selected (i.e.,  $\mathbf{S} = \mathbf{I}_M$ ), (2.36) becomes the original MVDR problem. The closed-form solution for (2.36), given by [75]

$$\mathbf{w}_o = \frac{(\mathbf{S} \mathbf{R}_x \mathbf{S})^\dagger \mathbf{S} \mathbf{a}(\theta_0)}{\mathbf{a}(\theta_0)^H \mathbf{S} (\mathbf{S} \mathbf{R}_x \mathbf{S})^\dagger \mathbf{S} \mathbf{a}(\theta_0)} \tag{2.37}$$

requires knowledge of  $\mathbf{R}_x$ , which is usually unavailable, and the computation of  $(\mathbf{S} \mathbf{R}_x \mathbf{S})^\dagger$ , which is computationally costly. To circumvent these problems, the same reasoning used for developing the CLMS algorithm [75] is adopted here. Thus, the following beamforming update equation is obtained:

$$\mathbf{w}(n+1) = \mathbf{P}_a(n) [\mathbf{w}(n) - \mu_w y^*(n) \mathbf{x}_s(n)] + \mathbf{f}_a(n) \tag{2.38}$$

where  $\mu_w$  is the step size,

$$\mathbf{P}_a(n) = \left[ \mathbf{I}_M - \frac{\mathbf{S}(n)\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0)\mathbf{S}(n)}{\mathbf{a}^H(\theta_0)\mathbf{S}(n)\mathbf{a}(\theta_0)} \right] \quad (2.39)$$

and

$$\mathbf{f}_a(n) = \frac{\mathbf{S}(n)\mathbf{a}(\theta_0)}{\mathbf{a}^H(\theta_0)\mathbf{S}(n)\mathbf{a}(\theta_0)}. \quad (2.40)$$

The normalized version of the CLMS [55], termed constrained CNLMS algorithm, provides faster convergence and lower misadjustment as compared with the original CLMS. Thus, taking into account this normalization idea, the CNLMS algorithm is obtained by dividing  $\mu_w$  by  $\mathbf{x}_s^H(n)\mathbf{P}_a(n)\mathbf{x}_s(n)$  in (2.38). Thereby,

$$\mathbf{w}(n+1) = \mathbf{P}_a(n) \left[ \mathbf{w}(n) - \frac{\mu_w y^*(n)\mathbf{x}_s(n)}{\mathbf{x}_s^H(n)\mathbf{P}_a(n)\mathbf{x}_s(n)} \right] + \mathbf{f}_a(n). \quad (2.41)$$

## 2.4.2 MVDR Antenna Selection

The proposed MVDR antenna selection method is based on finding the antenna subset that minimizes the variance of the beamforming output signal, without suppressing the gain towards the SOI direction. To this end, the following optimization problem:

$$\begin{aligned} & \underset{\mathbf{S}}{\text{minimize}} && \mathbf{w}^H \mathbf{S} \mathbf{R}_x \mathbf{S} \mathbf{w} \\ & \text{subject to} && [\mathbf{S}]_{m,m} \in \{0, 1\} \\ & && \text{Tr}(\mathbf{S}) = L \\ & && \text{Re}\{\mathbf{w}^H \mathbf{S} \mathbf{a}(\theta_0)\} = 1. \end{aligned} \quad (2.42)$$

Note that only the real part of  $\mathbf{w}^H \mathbf{S} \mathbf{a}(\theta_0)$  is constrained in (2.42), resulting in a loosen constraint as compared with the corresponding one in (2.36). This procedure can be done since such a constraint is sufficient to prevent  $\mathbf{S}$  from suppressing the SOI and the unit gain in the SOI direction is enforced by the beamforming update.

Similarly to (2.19), the MVDR antenna selection problem in (2.42) is also an NP-hard problem due to the binary constraints. Therefore, these constraints are relaxed by allowing  $[\mathbf{S}]_{m,m}$  to be a real number. Then, considering the auxiliary vector  $\mathbf{s}$  (diagonal vector of  $\mathbf{S}$ ) and matrix  $\mathbf{W} = \text{diag}(\mathbf{w})$ ,

the resulting optimization problem becomes

$$\begin{aligned}
& \underset{\mathbf{s} \in \mathbb{R}^M}{\text{minimize}} && \mathbf{s}^T \mathbf{W}^H \mathbf{R}_x \mathbf{W} \mathbf{s} \\
& \text{subject to} && \mathbf{1}^T \mathbf{s} = L \\
& && \text{Re}\{\mathbf{s}^T \mathbf{W}^H \mathbf{S} \mathbf{a}(\theta_0)\} = 1.
\end{aligned} \tag{2.43}$$

The second constraint in (2.43) is derived from the last constraint in (2.42) by considering the idempotent property of  $\mathbf{S}$ . In fact, matrix  $\mathbf{S}$  is not treated as an optimization variable in (2.43) and is used for ensuring unit gain for the SOI considering the *a priori* antenna selection.

To find the optimal solution to (2.43), the Lagrangian function is written as

$$\mathcal{L}(\mathbf{s}, \lambda, \alpha) = \mathbf{s}^T \mathbf{R}_{\text{rw}} \mathbf{s} + \lambda(\mathbf{1}^T \mathbf{s} - L) + \alpha(\mathbf{s}^T \mathbf{r}_a - 1) \tag{2.44}$$

where  $\lambda$  and  $\alpha$  are Lagrange multipliers,

$$\mathbf{R}_{\text{rw}} = \text{Re}\{\mathbf{W}^H \mathbf{R}_x \mathbf{W}\} \tag{2.45}$$

and

$$\mathbf{r}_a = \text{Re}\{\mathbf{W}^H \mathbf{S} \mathbf{a}(\theta_0)\}. \tag{2.46}$$

Next, taking the gradient of (2.44) with respect to  $\mathbf{s}$ , i.e.,

$$\nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda, \alpha) = 2\mathbf{R}_{\text{rw}} \mathbf{s} + \lambda \mathbf{1} + \alpha \mathbf{r}_a \tag{2.47}$$

and setting it equal to zero yields

$$\mathbf{s}_o = -\frac{1}{2} \mathbf{R}_{\text{rw}}^{-1} (\lambda \mathbf{1} + \alpha \mathbf{r}_a). \tag{2.48}$$

Finally, applying the constraints of (2.43) to (2.48) to obtain  $\lambda$  and  $\alpha$ , and substituting them back into (2.48) yields the optimal solution to (2.43), given as

$$\begin{aligned}
\mathbf{s}_o = & \left[ \frac{\mathbf{R}_{\text{rw}}^{-1}}{\mathbf{1}^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{1} \mathbf{r}_a^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{r}_a - \mathbf{1}^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{r}_a \mathbf{r}_a^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{1}} \right] \\
& \times [(L \cdot \mathbf{r}_a^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{r}_a - \mathbf{1}^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{r}_a) \mathbf{1} + (\mathbf{1}^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{1} - L \cdot \mathbf{r}_a^T \mathbf{R}_{\text{rw}}^{-1} \mathbf{1}) \mathbf{r}_a].
\end{aligned} \tag{2.49}$$

Similarly to the discussion in Section 2.3.2, the optimal solution given in (2.49) cannot be directly

used for obtaining the *a posteriori* selection matrix. In this case, considering the minimum disturbance criterion as in (2.28), the antenna selection is performed by setting the diagonal elements of  $\mathbf{S}$  corresponding to the  $L$  highest values of  $\mathbf{s}_o$  to one, while the remaining  $M - L$  diagonal elements are set to zero. Moreover, the evaluation of the closed-form solution in (2.49) requires knowledge of  $\mathbf{R}_x$  and demands the computation of  $\mathbf{R}_{rw}$  and its inverse. Therefore, an adaptive procedure based on the CLMS algorithm [55] for obtaining  $\mathbf{s}_o$  is proposed here. This procedure is derived by applying the steepest-descent method to (2.43), resulting in

$$\begin{aligned} \mathbf{s}(n+1) &= \mathbf{s}(n) - \mu_s \nabla_{\mathbf{s}} \mathcal{L}(\mathbf{s}, \lambda, \alpha) \\ &= [\mathbf{I}_M - 2\mu_s \mathbf{R}_{rw}] \mathbf{s}(n) - \mu_s \lambda \mathbf{1} - \mu_s \alpha \mathbf{r}_a. \end{aligned} \quad (2.50)$$

where  $\mu_s$  is the step size.

The Lagrange multipliers  $\lambda$  and  $\alpha$  are found by reinforcing the constraints of (2.43) with  $\mathbf{s}(n+1)$ , i.e., by solving the following linear system:

$$\begin{cases} \mathbf{1}^T \mathbf{s}(n+1) = L \\ \mathbf{s}^T(n+1) \mathbf{r}_a = 1. \end{cases} \quad (2.51)$$

The solution of (2.51) leads to

$$\lambda = \frac{[\mathbf{s}^T(n) - 2\mu_s \mathbf{s}^T(n) \mathbf{R}_{rw}] (\|\mathbf{r}_a\|^2 \mathbf{1} - \mathbf{r}_a) - L \|\mathbf{r}_a\|^2 + 1}{\mu_s (\|\mathbf{r}_a\|^2 M - 1)} \quad (2.52)$$

and

$$\alpha = \frac{[\mathbf{s}^T(n) - 2\mu_s \mathbf{s}^T(n) \mathbf{R}_{rw}] (M \mathbf{r}_a - \mathbf{1}) + L - M}{\mu_s (\|\mathbf{r}_a\|^2 M - 1)} \quad (2.53)$$

which are then plugged into (2.50) to obtain

$$\mathbf{s}(n+1) = \mathbf{P}_g [\mathbf{I}_M - 2\mu_s \mathbf{R}_{rw}] \mathbf{s}(n) + \mathbf{f}_g \quad (2.54)$$

with

$$\mathbf{P}_g = \mathbf{I}_M - \frac{\|\mathbf{r}_a\|^2 \mathbf{1} \mathbf{1}^T - \mathbf{1} \mathbf{r}_a^T - \mathbf{r}_a \mathbf{1}^T + M \mathbf{r}_a \mathbf{r}_a^T}{\|\mathbf{r}_a\|^2 M - 1} \quad (2.55)$$

and

$$\mathbf{f}_g = \frac{(L \|\mathbf{r}_a\|^2 - 1) \mathbf{1} + (M - L) \mathbf{r}_a}{\|\mathbf{r}_a\|^2 M - 1}. \quad (2.56)$$

Finally, one can further evaluate (2.54) using the instantaneous estimates of  $\mathbf{R}_{rw}$  and  $\mathbf{r}_a$ , which are given by  $\text{Re}\{\mathbf{W}^H(n) \mathbf{x}_s(n) \mathbf{x}_s^H(n) \mathbf{W}(n)\}$  and  $\text{Re}\{\mathbf{W}^H(n) \mathbf{S}(n) \mathbf{a}(\theta_0)\}$ , respectively.

---

**Algorithm 2.2:** Summary of the MVDR-AJASB Algorithm

---

```

1 Initialize:  $\mu_w, \mu_s, \varepsilon$ 
2  $\mathbf{P}_s = \left[ \mathbf{I}_M - \frac{\mathbf{1}\mathbf{1}^T}{M} \right]$ 
3  $\mathbf{w}(0) = [1 \ 0 \ \dots \ 0]^T$ 
4  $\mathbf{s}(0) = \frac{L}{M} \mathbf{1}$ 
5  $\mathbf{S}(0) = \begin{bmatrix} \mathbf{I}_L & \mathbf{0}_{L \times (M-L)} \\ \mathbf{0}_{(M-L) \times L} & \mathbf{0}_{(M-L) \times (M-L)} \end{bmatrix}$ 
6 for  $n = 0, 1, 2, \dots$  do
7    $y(n) = \mathbf{w}^H(n) \mathbf{x}_s(n)$ 
8    $\mathbf{P}_a(n) = \left[ \mathbf{I}_M - \frac{\mathbf{S}(n) \mathbf{a}(\theta_0) \mathbf{a}^H(\theta_0) \mathbf{S}(n)}{\mathbf{a}^H(\theta_0) \mathbf{S}(n) \mathbf{a}(\theta_0)} \right]$ 
9    $\mathbf{f}_a(n) = \frac{\mathbf{S}(n) \mathbf{a}(\theta_0)}{\mathbf{a}^H(\theta_0) \mathbf{S}(n) \mathbf{a}(\theta_0)}$ 
10   $\mathbf{w}(n+1) = \mathbf{P}_a(n) \left[ \mathbf{w}(n) - \frac{\mu_w y^*(n) \mathbf{x}_s(n)}{\mathbf{x}_s^H(n) \mathbf{P}_a(n) \mathbf{x}_s(n) + \varepsilon} \right] + \mathbf{f}_a(n)$ 
11   $\mathbf{W}(n) = \text{diag}[\mathbf{w}(n+1)]$ 
12   $\mathbf{R}_{\text{rw}} = \text{Re}\{\mathbf{W}^H(n) \mathbf{x}_s(n) \mathbf{x}_s^H(n) \mathbf{W}(n)\}$ 
13   $\mathbf{r}_a = \text{Re}\{\mathbf{W}^H(n) \mathbf{S}(n) \mathbf{a}(\theta_0)\}$ 
14   $\mathbf{P}_g = \mathbf{I}_M - \frac{\|\mathbf{r}_a\|^2 \mathbf{1}\mathbf{1}^T - \mathbf{r}_a \mathbf{r}_a^T - \mathbf{r}_a \mathbf{1}^T + M \mathbf{r}_a \mathbf{r}_a^T}{\|\mathbf{r}_a\|^2 M - 1}$ 
15   $\mathbf{f}_g = \frac{(L \|\mathbf{r}_a\|^2 - 1) \mathbf{1}_{M+(M-L)\mathbf{r}_a}}{\|\mathbf{r}_a\|^2 M - 1}$ 
16   $\mathbf{s}(n+1) = \mathbf{P}_g [\mathbf{I}_M - 2\mu_s \mathbf{R}_{\text{rw}}] \mathbf{s}(n) + \mathbf{f}_g$ 
17   $\mathbf{S}(n+1) = \text{select}[\mathbf{s}(n+1)]$ 
18 end

```

---

### 2.4.3 MVDR-based AJASB Algorithm

The proposed MVDR-based AJASB (MVDR-AJASB) algorithm consists of an iterative process that concurrently uses the beamforming strategy described in Section 2.4.1 and the antenna selection strategy from Section 2.4.2. Such an algorithm is summarized in Algorithm 2.2.

From Algorithm 2.2, one can notice that, in the first step, the received signal  $\mathbf{x}_s(n)$  is used for computing the beamforming output  $y(n)$ . Next, the beamforming vector  $\mathbf{w}(n)$  is updated through the CNLMS-based algorithm presented in Section 2.4.1. Once the *a posteriori* beamforming vector  $\mathbf{w}(n+1)$  is calculated, it is further used to compute  $\mathbf{s}(n+1)$  considering the algorithm proposed in Section 2.4.2. Finally, the antenna-selection matrix is evaluated by selecting the antennas corresponding to the  $L$  largest components of  $\mathbf{s}(n+1)$ , denoted in Algorithm 2.2 by the *select*  $[\mathbf{s}(n+1)]$  operation, as defined in Sec. 2.3.3. A regularization parameter  $\varepsilon$  has also been included in the update equations of  $\mathbf{w}(n)$  to prevent divisions by very small values.

Regarding the computational burden, the MVDR-based AJASB requires  $3M + 36L - 4$  sums,  $M + 39L + 18$  multiplications, and three divisions, plus at most  $(M - L)L$  real comparisons for the

*select*[ $\cdot$ ] operation. In contrast, the MVDR-based CNLMS beamforming requires  $24M - 4$  sums,  $24M + 8$  multiplications, and two divisions. Since  $M \gg L$ , the proposed MVDR-based AJASB also has lower computational complexity than its CNLMS counterpart.

Similar to the MMSE-AJASB algorithm in Section 2.3.3, the stochastic modeling required for an accurate convergence analysis is hard to carry out due to the coupling between the antenna selection matrix and beamforming vector. However, both CLMS and CNLMS algorithms considered in the proposed approach have guaranteed convergence when the step-size parameters are small [77]. Thereby, the convergence of the MVDR-AJASB algorithm is also ensured for small values of step size.

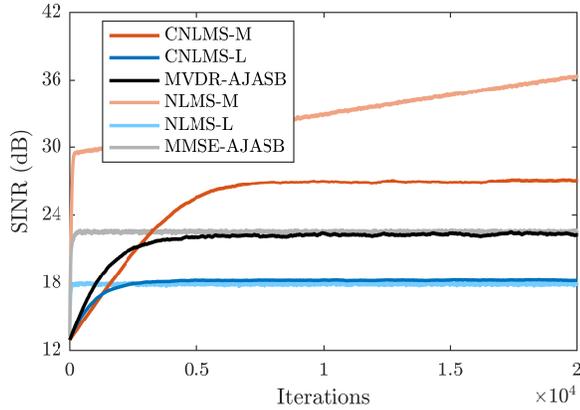
#### 2.4.4 Simulation Results

In this subsection, the performance of the proposed MVDR-AJASB algorithm is assessed and compared with that of the MVDR-based CNLMS beamforming algorithm without antenna selection. Two versions of the CNLMS beamforming algorithm are considered: one using all  $M$  antennas (CNLMS-M) and the other using only the first  $L$  antennas (CNLMS-L). As for the MMSE-based algorithm, it is important to highlight that other AJASB algorithms available in the open literature are not considered in the comparisons since they not abide by the limited number of RF chains, and thus, cannot be fairly compared with the proposed algorithm.

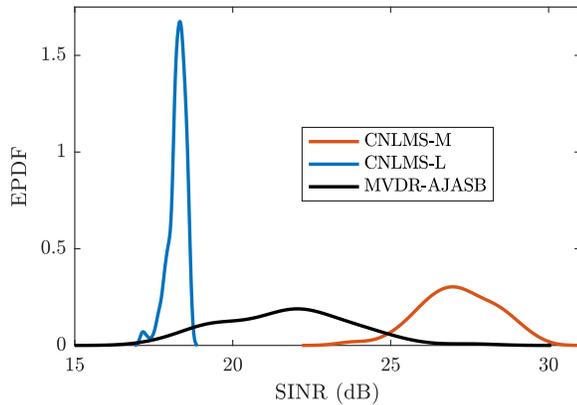
The same four examples presented in Section 2.3.4 are considered here. It is assumed a ULA with half-wavelength spacing between consecutive elements. The algorithms have their beamforming vectors initialized with  $\mathbf{w}(0) = [1 \ 0 \ \dots \ 0]^T$ , which corresponds to an omnidirectional radiation pattern over the azimuth plane. The auxiliary antenna-selection vector is initialized as  $\mathbf{s}(0) = (L/M)\mathbf{1}$  and the antenna-selection matrix  $\mathbf{S}$  is initialized so that the  $L$  leftmost antenna elements are selected. The symbols are generated from a zero-mean unit-variance Gaussian distribution, and the normalized power (relative to the noise variance) of the transmitted signals is 30 dB. Since the interfering signals come from co-channel cells, they are attenuated by 20 dB due to free-space path loss. As for the MMSE-based algorithms, the algorithms' performance is assessed in terms of the transient and steady-state SINR obtained from Monte Carlo simulations with 200 independent runs.

##### Example 1

This example considers a stationary environment where the subarray allocated to a given intra-cell user (SOI) is equipped with  $M = 64$  antennas and  $L = 4$  RF chains. It is assumed  $K = 6$  inter-cell users, with the SOI positioned at  $0^\circ$ , and five interferers located at  $-35^\circ$ ,  $-10^\circ$ ,  $12^\circ$ ,  $25^\circ$ , and  $40^\circ$ .



**Figure 2.8:** Example 1. SINR curves of the MVDR-based algorithms for  $M = 64$  and  $L = 4$ .

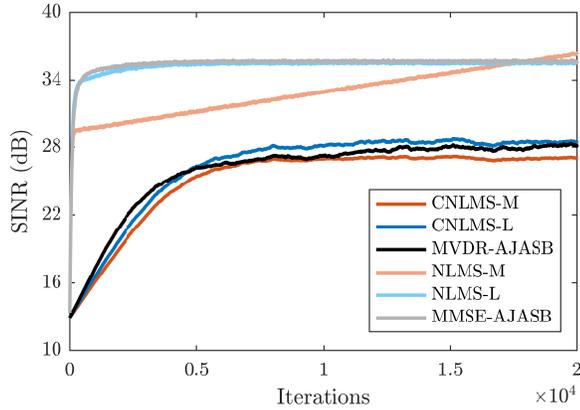


**Figure 2.9:** Example 1. EPDF curves for the MVDR-based algorithms with  $M = 64$  and  $L = 4$ .

The step-size parameters were adjusted to provide similar initial convergence rates, resulting in the following values:  $\mu = 0.002$  for both the CNLMS-M and CNLMS-L; and  $\mu_w = 0.002$  and  $\mu_s = 0.001$  for the MVDR-AJASB.

Fig. 2.8 shows the resulting SINR curves. For comparison, the results obtained with the MMSE-based algorithms are also plotted. From Fig. 2.8, it is observed that the proposed algorithm outperforms the CNLMS-L by nearly 4 dB in terms of average SINR. As expected, the CNLMS-M presents better performance than the other algorithms (due to the much larger number of RF chains used). Also, note that the MMSE-based algorithms have a much faster convergence rate. The EPDF of the steady-state SINR is shown in Fig. 2.9, from which one can observe that the proposed algorithm outperforms the CNLMS-L in most cases.

A second scenario, with the number of RF chains  $L = 8$ , is also considered. This case represents a fairly favorable scenario in which the SOI is positioned afar from the interfering signals and the



**Figure 2.10:** Example 1. SINR curves of the MVDR-based algorithms for  $M = 64$  and  $L = 8$ .

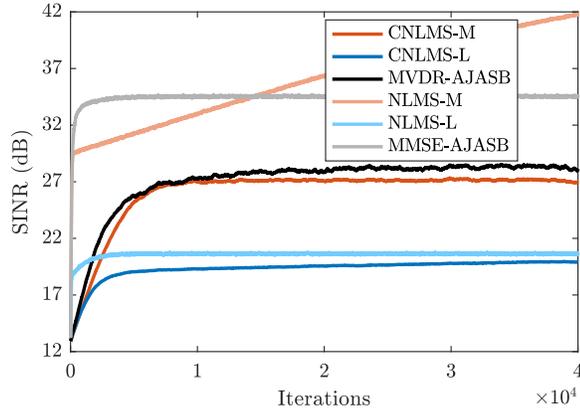
number of RF chains is larger than the number of users. The SINR curves are shown in Fig. 2.10. Note that the proposed algorithm can attain performance similar to that of the CNLMS-M while operating with 1/8 of RF chains.

### Example 2

This example assesses the proposed algorithms' performance in a scenario where the SOI is positioned close to one interferer. The system is assumed to operate with  $L = 8$  RF chains, the SOI is located at  $11^\circ$ , and the interferers' positions and all other parameters are the same as in Example 1. Here, the transient SINR and the radiation pattern obtained at the last iteration of a Monte Carlo run are assessed.

The SINR results are shown in Fig. 2.11. Note that the proposed algorithm is much more robust to unfavourable scenarios (such as when the SOI and interferer are closely positioned) than the CNLMS-L. When compared to the results from Example 1 with  $L = 8$  RF chains, shown in Fig. 2.10, one can observe that the proposed algorithm and the CNLMS-M can sustain nearly the same performance, while here, the CNLMS-L attains SINR more than 8 dB lower as compared to the favourable scenario in the Example 1.

The gains at all users' directions, obtained from the radiation patterns at the last iteration of a given Monte Carlo run of each algorithm, are summarized in Table 2.2. These results show that the attenuation at  $12^\circ$  (i.e., the direction of the interferer close to the SOI) is 8.78 dB for the CNLMS-M, 6.36 dB for the proposed MVDR-AJASB, and only 1.01 dB for the CNLMS-L. Additionally, note that all MVDR-based algorithms attain unit gain (0 dBi) in the SOI direction, as required by the MVDR criterium.



**Figure 2.11:** Example 2. SINR curves of the MVDR-based algorithms for  $M = 64$  and  $L = 8$ .

**Table 2.2**

Example 2. Gain (in dBi) at the direction of the SOI and the interferers for the MVDR-based algorithms.

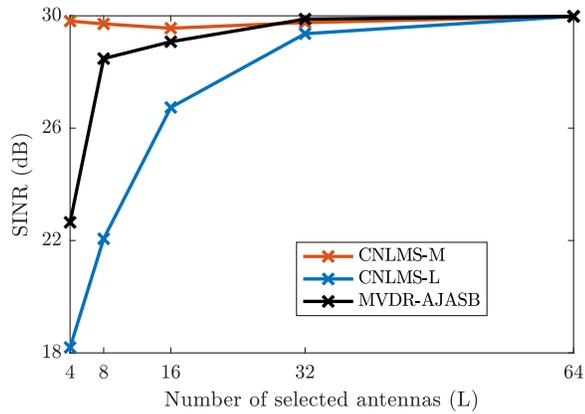
$\theta$	CNLMS-M	CNLMS-L	MVDR-AJASB
$11^\circ$	0.00	0.00	0.00
$-35^\circ$	-7.31	-6.06	-7.40
$-10^\circ$	-11.01	-8.57	-11.71
$12^\circ$	-8.78	-1.01	-6.36
$25^\circ$	-11.96	-6.53	-9.11
$40^\circ$	-3.88	-4.22	-4.49

### Example 3

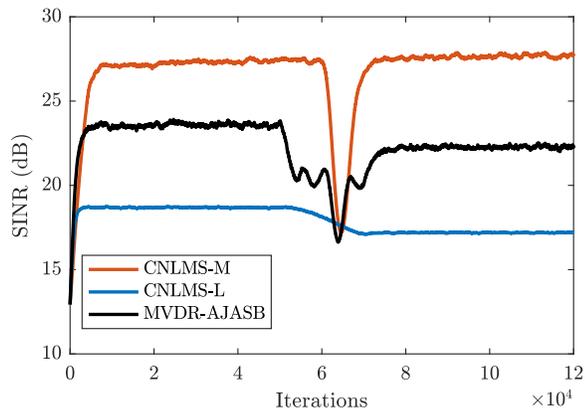
This example assesses the impact of the number of RF chains on the proposed algorithm's steady-state performance. A BS equipped with  $M = 64$  antenna elements is assumed, and the steady-state SINR is assessed when considering  $L = 4, 8, 16, 32,$  and  $64$  RF chains. The steady-state SINR is obtained by averaging the SINR values achieved after  $2 \times 10^6$  iterations of each Monte Carlo run. It is further assumed  $K = 6$  users, with the SOI positioned at  $5^\circ$ , and five interferers located at  $-42^\circ, -18^\circ, 25^\circ, 44^\circ,$  and  $6^\circ$  (the last one is positioned close to the SOI). The step sizes are chosen as  $\mu = 0.002$  for both the CNLMS-M and CNLMS-L, and  $\mu_w = 0.002$  and  $\mu_s = 0.001$  for the MVDR-AJASB. The results, depicted in Fig. 2.12, show that the MVDR-AJASB has nearly the same performance as the CNLMS-M when  $L \geq 16$  RF chains, meaning that almost the same performance can be obtained by using a number as low as one-fourth of RF chains.

### Example 4

The last example investigates the performance of the proposed algorithm under a non-stationary scenario. In such a scenario, the SOI is initially positioned at  $15^\circ$  and moves to  $0^\circ$  between iterations



**Figure 2.12:** Example 3. Steady-state SINR curves for different numbers of RF chains using the MVDR-based algorithms.



**Figure 2.13:** Example 4. SINR curves for the MVDR-based algorithms in the non-stationary scenario.

50,000 and 70,000 with constant angular speed. The interferers are positioned at  $-45^\circ$ ,  $-10^\circ$ ,  $5^\circ$ ,  $25^\circ$ , and  $40^\circ$ , and the corresponding subarray has  $M = 32$  antennas and  $L = 4$  RF chains, while the step sizes are the same as in Exp. 1.

Fig. 2.13 shows the obtained SINR curves. Similar to the MMSE-AJASB, the MVDR-AJASB algorithm has a clear advantage over its counterpart using the same number of RF chains (i.e., the CNLMS-L). Comparing the steady-state SINR before iteration 50,000 with that at the end of the process, one can note that the SINR falls by 8% for the CNLMS-L and less nearly 5.7% for the proposed MVDR-AJASB algorithm. Such results confirm the ability of the proposed algorithm to track the SOI and attenuate the interfering signals in non-stationary scenarios.

## 2.5 Concluding Remarks

In this chapter, two new adaptive algorithms to solve the JASB problem for uplink reception in mobile communication systems have been proposed. The first algorithm is based on an iterative alternating optimization procedure, in which the beamforming vector and the antenna selection matrix are computed by using the MMSE criterion. Next, a similar alternating optimization procedure based on the MVDR criterion is used to derive the second algorithm. Numerical simulation results showed that the proposed algorithms led to high SINR values in adverse scenarios, such as those with fewer RF chains than users, with interfering signals located at the SOI's vicinity, and non-stationary environments. Thereby, based on the obtained results, the proposed algorithms' effectiveness has been demonstrated.

It is worth noticing that the proposed MMSE- and MVDR-based AJASB algorithms require the transmission of a training sequence and the DOA estimation, respectively. Although the effects of these problems have been overlooked in this work, they have an important influence on practical deployments. Therefore, further investigation on this matter is recommended.

## Chapter 3

# Hybrid Beamforming for Narrowband mmWave Massive MIMO Systems

### 3.1 Introduction

The design of hybrid beamforming can be traced back to the work of Zhang *et al.* [59], where the authors proposed a phase-shifter-based hybrid beamforming for conventional MIMO. The design was intended as an alternative to antenna selection schemes, such as those studied in Chapter 2, since those schemes suffer from a performance loss caused by the low array gain when only a small number of antennas are selected. However, only after the seminal work of Ayach *et al.* [24], where authors broke the ground in leveraging the mmWave channel characteristics, hybrid beamforming started gaining significant attention in the mmWave massive MIMO community. Initially, the research works on hybrid beamforming design for mmWave massive MIMO considered only narrowband single-user (SU-MIMO) systems [24], [83]–[87]. Solving the hybrid precoder and combiner design problem is very challenging as it requires a joint optimization over multiple matrix variables under the non-linear and non-convex constant modulus constraint imposed by the phase-shifters on analog precoders and combiners. To solve the design problem, authors resort to sub-optimal techniques, which temporarily decouple either the precoder and combiner design, or the analog and digital beamforming design, and solve these problems separately.

In [24], Ayach *et al.* exploited the mmWave channels' sparse-scattering structure to formulate the hybrid beamforming design as a sparsity constrained matrix reconstruction problem. The authors sought to design the hybrid beamforming to maximize its spectral efficiency. By decoupling the precoder and combiner design, they showed that this problem could be cast as an approximation problem that minimizes the Frobenius norm of the difference between the optimal unconstrained beamforming (i.e. that finding the best projection of the optimal unconstrained beamforming onto the set of feasible hybrid beamforming) [24]. This problem consists of a matrix-factorization problem and is solved using the compressive-sensing-based orthogonal matching pursuit (OMP) algorithm. The solution requires an exhaustive search over an analog beamforming codebook (e.g. steering vectors), which may lead to high complexity when the codebook size increases or limit the performance when it decreases. Inspired by the same matrix-factorization problem, the authors in [83] also presented two alternating-minimization (AltMin) approaches for hybrid beamforming designs based on manifold optimization (MO) and phase-extraction (PE). The MO-based design can attain near-optimal performance, but its computational complexity limits its practical implementation. In contrast, the PE-based design trades off a slight performance loss for a lower computational complexity. Despite this performance loss, the PE-AltMin algorithm performs better than most of the currently existing algorithms.

In [84], [85], the authors proposed simple yet efficient PE-based hybrid beamforming designs, in which the analog beamforming is simply obtained by extracting the phases of the optimal unconstrained precoder. Specifically, in [84], the baseband beamforming is a constant block diagonal matrix (e.g., identity matrix), while in [85], the baseband precoder is obtained by the least-squares solution, aiming to approximate the optimal unconstrained precoder. Overall, the hybrid beamforming designs in [24], [83]–[85] are based on approximating the hybrid beamforming to its optimal unconstrained counterpart and require *a priori* computation of the optimal unconstrained beamforming.

Unlike the approximation-based methods above, [86] and [87] design the hybrid beamforming by directly tackling the rate maximization problem. In [86], authors propose a joint hybrid precoder and combiner design for the sum-rate maximization. Their method ensures orthogonality of the effective spatial multiplexing channel, leading to higher sum-rate performance. However, as in [24], this method also requires exhaustive search over analog precoder/combiner codebooks. In [87], the hybrid beamforming is formulated by extracting the contribution of each element of the analog beamforming to the system's rate and solved by using an iterative coordinate descent algorithm to update each of these elements successively.

The promising results obtained for SU-MIMO hybrid beamforming encouraged researchers to investigate its use in multiuser (MU-MIMO) systems [61], [88]–[97]. Early works on MU-MIMO hybrid beamforming design have considered the single data stream transmission for single-antenna [88], [89] and multi-antenna [90]–[93] mobile users. In these works, the analog beamforming is often designed to harvest the array gain by either using the equal gain transmission (EGT) [88], performing an exhaustive search over an analog beamforming codebook [90], [91], or extracting the phases of the channel’s principal singular vector pair [93], without accounting for the multiuser interference (MUI). A different approach has been taken in [89] and [92], where authors proposed analog beamforming designs to simultaneously maximize each user’s equivalent channel gain and mitigate the MUI using Gram-Schmidt-based approaches. On the other hand, the digital precoder is often designed by the zero-forcing (ZF) method. The ZF precoder allows a complete diagonalization of the effective channel, suppressing all inter-user interference. However, when multi-antenna receivers are considered, the complete channel diagonalization at the transmitter using the ZF precoder is sub-optimal since the receiver can also employ coordinate processing to help reduce the interference [98]. The alternative is a generalization of the ZF processing for multi-antenna users, known as block-diagonalization (BD) [98].

The MU-MIMO hybrid beamforming design for multiple data stream transmission has been considered in [61], [94]–[97]. In [94], the authors introduced the hybrid BD (Hy-BD) scheme, in which the analog combiner and precoder are designed by using exhaustive search over a discrete Fourier transform (DFT) codebook and EGT, respectively, in order to harvest the array gain. The digital precoder and combiner are designed by the BD method over the low-dimensional effective baseband channel. In [95], the authors presented a subspace projection aided BD scheme for fully-digital beamforming design using the angle of departure information and used an iterative matrix decomposition to obtain the hybrid precoder and combiner from the fully-digital solution. Authors in [96] devised a two-stage method, in which first, the analog precoder and combiner are designed to maximize the capacity of the effective baseband channel, and then, the digital precoder and combiner are obtained from the traditional BD. In [97], the authors have designed the analog beamforming using an approach similar to [87] and obtained digital beamforming from the traditional BD. Finally, the author in [61] introduced the hybrid regularized channel diagonalization (HRCDD) that uses exhaustive search over the array response vector codebook for designing the analog combiner, EGT for the analog precoder, and regularized channel diagonalization (RCD) for the digital precoder and combiner.

Note that the analog beamforming designs in [61], [88], [90], [91], [93]–[95], and [97] seek to

provide higher effective baseband channel gain, thus increasing the desired signal’s power at the receiver, but do not attempt to reduce inter-user and intra-user interferences. The interference mitigation is performed only at the digital beamforming stage. Moreover, although the methods in [89] and [92] can reduce MUI at the analog beamforming stage, they are designed for single data stream transmission and, therefore, do not account for intra-user interference. Furthermore, the methods in [95] and [61] rely on the perfect knowledge of the steering vectors of all propagation paths.

This chapter presents a novel joint hybrid precoder and combiner design method for mmWave massive MIMO systems. The proposed method consists of a two-stage approach, in which the analog and digital beamformers are designed separately. In the first stage, the analog beamforming is designed using a constrained channel decomposition approach, which provides higher effective baseband channel gain while also attempting to reduce inter-user and intra-user interference. Two distinct digital beamforming designs are considered in the second stage. For SU-MIMO systems, the digital precoder and combiner are obtained from the optimal SVD solution; and for MU-MIMO, the digital precoder and combiner are obtained from the RCD method. Simulation results show that the proposed design is able to attain near-optimal spectral efficiency and sum-rate, outperforming most of the hybrid beamforming designs in the literature.

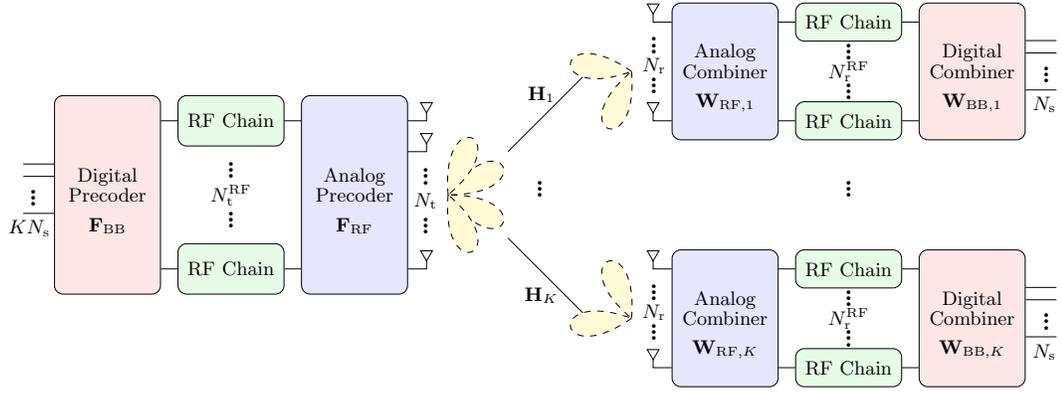
This chapter is organized as follows: Section 3.2 describes the system and channel models, and the problem formulation; Section 3.3 introduces the novel hybrid beamforming design for narrowband mmWave massive MIMO systems based on the proposed constrained channel decomposition approach; Section 3.4 demonstrates the numerical simulation results for different system’s setups and compares the proposed approaches with other designs in the literature; finally, Section 3.5 presents the concluding remarks of this chapter.

## 3.2 System Model and Problem Formulation

This section presents the system model, the channel model, and the problem formulation.

### 3.2.1 System Model

Here, the downlink of an MU-MIMO narrowband massive MIMO system with a hybrid beamforming structure, as depicted in Fig. 3.1, is considered. It is assumed that a base station having  $N_t$  antenna elements and  $N_t^{\text{RF}}$  RF chains simultaneously communicates with  $K$  users. Each user is equipped with  $N_r$  antenna elements and  $N_r^{\text{RF}}$  RF chains and supports  $N_s$  data streams. To ensure effective



**Figure 3.1:** System diagram for the downlink of a multiuser massive MIMO system with hybrid beamforming.

multi-stream communication, the system must satisfy  $N_s \leq N_r^{\text{RF}} \leq N_r$  and  $KN_s \leq N_t^{\text{RF}} \leq N_t$  [87]. At the base station, the hybrid precoder consists of an analog precoder matrix  $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t \times N_t^{\text{RF}}}$  and a digital precoder matrix  $\mathbf{F}_{\text{BB}} = [\mathbf{F}_{\text{BB},1} \cdots \mathbf{F}_{\text{BB},k} \cdots \mathbf{F}_{\text{BB},K}] \in \mathbb{C}^{N_t^{\text{RF}} \times KN_s}$ , where  $\mathbf{F}_{\text{BB},k} \in \mathbb{C}^{N_t^{\text{RF}} \times N_s}$  is the digital precoder matrix for the  $k$ -th user. The analog precoder is implemented using phase shifters and is subjected to a constant modulus constraint (i.e.  $|\mathbf{F}_{\text{RF}}(m,n)| = 1/\sqrt{N_t} \forall m,n$ ). The symbol vector  $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_K^T]^T \in \mathbb{C}^{KN_s \times 1}$ , where  $\mathbf{s}_k \in \mathbb{C}^{N_s \times 1}$  refers to the  $k$ -th user's symbols, is assumed independent with unity-variance components, i.e.  $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{KN_s}$ , and is precoded in the digital and analog domains before being transmitted, yielding the following transmitted signal

$$\mathbf{x} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\mathbf{s} \quad (3.1)$$

where  $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$  and  $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = P_{\text{T}}$  is the total transmitted power.

Assuming that the signal is transmitted over a narrowband flat-fading channel  $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ , where  $\mathbf{H}_k$  represents the fast-fading channel matrix for the  $k$ -th user, the signal received at the  $k$ -th user's antenna array,  $\mathbf{y}_k \in \mathbb{C}^{N_r \times 1}$ , is given by

$$\mathbf{y}_k = \sqrt{\rho_k}\mathbf{H}_k\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\mathbf{s} + \mathbf{n}_k \quad (3.2)$$

where  $\rho_k$  represent the large scale fading and  $\mathbf{n}_k \in \mathbb{C}^{N_r \times 1}$  is the noise vector, such that  $\mathbf{n}_k$  is i.i.d.  $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_r})$ . At the  $k$ -th user, the received signal  $\mathbf{y}_k$  is first processed by the analog combiner matrix  $\mathbf{W}_{\text{RF},k} \in \mathbb{C}^{N_r \times N_r^{\text{RF}}}$ , and then processed by the digital combiner matrix  $\mathbf{W}_{\text{BB},k} \in \mathbb{C}^{N_r^{\text{RF}} \times N_s}$ . The analog combiner is also implemented using phase shifters and is constrained to have a constant

modulus, such that  $|\mathbf{W}_{\text{RF},k}(m,n)| = 1/\sqrt{N_r} \forall m,n$ . The received symbol is expressed by

$$\hat{s}_k = \sqrt{\rho_k} \mathbf{W}_{\text{BB},k}^H \mathbf{W}_{\text{RF},k}^H \mathbf{H}_k \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BBS}} + \mathbf{W}_{\text{BB},k}^H \mathbf{W}_{\text{RF},k}^H \mathbf{n}_k. \quad (3.3)$$

By defining the aggregated multiuser channel as  $\mathbf{H} = [\mathbf{H}_1^T \mathbf{H}_2^T \cdots \mathbf{H}_K^T]^T$ , the aggregated analog combining matrix as  $\mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \cdots, \mathbf{W}_{\text{RF},K}\}$ , and the effective baseband channel of the  $k$ -th user as  $\bar{\mathbf{H}}_k = \mathbf{W}_{\text{RF},k}^H \mathbf{H}_k \mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_r^{\text{RF}} \times N_t^{\text{RF}}}$ , one can represent the entire multiuser effective baseband channel  $\mathbf{H}_{\text{eff}}$  as

$$\begin{aligned} \mathbf{H}_{\text{eff}} &= \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \\ &= [\bar{\mathbf{H}}_1^T \bar{\mathbf{H}}_2^T \cdots \bar{\mathbf{H}}_K^T]^T \\ &= \begin{bmatrix} \mathbf{W}_{\text{RF},1}^H & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{\text{RF},2}^H & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_{\text{RF},K}^H \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} \mathbf{F}_{\text{RF}}. \end{aligned} \quad (3.4)$$

The estimated symbol for the  $i$ -th data stream in the  $k$ -th user,  $\hat{s}_{k_i}$ , can be expressed as [96]

$$\begin{aligned} \hat{s}_{k_i} &= \sqrt{\rho_k} \mathbf{W}_{\text{BB},k}^H(i,:) \bar{\mathbf{H}}_k \mathbf{F}_{\text{BB}}(:,k_i) s_{k_i} + \sum_{j=1, j \neq i}^{N_s} \sqrt{\rho_k} \mathbf{W}_{\text{BB},k}^H(i,:) \bar{\mathbf{H}}_k \mathbf{F}_{\text{BB}}(:,k_j) s_{k_j} \\ &+ \sum_{m=1, m \neq k}^K \sum_{l=1}^{N_s} \sqrt{\rho_k} \mathbf{W}_{\text{BB},k}^H(i,:) \bar{\mathbf{H}}_k \mathbf{F}_{\text{BB}}(:,m_l) s_{m_l} + \mathbf{W}_{\text{BB},k}^H(i,:) \mathbf{W}_{\text{RF},k}^H \mathbf{n}_k \end{aligned} \quad (3.5)$$

where  $k_i = (k-1)N_s + i$  and  $s_{k_i}$  is the  $i$ -th element of  $\mathbf{s}_k$ .

Considering that Gaussian symbols are transmitted, the achieved spectral efficiency is given by [99]

$$R = \sum_{k=1}^K \log_2 \left( \left| \mathbf{I}_{N_s} + \frac{\rho_k}{\sigma_n^2} \mathbf{R}_n^{-1} \mathbf{W}_k^H \mathbf{H}_k \mathbf{F}_k \mathbf{F}_k^H \mathbf{H}_k^H \mathbf{W}_k \right| \right) \quad (3.6)$$

where  $\mathbf{F}_k \triangleq \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB},k}$ ,  $\mathbf{W}_k \triangleq \mathbf{W}_{\text{RF},k} \mathbf{W}_{\text{BB},k}$ , and

$$\mathbf{R}_n = \frac{\rho_k}{\sigma_n^2} \sum_{j \neq k}^K \mathbf{W}_k^H \mathbf{H}_k \mathbf{F}_j \mathbf{F}_j^H \mathbf{H}_k^H \mathbf{W}_k + \mathbf{W}_k^H \mathbf{W}_k$$

is the interference plus noise covariance matrix. However, note that the spectral efficiency of a MIMO system represents a performance upper bound, and when linear precoders and combiners are used for spatial multiplexing, the sum-rate is more reasonable performance metric [86], [100]. The

sum-rate is given by

$$R_{\text{sum}} = \sum_{k=1}^K \sum_{i=1}^{N_s} \log_2 (1 + \text{SINR}_{k_i}) \quad (3.7)$$

where  $\text{SINR}_{k_i}$  represents the signal-to-interference-plus-noise ratio of the  $i$ -th data stream of the  $k$ -th user, given by

$$\text{SINR}_{k_i} = \frac{S}{I + N} \quad (3.8)$$

with  $S$ ,  $I$ , and  $N$  being, respectively, the signal, interference, and noise power, given by

$$\begin{aligned} S &= \rho_k |\mathbf{W}_k(:, i)^H \mathbf{H}_k \mathbf{F}_k(:, i)|^2 \\ I &= \sum_{j=1, j \neq i}^{N_s} \rho_k |\mathbf{W}_k(:, i)^H \mathbf{H}_k \mathbf{F}_k(:, j)|^2 \\ &\quad + \sum_{l=1, l \neq k}^K \sum_{j=1}^{N_s} \rho_l |\mathbf{W}_k(:, i)^H \mathbf{H}_k \mathbf{F}_l(:, j)|^2 \\ N &= \sigma_n^2 \mathbf{W}_k(:, i)^H \mathbf{W}_k(:, i). \end{aligned}$$

### 3.2.2 Channel Model

Two distinct fast fading channel models are assumed: the Rayleigh model, which represents a rich scattering environment, and the Saleh-Valenzuela model, representing the sparse scattering environment of mmWave channels. In the Rayleigh model, each entry of the fast fading channel matrix  $\mathbf{H}_k$  follows an i.i.d. zero-mean complex Gaussian distribution, i.e.  $\mathbf{H}_k(m, n) \sim \mathcal{CN}(0, 1)$ .

On the other hand, the Saleh-Valenzuela channel model corresponds to the sum of the contributions of  $N_{\text{cl}}$  scattering clusters, each formed by  $N_{\text{ray}}$  propagation paths. The discrete-time narrowband flat-fading channel matrix of the  $k$ -th user is given by [24]

$$\mathbf{H}_k = \sqrt{\frac{N_r N_t}{N_{\text{cl}} N_{\text{ray}}}} \sum_{i=0}^{N_{\text{cl}}-1} \sum_{l=0}^{N_{\text{ray}}-1} \alpha_{il} \mathbf{a}_r(\phi_{il}^r, \theta_{il}^r) \mathbf{a}_t^H(\phi_{il}^t, \theta_{il}^t), \quad (3.9)$$

where  $\alpha_{il}$  corresponds to the complex gain of the  $l$ -th multipath ray in the  $i$ -th cluster, the vectors  $\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)$  and  $\mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)$  are the array response vectors of the receiver and transmitter, respectively, whereas  $\phi_{il}^t$  and  $\theta_{il}^t$  are the azimuth and elevation angles of departure (AoD), and  $\phi_{il}^r$  and  $\theta_{il}^r$  are the azimuth and elevation angles of arrival (AoA). The complex gain  $\alpha_{il}$  is assumed to be i.i.d.  $\mathcal{CN}(0, \sigma_{\alpha, i}^2)$ , where  $\sigma_{\alpha, i}^2$  is the average power of the  $i$ -th cluster. The average cluster powers are such that  $\sum_{i=1}^{N_{\text{cl}}} \sigma_{\alpha, i}^2 = \sqrt{N_r N_t / N_{\text{cl}} N_{\text{ray}}}$  [24]. The azimuth and elevation angles of departure  $\phi_{il}^t, \theta_{il}^t$  and arrival  $\phi_{il}^r, \theta_{il}^r$  for the  $N_{\text{ray}}$  paths in the  $i$ -th cluster are modeled as Laplacian distributed

random variable, with mean  $\phi_i^t, \theta_i^t, \phi_i^r, \theta_i^r$  uniformly-distributed over  $[-\pi, \pi)$ , and angular spread of  $\sigma_\phi^t, \sigma_\theta^t, \sigma_\phi^r, \sigma_\theta^r$ , respectively. A  $\sqrt{N} \times \sqrt{N}$  uniform square planar array (USPA), with  $N = N_t$  and  $N = N_r$ , respectively, for the transmitter and receiver, is assumed. The array response vector for the USPA is defined as [83]

$$\mathbf{a}(\phi, \theta) = \frac{1}{\sqrt{N}} \left[ 1, \dots, e^{j \frac{2\pi d}{\lambda} [h \sin(\phi) \sin(\theta) + v \cos(\theta)]}, \dots, e^{j \frac{2\pi d}{\lambda} [(\sqrt{N}-1) \sin(\phi) \sin(\theta) + (\sqrt{N}-1) \cos(\theta)]} \right]^T \quad (3.10)$$

where  $d$  and  $\lambda$  are the spacing between elements and the signal wavelength, respectively, and  $0 \leq h < \sqrt{N} - 1$  and  $0 \leq v < \sqrt{N} - 1$  are the indexes of the antenna element in the 2D plane.

### 3.2.3 Problem Formulation

Here, the design of hybrid precoders and combiners aims to maximize the sum-rate of MIMO system. This design problem is formulated as

$$\begin{aligned} & \max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, \mathbf{W}_{\text{RF},k}, \mathbf{W}_{\text{BB},k}} && R_{\text{sum}} \\ & \text{s.t.} && \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}} \\ & && \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\ & && \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = P_{\text{T}} \end{aligned} \quad (3.11)$$

where  $\mathcal{F}_{\text{RF}}$  and  $\mathcal{W}_{\text{RF}}$  are, respectively, the set of all feasible analog precoder and combiner matrices (i.e., all  $N_t \times N_t^{\text{RF}}$  and  $N_r \times N_r^{\text{RF}}$  matrices with constant modulus entries), and  $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = P_{\text{T}}$  ensures the total transmitted power constraint. For simplicity, it is assumed that  $N_t^{\text{RF}} = K N_r^{\text{RF}} = K N_s$ . Nonetheless, following [83]–[97], it is further assumed that perfect channel state information (CSI) is available. Although perfect channel estimation with hybrid beamforming architecture is very challenging, efficient methods based on compressive sensing [101] and Bayesian learning [102] have been proposed.

In SU-MIMO systems, the maximal sum-rate is attained when the data streams are transmitted by multiple non-interfering parallel sub-channels, which is achieved by designing precoders and combiners that diagonalize the effective channel. In a fully-digital scheme, this can be obtained by choosing the precoder and combiner, respectively, as the right and left singular vectors of the channel matrix and using water-filling to allocate the power of each data stream [103]. For MU-MIMO systems, although the optimal solution is not necessarily obtained by diagonalizing the effective

channel, it generally results in a near-diagonal channel [98]. Thus, the precoder and combiner design can be seen as an attempt to perform such diagonalization.

When a hybrid structure is employed, solving the sum-rate maximization problem, as in (3.11), is very challenging. The problem is non-linear, due to the coupling between the analog and digital beamforming, and non-convex, due to the constant modulus constraint imposed by the phase-shifters in the analog beamforming. In fact, even for a single-user system, designing hybrid beamforming to attain the globally optimal sum-rate is found to be intractable [24]. Therefore, practical designs often resort to sub-optimal solutions, such as the two-stage approaches in which the analog and digital beamforming are designed separately. However, in many of these approaches, such as in [61], [91], [94], the analog precoder and combiner are designed to increase the desired signal's power at the receiver (i.e., increase the squared elements in the effective baseband channel's diagonal) without attempting to reduce intra-user and inter-user interference (i.e., reduce the effective baseband channel's off-diagonal elements). The interference mitigation is left for the digital beamforming alone.

### 3.3 Constrained Channel Decomposition-based Hybrid Beamforming Design for MIMO Systems

This section proposes a novel hybrid precoder and combiner design for narrowband mmWave massive MIMO. The proposed design consists of two stages. In the first stage, the analog precoder and combiner are jointly designed using the proposed constrained channel decomposition approach, which explores the extra degree-of-freedom provided by the analog beamforming to harvest the array gain and to provide intra-user and inter-user interference suppression. This approach is very general and applies to both single-user and multi-user systems. In the second stage, the digital beamforming is designed considering two distinct approaches. First, for single-user systems, the digital beamforming is obtained from the optimal SVD-based solution considering the effective baseband channel. Next, for multi-user systems, the digital beamforming is obtained using the regularized channel diagonalization approach (RCD) [61]. The proposed design and its computational complexity analysis are presented in the following subsections.

### 3.3.1 Analog Precoder and Combiner Design

The analog precoder and combiner's design aims to diagonalize the effective baseband channel  $\mathbf{H}_{\text{eff}}$ , such that the effective gain of each data stream is maximized while the interference among different data streams is suppressed. This problem is equivalent to finding the best rank- $KN_s$  channel approximation  $\mathbf{H} \approx \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}$ , so that  $\mathbf{H}_{\text{eff}} = \mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{H}\mathbf{F}_{\text{RF}} \approx \mathbf{\Sigma}$ , where  $\mathbf{\Sigma}$  is a  $KN_s \times KN_s$  diagonal matrix while both  $\mathbf{W}_{\text{RF}}$  and  $\mathbf{F}_{\text{RF}}^{\text{H}}$  are orthonormal. After taking the constant modulus constraints associated with the analog precoder and combiner into account, the design problem is formulated as

$$\begin{aligned}
& \min_{\mathbf{F}_{\text{RF}}, \mathbf{W}_{\text{RF}}, \mathbf{\Sigma}} && \|\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 \\
& \text{s.t.} && \mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{W}_{\text{RF}} = \mathbf{I}_{KN_s} \\
& && \mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{F}_{\text{RF}} = \mathbf{I}_{KN_s} \\
& && \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\} \\
& && \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\
& && \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}}
\end{aligned} \tag{3.12}$$

where the third constraint is imposed by the system model, while the two last constraints correspond to the constant-modulus constraint imposed by the phase-shifters (without these three constraints, the problem corresponds to the classical low-rank matrix approximation problem). With proper algebraic manipulation, as demonstrated in Appendix A, the problem in (3.12) can be equivalently cast as

$$\begin{aligned}
& \max_{\mathbf{F}_{\text{RF}}, \mathbf{W}_{\text{RF}}} && \|\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{H}\mathbf{F}_{\text{RF}}\|_{\text{F}}^2 \\
& \text{s.t.} && \mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{W}_{\text{RF}} = \mathbf{I}_{KN_s} \\
& && \mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{F}_{\text{RF}} = \mathbf{I}_{KN_s} \\
& && \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\} \\
& && \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\
& && \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}}.
\end{aligned} \tag{3.13}$$

The problem in (3.13) is intractable due to the non-convex constraints. A heuristic algorithm for solving (3.13) through a series of  $KN_s$  successive rank-1 approximations is proposed to overcome such a hurdle. Each rank-1 approximation corresponds to the design of one analog precoder and

combiner vector pair. The rank-1 approximation problem is formulated as

$$\begin{aligned}
& \max_{\mathbf{f}_j, \mathbf{w}_j} \quad |\mathbf{w}_j^H \mathbf{H} \mathbf{f}_j| \\
& \text{s.t.} \quad \mathbf{f}_j^H \mathbf{f}_j = 1, \quad \mathbf{f}_j^H \mathbf{f}_l = 0 \quad \forall j \neq l \\
& \quad \quad \mathbf{w}_j^H \mathbf{w}_j = 1, \quad \mathbf{w}_j^H \mathbf{w}_l = 0 \quad \forall j \neq l \\
& \quad \quad \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\} \\
& \quad \quad |[\mathbf{f}_j]_m| = 1/\sqrt{N_t}, \quad \forall m \in \{1, \dots, N_t\} \\
& \quad \quad |[\mathbf{w}_j]_n| = 1/\sqrt{N_r}, \quad \forall n \in \{1, \dots, N_r\}
\end{aligned} \tag{3.14}$$

where  $\mathbf{F}_{\text{RF}} = [\mathbf{f}_1 \cdots \mathbf{f}_{KN_s}]$ ,  $\mathbf{W}_{\text{RF}} = [\mathbf{w}_1 \cdots \mathbf{w}_{KN_s}]$ , and  $\{j, l\} \in \{1, 2, \dots, KN_s\}$ . The problem in (3.14) is still non-convex and difficult to solve.

Problem (3.14) can be further simplified according to the following observations: i) the objective function and constraints in (3.14) are invariant to rotations in  $\mathbf{w}_i$  and  $\mathbf{f}_i$ , thereby replacing the objective function with  $\text{Re}(\mathbf{w}_i^H \mathbf{H}_{\text{res}}^{(i)} \mathbf{f}_i)$  will also lead to an optimal solution of (3.14); ii) given the block-diagonal structure of the aggregated analog combiner matrix, each analog combiner vector interacts only with the channel matrix of the same user and, in addition, only the orthogonality among analog combiner vectors within the same user needs to be ensured; lastly, iii) the orthogonality constraint, which guarantees the interference suppression, can be slightly relaxed since it can be further ensured by the digital precoder and combiner. Bearing these observations in mind, problem (3.14) is relaxed as

$$\begin{aligned}
& \max_{\mathbf{w}_{k,i}, \mathbf{f}_{k,i}} \quad \text{Re} \left( \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} \right) \\
& \text{s.t.} \quad \mathbf{w}_{k,i}^H \mathbf{w}_{k,i} = 1 \\
& \quad \quad \mathbf{f}_{k,i}^H \mathbf{f}_{k,i} = 1 \\
& \quad \quad |[\mathbf{f}_{k,i}]_m| = 1/\sqrt{N_t}, \quad \forall m \in \{1, \dots, N_t\} \\
& \quad \quad |[\mathbf{w}_{k,i}]_n| = 1/\sqrt{N_r}, \quad \forall n \in \{1, \dots, N_r\}.
\end{aligned} \tag{3.15}$$

where  $\mathbf{f}_{k,i}$  is the  $[(k-1)N_s + i]$ -th column of  $\mathbf{F}_{\text{RF}}$ ,  $\mathbf{w}_{k,i}$  is the  $i$ -th column of  $\mathbf{W}_{\text{RF},k}$ , with  $i \in \{1, \dots, N_s\}$  and  $k \in \{1, \dots, K\}$ , such that  $\mathbf{f}_{k,i}$  and  $\mathbf{w}_{k,i}$  are, respectively, the analog precoder and analog combiner corresponding to the  $i$ -th data stream in the  $k$ -th user. Note that, to preserve some level of orthogonality, the channel matrix in (3.14) has been replaced with the residual channel matrix  $\mathbf{H}_{\text{res}}^{(i)}$ , i.e., the channel matrix subtracted the contributions of the previously computed analog

precoder and combiner vectors. In particular,  $\mathbf{H}_{\text{res},k}^{(i)}$  represents the residual channel matrix from the  $k$ -th user.

The solution to problem (3.15) is obtained through the projected block coordinate descent method [104], [105] by cyclically seeking the block coordinate descent solution with respect to each block variable (e.g.,  $\mathbf{w}_{k,i}$  or  $\mathbf{f}_{k,i}$ ) while keeping the other block variable fixed. Thereby, each block coordinate problem is solved by relaxing the non-convex constant modulus constraint and then projecting the relaxed solution onto the non-convex set of constant modulus vectors. The relaxed block coordinate descent problems are, respectively,

$$\begin{aligned} \max_{\mathbf{w}_{k,i}} \operatorname{Re} \left( \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} \right) & \quad \text{and} \quad \max_{\mathbf{f}_{k,i}} \operatorname{Re} \left( \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} \right) \\ \text{s.t. } \mathbf{w}_{k,i}^H \mathbf{w}_{k,i} = 1 & \quad \quad \quad \text{s.t. } \mathbf{f}_{k,i}^H \mathbf{f}_{k,i} = 1. \end{aligned} \quad (3.16)$$

The solution to the problems in (3.16) is obtained using the Lagrange multiplier method. For instance, for the problem on the left-hand side, the Lagrangian function is defined as

$$\mathcal{L}(\mathbf{w}_{k,i}, \lambda) = \operatorname{Re} \left( \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} \right) - \lambda (\mathbf{w}_{k,i}^H \mathbf{w}_{k,i} - 1). \quad (3.17)$$

The optimal  $\mathbf{w}_{k,i}$  is then obtained by setting the gradient of  $\mathcal{L}(\mathbf{w}_{k,i}, \lambda)$  w.r.t.  $\mathbf{w}_{k,i}^*$  to zero, i.e.,

$$\nabla_{\mathbf{w}_{k,i}^*} \mathcal{L}(\mathbf{w}_{k,i}, \lambda) = \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} - \lambda \mathbf{w}_{k,i} = \mathbf{0}. \quad (3.18)$$

Eq. (3.18) yields  $\hat{\mathbf{w}}_{k,i} = \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} / \lambda$ , where  $\lambda$  is obtained to ensure  $\hat{\mathbf{w}}_{k,i}^H \hat{\mathbf{w}}_{k,i} = 1$ . The solution to the problem on the right-hand side in (3.16) is obtained through the same procedure. The solutions to the relaxed block coordinate descent problems are, respectively, given as

$$\hat{\mathbf{w}}_{k,i} = \frac{\mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i}}{\left\| \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i} \right\|} \quad \text{and} \quad \hat{\mathbf{f}}_{k,i} = \frac{\mathbf{H}_{\text{res},k}^{(i)H} \mathbf{w}_{k,i}}{\left\| \mathbf{H}_{\text{res},k}^{(i)H} \mathbf{w}_{k,i} \right\|} \quad (3.19)$$

and the projection of the relaxed solutions onto the constant modulus vector space are obtained by extracting the phases of the relaxed solutions, i.e., by making

$$\mathbf{w}_{k,i} = \frac{1}{\sqrt{N_r}} e^{j \angle \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i}} \quad \text{and} \quad \mathbf{f}_{k,i} = \frac{1}{\sqrt{N_t}} e^{j \angle \mathbf{H}_{\text{res},k}^{(i)H} \mathbf{w}_{k,i}}. \quad (3.20)$$

Note that the normalization factors in (3.19) do not need to be taken into account since they do not change the phases of the solution vector. Moreover, despite the non-convexity of the block

coordinate problems, Appendix B proves that the solutions in (3.20) are, in fact, optimal. Finally, the solution to (3.15) is obtained by iterating the equations in (3.20) until convergence.

After obtaining the pair of analog precoder and combiner vectors, the residual channel matrix for the  $k$ -th user is updated by simultaneously projecting the columns of  $\mathbf{H}_{\text{res},k}^{(i)}$  onto the subspace orthogonal to the analog combiner  $\mathbf{w}_{k,i}$  and their rows onto the subspace orthogonal to the analog precoder  $\mathbf{f}_{k,i}$  by making

$$\mathbf{H}_{\text{res},k}^{(\gamma+1)} = (\mathbf{I}_{N_r} - \mathbf{w}_{k,i}\mathbf{w}_{k,i}^H)\mathbf{H}_{\text{res},k}^{(\gamma)}(\mathbf{I}_{N_t} - \mathbf{f}_{k,i}\mathbf{f}_{k,i}^H) \quad (3.21)$$

while the residual channel matrix for the remaining users are updated by projecting their rows onto the subspace orthogonal to the analog precoder  $\mathbf{f}_{k,i}$  by

$$\mathbf{H}_{\text{res},j}^{(\gamma+1)} = \mathbf{H}_{\text{res},j}^{(\gamma)}(\mathbf{I}_{N_t} - \mathbf{f}_{k,i}\mathbf{f}_{k,i}^H), \quad \forall j \in \{1, \dots, K | j \neq k\}. \quad (3.22)$$

The update in (3.21) allows finding the analog precoder and combiner vectors for the  $(i+1)$ -th data stream in the  $k$ -th user nearly-orthogonal to, respectively, the analog precoder and combiner vectors for the  $i$ -th data stream of the same user. The update in (3.22), on the other hand, allows seeking analog precoders for the remaining users that are nearly-orthogonal to the analog precoder computed for the  $i$ -th data stream in the  $k$ -th user. Note that finding orthogonal analog precoders and combiners is very difficult when the constant modulus constraints are taken into account. Nonetheless, while the updates in (3.21) and (3.22) cannot assure orthogonality among the analog beamforming vectors, the near-orthogonality is good enough to provide intra-user and inter-user interference reduction since the interference suppression is reinforced at the digital beamforming stage. The procedure for solving (3.15) described herein is repeated until the analog beamforming for all users and data streams are designed. The proposed algorithm is summarized in Algorithm 3.1.

Note that different user scheduling schemes can be adopted when solving the successive rank-one approximation problem in (3.15). For instance, Alg. 3.1 seeks fairness among the users, and, therefore, the analog precoder and combiner vector pairs are designed for one data stream in each user at a time. An additional possibility is to follow a certain user priority scheme, thereby designing all the analog precoder and combiner vector pairs for the highest priority users to the lowest ones.

---

**Algorithm 3.1:** Joint Analog Precoder & Combiner Design for Narrowband MIMO

---

```

1 Initialize  $\gamma = 0$  and  $\mathbf{F}_{\text{RF}} = [ ]$ 
2 Initialize  $\mathbf{W}_{\text{RF},k} = [ ]$  and  $\mathbf{H}_{\text{res},k}^{(0)} = \mathbf{H}_k, \forall k = 1, \dots, K$ 
3 for  $i = 1 : N_s$  do
4   for  $k = 1 : K$  do
5     Initialize  $\mathbf{w}_{k,i}$  and  $\mathbf{f}_{k,i}$  randomly
6     Initialize  $\eta = 0; \delta^{(0)} = 1; \delta^{(1)} = 0$ 
7     while  $|\delta^{(\eta+1)} - \delta^{(\eta)}| \geq \varepsilon$  do
8        $\eta = \eta + 1$ 
9        $\hat{\mathbf{w}}_{k,i} = \mathbf{H}_{\text{res},k}^{(\gamma)} \mathbf{f}_{k,i}$ 
10       $\mathbf{w}_{k,i} = (1/\sqrt{N_r}) e^{j\angle \hat{\mathbf{w}}_{k,i}}$ 
11       $\hat{\mathbf{f}}_{k,i} = \mathbf{H}_{\text{res},k}^{(\gamma)\text{H}} \mathbf{w}_{k,i}$ 
12       $\mathbf{f}_{k,i} = (1/\sqrt{N_t}) e^{j\angle \hat{\mathbf{f}}_{k,i}}$ 
13       $\delta^{(\eta+1)} = \mathbf{w}_{k,i}^{\text{H}} \mathbf{H}_{\text{res},k}^{(\gamma)} \mathbf{f}_{k,i}$ 
14    end
15     $\mathbf{H}_{\text{res},k}^{(\gamma+1)} = (\mathbf{I}_{N_r} - \mathbf{w}_{k,i} \mathbf{w}_{k,i}^{\text{H}}) \mathbf{H}_{\text{res},k}^{(\gamma)} (\mathbf{I}_{N_t} - \mathbf{f}_{k,i} \mathbf{f}_{k,i}^{\text{H}})$ 
16    for  $j = 1 : K, j \neq k$  do
17       $\mathbf{H}_{\text{res},j}^{(\gamma+1)} = \mathbf{H}_{\text{res},j}^{(\gamma)} (\mathbf{I}_{N_t} - \mathbf{f}_{k,i} \mathbf{f}_{k,i}^{\text{H}})$ 
18    end
19     $\mathbf{W}_{\text{RF},k}(:, i) = \mathbf{w}_{k,i}$ 
20     $\mathbf{F}_{\text{RF}}(:, (k-1)N_s + i) = \mathbf{f}_{k,i}$ 
21     $\gamma = \gamma + 1$ 
22  end
23 end
24 return  $\mathbf{F}_{\text{RF}}, \mathbf{W}_{\text{RF}}$ 

```

---

### 3.3.2 Digital Precoder and Combiner Design for Single-User Systems

First, consider a single-user system (i.e.,  $K = 1$ ), where only intra-user interference is present. In this case, for a given analog precoder and combiner matrix, the optimal digital precoder and combiner matrices are obtained from the effective baseband channel's SVD, followed by the water-filling power allocation [103].

The effective baseband channel, i.e.,  $\mathbf{H}^{\text{eff}} = \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}$ , takes into account the effects of the analog precoder and combiner and has SVD as

$$\begin{aligned}
\mathbf{H}^{\text{eff}} &= \mathbf{U} \mathbf{\Sigma} \mathbf{V}^{\text{H}} \\
&= [\mathbf{U}_0 \quad \mathbf{U}_1] \begin{bmatrix} \mathbf{\Sigma}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1 \end{bmatrix} [\mathbf{V}_0 \quad \mathbf{V}_1]^{\text{H}}
\end{aligned} \tag{3.23}$$

where  $\mathbf{\Sigma}$  is a diagonal matrix with entries given by the singular values of  $\mathbf{H}^{\text{eff}}$  arranged in descending order, and the columns of  $\mathbf{U}$  and  $\mathbf{V}$  are, respectively, the left and right singular vectors associated with these singular values. Let  $\mathbf{\Sigma}_0$  be the diagonal matrix with the  $N_s$  largest singular values, and  $\mathbf{U}_0$  and  $\mathbf{V}_0$  the  $N_s$  right and left singular vector associated with these singular values, the optimal digital precoder and combiner are, respectively,

$$\mathbf{F}_{\text{BB}} = \mathbf{V}_0 \mathbf{P} \quad (3.24)$$

and

$$\mathbf{W}_{\text{BB}} = \mathbf{U}_0 \quad (3.25)$$

where  $\mathbf{P}$  is the diagonal power allocation matrix.

For equal power allocation (i.e., all data streams are transmitted with the same power), the power allocation matrix is set to  $\mathbf{P} = \sqrt{P_{\text{T}} \mathbf{I}_{N_s}} / \|\mathbf{F}_{\text{RF}} \mathbf{V}_0\|_{\text{F}}$  to ensure the total transmitted power  $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = P_{\text{T}}$  is attained. Therefore, the digital precoder is given by

$$\mathbf{F}_{\text{BB}} = \sqrt{P_{\text{T}}} \frac{\mathbf{V}_0}{\|\mathbf{F}_{\text{RF}} \mathbf{V}_0\|_{\text{F}}}. \quad (3.26)$$

For optimal power allocation, the power allocation matrix is computed by solving the water-filling problem [106],

$$\begin{aligned} \max_{\mathbf{P}} \quad & \sum_{i=1}^{N_s} \log \left( 1 + [\mathbf{P}]_{i,i} [\mathbf{\Lambda}]_{i,i} \right) \\ \text{s.t.} \quad & \sum_{i=1}^{N_s} [\mathbf{P}]_{i,i} \leq P_{\text{T}} \\ & [\mathbf{P}]_{i,i} \geq 0, \quad 1 \leq i \leq N_s \end{aligned} \quad (3.27)$$

whose solution is given by

$$[\mathbf{P}]_{i,i} = (\mu - [\mathbf{\Lambda}]_{i,i}^{-1})^+, \quad 1 \leq i \leq N_s \quad (3.28)$$

where  $[\mathbf{P}]_{i,i}$  is the power allocated to the  $i$ -th data stream,  $[\mathbf{\Lambda}]_{i,i}$  is the  $i$ -th diagonal of the effective channel  $\mathbf{\Lambda} = \mathbf{U}_0^{\text{H}} \mathbf{H}^{\text{eff}} \mathbf{V}_0$ ,  $(\lambda)^+ \triangleq \max(0, \lambda)$ , and  $\mu$  is the water level chosen to satisfy the total transmitted power constraint.

### 3.3.3 Digital Precoder and Combiner Design for Multi-User Systems

Unlike single-user systems, where the precoders and combiners are designed to increase the effective gain and reduce intra-user interference, multi-user systems also need to address MUI (i.e. inter-user

interference). Besides, in single-user systems, the optimal spectral efficiency can be attained using linear processing. This is not the case in multi-user systems, where optimal spectral efficiency can, in general, only be achieved using non-linear techniques (e.g. dirty-paper coding [107]), incurring increased complexity. Different sub-optimal linear processing techniques have been proposed to overcome such a drawback, such as the BD [98] and the RCD [61]. In contrast to the BD, which only addresses the MUI, the RCD approach can trade-off MUI suppression for noise suppression and transmit diversity, thus attaining a better performance in low-SNR scenarios. Here, the RCD approach is adopted to design the digital precoder and combiner.

The RCD's goal is to reduce the MUI plus noise and to optimize the system performance. Thus, the RCD design is obtained in two steps: the first step addresses the MUI plus noise suppression, and the second step performs the intra-user interference suppression and performance optimization. The digital precoder is written as

$$\mathbf{F}_{\text{BB}} = \beta \mathbf{F}_{\text{a}} \mathbf{F}_{\text{b}} \quad (3.29)$$

with  $\mathbf{F}_{\text{a}} = [\mathbf{F}_{\text{a}_1} \ \mathbf{F}_{\text{a}_2} \ \cdots \ \mathbf{F}_{\text{a}_K}] \in \mathbb{C}^{KN_s \times KN_s}$  and

$$\mathbf{F}_{\text{b}} = \begin{bmatrix} \mathbf{F}_{\text{b}_1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{F}_{\text{b}_K} \end{bmatrix} \in \mathbb{C}^{KN_s \times KN_s}$$

where  $\mathbf{F}_{\text{a}_k} \in \mathbb{C}^{KN_s \times N_s}$  is the primary digital precoder for MUI plus noise suppression,  $\mathbf{F}_{\text{b}_k} \in \mathbb{C}^{N_s \times N_s}$  is the secondary digital precoder used for performance optimization, and  $\beta$  is chosen to fulfill the total transmit power constraint in (3.11).

Thus, by defining  $\tilde{\mathbf{H}}_k = [\tilde{\mathbf{H}}_1^T \ \cdots \ \tilde{\mathbf{H}}_{k-1}^T \ \tilde{\mathbf{H}}_{k+1}^T \ \cdots \ \tilde{\mathbf{H}}_K^T]^T \in \mathbb{C}^{(K-1)N_s \times KN_s}$ , which aggregates the effective baseband channel of all users, except the  $k$ -th user, the primary precoder design is formulated as the MUI plus noise minimization problem [61]

$$\min_{\mathbf{F}_{\text{a}}} \mathbb{E} \left\{ \sum_{k=1}^K \|\tilde{\mathbf{H}}_k \mathbf{F}_{\text{a}_k}\|_{\text{F}}^2 + \frac{\|\mathbf{z}\|^2}{\beta^2} \right\} \quad (3.30)$$

where  $\|\tilde{\mathbf{H}}_k \mathbf{F}_{\text{a}_k}\|_{\text{F}}$  relates to the level of interference caused by the  $k$ -th user's signals to the remaining users,  $\mathbf{z}$  is the noise vectors at the input of the digital combiners, such that  $\mathbf{z} = \mathbf{W}_{\text{RF}}^H \mathbf{n}$  with  $\mathbf{n} = [\mathbf{n}_1^T \ \cdots \ \mathbf{n}_K^T]^T$  being i.i.d. with zero-mean and variance  $\sigma_n^2$  equal for all users, and  $\beta$  ensures the

total transmitted power, i.e.,

$$\beta^2 \sum_{k=1}^K \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{a}_k} \mathbf{F}_{\text{b}_k}\|_{\text{F}}^2 = P_{\text{T}}. \quad (3.31)$$

Assuming that  $\mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{F}_{\text{RF}} = \mathbf{I}_{KN_s}$  and  $\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{W}_{\text{RF}} = \mathbf{I}_{KN_s}$ , which holds with high probability for large-scale arrays [87], and that under the equal power allocation assumption, the secondary digital precoder  $\mathbf{F}_{\text{b}_k}$  is a unitary matrix (which indeed holds, as will be shortly demonstrated), one obtains

$$\begin{aligned} \beta^2 &= \frac{P_{\text{T}}}{\sum_{k=1}^K \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{a}_k} \mathbf{F}_{\text{b}_k}\|_{\text{F}}^2} \\ &= \frac{P_{\text{T}}}{\sum_{k=1}^K \text{Tr} [\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{a}_k} \mathbf{F}_{\text{b}_k} \mathbf{F}_{\text{b}_k}^{\text{H}} \mathbf{F}_{\text{a}_k}^{\text{H}} \mathbf{F}_{\text{RF}}^{\text{H}}]} \\ &= \frac{P_{\text{T}}}{\sum_{k=1}^K \text{Tr} [\mathbf{F}_{\text{a}_k}^{\text{H}} \mathbf{F}_{\text{a}_k}]} \end{aligned} \quad (3.32)$$

and

$$\begin{aligned} \text{E} \{ \|\mathbf{z}\|^2 \} &= \text{E} \{ \text{Tr} [\mathbf{z}\mathbf{z}^{\text{H}}] \} \\ &= \text{E} \{ \text{Tr} [\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{n}\mathbf{n}^{\text{H}} \mathbf{W}_{\text{RF}}] \} \\ &= \sigma_{\text{n}}^2 \text{Tr} [\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{W}_{\text{RF}}] = \sigma_{\text{n}}^2 KN_s. \end{aligned} \quad (3.33)$$

By plugging in the results from (3.32) and (3.33), problem (3.30) is rewritten as [108, Appendix I]

$$\min_{\mathbf{F}_{\text{a}}} \sum_{k=1}^K \text{Tr} \left[ \mathbf{F}_{\text{a}_k}^{\text{H}} \left( \tilde{\mathbf{H}}_k^{\text{H}} \tilde{\mathbf{H}}_k + \frac{KN_s \sigma_{\text{n}}^2}{P_{\text{T}}} \mathbf{I}_{KN_s} \right) \mathbf{F}_{\text{a}_k} \right] \quad (3.34)$$

whose solution is given by [109]

$$\mathbf{F}_{\text{a}_k} = \left( \tilde{\mathbf{H}}_k^{\text{H}} \tilde{\mathbf{H}}_k + \frac{KN_s \sigma_{\text{n}}^2}{P_{\text{T}}} \mathbf{I}_{KN_s} \right)^{-1}. \quad (3.35)$$

Next, assuming that the primary digital precoder minimizes all MUI, so that  $\mathbf{H}_{\text{eff}} \mathbf{F}_{\text{a}}$  is block diagonal, the secondary digital precoder and the digital combiner design can be viewed as optimizing  $K$  parallel SU-MIMO. Let the effective SU-MIMO channel for the  $k$ -th user  $\tilde{\mathbf{H}}_k \mathbf{F}_{\text{a}_k}$  and its SVD  $\tilde{\mathbf{H}}_k \mathbf{F}_{\text{a}_k} = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^{\text{H}}$ , the optimal secondary digital precoder and digital combiner for the sum-rate maximization under total transmitted power constraint are, respectively,

$$\mathbf{F}_{\text{b}_k} = \mathbf{V}_s \quad (3.36)$$

and

$$\mathbf{W}_{\text{BB},k} = \mathbf{U}_s. \quad (3.37)$$

Note that, since  $\mathbf{F}_{\text{b}_k}$  is a right singular vector matrix, the assumption that  $\mathbf{F}_{\text{b}_k}$  is unitary holds for equal power allocation. Finally, by plugging (3.31) into the expression in (3.29), the digital precoder

is given by

$$\mathbf{F}_{\text{BB}} = \sqrt{\frac{P_{\text{T}}}{\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{a}}\mathbf{F}_{\text{b}}\|_{\text{F}}^2}} \mathbf{F}_{\text{a}}\mathbf{F}_{\text{b}}. \quad (3.38)$$

Alternatively, for optimal power allocation, the water-filling algorithm can be used to distribute the power along the columns of the digital precoding matrix.

In contrast to the classical BD approach, in which the primary precoder forces the  $k$ -th user to transmit in the null-space of  $\tilde{\mathbf{H}}_k$ , providing a better MUI suppression [94], [98], in the RCD design, the primary digital precoder allows the  $k$ -th user transmits in the subspace spanned by all other users with power inversely proportional to  $\tilde{\mathbf{H}}_k$ 's singular values [108]. As a result, the RCD design trades-off MUI suppression for noise suppression and transmit diversity, leading to a better performance in low-SNR scenarios. For high SNR, the transmission tends to concentrate in the null-space of  $\tilde{\mathbf{H}}_k$ , and thus, the RCD precoder approaches the BD design.

### 3.3.4 Computational complexity

The iterative process used to design the analog precoder and combiner matrices has computational complexity order of  $\mathcal{O}\{KN_{\text{s}}N_{\text{t}}N_{\text{r}}(N_{\text{ite}} + K)\}$ , where  $N_{\text{ite}}$  is the number of iterations required to compute an analog precoder and combiner vector pair. The digital precoder and combiner's design requires the computation of the effective baseband channel, with complexity  $\mathcal{O}\{KN_{\text{s}}(N_{\text{t}}N_{\text{r}} + KN_{\text{t}}N_{\text{s}})\}$ .

In SU-MIMO systems (i.e.,  $K = 1$ ), the complexity of the analog precoder and combiner's design and the effective baseband channel's computation reduces to, respectively,  $\mathcal{O}\{N_{\text{s}}N_{\text{t}}N_{\text{r}}N_{\text{ite}}\}$  and  $\mathcal{O}\{N_{\text{s}}^2N_{\text{t}} + N_{\text{s}}N_{\text{t}}N_{\text{r}}\}$ . The digital precoder and combiner are obtained from the effective baseband channel's SVD, with complexity  $\mathcal{O}(N_{\text{s}}^3)$ , followed by the digital precoder normalization with complexity  $\mathcal{O}(N_{\text{s}}^2N_{\text{t}})$ . Therefore, under the assumption that  $N_{\text{s}} \leq N_{\text{r}}$  and  $N_{\text{s}} \leq N_{\text{t}}$  (which always holds for the proposed system), the overall complexity order of the proposed design for SU-MIMO is approximately  $\mathcal{O}\{N_{\text{s}}N_{\text{t}}N_{\text{r}}N_{\text{ite}}\}$ . For comparison, Table 3.1 shows the complexity of the proposed method along with other SU-MIMO hybrid beamforming designs in the literature.  $|\mathcal{F}|$  and  $|\mathcal{W}|$  are, respectively, the sizes of the analog precoder and combiner codebooks.

In MU-MIMO systems, the digital precoder and combiner design is obtained from the RCD method, with complexity  $\mathcal{O}\{K^4N_{\text{s}}^3\}$ , followed by the digital precoder normalization with complexity  $\mathcal{O}\{K^2N_{\text{s}}^2N_{\text{t}}\}$ . Thus, under the assumption that  $N_{\text{s}} \leq N_{\text{r}}$  and  $N_{\text{s}} \leq N_{\text{t}}$ , the overall complexity of the proposed design for MU-MIMO systems is approximately  $\mathcal{O}\{KN_{\text{s}}[K^3N_{\text{s}}^2 + N_{\text{t}}N_{\text{r}}(K + N_{\text{ite}})]\}$ . For comparison, the complexity of the proposed method and other existing MU-MIMO hybrid beamforming designs are shown in Table 3.2.

**Table 3.1**  
Computational Complexity of SU-MIMO Hybrid Beamforming Design Methods

Method	Computational complexity
Prop. Method	$\mathcal{O}\{N_t N_r N_s N_{\text{ite}}\}$
OMP [24]	$\mathcal{O}\{N_t N_r^2 + (N_t + N_r) N_s^3 + (N_t  \mathcal{F}  + N_r  \mathcal{W} ) N_s\}$
PE-AltMin [83]	$\mathcal{O}\{N_s^2 (N_t + N_r) N_{\text{ite}} + N_s N_t N_r\}$
HBF-LSAA [87]	$\mathcal{O}\{N_s^2 N_t N_r^2 N_{\text{ite}}\}$
SVD-PE [84]	$\mathcal{O}\{N_t N_r^2\}$
SVD-PE-HBF [85]	$\mathcal{O}\{N_t N_r^2 + N_r N_s^2\}$
SSRM-HBF [86]	$\mathcal{O}\{N_s^3 (N_t + N_r) + N_s N_t N_r  \mathcal{F}   \mathcal{W} \}$

**Table 3.2**  
Computational Complexity of MU-MIMO Hybrid Beamforming Design Methods

Method	Computational complexity
Prop. Method	$\mathcal{O}\{K N_s [K^3 N_s^2 + N_t N_r (K + N_{\text{ite}})]\}$
Hy-BD [94]	$\mathcal{O}\{K [K^3 N_s^3 + K N_s^2 N_t + N_t N_r^2]\}$
Hy-SBD [96]	$\mathcal{O}\{K [K^3 N_s^3 + K N_s^2 N_t + N_t N_r^2]\}$
EBF-HBF [97]	$\mathcal{O}\{K N_s^2 [K^3 N_s + K N_s^2 N_t + N_{\text{ite}} (N_t^2 + N_r^2)]\}$
HRCDD [61]	$\mathcal{O}\{K [K^3 N_s^3 + K N_s^2 N_t + N_t N_r^2 + N_s N_{\text{cl}} N_{\text{ray}}]\}$

### 3.4 Simulation Results

This section presents extensive simulation results to evaluate the proposed method, compare it with alternative designs available in the literature, and validate its effectiveness. The section is organized as follows: Sections 3.4.1 and 3.4.2 presents the simulation results, respectively, for SU and MU systems, considering both Rayleigh and mmWave channel models; Section 3.4.3 investigates the convergence of the proposed method; and finally, Section 3.4.4 presents further discussions on the results.

In all examples, optimal water-filling power allocation is considered for all methods. Moreover, the examples evaluate both the spectral efficiency and the sum-rate performances of the different hybrid beamforming designs. As previously discussed, the spectral efficiency represents an upper-bound on the achievable sum-rate, while the sum-rate itself provides a more reasonable metric when linear processing techniques are employed [86], [100]. The SNR is defined as  $\text{SNR} = P_T \rho_k / \sigma_n$ , where the total transmitted power is set to  $P_T = K N_s$  and  $\rho_k$  is assumed equal for all users. The simulation results are obtained by averaging the spectral efficiency and the sum-rate over  $10^3$

channel realizations. The stopping criteria in Alg. 3.1 is set to  $\varepsilon = 0.001$ , and the maximum number of iteration is set to  $N_{\text{ite}} = 100$ . For the mmWave channel model, it is assumed  $N_{\text{cl}} = 5$  and  $N_{\text{ray}} = 10$ , with average cluster power  $\sigma_{\alpha,i}^2 = 1$ , angular spread  $\sigma_{\phi}^{\text{t}} = \sigma_{\theta}^{\text{t}} = \sigma_{\phi}^{\text{r}} = \sigma_{\theta}^{\text{r}} = 10^\circ$ , and a USPA with antenna spacing  $d = \lambda/2$ .

### 3.4.1 Single-User Systems

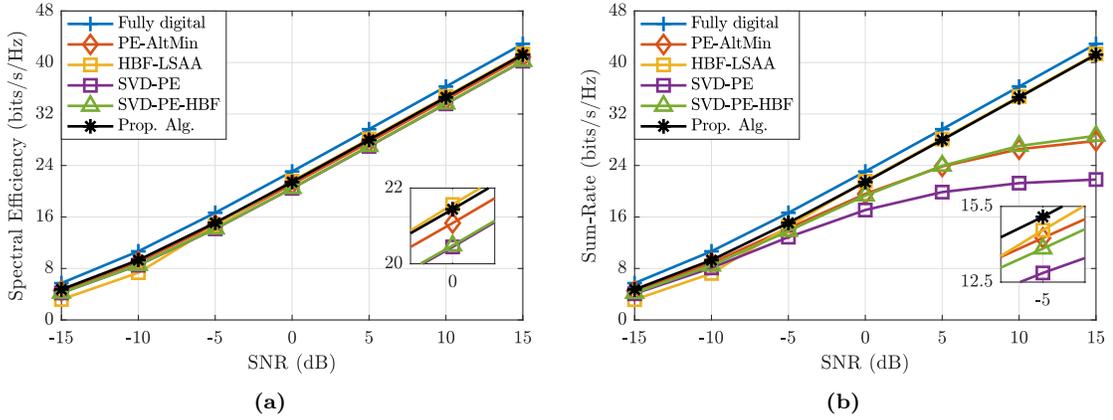
This section studies the performance of the proposed design in SU-MIMO systems. Two sets of examples are presented, considering Rayleigh and mmWave channels, respectively. The proposed design is compared to the optimal fully-digital solution, OMP [24], PE-AltMin [83], HBF-LSAA [87], SVD-PE [84], SVD-PE-HBF [85], and SSRM-HBF [86]. A beam-steering codebook containing the array response vectors of all propagation paths is considered for the OMP and the SSRM-HBF. Since these two designs are specifically tailored for the mmWave channel, they are not considered in the Rayleigh channel examples.

#### Rayleigh Channel

The first set of examples investigates the proposed hybrid beamforming design performance for SU-MIMO systems under Rayleigh channels.

The first example evaluates the spectral efficiency and the sum-rate for different SNR values, ranging from  $-15$  dB to  $15$  dB, for a  $64 \times 64$  MIMO system (i.e.,  $N_{\text{t}} = 64$  transmit and  $N_{\text{r}} = 64$  receive antennas), communicating through  $N_{\text{s}} = 4$  data streams. The results are shown, respectively, in Fig. 3.2a and Fig. 3.2b. Note that for SNR below  $0$  dB, the proposed design achieves the highest spectral efficiency among all hybrid beamforming designs. For higher SNR, the HBF-LSAA design slightly outperforms (with up to  $0.13$  bits/s/Hz higher spectral efficiency). On the other hand, note that the HBF-LSAA design has the worst performance in low SNR, with spectral efficiency around  $2$  bits/s/Hz lower than the proposed method at SNR =  $-10$  dB. The remaining designs achieve spectral efficiency lower than the proposed method at all SNR range. A similar result is obtained in terms of the sum-rate. However, note that although the PE-AltMin, SVD-PE, and SVD-PE-HBF can attain near-optimal spectral efficiencies, this does not reflect a high sum-rate. As shown in Fig. 3.2b, these methods have their sum-rates degraded with the increase in the SNR.

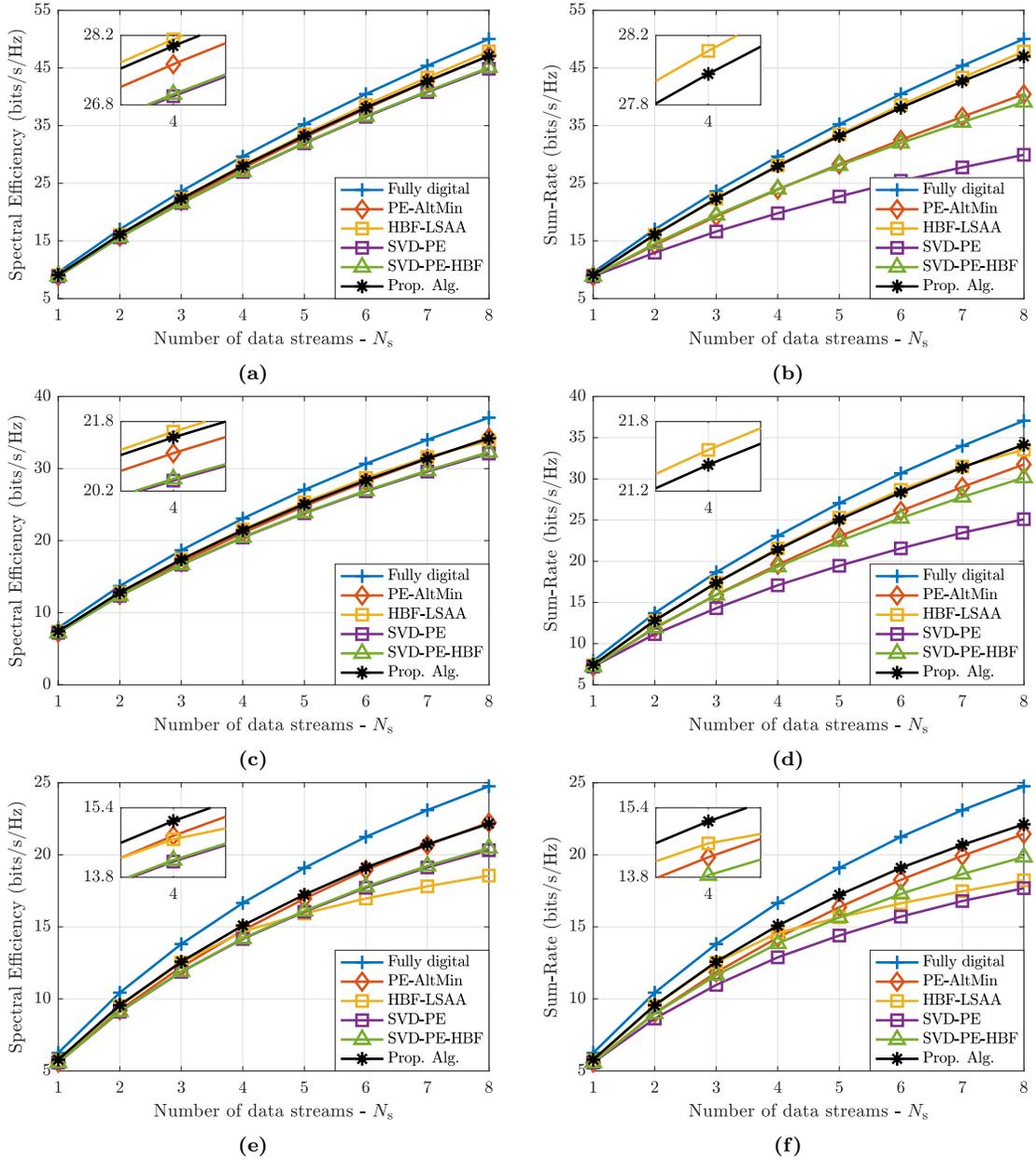
The second example investigates the impact of the number of data streams in the spectral efficiency and the sum-rate under different SNR. A  $64 \times 64$  MIMO system is considered. The results are depicted in Fig. 3.3, where the graphs in the left (i.e., (a), (c), and (e)) correspond to the spectral efficiency and those in the right (i.e., (b), (d), and (f)), to the sum-rate. Likewise, the graphs at



**Figure 3.2:** Example 1. SU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with  $N_t = N_r = 64$  antennas and  $N_s = 4$  data streams.

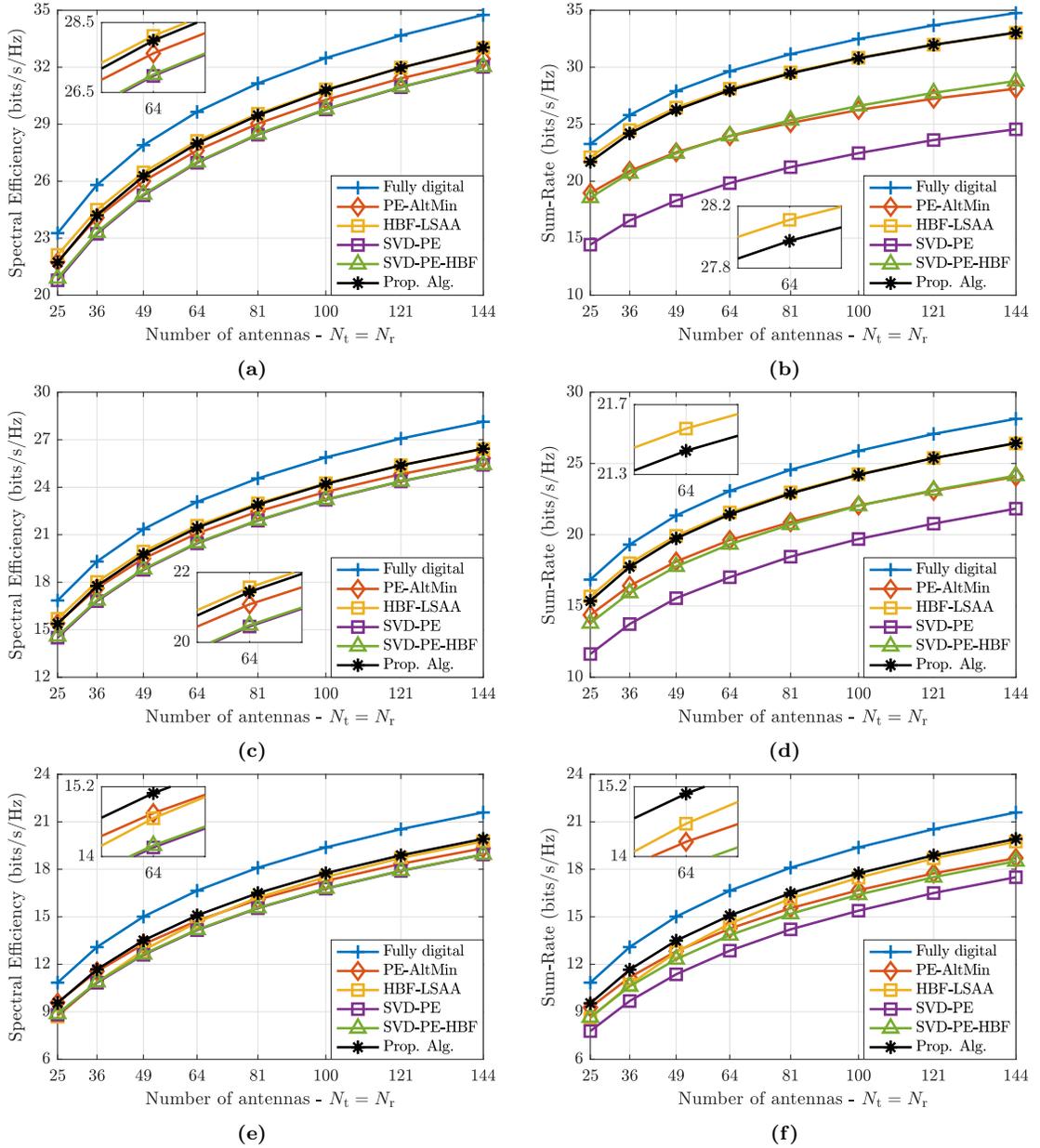
the top, center, and bottom correspond to SNR of 5 dB, 0 dB, and  $-5$  dB, respectively. From these results, one can observe that the proposed method, the PE-AltMin, and the HBF-LSAA have very similar spectral efficiency performance, particularly at SNR equals 0 dB and 5 dB, with the HBF-LSAA slightly outperforming the proposed method and with the PE-AltMin having the lowest performance among the three designs, especially at small numbers of data streams. However, observe that the performance of the HBF-LSAA is severely degraded at low SNR, achieving spectral efficiency nearly 3.6 bits/s/Hz lower than the proposed method at SNR =  $-5$  dB when transmitting eight data streams. Moreover, the SVD-PE and the SVD-PE-HBF perform much worse than the other designs. Similarly, when evaluating the sum-rate, note that the HBF-LSAA slightly outperforms the proposed method at high SNR, but it is highly degraded as the number of data streams increases at a low SNR. On the other hand, the PE-AltMin has its performance degraded at high SNR, achieving a sum-rate up to 6.6 bits/s/Hz lower than the proposed design at SNR = 5 dB when transmitting eight data streams. Overall, the proposed method has a consistent near-optimal performance for all numbers of data streams and SNR.

The third example investigates the impact of the number of antennas, assuming  $N_t = N_r$ . It is assumed that transmitter and receiver communicate through  $N_s = 4$  data streams. The results are depicted in Fig. 3.4, where, similarly to previous examples, the graphs in the left and right correspond to the spectral efficiency and sum-rate, respectively; and at the top, center, and bottom, they correspond to SNR of 5 dB, 0 dB, and  $-5$  dB, respectively. Similar to the previous results, note that the HBF-LSAA slightly outperforms the proposed method at SNR equal 0 dB and 5 dB, particularly when few antennas are used and, as the number of antennas grows, the proposed method



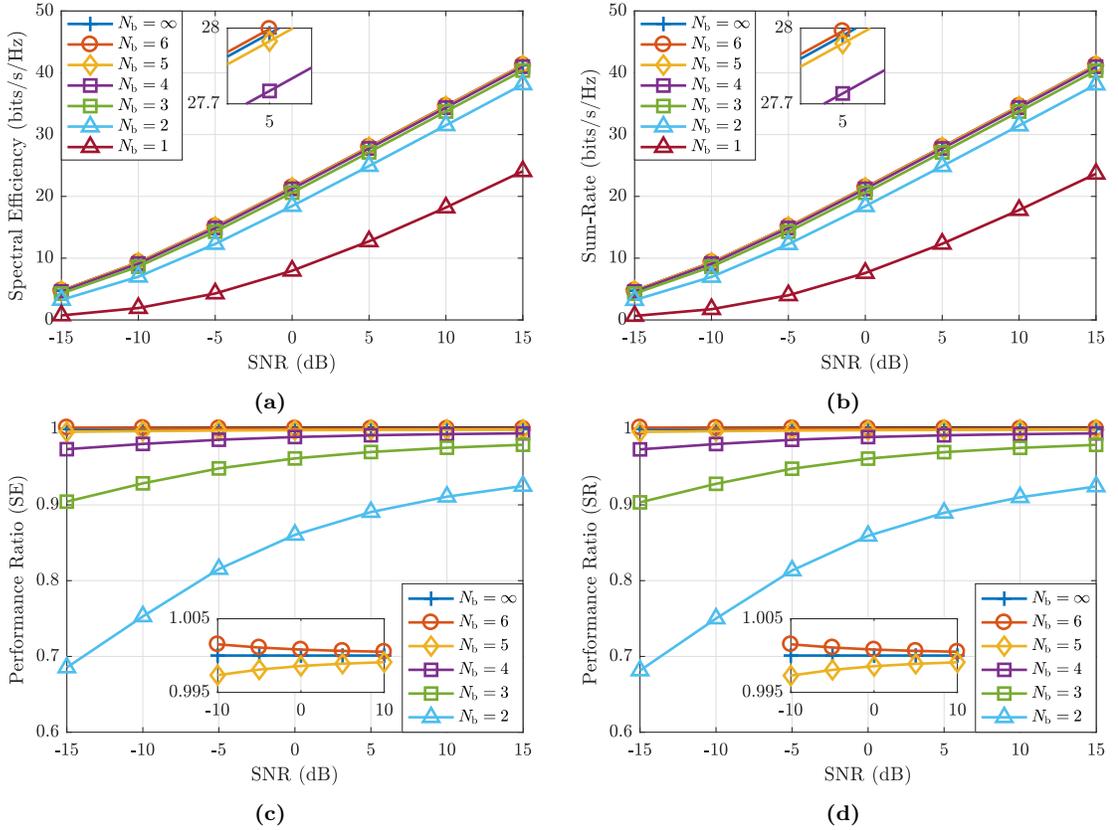
**Figure 3.3:** Example 2. SU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = N_r = 64$  antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

approaches the HBF-LSAA's performance, even surpassing it when  $N_t = N_r = 144$ . For SNR = -5 dB, the proposed design outperforms the HBF-LSAA irrespective of the number of antennas. The PE-AltMin, on the other hand, has its performance increasingly lower than the proposed method as the number of antennas increases. Moreover, the SVD-PE and the SVD-PE-HBF have significantly lower performances. These behaviours are observed for both the spectral efficiency and the sum-rate.



**Figure 3.4:** Example 3. SU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas  $N_t = N_r$  for a system communicating through  $N_s = 4$  data streams at SNR of 5 dB (top), 0 dB (center), and  $-5$  dB (bottom).

Finally, the effects of PS quantization are studied. The previous simulations have considered only unquantized (i.e., infinite resolution) phase shifters. However, quantized PS is preferred in practice, thanks to its lower energy consumption and hardware complexity [110]–[112]. Quantized PS has a finite number of feasible phase shifts, given by  $2^{N_b}$ , where the number of quantization bits  $N_b$  is usually between 4 and 6 bits [113]. To adjust the proposed algorithm for the quantized PS case, the closest quantized phases are chosen during the phase-extraction, performed in lines 10



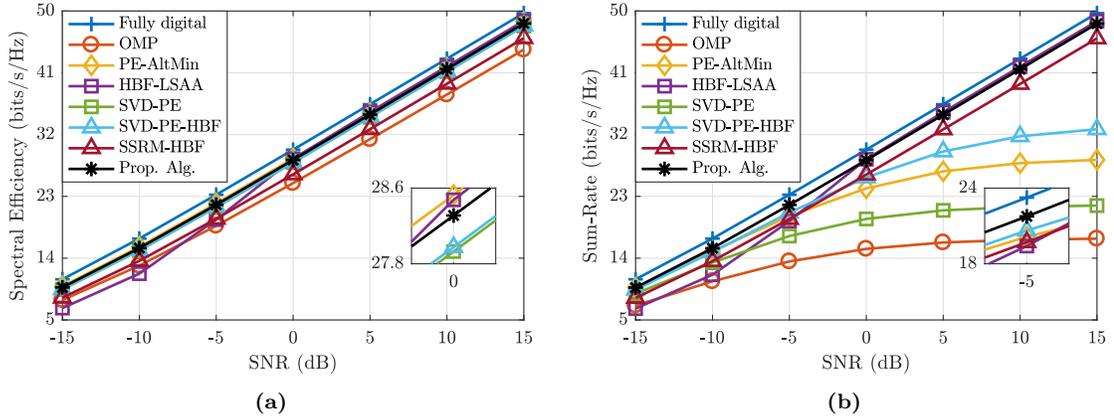
**Figure 3.5:** Example 4. SU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$ , and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).

and 12 of Alg. 3.1. In this last example, the same scenario as the first example is considered. The results are presented in Fig. 3.5, which shows the spectral efficiency (left) and the sum-rate (right) at the top, and the performance ratio between the quantized and unquantized analog beamforming, respectively, in terms of the spectral efficiency (left) and the sum-rate (right) at the bottom, for different numbers of quantization bits  $N_b$ . The results show that the proposed method attains more than 90% of its unquantized performance when using PS with as low as three quantization bits and more than 99.6% when using five or more quantization bits.

### mmWave Channel

The second set of examples investigates the proposed hybrid beamforming design performance for SU-MIMO systems under mmWave channels.

The first example considers a  $64 \times 64$  MIMO system, communicating through  $N_s = 4$  data streams. The spectral efficiency and the achieved sum-rate are evaluated for different SNR values, ranging from  $-15$  dB to  $15$  dB, and the results are shown, respectively, in Fig. 3.6a and Fig. 3.6b.

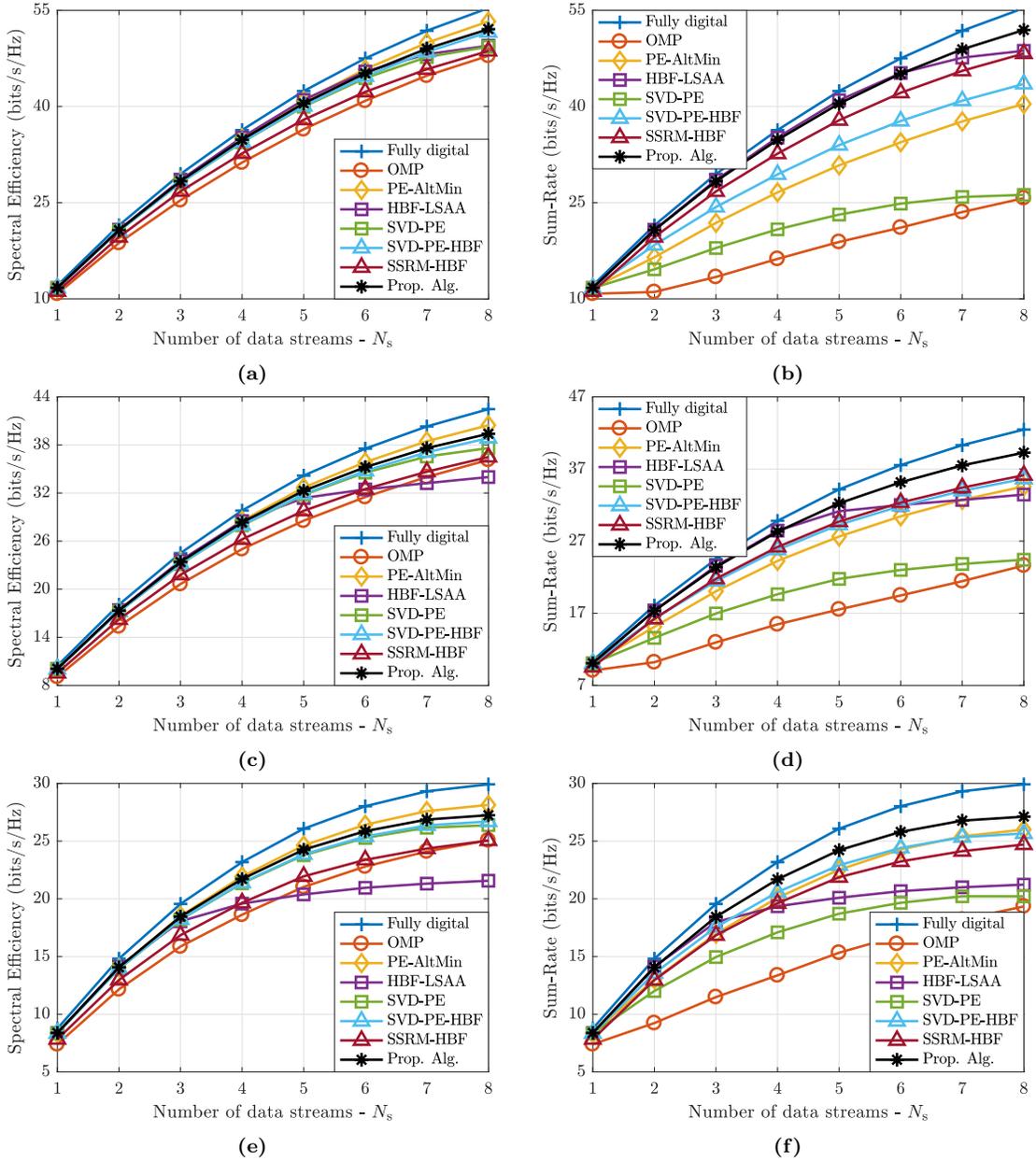


**Figure 3.6:** Example 1. SU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with  $N_t = N_r = 64$  antennas and  $N_s = 4$  data streams.

These results show that the proposed method consistently attains near-optimal performance in terms of both the spectral efficiency and the sum-rate. On the other hand, one can observe that the HBF-LSAA has poor performance at low SNR (as also observed for Rayleigh channels) and that, although the PE-AltMin can attain a spectral efficiency slightly higher than the proposed method, its sum-rate performance is severely affected as the SNR increases, achieving sum-rate nearly 20 bits/s/Hz lower than the proposed method at SNR = 15 dB. Moreover, one can note that both the OMP and the SSRM-HBF have lower spectral efficiency than the remaining designs, which could be explained by the fact that these methods have a reduced solution space for the analog precoder and combiner design problems (e.g., limited by the codebook).

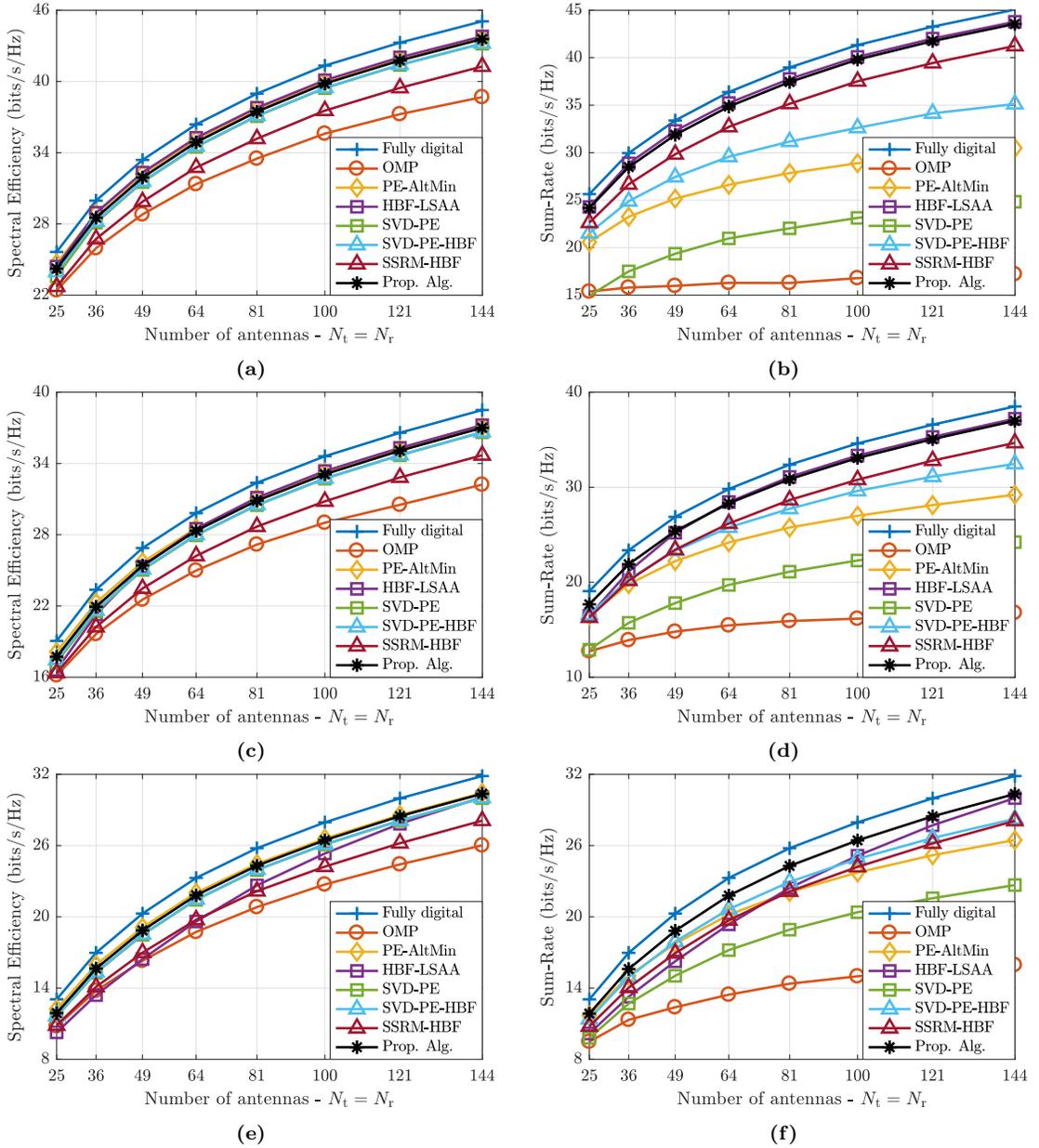
The second example evaluates the performance of a  $64 \times 64$  MIMO system when different numbers of data streams are considered. The results are presented in Fig. 3.7, which shows the spectral efficiency (left) and the sum-rate (right) for SNR of 5 dB, 0 dB, and  $-5$  dB, respectively, at the top, center and bottom graphs. From these results, one observes that although the HBF-LSAA design can marginally outperform other designs at high SNR, it is negatively impacted by the increase in the number of transmitted data streams, particularly at low SNR, even more so than in Rayleigh channels. Additionally, note that the PE-AltMin design can achieve the highest spectral efficiency in nearly all scenarios but fails to provide a high sum-rate. On the other hand, the proposed design can support a near-optimal performance in all scenarios, even outperforming the other methods in more critical scenarios, such as in low SNR or when a high number of data streams is transmitted.

The third example studies the impact of the number of antennas, assuming  $N_t = N_r$ , on the spectral efficiency and sum-rate of different hybrid beamforming designs when  $N_s = 4$  data streams are transmitted. Fig. 3.8 shows the spectral efficiency (left) and the sum-rate (right) for SNR of



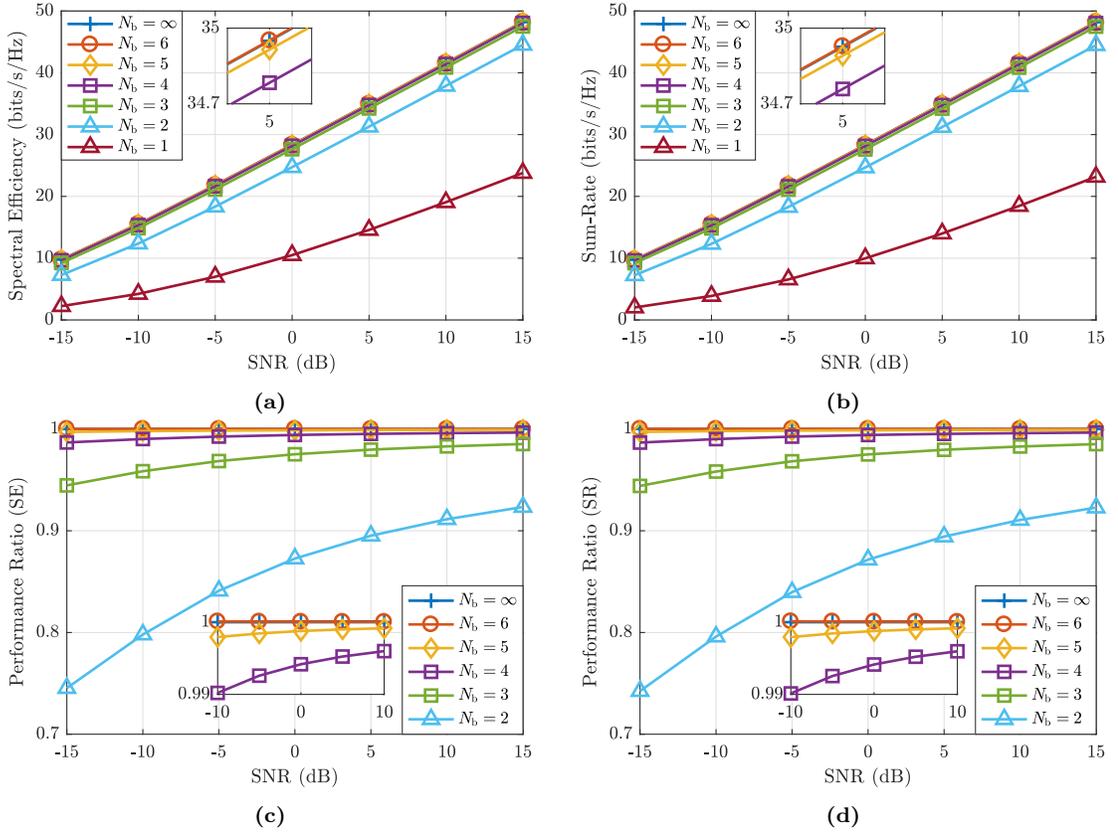
**Figure 3.7:** Example 2. SU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = N_r = 64$  antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

5 dB, 0 dB, and -5 dB, respectively, at the top, center and bottom graphs. The results show that the PE-AltMin and the HBF-LSAA designs marginally outperform the proposed method in some scenarios, achieving up to 0.45 bits/s/Hz higher spectral efficiency. However, the PE-AltMin design shows a much lower sum-rate than the proposed method, particularly at high SNR. For instance, using 144 transmit and receive antennas, the sum-rate attained by the PE-AltMin is 3.9 bits/s/Hz



**Figure 3.8:** Example 3. SU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas  $N_t = N_r$  for a system communicating through  $N_s = 4$  data streams at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom).

and 13.1 bits/s/Hz lower than that attained by the proposed method at SNR of -5 dB and 5 dB, respectively. It is worth noting that, although the HBF-LSAA can attain sum-rates slightly above those of the proposed method, particularly at high SNR, its performance is highly affected by the number of antennas, under-performing several other designs at low SNR and a low number of antennas. On the contrary, the proposed design attains near-optimal performance in all scenarios, particularly at the more critical ones, such as at low SNR and a low number of antennas.

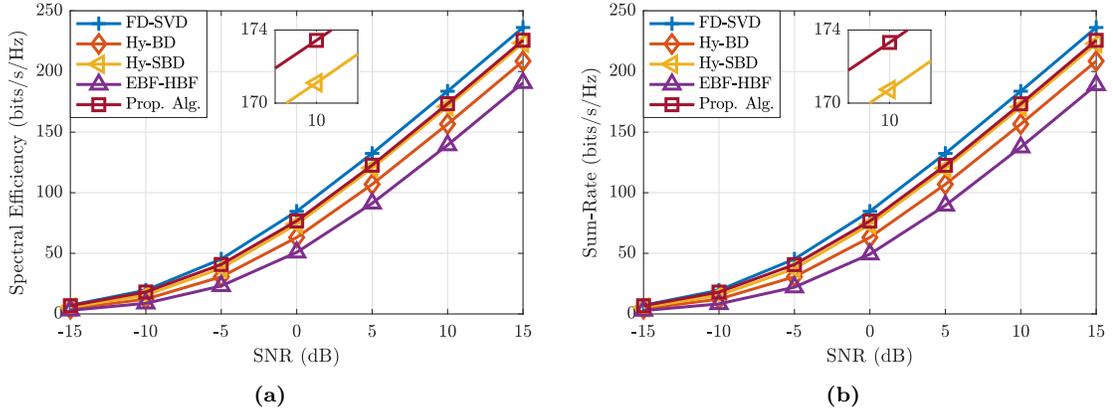


**Figure 3.9:** Example 4. SU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$  and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).

Lastly, the fourth example investigates the impact of the PS quantization on the proposed method's performance. The system parameters are the same as in the first example. The results are depicted in Fig. 3.9, which shows the spectral efficiency (left) and the sum-rate (right) at the top and the performance ratio between the quantized and unquantized analog beamforming, respectively, for each metric, at the bottom. These results show that the proposed method attains more than 94.4%, 98.7%, and 99.7% of its spectral efficiency and sum-rate for unquantized hybrid beamforming when using PS with, respectively,  $N_s = 3, 4,$  and  $5$  quantization bits; and attain the full unquantized hybrid beamforming performance using  $N_s = 6$  quantization bits.

### 3.4.2 Multi-User Systems

This section evaluates the performance of the proposed method in MU-MIMO systems. Similar to the previous section, two sets of examples are presented, considering, respectively, Rayleigh and mmWave channels. The proposed design is compared with the SVD-based fully-digital (FD-SVD)



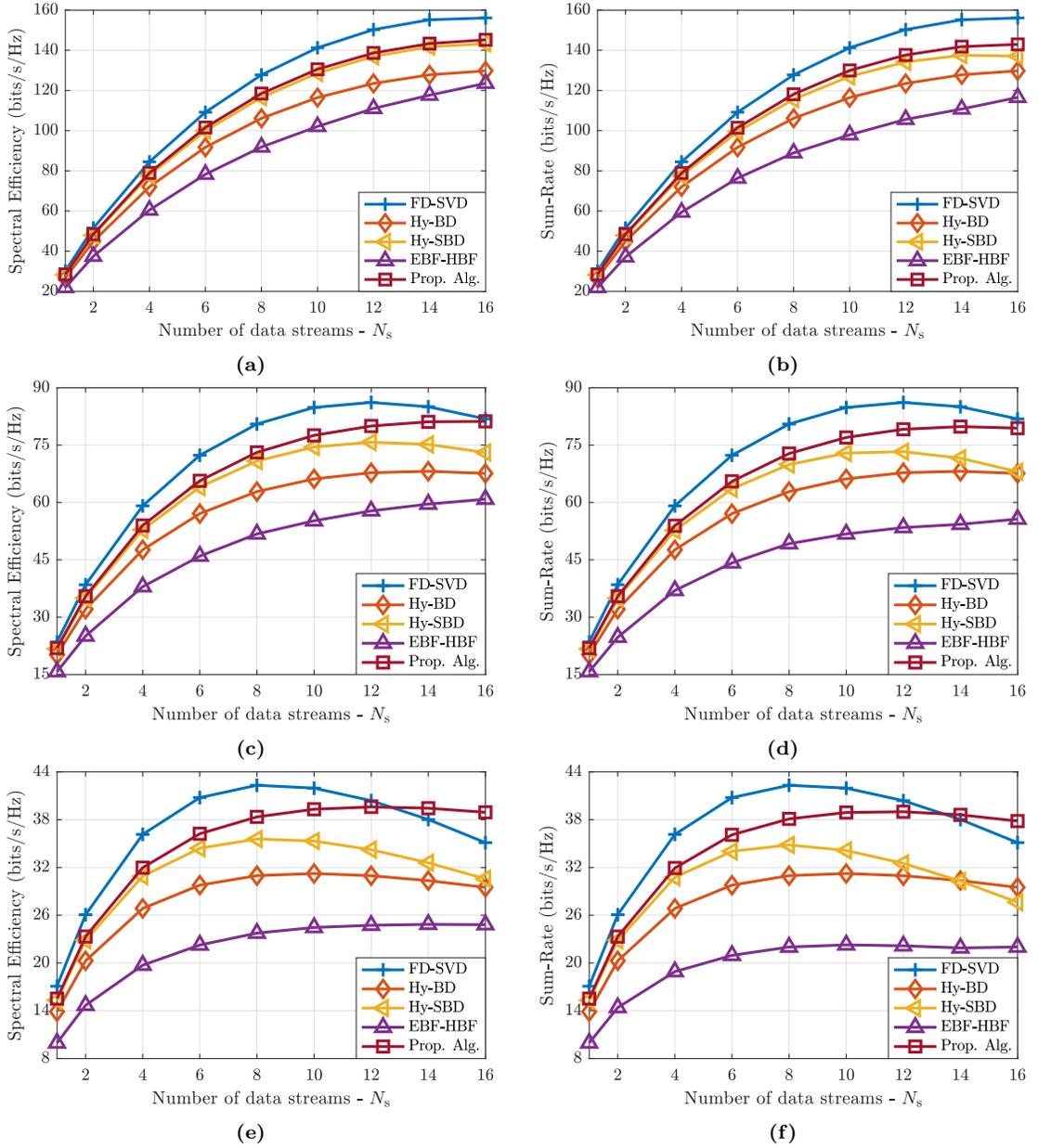
**Figure 3.10:** Example 1. MU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with  $N_t = 144$  transmit antennas and  $K = 8$  users, each having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams.

design in [96], hybrid BD (Hy-BD) [94], subspace-projection-aided BD (Hy-SBD) [96], equivalent baseband channel hybrid beamforming (EBF-HBF) [97], and hybrid regularized channel diagonalization (HRCDD) [61]. The SVD-based fully-digital (FD-SVD) design in [96] is used as a fully-digital benchmark, as it provides better performance than the traditional fully-digital BD [98]. However, note that the FD-SVD is not capacity-achieving, and hybrid beamforming designs may outperform the fully-digital design in some situations. Moreover, the HRCDD design requires the knowledge of the array response vectors for all propagation paths and can only be deployed for the mmWave channel model.

### Rayleigh Channel

The first set of example assesses the proposed hybrid beamforming design’s performance for MU-MIMO systems under Rayleigh channels.

The first example evaluates the spectral efficiency and the sum-rate under different SNR. An MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 8$  users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams is considered. The results are depicted in Fig. 3.10, which shows the spectral efficiency (left) and the sum-rate (right) for SNR ranging from  $-15$  dB to  $15$  dB. These results show that the proposed method outperforms all other hybrid beamforming designs, attaining spectral efficiency and sum-rate, respectively, up to  $3.1$  bits/s/Hz and  $3.3$  bits/s/Hz higher than the second-best performing design, the Hy-SBD. Compared with the Hy-BD and the EBF-HBF, the spectral efficiency improvement provided by the proposed method can go up to nearly  $17$  bits/s/Hz and  $35$  bits/s/Hz, respectively. The same is observed for the sum-rate.



**Figure 3.11:** Example 2. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = 144$  transmit antennas and  $K = 4$  users having  $N_r = 16$  receive antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

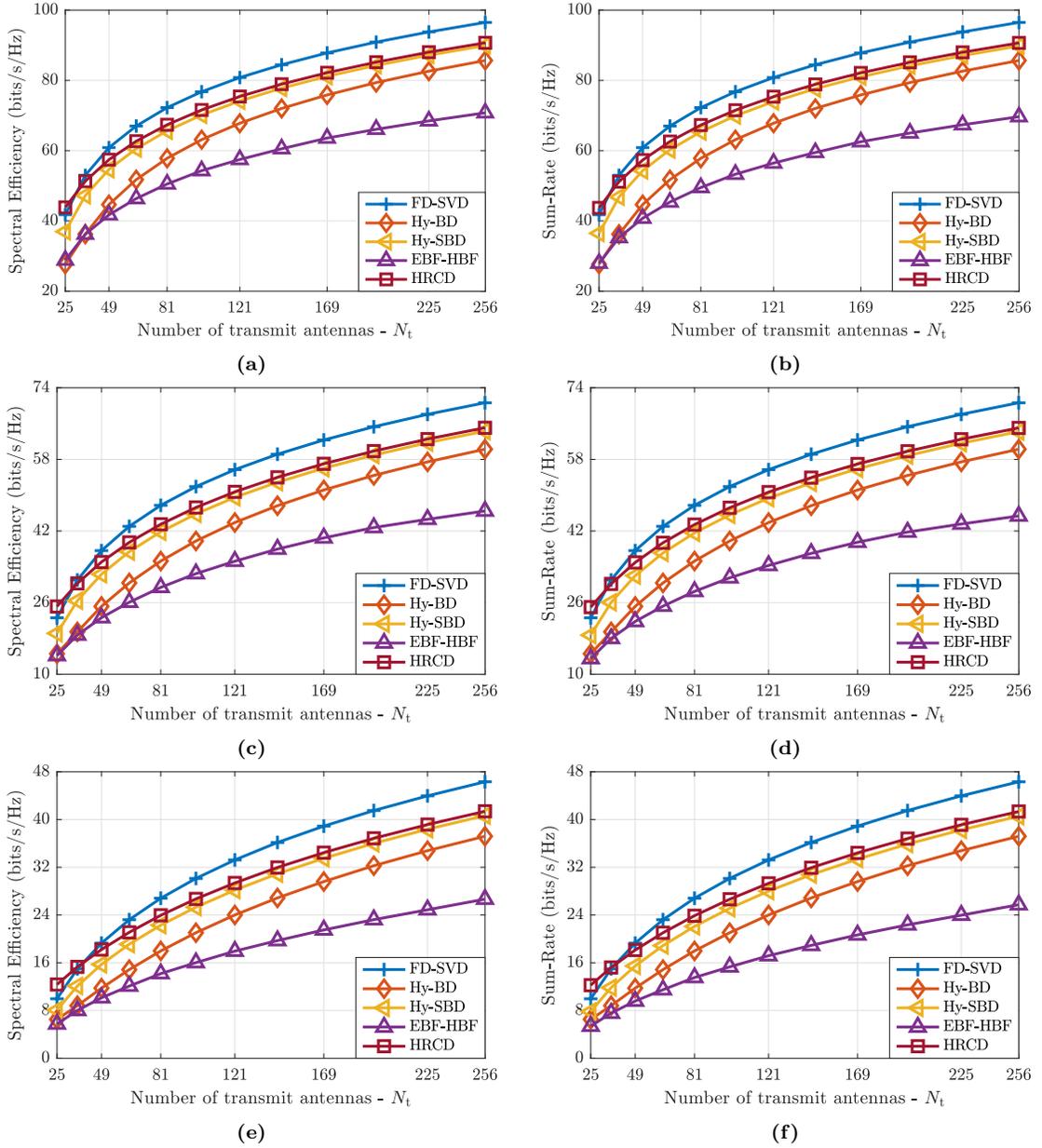
The second example investigates the impact of the number of transmitted data streams on the spectral efficiency and sum-rate performances. An MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 4$  users having  $N_r = 16$  receive antennas is considered. The results, presented in Fig. 3.11, show the spectral efficiency (left) and sum-rate (right) for different values of  $N_s$  at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). From Fig. 3.11, one can observe that the proposed approach outperforms all the other hybrid beamforming designs in all  $N_s$  range and all SNR; and

even the fully-digital design in more critical scenarios such as when transmitting a high number of data streams at low SNR. Additionally, one can note that the Hy-SBD, which has the second-best performance among the hybrid beamforming designs, has its performance deteriorated with the increase in the number of transmitted data streams, particularly at low SNR. At SNR = 5 dB, the proposed approach outperforms the Hy-SBD' spectral efficiency and sum-rate by up to, respectively, 1.9 bits/s/Hz and 5.8 bits/s/Hz; while at SNR = -5 dB, the proposed approach' spectral efficiency and sum-rate are up to 8.4 bits/s/Hz and 10.2 bits/s/Hz, respectively, higher than those of the Hy-SBD design.

The third example studies the impact of the number of transmit antennas on the spectral efficiency and sum-rate performances. An MU-MIMO system with  $K = 4$  users, each having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams, is considered. The results are shown in Fig. 3.12, following the same organization as in previous figures. From these results, one observes that, once again, the proposed method outperforms all other hybrid beamforming designs; the proposed method even outperforms the fully-digital solution at small arrays (e.g. when  $N_t = 25$ ). One also observes that the Hy-SBD, which again has the second-best performance among the hybrid beamforming designs, seems to approach the proposed method's performance, achieving spectral efficiency and sum-rate less than 1 bits/s/Hz lower than the proposed approach for large scale arrays (e.g.  $N_t \geq 196$ ). On the other hand, for small arrays, the proposed method outperforms the Hy-SBD' spectral efficiency and sum-rate by up to nearly 4 bits/s/Hz at SNR = -5 dB, and 7 bits/s/Hz at SNR = 5 dB.

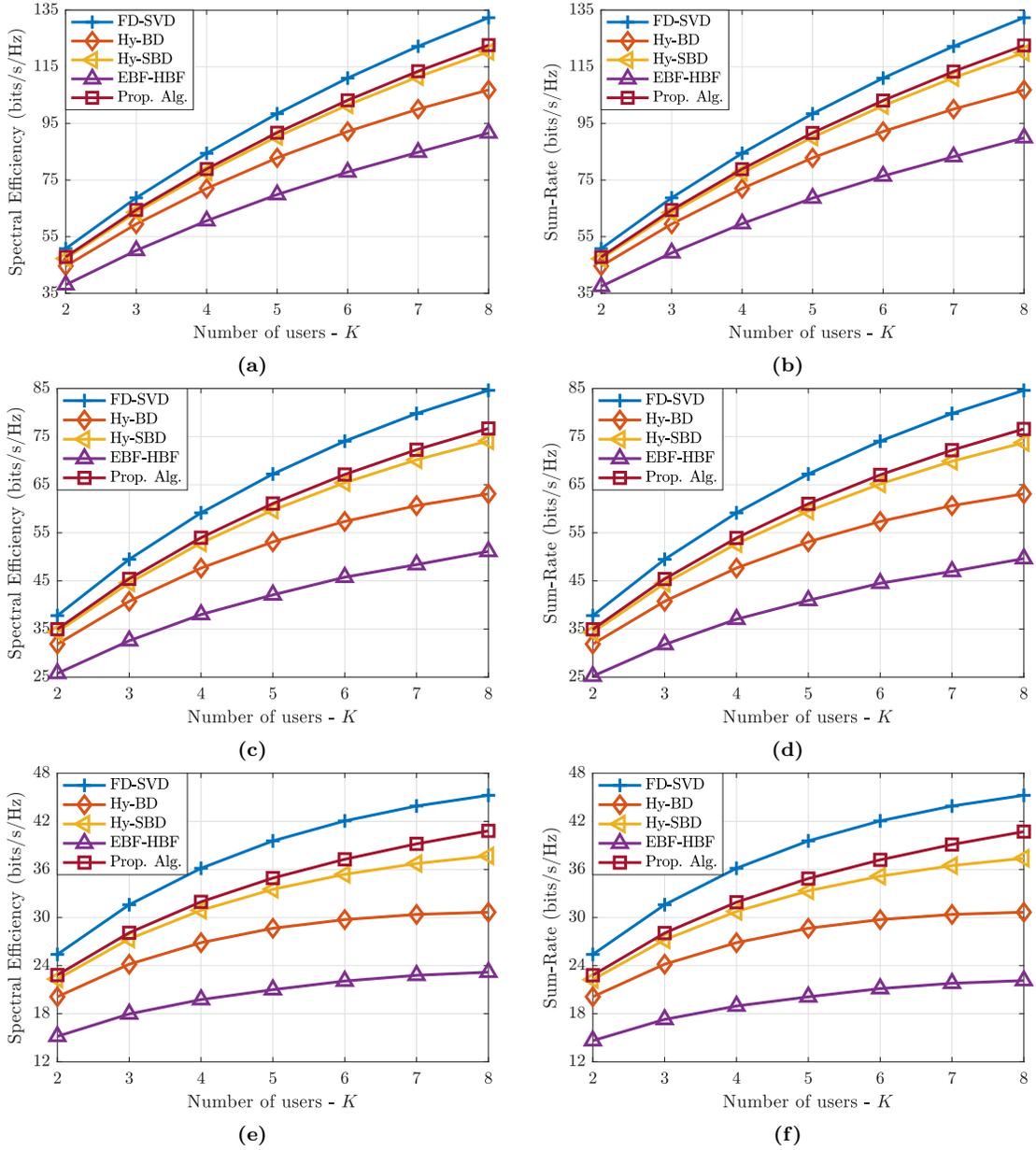
In the fourth example, the impact of the number of users on the spectral efficiency and sum-rate is investigated. An MU-MIMO system with  $N_t = 144$  transmit antennas and with users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams is considered. The results are shown in Fig. 3.13, which again follows the same organization as in previous figures. These results show that the proposed method outperforms other hybrid beamforming designs in all scenarios, especially the more critical ones, such as in low SNR or when communicating with many users. For instance, at SNR = 5 dB, the proposed approach attains spectral efficiency and sum-rate up to 2.3 bits/s/Hz and 2.6 bits/s/Hz, respectively, higher than the Hy-SBD, which has the second-best performance; at SNR = -5 dB, these performance gaps reach up to 3.1 bits/s/Hz and 3.3 bits/s/Hz, respectively.

Finally, the fifth example investigates the impact of PS quantization when considering the same scenario as in the first example. The results, depicted in Fig. 3.14, show that the proposed method can still attain nearly the same performance of the unquantized hybrid beamforming when practical



**Figure 3.12:** Example 3. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas  $N_t$  for a system with  $K = 4$  users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

PS quantization is taken into account. For instance, one observes that the proposed method achieves more than 93.0%, 97.4%, and 99.1% of its unquantized spectral efficiency and sum-rate, respectively, for  $N_b = 4, 5,$  and  $6$  quantization bits.

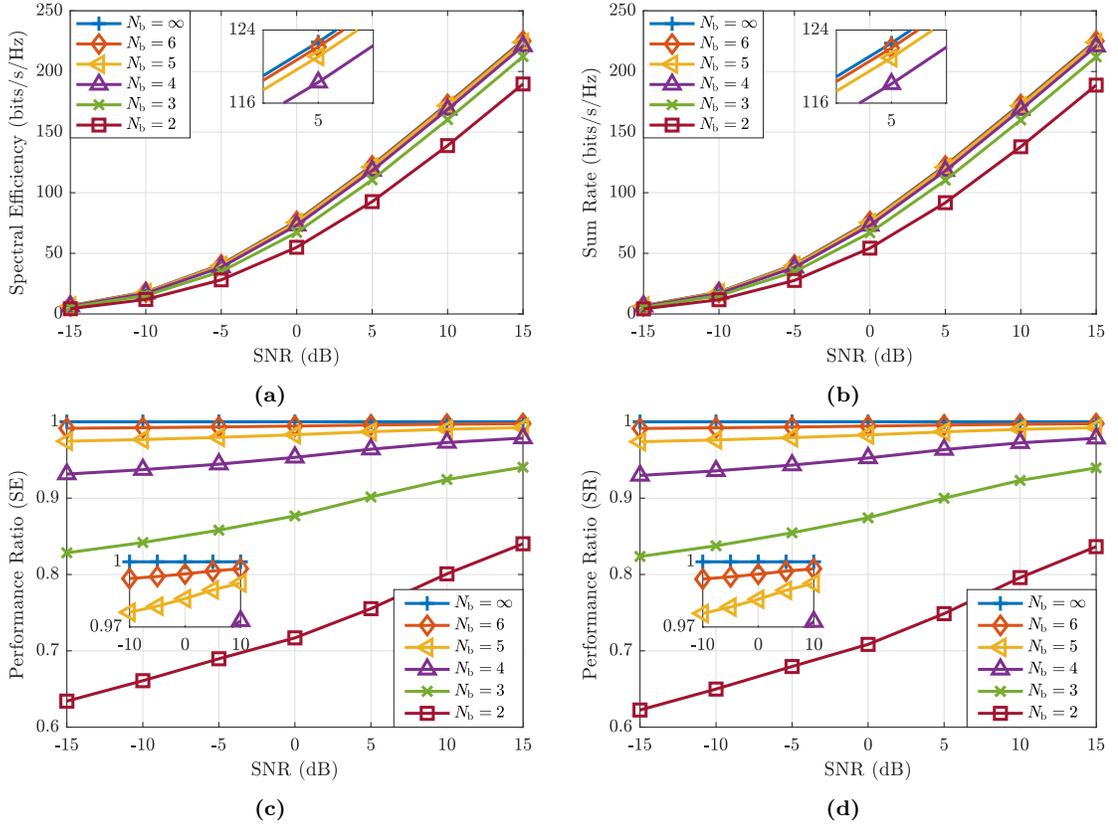


**Figure 3.13:** Example 4. MU-MIMO system and Rayleigh channel: Spectral efficiency (left) and sum-rate (right) vs. the number of users  $K$  for a system with  $N_t = 144$  transmit antennas and users having  $N_r = 16$  and communicating through  $N_s = 4$  data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

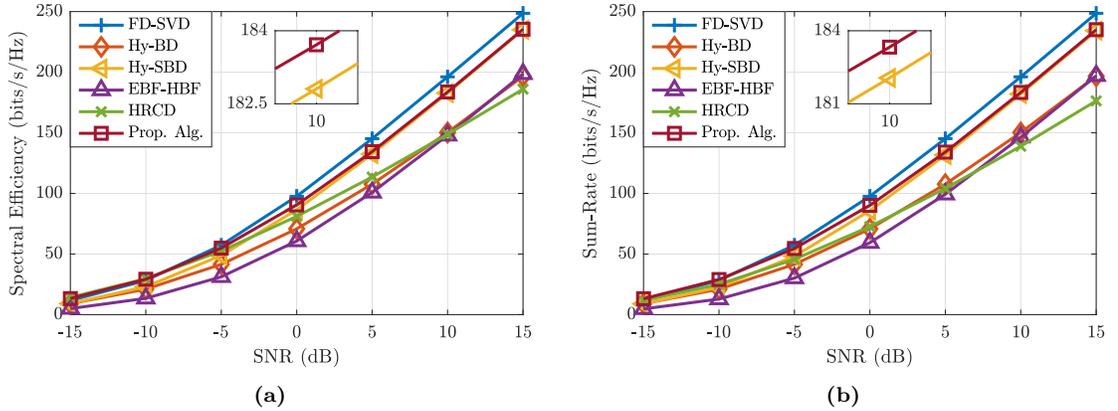
### MmWave Channel

The following set of examples evaluates the proposed hybrid beamforming design's performance and other designs in the literature for MU-MIMO systems under mmWave channels.

The first example assesses the spectral efficiency and the sum-rate of the designs under different SNR for an MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 8$  users having  $N_r = 16$



**Figure 3.14:** Example 5. MU-MIMO system and Rayleigh channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$  and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).



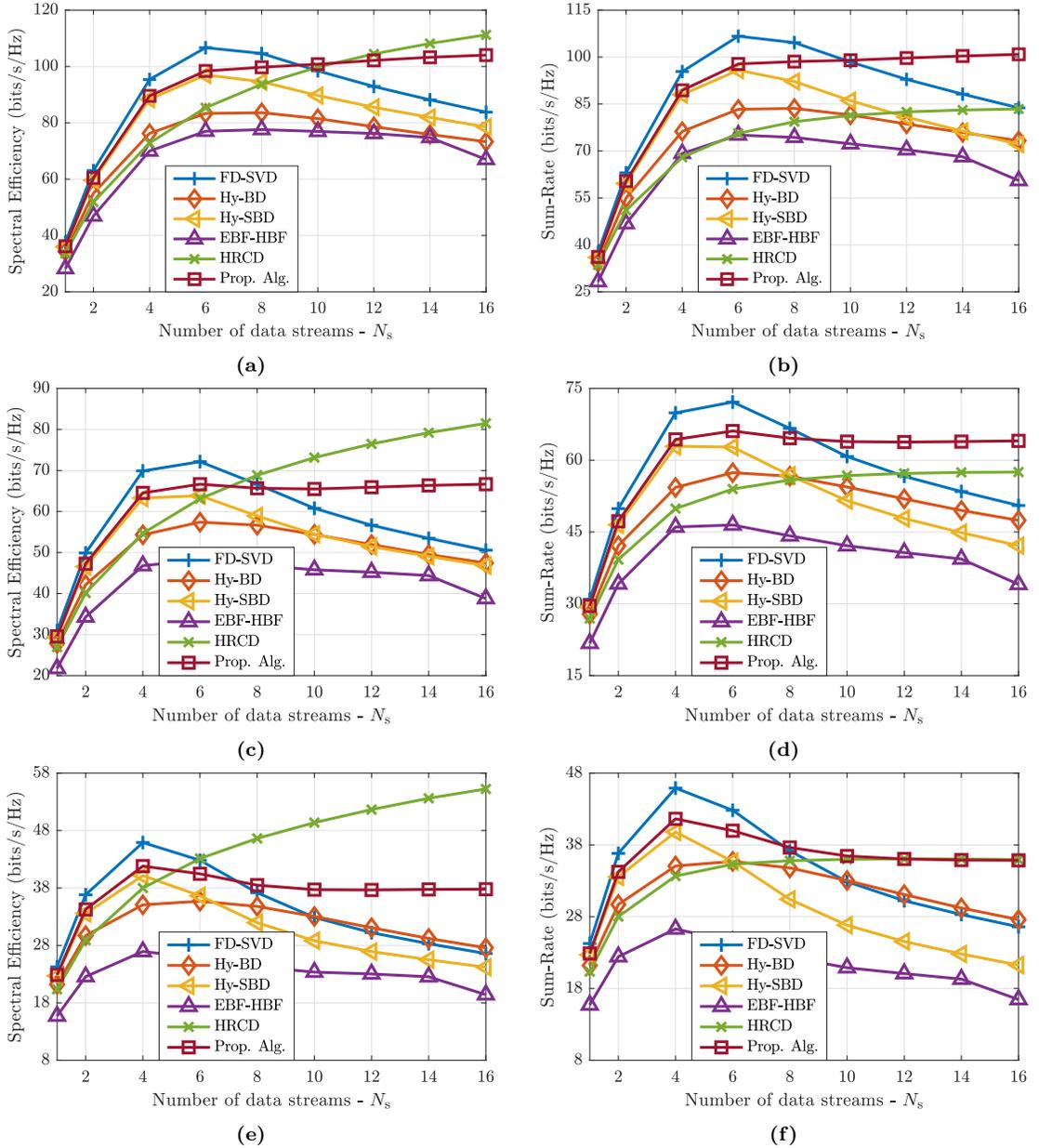
**Figure 3.15:** Example 1. MU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with  $N_t = 144$  transmit antennas and  $K = 8$  users, each having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams.

receive antennas and communicating through  $N_s = 4$  data streams. Fig. 3.15 shows the spectral efficiency (left) and the sum-rate (right) attained by the different designs. The results in Fig. 3.15

show that the proposed method outperforms the other hybrid beamforming methods in all scenarios in terms of the sum-rate. Note that the HRCD can slightly outperform the proposed method's spectral efficiency at low SNR, achieving up to 1 bits/s/Hz higher spectral efficiency at SNR = -15 dB; however, the performances of the HRCD, for both the spectral efficiency and the sum-rate, are negatively impacted by the increasing SNR, achieving up to nearly 37 bits/s/Hz lower spectral efficiency and sum-rate as compared to the proposed method at SNR = 15 dB. The Hy-SBD, which overall has the second-best performance among the hybrid beamforming designs, shows a similar behaviour as that shown for the Rayleigh channel, approaching the performance of the proposed method for high SNR; at low SNR, however, the proposed method can reach up to 6.1 bits/s/Hz and 6.4 bits/s/Hz higher spectral efficiency and sum-rate, respectively (which is nearly double the performance gap seen for Rayleigh channels).

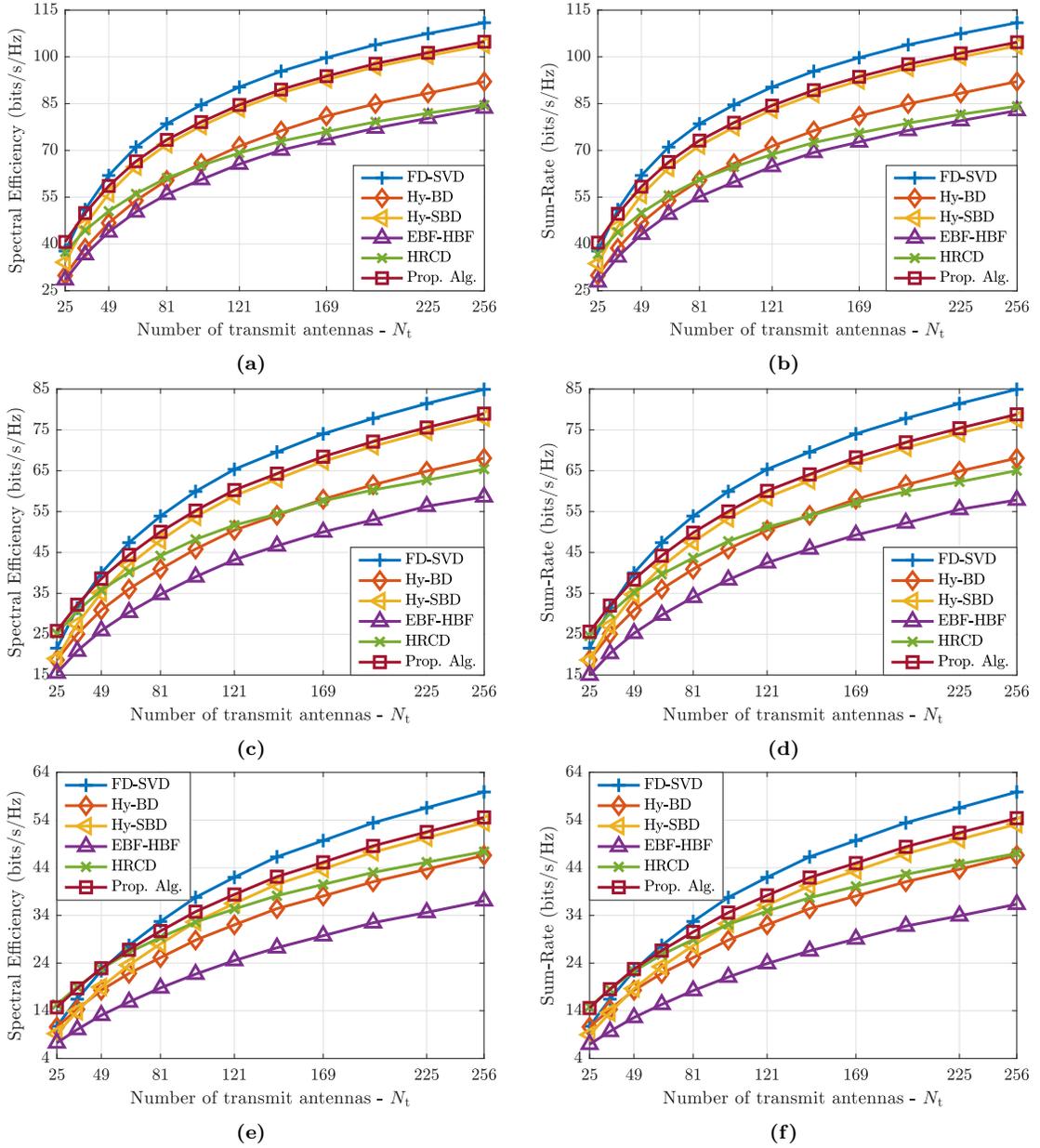
The second example investigates the impact of the number of data streams on the spectral efficiency and the sum-rate. An MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 4$  users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams is considered. The results, depicted in Fig. 3.16, present the spectral efficiency (left) and sum-rate (right) for different values of  $N_s$  at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom). From these results, one observes that the proposed method's spectral efficiency and sum-rate outperform the Hy-BD, the Hy-SBD, and the EBF-HBF in all the range of  $N_s$ , at all SNR considered in the example. On the other hand, the HRCD shows a better spectral efficiency than the proposed method at a high number of data streams; this, however, does not hold for the sum-rate, for which the proposed method outperforms the HRCD in nearly all scenarios. In fact, the HRCD performs worse than the other methods (except the EBF-HBF). Additionally, note that except for the HRCD, which also adopts the RCD precoding, all the other designs, which employ traditional BD digital precoding, have their performance severely degraded as the number of data streams grows. The traditional BD transmits each user's signal through the other users' null-space, and thus, by increasing the number of data streams, the dimension of the subspace each user can explore for transmission is reduced. In addition, note that this effect is much less present under the Rayleigh channel in Fig. 3.11 thanks to its rich scattering environment.

The third example evaluates the impact of the number of antennas on the spectral efficiency and the sum-rate for an MU-MIMO system with  $K = 4$  users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams. The results are presented in Fig. 3.17, organized as in the previous examples. The results show that the proposed design outperforms all other hybrid beamforming designs in nearly all scenarios. As observed for Rayleigh channels, for large scale



**Figure 3.16:** Example 2. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = 144$  transmit antennas and  $K = 4$  users having  $N_r = 16$  receive antennas at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

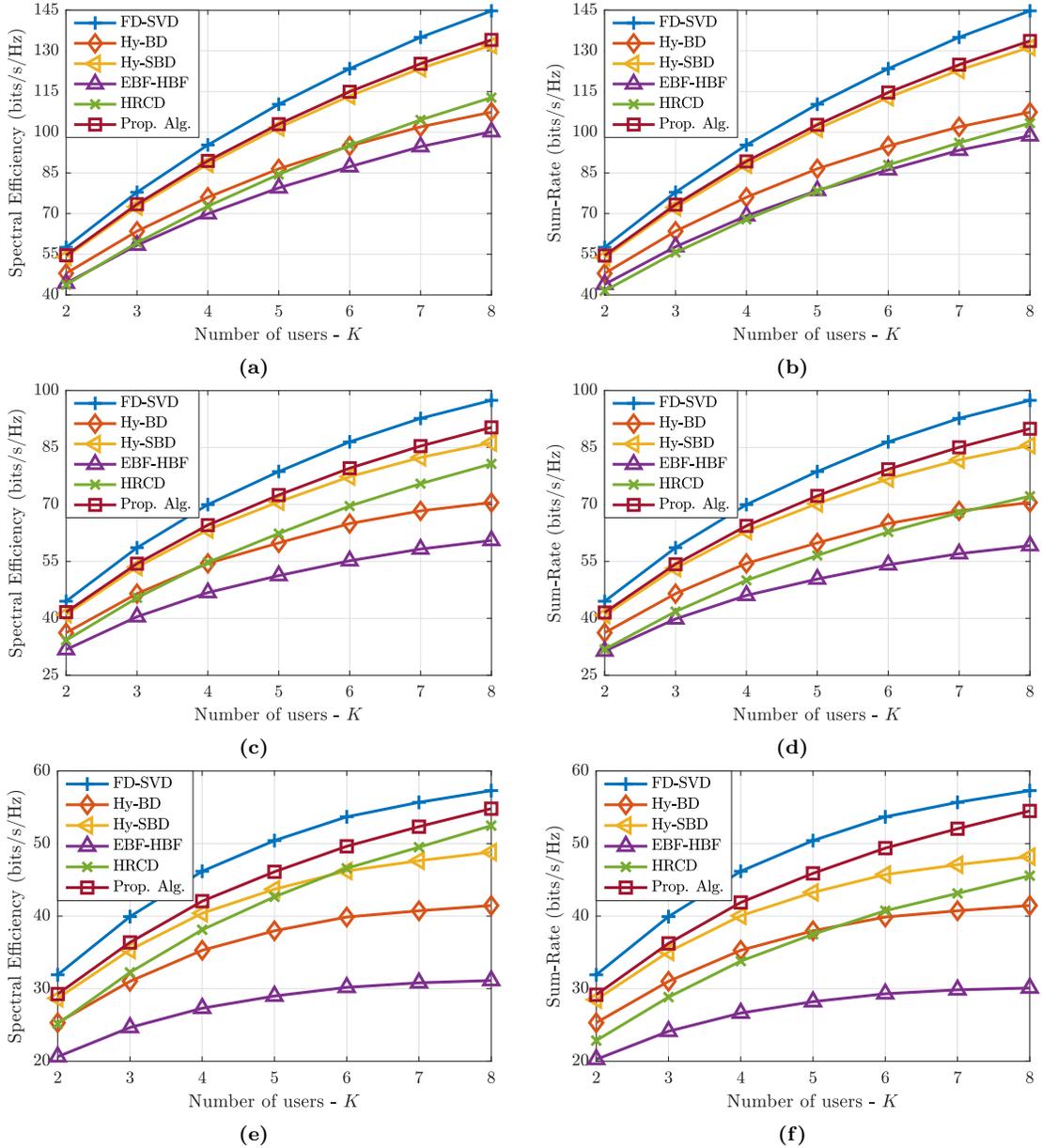
arrays, the Hy-SBD's performance approaches that from the proposed design (with, respectively, 1 bits/s/Hz and 1.2 bits/s/Hz lower spectral efficiency and sum-rate at SNR = 5 dB and  $N_t = 256$ ). In contrast, for small scale arrays, the Hy-SBD design shows a spectral efficiency and sum-rate is much lower than the proposed design (e.g., more than 6.5 bits/s/Hz lower spectral efficiency and sum-rate at SNR = 5 dB and  $N_t = 25$ ). It is also interesting to note that the HRC-D has a notably good performance at critical scenarios, such as in low SNR and a low number of antennas. However,



**Figure 3.17:** Example 3. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas  $N_t$  for a system with  $K = 4$  users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

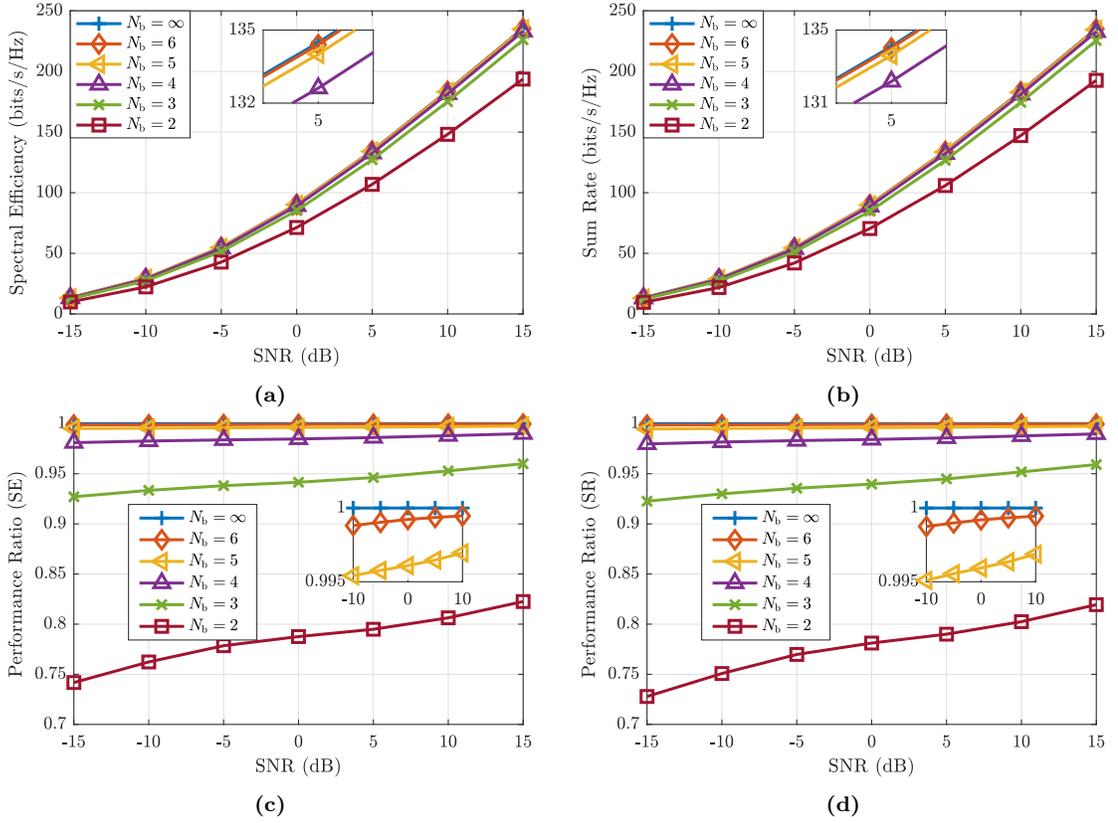
its performance decreases as the number of antennas or the SNR grows, reaching nearly 20 bits/s/Hz lower spectral efficiency and sum-rate than the proposed design at SNR = 5 dB and  $N_t = 256$ .

The fourth example investigates the effects of the number of users on the spectral efficiency and the sum-rate. An MU-MIMO system with  $N_t = 144$  transmit antennas and with users having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams is considered. The



**Figure 3.18:** Example 4. MU-MIMO system and mmWave channel: Spectral efficiency (left) and sum-rate (right) vs. the number of users  $K$  for a system with  $N_t = 144$  transmit antennas and users having  $N_r = 16$  and communicating through  $N_s = 4$  data streams at SNR = 5 dB (top), 0 dB (center), and  $-5$  dB (bottom).

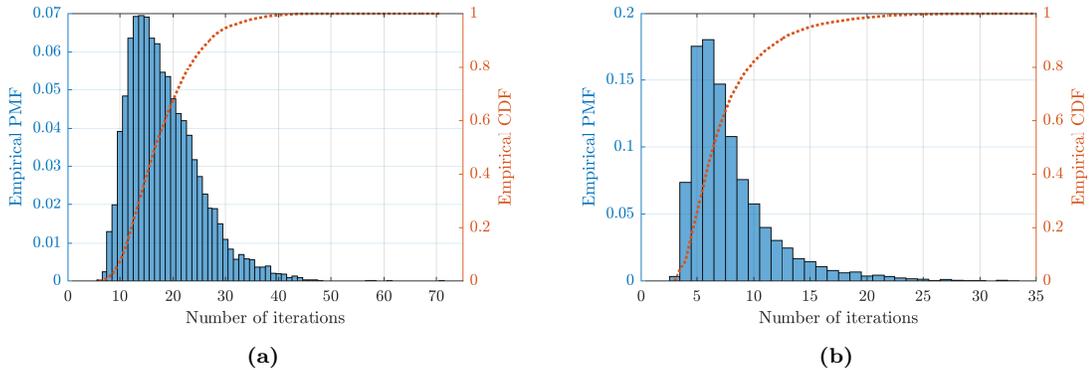
results, depicted in Fig. 3.18, show that the proposed design's spectral efficiency and sum-rate exceed those from other hybrid beamforming designs in all scenarios. Interestingly, note that the performances of the Hy-BD, Hy-SBD, and EBF-HBF begin to plateau as the number of users increases at a low SNR, while for the HRCD and the proposed method, they continue increasing. Similarly to the discussion in the second experiment, this is related to the fact that the Hy-BD, the Hy-SBD, and the EBF-HBF adopt the traditional BD precoding, which transmits the signals



**Figure 3.19:** Example 5. MU-MIMO system and mmWave channel: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$  and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).

exclusively through null-space of the other users, and thereby increasing the number of users reduces the dimension of the subspace each user can explore for transmission. In contrast, the HRC and the proposed method adopt the RCD approach, which can compromise the MUI interference suppression (e.g. by transmitting outside the other users' null-space) to support more transmit diversity. Again, note that this effect is also much less present under the Rayleigh channels, likely due to its rich scattering environment.

Lastly, the fifth example evaluates the PS quantization impact on the spectral efficiency and sum-rate attained by the proposed method. The scenario of the first example is considered. The results are presented in Fig. 3.19. From these results, one observes that using quantized PS with as low as  $N_b = 4$  quantization bits allows the proposed design to attain more than 97.8% of its unquantized performance, for both spectral efficiency and sum-rate. Similarly, it achieves more than 99.5% and 99.9% when  $N_b$  is 5 and 6, respectively.



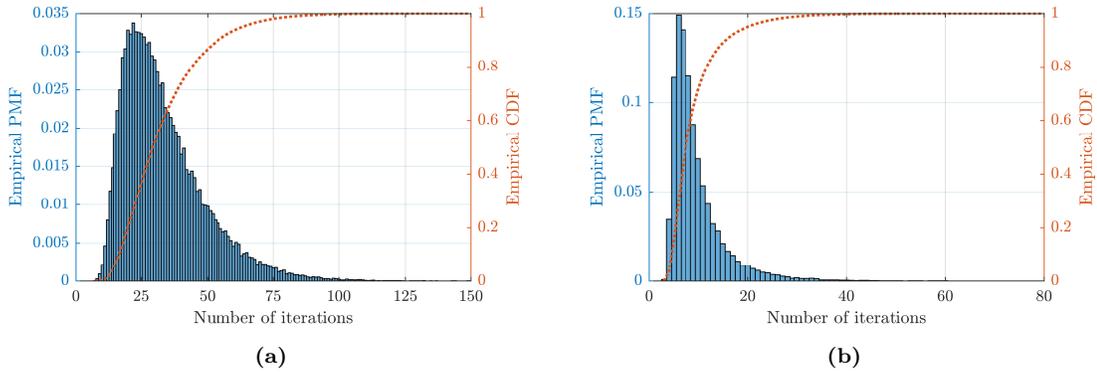
**Figure 3.20:** Convergence analysis for SU-MIMO systems. Empirical probability mass function (PMF) and empirical cumulative density function (CDF) of the number of iterations required for convergence of the proposed method considering a  $64 \times 64$  MIMO systems with  $N_s = 8$  in Rayleigh (a) and mmWave (b) channels.

### 3.4.3 Convergence Analysis

This section presents insights into the convergence of Algorithm 3.1. Recalling the proof of optimality for the block coordinate problems related to (3.15), presented in Appendix B, the solutions in (3.20) are optimal, and iterating between these solutions will lead to an increasingly better value for the objective function in (3.15). Therefore, the convergence is determined once the objective function ceases to increase. However, due to the non-convexity of (3.15), one cannot guarantee that this iterative process will converge to a global optimum solution, only to local ones. On the other hand, numerical simulation results can provide important insights into the convergence of the proposed algorithm.

Here, the number of iterations required for the algorithm to converge is evaluated. First, a  $64 \times 64$  SU-MIMO system transmitting  $N_s = 8$  data streams is considered. Fig. 3.20 shows the empirical probability mass function (PMF) and the empirical cumulative density function (CDF) of the number of iterations the inner-loop of Alg. 3.1 (i.e., lines 7-14) requires for convergence. The results in Fig. 3.20a and Fig. 3.20b correspond, respectively, to Rayleigh and mmWave channels. For Rayleigh channels, the proposed algorithm requires 18.4 iterations for convergence on average and converges within 26 and 38 iterations, respectively, in 90% and 99% of the trials. In contrast, for mmWave channels require, on average, 8 iterations and converges within 12 and 21 iterations, respectively, in 90% and 99% of the trials.

Next, an MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 8$  users, with each user having  $N_r = 16$  receive antennas and communicating through  $N_s = 8$  data streams, is considered. The empirical PMF and CDF are shown in Fig. 3.21a and Fig. 3.21b, respectively, for the Rayleigh



**Figure 3.21:** Convergence analysis for MU-MIMO systems. Empirical probability mass function (PMF) and empirical cumulative density function (CDF) of the number of iterations required for convergence of the proposed method considering a system with  $N_t = 144$  transmit antennas and  $K = 8$  users, each having  $N_r = 16$  receive antennas and communicating through  $N_s = 8$  data streams in Rayleigh (a) and mmWave (b) channels.

and mmWave channels. In this case, the results obtained for Rayleigh channel show that the proposed algorithm requires on average 33.4 iterations for convergence, and converges within 54 and 84 iterations, respectively, in 90% and 99% of the trials. Finally, for mmWave channels, the average number of iterations required by the proposed algorithm is 9.6, and the convergence is reached within 15 and 31, respectively, in 90% and 99% of the trials.

Nonetheless, extensive simulations have shown that these results in Fig. 3.20 and 3.21 remain fairly steady when varying system parameters, such as the number of antennas, users, or data streams; and have nearly the same shape when evaluated individually for each user or data stream. On the other hand, these results are highly dependent on the value of  $\varepsilon$ , the stopping criterion in Alg. 3.1, allowing the designer to trade off performance for computational complexity.

The results also show that the proposed algorithm converges much faster in mmWave channels than in Rayleigh channels. Despite the lack of an analytical convergence proof (mainly due to the difficulty of handling the non-convexity of (3.15)), it is possible to infer that this phenomenon is associated with the fast decay of the channel's singular values' energy, in contrast with the slowly decaying energy of the singular values in the Rayleigh channel, analogously to what happens for the unconstrained rank-1 approximation using the power iteration method [114]. Moreover, note that for mmWave channels, the algorithm is more likely to converge to the same (possibly the global optimal) solution, regardless of the initialization; while for Rayleigh channels, it may converge to different local optima, depending on the initialization.

### 3.4.4 Further Discussions

The results presented in Sec. 3.4.1 and Sec. 3.4.2 have demonstrated that the proposed hybrid beamforming design consistently attains near-optimal performance, outperforming other design methods for both the rich scattering (Rayleigh fading) and limited scattering (mmWave) environments in a wide range of scenario parameters.

The results obtained for SU-MIMO systems have shown that the proposed method and the HBF-LSAA have, in general, often similar spectral efficiency and sum-rate performances, with the HBF-LSAA marginally outperforming the proposed design in more favourable scenarios, such as in high SNR, but being negatively impacted in low SNR. Moreover, these negative impacts are even more prominent in mmWave channels. On the other hand, the PE-AltMin has attained high spectral efficiency performance in mmWave channels, outperforming both the HBF-LSAA and the proposed method in some scenarios. However, this does not translate in a high sum-rate as well, and in most cases, the sum-rate attained by the PE-AltMin is much lower than that attained by other methods, such as the SVD-PE-HBF, the SSRM-HBF, the HBF-LSAA, and the proposed one. It is important to note that, despite the high performance of the HBF-LSAA design, this method has the highest computational complexity. Considering a  $64 \times 64$  SU-MIMO system transmitting  $N_s = 4$  data streams and assuming  $N_{\text{ite}} = 20$  and  $|\mathcal{F}| = |\mathcal{W}| = 64$ , the computational complexity of the HBF-LSAA design is in the order of  $10^{10}$ , while all other designs have complexity in the order of  $10^5$ . Overall, the proposed design attain near-optimal spectral efficiency and sum-rate in all scenarios, standing out in more critical scenarios, such as in low SNR, a large number of data streams or a small number of antennas. Moreover, it does so with reasonable computational complexity.

For the MU-MIMO case, the proposed design has outperformed other methods in nearly all scenarios. In particular, the proposed design, and the HRCD (in mmWave channels), have shown noticeably better performances than the other designs in critical scenarios. These two designs adopt the RCD precoding, which trade off interference suppression for noise suppression and transmit diversity, whereas the other designs adopt the traditional BD precoder, which focuses exclusively on interference suppression. Nonetheless, despite attaining significantly higher spectral efficiency at large  $N_s$  values, the sum-rate performance of the HRCD design is generally below that attained by the proposed method. Moreover, considering an MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 4$  users with  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams, the computational complexity of all designs are in the order of  $10^5$ , except the EBF-HBF, which has complexity in the order of  $10^7$ .

It is essential to highlight that the simulations have shown that the proposed method can still perform very well when practical PS quantization constraints are taken into account. The results indicated that the proposed scheme could be implemented using practical quantized PS with virtually no performance loss.

Finally, it should be mentioned that the proposed design suggests the computation of the hybrid precoder and combiner be performed either on the transmitter or receiver, which however requires feedback to the other end and thus incurs extra overhead of communication resource. As such, it is worth further investigating techniques that can reduce the computational complexity and overhead of the proposed hybrid beamforming design.

### 3.5 Concluding Remarks

In this chapter, a novel hybrid beamforming design for mmWave massive MIMO has been proposed. The proposed method consists of a two-stage approach, in which the analog and digital beamforming design is performed separately. In the first stage, the analog precoder and combiner are designed using the proposed constrained channel decomposition approach, which aims to harvest the array gain provided by the massive MIMO system while simultaneously reducing the intra-user and inter-user interferences. At the second stage, two digital precoder and combiner designs are considered. For SU-MIMO systems, the digital precoder and combiner are obtained from the effective baseband channel's SVD. For MU-MIMO systems, the digital precoder and combiner are obtained using the regularized channel diagonalization method, which trades off MUI suppression for noise suppression and more transmit diversity, thus attaining a better performance in low-SNR scenarios or when the number of users or data streams grows. To confirm the effectiveness of the proposed design, numerical simulation results have been provided for a wide range of scenarios. Moreover, the performance of the proposed method under practical PS quantization is also considered.

## Chapter 4

# Hybrid Beamforming for OFDM mmWave Massive MIMO Systems

### 4.1 Introduction

The promising results obtained in initial works on narrowband mmWave massive MIMO hybrid beamforming design [24], [63], [64], [83], [87] encouraged researchers to explore beamforming techniques for frequency-selective broadband channels. In particular, a great deal of attention has been focused on orthogonal frequency-division multiplexing (OFDM) systems, where the frequency-selective channel is decomposed into a set of non-interfering parallel frequency-flat narrowband sub-channels, allowing precoders and combiners to be designed for each narrowband channel, i.e., each subcarrier [115]. Nonetheless, OFDM hybrid beamforming design is very challenging since the analog beamforming is shared among all subcarriers [83], [116]–[119].

The design of hybrid beamforming for SU-MIMO-OFDM systems has been investigated in [83], [116]–[119]. In [83], the authors extend the PE-AltMin method to present an alternating optimization approach to approximate the optimal digital solution by minimizing the sum of the distances between the hybrid beamforming and optimal unconstrained beamforming in all subcarriers. However, this approach requires *a priori* computation of the optimal unconstrained beamforming. In [116], the authors focus on the codebook design for analog and digital precoders to maximize the achievable mutual information in limited-feedback systems; their work focuses solely on the precoder, which is designed through exhaustive search over the codebooks. Under the assumption that the

channel covariance matrices at different subcarriers are asymptotically equal and share the same set of eigenvectors, the authors in [117] used the average channel covariance matrix over all subcarriers along with the narrowband analog beamforming design algorithm in [87] to design the analog beamforming for OFDM systems. The digital precoder was then designed on a per-subcarrier basis to maximize the overall spectral efficiency. In [118], the authors designed the analog precoder and combiner by extracting the phases of the eigenvectors of the averaged channel covariance matrix and the averaged conjugate-transposed-channel covariance matrix, respectively. Finally, in [119], the authors presented a beam selection approach, where the analog beamforming vectors are selected from orthogonal codebooks.

Hybrid beamforming designs for multi-user (MU) MIMO-OFDM systems have been proposed in [117], [120]–[123]. In [117], the authors first design the analog precoder by using the average channel covariance matrix and the algorithm presented for SU-MIMO system. The analog precoder is computed to improve the channel gain of all users without considering the inter-user interference. Next, authors design the digital precoder using the weighted MMSE (WMMSE) to manage the inter-user interference. In [120], the authors introduced two hybrid precoder design approaches. In the first approach, the authors adopted alternating optimization to iterate between the analog and digital precoder designs. The analog precoder was obtained through a manifold optimization framework, while two digital precoder designs for MMSE minimization were considered: one requiring a second-order cone programming solution, and the other, using an iterative WMMSE design. Given the high computational complexity of the first approach, the authors then introduced a low-complexity closed-form hybrid precoder, in which the analog precoder is obtained by extracting the phases of the covariance matrix’s eigenvectors and the digital precoder is calculated from the MMSE criteria. In [121], the authors also designed the analog precoder by extracting the phases of the covariance matrix’s eigenvectors, but the digital precoder was designed using the ZF and the regularized-ZF (RZF) approaches. However, the designs in [117], [120], [121] are limited to single-stream transmission.

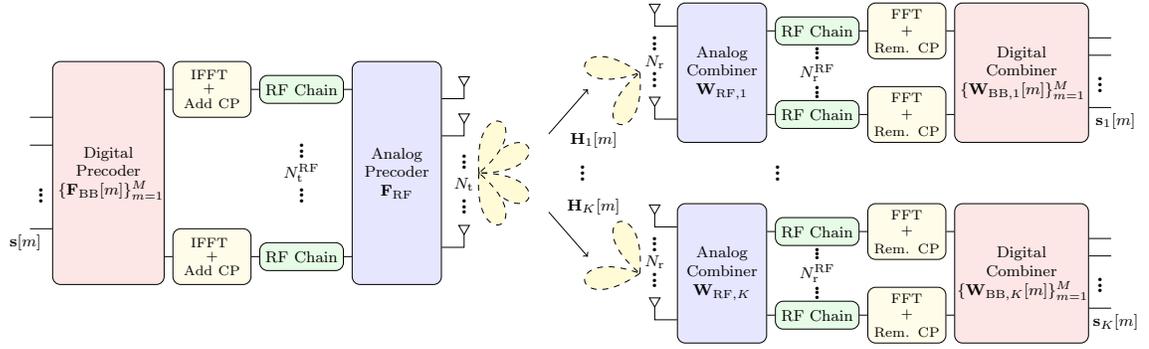
Multi-stream MU-MIMO-OFDM has been considered in a few recent works, such as [122], [123]. In [122], the authors split the hybrid precoder and hybrid combiner design and propose two methods to solve them separately. The first method designs the hybrid precoder/combiner by approximating it to the optimal unconstrained solution using a matrix factorization approach, similar to [83]. The second method uses alternate optimization to iterate between the analog and digital solutions, aiming to minimize the mean-squared error (MSE). Finally, in [123], the authors introduce a tensor-unfolding-based hybrid beamforming design to maximize the spectral efficiency. In this method,

the analog precoder and combiner are designed by extracting the phases from the channel tensor-unfolding' SVD in an attempt to maximize the gains between the base station and the mobile users. On the other hand, the digital precoders and combiners are obtained by maximizing the signal-to-leakage-plus-noise ratio (SLNR) via generalized eigendecomposition.

Given the multidimensional structure of OFDM channels, tensor decomposition emerged as a promising tool to handle OFDM system design. In [124], tensor decomposition was employed to design fully-digital precoders. In [125]–[127], tensor decomposition-based channel estimation techniques were proposed for OFDM-MIMO systems with hybrid beamforming architectures. Tensor decomposition has also been considered in hybrid beamforming design for narrowband multi-user systems [128]. However, it is yet to be explored in the OFDM hybrid beamforming design.

This chapter proposes a novel two-stage joint hybrid beamforming design method for mmWave massive MIMO-OFDM communication systems. At the first stage, a tensor representation of the OFDM channel is used to formulate the analog beamforming design as a constrained Tucker2 tensor decomposition problem. This formulation allows maximizing the sum of the effective baseband gains over every subcarrier while suppressing both intra-user interference and multi-user interference (MUI) within the same subcarrier. The proposed constrained Tucker2 decomposition is obtained by a block coordinate descent algorithm based on the projected alternate least-square (ALS) method. The digital precoder and combiner for SU systems are obtained from the optimal SVD-based solution, considering the effective baseband channel (which takes into account the effects of the analog beamforming). For MU systems, the digital beamforming is obtained using the regularized channel diagonalization (RCD) method [61]. This approach provides a balance between transmit diversity and noise and MUI suppression, in contrast to other MU-MIMO precoder and combiner designs, such as the traditional block diagonalization, which only addresses the MUI suppression; and therefore, it can perform more efficiently in low-SNR scenarios or the scenarios with many users or data streams. Extensive simulation results are presented to evaluate and compare the proposed design with other existing methods. These results show the proposed approach can outperform other methods in most scenarios and that such a high performance can also be achieved when more realistic hardware constraints, such as the quantization of PSs, are considered.

This chapter is organized as follows: Section 4.2 introduces the system and channel models and the problem formulation; Section 4.3 proposed the novel hybrid beamforming design for mmWave massive MIMO-OFDM systems based on the proposed constrained tensor decomposition approach; Section 4.4 presents the simulation results for different system's setups and compares the proposed design with other methods in the literature; finally, Section 4.5 presents the concluding remarks.



**Figure 4.1:** Downlink of an MU-MIMO-OFDM system with hybrid beamforming architecture.

## 4.2 System Model and Problem Formulation

This section introduces the system and channel models, and the problem statement.

### 4.2.1 System Model

Here, the downlink of an MU-MIMO-OFDM system with a hybrid beamforming structure, as shown in Fig. 4.1, is considered. In this system, a base station equipped with  $N_t$  antennas and  $N_t^{\text{RF}}$  RF chains is transmitting to  $K$  mobile users through  $M$  subcarriers. The mobile users are equipped with  $N_r$  antennas and  $N_r^{\text{RF}}$  RF chains, supporting  $N_s$  data-streams per subcarrier. The system satisfies  $N_s \leq N_r^{\text{RF}} \leq N_r$  and  $KN_s \leq N_t^{\text{RF}} \leq N_t$  to ensure effective multi-stream communication.

At the base station, the transmitted symbol vector is precoded on a per-subcarrier basis, using the digital precoder  $\mathbf{F}_{\text{BB}}[m] = [\mathbf{F}_{\text{BB},1}[m], \dots, \mathbf{F}_{\text{BB},K}[m]] \in \mathbb{C}^{N_t^{\text{RF}} \times KN_s}$ , where  $\mathbf{F}_{\text{BB},k}[m]$  is the digital precoder associated with the  $k$ -th user at the  $m$ -th subcarrier. The symbols then go through the OFDM modulation, where the cyclic prefix addition (CPA) and the inverse fast Fourier transform (IFFT) operations take place, assuming that the cyclic prefix's length is big enough to prevent inter-carrier interference (ICI) [115], [129]. The modulated symbols are then precoded on the analog domain, using the analog precoder  $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t \times N_t^{\text{RF}}}$ , common to all subcarriers. The phase shifters in the analog precoder imposes a constant modulus constraint, i.e.,  $|\mathbf{F}_{\text{RF}}(i, j)| = 1/\sqrt{N_t} \forall i, j$ . Moreover,  $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}[m]\|_{\text{F}}^2 = P_{\text{T}}$  represents the per-subcarrier total transmitted power at the  $m$ -th subcarrier.

At the  $k$ -th user, the received signal is combined in the analog domain using the analog combiner matrix  $\mathbf{W}_{\text{RF},k} \in \mathbb{C}^{N_r \times N_r^{\text{RF}}}$ , which is also constant modulus-constrained, such that  $|\mathbf{W}_{\text{RF},k}(i, j)| = 1/\sqrt{N_r} \forall i, j$ , and common to all subcarriers. The signal then goes through the OFDM demodulation,

which applies the fast Fourier transform (FFT) and the cyclic prefix removal (CPR) operations. The demodulated signal is finally combined in the digital domain through the combining matrix  $\mathbf{W}_{\text{BB},k}[m] \in \mathbb{C}^{N_r^{\text{RF}} \times N_s}$  on a per-subcarrier basis. Therefore, the received symbol is

$$\hat{\mathbf{s}}_k[m] = \sqrt{\rho_k} \mathbf{W}_{\text{BB},k}^{\text{H}}[m] \mathbf{W}_{\text{RF},k}^{\text{H}} \mathbf{H}_k[m] \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[m] \mathbf{s}[m] + \mathbf{W}_{\text{BB},k}^{\text{H}}[m] \mathbf{W}_{\text{RF},k}^{\text{H}} \mathbf{n}_k[m]. \quad (4.1)$$

where the transmitted symbol vector in the  $m$ -th subcarrier  $\mathbf{s}[m] = [\mathbf{s}_1^{\text{T}}[m], \mathbf{s}_2^{\text{T}}[m], \dots, \mathbf{s}_K^{\text{T}}[m]]^{\text{T}} \in \mathbb{C}^{KN_s \times 1}$  is assumed i.i.d. with zero-mean and  $\mathbb{E}[\mathbf{s}[m] \mathbf{s}^{\text{H}}[m]] = \mathbf{I}_{KN_s}$ , with  $\mathbf{s}_k[m] \in \mathbb{C}^{N_s \times 1}$  being the  $k$ -th user's symbols, and  $\mathbf{n}_k[m] \in \mathbb{C}^{N_r \times 1}$  is the noise vector, such that  $\mathbf{n}_k$  is i.i.d.  $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_r})$ . Since OFDM modulation is considered, the channel for each subcarrier is represented by a flat-fading narrowband channel  $\mathbf{H}_k[m] \in \mathbb{C}^{N_r \times N_t}$ , where  $\mathbf{H}_k[m]$  represents the  $k$ -th user's frequency-domain channel matrix at the  $m$ -th subcarrier.

By defining the aggregated analog combining matrix  $\mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\}$ , the aggregated channel matrix for the  $m$ -th subcarrier  $\mathbf{H}[m] = [\mathbf{H}_1^{\text{T}}[m], \mathbf{H}_2^{\text{T}}[m], \dots, \mathbf{H}_K^{\text{T}}[m]]^{\text{T}}$ , and the  $k$ -th user's effective baseband channel at the  $m$ -th subcarrier  $\bar{\mathbf{H}}_k[m] = \mathbf{W}_{\text{RF},k}^{\text{H}} \mathbf{H}_k[m] \mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_r^{\text{RF}} \times N_t^{\text{RF}}}$ , the aggregated multiuser effective baseband channel at the  $m$ -th subcarrier can then be represented as

$$\begin{aligned} \mathbf{H}_{\text{eff}}[m] &= \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H}[m] \mathbf{F}_{\text{RF}} \\ &= [\bar{\mathbf{H}}_1^{\text{T}}[m] \ \bar{\mathbf{H}}_2^{\text{T}}[m] \ \dots \ \bar{\mathbf{H}}_K^{\text{T}}[m]]^{\text{T}} \\ &= \begin{bmatrix} \mathbf{W}_{\text{RF},1}^{\text{H}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{W}_{\text{RF},K}^{\text{H}} \end{bmatrix} \begin{bmatrix} \mathbf{H}_1[m] \\ \vdots \\ \mathbf{H}_K[m] \end{bmatrix} \mathbf{F}_{\text{RF}}. \end{aligned} \quad (4.2)$$

Next, defining the hybrid precoder and combiner for  $k$ -th user at the  $m$ -th subcarrier, respectively, as  $\mathbf{F}_{k,m} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB},k}[m]$  and  $\mathbf{W}_{k,m} = \mathbf{W}_{\text{RF},k} \mathbf{W}_{\text{BB},k}[m]$ , and considering that Gaussian symbols are transmitted, the spectral efficiency for the  $m$ -th subcarrier is [94]

$$R_m = \sum_{k=1}^K \log \left| \mathbf{I}_{N_s} + \frac{\rho_k}{\sigma_n^2} \mathbf{R}_k^{-1} \mathbf{W}_{k,m}^{\text{H}} \mathbf{H}_k[m] \mathbf{F}_{k,m} \mathbf{F}_{k,m}^{\text{H}} \mathbf{H}_k^{\text{H}}[m] \mathbf{W}_{k,m} \right| \quad (4.3)$$

where

$$\mathbf{R}_k = \frac{\rho_k}{\sigma_n^2} \sum_{j=1, j \neq k}^K \mathbf{W}_{k,m}^{\text{H}} \mathbf{H}_k[m] \mathbf{F}_{j,m} \mathbf{F}_{j,m}^{\text{H}} \mathbf{H}_k^{\text{H}}[m] \mathbf{W}_{k,m} + \mathbf{W}_{k,m}^{\text{H}} \mathbf{W}_{k,m}$$

is the interference-plus-noise covariance matrix.

Following the discussion in 3.2.1, linear detection techniques are preferred in practical deployments due to the high computational complexity of nonlinear techniques. In this case, the  $N_s$  data streams are treated independently at the receiver and thus, the sum-rate of all data streams is a reasonable performance metric. The achievable sum-rate of the  $m^{\text{th}}$  subcarrier is

$$R_{\text{sum},m} = \sum_{k=1}^K \sum_{i=1}^{N_s} \log_2(1 + \text{SINR}_{k,i,m}) \quad (4.4)$$

where  $\text{SINR}_{k,i,m}$  is the SINR of the  $i$ -th data stream of the  $k$ -th user in the  $m$ -th subcarrier, defined as

$$\text{SINR}_{k,i,m} = \frac{S}{I + N} \quad (4.5)$$

with S, I, and N being, respectively, the signal, interference, and noise power, given by

$$\begin{aligned} S &= \rho_k |\mathbf{W}_{k,m}(:,i)^H \mathbf{H}_k[m] \mathbf{F}_{k,m}(:,i)|^2 \\ I &= \sum_{j=1, j \neq i}^{N_s} \rho_k |\mathbf{W}_{k,m}(:,i)^H \mathbf{H}_k[m] \mathbf{F}_{k,m}(:,j)|^2 \\ &\quad + \sum_{l=1, l \neq k}^K \sum_{j=1}^{N_s} \rho_l |\mathbf{W}_{k,m}(:,i)^H \mathbf{H}_k[m] \mathbf{F}_{l,m}(:,j)|^2 \\ N &= \sigma_n^2 \mathbf{W}_{k,m}(:,i)^H \mathbf{W}_{k,m}(:,i). \end{aligned}$$

## 4.2.2 OFDM Channel Model

In this chapter, only the mmWave channel model is considered. The mmWave channel is modeled with the extended Saleh-Valenzuela channel model, which combines the contributions of  $N_{\text{cl}}$  scattering clusters, each having  $N_{\text{ray}}$  propagation paths. For OFDM systems, the frequency-selective channel is separated into  $M$  narrowband frequency-flat sub-channels with non-interfering subcarriers. The frequency-domain channel matrix for the  $k$ -th user in the  $m$ -th subcarrier is given by [130]

$$\mathbf{H}_k[m] = \sqrt{\frac{N_r N_t}{N_{\text{cl}} N_{\text{ray}}}} \sum_{i=0}^{N_{\text{cl}}-1} \sum_{l=0}^{N_{\text{ray}}-1} \alpha_{il} \mathbf{a}_r(\phi_{il}^r, \theta_{il}^r) \mathbf{a}_t^H(\phi_{il}^t, \theta_{il}^t) e^{-j2\pi i \frac{(m-1)}{M}} \quad (4.6)$$

where, as in Chapter 3,  $\alpha_{il} \sim \mathcal{CN}(0, \sigma_{\alpha,i}^2)$  corresponds to the complex gain of the  $l^{\text{th}}$  scattering ray in the  $i^{\text{th}}$  cluster and  $\sigma_{\alpha,i}^2$  is the average power of the  $i^{\text{th}}$  cluster, such that  $\sum_{i=1}^{N_{\text{cl}}} \sigma_{\alpha,i}^2 = \sqrt{N_r N_t / N_{\text{cl}} N_{\text{ray}}}$  to ensures  $\mathbb{E}[\|\mathbf{H}_k[m]\|_{\text{F}}^2] = N_r N_t$ ;  $\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)$  and  $\mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)$  are, respectively, the receiver and transmitter array response vectors;  $\phi_{il}^t$  and  $\theta_{il}^t$  are the azimuth and elevation angles of departure

(AoD), and  $\phi_{il}^r$  and  $\theta_{il}^r$  are the azimuth and elevation angles of arrival (AoA), all following a Laplacian distribution, with mean  $\phi_i^t, \theta_i^t, \phi_i^r, \theta_i^r$  uniformly-distributed over  $[-\pi, \pi)$ , and angular spread of  $\sigma_\phi^t, \sigma_\theta^t, \sigma_\phi^r, \sigma_\theta^r$  [24], [87]. Furthermore, a  $\sqrt{N} \times \sqrt{N}$  uniform square planar array (USPA) with array response vector given in (3.10) is considered.

### 4.2.3 Problem Formulation

The objective of this chapter is to design hybrid beamforming for MIMO-OFDM systems to maximize its average sum-rate across all subcarrier, which is mathematically expressed as

$$\begin{aligned}
 & \max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}[m],} && \frac{1}{M} \sum_{m=1}^M R_{\text{sum},m} \\
 & \mathbf{W}_{\text{RF},k}, \mathbf{W}_{\text{BB},k}[m] && \\
 & \text{s.t.} && \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}} \\
 & && \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\
 & && \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[m]\|_{\text{F}}^2 = P_{\text{T}}
 \end{aligned} \tag{4.7}$$

where  $\mathcal{F}_{\text{RF}}$  and  $\mathcal{W}_{\text{RF}}$  represent the feasible analog precoder and combiner solutions, i.e., matrices with constant modulus entries. For convenience, it is assumed  $N_{\text{t}}^{\text{RF}} = KN_{\text{r}}^{\text{RF}} = KN_{\text{s}}$ . Furthermore, perfect channel state information (CSI) knowledge is also assumed. Although estimating the OFDM channel when using a hybrid beamforming structure is quite challenging, efficient OFDM channel estimation methods are presented in [122], [131]–[134] and references therein.

Following the discussion presented in Chapter 3, the maximal sum-rate is, in general, attained when the data streams are transmitted by multiple non-interfering parallel sub-channels, i.e., by designing precoders and combiners that diagonalize the effective channel. However, attaining such a diagonalization, which is already a difficult task in narrowband hybrid beamforming, is even more challenging in OFDM systems. In OFDM systems, the analog precoders and combiners must be designed to diagonalize the effective baseband channel at all subcarriers simultaneously.

Inspired by the constrained channel matrix decomposition approach for narrowband hybrid beamforming systems proposed in the previous chapter, the present chapter proposes a constrained tensor decomposition-based hybrid beamforming design for MIMO-OFDM systems. The proposed design can increase the effective baseband gain while also addressing the intra-user and inter-user interference at the analog beamforming level. Likewise, the digital beamforming is also designed to further mitigate the interference and improve the system's spectral efficiency.

## 4.3 Constrained Channel Tensor Decomposition-based Hybrid Beamforming Design for MIMO-OFDM Systems

A novel joint hybrid precoder and combiner design for mmWave massive MIMO-OFDM systems is proposed in this section. Like the narrowband design proposed in Chapter 3, this is a two-stage approach, where analog and digital beamforming are designed separately. Here, the analog precoder and combiner are designed by the proposed low-rank constrained Tucker2 decomposition, which aims to maximize the sum of the effective baseband channel gains at all subcarriers while suppressing interference among data streams within the same subcarrier by seeking orthogonality among the analog precoders/combiners of different data streams. Due to the phase-shifters' constant-modulus constraint, the orthogonality among the analog precoders/combiners, which ensures the effective channel's diagonalization, cannot be entirely attained by the analog beamforming alone. Therefore, at the second stage, the digital precoder and combiner are designed to further attain the complete diagonalization. Since the digital precoder and combiner are computed on a per-subcarrier basis, they are obtained using the narrowband designs presented in Sections 3.3.2 and 3.3.3, respectively, for single-user and multi-user systems.

### 4.3.1 Analog Precoder and Combiner Design

To design the analog precoder and combiner, a novel constrained Tucker2 tensor decomposition is proposed, in which an additional constant modulus constraint is imposed to the traditional Tucker decomposition (presented in Appendix C.2). For details on tensor notations and operations and the Tucker decomposition, the reader should refer to Appendix C.

First, let the channel tensor  $\mathcal{H} \in \mathbb{C}^{KN_r \times N_t \times M}$  be defined such that each subcarrier's aggregated channel matrix corresponds to a frontal slice of  $\mathcal{H}$ , i.e., the mode-1 matricization of  $\mathcal{H}$  is given as

$$\mathbf{H}_{(1)} = [\mathbf{H}[1] \quad \mathbf{H}[2] \quad \cdots \quad \mathbf{H}[M]]. \quad (4.8)$$

Likewise, let the effective channel tensor  $\mathcal{H}_{\text{eff}} \in \mathbb{C}^{KN_s \times KN_s \times M}$  be defined as

$$\mathcal{H}_{\text{eff}} = \mathcal{H} \times_1 \mathbf{W}_{\text{RF}}^H \times_2 \mathbf{F}_{\text{RF}}^T \times_3 \mathbf{I}_M \quad (4.9)$$

such that the  $m$ -th frontal slice of  $\mathcal{H}_{\text{eff}}$  is the effective baseband at the  $m$ -th subcarrier and the mode-1 matricization of  $\mathcal{H}_{\text{eff}}$  is given as

$$\mathbf{H}_{\text{eff}(1)} = [\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H}[1] \mathbf{F}_{\text{RF}} \quad \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H}[2] \mathbf{F}_{\text{RF}} \quad \cdots \quad \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H}[M] \mathbf{F}_{\text{RF}}]. \quad (4.10)$$

Similar to the design in Chapter 3, the proposed analog precoder and combiner design seeks the constant modulus constrained orthogonal factor matrices  $\mathbf{W}_{\text{RF}}$  and  $\mathbf{F}_{\text{RF}}$  that best approximate the channel tensor, i.e.,  $\mathcal{H} \approx \mathcal{S} \times_1 \mathbf{W}_{\text{RF}} \times_2 \mathbf{F}_{\text{RF}}^* \times_3 \mathbf{I}_M$ , so that  $\mathcal{H}_{\text{eff}} \approx \mathcal{S}$ , with  $\mathcal{S}_{(1)} = [\boldsymbol{\Sigma}[1] \cdots \boldsymbol{\Sigma}[M]]$ , where  $\boldsymbol{\Sigma}[m]$  is a  $KN_s \times KN_s$  diagonal matrix. This problem is written as the following constrained Tucker2 decomposition (or approximation) problem:

$$\begin{aligned} \min_{\mathbf{W}_{\text{RF}}, \mathbf{F}_{\text{RF}}, \mathcal{S}} \quad & \|\mathcal{H} - \mathcal{S} \times_1 \mathbf{W}_{\text{RF}} \times_2 \mathbf{F}_{\text{RF}}^* \times_3 \mathbf{I}_M\|^2 \\ \text{s.t.} \quad & \mathcal{S} \in \mathbb{C}^{KN_s \times KN_s \times M} \\ & \mathbf{W}_{\text{RF},k}^{\text{H}} \mathbf{W}_{\text{RF},k} = \mathbf{I}_{N_s} \\ & \mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{F}_{\text{RF}} = \mathbf{I}_{N_s} \\ & \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \cdots, \mathbf{W}_{\text{RF},K}\} \\ & \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\ & \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}}. \end{aligned} \quad (4.11)$$

As shown in Appendix C.2, problem (4.11) can be equivalently solved by

$$\begin{aligned} \max_{\mathbf{W}_{\text{RF}}, \mathbf{F}_{\text{RF}}} \quad & \|\mathcal{H} \times_1 \mathbf{W}_{\text{RF}}^{\text{H}} \times_2 \mathbf{F}_{\text{RF}}^{\text{T}} \times_3 \mathbf{I}_M\|^2 \\ \text{s.t.} \quad & \mathbf{W}_{\text{RF},k}^{\text{H}} \mathbf{W}_{\text{RF},k} = \mathbf{I}_{N_s} \\ & \mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{F}_{\text{RF}} = \mathbf{I}_{N_s} \\ & \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \cdots, \mathbf{W}_{\text{RF},K}\} \\ & \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\ & \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}}. \end{aligned} \quad (4.12)$$

However, due to the orthogonality and constant modulus constraints, finding the analog precoder and combiner matrices that solves (4.12) is an intractable problem.

To handle this issue, problem (4.12) is broken down into  $KN_s$  successive rank-(1, 1) constrained Tucker2 approximation subproblems, where each subproblem corresponds to the analog precoder and

combiner vector pair of one user's data stream. The rank-(1,1) constrained Tucker2 approximation subproblem is expressed by

$$\begin{aligned}
& \max_{\mathbf{w}_j, \mathbf{f}_j} \quad \left\| \mathcal{H} \times_1 \mathbf{w}_j^H \times_2 \mathbf{f}_j^T \times_3 \mathbf{I}_M \right\|^2 \\
& \text{s.t.} \quad |[\mathbf{w}_j]_n| = 1/\sqrt{N_r}, \quad \forall n \in 1, \dots, N_r \\
& \quad \mathbf{w}_j^H \mathbf{w}_j = 1 \\
& \quad \mathbf{w}_j^H \mathbf{w}_l = 0, \quad \forall j \neq l \\
& \quad \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\} \\
& \quad |[\mathbf{f}_j]_n| = 1/\sqrt{N_t}, \quad \forall n \in 1, \dots, N_t \\
& \quad \mathbf{f}_j^H \mathbf{f}_j = 1 \\
& \quad \mathbf{f}_j^H \mathbf{f}_l = 0, \quad \forall j \neq l
\end{aligned} \tag{4.13}$$

where  $\mathbf{F}_{\text{RF}} = [\mathbf{f}_1 \cdots \mathbf{f}_{KN_s}]$  and  $\mathbf{W}_{\text{RF}} = [\mathbf{w}_1 \cdots \mathbf{w}_{KN_s}]$  and  $j, l \in \{1, \dots, KN_s\}$ . However, the rank-(1,1) approximation problem in (4.13) is still intractable.

To further simplify problem (4.13), one can note that the aggregated analog combiner matrix has a block-diagonal structure, such that the analog combiners vectors of different users are already orthogonal, and thus, only the orthogonality among analog combiners of the same user must be guaranteed. Besides, these combiners only interact with the same user's channel matrices. Furthermore, the orthogonality constraint can be slightly relaxed since the digital combiner and precoder can further ensure the MUI interference suppression. Therefore, problem (4.13) can be recast as

$$\begin{aligned}
& \max_{\mathbf{w}_{k,i}, \mathbf{f}_{k,i}} \quad \left\| \mathcal{H}_{\text{res},k}^{(\gamma)} \times_1 \mathbf{w}_{k,i}^H \times_2 \mathbf{f}_{k,i}^T \times_3 \mathbf{I}_M \right\|^2 \\
& \text{s.t.} \quad |[\mathbf{w}_{k,i}]_n| = 1/\sqrt{N_r}, \quad \forall n \in 1, \dots, N_r \\
& \quad \mathbf{w}_{k,i}^H \mathbf{w}_{k,i} = 1 \\
& \quad |[\mathbf{f}_{k,i}]_n| = 1/\sqrt{N_t}, \quad \forall n \in 1, \dots, N_t \\
& \quad \mathbf{f}_{k,i}^H \mathbf{f}_{k,i} = 1
\end{aligned} \tag{4.14}$$

where the aggregated channel tensor  $\mathcal{H}$  is replaced by the residual channel tensor  $\mathcal{H}_{\text{res}}^{(\gamma)}$  in order to reinforce the relaxed orthogonality constraints,  $\mathcal{H}_{\text{res},k}^{(\gamma)}$  is a sub-tensor, having only the residual channel tensor corresponding to the  $k^{\text{th}}$  user, and  $\mathbf{w}_{k,i}$  and  $\mathbf{f}_{k,i}$  are the analog precoder and combiner for the  $i$ -th data stream of the  $k$ -th user, such that  $\mathbf{w}_{k,i}$  is the  $i$ -th column of  $\mathbf{W}_{\text{RF},k}$  and  $\mathbf{f}_{k,i}$  is the  $[(k-1)N_s + i]$ -th column of  $\mathbf{F}_{\text{RF}}$ , where  $k \in \{1, \dots, K\}$  and  $i \in \{1, \dots, N_s\}$ .

The solution to problem (4.14) is then obtained with the block coordinate descent-type alternate least square (ALS) method by alternately solving<sup>1</sup>

$$\begin{aligned}
& \max_{\mathbf{w}_{k,i}} \left\| \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i}) \right\|_{\text{F}}^2 \\
& \text{s.t.} \quad |[\mathbf{w}_{k,i}]_n| = 1/\sqrt{N_r}, \quad \forall n \in 1, \dots, N_r \\
& \quad \mathbf{w}_{k,i}^H \mathbf{w}_{k,i} = 1
\end{aligned} \tag{4.15}$$

and

$$\begin{aligned}
& \max_{\mathbf{f}_{k,i}} \left\| \mathbf{f}_{k,i}^H \mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}) \right\|_{\text{F}}^2 \\
& \text{s.t.} \quad |[\mathbf{f}_{k,i}]_n| = 1/\sqrt{N_t}, \quad \forall n \in 1, \dots, N_t \\
& \quad \mathbf{f}_{k,i}^H \mathbf{f}_{k,i} = 1
\end{aligned} \tag{4.16}$$

until convergence.

Problems (4.15) and (4.16) are still non-convex problems due to the constant modulus constraints. Therefore, feasible solutions can be obtained by relaxing the non-convex constraint, solving the relaxed convex problem, and finally projecting the solution onto the non-convex solution space. Thus, by dropping the constant modulus constraint, the relaxed alternate least square problems are, respectively,

$$\begin{aligned}
& \max_{\mathbf{w}_{k,i}} \left\| \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i}) \right\|_{\text{F}}^2 \\
& \text{s.t.} \quad \mathbf{w}_{k,i}^H \mathbf{w}_{k,i} = 1
\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}
& \max_{\mathbf{f}_{k,i}} \left\| \mathbf{f}_{k,i}^H \mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}) \right\|_{\text{F}}^2 \\
& \text{s.t.} \quad \mathbf{f}_{k,i}^H \mathbf{f}_{k,i} = 1.
\end{aligned} \tag{4.18}$$

The solution to the problems in (4.17) and (4.18) are obtained using the Lagrange multiplier

---

<sup>1</sup>Note that the objective function in (4.14), (4.15), and (4.16) are equivalent. This comes from the following two identities:

$$\begin{aligned}
\mathcal{Y} = \mathcal{H}_{\text{res},k}^{(\gamma)} \times_1 \mathbf{w}_{k,i}^H \times_2 \mathbf{f}_{k,i}^T \times_3 \mathbf{I}_M & \Leftrightarrow \mathbf{Y}_{(1)} = \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i}^T)^T = \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i}) \\
& \Leftrightarrow \mathbf{Y}_{(2)} = \mathbf{f}_{k,i}^T \mathbf{H}_{\text{res},k(2)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}^H)^T = \mathbf{f}_{k,i}^T \mathbf{H}_{\text{res},k(2)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}^*)
\end{aligned}$$

and

$$\begin{aligned}
\left\| \mathcal{H}_{\text{res},k}^{(\gamma)} \times_1 \mathbf{w}_{k,i}^H \times_2 \mathbf{f}_{k,i}^T \times_3 \mathbf{I}_M \right\|^2 &= \left\| \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i}) \right\|_{\text{F}}^2 \\
&= \left\| \mathbf{f}_{k,i}^T \mathbf{H}_{\text{res},k(2)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}^*) \right\|_{\text{F}}^2 = \left\| \mathbf{f}_{k,i}^H \mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i}) \right\|_{\text{F}}^2
\end{aligned}$$

method. For instance, for problem (4.17), the Lagrangian function is defined as

$$\mathcal{L}(\mathbf{w}_{k,i}, \lambda) = \mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H} \mathbf{w}_{k,i} - \lambda (\mathbf{w}_{k,i}^H \mathbf{w}_{k,i} - 1). \quad (4.19)$$

The optimal  $\mathbf{w}_{k,i}$  is then obtained by setting the gradient of  $\mathcal{L}(\mathbf{w}_{k,i}, \lambda)$  w.r.t.  $\mathbf{w}_{k,i}^*$  to zero, i.e.,

$$\nabla_{\mathbf{w}_{k,i}^*} \mathcal{L}(\mathbf{w}_{k,i}, \lambda) = \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H} \mathbf{w}_{k,i} - \lambda \mathbf{w}_{k,i} = \mathbf{0} \quad (4.20)$$

which leads to the following eigen-problem:

$$\mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H} \mathbf{w}_{k,i} = \lambda \mathbf{w}_{k,i}. \quad (4.21)$$

Analogously, the solution to (4.18) will also lead to an eigenproblem. Therefore, since the goal is to maximize the objective function, the optimal solution to problems (4.17) and (4.18) are, respectively, given by the principal eigenvectors (i.e., the eigenvector associated with the largest eigenvalue) of

$$\mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H}$$

and

$$\mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H) \mathbf{H}_{\text{res},k(2)}^{(\gamma)T}.$$

Solving these eigenproblems for each ALS iteration is computationally expensive, therefore, the solution can be approximated by a single iteration of the power-iteration method [114], which yields

$$\hat{\mathbf{w}}_{k,i} = \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H} \mathbf{w}_{k,i} \quad (4.22)$$

and

$$\hat{\mathbf{f}}_{k,i} = \mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H) \mathbf{H}_{\text{res},k(2)}^{(\gamma)T} \mathbf{f}_{k,i}. \quad (4.23)$$

The relaxed solutions in (4.22) and (4.23) are finally projected onto the set of constant modulus vectors, by extracting their phases through, respectively,

$$\mathbf{w}_{k,i} = \frac{1}{\sqrt{N_r}} e^{j\angle \hat{\mathbf{w}}_{k,i}} \quad \text{and} \quad \mathbf{f}_{k,i} = \frac{1}{\sqrt{N_t}} e^{j\angle \hat{\mathbf{f}}_{k,i}}. \quad (4.24)$$

After computing the analog precoder and combiner vector pair, the residual channel tensor is

---

**Algorithm 4.1:** Joint Analog Precoder & Combiner Design for MIMO-OFDM
 

---

```

1 Initialize  $\mathcal{H}_{\text{res}}^{(0)} = \mathcal{H}$  and  $\mathbf{W}_{\text{RF},k} = [ ]$ ,  $\forall k = 1, \dots, K$ 
2 Initialize  $\mathbf{F}_{\text{RF}} = [ ]$  and  $\gamma = 0$ 
3 for  $i = 1 : N_s$  do
4   for  $k = 1 : K$  do
5     Initialize  $\delta^{(0)} = 1$ ,  $\delta^{(1)} = 0$  and  $\eta = 0$ 
6     Randomly Initialize  $\mathbf{w}_{k,i}$  and  $\mathbf{f}_{k,i}$ 
7     while  $|\delta^{(\eta+1)} - \delta^{(\eta)}| \geq \varepsilon$  do
8        $\hat{\mathbf{w}}_{k,i} = \mathbf{H}_{\text{res},k(1)}^{(\gamma)} (\mathbf{I}_M \otimes \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H) \mathbf{H}_{\text{res},k(1)}^{(\gamma)H} \mathbf{w}_{k,i}$ 
9        $\mathbf{w}_{k,i} = \frac{1}{\sqrt{N_r}} e^{j\angle \hat{\mathbf{w}}_{k,i}}$ 
10       $\hat{\mathbf{f}}_{k,i} = \mathbf{H}_{\text{res},k(2)}^{(\gamma)*} (\mathbf{I}_M \otimes \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H) \mathbf{H}_{\text{res},k(2)}^{(\gamma)T} \mathbf{f}_{k,i}$ 
11       $\mathbf{f}_{k,i} = \frac{1}{\sqrt{N_t}} e^{j\angle \hat{\mathbf{f}}_{k,i}}$ 
12       $\eta = \eta + 1$ 
13       $\delta^{(\eta+1)} = \frac{1}{M} \left\| \mathcal{H}_{\text{res},k}^{(\gamma)} \times_1 \mathbf{w}_{k,i}^H \times_2 \mathbf{f}_{k,i}^T \times_3 \mathbf{I}_M \right\|^2$ 
14    end
15    Update  $\mathcal{H}_{\text{res}}$  according to (4.25) and (4.26)
16     $\gamma = \gamma + 1$ 
17     $\mathbf{W}_{\text{RF},k} = [\mathbf{W}_{\text{RF},k} \ \mathbf{w}_{k,i}]$ 
18     $\mathbf{F}_{\text{RF}} = [\mathbf{F}_{\text{RF}} \ \mathbf{f}_{k,i}]$ 
19  end
20 end
21 return  $\mathbf{F}_{\text{RF}}, \mathbf{W}_{\text{RF}}$ 

```

---

updated to remove the contribution of such a pair by setting

$$\mathbf{H}_{\text{res},k(1)}^{(\gamma+1)} = \mathbf{P}_{\mathbf{w}_{k,i}} \mathbf{H}_{\text{res},k(1)}^{(\gamma)} \mathbf{P}_{\mathbf{f}_{k,i}} \quad (4.25)$$

and

$$\mathbf{H}_{\text{res},j(1)}^{(\gamma+1)} = \mathbf{H}_{\text{res},j(1)}^{(\gamma)} \mathbf{P}_{\mathbf{f}_{k,i}}, \forall j \neq k \quad (4.26)$$

where  $\mathbf{P}_{\mathbf{w}_{k,i}} = [\mathbf{I}_{N_r} - \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H]$  and  $\mathbf{P}_{\mathbf{f}_{k,i}} = [\mathbf{I}_{N_t} - \mathbf{f}_{k,i} \mathbf{f}_{k,i}^H]$ . The expression in (4.25) projects the column space of  $\mathbf{H}_{\text{res},k(1)}^{(\gamma)}$  onto the subspace orthogonal to  $\mathbf{w}_{k,i}$ , and its row space onto the subspace orthogonal to  $\mathbf{f}_{k,i}$ . As a result, (4.25) removes the components of the  $k^{\text{th}}$  user's residual channel sub-tensor in the direction of  $\mathbf{w}_{k,i}$  and  $\mathbf{f}_{k,i}$ . On the other hand, in (4.26), the rows space of  $\mathbf{H}_{\text{res},j(1)}^{(\gamma)}$  ( $\forall j \neq k$ ) is projected onto  $\mathbf{f}_{k,i}$ 's orthogonal subspace, which, as a result, removes the components of the aggregated residual channel tensor of the remaining users in the direction of  $\mathbf{f}_{k,i}$ . The updated residual channel tensor is utilized to design the following analog precoder and combiner pair. This procedure is repeated until all the  $KN_s$  analog precoder and combiner vector pairs have been designed. The proposed analog beamforming design is summarized in Algorithm 4.1.

Note that using the single-power-iteration approximations in (4.22) and (4.23) instead of the eigenproblem solutions has no significant impact on the system performance; although, the proposed algorithm may require few more ALS iterations (see Sec. 4.4.4). Besides, the successive rank-(1,1) approximation subproblems in (4.13) can be solved according to distinct user scheduling schemes. In Alg. 4.1, the analog precoder and combiner are designed for one data stream of each user at a time, aiming to provide fairness among users. Alternatively, the analog precoder and combiner can be designed following a user priority schedule, by computing all the analog precoder and combiner pairs for each user at a time. Finally, note that the proposed constrained tensor decomposition is a general approach and can be used for both single-user and multi-user MIMO-OFDM systems. Different approaches for designing the digital beamforming for each of these systems are presented in the following sections.

### 4.3.2 Digital Precoder and Combiner Design for Single-User Systems

Having computed the analog precoder and combiner matrix (common to all subcarriers), the optimal digital precoder and combiner matrices for single-user systems are obtained from the effective baseband channel's SVD on a per-subcarrier basis. The effective baseband channel of the  $m^{\text{th}}$  subcarrier,  $\mathbf{H}_{\text{eff}}[m] = \mathbf{W}_{\text{RF}}^H \mathbf{H}[m] \mathbf{F}_{\text{RF}}$ , has SVD  $\mathbf{H}_{\text{eff}}[m] = \mathbf{U}_m \mathbf{\Sigma}_m \mathbf{V}_m^H$ . The digital precoder and combiner are designed by taking the  $N_s$  column vectors of  $\mathbf{V}_m$  and  $\mathbf{U}_m$  associated with the  $N_s$  largest singular values of  $\mathbf{H}_{\text{eff}}[m]$ .

Next, the digital precoder matrix is normalized, to ensure that  $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[m]\|_{\text{F}}^2 = P_{\text{T}}$ , by letting

$$\mathbf{F}_{\text{BB}}[m] = \sqrt{P_{\text{T}}} \frac{\mathbf{F}_{\text{BB}}[m]}{\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[m]\|_{\text{F}}} \quad (4.27)$$

and allocating the power according to the water-filling method, as described in Sec. 3.3.2.

### 4.3.3 Digital Precoder and Combiner Design for Multi-User Systems

The digital precoder and combiner design for multi-user systems is performed with the RCD method presented in Sec. 3.3.3, which can suppress intra-user and inter-user interference while also supporting noise suppression. The RCD precoder for the  $m$ -th subcarrier is given as

$$\mathbf{F}_{\text{BB}}[m] = \beta \mathbf{F}_{\text{a}}[m] \mathbf{F}_{\text{b}}[m] \quad (4.28)$$

with  $\mathbf{F}_a[m] = [\mathbf{F}_{a_1}[m] \mathbf{F}_{a_2}[m] \cdots \mathbf{F}_{a_K}[m]] \in \mathbb{C}^{KN_s \times KN_s}$  and

$$\mathbf{F}_b[m] = \begin{bmatrix} \mathbf{F}_{b_1}[m] & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{F}_{b_K}[m] \end{bmatrix} \in \mathbb{C}^{KN_s \times KN_s}$$

where  $\beta$  is selected to ensure the total transmit power constraint, i.e.,  $\beta^2 \|\mathbf{F}_{\text{RF}} \mathbf{F}_a[m] \mathbf{F}_b[m]\|_{\text{F}}^2 \leq P_{\text{T}}$ .

The primary precoder  $\mathbf{F}_{a_k}[m] \in \mathbb{C}^{KN_s \times N_s}$ , designed for MUI and noise suppression, is given by [109]

$$\mathbf{F}_{a_k}[m] = \left( \tilde{\mathbf{H}}_k^{\text{H}}[m] \tilde{\mathbf{H}}_k[m] + \frac{KN_s \sigma_n^2}{P_{\text{T}}} \mathbf{I}_{KN_s} \right)^{-1} \quad (4.29)$$

where  $\tilde{\mathbf{H}}_k[m] = [\bar{\mathbf{H}}_1^{\text{T}}[m] \cdots \bar{\mathbf{H}}_{k-1}^{\text{T}}[m] \bar{\mathbf{H}}_{k+1}^{\text{T}}[m] \cdots \bar{\mathbf{H}}_K^{\text{T}}[m]]^{\text{T}}$ ; and the secondary precoder  $\mathbf{F}_{b_k}[m] \in \mathbb{C}^{N_s \times N_s}$  and the digital combiner, designed to suppress intra-user interference and further optimize the system's performance, are given by

$$\mathbf{F}_{b_k}[m] = \mathbf{V}_s \quad (4.30)$$

and

$$\mathbf{W}_{\text{BB},k}[m] = \mathbf{U}_s \quad (4.31)$$

where  $\mathbf{U}_s$  and  $\mathbf{V}_s$  are, respectively, the right and left singular values of the SVD  $\bar{\mathbf{H}}_k[m] \mathbf{F}_{a_k}[m] = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^{\text{H}}$ .

Finally, to satisfy the per-subcarrier total transmitted power constraint  $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[m]\|_{\text{F}}^2 = P_{\text{T}}$ , the digital precoder is normalized by letting

$$\mathbf{F}_{\text{BB}}[m] = \sqrt{\frac{P_{\text{T}}}{\|\mathbf{F}_{\text{RF}} \mathbf{F}_a[m] \mathbf{F}_b[m]\|_{\text{F}}^2}} \mathbf{F}_a[m] \mathbf{F}_b[m] \quad (4.32)$$

and allocating the power through the water-filling algorithm.

#### 4.3.4 Computational complexity

The complexity order of the analog precoder and combiner design using the proposed constrained Tucker2 tensor decomposition, in Alg. 4.1, is approximately  $\mathcal{O}\{K[MN_t N_r N_s(K + N_{\text{ite}}) + N_t^2 N_s N_{\text{ite}}]\}$ , where  $N_{\text{ite}}$  is the maximum number of iterations needed for designing an analog combiner and precoder vector pair. Note that the complexity is mostly associated with the computations in lines 8, 10 and 13 of Alg. 4.1. Nevertheless, the structure of those operations can be explored to implement them in a parallel fashion, thereby reducing the computation time significantly.

**Table 4.1**

Computational Complexity Comparison for Single-User OFDM Hybrid Beamforming Design Methods

Method	Computational complexity order
Prop. Method	$MN_tN_rN_sN_{\text{ite}} + N_t^2N_sN_{\text{ite}}$
PE-AltMin [83]	$M[N_{\text{ite}}N_s^2(N_t + N_r) + N_tN_rN_s]$
HBF-LSAA [117]	$N_s^2N_t^2N_r^2N_{\text{ite}} + M(N_t^2N_r + N_r^2N_s)$
SS-SVD [118]	$MN_t(N_r^2 + N_s^2) + N_s(N_t^2 + N_r^2)$
ICSI-HBF [119]	$MN_s^2(N_t + N_r) + C^2N_s$

**Table 4.2**

Computational Complexity Comparison for Multi-User OFDM Hybrid Beamforming Design Methods

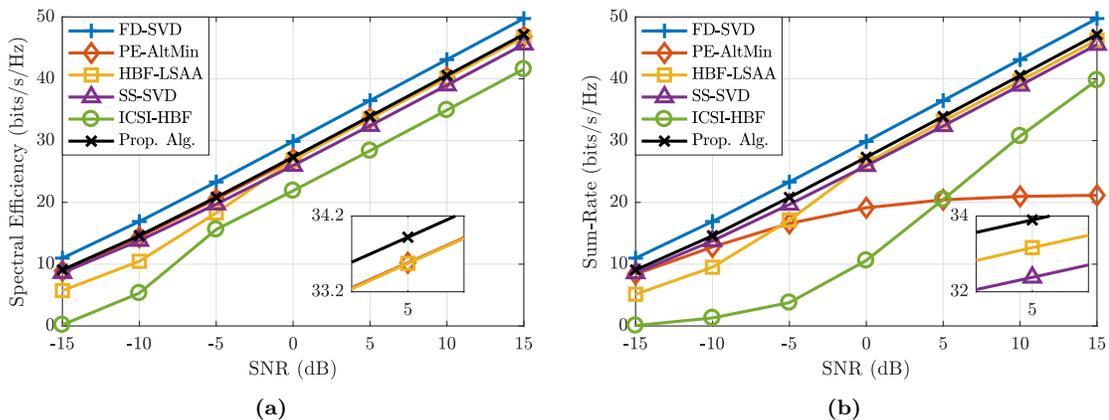
Method	Computational complexity order
Prop. Method	$K[MN_tN_rN_s(N_{\text{ite}} + K^2N_s) + N_t^2N_sN_{\text{ite}}]$
HD-PG-MMSE [122]	$MKN_{\text{ite}}[N_t^3 + N_r^3 + N_tN_r^2]$
TU-HBF [123]	$MK[K^3N_s^3 + N_t^2N_s + N_tN_r^2K]$

For single-user systems (i.e.,  $K = 1$ ), the complexity of Alg. 4.1 reduces to  $\mathcal{O}\{MN_tN_rN_sN_{\text{ite}} + N_t^2N_sN_{\text{ite}}\}$ . The digital beamforming, designed by computing the effective baseband channel, its SVD, and the digital precoder normalization, requires  $\mathcal{O}\{M(N_s^3 + N_sN_tN_r)\}$ . Therefore, considering that  $N_s \leq N_r, N_t$ , the proposed design's overall complexity reduces to  $\mathcal{O}\{MN_tN_rN_sN_{\text{ite}} + N_t^2N_sN_{\text{ite}}\}$ , approximately. For comparison, Table 4.1 shows the order of complexity of the proposed method and that in [83], [117]–[119], with  $C$  representing the codebook's cardinality in [119].

For multi-user systems, the digital beamforming requires computing the effective baseband channel, the RCD solution, and the digital precoder normalization. Such computations require  $\mathcal{O}\{MKN_sN_t[(KN_s)^2 + N_r]\}$ . Assuming  $N_s \leq N_r$  and  $KN_s \leq N_t$ , the proposed design's overall complexity, including the analog beamforming, is approximately  $\mathcal{O}\{K[MN_tN_rN_s(N_{\text{ite}} + K^2N_s) + N_t^2N_sN_{\text{ite}}]\}$ . Table 4.2 shows the complexity order of the proposed algorithm and those in [122] and [123]. It is worth noting that the number of iterations in the proposed algorithm is usually in the order of few tens, while for the HD-PG-MMSE, it is in the order of a few hundreds [122].

## 4.4 Simulation Results

This section provides numerical simulation results to assess the proposed method with comparison to other existing designs. This section is organized as follows: Sections 4.4.1 and 4.4.2 present the simulation results for SU and MU systems, respectively; Section 4.4.3 studies the convergence of the proposed method; Section 4.4.4 evaluates the effects of the single-iteration power method that led



**Figure 4.2:** Example 1. SU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. SNR for a system with  $N_t = N_r = 64$  antennas,  $N_s = 4$  data streams, and  $M = 512$  subcarriers.

to (4.22) and (4.23); Section 4.4.5 compares the narrowband design proposed in Chapter 3 with that proposed in this chapter when  $M = 1$ ; and lastly, Section 4.4.6 presents further discussions on the results.

The examples evaluate both the spectral efficiency and the sum-rate performances of the different hybrid beamforming designs obtained by averaging the results over 500 channel realizations. Optimal water-filling power allocation is considered for all methods. The SNR is defined as  $\text{SNR} = P_T \rho_k / \sigma_n$ , where the per-subcarrier total transmitted power is set to  $P_T = K N_s$ , with  $\rho_k$  assumed equal for all users. Only the mmWave scenarios are considered here. The mmWave channels are modeled by the extended Saleh-Valenzuela channel model with  $N_{cl} = 5$  scattering clusters and  $N_{ray} = 10$  propagation paths per cluster, angular spread  $\sigma_\phi^t = \sigma_\theta^t = \sigma_\phi^r = \sigma_\theta^r = 10^\circ$ , and a USPA with antenna spacing  $d = \lambda/2$  [24], [83]. The stopping criterion in Alg. 4.1 is set to  $\varepsilon = 0.001$  and the maximum number of iterations  $N_{ite} = 20$ .

#### 4.4.1 Single-User Systems

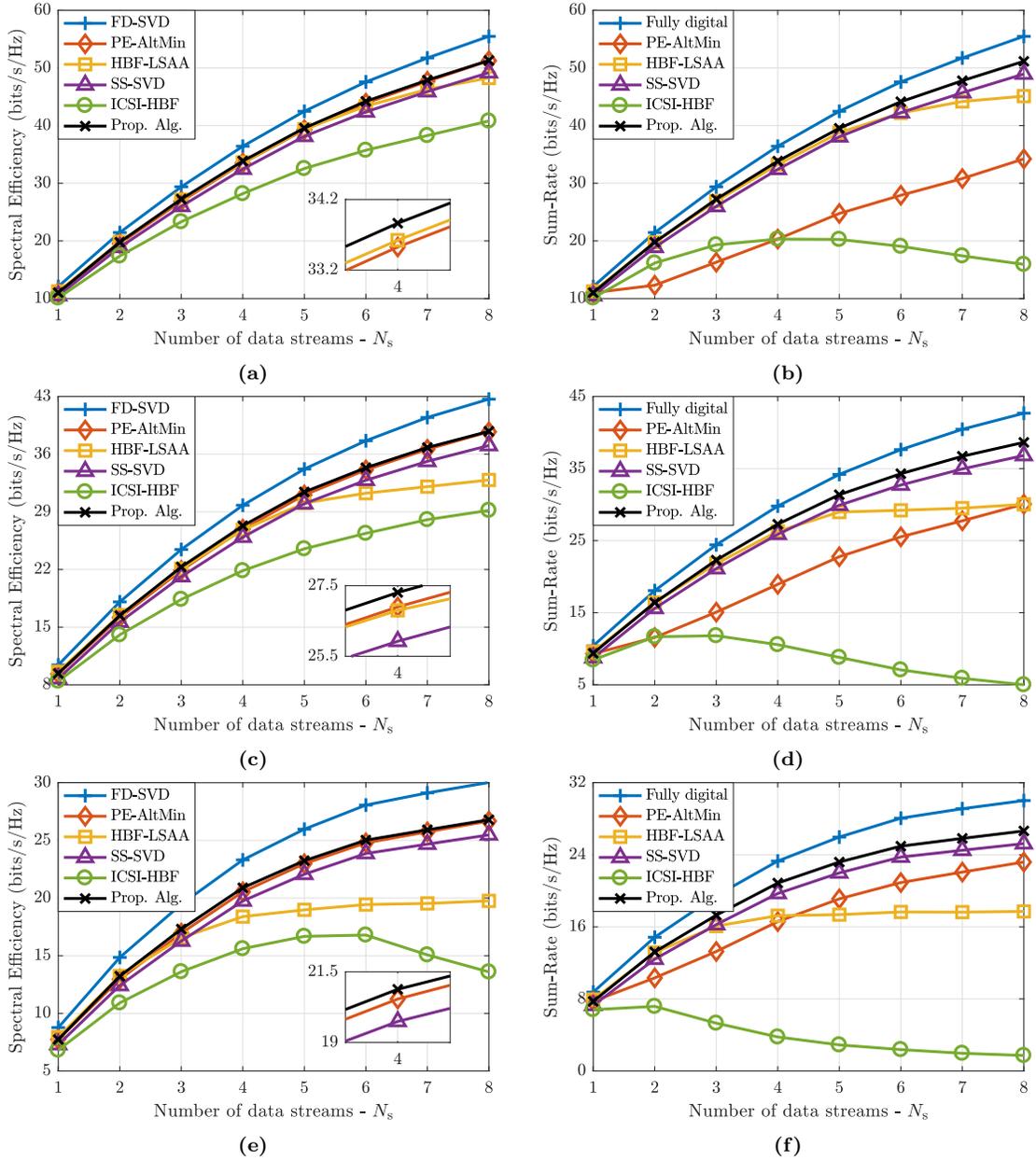
For single-user systems, the proposed design is compared with the fully-digital SVD-based beamforming (FD-SVD) and the hybrid beamforming designs from the PE-AltMin [83], HBF-LSAA [117], SS-SVD [118], and ICSI-HBF [119] (considering an orthogonal codebook with  $C = 64$ ). All examples consider an OFDM system with  $M = 512$  subcarriers.

The first example evaluates the spectral efficiency and the sum-rate for different SNR values for a  $64 \times 64$  SU-MIMO-OFDM system communicating through  $N_s = 4$  data streams. The results are shown in Fig. 4.2a and Fig. 4.2b, respectively. From these results, one observes that the proposed design outperforms all the other hybrid beamforming designs and has a consistent near-optimal

performance. Note that the PE-AltMin, which has the second-best spectral efficiency performance, has its sum-rate severely degraded at high SNR. For instance, the spectral efficiency achieved by the PE-AltMin design is between 0.17 bits/s/Hz and 0.34 bits/s/Hz lower than the proposed design, whereas its sum-rate reaches up to 26 bits/s/Hz lower. The HBF-LSAA design achieves near-optimal performance at high SNR, with spectral efficiency and sum-rate, respectively, around 0.4 bits/s/Hz and 0.8 bits/s/Hz lower than the proposed design; on the other hand, it has poor performance at low SNR with up to, respectively 4.1 bits/s/Hz and 5.1 bits/s/Hz lower spectral efficiency and sum-rate as compared to the proposed design. Unlike the PE-AltMin and the HBF-LSAA, which have their performances affected at either high or low SNR, the SS-SVD attains a consistent performance, but with spectral efficiency and sum-rate up to 1.6 bits/s/Hz lower than the proposed design. Finally, note that the ICSI-HBF has one of the worst performances (only outperforming the PE-AltMin' sum-rate at high SNR), likely due to the use of limited codebooks.

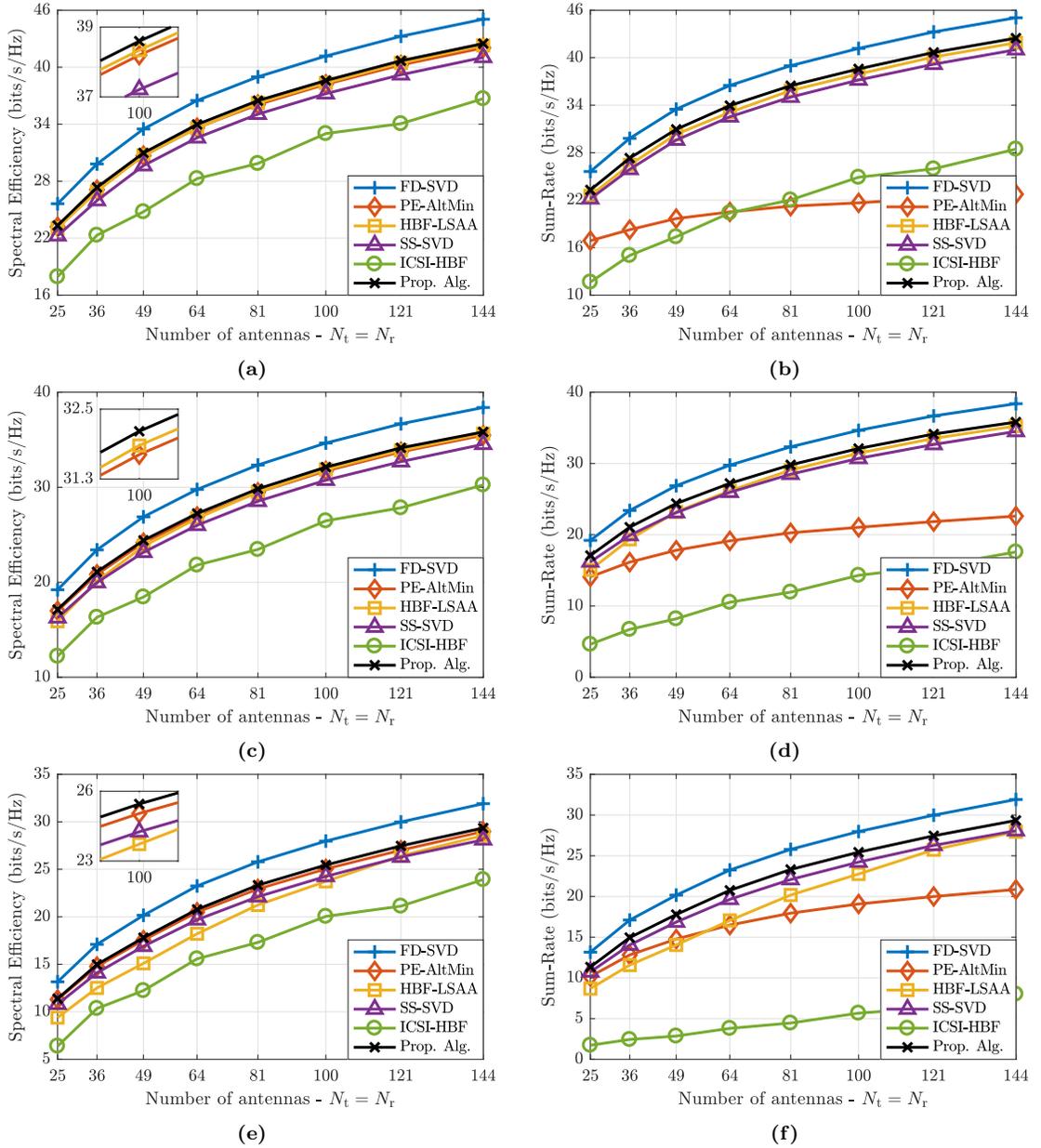
The second example evaluates the impact of the number of data streams on the spectral efficiency and sum-rate of a  $64 \times 64$  SU-MIMO-OFDM system. The results, depicted in Fig. 4.3, show the spectral efficiency (left) and the sum-rate (right) at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). Except at  $N_s = 1$ , where the PE-AltMin and HBF-LSAA designs slightly outperform the proposed design (with up to 0.2 bits/s/Hz higher spectral efficiency and sum-rate), the proposed method has outperformed other designs in all the other scenarios. Note that the PE-AltMin and the ICSI-HBF achieve sum-rates much lower than their respective spectral efficiencies. In particular, the PE-AltMin has the second-best spectral efficiency performance (between 0.2 bits/s/Hz and 0.4 bits/s/Hz lower than the proposed method); but achieve sum-rates up to 17 bits/s/Hz, 9 bits/s/Hz and 4 bits/s/Hz lower than the proposed design, respectively, at SNR = 5 dB, 0 dB, and -5 dB. Moreover, observe that the HBF-LSAA has its performance degraded as the number of data streams increases, particularly at lower SNR, achieving spectral efficiency and sum-rate up to 7.0 bits/s/Hz and 8.9 bits/s/Hz, respectively, lower than the proposed design at SNR = -5 dB. The SS-SVD, as in the previous example, attains a consistent performance in all scenarios; however, it achieves spectral efficiency and sum-rate up to, respectively, 2.1 bits/s/Hz and 2.2 bits/s/Hz lower than the proposed design.

The third example assesses the spectral efficiency and the sum-rate for different numbers of antennas, assuming  $N_t = N_r$ , for a system communicating through  $N_s = 4$  data streams in each subcarrier. The results are shown in Fig. 4.4, which depicts the spectral efficiency (left) and the sum-rate (right) at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom). In this example, the proposed design outperforms the other hybrid beamforming designs in all scenarios. In general, this example



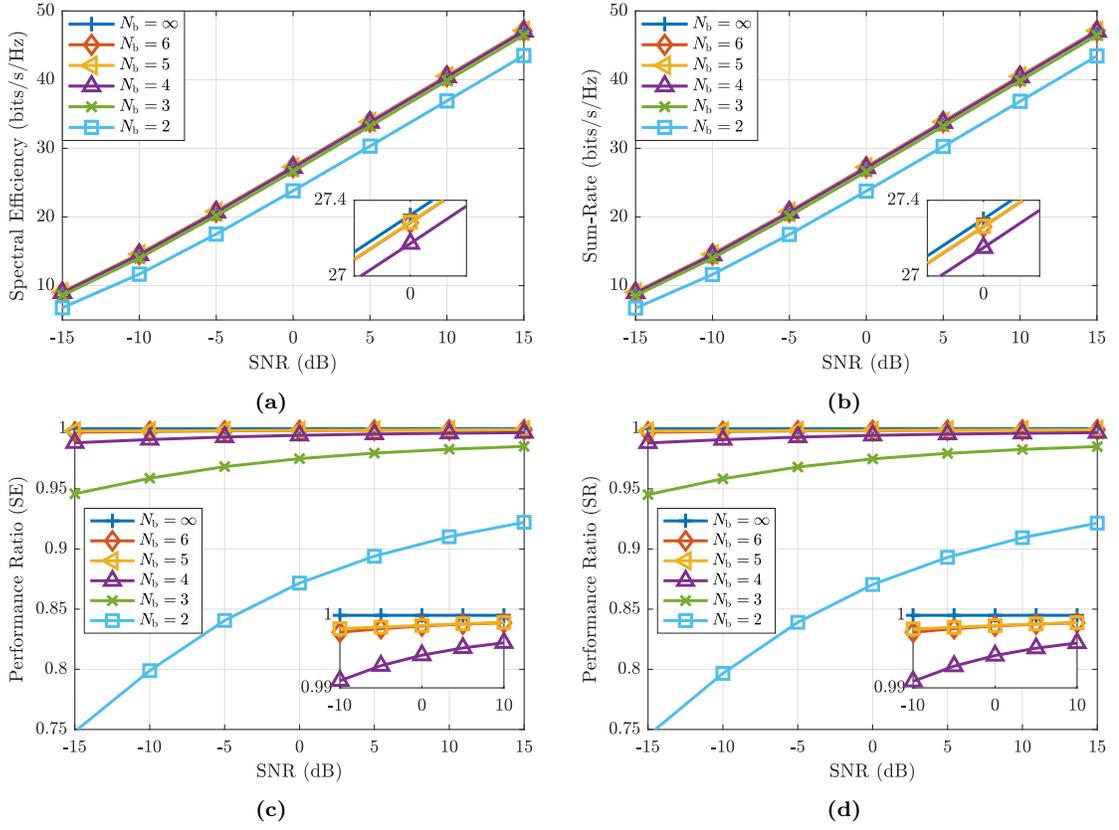
**Figure 4.3:** Example 2. SU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = N_r = 64$  antennas and  $M = 512$  subcarriers at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

confirms the previous results by showing that the PE-AltMin achieves high spectral efficiency but has its sum-rate significantly compromised, especially at high SNR; the HBF-LSAA achieves both near-optimal spectral efficiency and sum-rate at high SNR but has its performance degraded at low SNR, particularly when employing a small number of antennas; and that the SS-SVD attains a consistent near-optimal performance in all scenarios, although still lower than the proposed method.



**Figure 4.4:** Example 3. SU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit and receive antennas  $N_t = N_r$  for a system communicating through  $N_s = 4$  data streams and  $M = 512$  subcarriers at SNR = 5 dB (top), 0 dB (center), and -5 dB (bottom).

Lastly, the impact of the PS quantization on the proposed design is evaluated. As discussed in Chapter 3, practical deployments require using quantized PS to reduce the cost, energy consumption, and hardware complexity [110]–[112], commonly adopting PS with 4 to 6 quantization bits [113]. The phase-quantization operation are performed in lines 8 and 10 of Alg. 4.1 by selecting the nearest quantized phase. This example considers the same scenario as the first example. The

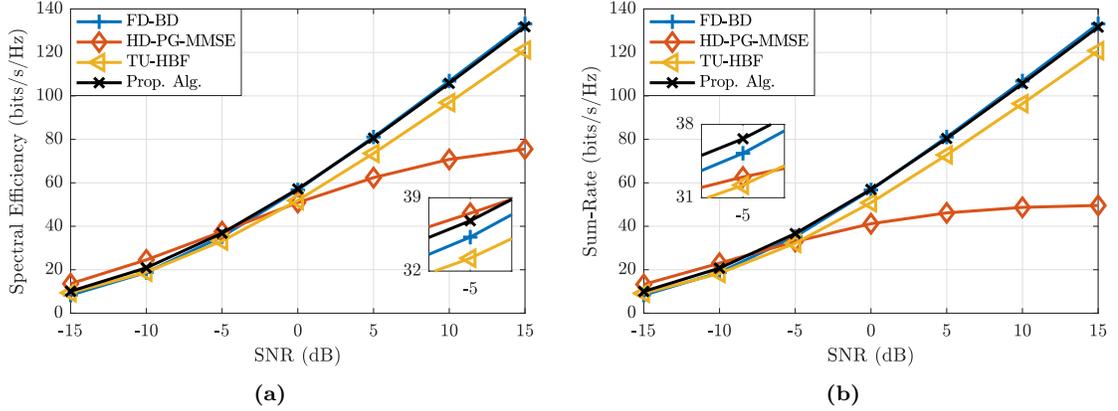


**Figure 4.5:** Example 4. SU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$  and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).

results in Fig. 4.5 show the spectral efficiency (left) and the sum-rate (right) at the top, and the performance ratio between the quantized and unquantized analog beamforming, respectively, in terms of the spectral efficiency and the sum-rate at the bottom, for different numbers of quantization bits  $N_b$ . Note that the proposed design can attain more than 94.5% and 98.8% of its unquantized-PS performance when employing PS with 3 and 4 quantization bits, respectively. For PS with more than 5 quantization bits, the proposed design attains at least 99.7% of its unquantized-PS performance.

#### 4.4.2 Multi-User Systems

For multi-user systems, the proposed hybrid beamforming design is compared to the fully-digital block diagonalization (FD-BD) [98] and the hybrid beamforming designs from HD-PG-MMSE [122] and TU-HBF [123]. However, note that the FD-BD is also sub-optimal and therefore, in certain situations, the hybrid beamforming designs may exceed its performance. Moreover, all examples consider an OFDM system with  $M = 32$  subcarriers. Although practical OFDM deployments often



**Figure 4.6:** Example 1. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. SNR for a system  $N_t = 144$  transmit antennas and  $K = 4$  users, each having  $N_r = 16$  receive antennas, communicating through  $N_s = 4$  data streams and  $M = 32$  subcarriers.

use much higher numbers of subcarriers (e.g., 512 or more), the analysis here is limited  $M = 32$  due to the computational time of the TU-HBF, which requires the SVD computation of very large matrices. The proposed design, however, readily extends to those cases.

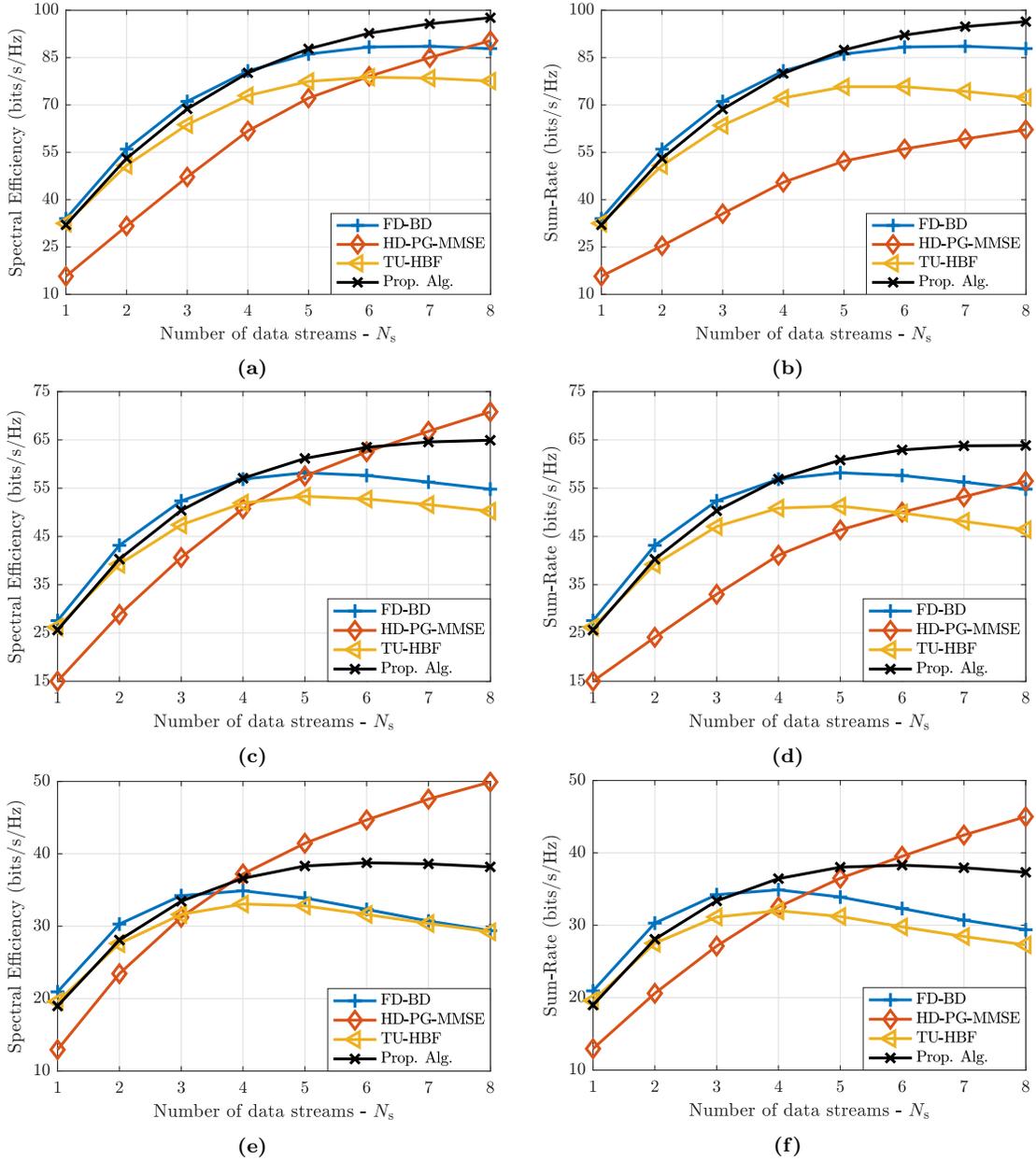
The first example evaluates the spectral efficiency versus the SNR for an MU-MIMO-OFDM system with  $N_t = 144$  transmit antennas and  $K = 4$  users with  $N_r = 16$  receive antennas, each communicating through  $N_s = 4$  data streams. The results are depicted in Fig. 4.6, which shows the spectral efficiency (a) and sum-rate (b) attained by all designs at SNR ranging from  $-15$  dB to  $15$  dB. From Fig. 4.6, one can observe that the proposed design attains very high performance in high SNR scenarios, reaching up to  $10.7$  bits/s/Hz higher spectral efficiency and  $10.8$  bits/s/Hz higher sum-rate than the second-best performing design, the TU-HBF, and up to  $53.3$  bits/s/Hz higher spectral efficiency and  $82.1$  bits/s/Hz higher sum-rate than the HD-PG-MMSE at SNR =  $15$  dB. However, at low SNR (e.g. SNR  $\leq -5$  dB), the proposed design is slightly outperformed by the HD-PG-MMSE, which achieves spectral efficiency and sum-rate up to, respectively,  $3.6$  bits/s/Hz and  $3.2$  bits/s/Hz higher than the proposed design. Nevertheless, despite the slightly better performance of the HD-PG-MMSE design in low-SNR scenarios, its performance is severely degraded in high SNR.

The second example assesses the spectral efficiency and the sum-rate versus the number of data streams  $N_s$  for an MU-MIMO-OFDM system with  $N_t = 144$ ,  $N_r = 16$ , and  $K = 4$ . Fig. 4.7 presents the spectral efficiency (left) and the sum-rate (right) attained by the different design at SNR =  $5$  dB (top),  $0$  dB (center), and  $-5$  dB (bottom). These results show that the proposed design outperforms the TU-HBF design whenever  $N_s \geq 2$ . When  $N_s = 1$ , both the spectral efficiency and the sum-rate attained by the TU-HBF are  $0.5$  bits/s/Hz ( $0.61$  bits/s/Hz) higher than those attained

by the proposed design at SNR = 5 dB (SNR = -5 dB). On the other hand, when  $N_s = 8$ , the proposed design attain spectral efficiency and sum-rate, respectively, 9 bits/s/Hz and 10 bits/s/Hz higher at SNR = -5 dB, and respectively, 20 bits/s/Hz and 24 bits/s/Hz higher at SNR = 5 dB. Moreover, note that the TU-HBF seems more affected by the increase in  $N_s$ , having its performance degraded once  $N_s$  grows past certain values. Compared to the HD-PG-MMSE design, the proposed method has a better performance at higher SNR or lower number of data streams. For instance, at SNR = 5 dB, the proposed method achieves 21.5 bits/s/Hz higher spectral efficiency when  $N_s = 3$  and 36 bits/s/Hz higher sum-rate when  $N_s = 6$ . On the other hand, at low SNR (e.g. SNR = -5 dB), the HD-PG-MMSE reaches up to 11.7 bits/s/Hz and 7.7 bits/s/Hz higher spectral efficiency and sum-rate, respectively. It is essential to highlight that the sum-rate attained by the HD-PG-MMSE design is, in general, much lower than the spectral efficiency. This indicates that linear detection schemes would not be able to explore the full potential of the HD-PG-MMSE design. Moreover, as further discussed in Sec. 4.4.6, the HD-PG-MMSE design has computational complexity much higher than the other methods.

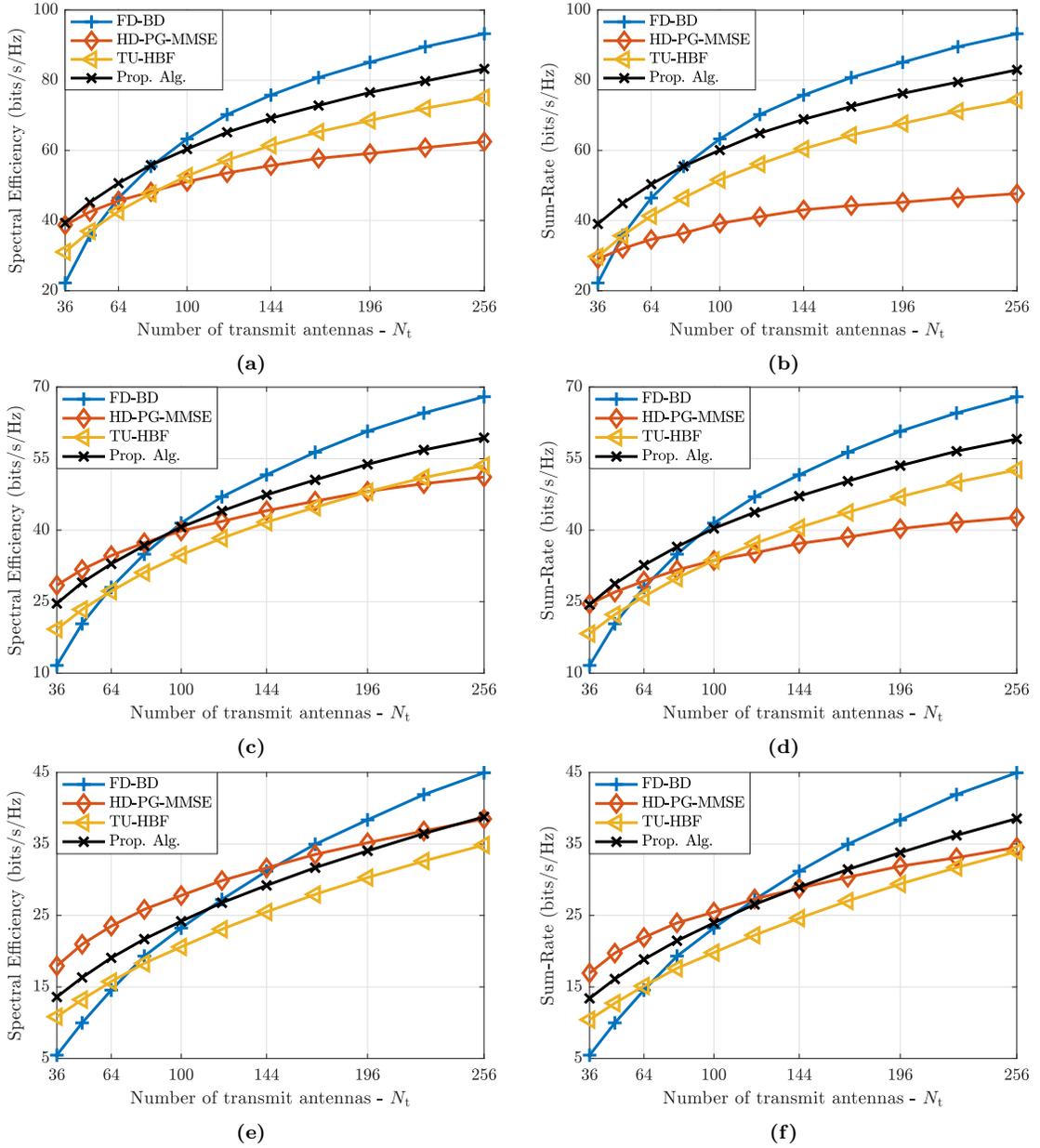
The third example evaluates the spectral efficiency and sum-rate for different numbers of transmit antennas  $N_t$ . An MU-MIMO-OFDM system with  $K = 4$  users, each having  $N_r = 9$  receive antennas and communicating through  $N_s = 4$  data streams is considered. The results are presented in Fig. 4.8, following the same organization as in the previous example. The proposed design outperforms the TU-HBF in all scenarios, attaining from 2.8 bits/s/Hz to 9.9 bits/s/Hz higher spectral efficiency and sum-rate. It also attains spectral efficiency and sum-rate up to, respectively, 20.7 bits/s/Hz and 35.3 bits/s/Hz higher than the HD-PG-MMSE at SNR = 5 dB. At lower SNR, the HD-PG-MMSE can achieve up to 4.6 bits/s/Hz and 3.6 bits/s/Hz higher spectral efficiency and sum-rate, respectively, as compared to the proposed design. However, note that the proposed design tends to outperform the HD-PG-MMSE as the number of transmit antennas grows.

Next, the fourth example investigates the spectral efficiency and the sum-rate performances under different numbers of users  $K$ . An MU-MIMO-OFDM system with  $N_t = 144$  transmit antennas, users having  $N_r = 16$  receive antennas, and communicating through  $N_s = 4$  data streams is considered. The results, depicted in Fig. 4.9, show the spectral efficiency and the sum-rate, following the organization as in previous examples. Once again, these results show that the proposed design greatly outperforms the other hybrid beamforming designs at high SNR. For instance, the proposed design attains a sum-rate up to 49.2 bits/s/Hz and 22.5 bits/s/Hz higher than the HD-PG-MMSE and the TU-HBF, respectively, at SNR = 5 dB. At SNR = -5 dB, the proposed design continues



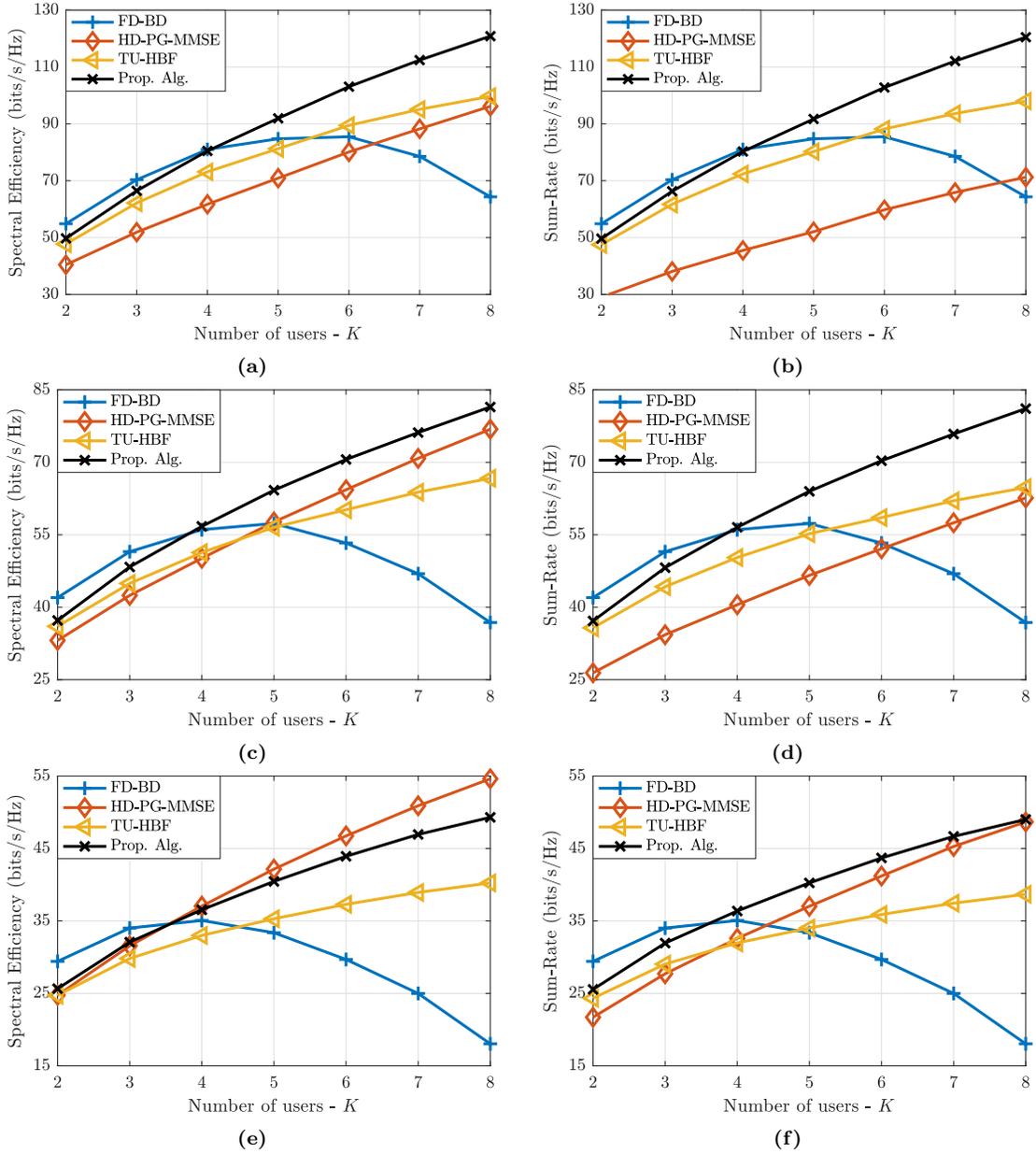
**Figure 4.7:** Example 2. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of data streams  $N_s$  for a system with  $N_t = 144$  transmit antennas and  $K = 4$  users, each having  $N_r = 16$  receive antennas, communicating through  $M = 32$  subcarriers at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom).

to outperform the TU-HBF, attaining, respectively, 9.1 bits/s/Hz and 10.3 bits/s/Hz higher spectral efficiency and sum-rate when  $K = 8$ ; on the other hand, the HD-PG-MMSE achieves spectral efficiency up to 5.3 bits/s/Hz higher than the proposed design, but fails to provide higher sum-rate.



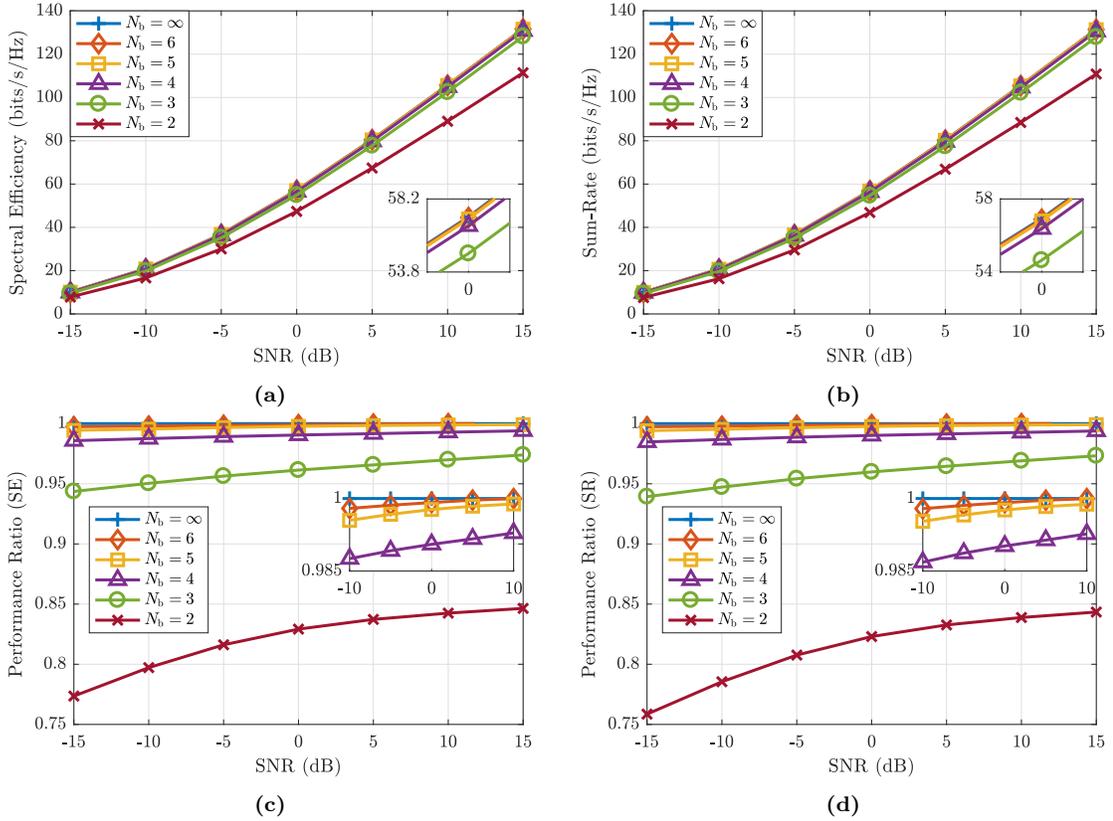
**Figure 4.8:** Example 3. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of transmit antennas  $N_t$  for a system with  $K = 4$  users having  $N_r = 9$  receive antennas and communicating through  $N_s = 4$  data streams and  $M = 32$  subcarriers at SNR of 5 dB (top), 0 dB (center), and -5 dB (bottom).

Lastly, the fifth example investigates the proposed method's performance under different phase-shifter quantization resolutions by considering the scenario of the first example. Fig 4.10 presents the spectral efficiency (left) and the sum-rate (right) at the top, and the performance ratio between the quantized and unquantized analog beamforming, respectively, in terms of the spectral efficiency



**Figure 4.9:** Example 4. MU-MIMO-OFDM system: Spectral efficiency (left) and sum-rate (right) vs. the number of users  $K$  for a systems with  $N_t = 144$  transmit antennas and users having  $N_r = 16$  receive antennas, communicating through  $N_s = 4$  data streams and  $M = 32$  subcarriers at SNR of 5 dB (top), 0 dB (center), and  $-5$  dB (bottom).

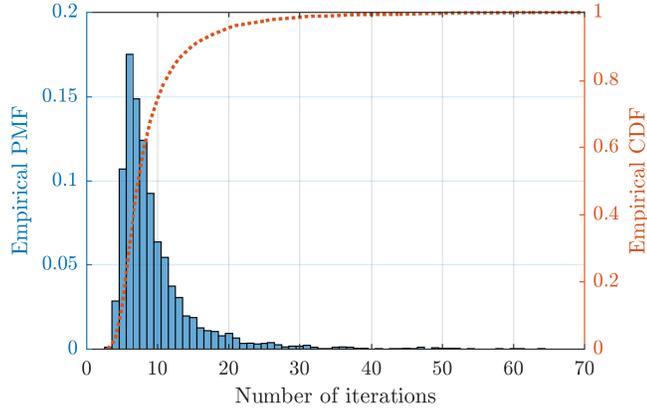
and the sum-rate at the bottom, for different numbers of quantization bits  $N_b$ . These results show that the proposed design can attain more than 99.4% and 99.7% of its unquantized-PS performance when employing PS with 5 and 6 quantization bits, respectively. In particular, when  $N_b = 6$ , the proposed design can even attain its full unquantized-PS performance at  $\text{SNR} \geq 10$  dB.



**Figure 4.10:** Example 5. MU-MIMO-OFDM system: Spectral efficiency (a) and sum-rate (b) vs. the SNR for different numbers of phase-shifter quantization bits  $N_b$  and the ratio between the quantized and unquantized PS performances in terms of spectral efficiency (c) and sum-rate (d).

#### 4.4.3 Convergence Analysis

This section presents additional simulation results that provide important insights into the convergence of Algorithm 4.1. First, it is important to highlight that, as briefly discussed in Appendix C.2.2, the ALS-based solution to the general Tucker decomposition problem is not unique nor guaranteed to converge to a global optimum [135]. This also holds for the proposed constrained Tucker2 decomposition in Alg. 4.1. In fact, for the proposed algorithm, the non-uniqueness of the solution is quite evident from the fact that any rotation in the analog precoder or the analog combiner vectors will not affect the objective function in (4.14). On the other hand, an analytical proof for the convergence is very difficult - if not impossible - due to the non-convex constant modulus constraint and the nonexistence of a derivative for the function that maps the solutions in (4.22) and (4.23) onto the constant modulus solution space. Nevertheless, the iterative solution of the equations in (4.24) leads to an increasingly better value for the objective function in (4.14), allowing the convergence of Alg. 4.1 to be determined once the objective function ceases to increase. The following examples



**Figure 4.11:** Empirical PMF and empirical CDF of the number of iterations needed for computing each analog beamforming vector pair, considering a SU-MIMO-OFDM system with  $N_t = N_r = 64$ ,  $N_s = 8$ , and  $M = 512$ .

show that Alg. 4.1 does converge within a reasonable number of iterations.

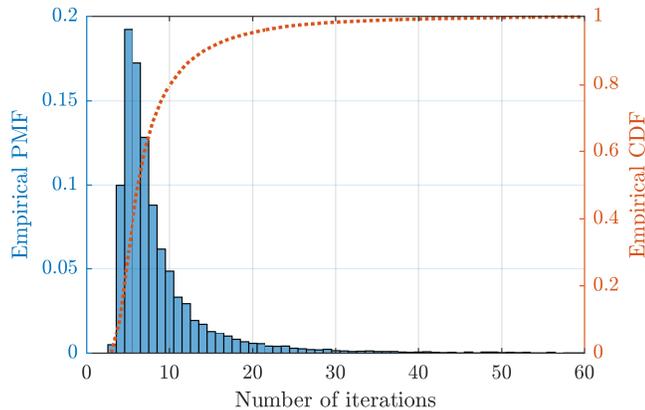
The first example considers a  $64 \times 64$  SU-MIMO-OFDM system communicating through  $M = 512$  subcarriers and  $N_s = 8$  data streams per subcarrier. Fig. 4.11 shows the empirical PMF and the empirical CDF of the number of iterations the inner-loop of Alg. 4.1 (i.e., lines 7-14) requires for convergence. These results show that the proposed algorithm requires, on average, 9.5 iterations for convergence, and converges within 15 and 35 iterations, respectively, in 90% and 99% of the trials.

The second example considers an MU-MIMO-OFDM system with  $N_t = 144$  transmit antennas and  $K = 8$  users, with each user having  $N_r = 16$  receive antennas and communicating through  $M = 32$  subcarriers and  $N_s = 4$  data streams per subcarrier. The empirical PMF and CDF of the number of iterations required for convergence are shown in Fig. 4.12. The proposed algorithm requires 8.7 iterations for convergence on average and converges within 15 and 38 iterations, respectively, in 90% and 99% of the trials.

Additional simulation results have shown the results in Fig. 4.11 and Fig. 4.12 are not significantly affected by the system’s parameters (e.g.,  $N_t$ ,  $N_r$ ,  $N_s$ , and  $K$ ). Moreover, the stopping criteria  $\varepsilon$  and  $N_{ite}$ , in Alg. 4.1, allows the designer to trade off computational complexity and convergence speed for better performance.

#### 4.4.4 Effects of Power Method Approximation

Recall that the solution to the relaxed ALS problems in (4.17) and (4.18) result in two eigenproblems, and to avoid such a complex computation, these solutions were approximated by a single iteration of the power method, leading to the expressions in (4.22) and (4.23). This example investigates the



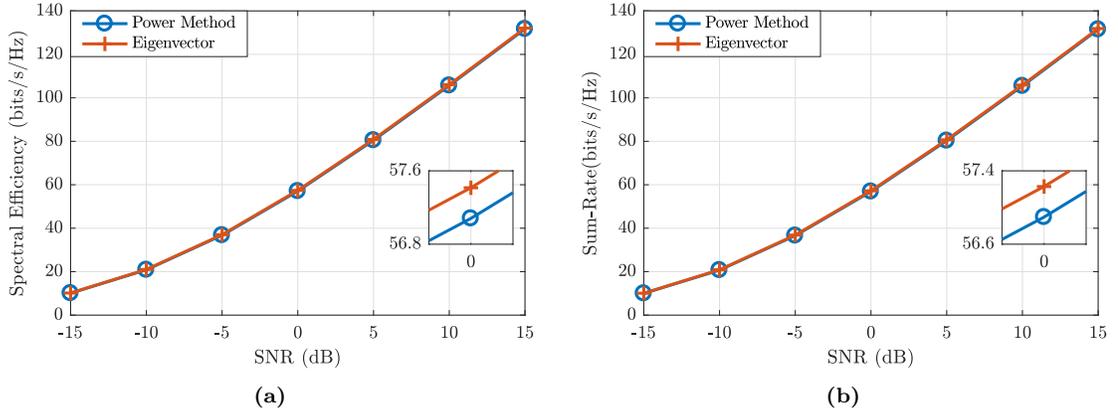
**Figure 4.12:** Empirical PMF and empirical CDF of the number of iterations needed for computing each analog beamforming vector pair, considering an MU-MIMO-OFDM system with  $N_t = 144$ ,  $N_r = 16$ ,  $K = 8$ , and  $N_s = 4$ , and  $M = 32$ .

effects of such approximation on the proposed algorithm’s performance and convergence.

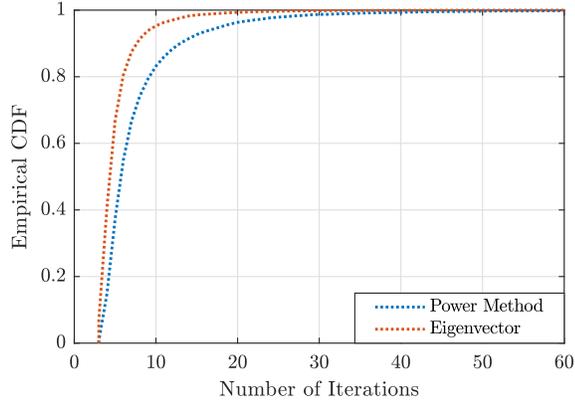
The example considers an MU-MIMO-OFDM system with  $N_t = 144$  transmit antennas and  $K = 4$  users, each having  $N_r = 16$  receive antennas, communicating through  $N_s = 4$  data streams in  $M = 32$  subcarriers. The results are shown in Fig. 4.13 and Fig. 4.14. Fig. 4.13 presents the spectral efficiency (left) and the sum-rate (right) attained by the proposed design when the principal eigenvector and the power method approximation are used. These results show that the proposed design can attain up to 0.4 bits/s/Hz higher spectral efficiency and sum-rate when the eigenproblem solution is used. Moreover, Fig. 4.14 shows the empirical CDF of the number of iteration required for convergence for both methods. The proposed design requires 6 iterations for convergence, on average, and converges within 8 iterations in more than 90% of the time when using the eigenproblem solution. On the other hand, when using the power method approximation, it requires, on average, 8 iterations, and converges within 13 iterations in 90% of the time.

Despite better performance and faster convergence when using the eigenproblem solution, note that the proposed single-iteration power method approximation can significantly reduce computational complexity. The single-iteration power method requires a matrix-vector multiplication, with complexity order  $\mathcal{O}\{N^2\}$ . In contrast, the complexity of the eigenproblem solution is bounded by  $\mathcal{O}\{N^3\}$  [136]. Alternatively, if the power-iteration method is used for computing the principal eigenvector (instead of solving the complete eigenproblem), this would still require multiple power iterations<sup>2</sup>.

<sup>2</sup>The power-iteration method requires approximately  $\mathcal{O}\{\frac{1}{\Delta} \log \frac{1}{\delta}\}$  iterations for convergence, where  $\Delta = \lambda_1 - \lambda_2$  is the gap between the largest and the second-largest eigenvalues, and  $\delta$  is the accuracy of the approximation measured by the squared sine or the angle between the true eigenvector and its approximation [137].



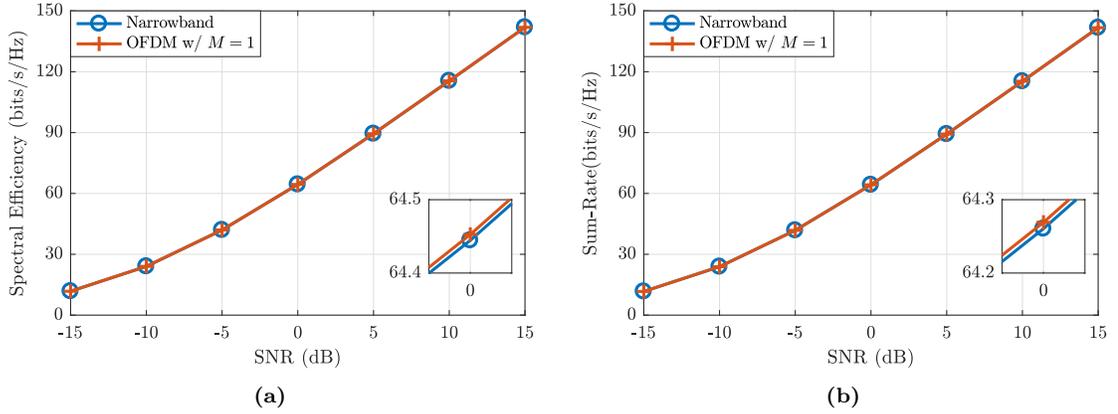
**Figure 4.13:** Effects of the power method approximation. Spectral efficiency (a) and sum-rate (b) vs. the SNR for an MU-MIMO-OFDM system with  $N_t = 144$  transmit antennas and  $K = 4$  users, each having  $N_r = 16$  receive antennas, communicating through  $N_s = 4$  data streams in  $M = 32$  subcarriers.



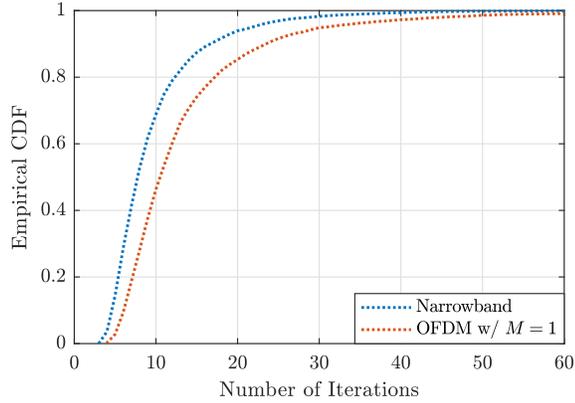
**Figure 4.14:** Effects of the power method approximation. Empirical CDF of the number of iterations required for computing each analog beamforming vector pair, considering an MU-MIMO system with  $N_t = 144$ ,  $N_r = 16$ ,  $K = 4$ , and  $N_s = 4$ .

#### 4.4.5 Comparison with the Narrowband Design in Chapter 3

This subsection compares the hybrid beamforming design for narrowband systems presented in Chapter 3 with the design proposed in this chapter when considering a narrowband system (i.e., when  $M = 1$ ). Despite the similarities between both designs, it is important to highlight that narrowband design in Chapter 3 is not a particular case of the design in this chapter. For instance, note that the objective function in (4.14) (from which originates Alg. 4.1) reduces to  $\|\mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i}\|^2$  when  $M = 1$ , while the objective function that originates the narrowband design in Chapter 3 is given by  $\text{Re}(\mathbf{w}_{k,i}^H \mathbf{H}_{\text{res},k}^{(i)} \mathbf{f}_{k,i})$ . Therefore, observe that any analog beamforming solution in the narrowband design is also a solution to the OFDM design (in fact, any rotation of such a solution, either in the analog precoder vector or the analog combiner vector, is also a solution to the OFDM



**Figure 4.15:** Comparison with the narrowband design in Chapter 3. Spectral efficiency (a) and sum-rate (b) vs. the SNR for an MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 4$  users, with each user having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams.



**Figure 4.16:** Comparison with the narrowband design in Chapter 3. Empirical CDF of the number of iterations required for computing each analog beamforming vector pair, considering an MU-MIMO system with  $N_t = 144$ ,  $N_r = 16$ ,  $K = 4$ , and  $N_s = 4$ .

design), but the reciprocal does not hold. Moreover, note that the OFDM design's computational complexity reduces to  $\mathcal{O}\{K[N_t N_r N_s (N_{\text{ite}} + K^2 N_s) + N_t^2 N_s N_{\text{ite}}]\}$  when  $M = 1$ . On the other hand, the narrowband design has complexity  $\mathcal{O}\{K[N_t N_r N_s (N_{\text{ite}} + K) + K^3 N_s^3]\}$ . Thus, since  $K^2 N_s \gg K$  and  $N_t^2 N_s N_{\text{ite}} \gg K^3 N_s^3$ , one can quickly note that the narrowband design in Chapter 3 has lower computational complexity (from one to two orders of magnitude lower when considering typical system parameters, such as those considered in the examples of this section).

An additional example is provided to obtain further insights on the differences between the narrowband and the OFDM designs. This example considers a narrowband MU-MIMO system with  $N_t = 144$  transmit antennas and  $K = 4$  users, with each user having  $N_r = 16$  receive antennas and communicating through  $N_s = 4$  data streams. Fig. 4.15 shows the spectral efficiency (left) and the sum-rate (right) attained for the narrowband design in Chapter 3 and this chapter's OFDM design

with  $M = 1$ . In general, the OFDM design attains a marginally better performance (with differences only in the second decimal place). Furthermore, Fig. 4.16 shows the empirical CDF of the number of iterations required by each algorithm. The narrowband design requires 10 iterations, on average, to reach the convergence, while the OFDM design requires, on average, 14 iterations. Besides, the narrowband and the OFDM design converge within, respectively, 16 and 23 iterations in 90% of the trials, and 35 and 57 iterations in 99% of the trials.

#### 4.4.6 Further Discussions

The results in Sec. 4.4.1 and Sec. 4.4.2 have demonstrated the proposed design's performance, respectively, for single-user and multi-user systems, under different scenarios.

The results obtained for single-user systems have shown that the proposed method consistently attained near-optimal spectral efficiency and sum-rate and outperformed the other hybrid beamforming designs in nearly all scenarios. In contrast, the HBF-LSAA has attained near-optimal performance only in favorable scenarios (such as in high SNR, with a large number of antennas, and a small number of data streams), failing to perform well under more critical scenarios; the PE-AltMin has attained near-optimal spectral efficiency in most scenarios, but has had its sum-rate negatively impacted, especially at high SNR; the SS-SVD has also attained a consistently good performance, although always below the proposed design's performance; and finally, the ICSI-HBF has significantly under-performed the other designs, likely due to the limitations imposed by the reduced codebook. It is also worth noting that the proposed design's computational complexity is similar to other existing methods. For the parameters of the first example in Sec. 4.4.1, the proposed method has complexity in the order of  $10^8$ . The ICSI-HBF, PE-AltMin, SS-SVD and the HBF-LSAA have complexity in the order of  $10^6$ ,  $10^7$ ,  $10^8$ , and  $10^9$ , respectively. Despite the lower complexity of the ICSI-HBF, it has significantly lower performance than the other methods.

For the multi-user case, the results in Sec. 4.4.2 have shown that the proposed design can also consistently attain high spectral efficiency and sum-rate. The proposed design has outperformed the TU-HBF in nearly all scenarios, performing particularly better at more critical conditions, such as when using a small number of antennas or transmitting many data streams. The HD-PG-MMSE has outperformed the proposed design in those critical scenarios, notably at low SNR; however, the HD-PG-MMSE has a much inferior performance at high SNR. Furthermore, the HD-PG-MMSE has a much higher computational complexity. For instance, considering an MU-MIMO-OFDM system with the same parameters as in the first example in Sec. 4.4.2, the proposed method has complexity in the order of  $10^8$ , while the TU-HBF and the HD-PG-MMSE have complexities in the order

of  $10^7$  and  $10^{10}$ , respectively. Moreover, the results have demonstrated that the proposed design can also perform well even when practical PS quantization is considered. For PS with 4 and 6 quantization bits, the performance loss remained under 1.4% and 0.25%, respectively, compared to the unquantized design.

Additional examples have provided further insights into the proposed design. In Sec. 4.4.3, examples have investigated the convergence of Alg. 4.1. The results showed that despite the non-guaranteed convergence to a globally optimal solution, the algorithm achieves a local optimal within a reasonably low number of iterations. Sec. 4.4.4 has shown that the proposed single-iteration power method simplification adopted to avoid solving the eigenproblem has, in fact, almost no effect on the design performance. While the simplification can demand a few more iterations, it significantly reduces the algorithm complexity. Finally, Sec. 4.4.5 has pointed out the differences between the narrowband design proposed in Chapter 3 and the design in this chapter when assuming a single subcarrier. Despite achieving nearly the same performance, the discussions in Sec. 4.4.5 revealed that the analog beamforming of both designs have different solution spaces (with the narrowband design's solution space being only a subset of the design in this chapter). Furthermore, it is shown that the narrowband design has lower computational complexity and converges within fewer iterations.

## 4.5 Concluding Remarks

This chapter proposed a novel hybrid beamforming design for selective-frequency mmWave massive MIMO-OFDM systems. The proposed method performs the analog and digital beamforming design in two separate stages. In the first stage, the analog precoder and combiner design are formulated as a constrained Tucker2 tensor decomposition problem, which can increase the effective baseband channel gain while also suppressing the MUI simultaneously over every subcarrier. In the second stage, the digital beamformers are computed on a per subcarrier basis, using different methods for single- and multi-user systems. For SU-MIMO systems, the digital precoder and combiner are obtained from the effective baseband channel's SVD, while for MU-MIMO systems, the digital precoder and combiner are obtained using the regularized channel diagonalization method. Numerical simulation results have confirmed the effectiveness of the proposed design, which outperforms other existing designs in nearly all scenarios. Results have also shown the proposed method can perform very well even when practical PS quantization is considered.

# Chapter 5

## Conclusions

This chapter summarizes the contributions of this thesis and points out future research directions.

### 5.1 Summary of the Contributions

In this thesis, different switching-network-based and phase-shifting-network-based hybrid beamforming are proposed for mmWave massive MIMO wireless communication systems. The main contributions of the thesis are summarized as follows.

In Chapter 2, two new adaptive algorithms for solving the joint antenna selection and beamforming problem for uplink reception in mobile communication systems have been proposed. Both algorithms are developed based on an alternating optimization procedure, in which the beamforming vector and the antenna selection matrix are computed iteratively through LMS-like update equations. The first algorithm is designed based on the MMSE criterion, aiming at reducing the error between the received signal and a reference signal (pilot sequence). On the other hand, the second algorithm is based on the MVDR criterion with an objective to reduce the beamformer's output power while ensuring unity gain towards the user of interest. The effectiveness of the proposed algorithms has been confirmed through numerical simulations. Simulation results have shown that the proposed algorithm provides extra degrees of freedom leading to higher SINR levels in adverse scenarios, such as those with fewer RF chains than users and with interfering signals located at the vicinity of the user of interest.

In Chapter 3, a novel phase-shifting-network-based hybrid precoder and combiner design based on the constrained channel decomposition has been proposed for narrowband mmWave massive MIMO wireless communication systems. The proposed design consisted of a two-stage approach

that splits the analog and digital beamforming designs. The analog beamforming is designed using the proposed constrained channel decomposition method, which provides the analog beamforming with the capability to harvest the array gain and suppress intra-user and inter-user interference. The proposed analog beamforming design is suitable for both single-user and multi-user systems. The digital beamforming, designed at the second stage, considers two different approaches: for single-user systems, the digital beamforming is obtained from the optimal SVD-based solution followed by optimal water-filling power allocation, and for multi-user systems, it is obtained from the regularized channel diagonalization approach. The regularized channel diagonalization provides a better compromise between multi-user interference suppression and noise suppression, resulting in a better performance in critical scenarios, such as in low SNR, a small number of antennas, or a large number of transmitted data streams. Numerical simulation results have confirmed the proposed design's effectiveness, which has attained near-optimal performance in a wide range of scenarios, outperforming most of the other existing designs. Simulation results have also demonstrated the proposed design's ability to perform well even under practical phase-shifter quantization constraints. Moreover, important insights on the convergence of the proposed constrained channel decomposition have been derived and confirmed through numerical examples.

Finally, in Chapter 4, a joint phase-shifting-network-based hybrid precoder and combiner design for frequency selective mmWave massive MIMO wireless communication systems is proposed. To address the frequency-selectivity, multi-carrier OFDM is considered. Similar to Chapter 3, the proposed design also splits the analog and digital beamforming designs into two stages. In the first stage, the analog precoder and combiners are designed using the proposed constrained Tucker2 tensor decomposition approach by leveraging the OFDM channel's multidimensional structure. The proposed tensor decomposition allows the analog beamforming to harvest the array gain and to suppress intra-user and inter-user interference simultaneously in every subcarrier. The digital precoder and combiner are obtained on a per-subcarrier basis, considering the optimal SVD-based solution for single-user systems and the regularized channel diagonalization approach for multi-user systems. Simulation results have shown that the proposed design outperforms other existing methods in nearly all scenarios, thus confirming its effectiveness. The results also demonstrated the proposed design's ability to perform well under practical phase-shifter quantization constraints. Additional simulation results have provided insights into the convergence of proposed constrained Tucker2 tensor decomposition, the effect of the power method approximation (adopted to avoid the computational complexity of eigenproblem's solutions), and the differences between the designs in this chapter and that in Chapter 3.

## 5.2 Future Work

The adaptive joint antenna selection and beamforming algorithms in Chapter 2 have filled a gap in the literature by presenting algorithms that consider the limited number of RF chains throughout the entire adaptive process. To further improve the performance of the proposed algorithms and to broaden their practical applicability, the following extensions are suggested:

- Include spatial multiplexing capabilities to accommodate multiple users sharing the same channel and/or RF hardware resources.
- Develop robust adaptive joint antenna selection and beamforming. Adaptive beamforming algorithms, in particular MVDR-based ones, are very sensitive to mismatches in the direction-of-arrival estimation and the array element positions. To avoid performance degradation, the optimization problems used to derive the proposed algorithms should be modified to incorporate constraints that provide robustness against these mismatches.
- Develop stochastic models for the MMSE-AJASB and the MVDR-AJASB algorithms. Such models can be used to derive expressions for evaluating the convergence rate, the steady-state SINR, and for establishing ranges of step-size that ensure proper operation of these algorithms.

The hybrid beamforming design methods proposed in Chapter 3 and 4 have attained promising results. The constrained matrix and tensor decompositions proposed in each of these chapters are derived from very general formulations that can be extended to more practical systems. Therefore, the following extensions are suggested:

- Include more practical hybrid beamforming architectures, such as partially connected analog beamforming architectures. Such architectures are known for their reduced power consumption compared with the architecture discussed in this thesis and would allow more energy-efficient designs.
- Develop robust approaches that can lower the reliance on perfect channel state information. The designs in these chapters have considered the availability of perfect channel estimates, which is often impractical despite being a common assumption in the hybrid beamforming literature. Therefore, incorporating the channel uncertainties into the design formulations would broaden their practical applicability.

# Bibliography

- [1] Ericsson, “5G systems - enabling the transformation of industry and society,” White Paper, Jan. 2017, pp. 1–14.
- [2] —, “5G wireless access: an overview,” White Paper, Apr. 2020, pp. 1–20.
- [3] Nokia, “The value of 5G services: consumer perceptions and the opportunity for CSPs,” White Paper, Jun. 2020, pp. 1–25.
- [4] McKinsey Global Institute, “Connected world: an evolution in connectivity beyond the 5G revolution,” Discussion Paper, Feb. 2020, pp. 1–100.
- [5] —, “Agriculture’s connected future: how technology can yield new growth,” White Paper, Jan. 2020, pp. 1–10.
- [6] D. Singh, “5G’s healthcare impact: 1 billion patients with improved access in 2030,” STL Partners, Executive Briefing, Oct. 2019, pp. 1–39.
- [7] D. Adib, “5G’s impact on manufacturing: \$740bn of benefits in 2030,” STL Partners, Executive Briefing, Oct. 2019, pp. 1–40.
- [8] International Telecommunication Union - ITU, “IMT vision – framework and overall objectives of the future development of IMT for 2020 and beyond,” Recommendation ITU-R M.2083-0, Sep. 2015, pp. 1–21.
- [9] —, “ITU-R M. [IMT-2020.TECH PERF REQ] - Minimum requirements related to technical performance for IMT2020 radio interface(s),” Report ITU-R M.2410-0, Nov. 2017, pp. 1–40.
- [10] 3GPP, “Study on scenarios and requirements for next generation access technologies,” Technical Report TR38.913, ver. 16.0.0, Release 16, Jul. 2020, pp. 1–40.
- [11] H. Viswanathan and M. Weldon, “The past, present, and future of mobile communications,” *Bell Labs Technical Journal*, vol. 19, pp. 8–21, 2014.
- [12] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
- [13] J. J. G. Andrews, S. Buzzi, W. Choi, *et al.*, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

- [14] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, “Millimeter-wave massive MIMO: the next wireless revolution?” *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, Sep. 2014.
- [15] L. Wei, R. Hu, Y. Qian, and G. Wu, “Key elements to enable millimeter wave communications for 5G wireless systems,” *IEEE Wireless Communications*, vol. 21, no. 6, pp. 136–143, Dec. 2014.
- [16] V. Chandrasekhar, J. Andrews, and A. Gatherer, “Femtocell networks: a survey,” *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, Sep. 2008.
- [17] D. Muirhead, M. A. Imran, and K. Arshad, “A survey of the challenges, opportunities and use of multiple antennas in current and future 5G small cell base stations,” *IEEE Access*, vol. 4, pp. 2952–2964, 2016.
- [18] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: potentials and challenges,” *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
- [19] M. Xiao, S. Mumtaz, Y. Huang, *et al.*, “Millimeter wave communications for future mobile networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.
- [20] T. S. Rappaport, S. Sun, R. Mayzus, *et al.*, “Millimeter wave mobile communications for 5G cellular: it will work!” *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [21] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, “Millimeter-wave communications: physical channel models, design considerations, antenna constructions, and link-budget,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2018.
- [22] A. Alejos, M. Sanchez, and I. Cuinas, “Measurement and analysis of propagation mechanisms at 40 GHz: viability of site shielding forced by obstacles,” *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3369–3380, Nov. 2008.
- [23] J. Zhang, X. Ge, Q. Li, M. Guizani, and Y. Zhang, “5G millimeter-wave antenna array: design and challenges,” *IEEE Wireless Communications*, vol. 24, no. 2, pp. 106–112, Apr. 2017.
- [24] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [25] S. Kutty and D. Sen, “Beamforming for millimeter wave communications: an inclusive survey,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 949–973, 2016.
- [26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005.
- [27] T. L. Marzetta, “How much training is required for multiuser MIMO?” In *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA: IEEE, 2006, pp. 359–363.
- [28] —, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

- [29] F. Rusek, D. Persson, Buon Kiong Lau, *et al.*, “Scaling up MIMO: opportunities and challenges with very large arrays,” *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [30] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [31] T. Marzetta, E. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [32] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: benefits and challenges,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [33] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [34] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, “An overview of signal processing techniques for millimeter wave MIMO systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [35] A. Alkhateeb, Jianhua Mo, N. Gonzalez-Prelcic, and R. W. Heath, “MIMO precoding and combining solutions for millimeter-wave systems,” *IEEE Communications Magazine*, vol. 52, no. 12, pp. 122–131, Dec. 2014.
- [36] A. Molisch, M. Win, and J. Winters, “Capacity of MIMO systems with antenna selection,” in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, Helsinki: IEEE, 2001, pp. 570–574.
- [37] —, “Reduced-complexity transmit/receive-diversity systems,” *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2729–2738, Nov. 2003.
- [38] A. Molisch and M. Win, “MIMO systems with antenna selection,” *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, Mar. 2004.
- [39] S. Sanayei and A. Nosratinia, “Antenna selection in MIMO systems,” *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
- [40] A. Molisch, M. Win, Yang-Seok Choi, and J. Winters, “Capacity of MIMO systems with antenna selection,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1759–1772, Jul. 2005.
- [41] A. Gorokhov, “Antenna selection algorithms for MEA transmission systems,” in *IEEE International Conference on Acoustics Speech and Signal Processing*, vol. 3, Orlando: IEEE, May 2002, pp. 2857–2860.
- [42] M. Gharavi-Alkhansari and A. Gershman, “Fast antenna subset selection in MIMO systems,” *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 339–347, Feb. 2004.

- [43] A. Dua, K. Medepalli, and A. Paulraj, "Receive antenna selection in MIMO systems using convex optimization," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2353–2357, 2006.
- [44] K. Elkhail, A. Kammoun, T. Y. Al-Naffouri, and M.-s. Alouini, "A blind antenna selection scheme for single-cell uplink massive MIMO," in *2016 IEEE Globecom Workshops*, IEEE, Dec. 2016, pp. 1–6.
- [45] S. Mahboob, R. Ruby, and V. C. Leung, "Transmit antenna selection for downlink transmission in a massively distributed antenna system using convex optimization," in *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, Victoria: IEEE, Nov. 2012, pp. 228–233.
- [46] I. Berenguer, X. Wang, and V. Krishnamurthy, "Adaptive MIMO antenna selection," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 1, Pacific Grove: IEEE, 2003, pp. 21–26.
- [47] I. Berenguer, X. Wang, and V. Krishnamurthy, "Adaptive MIMO antenna selection via discrete stochastic optimization," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4315–4329, Nov. 2005.
- [48] O. Mehanna, N. D. Sidiropoulos, and G. B. Giannakis, "Joint multicast beamforming and antenna selection," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2660–2674, May 2013.
- [49] O. T. Demir and T. E. Tuncer, "Multicast beamforming with antenna selection using exact penalty approach," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 1, Brisbane: IEEE, Apr. 2015, pp. 2489–2493.
- [50] S. He, Y. Huang, J. Wang, L. Yang, and W. Hong, "Joint antenna selection and energy-efficient beamforming design," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1–1, Sep. 2016.
- [51] X. Wang and J. Sheng, "Joint antenna selection and beamforming algorithms for physical layer multicasting with massive antennas," *Algorithms*, vol. 9, no. 2, p. 42, Jun. 2016.
- [52] J. F. de Andrade, M. L. R. de Campos, and J. A. Apolinario, "Sparse solutions for antenna arrays," in *XXIX Simpósio Brasileiro de Telecomunicações (SBrT 2011)*, Curitiba: Sociedade Brasileira de Telecomunicação, Oct. 2011, pp. 1–5.
- [53] J. F. de Andrade, M. L. R. de Campos, and J. A. Apolinario, "An  $L_1$ -constrained normalized LMS algorithm and its application to thinned adaptive antenna arrays," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver: IEEE, May 2013, pp. 3806–3810.
- [54] J. F. de Andrade, M. L. R. de Campos, and J. A. Apolinario, " $L_1$ -constrained normalized LMS algorithms for adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6524–6539, Dec. 2015.

- [55] W. Shi, Y. Li, L. Zhao, and X. Liu, “Controllable sparse antenna array for adaptive beamforming,” *IEEE Access*, vol. 7, pp. 6412–6423, 2019.
- [56] R. Mendez-Rial, C. Rusu, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath, “Channel estimation and hybrid combining for mmWave: phase shifters or switches?” In *2015 Information Theory and Applications Workshop (ITA)*, San Diego, CA, USA: IEEE, Feb. 2015, pp. 90–97.
- [57] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, “Hybrid MIMO architectures for millimeter wave communications: phase shifters or switches?” *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.
- [58] S. Payami, M. Ghoraiishi, M. Dianati, and M. Sellathurai, “Hybrid beamforming with reduced number of phase shifters for massive MIMO systems,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 4843–4851, Feb. 2018.
- [59] Xinying Zhang, A. Molisch, and Sun-Yuan Kung, “Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection,” *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.
- [60] A. F. Molisch, V. V. Ratnam, S. Han, *et al.*, “Hybrid beamforming for massive MIMO: a survey,” *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.
- [61] F. Khalid, “Hybrid beamforming for millimeter wave massive multiuser MIMO systems using regularized channel diagonalization,” *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 705–708, Jun. 2019.
- [62] G. M. Zilli, C. A. Pitz, E. L. O. Batista, R. Seara, and W.-P. Zhu, “An adaptive approach for the joint antenna selection and beamforming optimization,” *IEEE Access*, vol. 7, pp. 99 017–99 030, 2019.
- [63] G. M. Zilli and W.-P. Zhu, “Constrained-SVD based hybrid beamforming design for millimeter-wave communications,” in *Proc. IEEE 92nd Vehicular Technology Conference: VTC2020-Fall*, Victoria, BC: IEEE, Nov. 2020, pp. 1–5.
- [64] G. M. Zilli and W.-P. Zhu, “Constrained channel decomposition-based hybrid beamforming for mmWave massive MIMO systems,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1707–1720, 2020.
- [65] G. M. Zilli and W.-P. Zhu, “TUCKER2-based hybrid beamforming design for mmWave OFDM massive MIMO communications,” in *Proc. IEEE 93rd Vehicular Technology Conference: VTC2021-Spring*, Helsinki, Finland: IEEE, Apr. 2021, pp. 1–5.
- [66] C. A. Balanis and P. I. Ioannides, *Introduction to Smart Antennas*. San Rafael, CA: Morgan & Claypool Publishers, 2007, pp. 1–175.
- [67] J. Winters, “Smart antennas for wireless systems,” *IEEE Personal Communications*, vol. 5, no. 1, pp. 23–27, 1998.

- [68] A. Ghrayeb, “A survey on antenna selection for MIMO communication systems,” in *2006 2nd International Conference on Information & Communication Technologies*, vol. 2, Damascus, Syria: IEEE, Apr. 2006, pp. 2104–2109.
- [69] Xiang Gao, O. Edfors, Jianan Liu, and F. Tufvesson, “Antenna selection in measured massive MIMO channels using convex optimization,” in *2013 IEEE Globecom Workshops*, Atlanta: IEEE, Dec. 2013, pp. 129–134.
- [70] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Massive MIMO in real propagation environments: do all antennas contribute equally?” *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.
- [71] A. Alkhateeb, Y.-H. Nam, J. Zhang, and R. W. Heath, “Massive MIMO combining with switches,” *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 232–235, Jun. 2016.
- [72] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski, “Reduced switching connectivity for large scale antenna selection,” *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2250–2263, May 2017.
- [73] S. Payami, N. Mysore Balasubramanya, C. Masouros, and M. Sellathurai, “Phase shifters versus switches: an energy efficiency perspective on hybrid beamforming,” *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 13–16, Feb. 2019.
- [74] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, Feb. 2009.
- [75] O. Frost, “An algorithm for linearly constrained adaptive array processing,” *Proceedings of the IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.
- [76] J. A. Apolinário, S. Werner, P. S. R. Diniz, and T. I. Laakso, “Constrained normalized adaptive filters for CDMA mobile communications,” in *9th European Signal Processing Conference (EUSIPCO 1998)*, Rhodes Island: IEEE, 1998.
- [77] P. S. R. Diniz, *Adaptive Filtering*. Boston, MA, USA: Springer, 2013.
- [78] B. Hassibi and B. Hochwald, “How much training is needed in multiple-antenna wireless links?” *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [79] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [80] J. A. Zhang, T. T. Yang, and Z. Chen, “Under-determined training and estimation for distributed transmit beamforming systems,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1936–1946, Apr. 2013.
- [81] S. Noh, M. D. Zoltowski, and D. J. Love, “Training sequence design for feedback assisted hybrid beamforming in massive MIMO systems,” *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 187–200, Jan. 2016.
- [82] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.

- [83] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
- [84] S. Payami, M. Ghorraishi, and M. Dianati, "Hybrid beamforming for large antenna arrays with phase shifter selection," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7258–7271, Nov. 2016.
- [85] D. Zhang, P. Pan, R. You, and H. Wang, "SVD-based low-complexity hybrid precoding for millimeter-wave MIMO systems," *IEEE Communications Letters*, vol. 22, no. 10, pp. 2176–2179, Oct. 2018.
- [86] M. Li, Z. Wang, X. Tian, and Q. Liu, "Joint hybrid precoder and combiner design for multi-stream transmission in mmWave MIMO systems," *IET Communications*, vol. 11, no. 17, pp. 2596–2604, 2017.
- [87] F. Sotrabadi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
- [88] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, Dec. 2014.
- [89] J. Li, L. Xiao, X. Xu, and S. Zhou, "Robust and low complexity hybrid beamforming for uplink multiuser mmWave MIMO systems," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1140–1143, Jun. 2016.
- [90] A. Alkhateeb, R. W. Heath, and G. Leus, "Achievable rates of multi-user millimeter wave systems with hybrid precoding," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, London, United Kingdom: IEEE, Jun. 2015, pp. 1232–1237.
- [91] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
- [92] Z. Wang, M. Li, X. Tian, and Q. Liu, "Iterative hybrid precoder and combiner design for mmWave multiuser MIMO systems," *IEEE Communications Letters*, vol. 21, no. 7, pp. 1581–1584, Jul. 2017.
- [93] A. Li and C. Masouros, "Hybrid precoding and combining design for millimeter-wave multi-user MIMO based on SVD," in *2017 IEEE International Conference on Communications (ICC)*, IEEE, May 2017, pp. 1–6.
- [94] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 201–211, Jan. 2016.
- [95] R. Rajashekar and L. Hanzo, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1372–1384, Mar. 2017.

- [96] X. Wu, D. Liu, and F. Yin, "Hybrid beamforming for multi-user massive MIMO systems," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 3879–3891, Sep. 2018.
- [97] C. Hu, J. Liu, X. Liao, Y. Liu, and J. Wang, "A novel equivalent baseband channel of hybrid beamforming in massive multiuser MIMO systems," *IEEE Communications Letters*, vol. 22, no. 4, pp. 764–767, Apr. 2018.
- [98] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [99] A Goldsmith, S. Jafar, N Jindal, and S Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, Jun. 2003.
- [100] Chongning Na, Xiaolin Hou, and A. Harada, "Enhanced iterative max-sum-rate algorithm for linear MU-MIMO precoding," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, IEEE, Sep. 2011, pp. 1495–1499.
- [101] C.-R. Tsai, Y.-H. Liu, and A.-Y. Wu, "Efficient compressive channel estimation for millimeter-wave large-scale antenna systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 9, pp. 2414–2428, May 2018.
- [102] X. Cheng, C. Tang, and Z. Zhang, "Accurate channel estimation for millimeter-wave MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 5159–5163, May 2019.
- [103] G. Raleigh and J. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, Mar. 1998.
- [104] H.-J. M. Shi, S. Tu, Y. Xu, and W. Yin, "A primer on coordinate descent algorithms," Sep. 2016. [Online]. Available: <http://arxiv.org/abs/1610.00040>.
- [105] Y. Chi, Y. M. Lu, and Y. Chen, "Nonconvex optimization meets low-rank matrix factorization: an overview," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5239–5269, Oct. 2019.
- [106] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, Feb. 2005.
- [107] M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [108] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 953–961, Mar. 2008.
- [109] F. Khalid and J. Speidel, "Robust hybrid precoding for multiuser MIMO wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3353–3363, Jun. 2014.
- [110] J. C. Chen, "Hybrid beamforming with discrete phase shifters for millimeter-wave massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7604–7608, 2017.

- [111] Z. Wang, M. Li, Q. Liu, and A. L. Swindlehurst, “Hybrid precoder and combiner design with low resolution phase shifters in mmWave MIMO systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 256–269, Apr. 2018.
- [112] F. Dong, W. Wang, and Z. Wei, “Low-complexity hybrid precoding for multi-user mmWave systems with low-resolution phase shifters,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9774–9784, Oct. 2019.
- [113] A. Devices. “RF, microwave, and millimeter wave products: selection guide.” <https://www.analog.com/en/products/rf-microwave.html>. (2020), (visited on 07/20/2020).
- [114] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2013.
- [115] Bolckei, Gesbert, and Paulraj, “On the capacity of OFDM-based spatial multiplexing systems,” *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 225–234, Feb. 2002.
- [116] A. Alkhateeb and R. W. Heath, “Frequency selective hybrid precoding for limited feedback millimeter wave systems,” *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1801–1818, 2016.
- [117] F. Sotirani and W. Yu, “Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.
- [118] T.-H. Tsai, M.-C. Chiu, and C.-C. Chao, “Sub-system SVD hybrid beamforming design for millimeter wave multi-carrier systems,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 518–531, Jan. 2019.
- [119] H.-L. Chiang, W. Rave, T. Kadur, and G. Fettweis, “Hybrid beamforming based on implicit channel state information for millimeter wave links,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 326–339, May 2018.
- [120] J. Du, W. Xu, C. Zhao, and L. Vandendorpe, “Weighted spectral efficiency optimization for hybrid beamforming in multiuser massive MIMO-OFDM systems,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9698–9712, Oct. 2019.
- [121] Y. Liu and J. Wang, “Low-complexity OFDM-based hybrid precoding for multiuser massive MIMO systems,” *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 263–266, Mar. 2020.
- [122] J. P. Gonzalez-Coma, J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, L. Castedo, and R. W. Heath, “Channel estimation and hybrid precoding for frequency selective multiuser mmWave MIMO systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 353–367, May 2018.
- [123] D. Zhang, Y. Wang, X. Li, and W. Xiang, “Hybrid beamforming for downlink multiuser millimetre wave MIMO-OFDM systems,” *IET Communications*, vol. 13, no. 11, pp. 1557–1564, Jul. 2019.

- [124] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, “A constrained factor decomposition with application to MIMO antenna systems,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2429–2442, Jun. 2008.
- [125] Z. Zhou, J. Fang, L. Yang, *et al.*, “Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1524–1538, 2017.
- [126] D. C. Araujo, A. L. F. de Almeida, J. P. C. L. Da Costa, and R. T. de Sousa, “Tensor-based channel estimation for massive MIMO-OFDM systems,” *IEEE Access*, vol. 7, pp. 42 133–42 147, 2019.
- [127] M. Ruble and I. Guvenc, “Multilinear singular value decomposition for millimeter wave channel Parameter estimation,” *IEEE Access*, vol. 8, pp. 75 592–75 606, 2020.
- [128] L. Liu and Y. Tian, “Hybrid precoding based on tensor decomposition for mmWave 3D-MIMO systems,” in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, IEEE, Oct. 2017, pp. 1–6.
- [129] Chanhong Kim, Taeyoung Kim, and Ji-Yun Seol, “Multi-beam transmission diversity with hybrid beamforming for MIMO-OFDM systems,” in *2013 IEEE Globecom Workshops (GC Wkshps)*, IEEE, Dec. 2013, pp. 61–65.
- [130] J. Lee and Y. H. Lee, “AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing,” in *2014 IEEE International Conference on Communications (ICC)*, IEEE, Jun. 2014, pp. 5838–5842.
- [131] K. Venugopal, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath, “Channel estimation for hybrid architecture-based wideband millimeter wave systems,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [132] J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, K. Venugopal, and R. W. Heath, “Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 2946–2960, May 2018.
- [133] H. Kim, G.-T. Gil, and Y. H. Lee, “Two-step approach to time-domain channel estimation for wideband millimeter wave systems with hybrid architecture,” *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5139–5152, Jul. 2019.
- [134] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, “Wideband MIMO channel estimation for hybrid beamforming millimeter wave systems via random spatial sampling,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 1136–1150, Sep. 2019.
- [135] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, Jul. 2009.
- [136] V. Y. Pan and Z. Q. Chen, “The complexity of the matrix eigenproblem,” in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, ser. STOC '99, Atlanta, Georgia, USA: Association for Computing Machinery, 1999, 507–516.

- [137] P. Xu, B. He, C. D. Sa, I. Mitliagkas, and C. Re, “Accelerated stochastic power iteration,” in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, A. Storkey and F. Perez-Cruz, Eds., ser. Proceedings of Machine Learning Research, vol. 84, PMLR, Apr. 2018, pp. 58–67.
- [138] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd ed. New York, NY, USA: Springer, 2008.
- [139] T. G. Kolda, “Multilinear operators for higher-order decompositions,” Sandia National Laboratories, Albuquerque, NM, US and Livermore, CA, US, Tech. Rep., 2006.
- [140] L. De Lathauwer, B. De Moor, and J. Vandewalle, “A multilinear singular value decomposition,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, Jan. 2000.
- [141] —, “On the best rank-1 and rank-(R1, R2, ..., RN) approximation of higher-order tensors,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, Jan. 2000.

# Appendices

## Appendix A

### Obtaining (3.13) from (3.12)

Given the original problem in (3.12), one can obtain the optimal solution w.r.t.  $\mathbf{\Sigma}$  for a given  $\mathbf{F}_{\text{RF}}$  and  $\mathbf{W}_{\text{RF}}$  by solving

$$\begin{aligned}
 \min_{\mathbf{\Sigma}} \quad & \|\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 \\
 \text{s.t.} \quad & \mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{W}_{\text{RF}} = \mathbf{I}_{KN_s} \\
 & \mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{F}_{\text{RF}} = \mathbf{I}_{KN_s} \\
 & \mathbf{W}_{\text{RF}} = \text{blkdiag}\{\mathbf{W}_{\text{RF},1}, \dots, \mathbf{W}_{\text{RF},K}\} \\
 & \mathbf{W}_{\text{RF},k} \in \mathcal{W}_{\text{RF}} \\
 & \mathbf{F}_{\text{RF}} \in \mathcal{F}_{\text{RF}}.
 \end{aligned} \tag{A.1}$$

The cost function in (A.1) can be rewritten as

$$\begin{aligned}
 \|\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 &= \text{Tr}\{(\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}})^{\text{H}}(\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}})\} \\
 &= \text{Tr}\{\mathbf{H}^{\text{H}}\mathbf{H} - \mathbf{H}^{\text{H}}\mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}} - \mathbf{F}_{\text{RF}}\mathbf{\Sigma}^{\text{H}}\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{H} + \mathbf{F}_{\text{RF}}\mathbf{\Sigma}^{\text{H}}\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\} \\
 &= \text{Tr}\{\mathbf{H}^{\text{H}}\mathbf{H}\} - \text{Tr}\{\mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{W}_{\text{RF}}\mathbf{\Sigma}\} - \text{Tr}\{\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{\Sigma}^{\text{H}}\} \\
 &\quad + \text{Tr}\{\mathbf{\Sigma}^{\text{H}}\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{F}_{\text{RF}}\}.
 \end{aligned} \tag{A.2}$$

Assuming that the constraints in (A.1) are satisfied, the cost function in (A.2) reduces to

$$\|\mathbf{H} - \mathbf{W}_{\text{RF}}\mathbf{\Sigma}\mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 = \text{Tr}\{\mathbf{H}^{\text{H}}\mathbf{H}\} - \text{Tr}\{\mathbf{F}_{\text{RF}}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{W}_{\text{RF}}\mathbf{\Sigma}\} - \text{Tr}\{\mathbf{W}_{\text{RF}}^{\text{H}}\mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{\Sigma}^{\text{H}}\} + \text{Tr}\{\mathbf{\Sigma}^{\text{H}}\mathbf{\Sigma}\}. \tag{A.3}$$

Therefore, the optimal  $\Sigma$  is obtained by setting the derivative of (A.3) w.r.t.  $\Sigma^*$  to zero, i.e.,

$$\frac{\partial}{\partial \Sigma^*} \left( \|\mathbf{H} - \mathbf{W}_{\text{RF}} \Sigma \mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 \right) = -\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}} + \Sigma = \mathbf{0} \quad (\text{A.4})$$

which yields

$$\Sigma_{\text{opt}} = \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}. \quad (\text{A.5})$$

Next, by taking  $\Sigma_{\text{opt}}$  back into (A.3), one obtains

$$\begin{aligned} \|\mathbf{H} - \mathbf{W}_{\text{RF}} \Sigma_{\text{opt}} \mathbf{F}_{\text{RF}}^{\text{H}}\|_{\text{F}}^2 &= \text{Tr} \{ \mathbf{H}^{\text{H}} \mathbf{H} \} - \text{Tr} \{ \mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{H}^{\text{H}} \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}} \} \\ &\quad - \text{Tr} \{ \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{H}^{\text{H}} \mathbf{W}_{\text{RF}} \} + \text{Tr} \{ \mathbf{F}_{\text{RF}}^{\text{H}} \mathbf{H}^{\text{H}} \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}} \} \\ &= \|\mathbf{H}\|_{\text{F}}^2 - \|\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}\|_{\text{F}}^2 - \|\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}\|_{\text{F}}^2 + \|\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}\|_{\text{F}}^2 \\ &= \|\mathbf{H}\|_{\text{F}}^2 - \|\mathbf{W}_{\text{RF}}^{\text{H}} \mathbf{H} \mathbf{F}_{\text{RF}}\|_{\text{F}}^2. \end{aligned} \quad (\text{A.6})$$

Finally, since  $\|\mathbf{H}\|_{\text{F}}^2$  does not depend on the optimization variables of the original problem (3.12), replacing (A.6) into (3.12) leads to problem (3.13).

## Appendix B

# Proof of the optimality of (3.20)

By adopting a different formulation to problem (3.15) and using the first and second order optimality conditions, it can be demonstrated that the projected block coordinate solutions in (3.20) are, in fact, local optimal solutions to the non-convex block coordinate problems of (3.15). This result enables the derivation of important insights on the proposed algorithm's convergence. For conciseness, the indices in the vectors and matrices are omitted and only the solution to the block coordinate problem associated with the analog combiner vector is presented. Such a problem is obtained from (3.15) as

$$\begin{aligned} \max_{\mathbf{w}} \quad & \text{Re}(\mathbf{w}^H \mathbf{H}_{\text{res}} \mathbf{f}) \\ \text{s.t.} \quad & \|\mathbf{w}\| = 1 \\ & |[\mathbf{w}]_n| = 1/\sqrt{N_r}, \quad \forall n \in \{1, \dots, N_r\} \end{aligned} \tag{B.1}$$

Note that: i) the second constraint in (B.1) guarantees that the first constraint be attained; and ii) from the second constraint, the analog precoder vector can be expressed as

$$\mathbf{w} = \frac{1}{\sqrt{N_r}} [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_{N_r}}]^T = \frac{1}{\sqrt{N_r}} e^{j\boldsymbol{\theta}},$$

with  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_{N_r}]^T$ . Thus, by defining  $\mathbf{h} = [h_1, h_2, \dots, h_{N_r}]^T = \mathbf{H}_{\text{res}} \mathbf{f}$ , with  $h_i = |h_i| e^{j\phi_i}$ ,

the cost function in (B.1) can be reformulated as

$$\begin{aligned}
\mathcal{J} = \text{Re}\{\mathbf{w}^H \mathbf{H} \mathbf{f}\} &= \frac{1}{2} (\mathbf{w}^H \mathbf{H} \mathbf{f} + \mathbf{f}^H \mathbf{H}^H \mathbf{w}) \\
&= \frac{1}{2\sqrt{N_r}} \left( \sum_{i=1}^{N_r} e^{-j\theta_i} h_i + \sum_{i=1}^N e^{j\theta_i} h_i^* \right) \\
&= \frac{1}{2\sqrt{N_r}} \sum_{i=1}^{N_r} |h_i| \left( e^{-j(\theta_i - \phi_i)} + e^{j(\theta_i - \phi_i)} \right) \\
&= \frac{1}{\sqrt{N_r}} \sum_{i=1}^{N_r} |h_i| \text{Re}\{e^{-j(\theta_i - \phi_i)}\}
\end{aligned} \tag{B.2}$$

Therefore, the optimization problem in (B.1) is equivalent to the following unconstrained optimization problem

$$\arg \max_{\boldsymbol{\theta} \in \mathbb{R}^{N_r}} \sum_{i=1}^{N_r} |h_i| \text{Re}\{e^{-j(\theta_i - \phi_i)}\}. \tag{B.3}$$

The solution to problem (B.3) is obtained using the first- and second-order optimality conditions [138], i.e.,

$$\nabla_{\boldsymbol{\theta}} \mathcal{J} = \frac{1}{\sqrt{N_r}} \begin{bmatrix} |h_1| \text{Re}\{e^{-j(\theta_1 - \phi_1 + \pi/2)}\} \\ |h_2| \text{Re}\{e^{-j(\theta_2 - \phi_2 + \pi/2)}\} \\ \vdots \\ |h_{N_r}| \text{Re}\{e^{-j(\theta_{N_r} - \phi_{N_r} + \pi/2)}\} \end{bmatrix} = \mathbf{0} \tag{B.4}$$

and

$$\nabla_{\boldsymbol{\theta}}^2 \mathcal{J} = \frac{1}{\sqrt{N_r}} \text{diag} \left( \begin{bmatrix} |h_1| \text{Re}\{e^{-j(\theta_1 - \phi_1 + \pi)}\} \\ |h_2| \text{Re}\{e^{-j(\theta_2 - \phi_2 + \pi)}\} \\ \vdots \\ |h_{N_r}| \text{Re}\{e^{-j(\theta_{N_r} - \phi_{N_r} + \pi)}\} \end{bmatrix} \right) \prec \mathbf{0} \tag{B.5}$$

where  $\text{diag}(\mathbf{v})$  is a diagonal matrix with diagonal entries given by  $\mathbf{v}$  and  $\mathbf{A} \prec 0$  when  $\mathbf{A}$  is negative-definite. Together, these conditions yield multiple local optimal solutions, given as

$$\theta_i = \phi_i \pm 2\pi k, \quad \text{for } k = 0, 1, 2, \dots \tag{B.6}$$

Thus, by choosing the trivial solution ( $k = 0$ ), one obtains  $\theta_i = \phi_i$ , so that an optimal solution to problem (B.3) is given by

$$\boldsymbol{\theta} = \boldsymbol{\phi} = \angle\{\mathbf{H} \mathbf{f}\} \tag{B.7}$$

or equivalently, an optimal solution to problem (B.1) is given by

$$\mathbf{w} = \frac{1}{\sqrt{N_r}} e^{j\phi} = \frac{1}{\sqrt{N_r}} e^{j\angle \mathbf{H}_{\text{res}} \mathbf{f}}. \quad (\text{B.8})$$

# Appendix C

## Tensors Preliminaries

This appendix reviews the tensor notations and main operations adopted in this thesis and a brief description of the Tucker decomposition. For complete thorough overview of the tensor notation, operation, and other decompositions, readers are referred to [135].

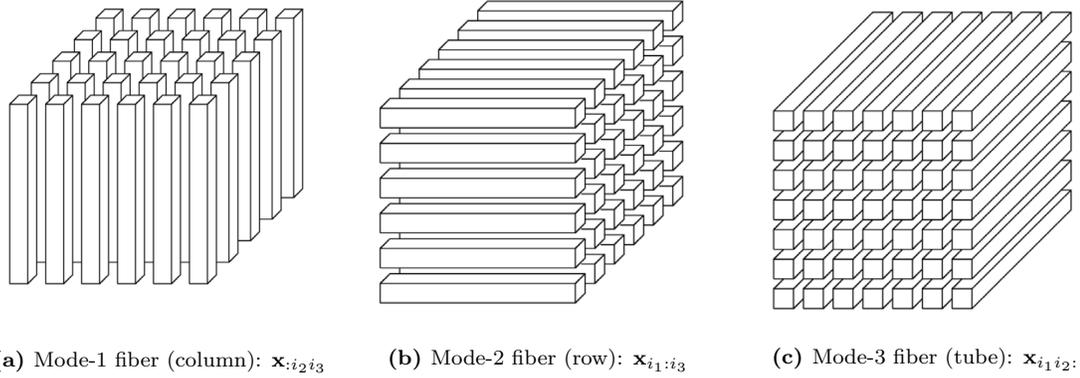
### C.1 Tensor Notations and Operations

Tensors are multidimensional arrays and an  $N$ -order tensor is formally described as the tensor product of  $N$  vector spaces, each having its own coordinate system [135]. Tensors can be thought of as a generalization of matrices: a vector is a first-order tensor; a matrix is a second-order tensor; and tensor of order three or higher are called higher-order tensors. Without loss of generality, this section focuses on the third-order tensors, since this is the case in the problems that this thesis aims to solve.

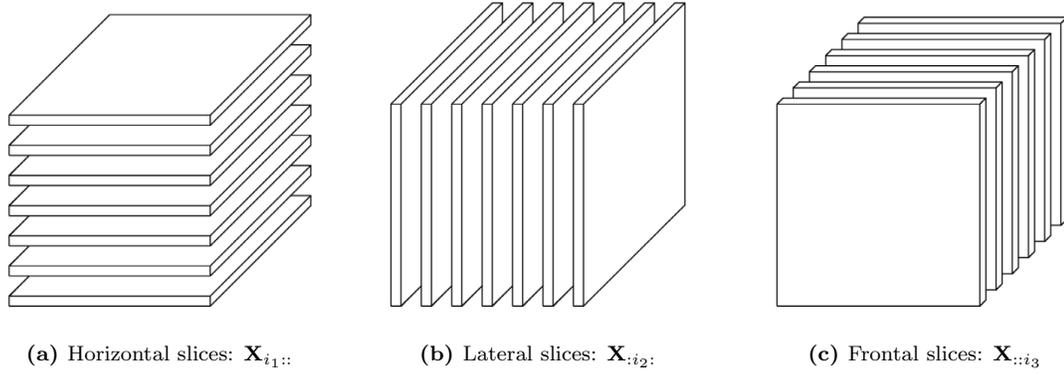
Thus, in the remaining of this section, consider a third-order tensor  $\mathcal{X}$ , for which the  $(i_1, i_2, i_3)$  element is denoted as  $x_{i_1 i_2 i_3}$ , with  $i_1 = 1, \dots, I_1$ ;  $i_2 = 1, \dots, I_2$ ; and  $i_3 = 1, \dots, I_3$ . The following subsections describe the fibers and slices, which are sub-arrays obtained from tensors; the tensor norm, matricization, and multiplication operations; and some other useful operations.

#### C.1.1 Fibers

Fibers are defined by fixing all index of a tensor but one. They are the analogous to rows and columns in a matrix. The fibers of a third-order tensor are named column, row, and tube, and are defined by  $\mathbf{x}_{:i_2 i_3}$ ,  $\mathbf{x}_{i_1 :i_3}$ , and  $\mathbf{x}_{i_1 i_2 :}$ , respectively, as shown in Fig. C.1.



**Figure C.1:** Fibers of a 3rd-order tensor [135]. Reproduced with permission from SIAM (Society for Industrial and Applied Mathematics).



**Figure C.2:** Slices of a 3rd-order tensor [135]. Reproduced with permission from SIAM.

### C.1.2 Slices

Slices are defined by fixing all index of a tensor but two. In a third-order tensors there can be horizontal, lateral, and frontal slices, as shown in Fig C.2. The horizontal, lateral, and frontal slices are, respectively, defined as by  $\mathbf{X}_{i_1::}$ ,  $\mathbf{X}_{:i_2:}$ , and  $\mathbf{X}_{::i_3}$ . For a more compact notation, the k-th frontal slice  $\mathbf{X}_{::k}$  can be denoted as  $\mathbf{X}_k$ .

### C.1.3 Inner Product and Norm

The inner product of two tensors  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$  and  $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$  is the sum of the products of their entries, defined as

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} x_{i_1 i_2 i_3} y_{i_1 i_2 i_3}^*.$$

The norm of a third-order tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$  is the square root of the sum of the squared magnitude of all its elements, i.e.,

$$\|\mathcal{X}\| = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} |x_{i_1 i_2 i_3}|^2}.$$

### C.1.4 Matricization

Matricization is the process of rearranging the elements of a tensor into a matrix. This process is also known as unfolding or flattening. The mode- $n$  matricization of a third-order tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$  is denoted by  $\mathbf{X}_{(n)}$  and arranges the mode- $n$  fibers as the columns of the resulting matrix.

The following example illustrate the matricization process. Consider a tensor  $\mathcal{X} \in \mathbb{C}^{4 \times 3 \times 2}$  with frontal slices

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} 13 & 14 & 15 \\ 16 & 17 & 18 \\ 19 & 20 & 21 \\ 22 & 23 & 24 \end{bmatrix}.$$

The mode-1, mode-2, and mode-3 matrices are, respectively

$$\mathbf{X}_{(1)} = \begin{bmatrix} 1 & 2 & 3 & 13 & 14 & 15 \\ 4 & 5 & 6 & 16 & 17 & 18 \\ 7 & 8 & 9 & 19 & 20 & 21 \\ 10 & 11 & 12 & 22 & 23 & 24 \end{bmatrix},$$

$$\mathbf{X}_{(2)} = \begin{bmatrix} 1 & 4 & 7 & 10 & 13 & 16 & 19 & 22 \\ 2 & 5 & 8 & 11 & 14 & 17 & 20 & 23 \\ 3 & 6 & 9 & 12 & 15 & 18 & 21 & 24 \end{bmatrix},$$

$$\mathbf{X}_{(3)} = \begin{bmatrix} 1 & 4 & 7 & 10 & 2 & 5 & 8 & 11 & 3 & 6 & 9 & 12 \\ 13 & 16 & 19 & 22 & 14 & 17 & 20 & 23 & 15 & 18 & 21 & 24 \end{bmatrix}.$$

Note that different column orderings can be adopted. In general, these orderings are not important, as long as it is consistent across all calculations [135].

### C.1.5 Tensor Multiplication

Here, only the  $n$ -mode product is considered, although different tensor products have been defined in the literature; see [135] and references therein.

The  $n$ -mode product refers to the product between a tensor and a matrix (or a vector) in mode  $n$ . Let a tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  and a matrix  $\mathbf{A} \in \mathbb{C}^{J \times I_n}$ . The  $n$ -mode product, denoted by  $\mathcal{X} \times_n \mathbf{A}$ , has entry-wise elements

$$[\mathcal{X} \times_n \mathbf{A}]_{i_1 \dots i_{n-1} j i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \dots i_N} a_{j i_n}$$

and is of size  $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$ . The  $n$ -mode product can also be denoted in terms of metricized tensors as

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{A} \Leftrightarrow \mathbf{Y}_{(n)} = \mathbf{A} \mathbf{X}_{(n)}.$$

### C.1.6 Other Useful Operations

Other operations are also useful when operating with tensor. These operations are defined below.

#### Outer Product

The outer product of two vectors  $\mathbf{a} \in \mathbb{C}^I$  and  $\mathbf{b} \in \mathbb{C}^J$ , denoted by  $\mathbf{a} \circ \mathbf{b}$ , is defined as

$$\mathbf{a} \circ \mathbf{b} = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \dots & a_1 b_J \\ a_2 b_1 & a_2 b_2 & \dots & a_2 b_J \\ \vdots & \vdots & \ddots & \vdots \\ a_I b_1 & a_I b_2 & \dots & a_I b_J \end{bmatrix}$$

where  $\mathbf{a} \circ \mathbf{b}$  is an  $I \times J$  matrix. The outer product can be extended (or concatenated) such that  $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ , assuming  $\mathbf{c} \in \mathbb{C}^K$ , will result in a third-order tensor  $\mathcal{X} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c} \in \mathbb{C}^{I \times J \times K}$ .

## Kronecker Product

The Kronecker product of matrices  $\mathbf{A} \in \mathbb{C}^{I \times J}$  and  $\mathbf{B} \in \mathbb{C}^{K \times L}$ , denoted by  $\mathbf{A} \otimes \mathbf{B}$ , is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \cdots & a_{IJ}\mathbf{B} \end{bmatrix}$$

where  $\mathbf{A} \otimes \mathbf{B}$  is an  $(IK) \times (JL)$  matrix.

A few important properties of the Kronecker product are also useful when operating with tensors:

- $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ ;
- $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$ ;
- $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$ ;
- Let  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$  and  $\mathbf{A}^{(n)} \in \mathbb{R}^{J_n \times I_n}$  for all  $n \in 1, \dots, N$ . Then, for any  $n \in 1, \dots, N$ , the following holds [139]:

$$\begin{aligned} \mathcal{Y} = \mathcal{X} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)} &\Leftrightarrow \\ \mathbf{Y}_{(n)} = \mathbf{A}^{(n)} \mathbf{X}_{(n)} \left( \mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)} \right)^T &. \end{aligned}$$

## C.2 The Tucker Decomposition

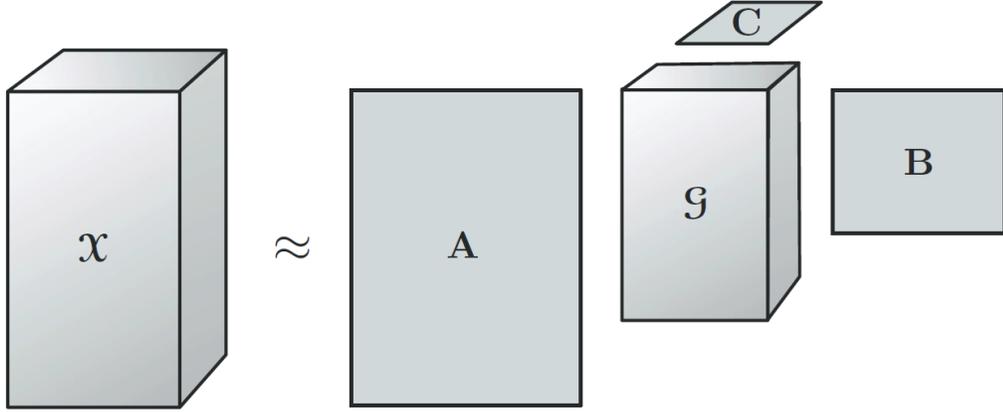
The Tucker decomposition<sup>1</sup> decomposes a tensor as a product of a core tensor and a factor matrices along each mode. The Tucker decomposition of a  $N$ -order tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$  is defined as [135]

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} g_{r_1 r_2 \cdots r_N} \mathbf{a}_{r_1}^{(1)} \circ \mathbf{a}_{r_2}^{(2)} \circ \cdots \circ \mathbf{a}_{r_N}^{(N)} \quad (\text{C.1})$$

where  $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$  for  $n \in 1, \dots, N$  are the factor matrices (which are usually orthogonal) and  $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times \cdots \times R_N}$  is the core tensor. Such a decomposition can be also expressed in its matricized version as

$$\mathbf{X}_{(n)} = \mathbf{A}^{(n)} \mathbf{G}_{(n)} (\mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)})^T. \quad (\text{C.2})$$

<sup>1</sup>The Tucker decomposition is also known by other names, such as  $N$ -mode PCA, higher-order SVD (HOSVD), and  $N$ -mode SVD; or the three mode factor analysis and three mode PCA, which are specific to third-order tensors [135].



**Figure C.3:** Tucker decomposition of a 3rd-order tensor [135]. Reproduced with permission from SIAM.

For a third-order tensor  $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ , the Tucker decomposition is depicted in Fig. C.3. In this case, the 3-way Tucker decomposition is mathematically expressed as

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} \mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r \quad (\text{C.3})$$

where  $\mathcal{G} \in \mathbb{C}^{P \times Q \times R}$  and  $\mathbf{A} \in \mathbb{C}^{I \times P}$ ,  $\mathbf{B} \in \mathbb{C}^{J \times Q}$ , and  $\mathbf{C} \in \mathbb{C}^{K \times R}$ .

An important variation of the Tucker decomposition is the Tucker2 decomposition of a third-order tensor, obtained by setting one of the factor matrix (e.g.  $\mathbf{C}$ ) to identity matrix. Thus, the Tucker2 decomposition can be expressed as

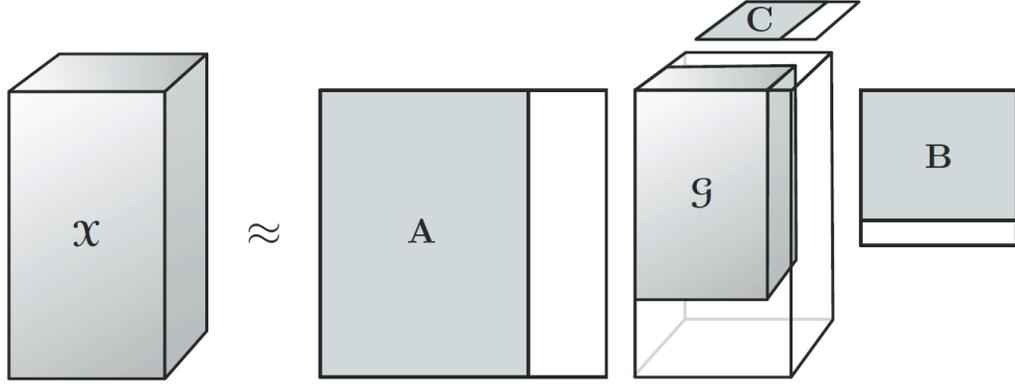
$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{I}_K \quad (\text{C.4})$$

such that  $\mathcal{G} \in \mathbb{C}^{P \times Q \times R}$  with  $R = K$  and  $\mathbf{I}_K$  is a  $K \times K$  identity matrix.

### C.2.1 The $n$ -Rank and the Truncated Tucker Decomposition

Given a  $N$ -order tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ , the  $n$ -rank of  $\mathcal{X}$ , denoted as  $\text{rank}_n(\mathcal{X})$ , is defined as the column rank of the mode- $n$  matricization  $\mathbf{X}_{(n)}$ . Thus, a tensor  $\mathcal{X}$  is said a rank- $(R_1, R_2, \dots, R_N)$  tensor if  $R_n = \text{rank}_n(\mathcal{X})$  for all  $n = 1, 2, \dots, N$ . It easily follows that  $R_n \leq I_n$  for all  $n = 1, 2, \dots, N$ .

For a given tensor  $\mathcal{X}$ , one can find an exact rank- $(R_1, R_2, \dots, R_N)$  Tucker decomposition, where  $R_n = \text{rank}_n(\mathcal{X})$ . On the other hand, if a rank- $(R_1, R_2, \dots, R_N)$  Tucker decomposition has  $R_n < \text{rank}_n(\mathcal{X})$  for one or more  $n$ , the Tucker decomposition will necessarily be inexact and, its computation, more cumbersome [135]. In such cases, the decomposition can be referred to as the



**Figure C.4:** Truncated Tucker decomposition of a 3rd-order tensor [135]. Reproduced with permission from SIAM.

truncated Tucker decomposition. The truncated Tucker decomposition of a third-order tensor is illustrated in Fig. C.4.

## C.2.2 Tucker Decomposition Computation

The rank- $(R_1, R_2, \dots, R_N)$  Tucker decomposition of a  $N$ -order tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  is defined as [135]

$$\mathcal{X} \approx \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_N \mathbf{A}^{(N)} \quad (\text{C.5})$$

where  $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$  and  $R_n < \text{rank}_n(\mathcal{X})$  for all  $n = 1, 2, \dots, N$ , and  $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times \dots \times R_N}$ . For the sake of tractability, most methods aim to find factor matrices that are orthonormal, i.e.,  $\mathbf{A}^{(n)}$  such that  $\mathbf{A}^{(n)\text{H}} \mathbf{A}^{(n)} = \mathbf{I}_{R_n}$  for all  $n = 1, 2, \dots, N$ . Thus, the decomposition in (C.5) is obtained by solving the following problem:

$$\begin{aligned} \min_{\mathcal{G}, \mathbf{A}^{(1)}, \dots, \mathbf{A}^{(N)}} & \left\| \mathcal{X} - \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_N \mathbf{A}^{(N)} \right\|^2 \\ \text{s.t.} & \mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times \dots \times R_N} \\ & \mathbf{A}^{(n)\text{H}} \mathbf{A}^{(n)} = \mathbf{I}_{R_n}, \forall n = 1, 2, \dots, N. \end{aligned} \quad (\text{C.6})$$

By extending the results in Appendix A to tensors, one can show that the solution to problem (C.6) can be equivalently obtained by solving

$$\begin{aligned} \max_{\mathbf{A}^{(1)}, \dots, \mathbf{A}^{(N)}} & \left\| \mathcal{X} \times_1 \mathbf{A}^{(1)\text{H}} \times_2 \mathbf{A}^{(2)\text{H}} \dots \times_N \mathbf{A}^{(N)\text{H}} \right\|^2 \\ \text{s.t.} & \mathbf{A}^{(n)\text{H}} \mathbf{A}^{(n)} = \mathbf{I}_{R_n}, \forall n = 1, 2, \dots, N. \end{aligned} \quad (\text{C.7})$$

Proof can be found in [135], [140], [141].

The solution of (C.7) can then be obtained by the alternating least squares (ALS) method, in which the problem is alternately solved for one factor matrix at a time, while fixing the others. Thereby, the solution for the factor matrix  $\mathbf{A}^{(n)}$  is obtained by solving [135], [140], [141]

$$\begin{aligned} \max_{\mathbf{A}^{(n)}} \quad & \left\| \mathcal{X} \times_1 \mathbf{A}^{(1)\text{H}} \times_2 \mathbf{A}^{(2)\text{H}} \cdots \times_N \mathbf{A}^{(N)\text{H}} \right\|^2 \\ \text{s.t.} \quad & \mathbf{A}^{(n)\text{H}} \mathbf{A}^{(n)} = \mathbf{I}_{R_n}. \end{aligned} \tag{C.8}$$

Finally, by using the identity

$$\left\| \mathcal{X} \times_1 \mathbf{A}^{(1)\text{H}} \times_2 \mathbf{A}^{(2)\text{H}} \cdots \times_N \mathbf{A}^{(N)\text{H}} \right\|^2 = \left\| \mathbf{A}^{(n)\text{H}} \mathbf{X}_{(n)} (\mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)})^* \right\|_{\text{F}}^2$$

problem (C.8) can be rewritten as

$$\begin{aligned} \max_{\mathbf{A}^{(n)}} \quad & \left\| \mathbf{A}^{(n)\text{H}} \mathbf{X}_{(n)} (\mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)})^* \right\|_{\text{F}}^2 \\ \text{s.t.} \quad & \mathbf{A}^{(n)\text{H}} \mathbf{A}^{(n)} = \mathbf{I}_{R_n}. \end{aligned} \tag{C.9}$$

It is worth noticing that Tucker decompositions are not unique and that the ALS solution is not guaranteed to converge to the global optimum of (C.6), but only to points where the objective function cease to decrease [135]. Moreover, different algorithmic solutions to (C.6) have been proposed. The most well known are, perhaps, the higher-order SVD (HOSVD) [140] and the higher-order orthogonal iteration (HOOI) [141].