



JACOBS
UNIVERSITY

Frame-theoretic Designs for Future Wireless Communications

by

Răzvan-Andrei Stoica

a Thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

Approved Dissertation Committee

Prof. Dr. Giuseppe Abreu
Jacobs University Bremen

Prof. Dr.-Ing. Werner Henkel
Jacobs University Bremen

Prof. Dr. George Alexandropoulos
National and Kapodistrian University of Athens

Date of defense: September 3rd, 2019

Department of Computer Science & Electrical Engineering

Statutory Declaration

Family Name, Given/First Name	STOICA, Răzvan-Andrei
Matriculation number	20329917
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	PhD-Thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

03.09.2019, Razvan-Andrei Stoica

Date, Signature

Abstract

Motivated by the historically proven depletion of communication resources, and respectively, by the lack of inherent redundancy necessary for robust and accurate signal processing and identification under orthogonal wireless communications systems, we use in this work over-complete redundant representations, *i.e.*, Frame Theory. To this point, such non-orthogonal signal decompositions provide alternative, yet robust representations, and respectively, aid signal processing of modern and future wireless communications systems beyond the limitations of their traditional orthogonal counterparts.

For instance, the millimeter wave (mmWave) initial access (IA) sparse channel estimation and training beamforming optimization problems are reduced to the generic frame design of the measurement matrix for compressed sensing (CS). Similarly, for non-orthogonal multiple access (NOMA) systems, the code-domain NOMA (CD-NOMA) transmission is abstracted as a generic linear model based on the synthesis linear operator of a frame. Under this abstraction, valuable information-theoretic insights regarding the optimum design of CD-NOMA multiplexing schemes are extracted.

Albeit different, both problems lead interestingly to common frame design principles based on incoherence, representational tightness and unit-normality frame-theoretic attributes. We highlight on the one hand that these properties are necessary in capturing the salient information present in sparse signals, by linking redundant representations to sufficient sparse recovery guarantees. It is shown therefore that these structures are generally performing very well in reconstructing sparse signals under the modern framework of CS signal processing. On the other hand, in NOMA fashion, it is shown that the same properties generate minimally interfering compressions of large amounts of user information spread across all available communication resources, where inherently the latter are considered to be less than the amount of served users. The CD-NOMA scheme thus designed is baptized as massively concurrent NOMA (MC-NOMA), given its dense and high system overloading potential. We prove moreover, that the MC-NOMA scheme is not only helping the receivers by minimizing interference, but also achieves optimum ergodic system sum-rates in terms of link capacity, outperforming the state of the art.

To provide a full system description, we complement these designs with newly proposed modern signal processing for the sparse recovery of the mmWave channel, or respectively, for the joint multi-user detection (MUD) on the receiving side of a MC-NOMA transceiver. Numerical experiments and software simulations confirm the general theoretic results and highlight, once more, the superiority of the proposed frame-theoretic designs against the existent art.

Acknowledgements

I would like to express my sincere gratitude to my advisor, Prof. Dr. Giuseppe Abreu, who has become a tremendous mentor to me over the course of the last five years we have been working together. I would like to thank him in particular for allowing me to grow as a researcher and for providing guidance and advice irrespective of the time of the day or his adjacent activities. He has been a patient academic support constant over the course of my work and his feedback has always been of incredible value to me, both technically, but also personally, leading to my development beyond research as well.

I would also like to thank the committee members, Prof. Dr.-Ing. Werner Henkel and Prof. Dr. George Alexandropoulos for agreeing to take part in the dissertation committee.

Additionally, I would like to extend my thanks to Prof. Koji Ishibashi, Hiroki Iimori, Dr. Stefano Severi and Takanori Hara with whom I have successfully collaborated and whose advice and feedback were always constructive leading to better research ideas from my side and improving always my research output. Their help was invaluable to me, and I can only hope that I have been able in exchange to return their contributions through my work and my comments during our joint research meetings.

Furthermore, I would like to express my special thanks to my parents, Flori and Cristi, who always supported me and provided me with all the necessary advice, but also with the freedom to go out on my own in the world and experience life as it is, with good and bad. They always were my biggest supporters, having sacrificed their happiness for mine and offering me the necessary balance, mental strength and mentality to always challenge myself and strive to persevere in anything I do. *Mamă, tată, vă mulțumesc!*

More good thoughts and appreciation are directed towards my grandparents, Stelian, Dumitru and Elena, who have raised me from a young age with good moral values that have steered me into becoming who I am today and will help me develop further in the future. On the same note, I extend my gratitude and admit my luck to having crossed paths along the years with passionate teachers and professors who have sparked my learning curiosity, inspired me and helped me as a pupil/student.

I would also like to thank my host parents, Inge and Wilfried, who have been an outstanding support. They have taught me how to better balance my own inflicted pressure and have given me the strength to move past difficult situations I have encountered over the last eight years since I met them.

Many feelings of appreciation go towards my friends who have always provided me with hope

and good spirits, energizing me: Ciprian, Cristian, Gică, Alex, Dominic, Mihai and Andrei. Similar gratitude is extended to my work friends Johnny, Ivan and Arne who have supported me during the last months and understood my overwhelming working schedule.

Last, but not at all least, I would like to express my deepest gratitude and humble appreciation for two special persons in my life: Leonie and George. They have always been the balance I needed over the course of this doctorate, helping me to release my steam and making me understand some of the mistakes I was doing along the way. *I can only say I am very lucky to have you both in my life!*

*To my late grandfathers,
Dumitru and Stelian.*

Contents

Statutory Declaration	i
Abstract	iii
Acknowledgements	v
List of Figures	ix
List of Tables	xi
List of Algorithms	xi
List of Acronyms	xi
Notation	xviii
1 Introduction	1
1.1 A Condensed History of Wireless Cellular Systems	2
1.2 Future Directions of Wireless Communications	5
1.3 Thesis Outline	6
1.4 Thesis Contributions	7
2 Intrinsic Orthogonality of Wireless Systems	9
2.1 Orthogonal Access	10
2.2 Orthogonal Frequency Division Multiplexing	13
2.3 A Simple Case Study: IEEE 802.11p	19
2.3.1 Channel Estimation via Self-Organizing Frequencies	23
2.3.2 Performance and Access Limitations	31
2.4 Conclusions	37
3 Non-orthogonality and Frame Theory	39
3.1 Preliminaries	39
3.2 Frame Theory Basics	42
3.3 Frame Transforms and Operators	46
3.4 Frame Properties and Associated Quantifiers	50
3.5 Special Frame Types and Their Designs	57
3.6 Frame Applications to Wireless Communications	63

3.7	Conclusions	66
4	Millimeter Wave Initial Access	69
4.1	System Model and Considerations	70
4.2	Channel Estimation as a Sparse Recovery Problem	73
4.3	Beamforming Management via Frame Theory	76
4.3.1	Sparse Recovery Guarantees and Requirements	76
4.3.2	A Parseval Tight Frame Approach	80
4.3.3	Optimally Incoherent Unit-norm Tight Frames	82
4.3.4	Kronecker-decomposable Practical Realizations	88
4.4	Beamforming Performance Evaluation	90
4.4.1	Beamforming Design Differences	90
4.4.2	Numerical Results and Comparisons	94
4.5	Enhanced Sparse Solver for Channel Estimation	100
4.6	Conclusions	105
5	Non-orthogonal Multiple Access Framework via Frame Theory	107
5.1	Core Idea and Preliminaries	108
5.2	A Classification of Existing Solutions and Practical Limitations	111
5.3	Frame-theoretic Representation of Non-orthogonal Access	115
5.4	Massively Concurrent Non-orthogonal Multiple Access Transmission	119
5.4.1	Capacity Analysis and Performance Guarantees	120
5.4.2	Transmitter Design via Incoherent Unit-norm Tight Frames	126
5.4.3	Achievable Sum-rate Performance	135
5.5	Massively Concurrent Non-orthogonal Multiple Access Detection	140
5.5.1	Maximum Likelihood Low-complexity Tree Search Decoding	140
5.5.2	Sparsely-structured Massive Scale Joint Detector	158
5.6	Conclusions	168
6	The 6G Horizon – Sense the Waveforms	171
6.1	Trends and Future Directions	171
6.2	Candidate Technologies and Expectations	172
6.3	Concluding Remarks	176
7	Epilogue	177
7.1	Main Conclusions	177
7.2	Outlook and Extensions	180
A	Normed, Inner Product and Hilbert Spaces	181
B	Own Publications	185
	Bibliography	187

List of Figures

1.1	Evolutionary pattern applicable to wireless communications.	2
1.2	From 1G to 5G and beyond – a historical snapshot.	4
2.1	TDMA orthogonal wireless channel access example	10
2.2	FDMA orthogonal wireless channel access example	11
2.3	DS-CDMA wireless channel access example	13
2.4	OFDM spectral packing and utilization example	17
2.5	Time-varying V2X channel impulse response realization.	20
2.6	IEEE 802.11p Tx – Rx system diagram.	21
2.7	Block diagram of the proposed SOF receiver	24
2.8	Adaptive MA filtering over real V2V channel at 3 dB SNR.	27
2.9	Adaptive MA filtering over real V2V channel at 18 dB SNR.	27
2.10	Adaptive MA filtering over real V2V channel at 30 dB SNR.	28
2.11	The SOF update rule for any fixed subcarrier k^*	29
2.12	Comparative BER, FER & MSE results of V2V _{EWO} at 300B payload.	32
2.13	Comparative BER, FER & MSE results of V2I _{UC} at 300B payload.	33
2.14	Comparative FER results of V2V _{EWO} and V2I _{UC} at 100B, 300B, 800B payloads.	34
2.15	Comparative FER results of V2V _{EWO} , V2I _{UC} and V2I _{SS} for different speeds.	35
2.16	Time-frequency analysis of error sources.	36
3.1	UNTF Mercedes-Benz and its UNF perturbed variant.	54
3.2	Redundancy function exemplary comparison.	54
3.3	Overview of frame types – equi-normality & tightness	62
3.4	Overview of frame types – equi-normality & incoherence	62
3.5	Noisy signal RMSE recovery performance under different redundant representations.	65
4.1	Simplified protocol diagram of initial access in mmWave systems.	72
4.2	Incoherent frame construction under geometric search space constraint.	86
4.3	Coherence PDF of joint beamforming measurement matrices	91
4.4	Coherence PDF of joint sensing matrices	93
4.5	NMSE performance of mmWave channel OMP/BPDN sparse recovery for three different training beamforming schemes – $M_T = M_R = 4$	96
4.6	NMSE performance of mmWave channel BPDN ℓ_1 -reweighted sparse recovery across different training beamforming schemes – $M_T = M_R = 5$	97

4.7	NMSE performance of mmWave channel BPDN ℓ_1 -reweighted sparse recovery across different training beamforming schemes – $M_T = M_R = 6$	98
4.8	NMSE performance of mmWave channel BPDN ℓ_1 -reweighted sparse recovery across different training beamforming schemes – $M_T = M_R = 7$	99
4.9	NMSE performance comparison: BPDN ℓ_1 -reweighted vs FPG ℓ_0 -approximated.	103
4.10	NMSE performance comparison for randomized and optimally incoherent-tight beamforming: BPDN ℓ_1 -reweighted vs FPG ℓ_0 -approximated	104
5.1	NOMA vs. OMA example: 2 users DL rate comparison	109
5.2	An illustration of the UL PDMA active user non-orthogonal multiplexing	114
5.3	An illustration of the SCMA active user non-orthogonal multiplexing	114
5.4	An illustration of a DL NOMA system realization	118
5.5	An example of an UL NOMA system	119
5.6	CSIDCO vs. QCSIDCO coherence PDF comparison for 8×14 frame	131
5.7	CSIDCO vs. QCSIDCO coherence PDF comparison for 20×32 frame	132
5.8	Normalized coherence intensity map of \mathbf{F} MC-NOMA frame	133
5.9	Normalized coherence intensity map of a DL MC-NOMA precoded channel	133
5.10	Normalized coherence intensity map of an UL MC-NOMA precoded channel	134
5.11	Theoretic spectrally normalized CCMC ergodic sum-rate for CD-NOMA DL systems	135
5.12	Theoretic spectrally normalized CCMC ergodic sum-rate for CD-NOMA UL systems	136
5.13	Spectrally normalized DCMC ergodic sum-rate for CD-NOMA DL systems	137
5.14	Spectrally normalized DCMC ergodic sum-rate for CD-NOMA UL systems	138
5.15	Uncoded QPSK DL optimum BER performance comparison of MC-NOMA, SCMA and PDMA.	139
5.16	Illustration of SE ordering and search optimization for a $(2, 3)$ bipolar overcomplete access system	145
5.17	Probabilistic tree pruning by δ -thresholding	150
5.18	BER performance of MC-NOMA $(M, K) = (4, 6)$ system with 4-QAM and 16-QAM inputs	152
5.19	Complexity of optimum joint detection for MC-NOMA systems with 4-QAM at SNR = 15dB, with $M = 4$ and $K = \{5, 6, \dots, 12\}$	154
5.20	BER performance for MC-NOMA systems with 4-QAM at SNR = 15dB, with $M = 4$ and $K = \{5, 6, \dots, 12\}$	155
5.21	BER performance of MC-NOMA for systems of size $(M, K) = \{(4, 6), (4, 8), (4, 12)\}$ over 4-QAM discrete constellations	156
5.22	Complexity of optimum joint detection for MC-NOMA systems with 4-QAM at SNR = 15dB of sizes $(M, M + 2)$ and $M = \{4, 5, \dots, 16\}$	157
5.23	Smooth and differentiable α -parameterized approximation of ℓ_0 -norm	161
5.24	MC-NOMA detection for $K = 48$ users over $M = 32$ resources	165
5.25	MC-NOMA detection for $K = 120$ users over $M = 80$ resources	166

5.26 Scalable MC-NOMA detection performance for systems with overloading factor of 150% and $M = \{32, 48, 64, 80\}$ available resources	167
--	-----

List of Tables

2.1 IEEE 802.11a/p communication parameters.	22
2.2 Update rule correction threshold for IEEE 802.11p constellations.	30
3.1 Mathematical summary of equi-normal & tight frame properties	63

List of Algorithms

1 Self-organizing frequencies receiver.	30
2 Alternating projections UNTF design.	61
3 Incoherent UNTF frame design	88
4 FPG mmWave channel estimator	102
5 Rubik-QR decomposition	148
6 Rubik probabilistic tree search with thresholding multi-user detector	151
7 Sparsely-structured ℓ_0 -FPG massive MUD via ADMM	164

List of Acronyms

1G first generation. ix, 2–4, 10
2G second generation. 2–4, 10
3G third generation. 3, 4, 69, 110
4G fourth generation. 3, 4, 9, 17, 69, 110, 172, 174
5G fifth generation. ix, 3–6, 9, 17, 70, 108, 171–176, 185
5G+ fifth generation (5G) and beyond. 6, 17, 70, 107, 171, 172, 176
6G sixth generation. 6–8, 171, 172, 174–176, 179

ADC analog-to-digital converter. 71

ADMM alternating direction method of multipliers. xi, 8, 162–166, 168, 169, 179

AI artificial intelligence. 4–6, 172–176, 179

AMP approximate message passing. 94

AoA angles of arrival. 73, 75, 76, 82, 83, 92, 96

AoD angles of departure. 73, 75, 76, 82, 83, 92, 96

AR augmented reality. 69, 117

AWGN additive white Gaussian noise. 12, 18, 19, 30, 40, 64, 73, 76, 96, 117, 118, 120, 149, 179

BER bit-error rate. x, 31, 34, 36, 37, 139, 152, 153, 155, 156, 165–167

BPDN basis pursuit denoising. ix, x, 74, 94–100, 103–105, 167, 178

BS base station. 11, 72, 108, 110–112, 117–119, 125, 150, 173, 174

C-ITS Cooperative Intelligent Transportation Systems. 37

CCMC continuous-input-continuous-output memoryless channel. x, 120, 122, 124, 126, 135–139, 178, 179

CD-NOMA code-domain non-orthogonal multiple access (NOMA). iii, x, 111–113, 115–117, 119, 121, 124, 126, 127, 132, 134–136, 139, 141, 168, 178, 179

CDMA code division multiple access. ix, 2, 3, 6, 12, 13, 60, 108, 110, 113, 115, 127, 140, 180

CDP Constructed Data Pilot. 31

CFR channel frequency response. 19, 22–26, 29–31, 36

CNR channel-to-noise ratio. 96, 110

COTS commercial off-the-shelf. 37

CP cyclic prefix. 16, 21

CS compressed sensing. iii, 7, 38, 66, 67, 74–77, 79, 80, 92, 94, 95, 97, 105, 107, 177, 178

CSI channel state information. 6–9, 25, 26, 29, 30, 34–36, 70–74, 102, 104, 111, 112, 120, 140, 175, 186

CSIDCO complex sequential iterative decorrelation via convex optimization (SIDCO). x, 87, 88, 92, 126–132, 158, 178, 189

CSIR channel state information (CSI) at the receiver. 140, 180

CSIT CSI at the transmitter. 6, 7, 120–122, 125, 126

DAC digital-to-analog converter. 71

DCMC discrete-input-continuous-output memoryless channel. x, 120, 122, 124, 126, 137–139, 178, 179

DFS depth-first search. 143, 144, 146, 153, 179

DFT Discrete Fourier Transform. 15, 18, 19, 83, 84, 187

DL downlink. x, 7, 8, 72, 108–111, 114, 117–127, 132–140, 150, 152, 153, 165, 168, 172, 173, 178, 179

DoF degrees-of-freedom. 122, 127, 140

DS direct-sequence. ix, 12, 13, 60

DSRC Dedicated Short Range Communication. 6, 9, 19, 22

EM expectation-maximization. 37

eMBB enhanced mobile broadband. 3, 70

ESE elementary signal estimator. 113, 141

ETF equiangular tight frames. 56, 57, 64, 85, 127, 132

FD full-duplex. 174, 175

FDMA frequency division multiple access. ix, 2, 3, 11, 13

FEC forward error correction code. 22, 24

FER frame error rate. 34

FFT fast-fourier transform. 15–17, 187

FISTA fast iterative soft-thresholding. 94

FP frame potential. 50–52, 57, 92, 127

FPG fractional programming. x, xi, 8, 100–105, 158, 161, 162, 164–169, 178, 179

FT Fourier Transform. 15, 23, 40, 187

Gbps Gigabit-per-second. 70, 71

GSD generalized sphere decoding. 142–144, 146–150, 152–154, 156, 157

i.i.d independent, identically distributed. 78, 95, 152

IA initial access. iii, 6, 7, 67, 71, 72, 76, 80, 88, 94, 105, 178

ICI intercarrier interference. 16–19, 22, 23

IDFT inverse Discrete Fourier Transform (DFT). 15

IDMA interleave division multiple access. 112, 113

IFFT inverse fast-fourier transform (FFT). 15, 16, 21

IFT inverse Fourier Transform (FT). 15

IGMA interleave-grid multiple access. 113

IIoT industrial Internet of Things (IoT). 175

IoT Internet of Things. 1, 3, 9, 69, 113, 175, 187

ISI Intersymbol Interference. 16, 22, 23

ITS Intelligent Transportation System. 19, 22

KD Kronecker-distributed. 89, 90, 92, 93, 97–100, 105

LAN local area network. 3, 17, 19

LDS low-density spreading. 113, 115, 132, 134

LLL Lenstra-Lenstra-Lovász. 144, 146, 150

LLR log-likelihood ratio. 113

LS least squares. 22–24, 31, 34

LTE long-term evolution. 3, 4, 9, 110, 172

MA moving average. ix, 25–28, 31, 110

MAP maximum a posteriori. 141

MC-NOMA massively concurrent NOMA. iii, x, xi, 7, 8, 122, 124–127, 131–141, 143, 144, 146, 149, 150, 152–161, 165–169, 175, 179, 180

MCS modulation coding scheme. 18, 21, 30

MIMO multiple-input multiple-output. 4, 5, 8, 37, 70–73, 91, 105, 121, 124, 141, 142, 147, 165, 172, 175

ML maximum likelihood. 7, 8, 22, 120, 122, 125, 139–141, 152, 153, 155, 156, 158–160, 167–169, 178, 179

MMSE minimum mean squared error. 112, 113, 165, 166, 188

MMSE-PIC minimum mean squared error (MMSE)-parallel interference cancellation (PIC). 113

MMSE-SIC MMSE-successive interference cancellation (SIC). 113

mMTC massive machine-type communications. 3, 117, 120, 175

mmWave millimeter wave. iii, ix–xi, 4–8, 67, 69–75, 77, 80, 82, 88, 91, 94–100, 102, 104, 105, 107, 158, 161, 172, 174, 175, 177, 178, 180

MPA message passing algorithm. 113, 141

MSE mean-square error. 65, 96, 189, 190

MUD multi-user detection. iii, xi, 112–114, 116, 120, 122, 125–127, 139–141, 143, 146, 150, 153, 158, 164–166, 168, 179

MUI mutual interference. 122, 126, 132

MUSA multi-user shared access. 113, 121, 122, 127

MUST multi-user simultaneous transmission. 110, 172

NCMA non-orthogonal coded multiple access. 113

NMSE normalized-mean-square error (MSE). ix, x, 96–100, 103–105

NOCA non-orthogonal coded access. 113

NOMA non-orthogonal multiple access. iii, x, 5–8, 67, 107–113, 115, 118, 119, 122, 123, 125, 126, 168, 174, 177–179, 186, 188

NR new radio. 171

OFDM orthogonal frequency division multiplexing. ix, 4, 6, 9, 13–24, 29–31, 34–37, 40, 66, 108, 116

OFDMA orthogonal frequency division multiple access. 3, 6

OFHSS orthogonal frequency hopping spread-spectrum. 13

OMA orthogonal multiple access. x, 109–111

OMP orthogonal matching pursuit. ix, 74, 94, 96, 97, 100

OTFS orthogonal time frequency space. 66

PA power amplifier. 71

PAPR peek-to-average-power-ratio. 17

PD-NOMA power-domain NOMA. 111, 112, 168

PDF probability density function. ix, x, 91–93, 123, 131, 132

PDMA pattern division multiple access. x, 113–115, 121, 122, 127, 132, 134–140, 179

PHY physical layer. 4–6, 9, 20, 22, 23, 31, 37, 174, 175, 179

PIC parallel interference cancellation. 112, 113, 188

PN pseudo-noise. 12, 60, 80, 112

PTF Parseval tight frame. 80–82, 88, 90, 92, 96–99, 105

QAM quadrature amplitude modulation. x, 143, 144, 150, 152, 156, 159, 166

QCSIDCO quadratic complex SIDCO (CSIDCO). x, 128–132, 158, 168, 178

QoS Quality of Service. 3, 5, 11, 71, 111, 117

QR QR. xi, 141, 147–149, 151, 179

QT quadratic transform. 101–103, 161, 162, 167, 178, 179

RE resource element. 6, 113–117, 121, 122, 124, 125, 137, 139, 156, 158, 168, 178

RF radio frequency. 4, 5, 17, 37, 70, 71, 73, 74, 172, 174, 175, 180

RIC restricted isometric constant. 78, 79

RIP restricted isometry property. 77–80, 84, 94

RMSE root-MSE. ix, 65

RSMA resource spread multiple access. 112, 127, 132

Rubik-PTST rubik probabilistic tree search with thresholding. xi, 150–158, 160, 168, 169, 179

Rx receiver. 7–9, 15, 16, 18–21, 39, 43, 70–74, 76, 82, 87–96, 103–105, 107, 113, 114, 124, 132, 140, 149, 172, 177, 178, 180

SAGE space-alternating generalized expectation-maximization. 36

SAMCDP SNR-assisted modified constructed data pilot. 31

SCMA sparse-coded multiple access. x, 113–115, 121, 122, 132, 134–139, 168, 179

SD sphere decoding. 141–144, 146, 149–151, 158, 179

SDR software defined radio. 175

SE Schnorr-Euchner. x, 144–146, 150, 151, 179

SESD Schnorr-Euchner sphere decoding. 144, 150, 152–157

SIC successive interference cancellation. 108–113, 188

SIDCO sequential iterative decorrelation via convex optimization. 86, 87, 186

SNR signal-to-noise ratio. ix, x, 19, 25, 26, 28–31, 34–36, 64, 65, 96, 103, 105, 112, 124, 138, 139, 143, 146, 147, 155, 156, 167, 178

SOAV sum of absolute value. 165–168

SOF self-organizing frequencies. ix, 24–26, 28–31, 34–37

SSC sum of squared correlations. 127

STA spectral-time averaging. 31

SVD singular value decomposition. 58, 60

TDL tapped-delay line. 19

TDMA time division multiple access. ix, 2, 3, 6, 10, 11, 37, 110

THz terahertz. 70

Tx transmitter. 7–9, 15, 18–22, 30, 70–74, 76, 82, 87–96, 103–105, 107, 108, 116, 117, 124, 132, 168, 172, 177, 178, 180

UE user equipment. 72, 118, 119, 150

UL uplink. x, 7, 8, 72, 110, 113–115, 117–120, 122, 125, 126, 132, 134–138, 140, 150, 152, 168, 172, 173, 175, 178, 179

ULA uniform linear array. 82, 83, 95, 96

UNF unit-norm frame. ix, 54–58, 83, 85, 87–89, 126

UNTF unit-norm tight frame. ix, xi, 53–55, 58, 59, 61, 64, 65, 79, 81–85, 87–90, 92, 121, 126, 131–135, 138, 152, 168, 177–179

URLLC ultra reliable low-latency communications. 3, 117, 120, 158, 175

V2I vehicle-to-infrastructure. 31

V2V vehicle-to-vehicle. ix, 26, 31

V2X vehicle-to-everything. ix, 6, 9, 19–23, 34, 35, 37, 117

VR virtual reality. 69, 117

- WAVE** wireless access in vehicular environments. 9
- WB** Welch bound. 56, 57, 85, 91–93, 126, 131, 132, 158
- WSMA** Welch-bound spreading multiple access. 127
- WSSUS** wide-sense stationary uncorrelated scattering. 19, 23

Notation

Unless otherwise specified the vector notation used in the sequel assumes column orientation.

\mathbb{R}	domain of real numbers
\mathbb{C}	domain of complex numbers
\mathbb{H}	generic Hilbert space associated domain
\mathbf{X}	generic matrix notation
\mathbf{x}	generic vector notation
$(\cdot)^H$	conjugate Hermitian operation on a vector/matrix
$(\cdot)^T$	transposition operation on a vector/matrix
$(\cdot)^*$	complex conjugate operation applied entry-wise
$\mathbf{1}_M$	vector of all 1s of length M
$\mathbf{0}_M$	vector of all 0s of length M
\mathbf{I}_M	$M \times M$ identity matrix
$\mathbf{0}_{M \times N}$	$M \times N$ all 0s matrix
\mathbf{X}^{-1}	inverse of the square matrix \mathbf{X} (if it exists)
\mathbf{X}^\dagger	Moore-Penrose left or right (context will clarify) pseudo-inverse of \mathbf{X}
$\text{diag}(\mathbf{x})$	diagonal matrix with \mathbf{x} on the main diagonal
$\text{diag}(\mathbf{X})$	column vector extracted from the main diagonal of matrix \mathbf{X}
$\text{trace}(\mathbf{X})$	trace of matrix \mathbf{X}
$\text{rank}(\mathbf{X})$	rank of matrix \mathbf{X}
$\det(\mathbf{X})$	determinant of square matrix \mathbf{X}
$\ker(\mathbf{X})$	null space of the linear operator \mathbf{X}
$\text{spark}(\mathbf{X})$	spark of matrix \mathbf{X}
$\text{vec}(\mathbf{X})$	vectorization operation of the matrix \mathbf{X} , <i>i.e.</i> $\text{vec}(\mathbf{X}) \triangleq [\mathbf{X}_1; \mathbf{X}_2; \dots; \mathbf{X}_M]$
$\text{vec}^{-1}(\mathbf{x})$	reciprocal operation of vectorization, reshaping \mathbf{x} back to \mathbf{X}
$\ \mathbf{x}\ _0$	ℓ_0 -norm of vector \mathbf{x}
$\ \mathbf{x}\ _1$	ℓ_1 -norm of vector \mathbf{x}
$\ \mathbf{x}\ _2$	ℓ_2 -norm of vector \mathbf{x}
$\ \mathbf{x}\ _\infty$	ℓ_∞ -norm of vector \mathbf{x}
$\ \mathbf{X}\ _2$	spectral norm of the matrix \mathbf{X}
$\ \mathbf{X}\ _F$	Frobenius norm of the matrix \mathbf{X}
\otimes	Kronecker product

- Hadamard entry-wise product
- ⊗ circular convolution linear operator
- $\mathbb{E}_{\mathbf{x}}[\cdot]$ expectation operator consider along the direction of the random variable \mathbf{x}
- $\Re[\cdot]$ real part entry-wise extractor of complex scalars, vectors and matrices
- $\Im[\cdot]$ imaginary part entry-wise extractor of complex scalars, vectors and matrices

(page intentionally left blank)

Chapter 1

Introduction

A retrospective look into the last three decades of communications outlines by far a clear, steady and increasingly strong interest in wireless technologies as enablers of mobile telephony, private/public networks, data exchange, and respectively, their derivative applications. The evolution of wireless communications systems during the course of this timespan has been undoubtedly astonishing by any means. To swiftly point out the latter, consider the bulky mobile phones of the mid-late 1980's accessible only to the executives, managers and high-end traders in comparison to the compact, multi-functional smartphones of today available to the “average Joe”, which are capable of packing more computing capacity than convenience desktop computers of early and mid 2000's.

Hence, on the one hand, the evolution of wireless communications was greatly influenced by the advances in computing capacity and hardware developments alike. On the other hand, the nowadays widespread use of both wireless and wired communications has also been motivated by the socio-economic ever-increasing requirements of access and communications rates. These latter collective needs of *better, faster, cheaper* connectivity came in fact as a by-product of the Internet and its *almost* limitless over-the-top applications. To consider a few, think about some of the already mundane ones such as emailing, messaging or documents sharing used daily throughout the world at a large scale. Nonetheless, albeit these are used frequently, they cumulate in today's world only a very tiny fraction of the traffic passed through wireless systems. The late developments of e-commerce, social networks, streaming services and the proliferation of Internet of Things (IoT) smart devices have taken over by a large margin the quota of data delivered by communication networks in the daily life of individuals. To top this up, these applications have their industrial counterparts operating under private corporate networks in factories and remote environments.

Given the above, the evolution pattern of wireless technologies and their subsequent applications may be portrayed in simple terms as follows. The wireless communications technological developments are enablers of countless new applications – some of which were priorly envisioned by the design of the wireless access and protocols, and some of which were not. The

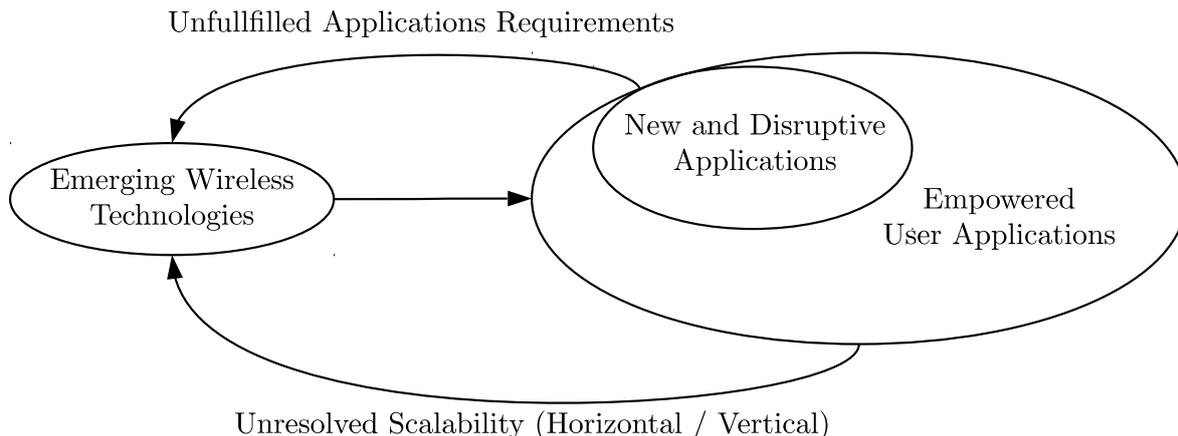


Figure 1.1: Evolutionary pattern applicable to wireless communications.

latter ones, the disruptive applications, are the ones that in turn catalyze and drive the evolution of next-generation wireless technologies. These events cycle periodically feeding the incremental evolution of wireless communications technologies from one iteration to the next. The above evolutionary pattern is outlined graphically by Figure 1.1.

The current dissertation aims therefore to provide humble design insights and concrete technology advances beyond the state of the art for the next iteration of wireless networks and communication systems.

1.1 A Condensed History of Wireless Cellular Systems

To better understand the requirements and technical features at the core of the previous and current iterations of wireless communications, let us take a short look over the evolution of mobile wireless systems. To this end, the first iteration of mobile telephony, coined as first generation (1G) [1,2], was in fact an analog cellular extension to cordless phones enabling voice services at low data rates. As a result, it was relying on *orthogonal* frequency division multiple access (FDMA) with circuit-switched control and resource allocation to provide service to its users. It was made obsolete by the increasing demand of access which depleted fast the resource capabilities of the 1G systems.

As a result, given the digital circuitry and miniaturization advances of the late 1980's, the second generation (2G) cellular based wireless networks were introduced as a replacement superseding the previous 1G counterparts. As fully digital cellular wireless systems, the 2G protocols expanded the FDMA resource allocation of its legacy and used as well resource elements over the time domain as *orthogonal* time division multiple access (TDMA) slots [3]. These changes enabled the 2G wireless systems to make a more efficient use of the wireless media by firstly optimizing the spectrum usage given the appropriate sampling of voice channels, and secondly, by optimizing the available time and frequency domain usage. Nonetheless, once again the popular widespread and rapid growth of mobile telephony services pushed the limits of 2G systems which started exploring an additional dimension, namely the code domain via code division

multiple access (CDMA) [1]. However, the latter was never fully leveraged by 2G networks which in the end coped and became tuned to properly address the voice service problem that the 1G systems failed to solve at large scale.

Motivated by the dawn of the Internet at scale, *i.e.* accessible to everyone, the third generation (3G) networks were designed to cope with both telephony and data services. Albeit the latter were possible over the 2G counterparts, they remained fixed to the level of short text messaging rather than upcoming electronic mail at the time. As a result, 3G networks were the first cellular wireless systems with layered services and resource allocation based on service policies tailored for both telephony and data. To this end, they incrementally incorporated and extended the CDMA, TDMA and FDMA *orthogonal* resource usage of legacy systems in offering communication rates capable of achieving up to several Mbps towards their latest releases.

The developments of wide-area networks focused around cellular communications systems were horizontally scaled by the wireless local area networks (LANs) based on the IEEE 802.11 family standard, and respectively, short-range counterparts, such as BluetoothTM for instance [4]. These combined advances enabled not only the migration of desktop Internet applications to the mobile phone ecosystem, but also heralded the emergence of pervasive mobile over-the-top applications, *e.g.* the *App Store*. Once more, the access and capacity limits of 3G systems were thus stretched, in this case, by the arrival of increasingly computationally capable mobile devices and proliferation of smartphones.

In the face of the latter disruptive emerging applications, the wireless communications industry and research community brought to life the fourth generation (4G) wireless networks as the first properly and purposefully built Internet based wireless systems capable of delivering a fully connected experience close to a desktop environment. From this perspective, 4G honed and bested the areas where 3G failed. In doing so, it used yet another *orthogonal* access scheme, *i.e.* orthogonal frequency division multiple access (OFDMA) [4–6], alongside and complementary to the legacy ones previously used. Furthermore improvements to 4G networks were integrated over the last years in the form of long-term evolution (LTE) addenda meant to boost the performance and scalability of original 4G systems. To this end, the latter succeeded in providing massive gains in capacity and end-to-end peak data rates bounded by 1 Gbps, capable of satisfying the needs of daily *human* users.

However, the 4G and its enhanced LTE equivalent were not designed in order to address hard requirements and traffic volumes that may surface given *things* – *e.g.* sensors, machines, cars, robots, drones – as network users. The latter, under the presumption of autonomy, may have in practice more stringent and hard Quality of Service (QoS) constraints that exceed the capacity of 4G in terms of bandwidth, connection density or end-to-end latency. In the light of the latter, and as a response to the disruptive unstoppable spread of such IoT cyber physical systems, the fifth generation (5G) wireless networks were proposed and researched over the course of the last 5 years. On such grounds, 5G categorized its connections into three different branches, namely: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC) and ultra reliable low-latency communications (URLLC). The three categories from above seek therefore to enforce the 5G IMT-2020 vision, [7], over the latest benchmark, *i.e.* the

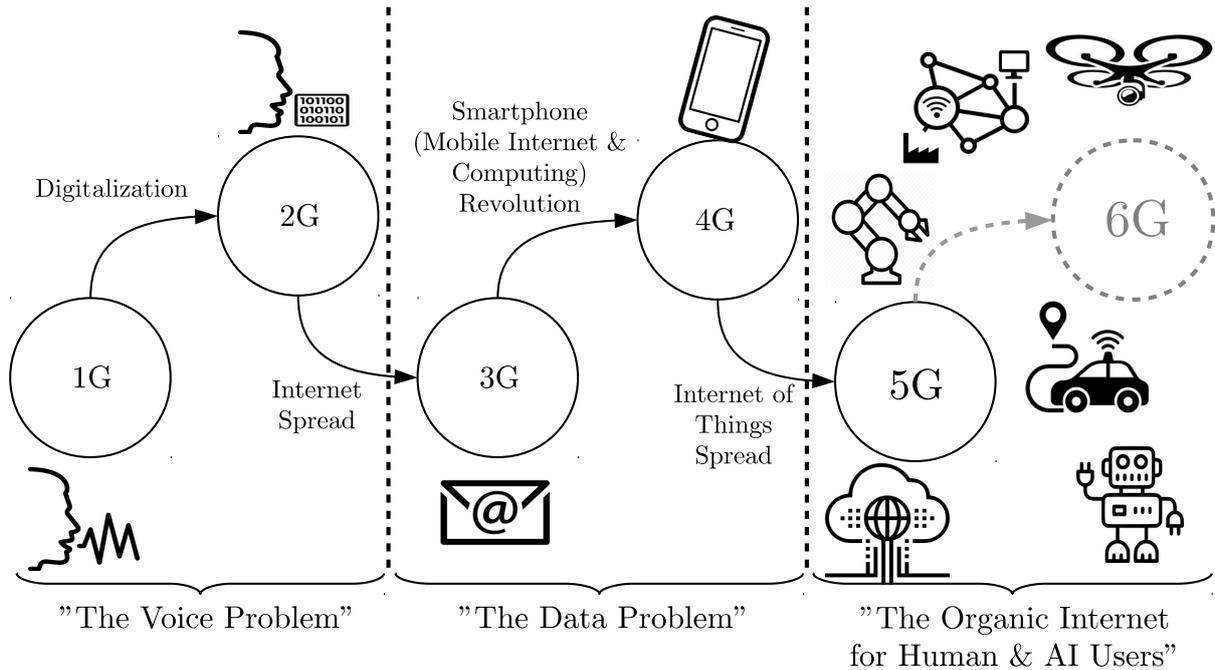


Figure 1.2: From 1G to 5G and beyond – a historical snapshot.

IMT-Advanced (4G-LTE) [8], across the planes of massive data rate, massive connectivity and reduced latency.

Given the latter categorization of serviced links and the network softwarization trends, 5G is expected therefore to provide the basis of a highly modular and scalable wireless communications platform capable of serving various applications needs. Albeit the high objectives, the 5G physical layer (PHY) technology is expected so far to rely heavily on established 4G *orthogonal* access and waveforms radiated over improved multiple-input multiple-output (MIMO) radio frequency (RF) links. As complements, networking virtualization, flexible PHY numerology of resource elements and the upcoming millimeter wave (mmWave) technologies are expected to aid 5G networks reach their ambitious goals according to the latest recently released standardization, Release-15 [9], and respectively, to its currently drafted successor, Release-16 [10], to be released in 2020 as a second phase to the realization of the IMT-2020 vision [7].

Following the past developments, it is important to remark the fact that each even iteration of a wireless communications system was in fact capable of acceptably solving its predecessor's main objectives as depicted in Figure 1.2. For instance, 2G solved the voice telephony that 1G aimed to provide a solution to by the disruptive means of digitalization. Similarly, 4G and LTE solved the data transmission conundrum that motivated 3G by using orthogonal frequency division multiplexing (OFDM) multicarrier waveforms, MIMO systems and better optimized, partly decentralized service policies via emergent network virtualization. Despite 5G's ambitious goals and capable tools to achieve them, it is to be expected that trully enabling a *virtually configurable and scalable* radio architecture would be superseded in the next 10-30 years by the rapid advances of machine learning and their digital beneficiaries, *the intelligent agents of tomorrow*, *i.e.* the artificial intelligence (AI) [11]. Under the premises of massively spread

digital intelligent entities, the 5G core network architecture may be capped by the sheer scale and heterogeneity of access, traffic and latency requirements.

Therefore, at the dawn of commercial 5G deployment it comes naturally to ask ourselves “*what are the next disruptive technologies that will enable 5G and beyond systems to reach the design goals of a globally scalable and optimized network for seamless Internet access to both humans and machines ?*”.

1.2 Future Directions of Wireless Communications

The “seamless Internet access” problem coined above refers to the features of high scalability and self-configuration that upcoming wireless communications systems need to fulfill in order to optimally satisfy the users’ needs. To this extent, the heterogeneity of users in the network will be a difficult problem to address on a global scale, requiring solutions not only on the upper network layers, but also at the physical layer, and respectively, at the application layer via inter-node cooperation. Concretely the “seamless Internet access” will become a distributed problem in future generations of wireless networks aiming to solve application-focused QoS requirements. On one hand, for the human users such needs may be summarized as “intuitive access” focusing on providing least amount of steps to get what one needs based on human like interactions, *e.g.* natural language processing, gesture recognition, facial expressions identification. On the other hand, for the AI users these needs would be reduced to distributed digital mitigation and agreement towards local and global optimizations satisfying the hard instantaneous requirements of the network nodes’ applications.

In essence, based on the above, the design of *context-aware cross-layer access and resource allocation* schemes would be crucial to satisfy such future scenarios. In the light of our PHY-focused research interests, two key technologies are identified as enablers of such future access schemes: mmWave and non-orthogonal multiple access (NOMA).

The mmWave systems are considered under the 5G umbrella as practical alternatives to evade the spectral clog in the sub-6 GHz spectra [12–14]. Given their high operating frequencies, mmWave candidate technologies operate in the frequency range of 24 – 300 GHz and are inherently broadband communication systems with high bandwidths. However these gains come at the cost of additional signal processing needed to radiate the waveforms over the highly attenuated mmWave operating frequencies. The solution to this problem has been the complementary utilization of MIMO architectures taking advantage of the short wavelengths of mmWave allowing for packing more antenna elements together at the level of the MIMO RF antenna arrays [14, 15].

Nonetheless, the high operating frequencies challenge the RF designs cost-effectiveness and efficiency of practical realizations of mmWave radios. To this extent, multiple feasible approaches to the RF design of mmWave systems were proposed ranging from fully digital to hybrid solutions. Recognizing the fact that hardware improvements over the next years may deem the RF design dilemma obsolete and enable fully digital designs, we identify the core problem of mmWave as being the *beamforming optimization and management*. To this end,

this is studied by this dissertation in the least advantageous context of no prior knowledge of channel state information (CSI) at the transmitter (CSIT) given the initial access (IA) scenarios of mmWave links.

The waveforms and PHY access has been insofar mainly based on orthogonal designs. In plain words, the maximum number of media resources used by the communication has always been bounded by the minimum number of physical resource element (RE) in the system, be it time, frequency, code or spatial dimensions. Historically as seen before, such systems were eventually capped by the ever-increasing demand for wireless access which led to the research and implementation of new access technologies, *e.g.* TDMA, CDMA, OFDMA [1, 2, 4]. Currently, 5G relies on the same orthogonal access and waveforms enhanced by configurable PHY numerology over the multicarrier signals transmitted [9]. At least with regards to 5G, this paradigm does not seem to change very soon even with the enhanced draft of the protocols, *i.e.*, Release 16. Despite the network virtualization and flexible waveform numerology, the 5G proposed access remains insofar based on the orthogonality concept.

It so follows that the expectation of 5G to be challenged to its limits by the increasing demand of traffic and access especially with the prevalence of AI is historically once more not far-fetched. Under these premises, we regard non-orthogonality and NOMA [16–18] as promising paradigms to enhance and truly scale the capabilities of 5G and beyond (5G+), and respectively, of future wireless communications systems, referred hereafter to as sixth generation (6G) networks. As a consequence, a second part of this thesis is dedicated to the study of feasible mathematically sound non-orthogonal transceiver and receiver designs under the context of NOMA as a scalable enabler of future wireless systems.

The mathematical framework under which the two main problems described above are studied is called Frame Theory [19–23]. Its focus is the study of non-orthogonal representations. Over the course of this dissertation, the connection between Frame Theory and the latter wireless communications topics is discussed and leveraged to propose two centralized frameworks, one for each problem, together with subsequent signal processing solutions.

1.3 Thesis Outline

The dissertation is structured as follows:

- *Chapter 2* discusses the fundamentals of conventional orthogonal access schemes in wireless communications. At a PHY-layer, the OFDM modulation is discussed in more depth to illustrate the orthogonality paradigm of current wireless systems waveforms. In addition, a practical study case is introduced and analyzed on the basis of Dedicated Short Range Communication (DSRC) [24] via the IEEE 802.11p protocol [25], currently in effect for the upcoming vehicle-to-everything (V2X) communications. The latter is used to illustrate some limitations of orthogonal designs related to access, achievable rate, and latency.
- *Chapter 3* provides a mathematical hiatus meant to provide a self-sustained reference to finite Frame Theory and related concepts of redundant representations. The core ideas

of non-orthogonality are introduced together with the formal definitions of Frame Theory and frames alike. In addition, the dual linear transforms of *analysis and synthesis* are discussed together with their effects of *expansion*, and respectively, *compression* of the equivalent signals. Frame properties, operators and special types of frames are covered as well. Lastly, a succinct perspective of immediate applications to wireless communications topics is provided at a high-level.

- *Chapter 4* treats the problem of mmWave IA channel acquisition and subsequent beamforming optimization and management. Firstly, the system model and main practical assumptions of the mmWave media and mmWave radios are discussed. Then the channel estimation problem associated with the first channel acquisition of a mmWave link is formulated according to the state-of-the-art as a sparse recovery problem. The joint problem of beamforming under no CSIT is next projected as a frame-theoretic optimization and design. This is firstly solved under theoretic assumptions and later discussed under its practical Kronecker-decomposable realization. Lastly, the circle is closed by the sparse recovery of the mmWave channel given sparsely-enhanced algorithms.
- *Chapter 5* researches the NOMA problem. At first, the core ideas of NOMA and a classification of existent solutions together with an overview of implementation limitations are presented. In the sequel, the linear system NOMA model is abstracted to a frame-theoretic representation for both uplink (UL) and downlink (DL). Under the abstract model, the design criteria of an optimal NOMA system via frames are discussed. The massively concurrent NOMA (MC-NOMA) transmitter (Tx) design solution to this problem is next presented together with its information-theoretic optimality and performance over the state of the art. The optimum receiver (Rx) design is then discussed and a low-complexity solution capable of maximum likelihood (ML) performance is proposed. Lastly, an alternative highly scalable and low-complexity joint detection scheme for massive MC-NOMA systems is proposed based on sparsely-structured detection.
- *Chapter 6* presents an exploratory perspective on the future trends of problems to be addressed by 6G networks. Candidate technologies based on the proposed schemes of this thesis are discussed together with expected outcomes of their prospective deployments.
- *Chapter 7* concludes the thesis and offers queues on future work points as interesting and worthwhile extensions of the research presented throughout the manuscript.

1.4 Thesis Contributions

The work performed during the course of this research has led to the following contributions listed in the sequel:

Topic: mmWave systems

- decomposed the compressed sensing (CS) mmWave CSI estimation problem, IA and subsequent beamforming to a frame-theoretic formulation;

- proposed a frame-based approach to optimization and management of training beamformers/combiners for optimum sensing of the sparse mmWave channel during the first stage of channel acquisition;
- studied the practical realization of optimum beamformers given their Kronecker factorization to realizable and achievable Tx/Rx counterparts;
- developed a new fractional programming (FPG) approximated ℓ_0 -norm sparse recovery algorithm as an alternative to state-of-the-art methods;
- performed synthetic evaluation of the proposed mmWave CSI estimation schemes and algorithms via PHY-link layer computer simulations.

Topic: NOMA systems

- proposed a generalized abstract framework for the treatment of UL/DL code-domain NOMA resource allocation and access schemes based on Frame Theory;
- formulated the code-domain linear precoding NOMA operation as a generalized frame design problem;
- introduced a new NOMA dense access scheme, *i.e.* MC-NOMA, based on an optimally designed frame capable of minimizing the associated system-inherent inter-user interference;
- proved the information-theoretic optimality of the scheme under practical conditions and outlined the gains over established state-of-the-art methods in both UL and DL scenarios;
- developed a low-complexity ML-capable tree search joint detector optimized by an algebraic search space reduction coupled with a hard probabilistic thresholding rule;
- developed a new formulation of MC-NOMA joint detection as an alternative probabilistic sparsely-structured ℓ_0 -norm problem widely-applicable to other MIMO systems alike;
- proposed a fast iterative sparsely-structured solver based on an ℓ_0 -approximated, FPG-relaxed formulation implemented via alternating direction method of multipliers (ADMM);
- performed extensive simulations to evaluate and outline the performance of the proposed transmission scheme and detection algorithms both from an information-theoretic perspective, but also from a joint detection perspective.

Topic: 6G and future wireless networks

- proposed an original vision towards the next generation of wireless networks as enablers of intelligent agents and their collaborative operation.

Moreover, large parts of the above contributions have been scientifically disseminated through publications, submissions for publication or presentations in specialized, peer-reviewed international IEEE journals, conferences and workshops on the three topics identified above. A complete list of the publications generated over the course of this PhD is given in Appendix B.

Chapter 2

Intrinsic Orthogonality of Wireless Systems

As detailed in the previous chapter, legacy and current wireless communications systems rely almost exclusively on orthogonal waveforms and access schemes. The core idea motivating such usages is quite simple and in fact practical for typical deployments prior to the IoT revolution. Concretely, it relies on the fact that orthogonally modulated information streams or access patterns are easily recoverable at the Rx side, to the extent that for linear systems, such as the wireless communications links, a matched filter could be applied to recover the radiated information. So, in plain words, the path of orthogonality for modulation and access over the wireless media has been conveniently chosen for an efficient and an economic information retrieval on the receiving end of the information.

To provide therefore a motivation and open the appetite to the non-orthogonal treaty of diverse PHY problems of future wireless networks across the planes of *CSI estimation*, *Tx design*, and respectively, *Rx design*, a brief discussion of conventional orthogonal concepts is firstly offered in this chapter.

The emphasis is placed furthermore on the OFDM waveforms and modulation which is largely employed nowadays across the communication architectures of 4G, LTE [5] and 5G [9], respectively. In the end of the chapter a specific case study of an orthogonal communication system is addressed based upon the IEEE 802.11p protocol [25]. This OFDM-based PHY communication stack is used for DSRC, or equivalently, wireless access in vehicular environments (WAVE) [24], by V2X communications across Europe [26–28] and North America [24, 29] alike.

This case study is thence used to showcase some of the practical challenges to be expected by upcoming networks across different harsh mobile environments and strict timing constraints,

Excerpts of this chapter are adapted and enhanced based on the selected article:

R.-A. Stoica, S. Severi and G. T. Freitas de Abreu, “A Self-Organizing Frequency Approach to 802.11p Channel Estimation,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1930-1942, July 2017. ©2017 IEEE.

and simultaneously, to outline some of the deficient performance characteristics that orthogonal communications systems may encounter. Last but not least, these remarks should therefore constitute a motivational step towards a non-orthogonal characterization of future wireless communication systems to follow in the next chapters.

2.1 Orthogonal Access

A simplified taxonomy of orthogonal media access in wireless systems may be easily formulated under the following four domains: *time*, *frequency*, *code*, and respectively, *space*. Concretely, each of these dimensions are orthogonalized via some sort of splitting or subspace separation allowing multiple users to access the media resources along a respective axis in an orthogonal manner ensuring therefore minimal interference relative to each other.

For instance the *time domain* orthogonal access, TDMA [4], is based on the very simple idea of channel splitting across the temporal dimension. This results in slotted time frames of exclusive access ensuring, upon distributed coordination policies, that each user is capable of using the wireless channel orthogonally to the other users with whom the media is being shared. The TDMA strategy dates back to the origins of 1G and 2G cellular systems [30] and relies on packetized access and contention resolution strategies [31] meant to solve potential packet collisions sourced by different users during the same time slots. As a consequence, the TDMA schemes rely heavily on Queuing Theory mechanisms and concepts [31,32] in order to mitigate channel contentions and assure temporal orthogonality. An example in this sense, is portrayed in Figure 2.1.

It is therefore clear that under excessive loads of the system, the queuing of users to be

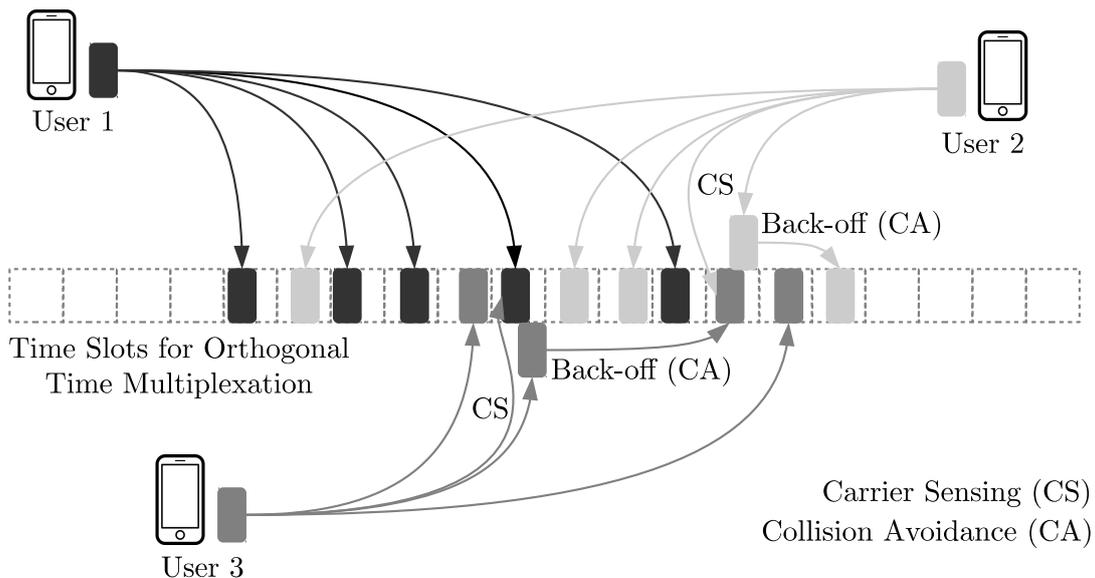


Figure 2.1: TDMA orthogonal wireless channel access example for 3 users. A simple carrier sensing and collision avoidance strategy is exemplified to outline the exclusive access to the wireless channel over a time slot.

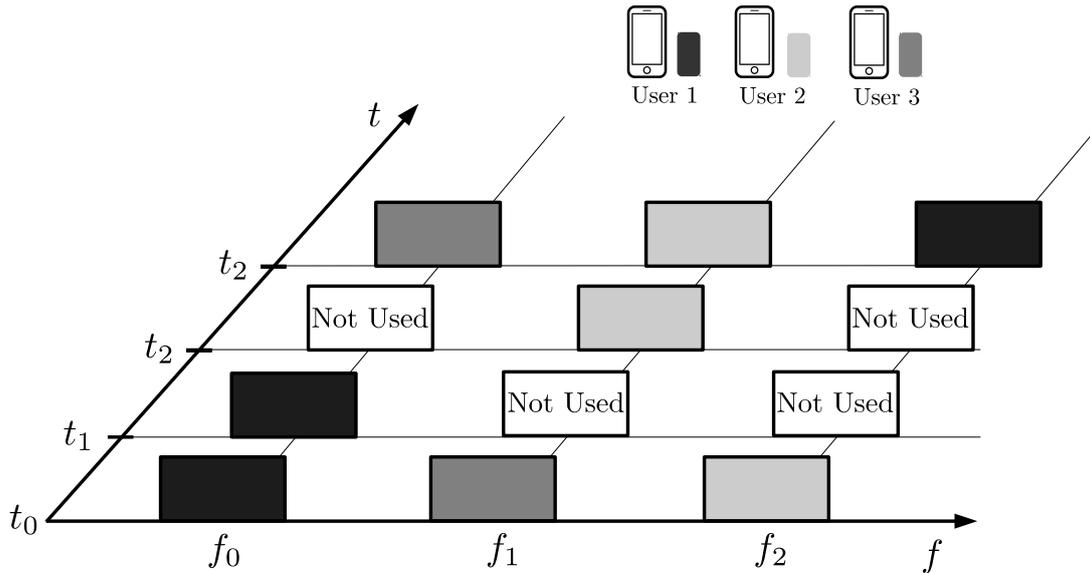


Figure 2.2: FDMA orthogonal wireless channel access example for 3 users given 3 available frequency resources over multiple transmissions instances.

served by a potential base station (BS) may increase catastrophically as the average packet serving time is lower than the average channel access rate of the user cohort. This leads to a complete bottleneck and total failure of the system in satisfying QoS requirements of incoming users which are postponed indefinitely once the capacity of the communication system has been reached.

The *frequency* based orthogonal access can be further split into single-carrier and multi-carrier strategies respectively. An old approach still widely used today to orthogonally split the frequency space of single-carrier wireless systems is represented by FDMA channel selection where users access the media over the total available bandwidth under different non-interfering subchannels. Under the classic Nyquist Sampling Theorem [33] it thus follows that these subchannels require a frequency separation of $2W$ for any signal of bandwidth W resulting in a maximum bandwidth efficiency of

$$\eta_{\text{FDMA}} \triangleq \frac{W}{2W} = \frac{1}{2}, \quad (2.1)$$

where the latter result holds by the same logic also for the case of the TDMA discussed above.

To this extent, an illustration of the single-carrier orthogonal splitting of the wireless channel and its subsequent access is displayed in Figure 2.2.

Based on the illustration from Figure 2.2 the depletion of resources under FDMA becomes apparent. To this end, consider the case where an additional user would like/need to use the wireless channel in the next transmission window, but none of the three previous users is freeing any of the available three frequency resources. This congestion leads to postponing or denying the service of the incoming user.

Of course that the simple deployment of individual orthogonal access schemes as singletons is naive and in practice the methods discussed in this section are in fact combined in order to

boost the user capacity of wireless systems. However, their individual treaty serve an illustrative purpose.

The *code domain* approach to multiple access over the wireless media, CDMA, relies on the core idea of *orthogonally spreading multiple users* information across the commonly used spectrum, hence it is also often named *spread-spectrum access* [4,34]. This concept is resumed to antipodal, *i.e* binary, orthogonal M -sequences multiplying each user's information at a higher rate than the symbol rate. Mathematically this concept resumes on the receiver end to a multi-user detection problem given by the received signal

$$r_{k,1} = \sum_{m=1}^M b_{k,1} \cdot E_{c,1} c_{m,1} + b_{k,2} \cdot E_{c,2} c_{m,2} + n_{m,1}, \quad (2.2)$$

where for simplicity of the exposition a 2-user scenario was considered and the notation used marks by $r_{k,1}$ the received signal over the additive white Gaussian noise (AWGN) channel at user 1, by $c_{m,\ell}$ the m -th chip of the spreading code of user ℓ , by $E_{c,\ell}$ the chip energy of the spreading code for user ℓ , by $b_{k,\ell}$ the k -binary information codeword of user ℓ , and respectively, by $n_{m,1}$ the AWGN of receiver 1 over the m -th chip interval.

Under the assumption of orthogonality of the sequences of different users, the optimal de-spreading strategy would resume to the application of matched filtering [34] at the receiver to obtain

$$y_{k,1} = \sum_{m=1}^M c_{m,1} \cdot (b_{k,1} \cdot E_{c,1} c_{m,1} + b_{k,2} \cdot E_{c,2} c_{m,2} + n_{m,1}) \quad (2.3a)$$

$$= \underbrace{b_{k,1} E_{c,1} \cdot \sum_{m=1}^M c_{m,1} c_{m,1}}_{\text{user 1 signal - desired}} + \underbrace{\sum_{m=1}^M c_{m,1} n_{m,1}}_{\text{filtered AWGN}} + \underbrace{b_{k,2} E_{c,2} \cdot \sum_{m=1}^M c_{m,1} c_{m,2}}_{\text{interference of 2 onto 1}} \quad (2.3b)$$

$$= b_{k,1} E_{c,1} M + \tilde{n}_{k,1} + b_{k,2} E_{c,2} \cdot \sum_{m=1}^M c_{m,1} c_{m,2}, \quad (2.3c)$$

where in the last step the filtered noise has been implicitly denoted as $\tilde{n}_{k,1}$.

Upon vectorization of the operations in Equation (2.3c), one may observe that the interference term is governed by the inner product between the spreading sequence of user 1, and the spreading sequence of user 2. It is thus clear that for orthogonal spreading such as conventional CDMA then this would cancel out as $\mathbf{c}_1 \perp \mathbf{c}_2$ and $|\langle \mathbf{c}_1, \mathbf{c}_2 \rangle| = 0$ respectively – where $\mathbf{c}_1, \mathbf{c}_2$ denote implicitly the stacked-up versions of the spreading codes of the 2 users considered.

In practice however, the generation of binary sequences with such attributes as envisioned by CDMA is constrained by their discreteness and dimensionality requirements of the applicable systems which increasingly challenges the design problem. As a result, a relaxed strategy is applied by the usage of pseudo-noise (PN) sequences [35] – artificially generated direct-sequence (DS) pseudo-random sequences with properties similar to white noise, high auto-correlation, and respectively, almost flat cross-correlations [34]. Consequently, it follows that $|\langle \mathbf{c}_1, \mathbf{c}_2 \rangle| \neq 0$, but small, and as a result, the interference term actually is complementing the receive noise of the

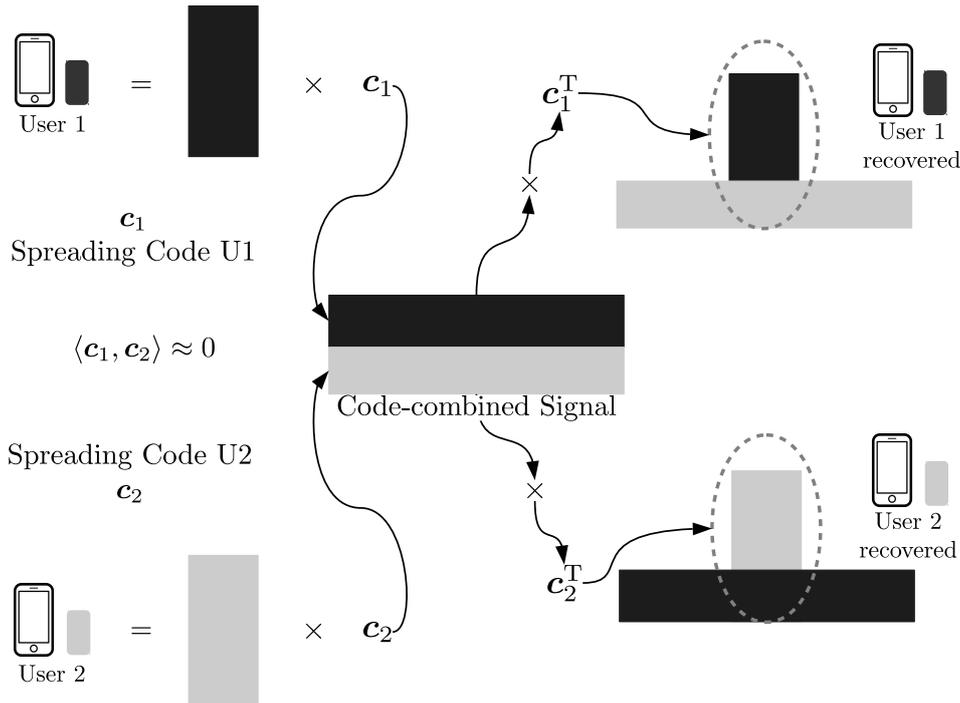


Figure 2.3: DS-CDMA orthogonal wireless channel access example for 2 users.

system, see for instance Figure 2.3. As an effect, practical CDMA systems suffer from the near-far problem since the matched filter receiver requires that the signal energies associated with each individual user component are actually closely the same. Power control is to this extent therefore a practical necessity of CDMA systems. To summarize in simple words, *there are no free gains*, and this can be seen for CDMA systems that provide a strong potential for high user capacity, but at the same time require strict power control to be able to achieve their theoretical gains [35].

An alternative application of CDMA robust against the near-far problem is represented by the application of the same spreading idea to the frequency domain subcarriers leading to orthogonal frequency hopping schemes [4, 34]. Concretely, similar DS *integer* spreading is applied to the different users in order to coordinate and spread orthogonally their individual temporal information over disjoint subcarriers leading to an orthogonal frequency hopping spread-spectrum (OFHSS) access [34]. However, the cost incurred by reinforcing robustness of CDMA by means of OFHSS is paid in terms of poor spectral efficiency. Intuitively, at any instance in time an orthogonally frequency-hopped channel allocation scheme over different users is equivalent in fact to any *orthogonal static* channel utilization by individual users over preassigned frequencies. As a result, this leads to the same bandwidth efficiency as for conventional single-carrier FDMA, *i.e.* following identity (2.1), $\eta_{\text{OFHSS}} \triangleq 0.5$.

2.2 Orthogonal Frequency Division Multiplexing

OFDM is a multi-carrier frequency division multiplexing modulation and access scheme

which relies on compacting information streams over orthogonal frequency subcarriers for enhanced spectral efficiency and throughput. In fact OFDM is inherently simple in terms of its core, in the sense that it relies on the orthogonality property between two sinusoids properly displaced in frequency domain. Concretely, consider two distinct symbols $s_k, s_\ell \in \mathbb{C}$ to be transmitted *in parallel* over the same temporal duration T_s with the corresponding passband forms

$$x_k(t) = s_k \cdot e^{j2\pi f_k t} \quad (2.4)$$

$$x_\ell(t) = s_\ell \cdot e^{j2\pi f_\ell t}, \quad (2.5)$$

for $t \in [0, T_s]$.

The two radiated signals may be compactly squeezed together under a multi-carrier transmission as long as the passband forms are interference-free in the frequency domain. As a result, it follows that the orthogonality condition

$$\langle x_k(t), x_\ell(t) \rangle \triangleq \frac{1}{T_s} \int_0^{T_s} x_k(t) x_\ell^*(t) dt = 0 \quad (2.6)$$

must be met for $t \in [0, T_s]$ and $k \neq \ell$ [5].

Let us denote for later convenience $f_k \triangleq f_0 + k\Delta f$ and similarly $f_\ell \triangleq f_0 + \ell\Delta f$. It follows by expanding the identity condition in (2.6) that

$$0 = \frac{s_k s_\ell^*}{T_s} \int_0^{T_s} e^{j2\pi f_0 t} e^{j2\pi k \Delta f t} \cdot e^{-j2\pi f_0 t} e^{-j2\pi \ell \Delta f t} dt \quad (2.7a)$$

$$= \frac{s_k s_\ell^*}{T_s} \int_0^{T_s} e^{j2\pi(k-\ell)\Delta f t} dt \quad (2.7b)$$

for $t \in [0, T_s]$ and $k \neq \ell$ iff

$$\Delta f = \frac{1}{T_s}, \quad (2.8)$$

or equivalently, in temporal domain T_s is long enough such that the identity in (2.8) is satisfied.

Remark furthermore that the inner product $\langle x_k(t), x_k(t) \rangle = \|s_k\|_2^2$ such that under normalized source symbols assumption the magnitude of each radiated symbol is unitary, leading to the fact that [5, 6]

$$\langle x_k(t), x_\ell(t) \rangle = \begin{cases} 1, & k = \ell \\ 0, & k \neq \ell \end{cases}. \quad (2.9)$$

Having outlined the basic principles of OFDM for the simple case of two parallel subcarrier streams, the generalized version of an OFDM modulated symbol follows naturally under the form [5]

$$x(t) = \sum_{k=0}^{N-1} s_k \cdot e^{j2\pi f_k t}, \quad \forall t \in [0, T_s] \quad (2.10)$$

where f_k is defined as before, *i.e.* $f_k = f_0 + k\Delta f$, and $\Delta f T_s = 1$ in the light of the orthogonality condition (2.9).

The demodulation procedure is moreover intuitive as well following the derivation so far. It is clear that by applying a matched correlation on the receiver side one can identify any symbol s_k multiplexed in $x(t)$. In other words, matched filtering with a filter bank corresponding to all subcarrier atoms $e^{j2\pi f_k t}$ shall demodulate the originally transmitted signals s_k . Mathematically, this resumes to the operations

$$\langle x(t), e^{j2\pi f_k t} \rangle = \frac{1}{T_s} \int_0^{T_s} \left(\sum_{\ell=0}^{N-1} s_\ell \cdot e^{j2\pi f_\ell t} \right) e^{-j2\pi f_k t} dt \quad (2.11a)$$

$$= \sum_{\ell=0}^{N-1} s_\ell \cdot \frac{1}{T_s} \int_0^{T_s} e^{j2\pi \ell \Delta f t} e^{-j2\pi k \Delta f t} dt \quad (2.11b)$$

$$= \sum_{\ell=0}^{N-1} s_\ell \cdot \langle e^{j2\pi \ell \Delta f t}, e^{j2\pi k \Delta f t} \rangle \quad (2.11c)$$

$$= s_k \quad (2.11d)$$

for all $k = \{0, 1, \dots, N-1\}$ symbols packed as an OFDM symbol of duration T_s .

It is straightforward to notice given the above analysis that the transforms associated with the OFDM continuous (passband) signal modulation and demodulation are in fact the inverse Fourier Transform (FT) (IFT) for the modulation, and respectively, the FT for the demodulation. Applying therefore appropriate sampling and downmixing, the equivalent discrete baseband formulations for the modulation and demodulation procedures involved in OFDM surface as [5]

$$\text{modulation: } x[n] = x(n\Delta T) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] \cdot e^{j2\pi f_k n \Delta T} \quad (2.12a)$$

$$= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] \cdot e^{j2\pi \frac{nk}{N}}, \quad n \in \{0, 1, \dots, N-1\} \quad (2.12b)$$

$$\text{demodulation: } X[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi f_k n \Delta T} \quad (2.13a)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi \frac{nk}{N}}, \quad k \in \{0, 1, \dots, N-1\} \quad (2.13b)$$

where for the sake of future convenience $X[k] \triangleq s_k$, and without loss of generality the sampling time $\Delta T = \frac{T_s}{N}$, and respectively, the base frequency $f_0 = 0$, such that the subcarrier harmonics become integer multipliers of Δk .

The baseband OFDM modulation and demodulation are therefore according to (2.12b) and (2.13b) nothing but the linear inverse Discrete Fourier Transform (DFT) (IDFT) (modulation), and respectively, DFT (demodulation) transforms. Additionally, these have been represented under their unitary normalized representation scaled by $\frac{1}{\sqrt{N}}$ for uniform and balanced transform energy dissipation between the OFDM Tx and Rx respectively. The latter transforms can be very efficiently implemented by their corresponding inverse fast-fourier transform (FFT) (IFFT) and FFT fast implementations following the optimization framework of the Cooley-Tuckey

radix-2 algorithm [36], reducing complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log_2 N)$.

Furthermore, OFDM is much more resilient to multipath propagation effects which are quite applicable in practice and affect especially single-carrier transmission methods. To increase robustness against multipath channels degradation and Intersymbol Interference (ISI) on the Rx side OFDM relies on the concept of cyclic prefix (CP) as a guard signal. The latter appends a short copy of the OFDM symbol itself to its beginning as a prefix, or to the end as a suffix, in order to combat the delay spread of the wireless multipath channels that may introduce ISI amongst the $X[k]$ symbols. This has been shown to be very effective in practice as long as the guard signal duration, *i.e.* T_g , is larger than the delay spread of the multipath channels, *i.e.* $T_g > \tau_s$ [5]. Since the CP is prefixed to the OFDM symbol it therefore increases the symbol duration such that the overall symbol duration is $T_s + T_g$, where $T_g \ll T_s$, which hardens the OFDM waveforms against multipath effects.

Another major advantage that OFDM in comparison to the previously discussed methods is in fact its spectral efficiency, and in particular, its bandwidth utilization efficiency [5,6]. This is clearly a byproduct of the tight packing of orthogonal harmonics as parallel subcarriers of information. Concretely, OFDM transmits information at a rate of $\frac{N}{T_s+T_g}$ utilizing a bandwidth of $(N + 2)\Delta f$, where the additional Δf spectra is in practice utilized as spectral separation at the side of the OFDM signal spectrum. It follows therefore that the bandwidth efficiency ratio for OFDM is given by

$$\eta_{\text{OFDM}} = \frac{\frac{N}{T_s+T_g}}{\frac{N+2}{T_s}} \tag{2.14a}$$

$$= \frac{N}{N + 2} \frac{T_s}{T_s + T_g} \tag{2.14b}$$

$$= \frac{1}{1 + \frac{2}{N}} \frac{1}{1 + \frac{T_g}{T_s}} \xrightarrow{N \rightarrow \infty, T_g \ll T_s} 1, \tag{2.14c}$$

which implies that asymptotically OFDM reaches perfect spectral utilization, whereas in practice since $N \gg 2$ and $T_g \ll T_s$ the achievable bandwidth efficiency is also close to the limit. An illustrative example of the time-domain achievable OFDM subcarriers packing together with the optimum sampling points is shown in Figure 2.4.

In the light of all the features briefly presented above, and respectively, of all the advantages that OFDM subsumes over its predecessors, *i.e.*

- high spectral efficiency as bandwidth is saved by orthogonal tight packing of information in frequency domain under an optimized narrowband intercarrier interference (ICI),
- optimized implementation of modulation and demodulation based on simple linear transforms achieved by IFFT and FFT computationally efficient algorithms,
- robust against harsh wireless multipath phenomena without need for advanced time domain equalization,
- reduced sensitivity to timing synchronization errors at the Rx side – demodulation reduced

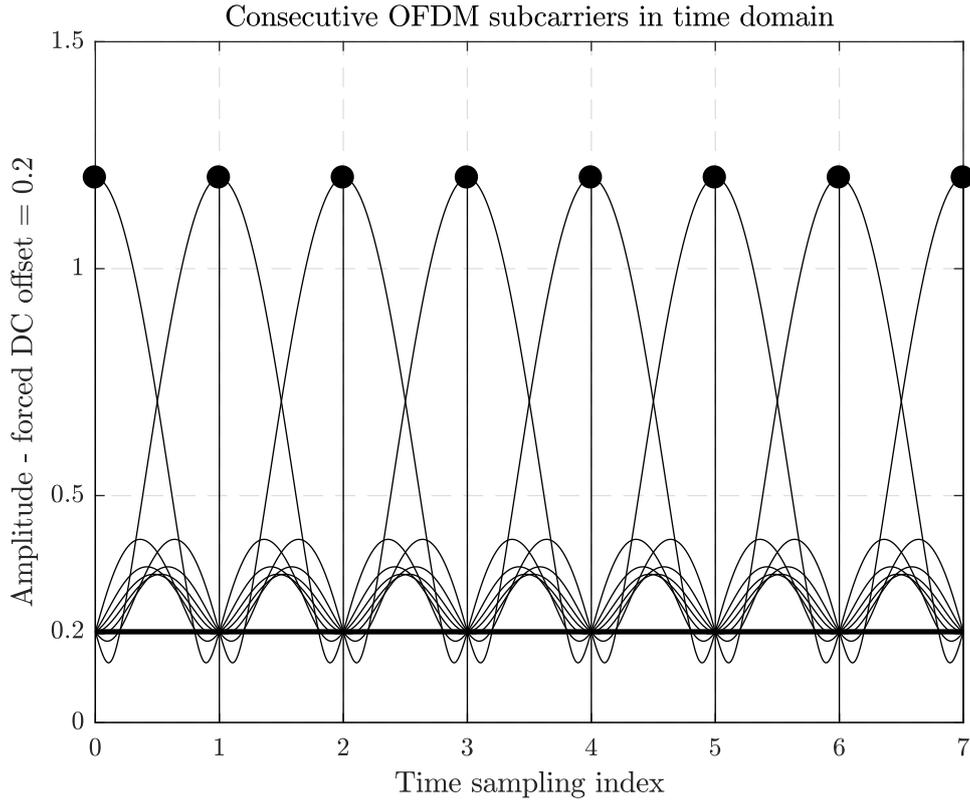


Figure 2.4: OFDM signal and subcarriers time representation and ideal sampling example.

to block retrieval of frequency encoded information via FFT.

OFDM became the principal high throughput and spectral efficient *physical transport waveform* and *multiple access* method of modern wireless communications from LAN to 4G and nowadays 5G/5G+.

Despite the favorable attributes OFDM presents as detailed above, there are a couple of areas where it suffers. These are as well inherited given the OFDM core concepts and represented respectively by:

- high peak-to-average-power-ratio (PAPR) upper bound linear in the number of used subcarriers N [5] requiring additional post-processing and dedicated RF circuitry to avoid inter-modulation interference.
- increased sensitivity to Doppler shifts and frequency offsets which may interfere with the OFDM orthogonality and cause ICI, and thus performance degradation.

Under symbol energy normalized assumptions, the first item is intuitively clear following (2.4) for instance. In other words, despite the fact that on average the symbol energy is unitary, *i.e.* $\mathbb{E}[\|s\|_2^2] = 1$, there may be symbols which deviate significantly from this value. Given the OFDM modulation technique where the information is carried over subcarriers directly, it follows that the dynamic range between peaks of the OFDM symbols may be large and in fact outside of the linear range of the RF amplification circuitry leading to clipping and non-linear

distortions [5]. This is a known problem of OFDM systems and multiple solutions have been researched and proposed to solve it, see for instance [37] and references therein.

The second item mentioned as a potential weakness of OFDM systems is caused by mobile environments in which either the Tx and /or Rx are mobile relative to each other or the scatterers are respectively. This mobility leads to Doppler effects which impose slight frequency shifts on the OFDM subcarriers potentially affecting their orthogonality and thus introducing ICI. This is the case in particular for complex geometric based wireless mobile channels, like most of the practical ones, that have complex Doppler spectra thus affecting non-linearly the OFDM subcarriers [5, 38–42]. ICI can however be avoided in typical systems by waveform numerology parameters and carrier frequency selection that exceed greatly the maximum Doppler spread, but the Doppler spectra effects are still affecting the subcarriers requiring to some extent a more involved symbol retrieval on the Rx side.

Having discussed the general concepts, advantages and potential weaknesses of OFDM let us next focus on the generic transmission system model. Consider next a data source which emits bits mapped through an aleatory modulation coding scheme (MCS) to a set of constellation symbols $\mathcal{M} \in \mathbb{C}$. It so follows that any OFDM symbol is formed by multiplexing N subcarriers modulated $X[k], k \in \{0, 1, \dots, N-1\}$ information symbols [5]. Let additionally $\mathbf{x} \triangleq [x[0], x[1], x[2], \dots, x[N-1]]^T$, $\mathbf{X} \triangleq [X[0], X[1], X[2], \dots, X[N-1]] \in \mathbb{C}^N$, and respectively, denote by $\mathbf{D}_N \in \mathbb{C}^{N \times N}$ the unitary N -DFT matrix. Thus, in matrix form (2.12b) becomes

$$\mathbf{x} = \mathbf{D}_N^H \mathbf{X}. \quad (2.15)$$

The linear baseband system model associated with an OFDM received symbol radiated over the channel $\mathbf{h} \in \mathbb{C}^N$ is resumed to the time domain equation

$$\mathbf{y} = \mathbf{h} \circledast \mathbf{x} + \mathbf{w}, \quad (2.16)$$

where \circledast denotes the circular convolution associated with the sampled and discretized linear convolution operator and \mathbf{w} represents the AWGN receive complex noise distributed according to $\mathcal{CN}(0, \sigma_w^2)$.

The circular convolution in (2.16) is rewritten as matrix multiplication by means of the circulant Toeplitz matrix circulant

$$\mathbf{C}_h = \begin{bmatrix} h[0] & h[N-1] & \dots & h[2] & h[1] \\ h[1] & h[0] & h[N-1] & & h[2] \\ \vdots & h[1] & h[0] & \ddots & \vdots \\ h[N-2] & & \ddots & \ddots & h[N-1] \\ h[N-1] & h[N-2] & \dots & h[1] & h[0] \end{bmatrix}, \quad (2.17)$$

such that the received time domain OFDM symbol is

$$\mathbf{y} = \mathbf{C}_h \mathbf{x} + \mathbf{w}. \quad (2.18)$$

Since circulant matrices are diagonalized by DFT unitary matrices [43], it results that the frequency domain OFDM symbol can in turn be written based upon the original frequency domain modulated symbols, *i.e.* \mathbf{X} , as

$$\mathbf{D}_N \mathbf{y} = \mathbf{D}_N \underbrace{\left(\mathbf{D}_N^H \overbrace{\text{diag}(\mathbf{H})}^{\text{CFR}} \mathbf{D}_N \right)}_{\text{upon diagonalization of } \mathbf{C}_h} \mathbf{x} + \underbrace{\mathbf{D}_N \mathbf{w}}_{\text{noise spectrum}} \quad (2.19a)$$

$$\mathbf{Y} = \mathbf{I}_N \text{diag}(\mathbf{H}) (\mathbf{D}_N \mathbf{x}) + \mathbf{W} \quad (2.19b)$$

$$= \text{diag}(\mathbf{H}) \mathbf{X} + \mathbf{W}. \quad (2.19c)$$

The above derivation is valid under the implicit assumption of no-ICI. Under this condition the circulant matrix in (2.17) is diagonalizable by the DFT transform vectors, and respectively, its diagonal entries are the channel frequency response (CFR) of the channel temporal sampled acquisition, *i.e.* $\mathbf{H} = \mathbf{D}_N \mathbf{h}$. In addition, since \mathbf{D}_N is an unitary transform the distribution and power of the AWGN noise is also preserved under Parseval's Theorem [2, 6], and hence, the demodulation to baseband frequency domain does not affect the signal-to-noise ratio (SNR).

In the light of all of the above, OFDM is therefore a highly efficient multiplexing scheme allowing for parallel processing of information streams as a tightly packed spectral representation enabled by the orthogonal DFT transform and its inverse.

2.3 A Simple Case Study: IEEE 802.11p

Let us now provide a compacted case study of an OFDM-based communication system, *i.e.*, IEEE 802.11p [25], in order to highlight that even one of the most efficient orthogonal waveform and access scheme is still challenged by practical applications to be expected in the future.

As a central technological part of Intelligent Transportation System (ITS), the ITS-G5 standard, namely 802.11p or DSRC, has been firstly proposed in [29] and later amended in [25], branching from the legacy indoor wireless LAN standard 802.11a [44]. The IEEE 802.11p is intended to act as the *de facto* technology for ITS, and hence, it is of high interest from the vehicular communication and access perspective [45].

The mobile outdoor channels associated with V2X communications are doubly-selective in time and frequency [41, 42, 45–50]. In fact, the V2X time-varying channels exhibit large delay spreads and high Doppler shifts given their interfering dynamic multipath components distorting the Tx radiated signals over the transmission paths to the Rx. Furthermore, from a stochastic perspective these effects generate in fact dynamic non-wide-sense stationary uncorrelated scattering (WSSUS) V2X channels dominated by fast fading effects and varying Doppler spread [49–52].

A common mathematical modeling of the IEEE 802.11p channels is thus resumed to their representation as tapped-delay line (TDL) filters with stochastically determined parameters describing the attenuation, the Doppler spread, and respectively, the time propagation delay

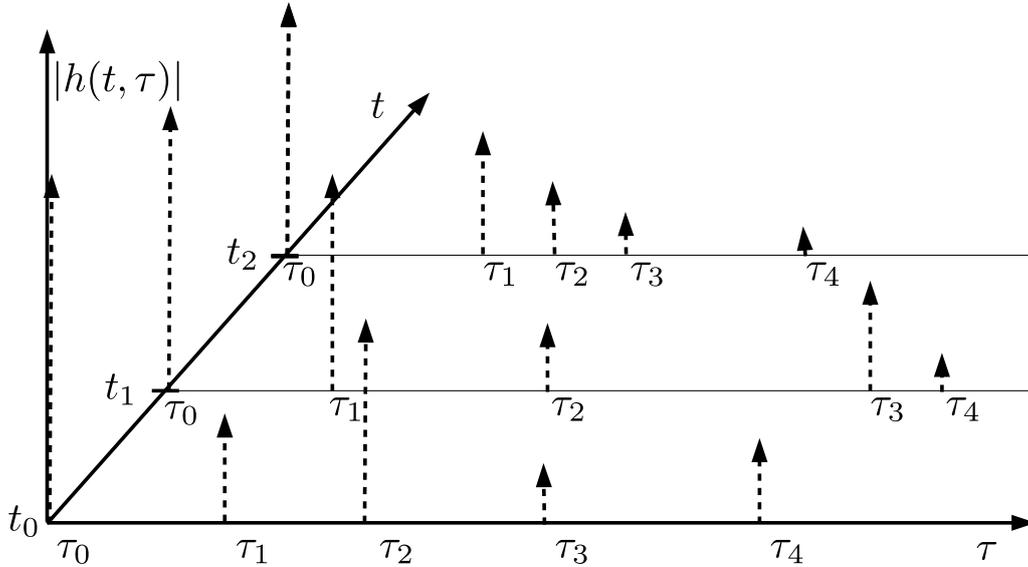


Figure 2.5: Channel impulse response sketch of a V2X time-varying mobile channel.

corresponding to each individual path [41, 42, 46, 48]. This model corresponds to the channel impulse response

$$h(t, \tau) = \sum_{p=1}^P h_p(t) e^{j2\pi\nu_p t} \delta(\tau - \tau_p(t)), \quad (2.20)$$

where for each p -th path the complex time-varying attenuation factor is denoted by $h_p(t)$, the Doppler shift projected on to the direction Rx is given by ν_p , and respectively, the time-varying propagation delay is marked as τ_p .

Furthermore, it is important to remark that the Doppler shift is in fact a consequence of the embedded system mobility, *i.e.*, the relative velocity between Tx and Rx along a given propagation path p such that

$$\nu_p \triangleq \frac{v_{\perp}}{c} f_c \quad (2.21)$$

with v_{\perp} being the projected relative velocity between Tx and Rx on the Rx motion direction. An illustrative sketch of such a discrete representation of a channel impulse response is given in Figure 2.5.

The PHY layer of IEEE 802.11p is based on the conventional OFDM waveforms [29] as described generically in the preceding Subsection. Usual to practical systems [53], the IEEE 802.11p PHY protocol provides additional mechanisms meant to provide robustness against channel induced errors. A complete functional diagram of an IEEE 802.11p [25] Tx-Rx link is thus presented in Figure 2.6.

Following the signal processing flowgraph of Figure 2.6, the inputs to the Tx are generated by a data source providing randomly bits to be coded, modulated and then transmitted as OFDM symbols. These bits are firstly scrambled using a linear shift feedback register to

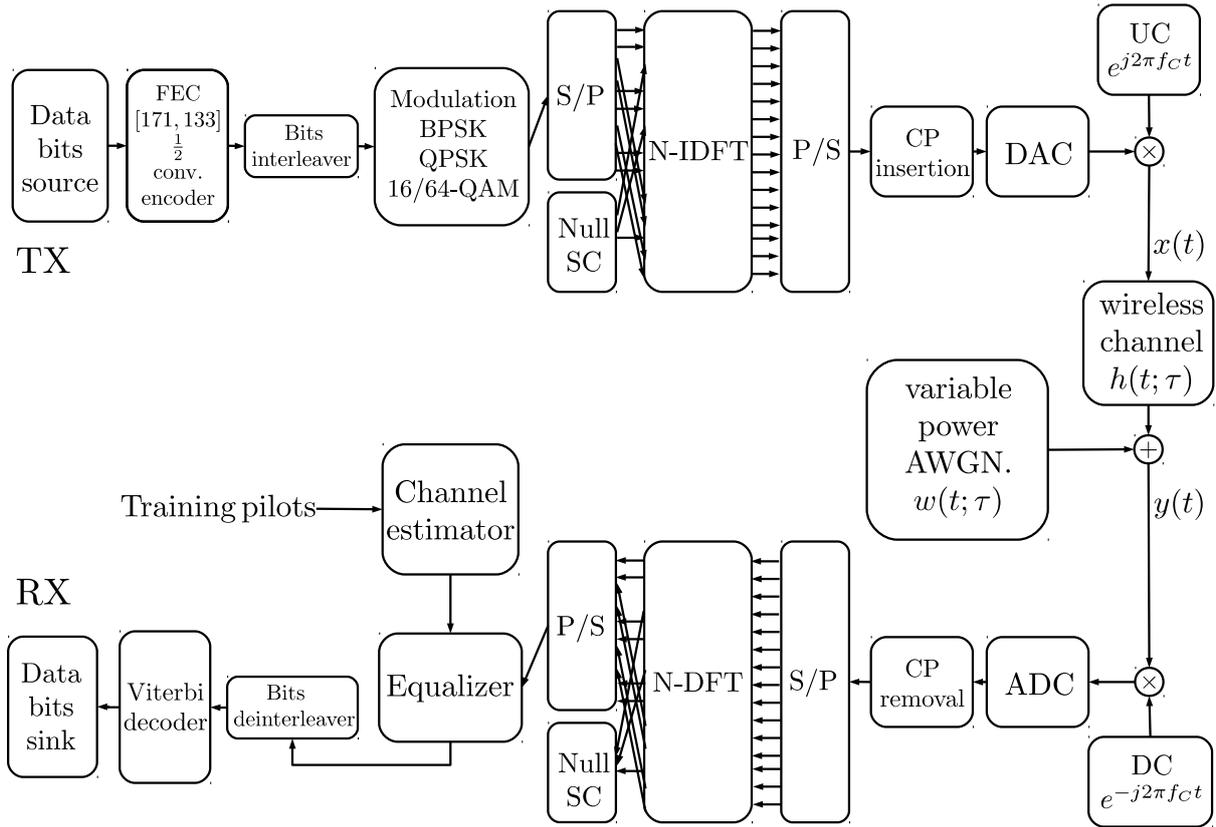


Figure 2.6: IEEE 802.11p Tx-Rx system diagram following specifications in [54]. ©2017 IEEE

ensure an uniform spreading of bits over an OFDM symbol regardless of the distribution of the data source, and therefore, maximize source information entropy. The output bitstream of the scrambler is then encoded using a convolutional code of constraint length $K = 7$ defined by the octal polynomials [171, 133]. This operation yields a rate of $R = 1/2$, but higher coding rates are attainable, i.e. $2/3, 3/4$, via puncturing [25]. The convolutional encoding is next followed by a block two-step permutation interleaver. The first permutation ensures that adjacent coded bits are mapped onto nonadjacent subcarriers, while the second one ensures that adjacent coded bits are mapped alternately onto less and more significant bits of the constellation, and so, potential burst errors on any of the OFDM subcarriers are mitigated [29]. The interleaving is followed by constellation mapping according to the desired data rate given the unitary powered constellations: BPSK, QPSK, 16QAM, and respectively, 64QAM. The 6 Mbps MCS corresponding to 1/2-coded QPSK is the protocol's default and the maximum data rate is of 27 Mbps [25]. The constellation mapping is followed by the OFDM modulation performed efficiently by the IFFT algorithm, whose outputs are later multiplexed and finally appended with a CP. Lastly, the obtained signals are upconverted to the passband frequency in the band of 5.9 GHz and sent over the available V2X channels [26, 27, 29].

On the other hand, the IEEE 802.11p Rx side is reciprocal to its Tx counterpart. Hence, upon preamble-based detection and synchronization the captured signal is downconverted to baseband and digitized. Then the CP is discarded and the OFDM information are sequentially retrieved via channel equalization and demodulation of each symbol.

Table 2.1: IEEE 802.11A/P COMMUNICATION PARAMETERS. ©2017 IEEE

Parameter	IEEE 802.11a [44] value	IEEE 802.11p [29] value
Carrier Frequency Band: f_c	5.2 GHz	5.9 GHz
Bandwidth: B	20 MHz	10 MHz
Constellation Modulations: \mathcal{M}	BPSK, QPSK, 16/64-QAM	
Code rates: R	1/2, 2/3, 3/4	
Data subcarriers: N_{dsc}	48	
Pilot subcarriers: N_{psc}	4	
Total subcarriers: N	64	
Carrier spacing: Δf	0.3125 MHz	0.15625 MHz
Baseband Sampling Time: T_s	0.05 μ s	0.1 μ s
OFDM Symbol Duration: T_N	3.2 μ s	6.4 μ s
Cyclic Prefix Duration: T_G	0.8 μ s	1.6 μ s
Total Symbol Duration: T_{OFDM}	4.0 μ s	8.0 μ s
Error Correction Coding: FEC	$K = 7, [171\ 133]$ Conv. Encoder	

The standard training symbols initialize the channel estimation block by providing an estimate of the CFR. This coherent initial estimate is further used in the equalization of upcoming OFDM symbols. After the equalization has been performed, the estimated symbols are demapped to bits which are further reciprocally deinterleaved in comparison to the similar Tx operation. The forward error correction code (FEC) decoding is implemented via the Viterbi decoder which returns the ML bit sequence for each OFDM symbol in the payload. Lastly, the output bitstream of the Viterbi decoder is descrambled using the same linear shift feedback register as in the Tx in order to retrieve the original source input bitstream. A summary of the system and waveform parameters that define the performance of IEEE 802.11p for DSRC in ITS applications over V2X channels is given in Table 2.1.

Empirical performance tests and field trials [45–47,55] proved that the IEEE 802.11p system parameters, see for a summary Table 2.1, are sufficient to ensure that the OFDM modulation is not affected in practice neither by ISI nor by ICI. In the light of Equation (2.19c) it follows that the CFR matrix of \mathbf{h} is diagonal, and as a result, simple least squares (LS) methods can be used to perform channel equalization and retrieve the transmitted OFDM symbols. In fact considering packetized bursty transmissions over the IEEE 802.11p PHY as described insofar, each subsequent i -th receive OFDM symbol and each of its subcarriers $k \in \{0, 1, \dots, N\}$ are obtained given (2.19c) as individually separated

$$Y_i[k] = H_i[k]X_i[k] + W_i[k], \quad (2.22)$$

where the subscript $0 \leq i \leq F - 1$ has been used to denote the temporal ordering of the individual OFDM symbols within a transmitted frame/packet of length F .

Hence, given Equation (2.22), a direct LS solution for the receive symbol leads to both channel estimation

$$\hat{H}_i[k] = \frac{Y_i[k]}{X_i[k]} \quad (2.23a)$$

$$= H_i[k] + \frac{W_i[k]}{X_i[k]} \quad (2.23b)$$

when $X_i[k]$ are considered known as either OFDM training pilots or estimated values, and respectively, to equalization

$$\hat{X}_i[k] = \frac{Y_i[k]}{H_i[k] + \Delta H_i[k]} \quad (2.24a)$$

$$= \frac{X_i[k]}{1 + \frac{\Delta H_i[k]}{H_i[k]}} + \frac{W_i[k]}{H_i[k] + \Delta H_i[k]} \xrightarrow{|\Delta H_i[k]| \ll |H_i[k]|} X_i[k] + \frac{W_i[k]}{H_i[k]} \quad (2.24b)$$

when $\hat{H}_i[k] = H_i[k] + \Delta H_i[k]$ is perfectly known *a priori*, *i.e.*, $\Delta H_i[k] = 0$, or practically estimated by the receiver.

As previously mentioned in (2.19), $H_i[k]$ represents the discretized CFR of \mathbf{h} , *i.e.*, $H_i[k] = H(iT_s, k\Delta f)$, where by FT

$$H(t, f) \triangleq \int_{-\infty}^{\infty} h(t, \tau) e^{-j2\pi f\tau} d\tau \quad (2.25a)$$

$$= \int_{-\infty}^{\infty} \left(\sum_{p=1}^P h_p(t) e^{j2\pi\nu_p t} \delta(\tau - \tau_p(t)) \right) e^{-j2\pi f\tau} d\tau \quad (2.25b)$$

$$= \sum_{p=1}^P h_p(t) e^{j2\pi\nu_p t} \int_{-\infty}^{\infty} \delta(\tau - \tau_p(t)) e^{-j2\pi f\tau} d\tau \quad (2.25c)$$

$$= \sum_{p=1}^P h_p(t) e^{j2\pi\nu_p t} e^{-j2\pi f\tau_p(t)}. \quad (2.25d)$$

However, as remarked earlier despite the fact that ISI and ICI are essentially mitigated by the IEEE 802.11p PHY numerology, the inherent Doppler and multipath fine distortions of the channel are still present and affect the received symbols, as exemplified by Equations (2.22) and (2.25d) above. Since in addition the time-varying V2X wireless channels are non-WSSUS [49–52], as mentioned above, the tracking of the CFR is further required over the duration of a packet in order to avoid the temporal deprecation of the initial coherent channel estimates, and hence, inaccurate equalization. To resolve the latter challenges, the focus is placed in the sequel on a low-latency solution to jointly estimate and track $H_i[k] \forall i, k$, in the form of a IEEE 802.11p-compliant receiver structure.

2.3.1 Channel Estimation via Self-Organizing Frequencies

A key observation derived in [40] and detailed in [39] regarding the non-WSSUS channels is

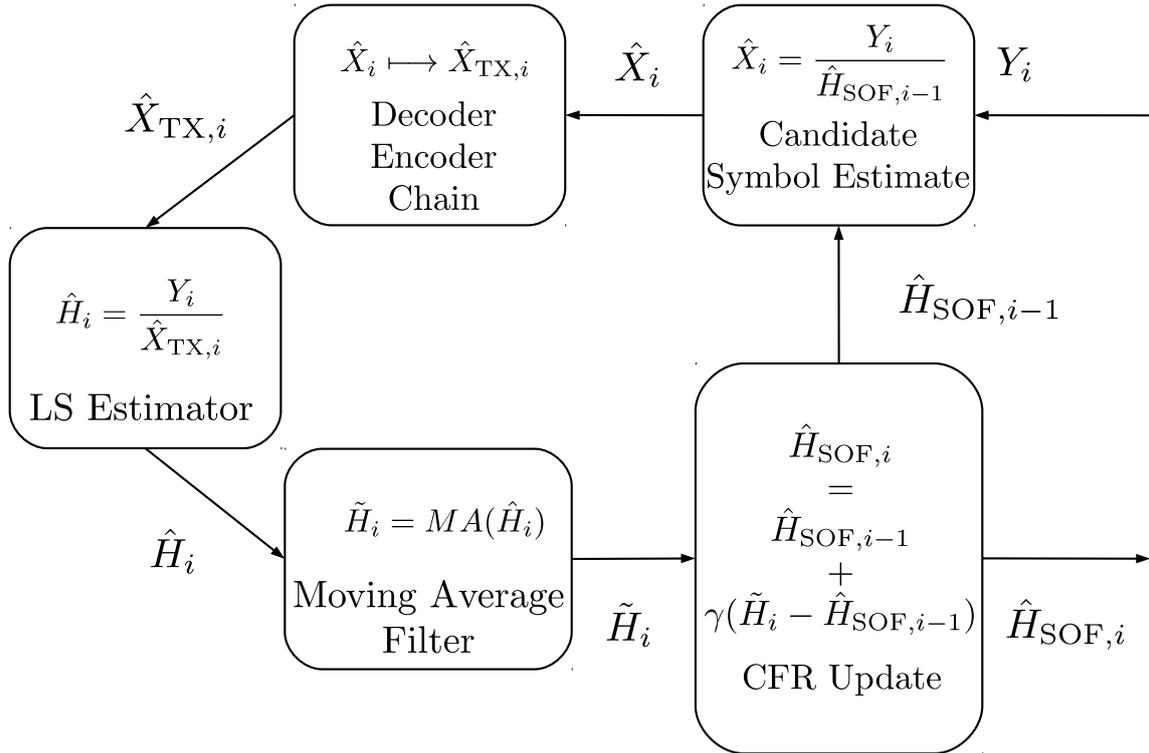


Figure 2.7: Block diagram of the proposed SOF joint channel estimator and receiver structure. ©2017 IEEE

that the *channel adjacent subcarriers* are in fact *correlated both in time and frequency domain*. In other words, there is a high level of similarity among adjacent subcarriers within the IEEE 802.11p receiver estimated channel for a limited time span over some sequential OFDM symbols.

The researched receiver and detailed hereafter is thus trying to leverage this critical insight. To this extent, individual estimated subbands of the channel frequency response are allowed to *organize themselves* based on local temporal and spectral correlations. This mechanism is then used to jointly track and estimate the wireless channel at the receiver over the course of a received packet. The receiver presented henceforth is coined as the self-organizing frequencies (SOF) receiver being firstly introduced in [56] and later substantially improved in [57].

The SOF is a sequential joint LS channel estimation and equalization algorithm with decision feedback. It contains also an embedded channel tracking structure based on linear time averaging controlled by CFR spectral and temporal correlations. For an enhanced decoding performance, the SOF equalization is checked and error corrected by one pass through a zero-padded Viterbi decoder before the estimated re-encoded symbols are used to reestimate the channel. The block diagram structure of the SOF receiver is summarized in Figure 2.7.

The first three blocks on the input side in Figure 2.7 represent graphically the forward-backward two-fold process over all subcarriers k of equalization of the i -th OFDM transmitted symbol, $X_i[k]$, and respectively, estimation of the i -th corresponding CFR, $H_i[k]$. The latter step is performed by feedback given the decoded, corrected and re-encoded OFDM symbol initially estimated. The FEC decoder/encoder chain is thus used in the receiver – similar to the

Turbo receiver concept [58] – to correct the possibly erroneous estimates of the transmitted symbols. The main causes of such errors are in practice poor previous channel estimates or high variations between current and previous real CSI realizations that failed to be properly temporally tracked. However, opposed to Turbo decoding, the SOF passes through its own decoder only once and relies on tracking the CFR estimates by temporal and spectral correlations in order to improve the accuracy of its channel estimates and inherently of the equalization. In the sequel, the last two blocks of the SOF are discussed in more details.

The *moving average (MA)* filter is used to further refine the channel estimates by *adjacent bands denoising*. This is possible as the noise samples in different subbands are statistically independent whilst the adjacent subcarrier complex values are highly correlated [45, 52]. The core filtering operation with length l is defined as

$$\tilde{H}_i[k] \triangleq \sum_{j=-\lfloor \frac{l}{2} \rfloor}^{\lfloor \frac{l}{2} \rfloor} \omega_j \hat{H}_i[k+j], \quad (2.26)$$

where $\lfloor \cdot \rfloor$ denotes the floor operation and ω_j represents the unit sum normalized tap values obtained from any filtering kernel.

The selected filtering kernel Kern_{MA} is the discretized and truncated Gaussian function. This kernel has been chosen on the one hand since the Gaussian distribution is in fact the highest entropic distribution with a known mean and variance [2]. On the other hand, it preserves mainly its central component, *i.e.*, the middle tap, while the other components decay to the side lobes of the filter. The sampling of the kernel is done following a truncation which symmetrically limits the side lobes of the core Gaussian continuous function and yields the taps

$$\text{Kern}_{\text{MA}}(k_0; \ell, \sigma, tr) = \left[\exp\left(-\frac{(k_0 - \ell)^2}{\sigma^2}\right) \right]_{tr}, \quad (2.27)$$

centered at frequency k_0 with spread σ .

The value of the truncation level is fixed by thresholding with $tr \ll 1$ such that the *bell* shape of the Gaussian function is preserved. To this end, $tr = 0.01$ has been found to be a good cutoff point for the Gaussian side lobes. Complementary to truncation, a maximum length is set as $l_{max} = 11$ in order to limit the filtering length, and so, the spectral leakage in adjacent subbands of the center frequency of interest k_0 .

Given all the above, the parameter σ yields therefore the spread of the filter taps across the subcarriers neighboring the central k_0 . The MA filter spread and implicitly, as previously detailed, the filter length l are controlled by σ which is modeled as a function of the SNR, hereby denoted as ρ . The connection between σ and the SNR is natural, as in low SNRs the noise is stronger and thus a longer filter is needed to denoise the subcarriers, whereas in high SNRs the noise is weak and spectral filtering may be reduced or drastically bypassed. Using the initial coherent piloting scheme the SNR can be easily estimated [59] as $\hat{\rho}$ for an entire packet. Under these assumptions, the following linear step-wise functional relation between the filter

length and the SNR estimates is proposed

$$\sigma(\hat{\rho}) = \begin{cases} \sqrt{\frac{-w_{up}^2}{\ln tr}}, & \hat{\rho} \leq 0 \\ \left(\sqrt{\frac{-w_{low}^2}{\ln tr}} - \sqrt{\frac{-w_{up}^2}{\ln tr}} \right) \frac{\hat{\rho}}{\rho_\sigma} + \sqrt{\frac{-w_{up}^2}{\ln tr}}, & \hat{\rho} \in (0, \rho_\sigma], \\ \sqrt{\frac{-1}{\ln tr}}, & \hat{\rho} > \rho_\sigma \end{cases} \quad (2.28)$$

with the lower and upper cutoff values w_{low}, w_{up} respectively defined as

$$w_{low} \triangleq \operatorname{argmax}_{\ell \in \mathcal{L}} R_{i, \hat{H}\hat{H}}[\ell] > 0.9, \quad (2.29)$$

$$w_{up} \triangleq \operatorname{argmax}_{\ell \in \mathcal{L}} R_{i, \hat{H}\hat{H}}[\ell] > 0.7, \quad (2.30)$$

in terms of the normalized CFR spectral correlation

$$R_{i, \hat{H}\hat{H}}[\ell] \triangleq \frac{|\sum_{k=-26}^{26-\ell} \hat{H}_i[k] \hat{H}_i^*[k+\ell]|}{\sum_{k=-26}^{26} |\hat{H}_i[k] \hat{H}_i^*[k]|}, \quad (2.31)$$

where integers $\ell \in \mathcal{L} \triangleq \{0, 1, \dots, \lfloor \frac{l_{max}}{2} \rfloor\}$.

The spectral discrete auto-correlation defined in Equation (2.31) is therefore a normalized measure of frequency selectivity relative to the possible MA filter lengths. It follows thus clearly based also on Equations (2.26) - (2.28) that if the SNR is high approaching the threshold ρ_σ , then the filtering is adaptively reduced more and more being finally disabled when the threshold ρ_σ is attained by the estimated SNR. The later SNR limit to disable the threshold has been experimentally set to 30 dB. Hence, these deterministic rules ensure that regardless of the shape and frequency selectivity of the real CSI the channel estimates are denoised while preserving the inter-carrier local correlations and geometrical representation without spectral leakage distortions.

Figures 2.8 - 2.10 show the adaptive MA filtering that each subcarrier undergoes under different SNRs. To outline the adaptation effects, the plots present the theoretical SNR parameterized filtering windows for the real channel model of V2V_{EW0}, *i.e.*, vehicle-to-vehicle (V2V) express-way oncoming at 104 km/h, based upon [41, 42], for low, medium and high SNRs.

The last block of the SOF receiver from Figure 2.7, the *update rule*, enables the channel tracking based on the exponential filtering

$$\hat{H}_{\text{SOF},i}[k] = \hat{H}_{\text{SOF},i-1}[k] + \gamma(\hat{\rho}) \left(\tilde{H}_i[k] - \hat{H}_{\text{SOF},i-1}[k] \right). \quad (2.32)$$

with an adaptation rate of

$$\gamma(\hat{\rho}) \triangleq c(\hat{\rho}) \cdot \frac{|\sum_{k=-26}^{26} \hat{H}_{\text{SOF},i-1}[k] \tilde{H}_i^*[k]|}{\sum_{k=-26}^{26} |\hat{H}_{\text{SOF},i-1}[k]|^2 + |\tilde{H}_i[k]|^2}, \quad (2.33)$$

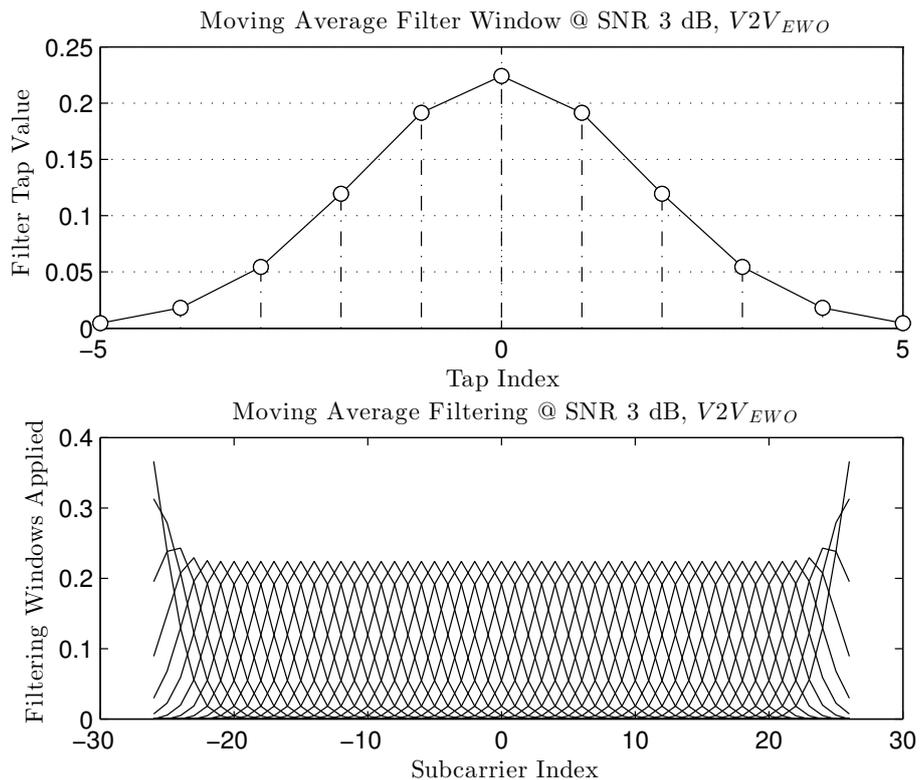


Figure 2.8: MA filtering examples — V2V channel (Express-Way Oncoming 104 km/h), [41]: $V2V_{EWO}$ @ 3 dB

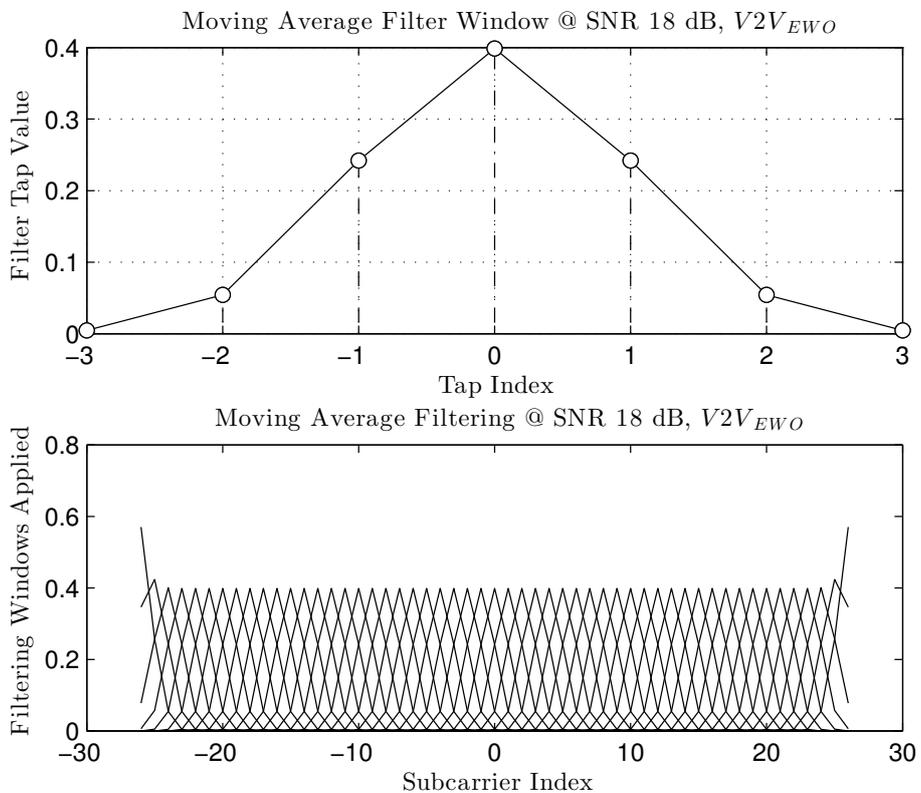


Figure 2.9: MA filtering examples — V2V channel (Express-Way Oncoming 104 km/h), [41]: $V2V_{EWO}$ @ 18 dB

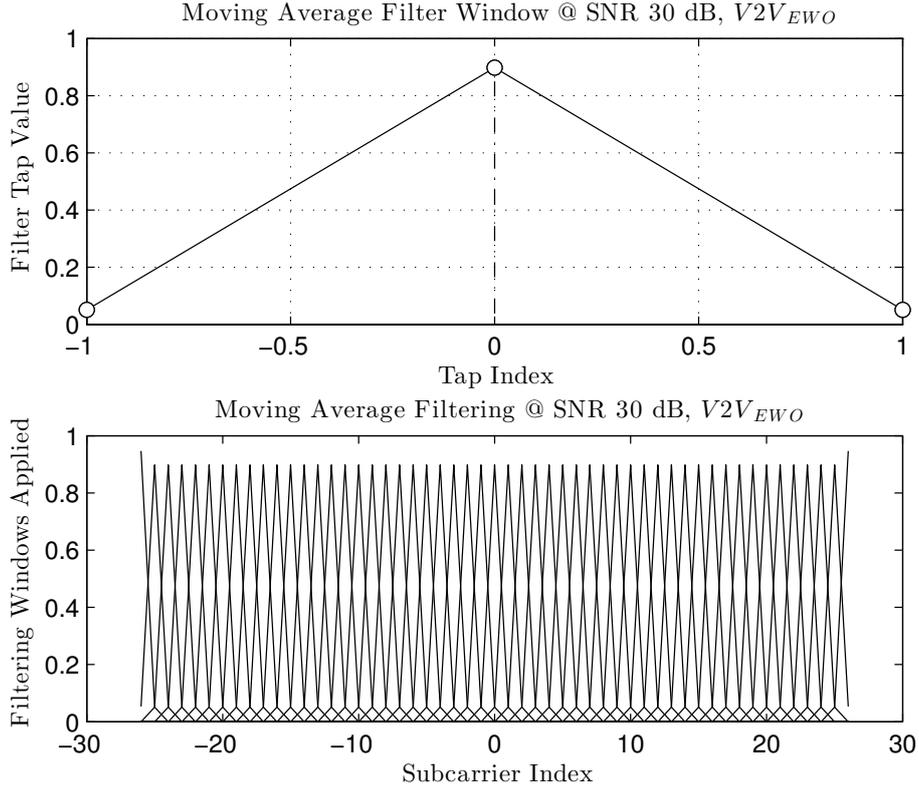


Figure 2.10: MA filtering examples — V2V channel (Express-Way Oncoming 104 km/h), [41]: $V2V_{EWO}$ @ 30 dB

and a confidence coefficient in *tracking* the new estimates $\tilde{H}_i[k]$ defined as

$$c(\hat{\rho}) = \begin{cases} 1 & \hat{\rho} \leq 0 \\ 1 + \frac{\hat{\rho}}{\rho_c} & \hat{\rho} \in (0, \rho_c] \\ 2 & \hat{\rho} > \rho_c \end{cases} . \quad (2.34)$$

The SOF update rule principles for each subcarrier are graphically displayed in Figure 2.11.

The initialization of the channel tracking, *i.e.*, the values of $H_{\text{SOF},0}[k]$, $\forall k$, are obtained based on the initial two repeated block training symbols, namely $X_{0,T_1}[k] = X_{0,T_2}[k] \triangleq X_{0,T}[k]$, $\forall k$, in the preamble of the IEEE 802.11p packets, such that

$$H_{\text{SOF},0}[k] = \text{MA} \left(\frac{Y_{0,T_1}[k] + Y_{0,T_2}[k]}{2X_{0,T}[k]} \right). \quad (2.35)$$

In addition, these two block training symbols from the preamble are used to obtain accurate levels of the average noise power at the receiver and thus also approximate the SNR [59] for the entire received packet in order to enable the adaptive filtering described previously.

Once more, a threshold ρ_c is used to mark the piece-wise continuous adaptive filtering and tracking of the new channel estimates. These are so considered to become more and more reliable with the increase of the estimated SNR. The latter measure is used as a result to decide whether it is best to track the channel by *jumping* directly to the current MA-filtered estimate

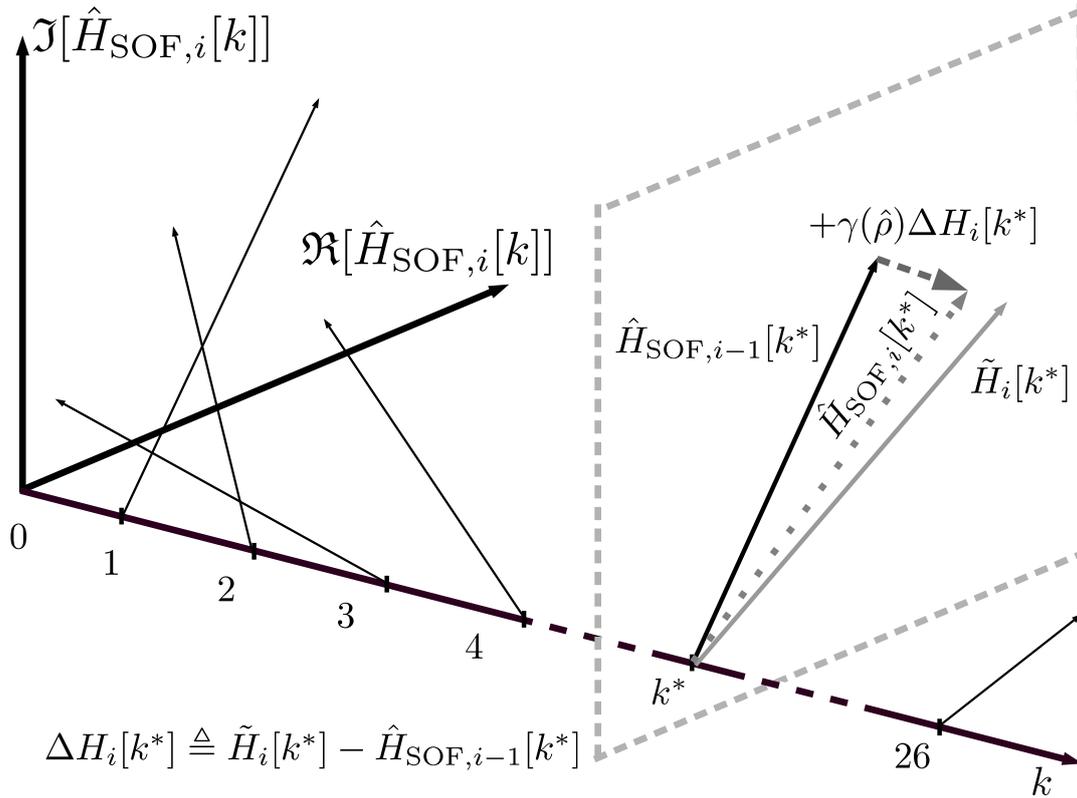


Figure 2.11: The IEEE 802.11p associated subcarrier channel feather and the corresponding SOF update rule (2.32) for any fixed subcarrier k^* over the i -th OFDM symbol. ©2017 IEEE

instead of linearly combining it with the previous estimate. Moreover, the temporal correlation of consecutive channel estimates is included by the adaptation rate $\gamma(\hat{\rho})$ via the second non-negative term of the product in Equation (2.33). The normalized correlation reaches in fact its maximum value of 0.5 for the limit case where $H_{\text{SOF},i}[k]$ and $\tilde{H}_i[k]$ are identical and is thus scaled by the confidence coefficient to maximum value of 1 for high SNR cases. In low SNR scenarios the adaptation rate is lowered implying a more intensive temporal averaging of the channel estimates meant to reduce the estimation noise. To this end the value of ρ_c has been set to be $\rho_c = 30$ dB.

Nonetheless, given all of the above, Equations (2.32) - (2.34) are effective just for small variations between consecutive CSI estimates, and thus still cannot avoid error propagation due to high temporal variations over the CFR that may appear. Hence, corrections of the CFR estimates may still be necessary along the entire length of a packet. This procedure is based therefore on the error-corrected symbols fed back to the estimator block of the SOF for each OFDM symbol and is enforced given that the SNR is reliable enough to *fully trust the decoded bits*. To summarize, the final estimate $\hat{H}_{\text{SOF},i}[k]$ for error-corrected data on the subcarrier k becomes

$$\hat{H}_{\text{SOF},i}[k] = \begin{cases} \tilde{H}_i[k] & , \hat{\rho} \geq \rho\hat{M} \\ \hat{H}_{\text{SOF},i-1}[k] + \gamma(\hat{\rho}) \left(\tilde{H}_i[k] - \hat{H}_{\text{SOF},i-1}[k] \right) & , \hat{\rho} < \rho\hat{M} \end{cases} \quad \forall k \text{ subcarrier error-corrected data.} \quad (2.36)$$

Algorithm 1 SELF-ORGANIZING FREQUENCIES RECEIVER.

-
- Inputs:** frequency domain received OFDM signals $Y_i[k] \forall i \in \{0, 1, \dots, F - 1\}$, MCS \mathcal{M} , training symbols $X_{0,T}[k]$, over all k OFDM subcarriers.
- Outputs:** equalized OFDM symbol estimates $X_i[k] \forall i \in \{1, 2, \dots, F - 1\}$, CFR as CSI estimates $H_{\text{SOF},i}[k] \forall i \in \{0, 1, \dots, F\}$ over all OFDM subcarriers k .
-
- 1: Estimate SNR, *e.g.* by method in [59]: $\hat{\rho}$
 - 2: Compute correction threshold (2.37): $\rho_{\mathcal{M}}$
 - 3: Compute confidence level (2.34): $c(\hat{\rho})$
 - 4: Compute initial CFR estimate from training symbols $X_{0,T}[k]$ by (2.35): $\hat{H}_{\text{SOF},0}[k]$
 - 5: **for** OFDM symbol $i = 1$ to F **do**
 - 6: Get OFDM symbol estimate: $\hat{X}_i[k] = \frac{Y_i[k]}{\hat{H}_{\text{SOF},i-1}[k]}$
 - 7: Deinterleave, correct, encode & interleave OFDM symbol estimate: $\hat{X}_i[k] \mapsto \hat{X}_{\text{Tx},i}[k]$
 - 8: Get intermediate channel estimate: $\hat{H}_i[k] = \frac{Y_i[k]}{\hat{X}_{\text{Tx},i}[k]}$
 - 9: Compute frequency correlation by (2.31): $R_{i,\hat{H}\hat{H}}[\ell]$
 - 10: Compute filtering window shape by (2.28): $\sigma(\hat{\rho})$
 - 11: Filter intermediate estimate by (2.26): $\tilde{H}_i[k] = \text{MA}(\hat{H}_i[k])$
 - 12: Compute time correlation and adaptation rate, (2.33): $\gamma(\hat{\rho})$
 - 13: Update SOF channel estimate: $\hat{H}_{\text{SOF},i}[k] = \hat{H}_{\text{SOF},i-1}[k] + \gamma(\hat{\rho})(\tilde{H}_i[k] - \hat{H}_{\text{SOF},i-1}[k])$
 - 14: **if** $\hat{\rho} > \rho_{\mathcal{M}}$ **then**
 - 15: $\hat{H}_{\text{SOF},i}[k] = \tilde{H}_i[k]$, $\forall k$ subcarriers with error detected & corrected symbols $\hat{X}_{\text{Tx},i}[k]$
 - 16: **end if**
 - 17: **end for**
-

Table 2.2: UPDATE RULE CORRECTION THRESHOLD FOR IEEE 802.11P CONSTELLATIONS.
 ©2017 IEEE

Constellation Scheme \mathcal{M}	$\rho_{\mathcal{M}}$ dB
BPSK	13.98
QPSK	16.99
16QAM	23.98
64QAM	30.21

The values of the *confidence SNR interval* $\rho_{\mathcal{M}}$ are computed additionally based on the used modulation schemes. They represent therefore *safety loci* around the constellation points spreading up to 5 times the constellation halved minimum distance under the assumption of receive complex AWGN, *i.e.*

$$\rho_{\mathcal{M}} = 10 \log_{10} \left(5^2 \left(\frac{d_{\min, \mathcal{M}}}{2} \right)^2 \right). \quad (2.37)$$

Assuming the valid IEEE 802.11p system modulation constellations, the values of $\rho_{\mathcal{M}}$ ac-

ording to (2.37) are listed in Table 2.2. Finally, to summarize, the SOF procedure is presented explicitly as Algorithm 1.

2.3.2 Performance and Access Limitations

The SOF receiver proposed previously has been evaluated over a complete PHY synthetic link compliant with the IEEE 802.11p standard as outlined in Figure 2.6. The simulations entailed multiple packets transmitted over stochastic channel realizations simulating real environment channels. To this extent, channels were implemented based on the stochastic specular models introduced in [41] following the measuring campaign over the 5 Ghz band detailed in [42]. The selected results discussed in the sequel rely on three of the six models from [41], namely:

- a V2V 300-400m express-way oncoming link for vehicles traveling at 104 km/h, [41] – *i.e.* $V2V_{EWO}$;
- a vehicle-to-infrastructure (V2I) 100m urban canyon channel for a vehicle moving at 32-48 km/h, [41] – *i.e.* $V2I_{UC}$;
- a V2I 100m suburban street link for a vehicle moving at 32-48 and 120 km/h, [41] – *i.e.* $V2I_{SS}$.

In this sense, the achievable bit-error rate (BER) performance of the IEEE 802.11p links in practical scenarios has been simulated and evaluated both for the SOF proposed method, but similarly also for existent alternatives. The state-of-the-art receivers considered for comparison were the conventional LS receiver, the spectral-time averaging (STA) receiver [60], and the family of Constructed Data Pilot (CDP) receivers, *i.e.* the CDP and SNR-assisted modified constructed data pilot (SAMCDP) detectors, from [61, 62].

In short, on one hand, the LS receiver outlines the achievable performance without channel tracking where the OFDM symbols are equalized just based on the initial channel acquisition given the training symbols, *i.e.* according to Equation (2.35). On the other hand, the STA and CDP receivers provide a typical decision feedback loop for channel re-estimation over each of the OFDM symbols in the received packets. The STA detector uses furthermore spectral MA in order to denoise the intermediate CFR estimates and then complements this filtering by temporal channel averaging between adjacent channel CFR acquisitions [60]. However, these procedures are completely non-adaptive and rely on fixed parameters. Lastly, the CDP receivers are rooted by the concept of validating the OFDM data as pilots based on the CFR time correlations [61, 62]. As a result, depending on the confidence in the constructed data pilots, decisions are made whether to update the CFR estimates or not for each subcarrier. Additionally the SAMCDP utilizes SNR estimation as a measure to improve the performance of the scheme for low-SNR cases by a more conservative update policy [62].

Performance evaluations across different channels, modulations and payload lengths are illustrated in Figures 2.12 - 2.15 for both the proposed SOF receiver, and respectively, the state-of-the-art methods detailed briefly above. In addition, as an ideal comparison curve, the

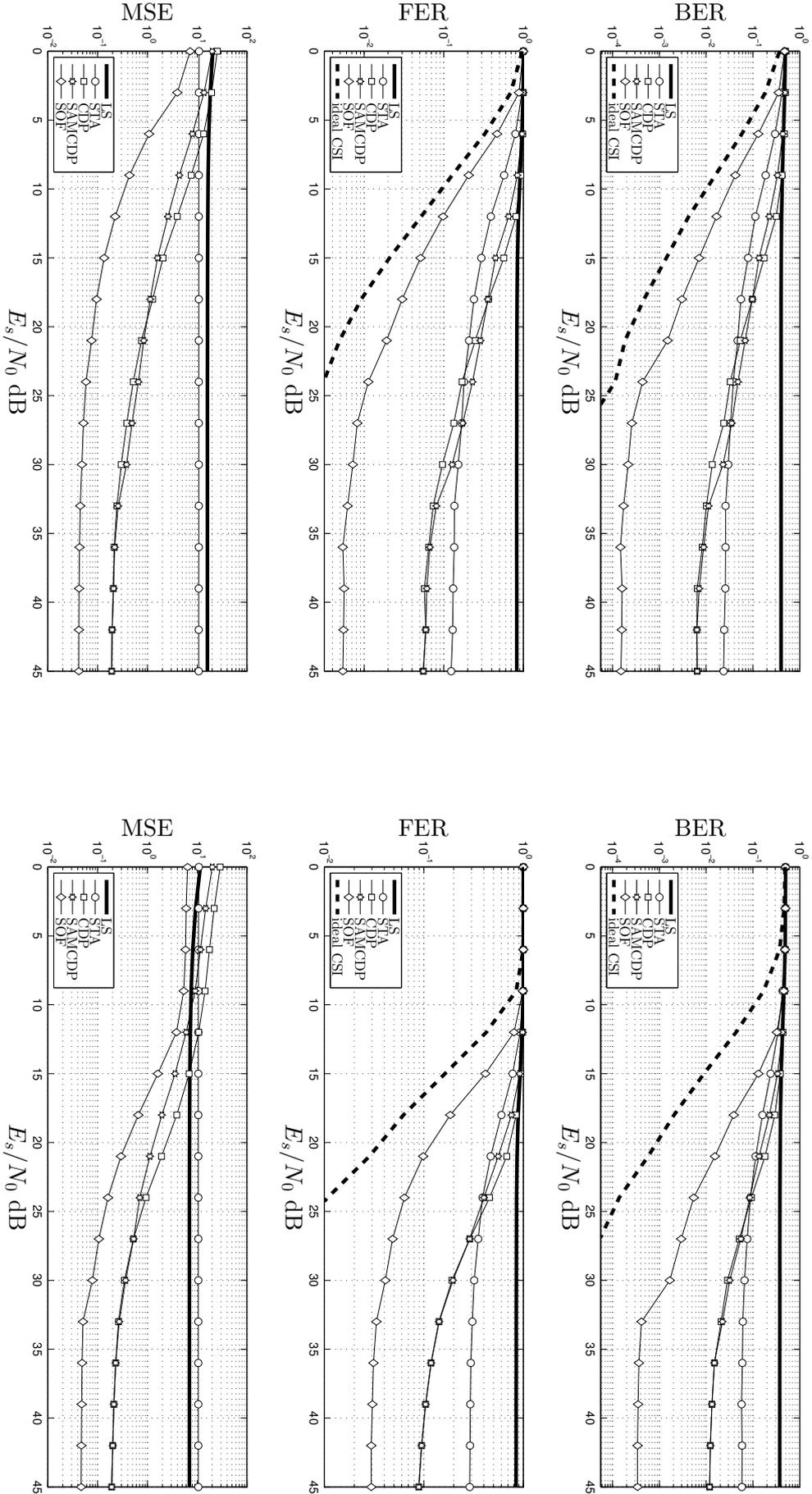


Figure 2.12: Comparative BER, FER & MSE results of V2V-EW0 @ 300B payload. Format: left column - QPSK-modulated symbols (i.e. 50 symbols frames); right column - 16QAM-modulated symbols (i.e. 25 symbols frames). ©2017 IEEE

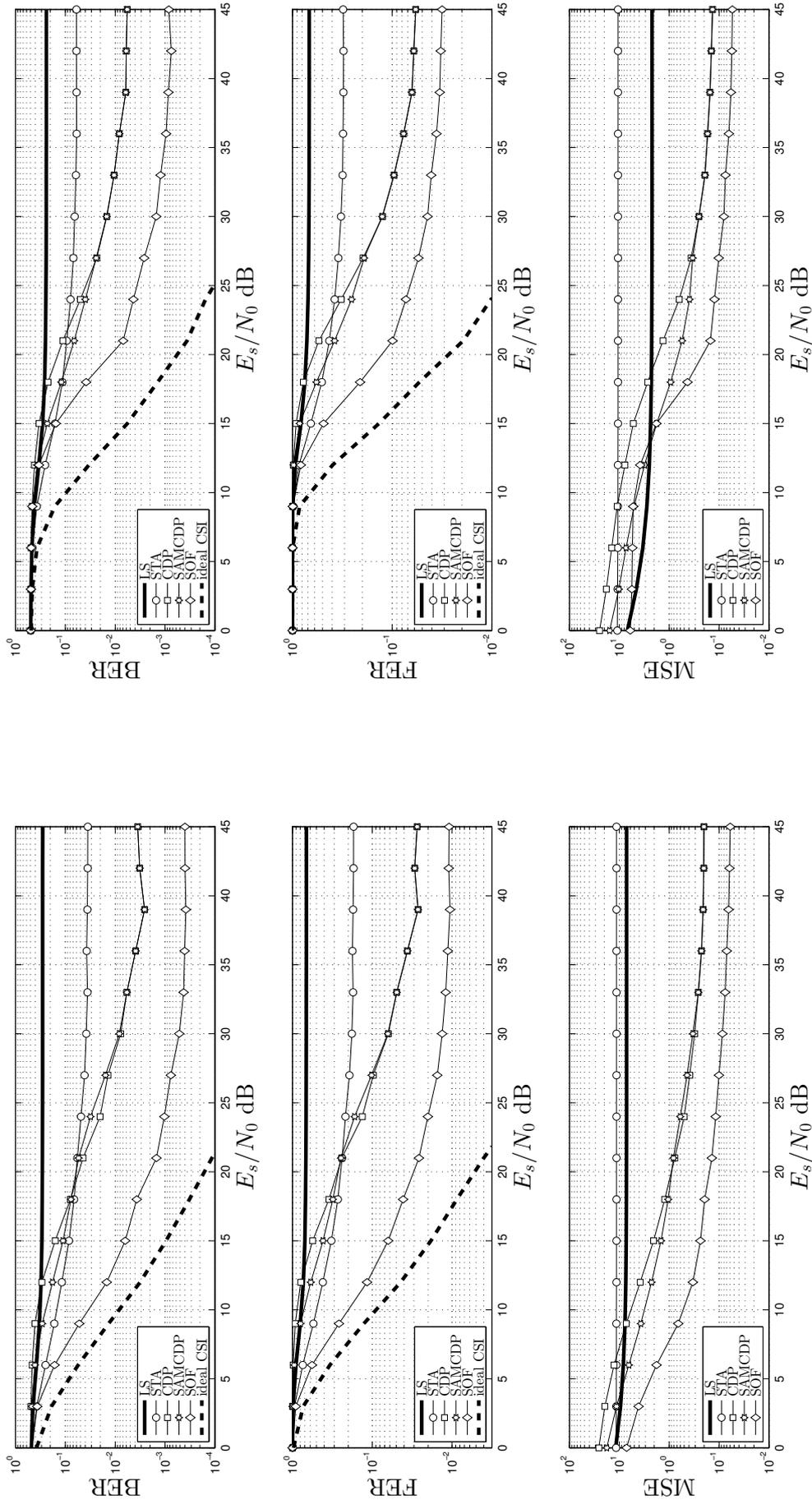


Figure 2.13: Comparative BER, FER & MSE results of V2IUC @ 300B payload. Format: left column - QPSK-modulated symbols (i.e. 50 symbols frames); right column - 16QAM-modulated symbols (i.e. 25 symbols frames). ©2017 IEEE

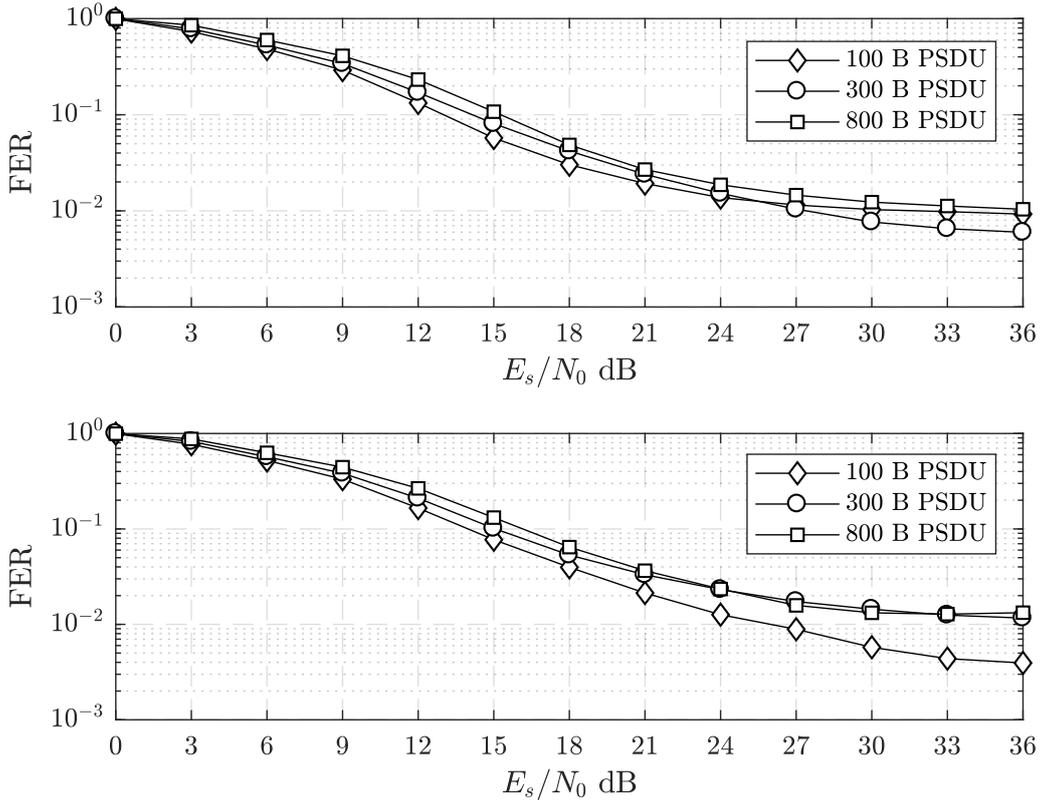


Figure 2.14: Comparative FER results of V2V_{EWO} and V2I_{UC} @ 100B, 300B, 800B payloads. Format: upper row - QPSK-modulated symbols simulated using V2V_{EWO}; lower row - QPSK-modulated symbols simulated using V2I_{UC}.

LS estimates based on *perfect CSI knowledge* for any OFDM symbol in any packet is provided as well. Based on these considerations and the plots from Figures 2.12 - 2.15, it is visible that the SOF obtains good systematic performance outperforming the state of the art and approaching the ideal equalization performance in the practical low to medium SNR regions across all different types of modulations, payloads and challenging environments considered. This suggests that the adaptive spectral and temporal filtering associated with the channel estimation side of the SOF is performing as intended by proper adjustment to V2X channels, and respectively, to IEEE 802.11p transmissions attributes. This performance is furthermore attainable at comparable complexity order to any of the other methods since the processing scales linearly with the packet length for all the considered methods as the dominant estimation and equalization procedures are reduced to simple LS. Moreover, inspecting the BER curves of the simple LS without tracking the channel evolution it is clear that relying on the initial channel estimates alone detection is practically infeasible in the rapid time-varying V2X environments.

From a practical standpoint, the less than 10 % frame error rate (FER) cutoff point usually aimed for in practice is attained by the SOF receiver within 3 dB of SNR distance relative to the ideal CSI based detection. Nonetheless, despite the demonstrated channel estimation and equalization BER performance of the proposed SOF method throughout the simulations of Figures 2.12 - 2.15, it is also clear that the receiver is still limited relative to the ideal CSI case. In particular, the proposed scheme exhibits still an error floor for the high SNRs which

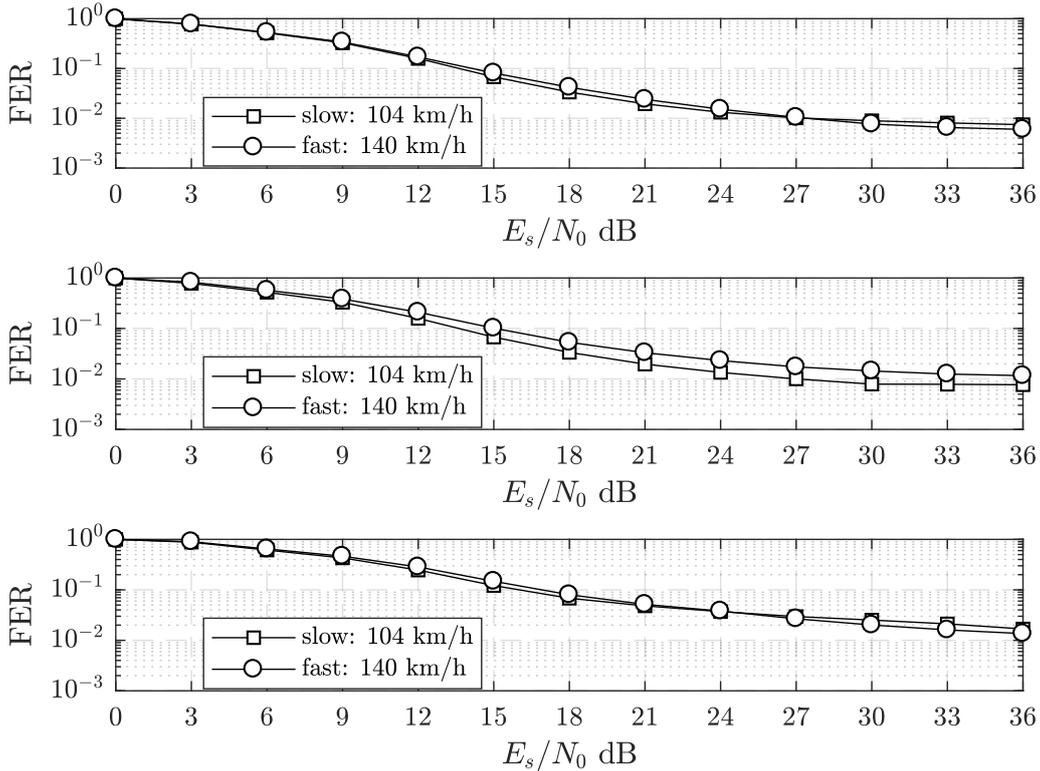


Figure 2.15: Comparative FER results of V2V_{EWO}, V2I_{UC} and V2I_{SS} @ *slow* and *fast* speeds. Format: upper row - QPSK-modulated symbols simulated using V2V_{EWO}; middle row - QPSK-modulated symbols simulated using V2I_{UC}; lower row - QPSK-modulated symbols simulated using V2I_{SS}.

underscores the fact that despite its processing efforts in tracking accurately the wireless channel the SOF receiver is unable to do so perfectly. In order to investigate the cause of this problem and to explain the observed asymptotic phenomena exhibited in the Figures 2.12 - 2.15, the transmission of several packets was closely monitored per OFDM symbol and channel evolution. These results are displayed in Figure 2.16.

The careful analysis of the errors has shown that systematically most of them are caused by local minima regions in the CSI magnitude as these are the points most vulnerable to noise, and hence, have the lowest subcarrier instantaneous SNRs. These valleys in the frequency response magnitude are a consequence of the multipath propagation and its subsequent frequency selectivity typical for V2X channels [41, 45]. Their variation over time comes however as an effect of the time variation and time-selectivity of the wireless media incurred under the mobile environments [50]. As a result, at the points where the channel is weak and poorly estimated the equalization procedure crosses the constellation symbol detection boundaries leading to erroneous detections.

Even though some of these erroneous symbols are corrected by the error-correction decoding embedded in the feedback mechanism of the SOF estimator, this may still prove to be insufficient. As illustrated in Figure 2.16, especially when the channel estimates are not accurate enough and the residual estimation noise is significant, it is a considerable risk to be unable to

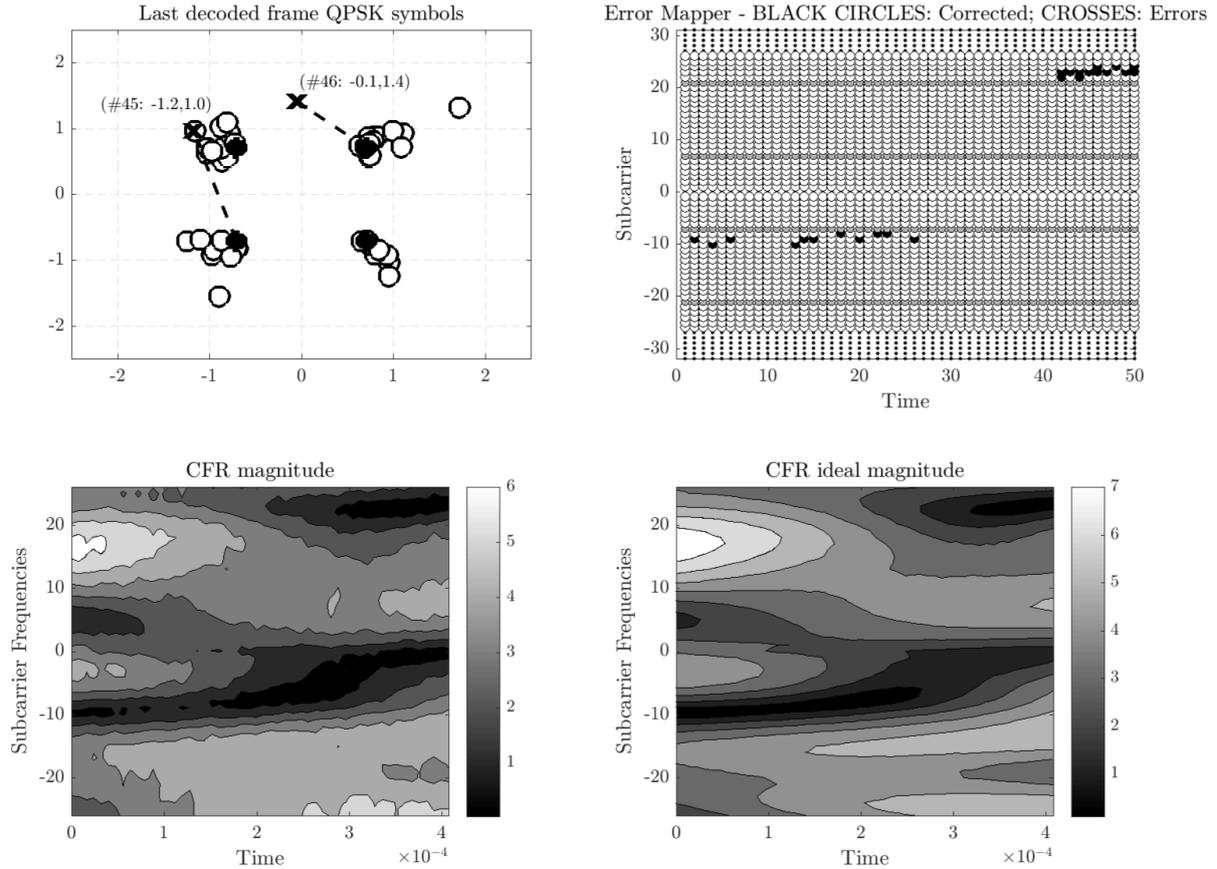


Figure 2.16: Time-frequency analysis of error sources — an example of V2I_{SS} @ 18 dB SNR, 120 km/h. Note: *white* dots are correctly decoded symbols, *black* are error corrected symbols and *crosses* are erroneous symbols. Format: upper left corner - equalization of last erroneous OFDM symbol; upper right corner - error mapper in time-frequency domain; lower left corner - SOF estimated CFR contour plot over the entire packet; lower right corner - ideal CSI contour plot over the entire packet.

properly equalize the transmitted symbols. Having multiple such valleys raises the chance of saturating the error-correction capability of the convolutional code used by the IEEE 802.11p standard. Failing to correct the OFDM data symbols leads in turn to inaccurate CFR estimates that may affect future symbols unless the subband SNR does not improve. These phenomena are compactly illustrated in Figure 2.16 for a realization of a 50 OFDM symbols packet radiated for a duration of 0.4 ms over the V2I_{SS} channel.

The observed BER performance and high SNR behavior suggests therefore that under the IEEE 802.11p parameters there are not enough levels of resolution to fully resolve the channel estimation under simple sequential processing of the OFDM symbols. This is a known fact given the 10 MHz reduced bandwidth of the scheme which is not able to distinguish in time-domain between individual propagation paths, and thus, this also results in a coarser CFR representation in comparison to wider bandwidth systems [47, 50]. To solve this problem, more involved schemes may rely on iteratively processing each OFDM symbol and/or packet, as for instance the Turbo receivers [58, 63] based on iterative decoding or the space-alternating generalized expectation-maximization (SAGE) receivers [64] based on the iterative space-leveraged

expectation-maximization (EM) algorithm. However, the latter introduce processing delay and their complexity may not be suitable to cheap commodity radios that can be embedded not only in cars, but also in other processing environments with less computational resources.

In practice, the early and current commercial off-the-shelf (COTS) equipment demonstrated only up to 6 – 12 Mbps communication rates under levels of reliability appropriate for latency-sensitive V2X in practical environments [65, 66], with the caveat that such rates are actually achievable in controlled experiments for 1 user exclusive access. Under the orthogonal TDMA access of the IEEE 802.11p however for multiple vehicles in a crowded traffic scenario the expected rates are much lower being usually limited to 100 – 500 Kbps [66]. Albeit the fact that such rates are sufficient for short cooperative awareness messages, they limit severely the applications that could later rely on the IEEE 802.11p protocol. For instance, under such low rates, the distribution of high definition local dynamic maps, temporal navigation layers and metadata in Cooperative Intelligent Transportation Systems (C-ITS) is restricted.

Furthermore, even though the slotted TDMA access of the protocol ensures access to an upper bound of 100 users served within 1 s, this capacity is severely lowered by practical mobile scenarios. For instance the US Department of transportation’s Technical Report [66] states that in fact under very common velocities of about 60 km/h the expected user capacity rate is of about 32 users served in fact. These combined low-performance figures in terms of *achievable BER*, *achievable rates*, and respectively, *achievable user access capacity* outline how even some of the most promising access schemes, *i.e.* OFDM, can be fundamentally challenged by real deployments at scale. Additionally, it is clear, based on the above exposition, that such effects are in fact rooted by the PHY access and waveforms that act as information carriers for the upper layers.

2.4 Conclusions

In the light of the recent case study from Section 2.3, one may think to solve the conundrum of V2X communications by improving upon the IEEE 802.11p system. And, in fact, some research has been devoted to this goal, as the SOF proposed receiver or the previously described methods throughout Subsection 2.3.1. Addressing this issue from a communications systems perspective, there will always be varying opinions of what can be done. For instance, the RF engineer would propose antenna diversity and MIMO processing on top of IEEE 802.11p, the coding theorist would improve the error-correction to state-of-the-art Turbo codes, whereas the wireless engineer would focus instead on improving the detection and access problems. Essentially, all the above solutions are yet limited by the IEEE 802.11p standard numerology and core OFDM-based constraints.

However, under a wider perspective encompassing also the introduction of this chapter, one may notice that historically all communication systems and access schemes have been inherently developed on the premises of orthogonality. Furthermore, all of them have failed to singularly scale or unboundedly adapt to the capacity and access problems users always increasingly imposed on them. As a result, a more fundamental worthwhile research question

should be *what is the limiting factor and how can we fundamentally escape it ?*

An intuitive short answer could follow as ...*orthogonality*. This is based upon the fact that orthogonal signal processing will always be constrained to preserving this innate principle, and thus, limited in scaling beyond what the physical signal space limits are. To escape this constraint, a natural approach should be in fact based on *relaxing the orthogonality principle* thus leading to *non-orthogonal signal processing*.

An immediate example of the latter could be provided in the form of modern CS techniques [67]. Analogously to the above idea, CS challenged the classical sampling theorem of Nyquist. Concretely, the idea behind CS being that for a signal that has bandwidth W but holds most of its information under a small fraction of that bandwidth its sampling can easily be applied under the Nyquist rate with still high probability of super-accurate reproduction of the original signal. The same principle is further applied to other fields such as data compression, image compression, video and signal processing to name a few.

From a mathematical standpoint this core idea of relaxing *orthogonality* is in other words equivalent to an extrapolation of orthogonal basis representations. The latter is in fact the field of study for *Frame Theory* [19, 20, 22, 23]. As an eclectic mathematical concept, Frame Theory has roots in the mathematical areas of harmonic and functional analysis, operator theory, linear algebra and matrix theory [19]. However, to the extent of this research, the focus shall be placed on *Finite Frame Theory* and its *applications to next-generation wireless non-orthogonal signal processing* in terms of waveform sensing and access paradigms.

Chapter 3

Non-orthogonality and Frame Theory

Having concluded Chapter 2 with the main motivations of wireless communications leading to non-orthogonal signal processing, the core mathematical background is detailed next. This has been developed during the last 30-40 years as a coherent mathematic formalism under the name of *Frame Theory*, [19,20], following the fundamental work of Daubechies, Grossman and Mayer in [68].

Nonetheless, quite interestingly, the core concept of non-orthogonality and therefore *redundant representations* can be traced back to Dennis Gabor's work related to the mathematical formulation of wireless communication systems [69–71]. Gabor researched the idea of a natural time-frequency description of information encoded and radiated over the air. Furthermore, a comprehensive analysis of the biological decoding of speech following the same organic time-frequency principles was also underlined from the Rx perspective. Gabor proposed in his work [69–71], unwillingly however at the time, a functional harmonic primitive set as an extrapolation of biorthogonal harmonic decompositions. This mathematical construct became later known under the umbrella of Frame Theory as the *Gabor frames* [19,22].

However, as Gabor frames are in fact specific time-frequency representations over infinite spaces, they actually are distant to the the scope of this dissertation, and therefore will not be discussed in further mathematical details. As a matter of fact, the focus is hereby placed on the finite-dimensional decompositions pertaining to finite Frame Theory [19].

3.1 Preliminaries

To bridge the gap between previously discussed orthogonal representations and their non-orthogonal frame-theoretic extensions, let us first take a short detour through the realm of biorthogonal projections. Consider next therefore the generic finite M -dimensional Hilbert

space¹ \mathbb{H}^M [72]. For wireless communications systems the Hilbert space \mathbb{H}^M is often reduced to its real, *i.e.*, \mathbb{R}^M , or complex, *i.e.*, \mathbb{C}^M , equivalents respectively. However, over the course of this chapter we maintain the generic \mathbb{H}^M notation for the sake of generality, and consider similarly \cdot^{H} to denote the Hermitian adjoint operation [72] over \mathbb{H}^M , equivalent to transposition for $\mathbb{H} = \mathbb{R}$, and respectively, conjugate transposition for $\mathbb{H} = \mathbb{C}$.

The generic goal of a robust signal processing representation is to provide an alternative signal description transforming any original \mathbf{x} to a vector of signal coefficients \mathbf{c} storing the same information, but robustly encoded against distortions. An instance of such a transformation is in fact the FT detailed previously in the case of OFDM. This provides the link between time and frequency domains adding for instance a better resistance against AWGN in its frequency representation relative to the time counterpart. The essence of Frame Theory is constituted in fact by the provision of a general description of the relationship between the signal \mathbf{x} representation and its subsequent coefficient \mathbf{c} domain.

Assume next that the two reciprocal mappings of information, *i.e.*, signal to coefficients $\mathbf{x} \mapsto \mathbf{c}$ (*the analysis operation*), and respectively, coefficients to signal $\mathbf{c} \mapsto \mathbf{x}$ (*the synthesis operation*), are linearly performed for $\mathbf{x} \in \mathbb{H}^M$, $\mathbf{c} \in \mathbb{H}^K$. To this extent, consider the analysis as being fully described by the set of K vectors $\{\mathbf{f}_k\}$, such that $\mathbf{f}_k \in \mathbb{H}^M$, and respectively,

$$c_k = \langle \mathbf{x}, \mathbf{f}_k \rangle \triangleq \mathbf{f}_k^{\text{H}} \mathbf{x}, \quad \forall k. \quad (3.1)$$

Stacking up the individual coefficients c_k in the K -dimensional vector \mathbf{c} and vectorizing Equation (3.1) the identity above may be compactly rewritten as

$$\mathbf{c} = \mathbf{F}^{\text{H}} \mathbf{x}, \quad (3.2)$$

with

$$\mathbf{F} \triangleq [\mathbf{f}_1 \quad \mathbf{f}_2 \quad \dots \quad \mathbf{f}_K] \quad (3.3)$$

implicitly defined as the *analysis matrix (linear operator)*.

Given (3.2), for the case that $K = M$ and \mathbf{F} is unitary, it results given (3.3) that the synthesis operation of recovering the original signal vector \mathbf{x} from the analysis coefficients \mathbf{c} is nothing else than the reciprocal operation, *i.e.*

$$\mathbf{x} = (\mathbf{F}^{\text{H}})^{-1} \mathbf{c} = \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{x}, \quad (3.4)$$

where $(\mathbf{F}^{\text{H}})^{-1} = \mathbf{F}$ is the unique *synthesis linear operator* under such an orthonormal basis scenario.

From an abstract dual perspective, the synthesis operation above may be described in terms of another set of K vectors $\{\tilde{\mathbf{f}}_k\}$ such that $\tilde{\mathbf{f}}_k \in \mathbb{H}^M$, $\forall k$. Relaxing the orthogonality principle on the vectors $\{\mathbf{f}_k\}$ above, equivalent to \mathbf{F}^{H} being non-unitary anymore, implies the fact that

¹A short review of linear algebra topics sufficient for a self-sustained comprehension of the manuscript is provided in Appendix A.

the synthesis operation cannot be achieved simply by means of a matched filtering with the left-adjoint multiplication as in Equation (3.4). Nevertheless, since $\{\mathbf{f}_k\}$ spans the space \mathbb{H}^M , the analysis matrix \mathbf{F}^H is in fact full-rank and square, recall that the assumption $K = M$ is nonetheless present, so that it still has a unique matrix inverse $(\mathbf{F}^H)^{-1}$. As a result of this relaxation, the synthesis operation from before must be rewritten in terms of an additional *dual spanning set* $\{\tilde{\mathbf{f}}_k\}$ such that

$$\mathbf{x} = (\mathbf{F}^H)^{-1} \underbrace{\mathbf{F}^H \mathbf{x}}_{\mathbf{c}} = \sum_{k=1}^K \langle \mathbf{x}, \mathbf{f}_k \rangle \tilde{\mathbf{f}}_k. \quad (3.5)$$

In the light of Equation (3.5), $\{\tilde{\mathbf{f}}_k\}$ is consequently the *unique dual set* of $\{\mathbf{f}_k\}$, where the vectors $\tilde{\mathbf{f}}_k$ are the columns of $(\mathbf{F}^H)^{-1}$. Furthermore, upon the linear inverse definition

$$\langle \tilde{\mathbf{f}}_k, \mathbf{f}_\ell \rangle = \begin{cases} 1, & k = \ell \\ 0, & k \neq \ell \end{cases}, \quad (3.6)$$

which makes the set pair $(\{\tilde{\mathbf{f}}_k\}, \{\mathbf{f}_k\})$ *biorthonormal*. As a result, in this case the synthesis is not a combined result of the orthogonal contributions from the individual projections onto $\{\mathbf{f}_k\}$ itself, but in fact of the linearly combined dual set contributions $\{\tilde{\mathbf{f}}_k\}$ weighted properly by the analysis coefficients of \mathbf{x} with their dual orthogonal pairs, *i.e.*, $\{\mathbf{f}_k\}$ as detailed mathematically by (3.5). In other words, simply put for $K = M$:

- orthonormality implies that $\{\tilde{\mathbf{f}}_k\} \triangleq \{\mathbf{f}_k\}$, and moreover, conditions (3.6) hold $\forall k$;
- biorthonormality resumes to finding the unique inverse and its corresponding column vectors $\{\tilde{\mathbf{f}}_k\}$ given any linearly dependent set $\{\mathbf{f}_k\}$.

The biorthonormal relaxation comes however at a price. To better visualize the latter consider the energy of the projected signal in the coefficients domain. On the one hand, for the orthonormal case, the identity

$$\|\mathbf{c}\|_2^2 = \langle \mathbf{c}, \mathbf{c} \rangle = \mathbf{c}^H \mathbf{c} = \mathbf{x}^H \mathbf{F} \mathbf{F}^H \mathbf{x} = \mathbf{x}^H \mathbf{I}_M \mathbf{x} = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|_2^2 \quad (3.7)$$

holds given the fact that \mathbf{F}^H is unitary and as a result the combined analysis and synthesis operations are *energy preserving*.

On the other hand, in the biorthonormal case

$$\|\mathbf{c}\|_2^2 = \langle \mathbf{c}, \mathbf{c} \rangle = \mathbf{c}^H \mathbf{c} = \mathbf{x}^H \mathbf{F} \mathbf{F}^H \mathbf{x} \quad (3.8)$$

such that $\mathbf{F} \mathbf{F}^H \neq \mathbf{I}_M$.

In plain words, Equation (3.8) outlines the fact that the analysis and synthesis operations over biorthonormal bases are *not norm preserving*. In fact, the latter comes as a corollary of Rayleigh-Ritz Theorem [43] summarized below.

Theorem 3.1.1 (Rayleigh-Ritz Theorem). *Let \mathbf{A} be a $M \times M$ self-adjoint matrix and \mathbf{x} a M -dimensional vector with its real eigenvalues ordered as*

$$\lambda_{\min} \equiv \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N \equiv \lambda_{\max}. \quad (3.9)$$

Then for any $\mathbf{x} \in \mathbb{H}^M$

$$\lambda_{\min} \|\mathbf{x}\|_2^2 \leq \mathbf{x}^H \mathbf{A} \mathbf{x} \leq \lambda_{\max} \|\mathbf{x}\|_2^2. \quad (3.10)$$

Proof: The result is immediate and follows upon the fact that \mathbf{A} is diagonalized by unitary transforms such that $\mathbf{A} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H$. As a consequence, $\mathbf{x}^H \mathbf{A} \mathbf{x}$ is equivalent to $\mathbf{x}^H \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H \mathbf{x}$ which is furthermore identical to

$$\lambda_{\min} \cdot \mathbf{x}^H \mathbf{Q} \mathbf{Q}^H \mathbf{x} \leq \mathbf{x}^H \mathbf{A} \mathbf{x} \leq \lambda_{\max} \cdot \mathbf{x}^H \mathbf{Q} \mathbf{Q}^H \mathbf{x}. \quad (3.11)$$

Given the fact that \mathbf{Q} is unitary, the conclusion immediately follows under the above inequalities. \square

In the light of the above theorem, it is thus clear that the relaxed biorthonormal representation leads to a coefficient decomposition \mathbf{c} of any \mathbf{x} such that

$$\lambda_{\min} \|\mathbf{x}\|_2^2 \leq \|\mathbf{c}\|_2^2 = \mathbf{x}^H \mathbf{F} \mathbf{F}^H \mathbf{x} \leq \lambda_{\max} \|\mathbf{x}\|_2^2. \quad (3.12)$$

where λ_{\min} and λ_{\max} represent respectively the lowest and largest eigenvalues of the matrix $\mathbf{F} \mathbf{F}^H$.

Furthermore, since $\mathbf{F} \mathbf{F}^H$ is itself full-rank and bounded, *i.e.*, $\|\mathbf{F} \mathbf{F}^H\|_{\mathbb{F}} < \infty$, it follows that the eigenvalues themselves are bounded and positive, *i.e.*, $0 < \lambda_{\min} \leq \lambda_{\max} < \infty$. Thus, the biorthonormal relaxation described above comes at a price of not preserving the energy of all the signals \mathbf{x} between the signal and coefficient domains albeit its direct and reciprocal transforms are *norm-bounded* as described in (3.12). This in turn incurs the fact that there may exist $\mathbf{x} \in \mathbb{H}^M$ such that the energy of their coefficient representation \mathbf{c} is lower relative to their original signal in comparison to others.

This latter remark suggests the fact that albeit providing more degrees of freedom, biorthonormal bases may lower the signal representational energy which could affect the analysis and synthesis operations in noisy or lossy environments such as the ones of wireless communications systems. To address this weakness, the additional requirement for increased signal robustness needs to be complementary fulfilled. And, as a matter of fact, this comes as an extra relaxation under the form of $K > M$.

3.2 Frame Theory Basics

As outlined earlier, relaxing the orthogonality to biorthonormal bases may provide some energy deficiencies which need to be accounted for from a signal processing perspective by increased redundancy. In practical systems this is often the case that due to noise, to non-

ideal mathematical models, to physical distortions or to non-recoverable impairments which are superimposed over the signal space. Therefore, it would be highly desirable if one could express a signal \mathbf{x} in terms of a redundant decomposition such as even if some signal coefficients in \mathbf{c} are corrupted beyond recovery one can still reconstruct the original signal. This is a common practice for instance in channel and source coding where redundancy is introduced as a mechanism for coping with forward channel effects during transmissions and at the Rx end. Nevertheless, a similar strategy can be used at the waveform level on the raw signals in order to provide the necessary mechanisms and failure domain safeties to redundantly represent transmission signals for robust and efficient communications in harsh environments (*e.g.*, indoor offices, indoor industrial hangars, urban canyons, time-varying mobile fading channels like the ones discussed in the IEEE 802.11p case study from Section 2.3 *etc.*).

This latter idea is the core motivation of Frame Theory and its applicability to modern communication systems' problems. To this end, expanding the K -dimensional coefficient space such that $K > M$ is employed hereafter. This leads to a completely non-orthogonal but redundant representational framework of the signal space [19, 22].

Let us reinspect the synthesis operation since the analysis transform depicted in Equation (3.2) remains essentially the same. It should be already clear at this point that biorthonormal bases are in fact a subset of redundant bases, whereas orthogonal bases are a subset of the biorthonormal ones. As a consequence, it follows that the synthesis operation in case of the redundant bases implied by $K > M$ is in fact mathematically similar to the case of biorthonormal bases, *i.e.*

$$\mathbf{x} = \sum_{k=1}^K \langle \mathbf{x}, \mathbf{f}_k \rangle \tilde{\mathbf{f}}_k = \sum_{k=1}^K c_k \tilde{\mathbf{f}}_k, \quad (3.13)$$

with the key distinction that the dual set $\tilde{\mathbf{f}}_k$ cannot be obtained anymore uniquely by simple inversion as $\mathbf{F}^H \in \mathbb{H}^{M \times K}$ is a non-square, rectangular matrix, and hence, non-invertible.

Under the above considerations, the perfect reconstruction of \mathbf{x} is obtained by means of the Moore-Penrose left pseudoinverse [43] of \mathbf{F}^H , namely

$$\mathbf{F}^\dagger \triangleq (\mathbf{F}\mathbf{F}^H)^{-1} \mathbf{F} \quad (3.14)$$

iff the columns of \mathbf{F} , *i.e.*, $\{\mathbf{f}_k\}$ span the space \mathbb{H}^M .

Nonetheless, the caveat is represented by the non-unique representation. In details, the left pseudoinverse is not unique [73, Ch. 2, Th. 2] being easily parameterizable, as for instance

$$\mathbf{F}_B = \mathbf{F}^\dagger + \mathbf{B} \left(\mathbf{I}_K - \mathbf{F}^H \mathbf{F}^\dagger \right), \quad (3.15)$$

for any generic matrix $\mathbf{B} \in \mathbb{H}^{M \times K}$.

The above parameterization of the left pseudoinverse is immediate by right multiplication with \mathbf{F}^H . In addition, considering all of the above it is clear that a family of dual sets $(\{\mathbf{f}_k\}, \{\tilde{\mathbf{f}}_k\})$ may be thus constructed starting from $\{\mathbf{f}_k\}$ where $\{\tilde{\mathbf{f}}_k\}$ are nothing but the columns of the parameterized left pseudoinverse from Equation (3.15). The *canonical left pseu-*

doinverse given by Equation (3.14) generates the *canonical dual set* $\{\tilde{\mathbf{f}}_k\}$ such that

$$\tilde{\mathbf{f}}_k = (\mathbf{F}\mathbf{F}^H)^{-1} \mathbf{f}_k, \quad \forall k. \quad (3.16)$$

Definition 3.2.1. A set of vector elements $\{\mathbf{f}_k\}$, such that $\mathbf{f}_k \in \mathbb{H}^M, k = \{1, 2, \dots, K\}$ is a frame for the Hilbert space \mathbb{H}^M , compactly denoted as $\mathbf{F} \triangleq [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K]$ if

$$\alpha \|\mathbf{x}\|^2 \leq \|\mathbf{c}\|_2^2 = \sum_{k=1}^K |\langle \mathbf{x}, \mathbf{f}_k \rangle|^2 = \|\mathbf{F}^H \mathbf{x}\|_2^2 \leq \beta \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{H}^M, \quad (3.17)$$

$\alpha, \beta \in \mathbb{R}$, such that $0 < \alpha \leq \beta < \infty$ are valid scalars called generically frame bounds.

The largest lower bound α and the smallest upper bound β represent the (tightest possible) frame bounds. The set $\{\tilde{\mathbf{f}}_k\}$ associated with the frame $\{\mathbf{f}_k\}$ which leads to the reconstruction of any \mathbf{x} as

$$\mathbf{x} = \sum_{k=1}^K \langle \mathbf{x}, \mathbf{f}_k \rangle \tilde{\mathbf{f}}_k \quad (3.18)$$

is the dual frame of $\{\mathbf{f}_k\}$ for the Hilbert space \mathbb{H}^M where $K \geq M$.

Needless to say that similar to the case of biorthonormal bases, the frame bounds are actually related to the spectral decomposition of the matrix $\mathbf{F}\mathbf{F}^H$. To this extent, it follows that $\alpha = \lambda_{\min}$, *i.e.*, the smallest eigenvalue of $\mathbf{F}\mathbf{F}^H$, and respectively, $\beta = \lambda_{\max}$, *i.e.*, the largest eigenvalue of $\mathbf{F}\mathbf{F}^H$. On the one hand, as for the biorthonormal case earlier discussed, a strictly positive lower frame bound, *i.e.*, $\alpha > 0$ ensures that in fact a left pseudoinverse exists for any coefficient linear transform \mathbf{F}^H . The same bound assures that additionally the completeness of a frame set since the only way $\|\mathbf{c}\|_2^2 = 0$ is for \mathbf{x} itself to be $\mathbf{x} = \mathbf{0}$. On the other hand, the finite upper bound, *i.e.* $\beta < \infty$ for any $\mathbf{x} \in \mathbb{H}^M$ implies the fact that the analysis linear operation described by the frame vectors is bounded², and consequently, continuous³ [23].

As far as original signal recovery of \mathbf{x} is concerned given its signal coefficients \mathbf{c} , the reconstruction procedure as seen above is not unique to frame \mathbf{F} , but unique to a fixed dual frame set pair $(\{\mathbf{f}_k\}, \{\tilde{\mathbf{f}}_k\})$, where the latter dual frame $\{\tilde{\mathbf{f}}_k\}$ has been constructed based on the core assumptions described throughout Equations (3.13) - (3.15). However, since the computation of the pseudoinverse is resource expensive and potentially numerically unstable, a cohort of iterative approximation based algorithms have been developed as alternatives. For instance, one simple such algorithm is the *Frame Algorithm* [19, Prop. 1.18] summarized below.

Proposition 3.2.1 (Frame Algorithm). Let $\{\mathbf{f}_k\}$ be frame \mathbf{F} of K vectors in the Hilbert space \mathbb{H}^M with frame bounds α, β . Then any signal $\mathbf{x} \in \mathbb{H}^M$ can be reconstructed from its frame coefficients' representation $\mathbf{c} = \mathbf{F}^H \mathbf{x}$ given the following iterative steps

$$\hat{\mathbf{x}}^{(i)} = \hat{\mathbf{x}}^{(i-1)} + \frac{2}{\alpha + \beta} \cdot \mathbf{F} \left(\mathbf{c} - \mathbf{F}^H \hat{\mathbf{x}}^{(i-1)} \right), \quad (3.19)$$

²A linear operator $A : \mathbb{H} \rightarrow \mathbb{H}'$ is bounded if $\exists c < \infty$ s.t. $\forall x \in \mathbb{H}, \|Ax\| \leq c\|x\|$.

³A linear operator $A : \mathbb{H} \rightarrow \mathbb{H}'$ is continuous for $x_0 \in \mathbb{H}$ if $\forall \epsilon > 0 \exists \delta > 0$ s.t. $\forall x \in \mathbb{H}, \|x - x_0\| < \delta$ then $\|Ax - Ax_0\| < \epsilon$. A is continuous over \mathbb{H} if A continuous point-wise $\forall x_0 \in \mathbb{H}$.

where $\hat{\mathbf{x}}^{(i)}$ denotes the estimate at the i -th iteration for all $i \geq 0$ and $\hat{\mathbf{x}}^{(0)} \triangleq \mathbf{0}$.

The sequence $\hat{\mathbf{x}}^{(i)} \in \mathbb{H}^M, i \geq 0$ is guaranteed to converge to $\mathbf{x} \in \mathbb{H}$ at the rate of

$$\|\mathbf{x} - \hat{\mathbf{x}}^{(i)}\|_2 \leq \left(\frac{\beta - \alpha}{\alpha + \beta}\right)^i \|\mathbf{x}\|_2, \quad i \geq 0. \quad (3.20)$$

Proof [19, Prop. 1.18]: For any $\mathbf{x} \in \mathbb{H}^M$,

$$\left\langle \mathbf{I}_M - \frac{2}{\alpha + \beta} \mathbf{F} \mathbf{F}^H \mathbf{x}, \mathbf{x} \right\rangle = \|\mathbf{x}\|_2^2 - \frac{2}{\alpha + \beta} \sum_{k=1}^K |\langle \mathbf{x}, \mathbf{f}_k \rangle|^2 \leq \|\mathbf{x}\|_2^2 - \frac{2\alpha}{\alpha + \beta} \|\mathbf{x}\|_2^2 = \frac{\beta - \alpha}{\alpha + \beta} \|\mathbf{x}\|_2^2, \quad (3.21)$$

where the inequality is a consequence of the frame definition given (3.17).

Analogously, it holds that

$$\frac{\beta - \alpha}{\alpha + \beta} \|\mathbf{x}\|_2^2 \leq \left\langle \mathbf{I}_M - \frac{2}{\alpha + \beta} \mathbf{F} \mathbf{F}^H \mathbf{x}, \mathbf{x} \right\rangle \quad (3.22)$$

and so it follows that

$$\left\| \mathbf{I}_M - \frac{2}{\alpha + \beta} \mathbf{F} \mathbf{F}^H \right\| \leq \frac{\beta - \alpha}{\alpha + \beta}. \quad (3.23)$$

Considering an equivalent form of the iteration (3.19) as

$$\hat{\mathbf{x}}^{(i)} = \hat{\mathbf{x}}^{(i-1)} + \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H (\mathbf{x} - \hat{\mathbf{x}}^{(i-1)}), \quad (3.24)$$

it results that

$$\mathbf{x} - \hat{\mathbf{x}}^{(i)} = \left(\mathbf{I} - \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H \right) (\mathbf{x} - \hat{\mathbf{x}}^{(i-1)}) \quad (3.25a)$$

$$= \left(\mathbf{I} - \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H \right)^i (\mathbf{x} - \hat{\mathbf{x}}^{(0)}) \quad (3.25b)$$

$$= \left(\mathbf{I} - \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H \right)^i \mathbf{x} \quad (3.25c)$$

The convergence rate follows as

$$\|\mathbf{x} - \hat{\mathbf{x}}^{(i)}\|_2 = \left\| \left(\mathbf{I} - \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H \right)^i \mathbf{x} \right\|_2 \quad (3.26a)$$

$$\leq \left\| \left(\mathbf{I} - \frac{2}{\alpha + \beta} \cdot \mathbf{F} \mathbf{F}^H \right) \right\|^i \|\mathbf{x}\|_2 \quad (3.26b)$$

$$\leq \left(\frac{\beta - \alpha}{\alpha + \beta} \right)^i \|\mathbf{x}\|_2 \quad (3.26c)$$

□

One can further intuitively observe that the reconstruction proposed by the Frame Algorithm is in fact reduced to an iterative weighted matched filter \mathbf{F} applied to the reconstruction residual, *i.e.* $\mathbf{c} - \mathbf{F}^H \hat{\mathbf{x}}^{(i-1)}$, during each iteration $i \geq 1$. Given the sensitivity of the convergence

rate (3.20) to the frame bounds α, β and the ratio $\frac{\beta-\alpha}{\alpha+\beta}$, it is clear that slow convergence may be exhibited for frames with increasingly higher displacement $\beta - \alpha$, which resumes to the spread of the eigenvalues of $\mathbf{F}\mathbf{F}^H$. To counteract this problem additional iterative frame reconstruction algorithms exist amending the core idea of the Frame Algorithm in order to improve the convergence rate. For instance, some of these algorithms amongst others [19] are the Chebyshev Algorithm [74] and the Conjugate Gradient Method [74].

3.3 Frame Transforms and Operators

Having laid down the foundations and basic concepts of Frame Theory, let us now extend the analytic perspective by introducing some operators meant to ease the discussion of different frame attributes and properties. Albeit explicitly formulated yet not formally defined, a first such linear transform is represented by the *frame operator*.

Definition 3.3.1 (The Frame Operator). *Let there be a frame \mathbf{F} containing K frame vectors $\{\mathbf{f}_k\} \in \mathbb{H}^M$, then the associated frame operator is defined as*

$$\mathbf{S} \triangleq \mathbf{F}\mathbf{F}^H \in \mathbb{H}^{M \times M}. \quad (3.27)$$

Rewriting the norm of the coefficients \mathbf{c} in terms of the frame operator \mathbf{S} , one obtains

$$\sum_{k=1}^K |\langle \mathbf{x}, \mathbf{f}_k \rangle|^2 = \mathbf{x}^H \mathbf{F}\mathbf{F}^H \mathbf{x} = \mathbf{x}^H \mathbf{S} \mathbf{x} = \langle \mathbf{S} \mathbf{x}, \mathbf{x} \rangle \quad (3.28)$$

and as a consequence the inequalities (3.17) of the frame definition can be easily recast as

$$\alpha \|\mathbf{x}\|_2^2 \leq \langle \mathbf{S} \mathbf{x}, \mathbf{x} \rangle \leq \beta \|\mathbf{x}\|_2^2, \quad (3.29)$$

and respectively,

$$\alpha \mathbf{I}_M \preceq \mathbf{S} \preceq \beta \mathbf{I}_M, \quad (3.30)$$

where the inequalities (3.30) have been obtained from (3.29) by rewriting the left and right bounds as $\langle \alpha \mathbf{I}_M \mathbf{x}, \mathbf{x} \rangle$, and respectively, $\langle \beta \mathbf{I}_M \mathbf{x}, \mathbf{x} \rangle$.

Based on the above, it follows then immediately that the matrices $(\mathbf{S} - \alpha \mathbf{I}_M)$ and $(\beta \mathbf{I}_M - \mathbf{S})$ are positive semi-definite, and as seen before by means of the Rayleigh-Ritz Theorem,

$$\lambda_{\min}(\mathbf{S}) \|\mathbf{x}\|_2^2 \leq \langle \mathbf{S} \mathbf{x}, \mathbf{x} \rangle \leq \lambda_{\max}(\mathbf{S}) \|\mathbf{x}\|_2^2, \quad (3.31)$$

where $\alpha = \lambda_{\min}(\mathbf{S})$ is the smallest eigenvalue of \mathbf{S} , and $\beta = \lambda_{\max}(\mathbf{S})$ is the largest eigenvalue of \mathbf{S} , respectively.

Moreover, the frame operator \mathbf{S} has a series of special structural attributes [19, 20, 23] summarized briefly as follows.

Proposition 3.3.1. *The frame operator \mathbf{S} satisfies in general the properties⁴:*

1. \mathbf{S} is linear and bounded;
2. $\mathbf{S}^H = \mathbf{S}$, i.e. \mathbf{S} is self-adjoint;
3. $\langle \mathbf{S}\mathbf{x}, \mathbf{x} \rangle > 0 \forall \mathbf{x} \in \mathbb{H}^M, \mathbf{x} \neq \mathbf{0}$, i.e. \mathbf{S} is positive definite;
4. $\mathbf{S} = \mathbf{S}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}$, i.e. \mathbf{S} has a unique self-adjoint, positive definite square root $\mathbf{S}^{\frac{1}{2}}$

Given Proposition 3.2.1 and the definition of the frame operator is now clear that \mathbf{S} plays an important role in the recovery characteristics of the original signal \mathbf{x} given its coefficients \mathbf{c} . Let us now reinterpret the reconstruction via the frame operator \mathbf{S} . As seen in (3.16), the canonical dual frame is given based on the inverse of the self-adjoint positive-definite \mathbf{S} , i.e.

$$\tilde{\mathbf{f}}_k = \mathbf{S}^{-1}\mathbf{f}_k. \quad (3.32)$$

Since $\{\tilde{\mathbf{f}}_k\}$ is the canonical dual frame of $\{\mathbf{f}_k\}$ over \mathbb{H}^M , the synthesis dual operator is obtained by (3.32) as

$$\tilde{\mathbf{F}} = \mathbf{S}^{-1}\mathbf{F}. \quad (3.33)$$

Considering now, the dual problem of starting with $\{\tilde{\mathbf{f}}_k\}$ as a frame over \mathbb{H}^M , it follows that the associated analysis operator in this case becomes

$$\tilde{\mathbf{F}}^H = (\mathbf{S}^{-1}\mathbf{F})^H = \mathbf{F}^H\mathbf{S}^{-1}, \quad (3.34)$$

where implicitly the fact that \mathbf{S} is self-adjoint has been leveraged in the derivation.

The following theorem stems from this point in the light of the canonical dual frame representation under \mathbf{S} .

Theorem 3.3.2. *Let $\{\mathbf{f}_k\}$ be a frame \mathbf{F} of K vectors over the Hilbert space \mathbb{H}^M with bounds α, β and frame operator \mathbf{S} . Assume $\{\tilde{\mathbf{f}}_k\}$ to be the canonical dual frame of $\{\mathbf{f}_k\}$ given by (3.32). Then, the following claims hold.*

1. The frame operator $\tilde{\mathbf{S}}$ associated to the canonical dual frame $\tilde{\mathbf{F}}$ out of $\{\tilde{\mathbf{f}}_k\}$ is called the canonical dual frame operator and satisfies

$$\tilde{\mathbf{S}} = \mathbf{S}^{-1}. \quad (3.35)$$

2. The set of K vectors $\{\tilde{\mathbf{f}}_k\}$ is itself a frame $\tilde{\mathbf{F}}$ for the Hilbert space \mathbb{H}^M with lower bound $\tilde{\alpha} = 1/\beta$ and upper bound $\tilde{\beta} = 1/\alpha$ and analysis operator $\tilde{\mathbf{F}}^H$ obtained via (3.32).
3. All the general properties of frame operators from Proposition 3.3.1 hold for \mathbf{S}^{-1} as well.

⁴The proofs of the claims are either immediately implied by the definition of the frame operator or trivially based on the already exposed equations and corollaries of well-known linear algebra results and theorems. Hence, they are skipped here, but the interested reader may refer for further details to [23, Th. 3.9]

Proof: Only the first claim is proven as the second one is in fact a corollary of the first given the frame and frame operator definitions, whereas the third one it is clear once more under Proposition 3.3.1 given the fact that \mathbf{S}^{-1} is a frame operator as well.

Thus, expand and rewrite $\tilde{\mathbf{S}}\mathbf{x}$ for any $\mathbf{x} \in \mathbb{H}^M$ as

$$\tilde{\mathbf{S}}\mathbf{x} = \sum_{k=1}^K \langle \mathbf{x}, \tilde{\mathbf{f}}_k \rangle \cdot \tilde{\mathbf{f}}_k \quad (3.36a)$$

$$= \sum_{k=1}^K \langle \mathbf{x}, \mathbf{S}^{-1} \mathbf{f}_k \rangle \cdot \mathbf{S}^{-1} \mathbf{f}_k \quad (3.36b)$$

$$= \mathbf{S}^{-1} \sum_{k=1}^K \langle \mathbf{S}^{-1} \mathbf{x}, \mathbf{f}_k \rangle \cdot \mathbf{f}_k \quad (3.36c)$$

$$= \mathbf{S}^{-1} \mathbf{S}(\mathbf{S}^{-1} \mathbf{x}) \quad (3.36d)$$

$$= \mathbf{S}^{-1} \mathbf{x}. \quad (3.36e)$$

□

To this point, is therefore clear that the canonical duality of frames is in fact a reciprocity relationship between a set frame $\{\mathbf{f}_k\}$ and its canonical dual $\{\tilde{\mathbf{f}}_k\}$, as these can be interchanged freely to obtain appropriate signal expansions

$$\mathbf{x} = \tilde{\mathbf{F}}\mathbf{F}^H \mathbf{x} = \sum_{k=1}^K \langle \mathbf{x}, \mathbf{f}_k \rangle \tilde{\mathbf{f}}_k, \quad (3.37)$$

$$\mathbf{x} = \mathbf{F}\tilde{\mathbf{F}}^H \mathbf{x} = \sum_{k=1}^K \langle \mathbf{x}, \tilde{\mathbf{f}}_k \rangle \mathbf{f}_k \quad (3.38)$$

at the expense of energy reciprocity relative to the chosen representation as outlined by Theorem 3.3.2.

Albeit the fact that the frame operator is tightly related to the analysis, synthesis and representation energy of the associated frame-theoretic transforms, it has however no characterizing role to describe the spread of the frame elements relative to each other over the available space \mathbb{H}^M . In other words, the frame operator cannot describe the relative angular displacement in a geometric sense between each pair of frame vectors. To fix this lack of representation, the Gramian operator is used.

Definition 3.3.2 (The Frame Gramian). *Let there be a frame \mathbf{F} containing K frame vectors $\{\mathbf{f}_k\} \in \mathbb{H}^M$, then the associated frame Gram operator is defined as*

$$\mathbf{G} \triangleq \mathbf{F}^H \mathbf{F} \in \mathbb{H}^{K \times K}, \quad (3.39)$$

where the individual entries of \mathbf{G} are given respectively by the inner products of all the frame pairs as

$$g_{k\ell} = \langle \mathbf{f}_k, \mathbf{f}_\ell \rangle = \mathbf{f}_\ell^H \mathbf{f}_k. \quad (3.40)$$

It is easy therefore to remark that in the case of orthonormal bases the Gram matrix is actually reduced to the identity matrix, *i.e.*, $\mathbf{G} = \mathbf{I}_K$. However, in the case of frames, more precisely for $K > M$, the Gram matrix itself cannot be reduced to the identity given the inherent overloading which prevents the frame vectors to be mutually orthogonal. Similarly to the frame operator \mathbf{S} , the Gramian \mathbf{G} has some attributes worth listing.

Proposition 3.3.3. *The frame Gram operator \mathbf{G} satisfies in general the properties:*

1. \mathbf{G} is linear and bounded;
2. $\mathbf{G}^H = \mathbf{G}$, *i.e.* \mathbf{G} is self-adjoint;
3. \mathbf{G} is rank-deficient and moreover $\text{rank}(\mathbf{G}) = M$.
4. $\langle \mathbf{G}\mathbf{x}, \mathbf{x} \rangle \geq 0 \forall \mathbf{x} \in \mathcal{H}^M, \mathbf{x} \neq \mathbf{0}$, *i.e.* \mathbf{G} is positive semi-definite;

Proof: The first statement is a consequence of the definition of \mathbf{G} . To this extent, linearity and boundedness of \mathbf{G} result from the fact that $\mathbf{G} = \mathbf{F}^H \mathbf{F}$, *i.e.* \mathbf{G} is obtained by multiplication of a linear bounded operator (this is the case given the definition of a frame) and its adjoint.

Secondly, to notice that \mathbf{G} is self-adjoint let

$$\mathbf{G}^H = (\mathbf{F}^H \mathbf{F})^H = \mathbf{F}^H \mathbf{F} = \mathbf{G}. \quad (3.41)$$

Thirdly, recall the general product rank inequality and apply it to \mathbf{G}

$$\text{rank}(\mathbf{G}) \leq \min(\text{rank}(\mathbf{F}), \text{rank}(\mathbf{F}^H)) = M, \quad (3.42)$$

where the last equality follows given the frame definition and strictly positive frame lower bound α .

Furthermore, given that the $\alpha > 0$ and \mathbf{F} spans the space \mathbb{H}^M , it also follows by means of Sylvester's rank inequality [43] that

$$\underbrace{\text{rank}(\mathbf{F}^H)}_M + \underbrace{\text{rank}(\mathbf{F})}_M - M \leq \text{rank}(\mathbf{G}) \quad (3.43)$$

$$M \leq \text{rank}(\mathbf{G}) \quad (3.44)$$

leading to the fact that $\text{rank}(\mathbf{G}) = M$, which is thus rank-deficient since $\mathbf{G} \in \mathbb{H}^{K \times K}$.

The last property, the semi-definiteness, follows by means of the Rayleigh-Ritz Theorem since given that \mathbf{G} is self-adjoint its eigenvalues are real and moreover

$$\langle \mathbf{G}\mathbf{x}, \mathbf{x} \rangle \geq \lambda_{\min} \mathbf{x} = 0, \quad \forall \mathbf{x} \in \mathbb{H}^K, \mathbf{x} \neq \mathbf{0}. \quad (3.45)$$

The last equality above is in fact a consequence of the rank-deficiency exhibited by \mathbf{G} and previously proven.

□

3.4 Frame Properties and Associated Quantifiers

Provided the specialized linear operators, *i.e.* the frame operator and the frame Gramian, able to characterize the attributes of any frame \mathbf{F} , three quantifiers of frame properties are next introduced: the *frame redundancy*, the *frame potential*, and respectively, the *frame mutual coherence*. These are discussed in detail together with frame categories exploiting some of the properties described.

Definition 3.4.1 (Frame Redundancy). *Given a frame $\mathbf{F} \in \mathbb{H}^{M \times K}$ formed of K frame vectors $\{\mathbf{f}_k\} \in \mathbb{H}^M$, its redundancy is defined as the ratio*

$$\rho \triangleq \frac{K}{M}. \quad (3.46)$$

Upon the above definition it so follows that in fact redundancy is a measure of overcompleteness of any frame \mathbf{F} . In other words, the overloading of the original signal space \mathbb{H}^M by K frame vectors represents the expected redundancy of the frame analysis representation, and conversely, the robustness gained under the reciprocal synthesis operation.

Definition 3.4.2 (Frame Potential, [75]). *Let \mathbf{F} be a frame with K normalized frame vectors, *i.e.* $\|\mathbf{f}_k\|_2 = 1, \forall \mathbf{f}_k \in \mathbb{H}^M$. The frame potential (FP) associated with the frame \mathbf{F} is given by*

$$\text{FP}(\mathbf{F}) = \sum_{k=1}^K \sum_{\ell=1}^K |\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|^2 = \sum_{k=1}^K \sum_{\ell=1}^K |g_{k\ell}|^2 = \|\mathbf{G}\|_{\mathbf{F}}^2. \quad (3.47)$$

The FP is meant to quantify based on quantum physics and measurements the frame force field [75] which falls out of the scope of the current work. However, one can intuitively remark based on the definition that the FP is in fact a measure quantifying how close is a frame to an analogous orthogonal representation. This is intuitively clear given the fact that the minimizer of $\|\mathbf{G}\|_{\mathbf{F}}$, and hence of the FP, is attained over diagonal and normalized Gram matrices. Given the fact that for any $K > M$, \mathbf{G} cannot achieve a purely diagonal structure, and so, FP measures as an ensemble the representational energy dispersed by a normalized frame \mathbf{F} for any vectors over a unit M -ball.

From a purely mechanical standpoint, the equilibrium of the representational system described by \mathbf{F} is achieved for minimizers of the Equation (3.47). Similarly, under an energy-only argument, the idea is to minimize the FP such that same signal space vectors are represented algebraically under an average optimized energy usage regardless of their position in \mathbb{H}^M . The upper intuitive remark is made clear by the main results of [75] summarized below.

Theorem 3.4.1 ([75, Th. 6.2]). *Given the frame $\mathbf{F} \in \mathbb{H}^{M \times K}$ with normalized frame vectors and the FP measure described by the functional (3.47), $\text{FP}(\mathbf{F})$, then the following results hold:*

$$\max\left(K, \frac{K^2}{M}\right) \leq \text{FP}(\mathbf{F}) \leq K^2. \quad (3.48)$$

Proof: The right-hand-side inequality of (3.48) is quite loose and thus implied directly by

means of the Cauchy-Schwarz classic inequality, and respectively, by the fact that $\|\mathbf{f}_k\| = 1, \forall k$, such that

$$\text{FP}(\mathbf{F}) = \sum_{k=1}^K \sum_{\ell=1}^K |\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|^2 \leq \sum_{k=1}^K \sum_{\ell=1}^K \|\mathbf{f}_k\|_2^2 \|\mathbf{f}_\ell\|_2^2 = \sum_{k=1}^K \sum_{\ell=1}^K 1 = K^2. \quad (3.49)$$

The left-hand-side inequality is however tight. One should first note that

$$\text{FP}(\mathbf{F}) = K + \underbrace{\sum_{k=1}^K \sum_{\ell \neq k} |\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|^2}_{\geq 0} \geq K. \quad (3.50)$$

As a matter of fact, the second term from above, $\sum_{k=1}^K \sum_{\ell \neq k} |\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|^2$, is lower bounded by 0 for the case where $K = M$. This case involves the orthogonality of the frame vectors, which corroborated with the normalization condition made clear in (3.50) results in \mathbf{F} being any orthonormal bases.

For $K > M$, as previously seen, the summation $\sum_{k=1}^K \sum_{\ell \neq k} |\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|^2$ is strictly positive. As a result, the lower bound in this case is different from K . To this end, using the identity $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$, it follows that

$$\|\mathbf{G}\|_{\mathbf{F}}^2 = \text{trace}(\mathbf{G}^{\mathbf{H}}\mathbf{G}) \quad (3.51a)$$

$$= \text{trace}(\mathbf{F}^{\mathbf{H}}\mathbf{F}\mathbf{F}^{\mathbf{H}}\mathbf{F}) \quad (3.51b)$$

$$= \text{trace}(\mathbf{F}\mathbf{F}^{\mathbf{H}}\mathbf{F}\mathbf{F}^{\mathbf{H}}) \quad (3.51c)$$

$$= \text{trace}(\mathbf{S}^{\mathbf{H}}\mathbf{S}) \quad (3.51d)$$

$$= \sum_{m=1}^M \lambda_m^2 \quad (3.51e)$$

Minimizing FP is therefore equivalent to minimizing (3.51e). In addition, following the definitions of the frame linear operators, it follows that $\text{trace}(\mathbf{S}) = \text{trace}(\mathbf{F}\mathbf{F}^{\mathbf{H}}) = \text{trace}(\mathbf{F}^{\mathbf{H}}\mathbf{F}) = \text{trace}(\mathbf{G}) = K$ given the normalized frame vectors. As a result, finding the tight lower bound of FP is reduced to solving the optimization problem

$$\min_{\lambda_m, m=\{1,2,\dots,M\}} \sum_{m=1}^M \lambda_m^2 \quad (3.52a)$$

$$\text{subject to } \lambda_m > 0, \forall m \quad (3.52b)$$

$$\sum_{m=1}^M \lambda_m = K. \quad (3.52c)$$

The latter is in fact the classical problem of finding the point closest to the origin on the hyperplane \mathbb{H}^M constrained to a fixed cumulative length of its coordinates. The solution to the latter is known, [76], and obtained as a unique minimizer by setting all the coordinates equal, *i.e.* $\lambda_m = \frac{K}{M} > 0, \forall m$.

As a consequence, the minimum value of (3.52) is attained for

$$\sum_{m=1}^M \left(\frac{K}{M}\right)^2 = M \cdot \frac{K^2}{M^2} = \frac{K^2}{M}, \quad (3.53)$$

which hence proves the left-hand-side inequality for the case of $K > M$ and concludes the proof under the same core arguments as of [75].

□

Theorem 3.4.1 has an immediate corollary given the frame definition and frame operator.

Corollary 1. *The unit-normalized frames attaining the minimum FP are*

- i). *orthonormal bases under the assumption of $K = M$, or*
- ii). *redundant bases under the assumption of $K > M$ such that $\alpha = \beta = \frac{K}{M}$ given the Definition 3.2.1.*

In other words, Corollary 1 states that the minimizers of the FP are any orthonormal bases or generalizations thereof which tighten the representational energy to the optimal point where $\alpha = \beta$ for any given vector \mathbf{x} from the signal space \mathbb{H}^M . This interpretation highlights the fact that the FP is a quantifiable measure that can be leveraged in determining how close a frame is to an analogous tight energy representation, achievable in classic terms via orthonormal bases. Consequently, two important frame properties span of the FP corollary, namely *tightness* and *equi-normality*.

Definition 3.4.3 (Tightness). *A frame \mathbf{F} of K frame vectors spanning \mathbb{H}^M is called tight frame given that it satisfies the frame definition tightly with equality, i.e.*

$$\alpha \|\mathbf{x}\|_2^2 = \|\mathbf{F}\mathbf{x}\|_2^2 = \beta \|\mathbf{x}\|_2^2, \quad \forall \mathbf{x} \in \mathbb{H}^M, \quad (3.54)$$

leading to $0 < \alpha = \beta < \infty$.

Definition 3.4.4 (Equi-normality). *A frame \mathbf{F} of K frame vectors spanning \mathbb{H}^M is called equi-normal frame given that all its frame vectors \mathbf{f}_k satisfy the fact that*

$$\|\mathbf{f}_k\|_2^2 = \zeta^2 > 0 \quad (3.55)$$

Using the inequalities (3.30), tightness diagonalizes the frame operator, i.e.

$$\mathbf{S} = \alpha \mathbf{I}_M. \quad (3.56)$$

Needless to say at this point, that the normalized frame vectors condition implicit in the FP definition is nothing but a particular instance of equi-normality, i.e., *unit-normality* such that $\zeta = 1$. In conjunction with the tightness condition and its result from Equation (3.56), this leads to the fact that

$$\alpha = \beta = \frac{K}{M}, \quad (3.57)$$

and consequently,

$$\mathbf{S} = \frac{K}{M} \mathbf{I}_M \quad (3.58)$$

for any unit-norm tight frame (UNTF) \mathbf{F} .

In the light of the latter results, it makes sense to critically revisit at this point the redundancy Definition 3.4.1. Although correct in capturing qualitatively the typical meaning of redundancy in communication systems through a classical Coding Theory approach, this definition is too naive in general as it does not capture the quantitative features of frame theoretic representations specific to particular signals. In fact, during the last years an alternative view of redundancy which better captures the features of frames has gathered more interest [19, Ch. 1.4, pp. 41]. To this end, in [77] the *redundancy function*

$$\rho_{\mathbf{F}}(\mathbf{x}) \triangleq \sum_{k=1}^K \left| \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|_2}, \mathbf{f}_k \right\rangle \right|^2 \quad (3.59)$$

was introduced in order to better capture both the qualitative and quantitative effects of a $\mathbf{F} \in \mathbb{H}^{M \times K}$ overloaded representations on any signal space vector $\mathbf{x} \in \mathbb{H}^M$.

Based on the frame definition and inequalities chain (3.17), it is easy to remark the fact that the redundancy function is a measure that computes the norm of the frame projection, *i.e.* analysis operation, for all the signal vectors on the surface of the M -sphere in \mathbb{H}^M . Under this normalization, it is therefore possible to visualize the different redundancies experienced by all the signal space vectors \mathbf{x} which can be obtained by appropriate scaling of the multidimensional sphere. By means of (3.17), it follows that the redundancy function is finite, and respectively, lower and upper bounded by

$$\rho^- = \min_{\mathbf{x} \in \mathbb{H}^M, \|\mathbf{x}\|=1} \rho_{\mathbf{F}}(\mathbf{x}) = \alpha \quad (3.60)$$

$$\rho^+ = \max_{\mathbf{x} \in \mathbb{H}^M, \|\mathbf{x}\|=1} \rho_{\mathbf{F}}(\mathbf{x}) = \beta, \quad (3.61)$$

in a similar fashion to the lower and upper frame bounds.

Furthermore, tight frames imply $\rho^- = \rho^+$, and such frames are said to exhibit *uniform redundancy* [77]. In the case of UNTF, the latter holds true with the added fact that $\rho^- = \rho^+ = \rho = \frac{K}{M}$ which is the classical definition of redundancy. Thus it is important to note the role of tight frames in providing uniform robustness against distortions by means of uniform redundancy.

Example 3.4.1. Consider for instance a special case of an UNTF, the Mercedes-Benz frame [19] in $\mathbb{R}^{2 \times 3}$:

$$\mathbf{F}_{\text{MB}} = \begin{bmatrix} 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ 1 & -\frac{\sqrt{1}}{2} & -\frac{\sqrt{1}}{2} \end{bmatrix} \quad (3.62)$$

and its slightly perturbed relative

$$\mathbf{F}_{\text{PMB}} = \begin{bmatrix} \cos\left(\frac{\pi}{2} - 0.1\right) & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \sin\left(\frac{\pi}{2} - 0.1\right) & -\frac{\sqrt{1}}{2} & -\frac{\sqrt{1}}{2} \end{bmatrix}. \quad (3.63)$$

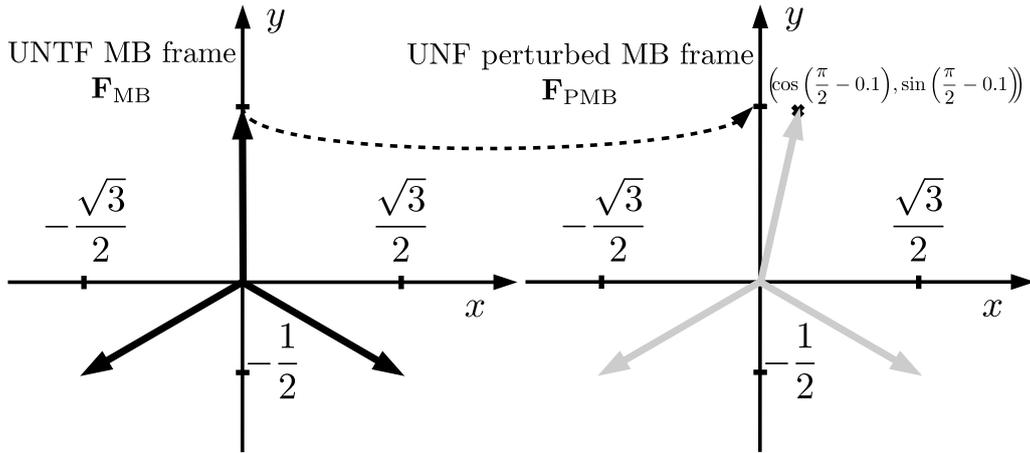


Figure 3.1: Outline of the frame vectors of the UNTF Mercedes-Benz frame (in black), and respectively, of the frame vectors of the unit-norm frame (UNF) perturbed Mercedes-Benz frame used in the current frame redundancy and redundancy function example.

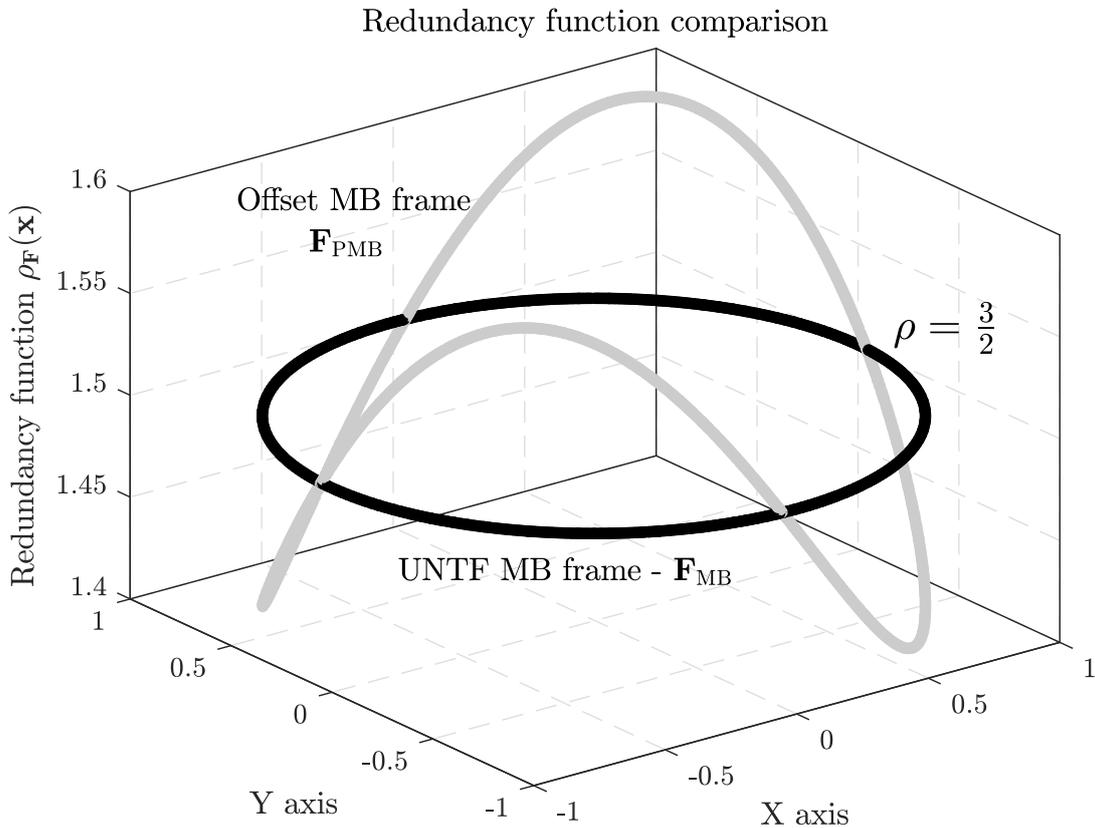


Figure 3.2: Redundancy function (3.59) exemplary comparison between an UNTF – the Mercedes-Benz frame (in black) – and its slightly offset – one-component rotated clockwise by approx 5.73° (in gray) – unit-norm frame variant.

The two frames are drawn together for comparison in Figure 3.1 above. Additionally, the redundancy function images corresponding to the two frames introduced above are computed over the surface of an unitary 2-dimensional sphere, i.e. an unitary ring, and plotted side-by-side in Figure 3.2 as well.

As expected, given the fact that \mathbf{F}_{MB} is an UNTF frame, its redundancy measure is uniform at the value $\rho(\mathbf{x}) = \frac{K}{M} = \frac{3}{2}$, and therefore similar to the redundancy Definition 3.4.1. On the other hand, the slight misplacement of one of the elements of Mercedes-Benz frame in \mathbf{F}_{PMB} generates a disturbed, non-tight but still unit-norm frame. As a consequence, in this case, as illustrated in Figure 3.2, the redundancy function varies with \mathbf{x} and so the robustness efficiency of \mathbf{F}_{PMB} varies depending on the processed signal as well.

We remark however that this *signal-dependent redundancy* is not necessarily a weakness, but a general feature of frames. In fact, this attribute can be exploited for instance to gain higher robustness to signals that are more susceptible to localized distortions.

The illustrative example in the prequel did not only underscore the fine details of redundancy and tightness, but also made clear the fact that insofar the analogy of frames to orthogonal bases has resumed to the representational energy and its subsequent tightness. To this point no concept has tackled the problem of orthogonality, or better phrased, lack thereof in redundant representations such as frames. Geometrically, this latter attribute is resumed to the angular displacements between any two frame vector elements. In orthogonal sense, this angular displacement is globally minimized across any disjoint pair of such vectors such that their inner product is nulled. However, this has no correspondence in terms of redundant bases, *i.e.*, for $K > M$, and to this extent it is relaxed to the concept of coherence and subsequently quantified by *mutual coherence*, [20, 78] and [19, Ch. 6].

Definition 3.4.5 (Mutual Coherence). *Consider a frame $\mathbf{F} \in \mathbb{H}^{M \times K}$. Its mutual coherence is given by*

$$\mu(\mathbf{F}) \triangleq \max_{1 \leq k < \ell \leq K} \frac{|\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|}{\|\mathbf{f}_k\|_2 \|\mathbf{f}_\ell\|_2} = \max_{1 \leq k < \ell \leq K} \frac{|\mathbf{f}_\ell^H \mathbf{f}_k|}{\|\mathbf{f}_k\|_2 \|\mathbf{f}_\ell\|_2}. \quad (3.64)$$

The mutual coherence is therefore a quantifier of maximum correlation among all the frame pairs taken out of the frame vectors. Additionally, the normalization step in the definition of mutual coherence inflicts a geometric interpretation by means of the angular representation of the inner product. For instance, both for the real and complex space cases, $\mathbb{H} = \mathbb{R}$ or $\mathbb{H} = \mathbb{C}$, the normalization ensures that the mutual coherence becomes in fact a quantization of the maximum angular displacement

$$\mu(\mathbf{F}) = \max_{1 \leq k < \ell \leq K} \frac{|\mathbf{f}_\ell^H \mathbf{f}_k|}{\|\mathbf{f}_k\|_2 \|\mathbf{f}_\ell\|_2} = \max_{1 \leq k < \ell \leq K} |\angle \cos(\mathbf{f}_k, \mathbf{f}_\ell)|. \quad (3.65)$$

Without any doubt in the case of UNFs, the mutual coherence is further reduced given the fact that $\|\mathbf{f}_k\|_2 = 1, \forall k$ to

$$\mu(\mathbf{F}) = \max_{1 \leq k < \ell \leq K} |\mathbf{f}_\ell^H \mathbf{f}_k| = \max_{1 \leq k < \ell \leq K} |g_{k\ell}|, \quad (3.66)$$

such that it can be fully described by the entries of the Gram matrix of the frame, *i.e.* \mathbf{G} .

Following this geometric argument, it thus yields that indeed the mutual coherence is naturally upper bounded such that

$$\mu(\mathbf{F}) \leq \angle \cos(0^\circ) = 1, \quad (3.67)$$

achievable whenever any pair of two distinct frame vectors are collinear.

Theorem 3.4.2. *Let \mathbf{F} be a frame of K vectors over \mathbb{H}^M . Then its mutual coherence is lower bounded by*

$$\mu(\mathbf{F}) \geq \sqrt{\frac{K-M}{M(K-1)}}. \quad (3.68)$$

The inequality above is met tightly with equality iff \mathbf{F} is an equiangular tight frames (ETF), i.e.,

$$\frac{|\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle|}{\|\mathbf{f}_k\|_2 \|\mathbf{f}_\ell\|_2} = \sqrt{\frac{K-M}{M(K-1)}}, \forall k \neq \ell, \quad (3.69)$$

where additionally, \mathbf{S} is also diagonal.

However, the existence of such minimally coherent frames is restricted by the frame redundancy such that the equality in (3.68) can only hold as long as

$$K \leq \frac{M(M+1)}{2} \text{ for } \mathbb{H} = \mathbb{R} \quad (3.70)$$

$$K \leq M^2 \text{ for } \mathbb{H} = \mathbb{C} \quad (3.71)$$

Proof: A complete proof of the lower bound in (3.68), also known as the Welch bound (WB) [78], is provided in [79,80]. Furthermore, the dimensionality constraints (3.70), (3.71) are obtained via [81, Tab. II].

As a consequence, a sketch proof is provided to the equivalence between ETF and satisfying the mutual coherence bound in (3.68) with equality. To this end, consider without loss of generality \mathbf{F} to be an UNF. Based upon (3.50) and (3.51) it follows that

$$\sum_{k=1}^K \sum_{\ell > k} |g_{k\ell}|^2 \geq \frac{K(K-M)}{2M}. \quad (3.72)$$

Considering the fact that there are $\frac{K(K-1)}{2}$ unique pairings drawn out of $[0, 1]$ on the left-hand side of (3.72), it follows that the normalized average squared correlation of any frame vector pair is bounded as

$$\overline{|g_{k\ell}|^2}_{k \neq \ell} \geq \frac{K-M}{M(K-1)}. \quad (3.73)$$

Since the mutual coherence is defined as the maximum correlation, see Definition 3.4.5, and the average correlation is itself lower bounded by the WB as outlined in (3.73), it is therefore implied that the minimum mutual coherence is attained as the ensemble of correlations meets uniformly with equality (3.73). As an effect, this is equivalent to

$$|g_{k\ell}| = \sqrt{\frac{K-M}{M(K-1)}} \quad (3.74)$$

which implies equiangularity and tightness as also (3.72) is met with equality.

□

An example of an ETF is in fact the Mercedes-Benz frame depicted already graphically in Figure 3.1 as \mathbf{F}_{MB} . However, despite numerically computed optimum incoherent ETFs for smaller dimensions, see [78] and references therein for more details, the universal existence of ETFs for any redundancy levels is without proof to date, and hence, a conjecture of Frame Theory [78]. To this end, a more general concept has been introduced during the last decades stemming from other areas of Mathematics following the *Grassmannian packing problem* or *equiangular lines and point sets* [78].

Definition 3.4.6 (Grassmannian frames). *A frame \mathbf{F} composed of K -vectors belonging to \mathbb{H}^M is called a Grassmannian frame if it solves the optimization problem*

$$\min_{\mathbf{F}} \quad \mu(\mathbf{F}) \quad (3.75a)$$

$$\text{subject to} \quad \|\mathbf{f}_k\|_2 = 1 \quad \forall k. \quad (3.75b)$$

Grassmannian frames therefore minimize the maximum correlation among any frame vectors pair across all the UNF with the same redundancy over $\mathbb{H}^{M \times K}$. Obviously the unit-normalization condition is applied given just mere mathematical convenience as any frame could be normalized relative to its frame vector elements. Additionally, upon Theorem 3.4.2, it results that Grassmannian frames represent in fact a frame superset enclosing the potentially existent ETFs that reach the WB in (3.68) with equality. Such frames are in fact named in the context of Grassmannian frames as nothing but *optimal Grassmannian frames* [78].

As previously mentioned, an universal constructive solution to the optimization problem (3.75a) is to this point unknown which makes the construction of Grassmannian frames of particular interest in the Frame Theory literature, *e.g.*, [19, 78, 82] and references therein. The problem is in fact relaxed in practice to generically finding frames capable of achieving maximum correlation levels among all their element pairs approaching the Welch bound. From this perspective, such frames and their design are coined as *frame incoherence* or, respectively, the *incoherent frame design problem* [19, 82].

3.5 Special Frame Types and Their Designs

As observed in the previous section, the redundancy introduced by overcomplete bases (frames with $K > M$) affects both the representation projected energy, but also the correlation among all the frame element pairs. Two fundamental properties address the latter two, tightness and coherence. These are quantified and therefore numerically summarized respectively by the FP and mutual coherence previously discussed. In addition, frames designed on top of such attributes are named as *tight*, and respectively, *incoherent*, where *equi-normality*, in particular unitary normalization, has been inherently considered for mathematical convenience and ease of the discussion.

Nevertheless, despite exposing most of the fundamental properties of redundancy, tightness and incoherence as core frame characteristics, no explicit construction methods were presented to build frames encompassing such properties. In other words, insofar no recipe has been provided

for instance to build UNTFs or low-coherent UNFs vector ensembles. Since such frames are desirable in practice, their design principles are detailed in the sequel of this section. To this end, we reduce the generic Hilbert space \mathbb{H} treated so far to the real and complex, *i.e.*, \mathbb{R} , and respectively, \mathbb{C} , particular to signals in wireless communications.

Most of the UNTF constructions rely on the spectral attributes of the frame operator \mathbf{S} associated with such frames, *i.e.*, upon the fact that \mathbf{S} is diagonalized given (3.58). To this end, spectral based projections are employed in order to force the frame operator of a given frame to the frame operator of an UNTF. The equivalence between diagonal \mathbf{S} and tight frames is then leveraged upon proper scaling to obtain the desired UNTF. The above projection design paradigm is in fact motivated by a general result of Frame Theory providing an analytic approach to generating the Frobenius norm nearest tight frame to any arbitrary starting frame.

Theorem 3.5.1. *Let \mathbf{F}_0 be a $M \times K$ matrix denoting a frame with K vectors in a M -dimensional Hilbert space $\mathbb{H} = \mathbb{R}$ or $\mathbb{H} = \mathbb{C}$ with the singular value decomposition (SVD) given by $\mathbf{U}\Sigma\mathbf{V}^H$. The closest α -tight frame to \mathbf{F}_0 , such that*

$$\underset{\mathbf{F}}{\operatorname{argmin}} \quad \|\mathbf{F}_0 - \mathbf{F}\|_F, \quad (3.76a)$$

$$\text{subject to} \quad \mathbf{F}\mathbf{F}^H = \alpha\mathbf{I}_M \quad (3.76b)$$

is given by $\sqrt{\alpha}\mathbf{U}\underline{\mathbf{V}}^H$ or equivalently by $\sqrt{\alpha}(\mathbf{F}\mathbf{F})^{-\frac{1}{2}}\mathbf{F}$.

Note that in the notation above the underlined matrices, *e.g.*, $\underline{\mathbf{V}}$, mark for convenience the structures that require particular truncation or trivial expansion (zeroed rows/columns) for compliant dimensionality under matrix multiplication given the fact that K may be larger than M .

Proof: In literature the proof above relies on a more general result from linear algebra, *i.e.*, the Schur-Horn Theorem [83], see for instance [19, 78] or their cross-references. However, an alternative straight-forward proof may be listed under the remark that in fact the linear transform above, $\underline{\mathbf{U}}\underline{\mathbf{V}}^H$, is nothing but the unitary part of the polar decomposition [84] of the matrix \mathbf{F}_0 .

Concretely, given the SVD decomposition of the original frame \mathbf{F}_0 , this can be rewritten in conjunction with the general polar factorization of matrices as

$$\mathbf{F}_0 = \mathbf{U}\Sigma\mathbf{V}^H \quad (3.77a)$$

$$= \underbrace{\mathbf{U}\Sigma\mathbf{U}^H}_{\mathbf{P} \succeq 0} \underline{\mathbf{U}}\underline{\mathbf{V}}^H \quad (3.77b)$$

$$= \mathbf{P} \cdot \underline{\mathbf{U}}\underline{\mathbf{V}}^H. \quad (3.77c)$$

Thus, \mathbf{P} is a square positive semi-definite matrix containing the power information of \mathbf{F}_0 , whereas $\underline{\mathbf{U}}\underline{\mathbf{V}}^H$ is the polar unitary rotation linear operator. Based on this remark and on the

fact that $\text{rank}(\mathbf{F}_0) = M$ as a frame, \mathbf{P} is uniquely given as [84]

$$\mathbf{P} = (\mathbf{F}_0 \mathbf{F}_0^H)^{\frac{1}{2}}, \quad (3.78)$$

and as a matter of fact, by means of the frame operator properties in Proposition 3.3.1, also strictly positive definite, *i.e.*, $\mathbf{P} \succ 0$, and hence, invertible.

As a consequence, applying the latter to (3.77c) yields the unique

$$\mathbf{U} \underline{\mathbf{V}}^H = (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \mathbf{F}_0. \quad (3.79)$$

In the light of the above, we can remark that the newly obtained frame $\mathbf{F} = \sqrt{\alpha} (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \mathbf{F}_0$ is in fact an α -tight frame. This is the case as its associated frame operator is given by

$$\mathbf{S} = \sqrt{\alpha} (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \mathbf{F}_0 \cdot \sqrt{\alpha} \left((\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \mathbf{F}_0 \right)^H \quad (3.80a)$$

$$= \alpha (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \mathbf{F}_0 \mathbf{F}_0^H (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \quad (3.80b)$$

$$= \alpha (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} (\mathbf{F}_0 \mathbf{F}_0^H)^{\frac{1}{2}} (\mathbf{F}_0 \mathbf{F}_0^H)^{\frac{1}{2}} (\mathbf{F}_0 \mathbf{F}_0^H)^{-\frac{1}{2}} \quad (3.80c)$$

$$= \alpha \mathbf{I}_M. \quad (3.80d)$$

The fact that the above is the global minimizer of the problem (3.76) follows as a direct corollary of low-rank matrix approximation problems and of the Eckart–Young–Mirsky Theorem [43, 85].

□

Theorem 3.5.1 is a very powerful result which enables the generation of tight frames starting from any generic frame \mathbf{F}_0 . Given the implied approximation of the construction it thus follows that the initial frame in turn determines partially the structure of the tight byproduct. However, potential desirable properties such as equi-normality or incoherence may be lost under application of Theorem 3.5.1. To circumvent such a desideratum, an alternative projections may be necessary in order to force the desired structures onto the frame Gram matrix \mathbf{G} , and respectively, the frame operator \mathbf{S} which compact and describe the frame properties as previously seen.

In fact, the idea of additional projections is a common approach [86, 87] meant to satisfy different requirements by means of iteratively alternating projections over convex sets or manifolds belonging to the spaces corresponding to the desired frame operators and Gram matrices. The corresponding sets for instance for the case of UNTF incoherent frames are given as follows.

- *The unit-normality set* \mathbb{S}_U , *i.e.*, the set of $K \times K$ Gram matrices with *unit diagonals* is

$$\mathbb{S}_U \triangleq \left\{ \mathbf{G} \in \mathbb{H}^{K \times K} \mid \mathbf{G} = \mathbf{G}^H, g_{ii} = 1 \forall i \in \{1, \dots, K\} \right\}. \quad (3.81)$$

If $\mathbf{G} = \mathbf{F}^H \mathbf{F}$ is the $K \times K$ Gram of any frame \mathbf{F} , the projection of \mathbf{G} onto the set \mathbb{S}_U is

given as

$$\mathcal{P}_U(\mathbf{G}) \triangleq \mathbf{F}_U^H \mathbf{F}_U \quad \text{where} \quad \mathbf{F}_U \triangleq \left[\frac{\mathbf{f}_1}{\|\mathbf{f}_1\|_2}, \dots, \frac{\mathbf{f}_K}{\|\mathbf{f}_K\|_2} \right]. \quad (3.82)$$

- *The incoherence set \mathbb{S}_I , i.e., the set of $K \times K$ symmetric matrices with bounded off-diagonals is*

$$\mathbb{S}_I \triangleq \left\{ \mathbf{G} \in \mathbb{H}^{K \times K} \mid \mathbf{G} = \mathbf{G}^H, |g_{ij}| \leq \gamma_I, \forall (i \neq j) \in \{1, \dots, K\} \right\}. \quad (3.83)$$

This projection can be performed simply by replacing the off-diagonal elements g_{ij} of \mathbf{G} that are larger than the bound by the bound itself while preserving the sign geometry, yielding

$$\mathcal{P}_I(\mathbf{G}) \triangleq \left\{ \mathbf{G} \mid g_{ij} = \begin{cases} \gamma_I \cdot \text{sgn}(g_{ij}), & \text{if } |g_{ij}| > \gamma_I, \\ g_{ij}, & \text{otherwise.} \end{cases} \right\} \quad (3.84)$$

The idea behind this operation is to limit highly correlated frame vectors by thresholding, idea firstly introduced by Elad in [86], and later, refined in [87]. A suitable threshold allowing for convergence of the alternating sequence of projections has been proposed based upon the highly decorrelated PN sequences used in DS-CDMA spreading codes [78]. Thus, for frame vectors of length M , the bound $\sqrt{\frac{1}{M}}$ was shown to lead to convergence whilst also lowering the correlation levels of frame vectors [87].

- *The frame set \mathbb{S}_F , i.e., the set of $K \times K$ symmetric matrices with rank M is*

$$\mathbb{S}_F \triangleq \left\{ \mathbf{G} \in \mathbb{H}^{K \times K} \mid \mathbf{G} = \mathbf{G}^H, \text{rank}(\mathbf{G}) = M \right\}. \quad (3.85)$$

This projection is necessary to satisfy the fact that the underlying frame corresponding to the projected Gram matrix \mathbf{G} is still a frame spanning the space \mathbb{H}^M . This resumes to the projection

$$\mathcal{P}_F(\mathbf{G}) = \underline{\mathbf{U}}|_{K \times M} \cdot \underline{\mathbf{\Sigma}}|_{M \times M} \cdot \underline{\mathbf{V}}|_{K \times M}^H, \quad (3.86)$$

where the SVD $\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ is implied and $\underline{\cdot}|_{K \times M}$ denotes a $K \times M$ upper-left truncation.

- *The tightness set \mathbb{S}_T , i.e., the set of $\frac{K}{M}$ -tight frames with $K \times K$ Gram matrices with spectrum is*

$$\mathbb{S}_T \triangleq \left\{ \mathbf{G} \in \mathbb{H}^{K \times K} \mid \mathbf{G} = \mathbf{G}^H, \lambda_i(\mathbf{G}) = \begin{cases} \frac{K}{M} \forall i \in \{1, \dots, M\} \\ 0 \forall i \in \{M+1, \dots, K\} \end{cases} \right\}. \quad (3.87)$$

Finally, if $\mathbf{G} = \mathbf{F}^H \mathbf{F}$ is the $K \times K$ Gram of a frame \mathbf{F} , the projection onto \mathbb{S}_T is achieved by means of Theorem 3.5.1 as

$$\mathcal{P}_T(\mathbf{G}) \triangleq \mathbf{F}_T^H \mathbf{F}_T \quad \text{where} \quad \mathbf{F}_T \triangleq \sqrt{\frac{K}{M}} (\mathbf{F} \mathbf{F}^H)^{-\frac{1}{2}} \mathbf{F}. \quad (3.88)$$

Algorithm 2 ALTERNATING PROJECTIONS UNTF DESIGN [87].

Inputs: M = dimension of the Hilbert space, K = number of frame vectors

Outputs: \mathbf{F}_* UNTF $\in \mathbb{H}^{M \times K}$

- 1: Generate random initial frame $\mathbf{F}_0 \leftarrow \text{rand}(M \times K)$
 - 2: Compute initial $\frac{K}{M}$ -tight Gram matrix $\mathbf{G}_T^{(0)} \leftarrow \mathcal{P}_T(\mathbf{F}_0^H \mathbf{F}_0)$
 - 3: **repeat**
 - 4: $\mathbf{G}_U^{(n)} \leftarrow \mathcal{P}_U(\mathbf{G}_T^{(n)})$ ▷ Projection over the set \mathbb{S}_U
 - 5: $\mathbf{G}_I^{(n)} \leftarrow \mathcal{P}_D(\mathbf{G}_I^{(n)})$ ▷ Projection over the set \mathbb{S}_I
 - 6: $\mathbf{G}_F^{(n)} \leftarrow \mathcal{P}_R(\mathbf{G}_F^{(n)})$ ▷ Projection over the set \mathbb{S}_F
 - 7: $\mathbf{G}_T^{(n)} \leftarrow \mathcal{P}_T(\mathbf{G}_R^{(n)})$ ▷ Projection over the set \mathbb{S}_T
 - 8: **until** convergence or maximum iterations reached
 - 9: $\mathbf{F}_* \leftarrow \mathbf{F}_T^{(n)}$ ▷ Variation 1 – *alternating projection*
 - or
 - 10: $\bar{\mathbf{G}} \leftarrow \frac{1}{4}(\mathbf{G}_U^{(n)} + \mathbf{G}_D^{(n)} + \mathbf{G}_R^{(n)} + \mathbf{G}_T^{(n)}) = \bar{\mathbf{U}}\bar{\mathbf{\Lambda}}\bar{\mathbf{U}}^H$
 - $\mathbf{F}_* \leftarrow \bar{\mathbf{\Lambda}}^{\frac{1}{2}}\bar{\mathbf{U}}$ ▷ Variation 2 – *average alternating projection*
-

In the light of all the projections above, the universal alternating projection approach [87] to generate approximate UNTFs can be concisely described by the pseudo-code listed under Algorithm 2.

Needless to say that the alternating projection method introduced above is not unique in generating UNTFs. There exist also other methods in the literature based on structured approaches to constructively and algebraically generate sparse UNTFs, the Spectral tetris algorithm [19], based on the similar properties embedded within the corresponding \mathbf{S} and \mathbf{G} frame operators. Alternatively, convex optimization methods have been very recently proposed to iteratively decorrelate arbitrary frames and reach highly incoherent redundant representations up to any dimensionality [82, 88, 89]. However, once again the latter works rely on the same inherent Gram matrix mathematical properties displayed and discussed above.

Finally, the dichotomy of different types of frames involving introduced properties of equinormality, tightness and incoherence is summarized graphically as Venn diagrams under Figures 3.3 and 3.4. On the one hand, the equinormality and tightness properties are brought together under Figure 3.3, and the mathematical relevant properties of frames exhibiting such attributes are listed under Table 3.1. The classification is based upon [22, Table A.3] and outlines the main properties of the identified frame classes with respect to the bounds of the inequality (3.17), to the frame elements and to their respective norms.

On the other hand, the geometric property of incoherence is paired with equinormality in Figure 3.4. The latter have a much coarser resolution compared to the first categorization since the research of frame incoherence is currently a popular and open topic among mathematicians and engineers alike, and as a result, a deep nomenclature beyond the level presented so far has not been so far broadly standardized [19, 78, 82].

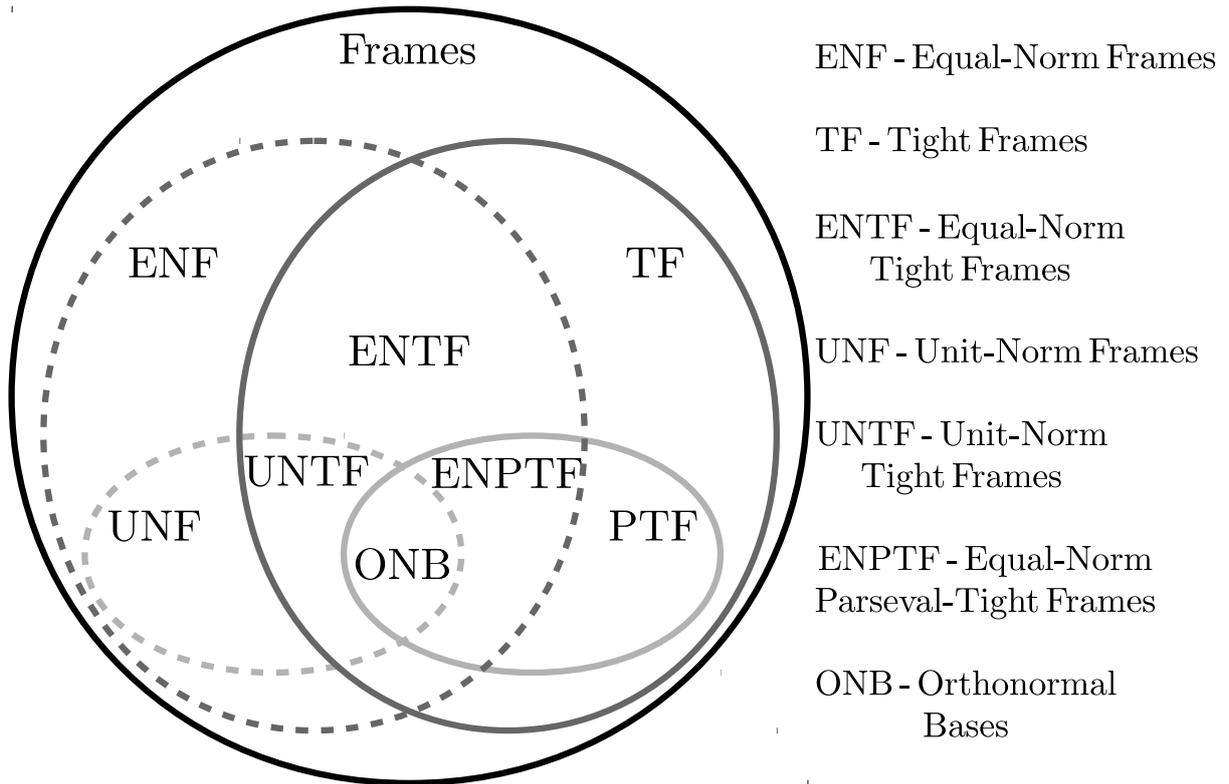


Figure 3.3: Overview of frame types – equi-normality & tightness.

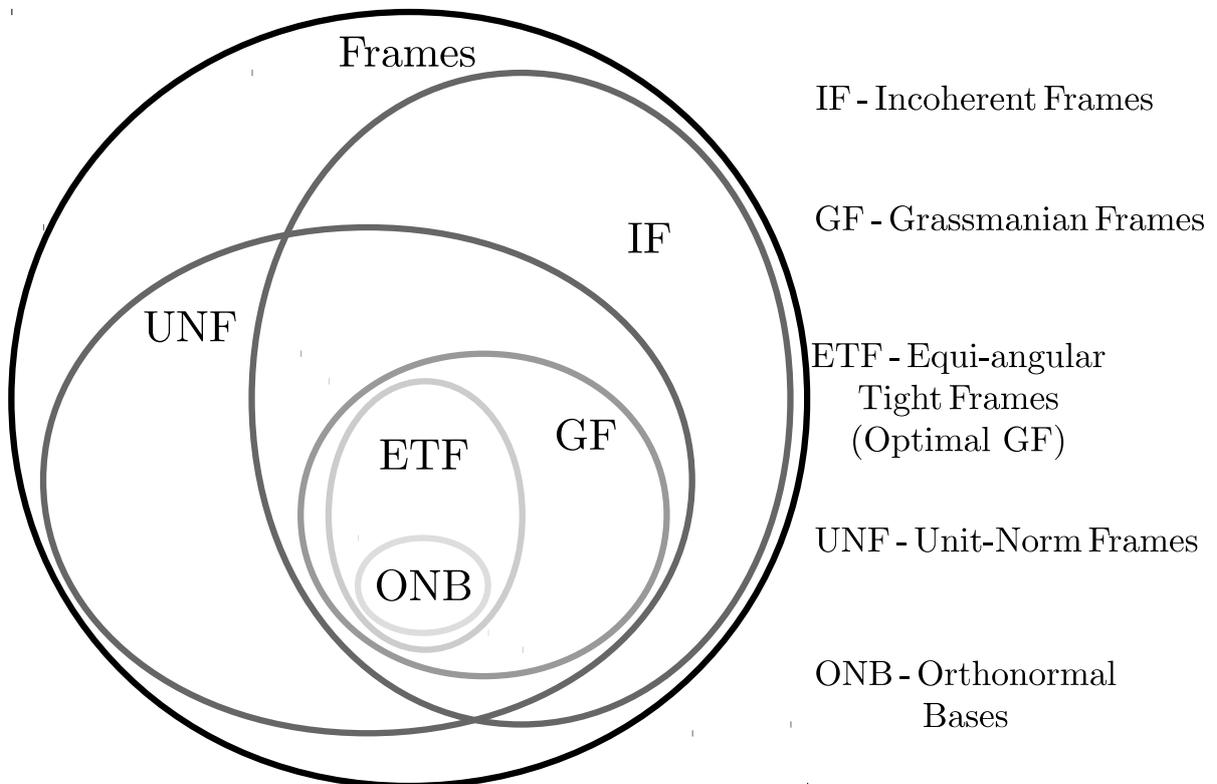


Figure 3.4: Overview of frame types – equi-normality & incoherence.

Table 3.1: MATHEMATICAL SUMMARY OF PROPERTIES OF EQUI-NORMAL & TIGHT FRAMES.

Frame	Constraints	Properties
Generic Frame	$\{\mathbf{f}_k\}_{k=\{1,2,\dots,K\}}$ linearly dependent spanning set for \mathbb{H}^M	$\alpha\ \mathbf{x}\ _2^2 \leq \sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 \leq \beta\ \mathbf{x}\ _2^2$ $\alpha\mathbf{I}_M \preceq \mathbf{S} \preceq \beta\mathbf{I}_M$ $\text{trace}(\mathbf{S}) = \sum_{i=1}^M \lambda_i = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle$
Equal-Norm Frame	$\ \mathbf{f}_k\ _2 = \zeta,$ $\forall k = \{1, 2, \dots, K\}$	$\alpha\ \mathbf{x}\ _2^2 \leq \sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 \leq \beta\ \mathbf{x}\ _2^2$ $\alpha\mathbf{I}_M \preceq \mathbf{S} \preceq \beta\mathbf{I}_M$ $\text{trace}(\mathbf{S}) = \sum_{i=1}^M \lambda_i = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle = K\zeta^2$
Tight Frame	$\alpha = \beta$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \alpha\ \mathbf{x}\ _2^2$ $\mathbf{S} = \alpha\mathbf{I}_M$ $\text{trace}(\mathbf{S}) = M\alpha = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle$
Parseval Tight Frame	$\alpha = \beta = 1$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \ \mathbf{x}\ _2^2$ $\mathbf{S} = \mathbf{I}_M$ $\text{trace}(\mathbf{S}) = M = \sum_{k \in \mathcal{K}} \langle \mathbf{f}_k, \mathbf{f}_k \rangle$
Equal-Norm Tight Frame	$\alpha = \beta$ $\ \mathbf{f}_k\ _2 = \zeta,$ $\forall k = \{1, 2, \dots, K\}$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \alpha\ \mathbf{x}\ _2^2$ $\mathbf{S} = \alpha\mathbf{I}_M$ $\text{trace}(\mathbf{S}) = M\alpha = \sum_{k \in \mathcal{K}} \langle \mathbf{f}_k, \mathbf{f}_k \rangle = K\zeta^2$ $\alpha = \frac{K}{M}\zeta^2$
Unit-Norm Tight Frame	$\alpha = \beta$ $\ \mathbf{f}_k\ _2 = 1,$ $\forall k = \{1, 2, \dots, K\}$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \alpha\ \mathbf{x}\ _2^2$ $\mathbf{S} = \alpha\mathbf{I}_M$ $\text{trace}(\mathbf{S}) = M\alpha = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle = K$ $\alpha = \frac{K}{M}$
Equal-Norm Parseval Tight Frame	$\alpha = \beta = 1$ $\ \mathbf{f}_k\ _2 = \zeta,$ $\forall k = \{1, 2, \dots, K\}$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \ \mathbf{x}\ _2^2$ $\mathbf{S} = \mathbf{I}_M$ $\text{trace}(\mathbf{S}) = M = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle = K\zeta^2$ $\zeta = \sqrt{\frac{M}{K}}$
Unit-Norm Parseval Tight Frame \Downarrow Orthonormal Basis	$\alpha = \beta = 1$ $\ \mathbf{f}_k\ _2 = 1,$ $\forall k = \{1, 2, \dots, K\}$ \Updownarrow $\langle \mathbf{f}_k, \mathbf{f}_\ell \rangle = \begin{cases} 1, & k = \ell \\ 0, & k \neq \ell \end{cases}$	$\sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle ^2 = \ \mathbf{x}\ _2^2$ $\mathbf{S} = \mathbf{I}_N$ $\text{trace}(\mathbf{S}) = M = \sum_k \langle \mathbf{f}_k, \mathbf{f}_k \rangle = K$ \Updownarrow $M = K$

3.6 Frame Applications to Wireless Communications

Naturally, by means of their definition and extrapolation of orthonormal bases to redundant spaces, frames are a salient and pervasive part of the signal domain processing pertaining to linear communication systems. As a core representational algebraic construct, frames are not necessarily cast into the spotlight of wireless signal processing albeit being a fundamental part of the latter. To name a couple of areas where frames are in fact more visible, it is worth

mentioning their applicability to Coding Theory, and respectively, Signal Representation Theory and advances in these fields.

For instance, as the study of redundant linear representations, Frame Theory represents the mathematical backbone of finite error-correction and forward-error correction codes. In fact, the incoherent frames discussed in the previous sections, and in particular the Grassmannian frames, constitute the solution behind spherical code designs and their associated sphere packing and kissing number problems [19, 22, 78]. To this end, such incoherent frames, particularly the optimal ETFs, provide the optimum redundancy and recovery guarantees for erasure and noisy distortions over lossy channels [19, 90].

Example 3.6.1. *In order to illustrate the efficacy of frames in general, and respectively, of incoherent frames in particular in adding robustness to signals in noisy or distorted conditions over the given signal space, consider the present example. Consider the AWGN channel modeled stochastically by the normal distribution $\mathbf{w} \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_M)$ such that any signal $\mathbf{x} \in \mathbb{R}^M$ is distorted as follows*

$$\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{w}, \quad (3.89)$$

leading to an SNR of the direct signal representation given by

$$\text{SNR} = \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\sigma_{\mathbf{w}}^2}. \quad (3.90)$$

Consider an alternative, redundant, frame-based representation of the signal \mathbf{x} as $\mathbf{c} = \mathbf{F}^H \mathbf{x}$. The matrix \mathbf{F} is thus either a randomly generated frame, or respectively, an incoherent UNTF following the average alternating projection method [87] listed under Algorithm 2. Over the same AWGN channel the redundant representation becomes

$$\tilde{\mathbf{c}} = \mathbf{c} + \boldsymbol{\varepsilon}, \quad (3.91)$$

where $\boldsymbol{\varepsilon}$ is once more modeled by an AWGN stochastic process with noise power σ_ε^2 such that

$$\text{SNR} = \frac{\mathbb{E}[\|\mathbf{x}\|_2^2]}{\sigma_{\mathbf{w}}^2} = \frac{\mathbb{E}[\|\mathbf{c}\|_2^2]}{\sigma_\varepsilon^2} \implies \sigma_\varepsilon^2 = \sigma_{\mathbf{w}}^2 \frac{\mathbb{E}[\|\mathbf{c}\|_2^2]}{\mathbb{E}[\|\mathbf{x}\|_2^2]}. \quad (3.92)$$

Furthermore, assume for sake of simplicity and without loss of generality that the signal has unitary average power, i.e. $\mathbb{E}[\|\mathbf{x}\|_2^2] = 1$. Moreover, by invoking the frame definition inequality (3.17) combined with the fact that the lower and upper frame bounds are in fact the lowest, and respectively, highest singular values of \mathbf{S} as previously illustrated, it follows that for a fair comparison the following condition must be satisfied

$$\sigma_\varepsilon^2 \approx \sigma_{\mathbf{w}}^2 \frac{\alpha + \beta}{2}, \quad (3.93)$$

where the approximate sign is used in reference to fact that the distribution of $\|\mathbf{c}\|_2^2$ between the frame bounds α and β is generally speaking not uniform, such that $\mathbb{E}[\|\mathbf{c}\|_2^2]$ is not truly equal to $\frac{\alpha + \beta}{2}$ but approaching it closely.

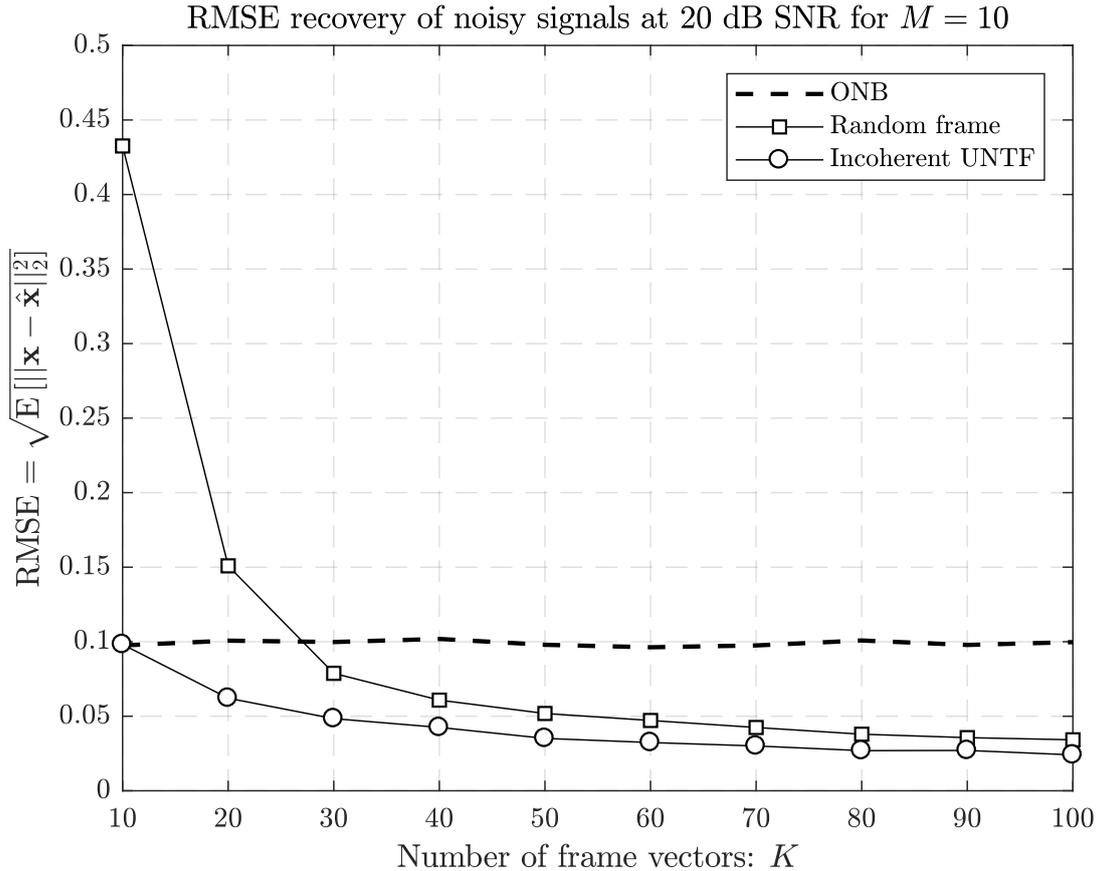


Figure 3.5: Noisy signal MSE recovery performance under different redundant representations for a fixed SNR level of 20 dB and frames of dimensionality $M = 10$ and $10 \leq K \leq 100$.

To illustrate the gains in terms of root-mean-square error (MSE) (RMSE) of the redundant bases over the conventional “orthogonal” representation, Figure 3.5 was generated for randomly selected frames, and respectively, incoherent UNTFs. The results in terms of noise reduction obtained via a purely random frame design outline that even with non-optimized design gains over the conventional non-redundant representation are obtained. However, in such a case a relatively high degree of redundancy is necessary to achieve such gains and improve over the latter. On the other hand, the incoherent UNTF is already equal to the ONB signal representation improving directly the SNR with any additional degree of redundancy added as yet another frame coefficient dimension. This result is intuitively immediate as incoherence spreads more uniformly the frame coefficients over the available signal space, whereas tightness ensures projection energy invariance regardless of the inputs or noisy distortions.

Another area of research where Frame Theory is a primary driver is the adaptive signal representation field. Under this umbrella, wavelets, Weyl-Heisenberg frames, or alternatively Gabor frames, [91] and their subsequent linear transforms may be considered [19, 22]. For instance, the Gabor frames [69] constitute a fundamental framework to decompose the inherent time-frequency information of an underlying communication system over a singular joint domain viewed physically as a time-frequency description. Under Heisenberg’s uncertainty principle, the best resolution in terms of time-frequency representation can be attained by modulation

with time-shifted versions of Gaussian probability functions [69]. Nevertheless, the most popular usage of Gabor frames in communications may be seen as a consequence of the emergence of OFDM, which is similarly a structure of time-frequency signal decomposition based on highly structured equi-spaced orthogonal subcarriers as seen in Section 2.2. In this particular case, the modulation Gaussians introduced by Gabor in his initial work lead to delta impulses, maintaining the initial probability distribution function feature of the Gabor frames [92].

A Gabor frame is formally and parametrically defined as the collection of functionals $\{f_k\}_{k=0}^K$

$$f_{k,\ell} = e^{-\frac{j2\pi k\ell}{K}} f_{0,\ell}, \quad (3.94)$$

where k denotes the frame element, $\ell \in \mathbb{Z}$ is a discrete time index and $f_{0,\ell}$ is the parametric prototype windowing or frame function.

Upon relaxing the orthogonality condition, it has been shown that Gabor frames are capable of alleviating the challenges of fast-fading mobile channels as the ones presented in the case study from Section 2.3. In fact, it has been proven in [92] that the Gabor-relaxed OFDM non-orthogonal representation leads to the optimum signal representation under rapidly-varying channels with multipath propagation. Paradoxically, this work led to the development of a new orthogonal modulation scheme, orthogonal time frequency space (OTFS) [93, 94], which leverages the fact that such channels are in fact sparse in the delay-Doppler domain, and hence, less affected by inherent high-mobility and multipath transmission effects and better suited for carrying information. The modulation is achieved by means of Heisenberg transform as an instantiation of Gabor frames [93] translating information from the delay-Doppler domain to the time-frequency domain common to physically realizable communication systems [94].

Similar representational frame-theoretic structures are also *wavelets*. These represent signals by decompositions into time-shifted and dilated versions of a particular *mother wavelet* function $f(t)$. In practice, wavelets are in fact discrete filterbanks tuned to extract particular signal information [20]. Nonetheless, as a representation tool they synthesize inherent signal information based on shifted and scaled versions of a particular wavelet function. They can be used with ease in complementarity to compress / decompress information, as well as, to extract information by exploiting the signal space features through the mother wavelet, having found applications to CS imaging, signal and video compressing or processing [19]. The interested reader can find more in-depth explanations and reviews of the fundamentals of Gabor frames and wavelets in [19, 20], and respectively, [95, 96] for instance.

3.7 Conclusions

In summary, Frame Theory provides core strong mathematical tools which can be leveraged well in the problems of signal decomposition and waveform synthesis. As a consequence, all the subsequent signal processing techniques building upon such strategies can benefit from the use of frame-theoretic approaches. As opposed to conventional signal processing, frames introduce in their decompositions redundant signal descriptions which can either leverage the signal space

or act robustly against destructive impairments occurring in practice as previously exemplified.

However, outside of the redundant representational realm, Frame Theory exhibit also compressive features that come into play in other areas of research that have recently gained increasingly more attention by their fundamental concepts and signal processing techniques, namely CS [97], and respectively, NOMA [17, 18]. Having laid down the fundamentals of Frame Theory, we shall venture next in applying such concepts to next-generation wireless communication systems such as mmWave and NOMA modems. Therefore, we study in the next two chapters two fundamental problems pertaining to such systems.

On one hand, we review and address the problem of mmWave IA and incipient beamforming management via frame-theoretic sparsely-enhanced CS channel recovery. On the other hand, we propose a full-chain, transmit and receive NOMA scheme on top of dense complex frames capable of optimizing data rates under efficient spectrum utilization.

Chapter 4

Millimeter Wave Initial Access

Over the course of last decade the wireless data rate demand has doubled almost every 18 months [98], leading to an exponential growth in the required service rates. As a result, the core technologies underlying the past and current wireless communications were always pushed to their physical limits eventually becoming saturated and depleted of communication resources as detailed shortly in Chapters 1 and 2.

The frequency spectrum, as the core resource of wireless communications, has been undoubtedly affected as well. The accentuated scarcity of the sub 6-GHz frequencies comes therefore as no surprise under such perpetually increasing demands of data rates experienced by modern wireless communications [13].

It is important to note however that such demands are not due only to mobile telephony or 3G/4G specific data traffic. In fact, these will be complemented if not heavily surpassed by uprising real-time data driven applications relying on data-hungry transmissions. Such examples are for instance virtual reality (VR)/augmented reality (AR) applications, efficient traffic flow management via smart connected vehicles, multi-viewpoint 4K live streaming for sports and electronic games or real-time data-driven massive IoT sensing networks. These, together with many other use cases not yet envisioned, are expected to push the global data consumption up to 175 zettabytes by 2025 [99], while also imposing new challenges related to massive wireless heterogeneous connectivity and low latency communications [100].

In response to the above context as to avoid a potential technological bottleneck, mmWave systems have been successfully proposed over the last years as a prospective solution [13, 100] for the future of high bandwidth mobile communications. These aim to make use of the soon to

Excerpts of this chapter are adapted and enhanced based on the selected articles:

R.-A. Stoica, G. T. Freitas de Abreu and H. Iimori, “A Frame-Theoretic Scheme for Robust Millimeter Wave Channel Estimation,” in *IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Aug. 2018, pp. 1-6. ©2018 IEEE

R.-A. Stoica, G. T. Freitas de Abreu, “Frame-theoretic Precoding and Beamforming Design for Robust mmWave Channel Estimation,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2019, to appear. ©2019 IEEE

be licensed¹ millimeter frequency bands, *i.e.* the portion of the spectrum formally referred to the bands between 24 GHz and 300 GHz, for the realization of fast, multi-Gigabit-per-second (Gbps) data transfers over wireless environments. Most probably, such endeavors in the high frequency bands will open further the interest and study field of submillimeter communications, *i.e. terahertz (THz) communications*, thus filling the gap between what has recently become standalone mmWave communications to visible light communications systems [101]. Taking therefore into account the very similar physical propagation properties of RF mmWave systems to those in higher bands, such as THz or optical communications, it is hence of high interest to study the mmWave not from their isolated perspective and success in 5G and 5G+ alone, but, furthermore, also in the light of next-generation networks to come.

As remarked by the power of example in Section 2.3, for the case of IEEE 802.11p protocol, the performance of a communication link is generally mirrored by the particularities of its physical RF propagation under the assumed communication model. To this end, the mathematical study and modeling of the mmWave channels is similarly a key component of any mmWave system. As a result, this Chapter is dedicated to the study of channel estimation and channel recovery prior to the establishment of communication links, *i.e.* under the stage of *initial access* or *early access*, where it is assumed that both the Tx and Rx have no appropriate knowledge of instantaneous or statistical CSI.

4.1 System Model and Considerations

Besides the argument of still available free frequency spectrum, mmWave communications present other attributes that make them highly desirable for eMBB 5G/5G+ applications, but also for low-latency communications.

However, note that first and foremost, mmWave are expected to suffer from severe signal path attenuation given the Friis free-space propagation law [2, 102], since the received signal power decreases proportionally to the squared carrier wavelength. Furthermore, as opposed to sub-6 GHz, mmWave may also experience path blockages due to even small or thin objects, such as people or windows. Also, mmWave is highly prone to diffraction and absorption under atmospheric phenomena such as severe rain or fog [103] given the millimeter-level propagation wavelength and specific gaseous mixes resonance frequencies.

Nonetheless, *the curse of the millimeter wavelengths* is also a blessing [12, 15]. In fact, the reduced wave propagation periodicity allows Tx and Rx architectures to employ larger antenna arrays than for traditional sub-6 GHz systems, which in conjunction with specific MIMO techniques are capable of combating the path loss effects previously described [15]. For instance, already large-scale prototype RF mmWave interfaces have demonstrated robust and reliable high data rate wireless transfers in controlled environments [104].

The MIMO techniques together with their potential massive scale realizations come furthermore with the added benefit of high directionality communication beams. This latter advantage

¹In most of the countries around the globe the unlicensed mmWave spectrum is at the moment of the writing of this dissertation undergoing the operator auction for licensing deals.

can be thus further leveraged into controlling and steering communications depending on the users to be served such that various QoSs and interference avoidance constraints are jointly satisfied thus optimizing potential communication rates.

All of the above features to be exploited on top of the mmWave links come packaged as a broadband communication system capable of fast communications given the wideband spectrum available above 24 GHz. Furthermore, as a result of such high-bandwidths and carrier frequencies, the associated RF architectures of mmWave transceivers have been a central point of research over the last years. Central to MIMO systems in sub-6 GHz, the fully digital beamforming designs are known to be the optimum [15]. However, these have been coined by the industry and academia alike as being infeasible for mmWave radios given their high costs of implementation motivated by the high bandwidth accurate analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) respectively, but also by the high-precision large dynamic-range power amplifiers (PAs) necessary to implement the RF chains [105, 106]. To solve this cost issue and to reduce the number of RF chains while maintaining a large number of antennas, the idea of a hybrid RF beamforming architecture has been universally proposed [14, 107, 108]. The core mechanism therefore proposed to combine the digital baseband processing given a reduced amount of RF chains with the analog processing realized by cheaper phase shifters. Considered as a breakthrough method for the feasibility of mmWave radios implementation, the idea motivated further research on several new hybrid beamforming techniques. Amongst such works, of importance are for instance the *manifold optimization alternating minimization* proposed in [109], the *fixed phase shifter alternating minimization* proposed in [110] or the *hybrid design by least squares relaxation* method [111], respectively.

However, across the discussed literature above, it can be noticed a recurring theme, *i.e.* that such hybrid beamforming designs are in fact heavily influenced by the quality of CSI. In other words, still many such contributions rely in fact on perfect CSI knowledge at either Tx/Rx or both. This remark does not go to say that hybrid beamforming schemes subjected to imperfect CSI do not exist. There are in fact articles that studied in depth during recent years the imperfect CSI based hybrid beamforming, *e.g.* [14, 112, 113]. The latter showed in fact the losses and performance decays resulted as the limitations of such practical imperfections. To this end, the problem of accurate channel estimation schemes for empowering hybrid beamforming designs and establishing practical mmWave communications links capable of robust multi-Gbps communications is currently a very popularly researched topic.

The beamforming and rapid channel changes of the mmWave propagation paths require a joint directional communication estimation and tracking similar at its core to the concepts exposed in Section 2.3 for the case of IEEE 802.11p communications. To this end, the procedure of estimation is in practice reduced to the idea of beamforming management. The latter transmit/receive vectors need to be controlled to maximize the gain and therefore sensing of the channel realizations, but similarly, also to adapt to the experienced changes given the multipath propagation and non-stationarity of scatterers or communication nodes. Therefore, the first steps of the communication establishment process, *i.e.* the IA, is quintessential.

The directional link establishment and subsequent IA can be summarized at a high-level [114]

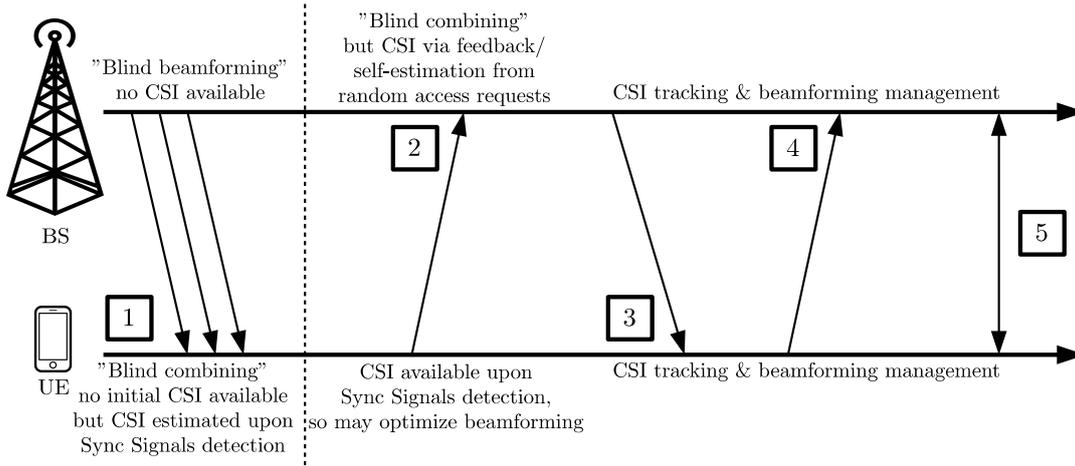


Figure 4.1: Simplified protocol diagram of initial access in mmWave systems [114].

based on the protocol time diagram in Figure 4.1. The stages are summarized therefore as:

1. *Synchronization signal detection* where each cell periodically broadcasts over its sectors synchronization signals allowing user equipment (UE) to detect the BS and recover down-link frame timings.
2. *Random access preamble selection & transmission* where the UE accesses one of the existent dedicated slots for random access messages previously synchronized upon and radiates information to the BS.
3. *Random access response* in which the BS acknowledges the random access index requested by the UE and grants subsequent access.
4. *Connection request* of the UE to the BS for scheduling and establishing a communication link.
5. *Scheduled communication* between UE and BS in UL or DL depending on various application needs.

As mentioned earlier, in the light of Figure 4.1, it is now clearer that the very first steps of the IA cell broadcast advertising and respectively, random access request are the most sensitive given the fact that full beamforming gain on both BS/UE sides is not possible. At most, only after the first step the UE can estimate the DL channel and subsequently adjust its beamformers, whilst similarly, after only the second step the same can be said about the BS. As a result, the training beamforming prior to these optimizations is in fact a fundamental problem given the propagation effects of mmWave systems and needs to be treated jointly with the channel estimation procedure in IA. The objective of the next sections is therefore the critical mathematical treaty of the first step of IA from the joint perspective of channel estimation and optimized training beamforming.

The mmWave initial CSI estimation undergoes a MIMO mmWave communication link given a Tx radio equipped with T transmit antennas, and respectively, a Rx one with R receive antennas. Furthermore, it is assumed that the channel estimation is aided by prescribed training

symbols, beamformers and combiners. Concretely, the setup is based on M_T training beamforming vectors on the Tx side transmitting a known training signal sequence $\mathbf{S} \in \mathbb{C}^{M_T \times M_T}$, and respectively, on M_R combining vectors on the Rx side applied to each training beamformer. The received and sensed signal matrix $\mathbf{Y} \in \mathbb{C}^{M_R \times M_T}$ is next used to estimate the MIMO wireless channel $\mathbf{H} \in \mathbb{C}^{R \times T}$.

The mathematical model of the system is given by

$$\mathbf{Y} = \mathbf{V}^H \mathbf{H} \mathbf{U} \mathbf{S} + \mathbf{N}, \quad (4.1)$$

where $\mathbf{U} \triangleq [\mathbf{u}_1, \dots, \mathbf{u}_{M_T}] \in \mathbb{C}^{T \times M_T}$ and $\mathbf{V} \triangleq [\mathbf{v}_1, \dots, \mathbf{v}_{M_R}] \in \mathbb{C}^{R \times M_R}$ are the precoding and the combining training beamforming vectors respectively.

The thermal noise subsequent to typical hardware impairments of RF chains is denoted by $\mathbf{N} \in \mathbb{C}^{M_R \times M_T}$ and modeled by the circularly symmetric complex AWGN, and \mathbf{H} is the mmWave channel matrix whose estimation is the target.

Based on the current literature on mmWave channel modeling, *i.e.* [107, 115–120], the wireless medium \mathbf{H} is mathematically described by

$$\mathbf{H} = \sqrt{\frac{TR}{L}} \sum_{\ell=1}^L \gamma_{\ell} \mathbf{a}_r(\phi_{\ell}^r) \mathbf{a}_t^H(\phi_{\ell}^t), \quad (4.2)$$

where L denotes the number of propagation paths, $\gamma_{\ell} \sim \mathcal{CN}(0, \sigma_{\gamma}^2)$ is the complex gain of the ℓ -th path, and respectively, $\mathbf{a}_r(\phi_{\ell}^r)/\mathbf{a}_t(\phi_{\ell}^t)$ are the array response vectors at the receiver/transmitter, with subsequent angles of arrival (AoA)/angles of departure (AoD)² denoted by $\phi_{\ell}^r, \phi_{\ell}^t \in [0, 2\pi]$.

An advanced high-resolution representation of the above channel based on clusterized propagation paths following the broadband Saleh-Valenzuela model [121],

$$\mathbf{H} = \sqrt{\frac{TR}{LC}} \sum_{k=1}^C \sum_{\ell=1}^L \gamma_{k,\ell} \mathbf{a}_r(\phi_{k,\ell}^r) \mathbf{a}_t^H(\phi_{k,\ell}^t), \quad (4.3)$$

has also been considered in the literature [14, 122], where C denotes the small number of clusters, and L marks similarly as above the separable paths.

Nonetheless, given the communication context in which CSI estimation is attempted, and the fact that prior knowledge of statistical CSI or quasi-optimized beamforming is not available, there is no point in aiming at a high-resolution channel detection on top of (4.3), but rather focus on the “narrowband” equivalent from Equation (4.2).

4.2 Channel Estimation as a Sparse Recovery Problem

Unfortunately, the fast channel variations due to the high carrier frequencies and high band-

²For simplicity of exposition and notation but without loss of generality the AoA/AoD are given by the azimuth direction characteristic to linear phased arrays, rather than considering also elevation specific to planar phased array surfaces.

widths characteristic to mmWave systems impede the time domain sampling and representation of the communication channel \mathbf{H} . On the other hand however, it has been shown in [107] that the mmWave channel is in fact geometrically sparse over the spatial RF antenna array domain. As a result, the channel estimation can be casted as a CS sparse vector recovery problem on top of the equivalent spatial channel representation [107]

$$\mathbf{H} = \mathbf{A}_R \mathbf{H}_\gamma \mathbf{A}_T^H, \quad (4.4)$$

where the following matrices have been implicitly and compactly defined

$$\mathbf{A}_R \triangleq [\mathbf{a}_r(\phi_1^r), \dots, \mathbf{a}_r(\phi_L^r)] \quad (4.5)$$

$$\mathbf{A}_T \triangleq [\mathbf{a}_t(\phi_1^t), \dots, \mathbf{a}_t(\phi_L^t)] \quad (4.6)$$

$$\mathbf{H}_\gamma \triangleq \sqrt{\frac{TR}{L}} \text{diag}(\gamma_1, \dots, \gamma_L) \quad (4.7)$$

By such means and under the influence of CS literature, the mmWave channel estimation problem has been thus reduced to different pursuit algorithms, *e.g.* orthogonal matching pursuit (OMP) [123, 124], basis pursuit denoising (BPDN) [125], as for instance in [120, 126, 127]. More recently, matrix norm recasting of the sparse recovery pursuit problems has also been addressed as a mean of CSI estimation via matrix completion of jointly sparse and low-rank structures with side information [128–131]. Albeit very interesting and promising, these approaches either assume particular RF architectures, requiring for instance additional high-speed RF switches on the RF paths, or consider just randomized non-optimal training beamforming [114].

Let us next derive the complete mathematical formulation of the mmWave channel estimation CS sparse recovery problem. Firstly, without loss of generality it is assumed that the training piloting symbols \mathbf{S} are in fact orthogonal and normalized, *i.e.* $\mathbf{S} = \mathbf{I}_{M_T}$. In effect this detaches the channel estimation problem from the symbol piloting and shifts the piloting effort to the beamspace. As a result, the beamforming management and optimization for training during channel acquisition becomes key to an accurate estimation. Given the previous assumptions the linear system in (4.1) is vectorized by the Kronecker product and its algebraic rules [43] to the identity

$$\mathbf{y} \triangleq \text{vec}(\mathbf{Y}) \quad (4.8)$$

$$= \text{vec}(\mathbf{V}^H \mathbf{A}_R \mathbf{H}_\gamma \mathbf{A}_T^H \mathbf{U}^T) + \text{vec}(\mathbf{N}), \quad (4.9)$$

$$= (\mathbf{U}^T \otimes \mathbf{V}^H)(\mathbf{A}_T^* \otimes \mathbf{A}_R) \text{vec}(\mathbf{H}_\gamma) + \text{vec}(\mathbf{N}), \quad (4.10)$$

such that the receive matrix \mathbf{Y} is reshaped to $\mathbf{y} \in \mathbb{C}^{M_R M_T \times 1}$.

To decouple the processing from the antenna signal space, but also to discretize and to build the CS dictionary matrix, an uniform gridding and quantization of the antenna space to the virtual beamspace has been proposed in the seminal work of [107]. Thus, the angular antenna domain is gridded and quantized both at the Tx and at the Rx using G_T , and respectively, G_R

quantization levels, such that $\mathbf{A}_T, \mathbf{A}_R$ are mapped to their virtually digitized correspondents

$$\mathbf{A}_R \mapsto \mathbb{A}_R \triangleq [\mathbf{a}_{q,r}(\theta_0^r), \dots, \mathbf{a}_{q,r}(\theta_{G_R-1}^r)] \in \mathbb{C}^{R \times G_R} \quad (4.11)$$

$$\mathbf{A}_T \mapsto \mathbb{A}_T \triangleq [\mathbf{a}_{q,t}(\theta_0^t), \dots, \mathbf{a}_{q,t}(\theta_{G_T-1}^t)] \in \mathbb{C}^{T \times G_T}, \quad (4.12)$$

given that the transformed AoA/AoD beamspace steering vectors $\mathbf{a}_{q,r}/\mathbf{a}_{q,t}$ radiate signals over the following discrete angular inputs

$$\theta_{g_r}^r \triangleq \frac{2\pi \cdot g_r}{G_R}, \quad \forall g_r \in \{0, 1, \dots, G_R - 1\} \quad (4.13)$$

$$\theta_{g_t}^t \triangleq \frac{2\pi \cdot g_t}{G_T}, \quad \forall g_t \in \{0, 1, \dots, G_T - 1\}. \quad (4.14)$$

Provided the binning and quantization operations from above, the equivalent channel representation over the discretized spatial space is returned by

$$\mathbf{H} \approx \mathbb{A}_R \mathbb{H}_\gamma \mathbb{A}_T^H, \quad (4.15)$$

up to the accuracy of the binned AoAs and AoDs scatterers, *i.e.* \mathbb{A}_R and \mathbb{A}_T .

As a consequence of Equation (4.15), the physical gain wireless channel diagonal component $\mathbf{H}_\gamma \in \mathbb{C}^{L \times L}$ is further sparsified to its equivalent version $\mathbb{H}_\gamma \in \mathbb{C}^{G_R \times G_T}$. The latter \mathbb{H}_γ contains in fact just L non-zero values corresponding to the grid positions $(i, j$ in $(\mathbb{A}_R, \mathbb{A}_T)$ where the discrete beam angles $\theta_{g_r}^r$ and $\theta_{g_t}^t$ are minimally close to the true physical AoAs and AoDs realizations, namely

$$\mathbb{H}_\gamma(i, j) = \gamma_\ell \Leftrightarrow \begin{cases} \|\theta_i - \phi_\ell^r\|_2 < \|\theta_{g_r \neq i} - \phi_\ell^r\|_2 \\ \|\theta_j - \phi_\ell^t\|_2 < \|\theta_{g_t \neq j} - \phi_\ell^t\|_2 \end{cases}, \quad (4.16)$$

for all the paths $\ell \in \{1, 2, \dots, L\}$.

A L -sparse representation of the original channel matrix \mathbf{H} over the spatial beamspace domain is therefore obtained provided that the quantization is sufficiently fine, *i.e.* $G_T \times G_R \gg L$. The complete noisy sparse linear representation associated therefore with the mmWave channel estimation problem is obtained as

$$\mathbf{y} = \underbrace{(\mathbf{U}^T \otimes \mathbf{V}^H)}_{\Phi} \underbrace{(\mathbb{A}_T^* \otimes \mathbb{A}_R)}_{\Psi} \underbrace{\text{vec}(\mathbb{H}_\gamma)}_{\mathbf{h}} + \underbrace{\text{vec}(\mathbf{N})}_{\mathbf{n}}, \quad (4.17)$$

where in the light of the CS nomenclature, [97], the following matrices and vectors are identified:

- $\Phi \triangleq (\mathbf{U}^T \otimes \mathbf{V}^H) \in \mathbb{C}^{M_T M_R \times T R}$ as the *measurement matrix*;
- $\Psi \triangleq (\mathbb{A}_T^* \otimes \mathbb{A}_R) \in \mathbb{C}^{T R \times G_T G_R}$ as the *sparse dictionary*;
- $\mathbf{h} \triangleq \text{vec}(\mathbb{H}_\gamma) \in \mathbb{C}^{G_T G_R \times 1}$ as the *L -sparse channel vector*.

Under the above sparse representation (4.17), it is immediate that the mmWave channel estimation problem is in fact equivalent to the generic noisy sparse recovery optimization problem

with equality constraint of CS, *i.e.*

$$\underset{h \in \mathbb{C}^{G_T G_R}}{\text{minimize}} \quad \|h\|_0, \quad (4.18a)$$

$$\text{subject to} \quad y = \underbrace{\Phi \Psi}_{\Omega} h + \mathbf{n}, \quad (4.18b)$$

which is solvable under the premises of the CS framework [97], having additionally denoted the subsequent *sensing matrix* as $\Omega \in \mathbb{C}^{M_T M_R \times G_T G_R}$.

4.3 Beamforming Management via Frame Theory

Since $\Phi = \mathbf{U}^T \otimes \mathbf{V}^H$, it is clear based on (4.17) - (4.18b) that the channel sampling and subsequent recovery accuracy is directly influenced jointly by the training beamforming patterns at the Tx side, *i.e.* \mathbf{U} , and at the Rx side *i.e.* \mathbf{V} . Furthermore, since the channel \mathbf{H} is represented under its quantized spatial beamspace model (4.11) - (4.15), the sparse dictionary $\Psi = \mathbb{A}_T^* \otimes \mathbb{A}_R$ is also central to the achievable estimation performance, and these remarks are visible by virtue of the sensing matrix $\Omega \triangleq \Phi \Psi$. However, the sparse dictionary Ψ is in fact subjected to the fixed harmonic geometry of the existent antenna arrays at the Tx/Rx side, and furthermore composed on top of the quantized AoDs/AoAs components respectively. So, from a practical perspective only the grid resolution could be influenced for Ψ , but not its geometric structure.

It follows therefore naturally to investigate the question: *what are the optimum training Tx/Rx beamforming vectors*. To this objective, as previously explained, the IA sensitive first communication round is the problem context, where *no a priori or robust contextual information is available* to enhance the beamforming. The central problem treated therefore next is of the *distributed optimized beamforming design with no side information* for initial channel estimation.

4.3.1 Sparse Recovery Guarantees and Requirements

Throughout this subsection the generic CS noisy linear system is considered

$$\mathbf{r} = \underbrace{\Phi_s \Psi_s}_{\Omega_s} \mathbf{s} + \mathbf{n}, \quad (4.19)$$

for the sake of upcoming discussion; \mathbf{r} is the M -dimensional compressed measured vector, \mathbf{s} is the K -dimensional original undersampled signal vector, Ψ_s is the sparsifying dictionary projection, Φ_s is the measurement matrix, \mathbf{n} is the AWGN distortion within the system, and respectively, $\Sigma_s \triangleq \{\mathbf{s} \mid \|\mathbf{s}\|_0 = s\}$.

Moreover, the sparse recovery is assumed given the context of canonical ℓ_0 -norm,

$$\underset{\mathbf{s}}{\text{minimize}} \quad \|\mathbf{s}\|_0, \quad (4.20a)$$

$$\text{subject to} \quad \mathbf{r} = \Omega_s \mathbf{s}, \quad (4.20b)$$

and respectively, ℓ_1 -norm minimization

$$\underset{\mathbf{s}}{\text{minimize}} \quad \|\mathbf{s}\|_1, \quad (4.21a)$$

$$\text{subject to} \quad \mathbf{r} = \mathbf{\Omega}_s \mathbf{s}. \quad (4.21b)$$

problems.

Unfortunately, universally strong and quantifiable sparse recovery guarantees do not exist for noisy linear systems [132] such as the underlying mmWave system model currently under scrutiny. In spite of this deficiency however, s -sparse recovery guarantees are spread throughout the CS literature regarding noiseless linear systems, see for instance [67, 97, 132–134] and references therein for an exhaustive treatment of subsequent sparse reconstruction requirements. Despite the strong assumptions impractical to real-life noisy systems of these references, they provide powerful and intuitive perspectives of the sparse recovery limits and how associated sensing matrices for CS should be designed. For instance, a natural perspective into the uniqueness of the s -sparse recovery of \mathbf{s} stems from the spark of $\mathbf{\Omega}$, *i.e.*, $\text{spark}(\mathbf{\Omega})$.

Definition 4.3.1. *The spark of a matrix $\mathbf{A} \in \mathbb{H}^{M \times K}$ is defined as the minimum number of linearly dependent column vectors.*

The connection between the s -sparse uniqueness and the spark is made clear by the following Theorem and its proof.

Theorem 4.3.1 ([133, Corr. 1]). *Given $\mathbf{\Omega}$ a $M \times K$ sensing matrix, then there exists an unique s -sparse solution to the linear system $\mathbf{r} = \mathbf{\Omega} \mathbf{s}$ if and only if*

$$\text{spark}(\mathbf{\Omega}) > 2s \quad (4.22)$$

Proof: Consider the forward implication. Given that exists s -sparse unique solution \mathbf{s} , assume that $s \geq \frac{1}{2}\text{spark}(\mathbf{\Omega})$. Then, since the $\text{spark}(\mathbf{\Omega}) \in \{2, \dots, M + 1\}$ can equivalently be seen as

$$\text{spark}(\mathbf{\Omega}) = \min\{s \mid \ker(\mathbf{\Omega}) \cap \Sigma_s \neq \{\mathbf{0}\}\}, \quad (4.23)$$

there exists $\mathbf{q} \neq \mathbf{0} \in \ker(\mathbf{\Omega})$ such that $\|\mathbf{q}\|_0 \leq 2s$, which means that there exist two different vectors \mathbf{s}, \mathbf{s}' such that $\mathbf{q} = \mathbf{s} - \mathbf{s}'$, and respectively, $\|\mathbf{s}\|_0 = \|\mathbf{s}'\|_0 \leq s$. As a consequence, $\mathbf{\Omega} \mathbf{s} = \mathbf{\Omega} \mathbf{s}'$ which violates the uniqueness constraint, and hence the forward direction is proven.

The reverse implication follows. Under the fact that (4.22) holds true, it follows that $\text{spark}(\mathbf{\Omega}) > 2s$. Consider \mathbf{s}, \mathbf{s}' solutions with $\|\mathbf{s}\|_0 \leq s, \|\mathbf{s}'\|_0 \leq s$ such that $\mathbf{r} = \mathbf{\Omega} \mathbf{s} = \mathbf{\Omega} \mathbf{s}'$. It follows therefore that $\|\mathbf{s} - \mathbf{s}'\|_0 \leq 2s < \text{spark}(\mathbf{\Omega})$. Consequently, $\mathbf{s} = \mathbf{s}'$ which proves the reverse direction.

□

The above theorem guarantees successful unique recovery of \mathbf{s} , under the noiseless minimization of $\|\mathbf{s}\|_0$ to its global minimum s . A more involved and rich condition for the sparse recovery linking the recovery via ℓ_0 -norm and ℓ_1 -norm minimization and therefore assuring their equivalence [135] is the restricted isometry property (RIP) introduced in [136].

Definition 4.3.2 (Restricted isometry property). *Let $\mathbf{\Omega}$ be an $M \times K$ sensing matrix. Then $\mathbf{\Omega}$ has the restricted isometry property of order s , if there exists a $\delta_s \in (0, 1)$ such that*

$$(1 - \delta_s)\|\mathbf{s}\|_2^2 \leq \|\mathbf{\Omega}_s \mathbf{s}\| \leq (1 + \delta_s)\|\mathbf{s}\|_2^2 \quad (4.24)$$

for any $\mathbf{s} \in \Sigma_s$.

Based on the similarity between the RIP Definition 4.3.2 the Frame Definition 3.2.1 (up to transposition of $\mathbf{\Omega}_s$), it is easy to remark that in fact the sensing matrix needs to be a sensing frame spanning the measurements space of dimensionality M . Centering the bounds in (4.24) around 1 is purely a matter of analytic convenience. In fact, any bounds $0 < \alpha_s, \beta_s < \infty$ could be used in (4.24) above such that

$$\alpha_s \leq \|\mathbf{s}\|_2^2 \leq \|\mathbf{\Omega}_s \mathbf{s}\| \leq \beta_s \|\mathbf{s}\|_2^2, \quad (4.25)$$

for all $\mathbf{s} \in \Sigma_s$, as under appropriate rescaling of $\mathbf{\Omega}_s$ the bounds can be refit and recentered around 1.

The latter remark is easy to prove as for the general rescaling of $\mathbf{\Omega}_s$ in (4.25) above by

$$\tilde{\mathbf{\Omega}}_s = \sqrt{\frac{2}{\alpha_s + \beta_s}} \mathbf{\Omega}_s \quad (4.26)$$

the byproduct $\tilde{\mathbf{\Omega}}_s$ still satisfies the RIP property (4.24) such that

$$\delta_s = \frac{2\alpha_s}{\alpha_s + \beta_s} \in (0, 1). \quad (4.27)$$

The RIP property is therefore, as seen in the case of (3.17), a way to express the requirements on the minimum and maximum singular values of *all* the $M \times s$ submatrices contained within $\mathbf{\Omega}_s$. This is achieved based on the constant $\delta_s \in (0, 1)$, namely the restricted isometric constant (RIC) [97], which stands for the smallest non-negative scalar under which the inequalities in (4.24) are fulfilled. As an example, for $\delta_s \approx 0$ the sparse recovery guarantee of any s -sparse vectors is ensured provided that $\mathbf{\Omega}_s$ is as close as possible to an orthogonal representation [67, 136], whereas on the other hand, as $\delta_s \approx 1$ the guarantee of sparse recovery vanishes. Intuitively, the latter remark is very clear as for the extreme case $\delta_s = 1$ there exists an s -sparse vector $\mathbf{s} \in \Sigma_s \cap \ker(\mathbf{\Omega}_s)$ whose support cannot be recovered anymore by $\mathbf{\Omega}_s$. It follows therefore, that the RIP criterion is one sufficient measure for robust sparse recovery given a sparsity level s for any prospective sensing matrix $\mathbf{\Omega}_s$.

Nonetheless, trying to verify the RIP for a fixed $M \times K$ sensing matrix $\mathbf{\Omega}_s$ becomes in fact computationally intractable being a combinatorial problem. To this end, random matrix theory has been used to propose generic matrices capable of providing uniform guarantees on the base of the RIP [135] for both ℓ_0 - or ℓ_1 -centered recovery of \mathbf{s} . It is therefore known that independent, identically distributed (i.i.d) Gaussian, binary matrices, alongside subsampled Fourier matrices offer such guarantees, fulfilling the RIP with high probability for low values of δ_s , thus assuring uniform robust recovery for low number of measurements M [67, 135]. The salient signal information is thus generally well captured asymptotically by Gaussian randomized sensing or

harmonic structures that are capable of uniformly spreading the sampling of measurements over the entire sparse signal space. However, the asymptotic inherent component may not be applicable to practical systems with reduced signal space dimensionality. As a result, more general conditions for generally good sensing matrices are desirable, and one of these is in fact tightness, more concretely, UNTF sensing [137, 138]. Tightness, in general, and UNTFs, in particular, represent good sensing matrices given they are capable of enforcing the RIP property uniformly across various levels of sparsity s given that their M rows are orthogonal.

However, UNTFs alone may fail to capture the available salient information of the subsampled signals since the RIP is only sufficient for successful recovery. To address this issue, an additional general requirement on the sensing matrix is *incoherence*. As a matter of fact, this property is actually related to the spark of the sensing matrix by means of the general inequality

$$\text{spark}(\mathbf{\Omega}_s) \geq 1 + \frac{1}{\mu(\mathbf{\Omega}_s)} \quad (4.28)$$

as a corollary of the Gershgorin's Circle Theorem [135].

In (4.28) above, $\mu(\mathbf{\Omega}_s)$ is nothing else but the mutual coherence (3.64) introduced in Section 3.4. This inequality is also the proof behind [133, Th. 12], listed for convenience below.

Theorem 4.3.2 ([133, Th. 12]). *If*

$$s < \frac{1}{2} \left(1 + \frac{1}{\mu(\mathbf{\Omega}_s)} \right) \quad (4.29)$$

then for all measurement vectors \mathbf{r} exists a unique signal $\mathbf{s} \in \Sigma_s$ solving the linear system $\mathbf{r} = \mathbf{\Omega}_s \mathbf{s}$.

The unique sparse recovery guarantee behind Theorem 4.3.2 is applicable both to the ℓ_0 - and ℓ_1 -associated recovery problems, although the theorem does not imply that the solutions also coincide, [133].

Furthermore, the mutual coherence can similarly be linked to the RIC of any sensing matrix $\mathbf{\Omega}_s$ with unit-norm columns that meets the RIP of sparsity order s under the identity [135, Lem. 1.5]

$$\delta_s = (s - 1)\mu(\mathbf{\Omega}_s), \quad \forall s < \frac{1}{\mu(\mathbf{\Omega}_s)}. \quad (4.30)$$

Lastly, since often $\mathbf{\Omega}_s$ is factorized into the product of a measurement matrix and a sparsifying dictionary, it is desirable therefore that $\mu(\mathbf{\Phi}_s \mathbf{\Psi}_s)$ is in fact equivalently minimized in terms of worst-case coherence. However, mutual coherence varies under plain matrix multiplication and does not preserve any particular mathematical structure or properties either. As a result, to theorize the CS basics, Candès introduced the *relative coherence* concept for orthonormal bases [67],

$$\mu_r(\mathbf{\Phi}_\perp, \mathbf{\Psi}_\perp) \triangleq \sqrt{K} \max_{1 \leq i, j \leq K} |\mathbf{\Phi}_\perp^H \mathbf{\Psi}_\perp|, \quad (4.31)$$

where the $K \times K$ square matrices $\mathbf{\Phi}_\perp, \mathbf{\Psi}_\perp$ denote the orthonormal linear projections of the *complete linear measurement system* to be subsampled under CS techniques.

Under straightforward linear algebraic computations [67], it follows that $1 \leq \mu_r(\Phi_\perp, \Psi_\perp) \leq \sqrt{K}$. Since the problem of minimizing (4.31) is once again combinatorial in nature, the general strategies adopted across CS literature [19, 67, 134, 135, 137, 138] are to *use harmonic structures for sparse projections* of the original, potentially non-sparse, signals in conjunction with *randomized incoherent measurement matrices* capable of extracting the salient information necessary for the sparse signal recovery.

Motivated by this CS hiatus, let us return now to the problem in hand listed as (4.18). Recall that Ψ is in fact jointly composed of the Kronecker product of harmonically structured matrices by following the phased antenna array designs, and sampled in the beamspace domain. To this end, Ψ is a highly harmonic structure, a frame with particular properties yet to be discovered and discussed in the sequel. For now, acknowledge just the fact that Ψ contains the high-level harmonic structure desired in relation to the relative coherence concept [67], and, as a consequence upon the above arguments, the optimization may solely target Φ independently of Ψ .

Having identified Φ , Ψ , and respectively, their product Ω as frames, the core frame-theoretic ideas to optimize and manage the training beamforming are enumerated below in the light of the previous argumentation:

1. Φ must be as incoherent as possible to reduce the contamination of salient information extracted via the projection of the dictionary Ψ , and thus enhance sparsity;
2. Ω must satisfy the RIP with tight bounds, *i.e.* $\delta_s \ll 1$, (or the equivalent relaxation, $\Omega\Omega^H$ should approach a diagonal structure – *i.e.*, Ω should be tight frame, [137, 139]) to provide robustness to the sparse recovery via either ℓ_0 -norm or ℓ_1 -norm minimization routines.

In the mmWave channel estimation literature, the problem of initial beamforming optimization during the first step of IA considered hereafter has not been regarded in depth. In fact, standards and research alike use sector sweeps to precompute the dominant signal direction for IA and then this is followed by successive tracking and beamforming management [114], thus increasing latency. Alternatively, in [127] the time-domain training sequences leading to sub-optimal performance have been derived. Similarly in [122, 140] training beamformers spanned by PN-sequences have been proposed, whereas typical massive random beamforming has been generally utilized as a blunt force but practical and good solution [114]. However, all these schemes failed to provide a closed circle and a mathematical explanation of their inherent design criteria for the training sequences. The latter are in fact directly linked for Frame Theory and CS reconstruction guarantees. All approaches from above targeted actually incoherent training constructions of either time symbols or beamspace vectors but failed to attain optimal structures given the available space dimensions.

4.3.2 A Parseval Tight Frame Approach

One step in the right direction highlighted by the previously listed rules has been presented in [126]. The authors promote the idea of a Parseval tight frame (PTF)-based optimization for

the entire sensing matrix $\mathbf{\Omega}$ in order to determine the optimized joint training beamforming vectors and provide a lower bound of the achievable estimation accuracy. Recall furthermore from Table 3.1 and Figure 3.3 that a PTF is both tight and Parseval energy preserving, *i.e.*, $\alpha = \beta = 1$.

Mathematically, this approach, named hereby *PTF-beamforming* for later comparison convenience, can be summarized to solving the optimization problem [126]

$$\underset{\mathbf{\Phi}}{\text{minimize}} \quad \|\mathbf{\Phi}\|_{\text{F}}^2 \quad (4.32a)$$

$$\text{subject to} \quad \mathbf{\Phi}\mathbf{\Psi}\mathbf{\Psi}^{\text{H}}\mathbf{\Phi}^{\text{H}} = \mathbf{I}_{M_{\text{T}}M_{\text{R}}}. \quad (4.32b)$$

Interestingly, the above problem has actually an analytic closed-form global solution derived in [138] as

$$\mathbf{\Phi}^* = \mathbf{P}_{\perp}\mathbf{\Lambda}_{\mathbf{\Psi}}\mathbf{I}_{TR}\mathbf{W}_{\mathbf{\Psi}}^{\text{H}}, \quad (4.33)$$

under the assumptions that \mathbf{P}_{\perp} is *any arbitrary orthonormal matrix*, $\mathbf{W}_{\mathbf{\Psi}}$ is the left singular matrix of $\mathbf{\Psi}$, and respectively, $\mathbf{\Lambda}_{\mathbf{\Psi}}$ is given by the identity

$$\mathbf{\Lambda}_{\mathbf{\Psi}} \triangleq \begin{bmatrix} \frac{1}{\lambda_1} & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\lambda_2} & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{\lambda_{M_{\text{T}}M_{\text{R}}}} & 0 & \dots & 0 \end{bmatrix}_{M_{\text{T}}M_{\text{R}} \times TR} \quad (4.34)$$

such that the main diagonal elements are the reciprocals of $\lambda_1, \dots, \lambda_{M_{\text{T}}M_{\text{R}}}$, *i.e.*, the $M_{\text{T}}M_{\text{R}}$ highest singular values of $\mathbf{\Psi}$ ordered non-ascendingly, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{M_{\text{T}}M_{\text{R}}} > 0$.

The solution (4.33) however yields only a tight sensing matrix not normalized relative to the initial objective of a PTF $\mathbf{\Omega}$, [138]. As a consequence, an additional normalization step is necessary to satisfy Parseval's Theorem. Therefore, to satisfy this requirement the obtained solution matrix $\mathbf{\Phi}^*$ from (4.34) is renormalized given its Frobenius norm to yield the final measurement matrix [138]

$$\mathbf{\Phi}_{\text{PTF}} = \frac{\sqrt{TR}\mathbf{\Phi}^*}{\|\mathbf{\Phi}^*\|_{\text{F}}}, \quad (4.35)$$

where for future differentiation of beamforming strategies the PTF notation was abused to mark the fact that this measurement matrix generates a PTF compliant sensing matrix $\mathbf{\Omega}$.

Note that the PTF-beamforming implies by design that the measurement matrix $\mathbf{\Phi}_{\text{PTF}}$ is optimized with the minimized sensing energy of $\|\mathbf{\Phi}_{\text{PTF}}\|_{\text{F}}^2 = TR$ given Equation 4.35.

Provided the above derivation, it follows that the optimized final sensing matrix fulfills therefore the second design criterion listed in the end of the previous subsection. As seen already throughout Chapter 3, albeit the fact that tightness implies some reduction of coherence, it does not lead to incoherent frames close to the Welch bound (3.68) of mutual coherence.

Remark 1. *If the fixed dictionary $\mathbf{\Psi}$ is an UNTF, then the above PTF-beamforming is itself a tight frame such that the sensing matrix is PTF, [138]. This remark can easily be proven given*

the constraint (4.32b) that (4.35) solves, *i.e.*, $\Phi_{\text{PTF}} \Psi \Psi^H \Phi_{\text{PTF}}^H = \mathbf{I}_{M_T M_R}$. Upon the Frame Theory Chapter 3, an UNTF Ψ associated frame operator is $\mathbf{S}_\Psi = \frac{G_T G_R}{T R} \mathbf{I}_{T R}$. Thus, the frame operator of Φ_{PTF} must satisfy $\mathbf{S}_{\Phi_{\text{PTF}}} = \frac{T R}{G_T G_R} \mathbf{I}_{M_T M_R}$, which is equivalent to Φ_{PTF} itself being a tight frame according to Table 3.1. However, since there is no equivalence between second and third columns of Table 3.1, there is no implicit conclusion to be drawn on the norm of the individual column vectors of Φ_{PTF} .

4.3.3 Optimally Incoherent Unit-norm Tight Frames

A new beamforming scheme is next proposed as a lower bound of achievable performance transcending the design PTF-beamforming design. To achieve this, both criteria identified and listed at the end of Subsection 4.3.1 regarding the design of optimized beamforming for a sparsity enhanced mmWave channel estimation are used, *i.e.*, incoherence of Φ and tightness of Ω .

Let us first remark that the available sparse dictionary Ψ formed by the quantized AoDs /AoAs components is geometrically fixed given the antenna arrays geometry at the Tx/Rx side respectively. Nevertheless, something fundamental can be mentioned regarding its structure and geometry under the beamspace domain. Consider for the sake of exposition uniform linear array (ULA) antennas. The logic and arguments that follow are also applicable to planar arrays as Kronecker products of their corresponding ULA partitions [107, 141, 142]. The steering vectors for ULA are defined as [142]

$$\mathbf{a}_r(\phi_\ell^r) = \frac{1}{\sqrt{R}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_\ell^r)}, \dots, e^{-j2\pi(R-1) \frac{d}{\lambda} \sin(\phi_\ell^r)} \right]^T \quad (4.36)$$

$$\mathbf{a}_t(\phi_\ell^t) = \frac{1}{\sqrt{T}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_\ell^t)}, \dots, e^{-j2\pi(T-1) \frac{d}{\lambda} \sin(\phi_\ell^t)} \right]^T, \quad (4.37)$$

for all the propagation paths ℓ of the channel \mathbf{H} , such that the subsequent incidence/departure angles are given by ϕ_ℓ^r and ϕ_ℓ^t respectively, and d marks the inter-element antenna spacing.

Consider furthermore the *normalized spatial frequency*, simply defined as

$$\vartheta \triangleq \frac{d}{\lambda} \sin(\alpha) \quad (4.38)$$

for any of the vectors above, such that $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ to avoid the North/South ambiguity specific to ULAs.

The expressions in Equations (4.36) and (4.37) become therefore periodic in ϑ with period $\frac{2d}{\lambda}$ and are explicitly given as

$$\mathbf{a}_r(\vartheta_\ell^r) = \frac{1}{\sqrt{R}} \left[1, e^{-j2\pi \vartheta_\ell^r}, \dots, e^{-j2\pi(R-1) \vartheta_\ell^r} \right]^T \quad (4.39)$$

$$\mathbf{a}_t(\vartheta_\ell^t) = \frac{1}{\sqrt{T}} \left[1, e^{-j2\pi \vartheta_\ell^t}, \dots, e^{-j2\pi(T-1) \vartheta_\ell^t} \right]^T, \quad (4.40)$$

for $\vartheta_\ell^r, \vartheta_\ell^t \in \left[-\frac{d}{\lambda}, \frac{d}{\lambda}\right]$, $\forall \ell$.

Linking the steering vectors obtained above as Equations (4.39) and (4.40) with the discrete

grid points of the channel estimation problem, both at Rx, (4.13), as well as at Tx, (4.14) and considering the inter-element spacing³ $d = \frac{\lambda}{2}$, it is easy to remark that in fact these correspond to the appropriate sampling

$$\theta_{g_r} = 2\pi g_r \Delta\vartheta_r \quad (4.41)$$

$$\theta_{g_t} = 2\pi g_t \Delta\vartheta_t \quad (4.42)$$

where implicitly the uniform sampling intervals $\Delta\vartheta_r \triangleq \frac{1}{G_R}$, and respectively, $\Delta\vartheta_t \triangleq \frac{1}{G_T}$ have been defined $\forall g_r = \{0, 1, 2, \dots, G_R - 1\}$, $g_t = \{0, 1, 2, \dots, G_T - 1\}$.

The spatial sensing AoA/AoD matrices quantized as their beamspace equivalents can be explicitly written by means of Equations (4.11), (4.12), (4.41) and (4.42) as

$$\mathbb{A}_R = \frac{1}{\sqrt{R}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & w_r & \dots & w_r^{(G_R-1)} \\ 1 & w_r^2 & \dots & w_r^{2(G_R-1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & w_r^{(R-1)} & \dots & w_r^{(R-1)(G_R-1)} \end{bmatrix} \in \mathbb{C}^{R \times G_R}, \quad (4.43)$$

and respectively,

$$\mathbb{A}_T = \frac{1}{\sqrt{T}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & w_t & \dots & w_t^{(G_T-1)} \\ 1 & w_t^2 & \dots & w_t^{2(G_T-1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & w_t^{(T-1)} & \dots & w_t^{(T-1)(G_T-1)} \end{bmatrix} \in \mathbb{C}^{T \times G_T}, \quad (4.44)$$

where the twiddle factors have been defined as $w_r \triangleq e^{-j\frac{2\pi}{G_R}}$, $w_t \triangleq e^{-j\frac{2\pi}{G_T}}$, and respectively, the normalization factors $\frac{1}{\sqrt{R}}$, $\frac{1}{\sqrt{T}}$ ensure unit-normality of the column vectors.

It is worth noting at this point that by Definition 3.2.1, \mathbb{A}_R and \mathbb{A}_T are UNFs since $R \leq G_R$ and $T \leq G_T$. Furthermore, the above matrices are in fact truncated and renormalized DFT matrices and, as a result, harmonic frames belonging to an infinite family of harmonic Fourier matrices with special tightness and coherence properties [143, 144]. As a consequence, $\mathbb{A}_R, \mathbb{A}_T$ are *geometric harmonic UNTFs* built on top of simplex manifolds [19], or equivalently, on harmonic Fourier groups [143, 144].

One can generally construct in fact for any dimensional pair (M, K) an UNTF $\mathbf{F} \in \mathbb{C}^{M \times K}$ starting from the K -DFT matrix or its conjugate by truncation of last $(K - M)$ rows and renormalization to unit-norm of the frame columns as above [19, 143].

Thence, both discretized phased ULAs \mathbb{A}_R , and respectively, \mathbb{A}_T are UNTFs. This remark is in fact a general result and valid as well for the planar antenna arrays which are similarly obtained by associated Kronecker products of azimuth and elevation angles as captured by the

³This is the critical inter-element spacing often preferred in practical deployments of linear-based phased arrays since it allows for maximum number of excitable, uncorrelated spatial analog modes on the array apertures [142].

embedded ULA components and their subsequent steering vectors [129, 141, 142].

However, the fixed beamspace dictionary $\Psi \in \mathbb{C}^{TR \times G_T G_R}$ is obtained upon (4.17) as the Kronecker product $\mathbb{A}_T^* \otimes \mathbb{A}_R$. Upon this decomposition, the following new general result stating the *invariability of tightness and equi-normality over Kronecker products* is introduced.

Lemma 4.3.3. *Given UNTFs $\mathbf{A} \in \mathbb{C}^{M \times N}$, $\mathbf{B} \in \mathbb{C}^{P \times Q}$ with subsequent frame operators $\mathbf{S}_\mathbf{A}$ and $\mathbf{S}_\mathbf{B}$, the Kronecker-distributed frame $\mathbf{C} \triangleq \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{MP \times NQ}$ has the following properties:*

i) $\mathbf{S}_\mathbf{C} = \mathbf{S}_\mathbf{A} \otimes \mathbf{S}_\mathbf{B}$;

ii) \mathbf{C} is UNTF.

Proof: The first item is proven algebraically by the identities below:

$$\mathbf{S}_\mathbf{C} \triangleq \mathbf{C}\mathbf{C}^H \tag{4.45a}$$

$$= (\mathbf{A} \otimes \mathbf{B})(\mathbf{A} \otimes \mathbf{B})^H \tag{4.45b}$$

$$= (\mathbf{A} \otimes \mathbf{B})(\mathbf{A}^H \otimes \mathbf{B}^H) \tag{4.45c}$$

$$= (\mathbf{A}\mathbf{A}^H) \otimes (\mathbf{B}\mathbf{B}^H) \tag{4.45d}$$

$$\triangleq \mathbf{S}_\mathbf{A} \otimes \mathbf{S}_\mathbf{B}. \tag{4.45e}$$

The second attribute results easily by the algebraic definition of the Kronecker product and of fact that both \mathbf{A}, \mathbf{B} are UNTFs. The former ensures the unit-normality of the columns of \mathbf{C} given the unit-normality of the columns of \mathbf{A} and \mathbf{B} respectively, whereas the latter ensures that the frame operator of \mathbf{C} becomes

$$\mathbf{S}_\mathbf{C} = \mathbf{S}_\mathbf{A} \otimes \mathbf{S}_\mathbf{B} = \frac{N}{M}\mathbf{I}_M \otimes \frac{Q}{P}\mathbf{I}_P = \frac{NQ}{MP}\mathbf{I}_{MP}, \tag{4.46}$$

implying that $\mathbf{C} \in \mathbb{C}^{MP \times NQ}$ is UNTF following the frame-theoretic preliminaries of Section 3.4.

□

Obviously the above lemma is applicable generally to equi-normalized tight frames as well, as the unit-normality has been chosen as normalization just as mere convenience and without loss of generality.

Applying the results of Lemma 4.3.3 to Ψ and using the arguments previously introduced that \mathbb{A}_T and \mathbb{A}_R are UNTFs, it becomes clear that actually Ψ is itself an UNTF. This holds true since \mathbb{A}_T^* is still an UNTF given the invariance of these properties to element-wise conjugation. Additionally, Ψ has a highly harmonic structure as desired from the perspective of recovery guarantees as per Subsection 4.3.1, obtained by virtues of the DFT structure embedded in the discrete matrices \mathbb{A}_T and \mathbb{A}_R respectively.

Reusing Remark 1 of Subsection 4.3.2 and the result from above, it is now clear that if Φ is designed as an UNTF, then Ω is going to be itself an UNTF, and hence, $\Omega\Omega^H = \frac{G_T G_R}{M_T M_R}\mathbf{I}_{M_T M_R}$ as desired by our second design criterion, thus satisfying the RIP. Furthermore, given the UNTF harmonic fixed dictionary Ψ it also becomes apparent in the light of Subsection 4.3.1 that to

enhance the sparse recovery, the first design criterion of incoherence should be enforced on the measurement matrix Φ . In summary thence, the joint training beamformers measurement matrix Φ should be in fact an *incoherent UNTF*.

As seen in Section 3.4, the problem of building highly incoherent and tight finite frames is an intense research topic within both frame-theoretic and signal processing communities. Having discussed the alternating projection core concepts [87] and the associated algorithms compacted in Algorithm 2, it was remarked that such methods fail to return frames with low-coherence approaching the WB as they are mostly concerned with tightness and address the coherence level only macroscopically via Gramian based projections. There are however alternative Group based methods of building incoherent and tight frames stemming from classic algebraic harmonic families such as the one of Fourier matrices [143,144]. Concrete examples in fact are represented to this end by \mathbb{A}_T and \mathbb{A}_R respectively. Albeit the latter approach being even capable of achieving optimal ETF designs, this is constrained to specific strong coprime dimensionality pairs [143], and as a result, is impractical to generic settings.

In the sequel, an alternative solution to the prior art capable of generating both incoherent and UNTFs is however proposed. The incoherence side of the method is rooted by a newly proposed convex optimization extensible framework to iteratively and successively decorrelate frame vectors capable of achieving low-coherence for essentially any dimensions (M, K) [82, 88, 89]. Whereas the tightness is assured finally by application of Theorem 3.5.1 discussed in detail in Chapter 3.

Concretely, low-coherence frames are constructed iteratively starting from any randomly fixed frame $\tilde{\mathbf{F}} \in \mathbb{C}^{M \times K}$ whose frame vectors have been subsequently normalized to unit-norm. The incoherence is then enforced by successive geometric based decorrelation of frame vectors. To this end, a vector $\mathbf{f}_k, \forall k$ is successively picked as the target frame vector undergoing optimization. Since in geometrical sense, the mutual coherence of an UNFs such as $\tilde{\mathbf{F}}$ is in fact

$$\mu(\tilde{\mathbf{F}}) = \max_{\tilde{\mathbf{F}}_{\text{UNF}}} \left| \cos \left(\angle \left(\tilde{\mathbf{f}}_k, \tilde{\mathbf{f}}_\ell \right) \right) \right|, \quad k \neq \ell, \quad (4.47)$$

it follows that one can directly minimize this by optimizing the associated maximum angular correlation of each targeted frame vector \mathbf{f}_k with respect to the other distinct frame vector elements $\tilde{\mathbf{f}}_\ell, \ell \neq k$, *i.e.*,

$$\min_{\mathbf{f}_k} \max_{\tilde{\mathbf{f}}_\ell, \ell \neq k} \left| \cos \left(\angle \left(\mathbf{f}_k, \tilde{\mathbf{f}}_\ell \right) \right) \right|, \quad (4.48)$$

across all k frame vectors [82, 89].

Based on (4.48), one may remark that the search space available for these optimization problem is in fact unconstrained and subject to the entire multi-dimensional space \mathbb{C}^M . To be able to solve the initial problem (4.47) by iteratively applying (4.48) for successive target frame vectors, the search space of each target vector \mathbf{f}_k needs to exclude all the other frame vectors $\tilde{\mathbf{f}}_\ell$ in order to avoid accidental collinearities that would become catastrophic and increase in fact mutual coherence. Rusu et al. realized this constraint by an M -dimensional ball packing and search space constraint [89]. To construct such search regions it was proposed firstly in [88] that

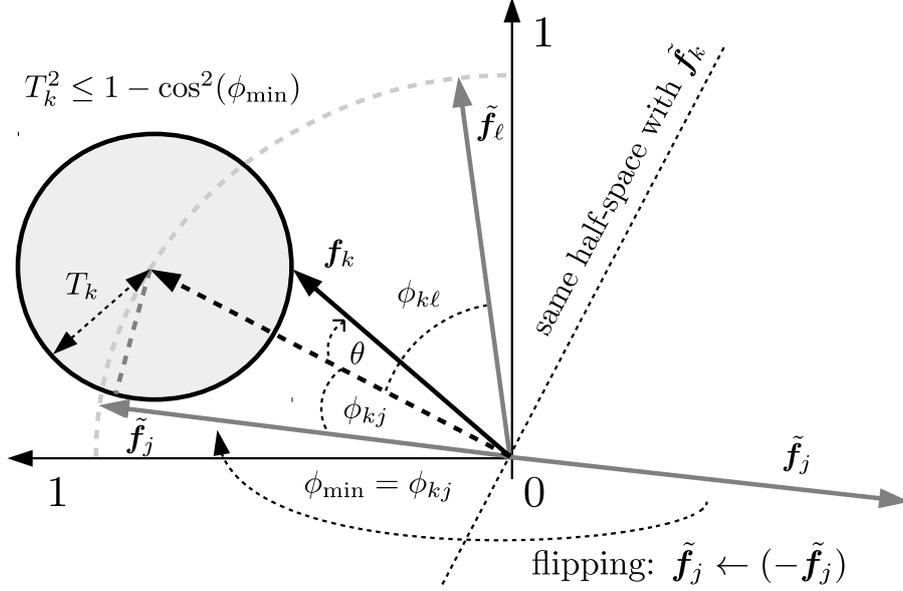


Figure 4.2: Illustration of the SIDCO [82, 89] search space constraint of problem (4.50). To this end, both the frame vector flipping and the search ball $\|\mathbf{f}_k - \tilde{\mathbf{f}}_k\|_2^2 \leq T_k^2$ geometry are displayed.

the mutual coherence flip invariance should be used. The flip invariance property of coherence is rooted by the trivial identity

$$\left| \cos \left(\angle \left(\mathbf{f}_k, \tilde{\mathbf{f}}_\ell \right) \right) \right| = \left| \cos \left(\angle \left(\mathbf{f}_k, -\tilde{\mathbf{f}}_\ell \right) \right) \right| \quad (4.49)$$

which furthermore generally preserves the coherence, and respectively, mutual coherence properties of any frame irrespective of the flipping of any frame vectors.

The flip operation has been imposed in [82, 88, 89] in order to force all frame vectors to be in the same half-space as the target frame vector undergoing optimization, *i.e.*, \mathbf{f}_k . Then, a M -ball search space with adaptive search radius T_k has been defined to compactly define the core optimization problem solved by the framework of *sequential iterative decorrelation via convex optimization (SIDCO)* as [82],

$$\underset{\mathbf{f}_k \in \mathbb{C}^M}{\text{minimize}} \quad \|\tilde{\mathbf{F}}_k^H \mathbf{f}_k\|_\infty, \quad (4.50a)$$

$$\text{subject to} \quad \|\mathbf{f}_k - \tilde{\mathbf{f}}_k\|_2^2 \leq T_k^2, \quad (4.50b)$$

with the \mathbf{f}_k -pruned frame implicitly denoted as

$$\tilde{\mathbf{F}}_k \triangleq \left[\tilde{\mathbf{f}}_1, \dots, \tilde{\mathbf{f}}_\ell, \dots, \tilde{\mathbf{f}}_K \right], \quad \ell \neq k. \quad (4.51)$$

The search space geometric constraint (4.50b) for \mathbf{f}_k is illustrated in Figure 4.2. The flip property is as well jointly illustrated for the frame vector $\tilde{\mathbf{f}}_j$. It is thus clear that in order to avoid collinearity, any $\tilde{\mathbf{f}}_\ell$ needs to fall just outside the search M -ball of \mathbf{f}_k . Therefore, the

search radius T_k is in fact geometrically bounded, as shown in Figure 4.2, by

$$T_k^2 \leq 1 - \cos^2(\phi_{\min}), \quad (4.52)$$

which is in fact equivalent to the algebraic equality

$$T_k^2 \leq 1 - \max_{\ell \neq k} |g_{k\ell}|^2, \quad (4.53)$$

since $\tilde{\mathbf{F}}$ is UNF, and so, the radius of the M -ball search space, [89],

$$T_k^2 = 1 - \max_{\ell \neq k} |g_{k\ell}|^2. \quad (4.54)$$

As outlined in Figure 4.2, it can be seen that the optimized frame vector \mathbf{f}_k obtained as a result of solving (4.50) may not be unitary as it is constrained to the surface of the M -ball search region [89]. Consequently, a posterior normalization step is necessary to enforce the norm-1 constraint as

$$\tilde{\mathbf{f}}_k = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|_2}, \quad (4.55)$$

and recomplete the UNF after an iteration, [89], as $\tilde{\mathbf{F}}$.

This SIDCO framework [82,89] able to generate incoherent frames using iteratively a simple yet effective convex optimization problem (4.50), has been recently enhanced in [82] to different signal spaces amongst which also the complex domain. Therefore, the authors of [82] coined the scheme applied to the unconstrained \mathbb{C} space, as above, *complex SIDCO (CSIDCO)*. In the light of its demonstrated scalability in comparison with the other discussed alternatives, this method is next utilized to generate our desired frame as the measurement matrix Φ .

The CSIDCO solution obtained by solving the problem (4.50) over all frame vectors \mathbf{f}_k for I iterations leads thus to the construction of a locally optimized incoherent UNF \mathbf{F}_* . CSIDCO is of course also capable of producing Grassmannian frames as optimally incoherent outputs, however there are no guarantees to this end [82]. Unfortunately, the CSIDCO by-product is not a tight frame, as outlined in [82], and as a result, to fulfill our Φ design objectives, the associated problem of tightening

$$\underset{\Phi \in \mathbb{C}^{M_T M_R \times T R}}{\text{minimize}} \quad \|\Phi - \mathbf{F}_*\|_F \quad (4.56a)$$

$$\text{subject to} \quad \Phi \Phi^H = \frac{K}{M} \mathbf{I}_{M_T M_R} \quad (4.56b)$$

is yet to solve, where the equality constraint aims to enforce the UNTF condition according to Table 3.1, given that \mathbf{F}_* is an UNF already.

The analytic solution to problem (4.56) is given by means of Theorem 3.5.1 and is obtained under appropriate scaling as

$$\Phi_{\text{OPT}} = \sqrt{\frac{K}{M}} (\mathbf{F}_* \mathbf{F}_*^H)^{-\frac{1}{2}} \mathbf{F}_*. \quad (4.57)$$

Therefore the CSIDCO incoherence and unit-normality is approximated via the polar de-

Algorithm 3 INCOHERENT UNTF FRAME DESIGN

Inputs: Signal space dimension M , number of frame vectors K and number of iterations I
Output: Incoherent UNTF $\mathbf{F} \in \mathbb{C}^{M \times K}$, incoherent UNF $\mathbf{F}_* \in \mathbb{C}^{M \times K}$

1: Generate initial random matrix: $\tilde{\mathbf{F}}_{(0)} \in \mathbb{C}^{M \times K}$

2: Tighten $\tilde{\mathbf{F}}_{(0)}$ using Theorem 3.5.1:

$$\tilde{\mathbf{F}}_{(0)} = \sqrt{\frac{K}{M}} \left(\tilde{\mathbf{F}}_{(0)} \tilde{\mathbf{F}}_{(0)}^H \right)^{-\frac{1}{2}} \tilde{\mathbf{F}}_{(0)} \quad \triangleright \text{(Optional) speeds up CSIDCO as implicitly}$$

\triangleright lowers coherence of $\tilde{\mathbf{F}}_{(0)}$, see Section 3.5

3: Normalize all column vectors of $\tilde{\mathbf{F}}_{(0)}$ to norm 1:

$$\tilde{\mathbf{f}}_{(0),k} = \frac{\tilde{\mathbf{f}}_{(0),k}}{\|\tilde{\mathbf{f}}_{(0),k}\|_2}, \quad \forall k$$

4: Iteratively optimize and update $\tilde{\mathbf{F}}_{(0)}$ into $\tilde{\mathbf{F}}_{(i)}$ by solving (4.50), with T_k updated via Equation (4.54), $\forall k$ at each iteration $i \leq I$ \triangleright CSIDCO, [82]

5: Cache optimized incoherent UNF: $\mathbf{F}_* = \tilde{\mathbf{F}}_{(I)}$

6: Tighten incoherent UNF \mathbf{F}_* into an approximate incoherent UNTF \mathbf{F} :

$$\mathbf{F} = \sqrt{\frac{K}{M}} \left(\mathbf{F}_* \mathbf{F}_*^H \right)^{-\frac{1}{2}} \mathbf{F}_*$$

composition of Theorem 3.5.1 to an incoherent UNTF representation denoted hereafter Φ_{OPT} as *optimally incoherent-tight joint beamforming* of both Tx and Rx training beamformers. The entire algorithm combining the CSIDCO procedure [82] and the tight incoherent UNTF approximation detailed so far is listed as Algorithm 3. The latter returns finally the optimized beamforming tight frame Φ_{OPT} whose frame vectors are approximately unit-normed and mutual coherence is slightly increased in comparison to $\mu(\mathbf{F}_*)$ upon the tightening operation (4.57). Nevertheless, the result of Φ_{OPT} of Algorithm 3 fulfills numerically and practically the requirements of generating an incoherent UNTF joint beamforming matrix Φ as desired.

4.3.4 Kronecker-decomposable Practical Realizations

It is important to note that both optimized training methods, *i.e.*, the PTF-beamforming from Subsection 4.3.1 and the optimally incoherent-tight beamforming from Subsection 4.3.2, are proposed on top of the joint measurement matrix Φ , rather than separately as expected on the Tx, and respectively, on the Rx side. As a consequence, they are both practically unfeasible to real mmWave systems given the system model (4.17). Despite this feasibility issue however, they serve in fact as idealized benchmarks and therefore as upper bounds on the achievable sparse recovery performance of practical mmWave channel estimation.

To elaborate, the frame-theoretic methods discussed so far correspond actually to the impractical scenario where the Tx and the Rx would need to perform a perfectly synchronous beamforming management, where the Rx needs to have synchronized full-knowledge of the Tx training beamformers at any step during the channel estimation process. This core requirement is of course infeasible in practice, especially in the context of first link establishment and channel estimation during the considered IA scenario. Thus, given the assumptions of a practical

deployment where during the early stages of channel acquisition and piloting the Tx and Rx have no prior knowledge of the channel, and thus, independently need to manage their training beamforming vectors, the measurement matrix Φ needs therefore to be Kronecker-decomposable as $\mathbf{U}^T \otimes \mathbf{V}^H$ upon (4.17). Having introduced earlier Lemma 4.3.3 regarding the preservation of equi-normality and tightness over the Kronecker product of frames, the Kronecker decomposability of incoherent frames and *the tractability of incoherence over the Kronecker multiplication* is next addressed by the results of Lemma 4.3.4.

Lemma 4.3.4. *Given UNFs $\mathbf{A} \in \mathbb{C}^{M \times N}$, $\mathbf{B} \in \mathbb{C}^{P \times Q}$ with subsequent Gram matrices $\mathbf{G}_\mathbf{A}$ and $\mathbf{G}_\mathbf{B}$, the Kronecker-distributed frame $\mathbf{C} \triangleq \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{MP \times NQ}$ has the following properties:*

- i) $\mathbf{G}_\mathbf{C} = \mathbf{G}_\mathbf{A} \otimes \mathbf{G}_\mathbf{B}$;
- ii) \mathbf{C} is an UNF;
- iii) $\mu(\mathbf{C}) = \max(\mu(\mathbf{A}), \mu(\mathbf{B}))$.

Proof: Property i) is derived immediately by straightforward algebraic computation as

$$\mathbf{G}_\mathbf{C} \triangleq \mathbf{C}^H \mathbf{C} \tag{4.58a}$$

$$= (\mathbf{A} \otimes \mathbf{B})^H (\mathbf{A} \otimes \mathbf{B}) \tag{4.58b}$$

$$= (\mathbf{A}^H \otimes \mathbf{B}^H) (\mathbf{A} \otimes \mathbf{B}) \tag{4.58c}$$

$$= (\mathbf{A}^H \mathbf{A}) \otimes (\mathbf{B}^H \mathbf{B}) \tag{4.58d}$$

$$\triangleq \mathbf{G}_\mathbf{A} \otimes \mathbf{G}_\mathbf{B}. \tag{4.58e}$$

The second property follows from the first one by the definition of the Kronecker product, and respectively, by the fact that \mathbf{A} and \mathbf{B} are themselves unit-norm frames, and as a result, $\text{diag}(\mathbf{G}_\mathbf{A}) = \text{diag}(\mathbf{G}_\mathbf{B}) = \mathbf{1}_{NQ}$. As a consequence, the diagonal of $\mathbf{G}_\mathbf{C}$ is therefore also filled by 1's, which is equivalent to the frame vectors \mathbf{c}_k having the norm $\|\mathbf{c}_k\|_2 = 1, \forall k = \{1, \dots, NQ\}$.

Lastly, property iii) is similarly immediate based on the fact that the absolute value of each non-diagonal entry of $\mathbf{G}_\mathbf{C}$ is in fact upper-bounded by the product of the absolute value of each entry of $\mathbf{G}_\mathbf{A}$ multiplied by the absolute value of each entry of $\mathbf{G}_\mathbf{B}$. Moreover since the magnitudes of the entries of the Gramians are confined to the interval $[0, 1]$ with their diagonal entries always unitary powered, as \mathbf{A}, \mathbf{B} are UNF, it so follows that $\mu(\mathbf{C}) = \max(\mu(\mathbf{A}), \mu(\mathbf{B}))$.

□

An immediate consequence of Lemma 4.3.4 is the fact that coherence of a Kronecker product is therefore increased up to the highest individual coherence of its terms. This comes as no surprise given the Kronecker product embedded structure and subsequent loss of degrees of freedom.

Lemmas 4.3.3 and 4.3.4 therefore link on one hand the UNTF, and on the other hand, the incoherence invariance over *any Kronecker product* of frames exhibiting such properties. Based on these results, a new Kronecker-distributed (KD) frame design is proposed as a practical alternative for Tx/Rx optimized training beamforming vectors. Thus, the Tx training beamformers

\mathbf{U}^T , and respectively, the Rx training combiners \mathbf{V}^H are generated independently at the Tx, respectively at the Rx, as incoherent UNTFs following Algorithm 3. As a result, this would coincide to the realization of Φ_{KD} as a KD joint measurement matrix formed by the Kronecker product of the aforementioned beamforming Tx/Rx vectors. We refer to the later strategy of distributed optimized beamforming for the initial channel acquisition as the *KD-beamforming*. Remark to this end that KD-beamforming is in fact a practical realization of the optimally incoherent-tight joint beamforming scheme suggested in Subsection 4.3.3, employing the same core concepts but independently at both the Tx and Rx ends.

Recall that upon Remark 1 of Subsection 4.3.1 Φ_{PTF} is a tight frame and by means of Lemma 4.3.3 is similarly factorizable as a Kronecker product of two appropriately dimensioned tight frames Tx/Rx beamforming matrices. However, to maintain the PTF design, the Rx would require prior knowledge of the Tx beamforming to equalize the former in real-time and reach the equivalent of the joint PTF-beamforming. Once more, this is infeasible, and due to the practicality motivation behind this subsection this approach has been abandoned.

In summary, the distributed design scheme of the equivalent KD measurement matrix Φ_{KD} is thus listed as a high-level sequence of steps below.

1. Generate an incoherent UNTF, \mathbf{U}_{KD}^T , at the Tx side using Algorithm 3 without any prior knowledge of the joint fixed dictionary Ψ .
2. Generate independently of the first frame at the Tx an incoherent UNTF, \mathbf{V}_{KD}^H , at the Rx side using similarly Algorithm 3.
3. Based on the Lemmas 4.3.3 and 4.3.4, the equivalent $\Phi_{\text{KD}} \triangleq \mathbf{U}_{\text{KD}}^T \otimes \mathbf{V}_{\text{KD}}^H$ is implicitly obtained and the estimation model described by (4.17) can be applied.

4.4 Beamforming Performance Evaluation

Let us now have a critical look into the three frame-theoretic beamforming schemes discussed in the previous section. Furthermore, alongside the optimized PTF-beamforming [138], also the practically utilized random beamforming [114] shall be considered as a state-of-the-art realizable reference.

4.4.1 Beamforming Design Differences

Next the main design differences among the three frame-theoretic driven beamforming schemes, *i.e.*, Φ_{PTF} discussed in Subsection 4.3.1, Φ_{OPT} proposed in Subsection 4.3.2, and Φ_{KD} proposed in Subsection 4.3.3, are discussed. Comparisons are made to individual realizations of the randomized Gaussian beamforming, denoted hereafter for convenience Φ_{RND} . Following the theoretical system modeling from Section 4.2, $\Phi_{\text{RND}} = \mathbf{U}_{\text{RND}}^T \otimes \mathbf{V}_{\text{RND}}^H$, where actually the randomized beamforming vectors are in fact the Tx side beamformers, and respectively, the Rx side combiners. For a fair comparison across all numerical results in the sequel, the random beamforming was scaled such that $\mathbb{E}[\|\Phi_{\text{RND}}\|_F^2] = TR$ to match the sensing costs of the other

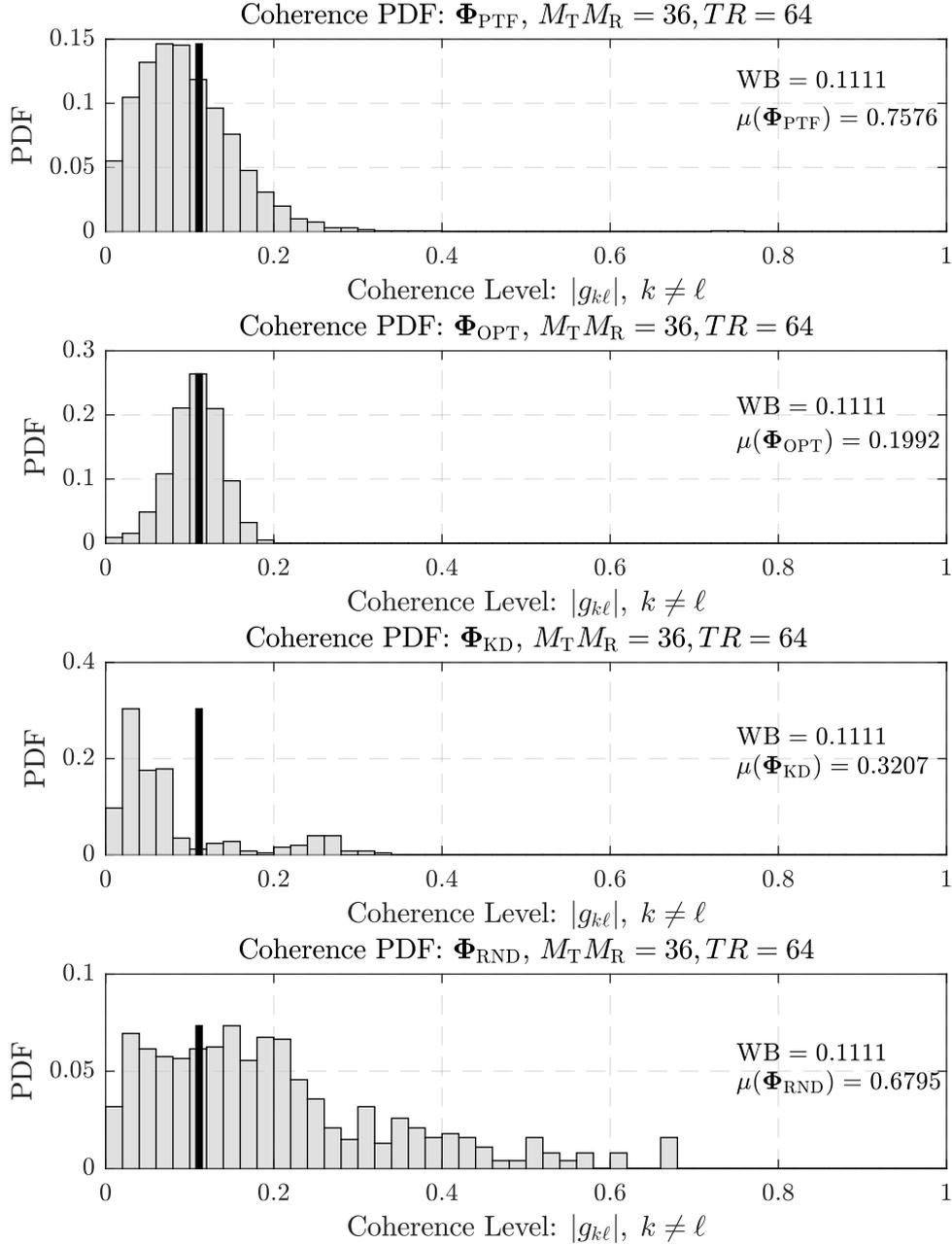


Figure 4.3: Normalized coherence PDF of the measurement matrix Φ for different beamforming schemes given 6/8 training beamformers at each Tx/Rx. The corresponding WB of mutual coherence for $\Phi \in \mathbb{C}^{36 \times 64} = 0.1111$ is plotted as a horizontal black bar for comparison. Furthermore, mutual coherence values for each Φ are also listed inline for numerical comparison.

Φ matrices. The fact that $\|\Phi_{\text{PTF}}\|_{\text{F}}^2 = \|\Phi_{\text{OPT}}\|_{\text{F}}^2 = \|\Phi_{\text{KD}}\|_{\text{F}}^2 = TR$ is trivial and follows on one hand for Φ_{PTF} given (4.35), whereas on the other hand, for the UNTF $\Phi_{\text{OPT}}, \Phi_{\text{KD}}$ the equality is immediate given Table 3.1, since $\|\Phi\|_{\text{F}}^2 = \text{trace}(\Phi\Phi^{\text{H}}) = \text{trace}(\mathbf{S}_{\Phi}) = TR$ for UNTFs.

To visualize the main criteria identified in Subsection 4.3.1 as desirable for good recovery guarantees, the Figures 4.3 - 4.4 were rendered outlining the probability density functions (PDFs) of the normalized Gramians of the measurement matrices and their corresponding sensing matrices alike across different beamforming strategies. The enforced normalization

follows in fact the Definition 3.4.5 of the mutual coherence and provides a fair comparison across all comparison points. The displayed frames and their coherence PDFs correspond to MIMO mmWave systems with $T = R = 8$ Tx/Rx antennas, training Tx/Rx space dimensionality of $M_T = M_R = 6$, and respectively, a virtual beamspace resolution based on the uniform angular grid quantization of $G_T = G_R = 10$ on the Rx side.

As it can be observed in Figure 4.3, the measurement frames Φ_{PTF} , Φ_{OPT} and Φ_{KD} are all on average incoherent. However, given that Φ_{PTF} is actually optimized to minimize the sensing measurement energy under the constraint of obtaining a PTF sensing matrix, its low coherence is a consequence of the tightness therefore resulted and mentioned in Remark 1. Due to this, Φ_{PTF} has a long tail distribution of coherence leading to a high mutual coherence, *i.e.*, $\mu(\Phi_{\text{PTF}}) = 0.7576$. On the contrary, the Φ_{OPT} proposed method addresses this issue via the CSIDCO incoherence optimization, but at the same time also enforces tightness and implicitly minimizes to the optimum point the FP. This results in an almost bell-shaped distribution with the average coherence centered at the WB of mutual coherence but with a low variance such that the worst-possible angular correlation is 0.1992, quite close to the WB as outlined in Figure 4.3.

Since the former two beamforming strategies represent jointly optimized but impractical strategies acting as mere upper bounds of achievable performance alternative practical realizations are also illustrated via the KD-beamforming and randomized beamforming. Given that both are in fact Kronecker products of Tx/Rx beamforming vectors, it is visible that their correlations are widely spread in comparison to the previously discussed strategies. However, the KD-beamforming achieves a much better performance of coherence as seen from Lemmas 4.3.3 and 4.3.4 by utilizing Tx/Rx incoherent UNTF beamforming, and thus, fulfills both tightness and incoherence conditions in a distributed manner as already outlined through Subsection 4.3.4. In contradiction, the randomized beamforming simply spreads out the coherence of its joint measurement frame representation based on the Kronecker product whilst at the same time missing out in enforcing a tight frame representation given its generation principles.

However, as already pointed out in Subsection 4.3.1, there are no general results or guarantees of preserving coherence values under matrix multiplication, especially in the case of multiplication with another frame, albeit UNTF as Φ . To point out this fact, but also highlight the coherence transformation yielded under multiplication for Ω , the coherence PDF of the sensing matrices corresponding of the previous measurement matrices are plotted in Figure 4.4. The sensing matrices are all obtained under multiplication of their corresponding measurement frames with the $\Psi \in \mathbb{C}^{64 \times 100}$ harmonic frame dictionary resulted given the considered 10×10 quantization grid at the Rx and all its possible AoA/AoD quantized combinations.

Based on Figure 4.4 and on Figure 4.3 an interesting high-level effect that the UNTF Ψ dictionary has over all sensing matrices obtained regardless of the measurement beamforming strategy is the averaged decorrelation of measurement frames vector elements. This happens naturally via multiplication with the harmonic UNTF Ψ which as a result smears out the measurement matrices distributions originally depicted under Figure 4.3.

Under this effect, the better incoherence characteristics of Φ_{OPT} over Φ_{PTF} are reduced as

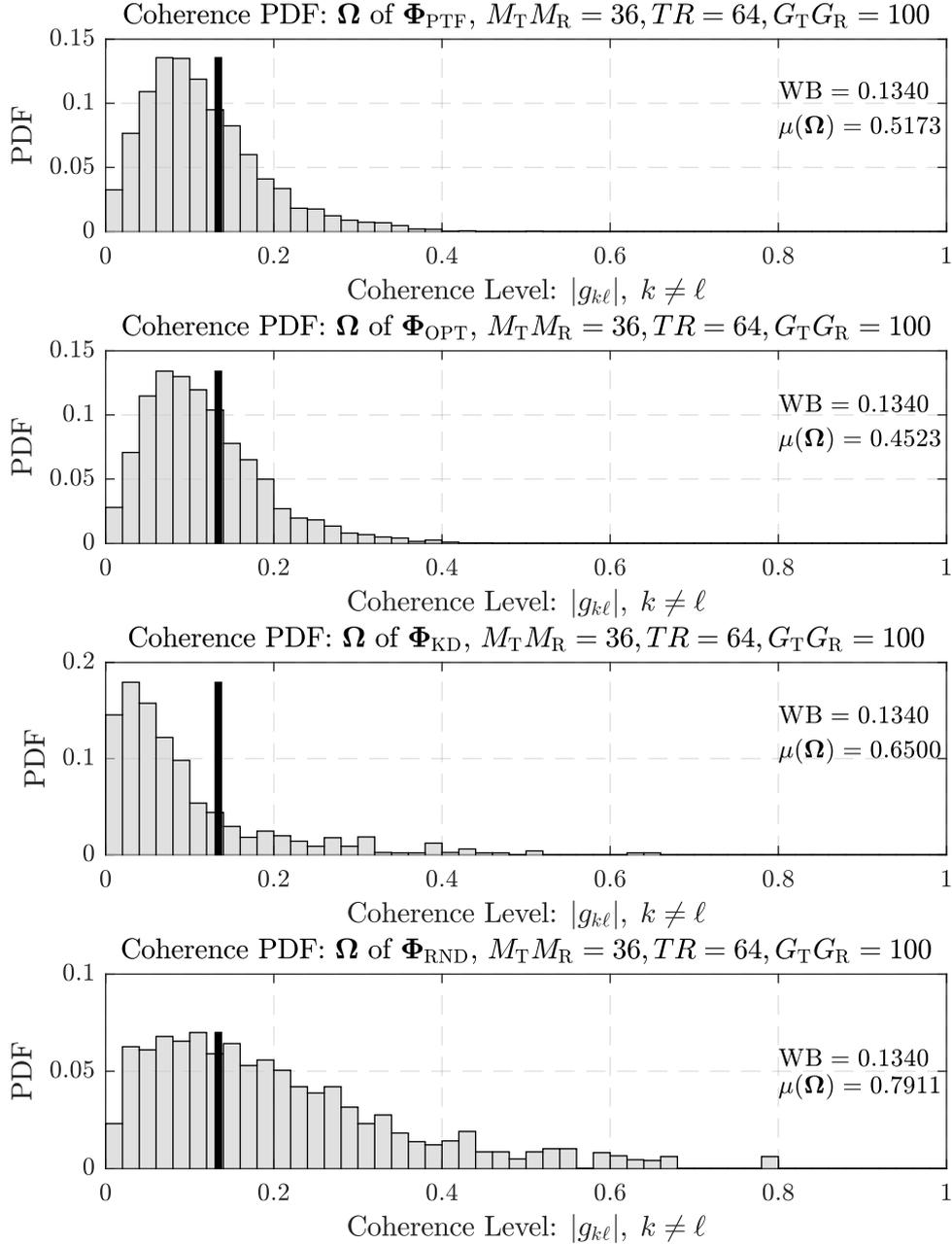


Figure 4.4: Normalized coherence PDF of the sensing matrix Ω corresponding to different beamforming schemes given 6/8 training beamformers at each Tx/Rx. The corresponding WB of mutual coherence for $\Omega \in \mathbb{C}^{36 \times 100} = 0.1340$ is plotted as a horizontal black bar for comparison. Furthermore, mutual coherence values for each Ω are also listed inline for numerical comparison.

both frames lead still to tight representations, but with an increased coherence. Nevertheless, even so, as pointed out by asymptotic analyses of CS, [134–136], high incoherence for the measurement matrix is still desirable as under harmonic dictionary multiplication it still leads to incoherent sensing. This is the case as the long tail of Φ_{PTF} coherence distribution is preserved and increased to the worst-case coherence level of 0.5173, whereas in comparison the Φ_{OPT} is limited to 0.4523. Simultaneously, both sensing matrices though have similar distributions of the coherence levels.

The same general remarks can be analogously made for the practical achievable sensing matrices corresponding to KD, and respectively, randomized beamforming. Once more, the superiority of the proposed KD-beamforming in terms of both tightness and coherence over the randomized beamforming is maintained as outlined by the mutual coherence values inlined in Figure 4.4, and respectively, by the tightness preservation under the Kronecker product following Lemma 4.3.3. Related to tightness, it is trivial to see following the same lemma that as long as Φ_{RND} is not tight, then its resulted sensing matrix is also not tight. Nevertheless, it is stressed once more that tightness is not completely equivalent to the RIP, and in fact, randomized measurement matrices achieve asymptotically the latter property with high probability, [135], as elaborated earlier.

These effects discussed above punctually to a particular realization of a mmWave system are general and implied by the accompanying beamforming design across various instantiations and parameterizations within the model constraints discussed so far. The general trend is that they are more accentuated, and thus influence more the performance, for higher redundant systems where the overloading of both the virtually quantized beamspace and the training beamforming signal space is predominant, in turn providing better robustness necessary to improve fast IA channel acquisition under less number of measurements than the state of the art.

4.4.2 Numerical Results and Comparisons

To provide a comparison regarding the recovery performance of the proposed frame-theoretic beamforming schemes, the simulation of the mmWave Tx-Rx model (4.1) corresponding to the CS linear system (4.17) is performed. As seen previously, the canonical ℓ_0 -sparse recovery problem (4.18) solving (4.17) is combinatorial in nature [67], and so, NP-hard. To alleviate this issue and provide computationally tractable solutions the literature of CS solvers provides a multitude of solutions [132], like variations of greedy matching pursuit approaches, *e.g.* OMP [123,124], basis pursuit ℓ_1 -convex optimization methods, *e.g.* BPDN [125], reweighted- ℓ_1 BPDN [145] and iteratively approximated algorithms such as fast iterative soft-thresholding (FISTA) [146], approximate message passing (AMP) [147] and other variations thereof.

We have considered in this work two classical solvers, *i.e.*, the OMP, and respectively, the BPDN with its ℓ_1 -reweighting flavor. These selections are motivated given their mathematical elegance, thorough literature coverage and available strong recovery guarantees already discussed in Subsection 4.3.1 fully applicable to them. Based on these choices, we show next the universality and scalability of the proposed frame-theoretic beamforming methods as associated measurement matrices for the mmWave channel estimation problem. To this extent, we leverage the guaranteed performance of the selected algorithms in order to outline the gains obtained simply by means of beamforming optimization in the context of our problem.

The OMP [124], was in fact used as the core algorithm behind the mmWave channel estimation in the works of [107,120], as it solves the problem (4.18) by recasting it to

$$\underset{h \in \mathbb{C}^{G_T G_R}}{\text{minimize}} \quad \|y - \mathbf{\Omega}h\|_2, \quad (4.59a)$$

$$\text{subject to} \quad \|h\|_0 = s. \quad (4.59b)$$

The OMP solution is then obtained by iteratively and greedily increasing the solution support vector to the expected sparsity level s , based on the most correlated sensing column vector to the residual. Simultaneously, during each iteration the residual is decreased by orthogonal matching of the support contributions to the residual via simple least squares [123, 124].

On the other hand, the BPDN algorithm [125] uses the ℓ_0 - ℓ_1 relaxation [67, 134, 135] typical to CS in order to alternatively solve the convex ℓ_1 -problem

$$\underset{h \in \mathbb{C}^{G_T G_R}}{\text{minimize}} \quad \|h\|_1, \quad (4.60a)$$

$$\text{subject to} \quad \|y - \underbrace{\mathbf{\Phi}\mathbf{\Psi}}_{\mathbf{\Omega}}h\|_2 \leq \delta, \quad (4.60b)$$

instead of the original (4.18).

Alternative forms to (4.60) exist under the unconstrained Lasso ℓ_1 -regularized form [124, 125], however performance guarantees are similar given both representations. Additionally however, as pointed out in [126], the sparsity-enhanced BPDN ℓ_1 -reweighted is a better suited candidate to noisy recovery problems, and hence, was also applied to the original problem hereby considered as well. The idea behind the ℓ_1 -reweighting came from [124] and sought to improve the ℓ_0 - ℓ_1 equivalence given noisy scenarios where the latter suffered. This has been done based on linearly weighting the BPDN original problem and solving the reweighted form

$$\underset{h \in \mathbb{C}^{G_T G_R}}{\text{minimize}} \quad \|\mathbf{W}h\|_1, \quad (4.61a)$$

$$\text{subject to} \quad \|y - \underbrace{\mathbf{\Phi}\mathbf{\Psi}}_{\mathbf{\Omega}}h\|_2 \leq \delta, \quad (4.61b)$$

iteratively up to a certain number of maximum iterations I_{\max} or convergence.

The diagonal weighting matrix \mathbf{W} in (4.61a) is initialized to identity. This results in fact in the simple BPDN solution after just one iteration of the ℓ_1 -reweighted BPDN algorithm [145]. Then the diagonal entries of \mathbf{W} are updated after every i -th iteration of (4.61) as

$$w_j = \frac{1}{|\hat{h}_j^{(i)}| + \epsilon} \quad \forall j, \quad (4.62)$$

where parameter $0 < \epsilon \ll 1$ provides numerical stability against 0 entries of $\hat{h}_i^{(i)}$ [124].

Remark 2. Notice that upon the sparse estimation of h as \hat{h} , the matrix channel representation \mathbf{H} needs to be reconstructed as well. This is achieved by inverting the vectorization and applying the model (4.17) under the existing quantization resolution, resulting into the final estimate

$$\hat{\mathbf{H}} = \mathbf{A}_R \underbrace{\left(\text{vec}^{-1}(\hat{h}) \right)}_{\hat{\mathbf{H}}_\gamma} \mathbf{A}_T^H. \quad (4.63)$$

The upcoming simulations have been generated based on the mmWave system described

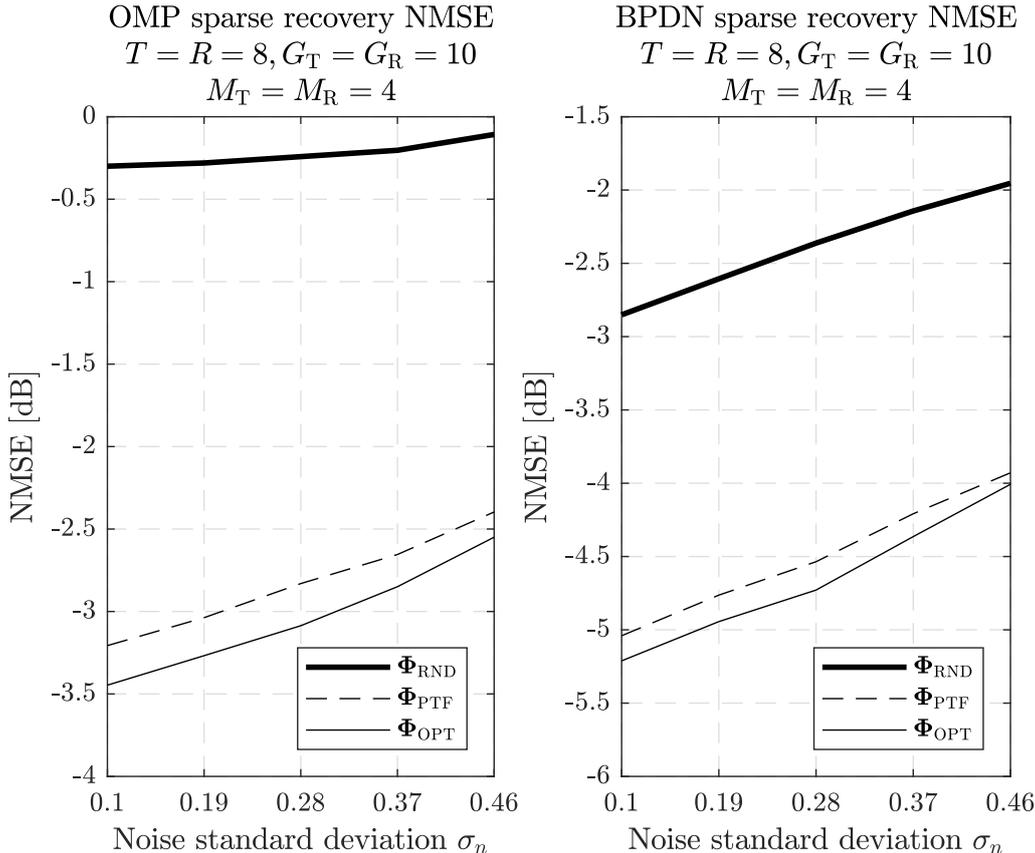


Figure 4.5: NMSE performance of mmWave channel sparse recovery using OMP [123] (left) / BPDN [125] (right) for $M_T = M_R = 4$ across three different training beamforming schemes: Φ_{RND} randomized Gaussian beamforming [114, 135], PTF-beamforming Φ_{PTF} [126, 138], and respectively, optimally incoherent-tight beamforming Φ_{OPT} proposed in Subsection 4.3.3.

insofar throughout this chapter and roughly summarized by Equations (4.1) and (4.17). Concretely, the channel was modeled under premises of (4.2) with $L = 3$ main propagation paths each with gains modeled as i.i.d complex normally random variables following the distribution $\mathcal{CN}(0, \sigma_\gamma^2)$ and $\sigma_\gamma^2 = 1$. Furthermore, the number of Rx/Tx antennas was set to $R = T = 8$ under the ULA configuration already previously covered such that their antenna array steering vectors on the Rx/Tx side was given by Equations (4.36) and (4.37) respectively. The inter-element antenna spacing of both antenna arrays was set to $d = \frac{\lambda}{2}$, and the AoA/AoD were stochastically generated uniformly over the interval $[0, 2\pi]$. The virtual beamspace was sampled under a grid of 10×10 , *i.e.*, the number of bins for AoA was set to $G_R = 10$, whereas the number of bins for AoD was respectively set to $G_T = 10$ on the Rx side. Lastly, the number of training beamforming elements M_T and M_R were varied synchronously in the set $\{4, 5, 6, 7\}$ to outline different beamforming diversity factors and achievable performance.

The AWGN noise vector realizations in (4.17) were generated as i.i.d randomized samples distributed according to the complex multivariate Gaussian distribution $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{M_T M_R})$ under various power levels σ_n^2 . In addition, given the imposed average unit-energy normalization of the channel model (4.2) radiated over the ULA steering vectors, it follows that the average SNR, or equivalently in this case channel-to-noise ratio (CNR), is actually returned by the reciprocal

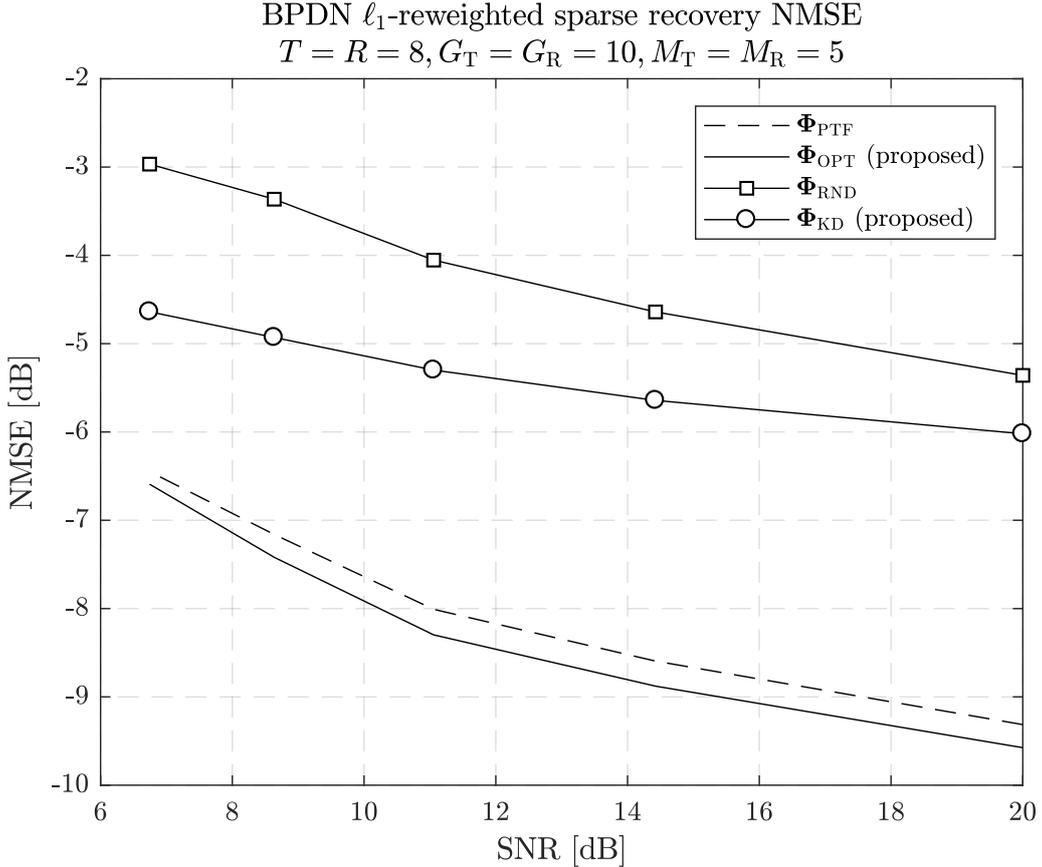


Figure 4.6: NMSE performance of mmWave channel sparse recovery using BPDN ℓ_1 -reweighted [134] for $M_T = M_R = 5$ across four different training beamforming schemes: PTF-beamforming Φ_{PTF} [126, 138], optimally incoherent-tight beamforming Φ_{OPT} proposed in Subsection 4.3.3, Φ_{RND} randomized Gaussian beamforming [114, 135], and respectively, KD-beamforming Φ_{KD} proposed in Subsection 4.3.4.

of the noise power level, *i.e.*, $\frac{1}{\sigma_n^2}$.

Lastly, the quantitative performance analysis was based on the normalized-MSE (NMSE) channel estimation

$$\text{NMSE} = \mathbb{E} \left[\frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_{\text{F}}^2}{\|\mathbf{H}\|_{\text{F}}^2} \right] \quad (4.64)$$

as this residual captured both the sparse values related errors, but also the sparse support errors across the variously employed estimation combinations of solver and measurement matrix.

In Figure 4.5, the gains between the two idealized, practically infeasible, frame-theoretic beamforming schemes and the regular random beamforming are outlined for both OMP and simple BPDN algorithms. For these baseline scenarios, the performance increase of PTF-beamforming, and respectively, of optimally incoherent-tight beamforming with respect to both algorithms relative to the randomized beamforming is clearly visible and in accordance to the expectations given the previously outlined design differences among the schemes. The same can be said also for the performance gains of the proposed method of Subsection 4.3.3, over the PTF-beamforming [126] covered from the Frame Theory perspective in Subsection 4.3.2.

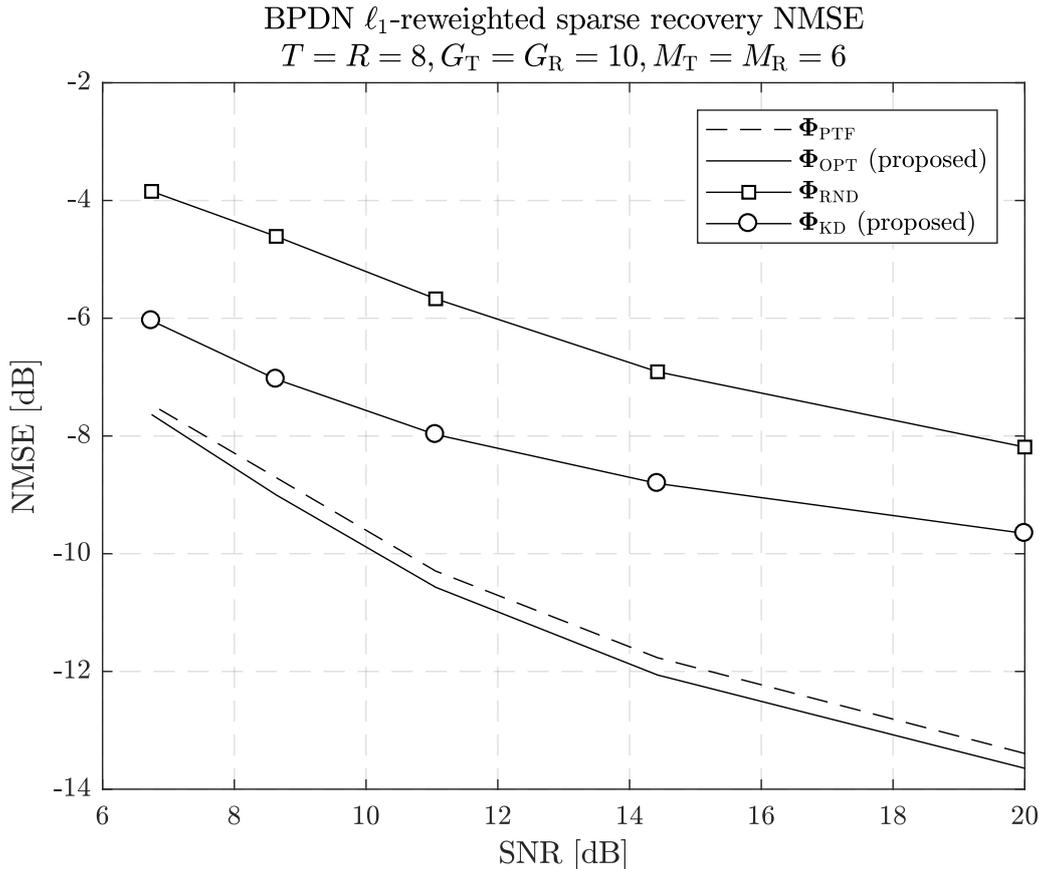


Figure 4.7: NMSE performance of mmWave channel sparse recovery using BPDN ℓ_1 -reweighted [134] for $M_T = M_R = 6$ across four different training beamforming schemes: PTF-beamforming Φ_{PTF} [126, 138], optimally incoherent-tight beamforming Φ_{OPT} proposed in Subsection 4.3.3, Φ_{RND} randomized Gaussian beamforming [114, 135], and respectively, KD-beamforming Φ_{KD} proposed in Subsection 4.3.4.

An additional remark to note is the difference in NMSE performance between the two algorithms. As expected, the simple BPDN algorithm outperforms OMP as it is much better suited to noisy sparse recovery than the latter [125, 126], being one of the best solvers on average across various conditions in the CS literature [132, 145]. Throughout the simulations, the δ residual estimate necessary for BPDN variants has been considered to be nothing but an upper bound of the expected noise magnitude, *i.e.*, $\delta > \mathbb{E}[\|n\|_2] = \sigma_n \sqrt{M_T M_R}$, such that the sparse solution would lie inside the residual hypersphere.

To outline the efficiency and scalability of the proposed optimal incoherent-tight beamforming against the PTF-beamforming on one hand, and respectively, of the practically realizable KD-beamforming versus the randomized Gaussian beamforming on the other hand, plots in Figures 4.6 - 4.8 were generated across various numbers of training sequences, *i.e.*, M_T, M_R . Concretely, M_T, M_R were set such that $M_T = M_R$ across all plots with values of $\{5, 6, 7\}$. In addition, based on the fact that under weighted sparsity enhancement BPDN ℓ_1 -reweighted is marginally superior to BPDN with each iteration [145], the ℓ_1 -reweighted variant was subsequently employed in order to highlight the potentially maximum achievable performance across

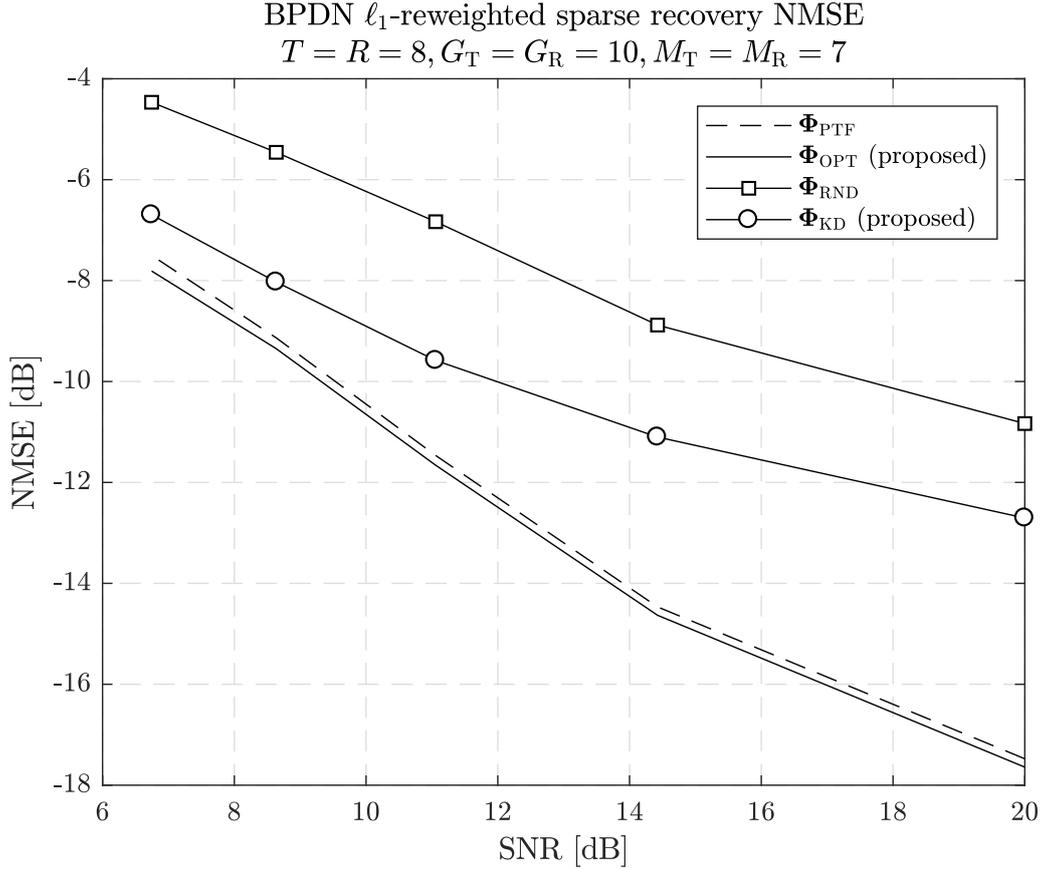


Figure 4.8: NMSE performance of mmWave channel sparse recovery using BPDN ℓ_1 -reweighted [134] for $M_T = M_R = 5$ across four different training beamforming schemes: PTF-beamforming Φ_{PTF} [126, 138], optimally incoherent-tight beamforming Φ_{OPT} proposed in Subsection 4.3.3, Φ_{RND} randomized Gaussian beamforming [114, 135], and respectively, KD-beamforming Φ_{KD} proposed in Subsection 4.3.4.

all scenarios and beamformers. Its maximum numbers of iterations was thus set to $I_{\max} = 5$.

Once more, the incremental gains obtained by the optimal incoherent-tight proposed training beamforming scheme against the PTF-beamforming state of the art are visible throughout Figures 4.6 - 4.8. However, the effects of increasing the number of beamformers, hence of measurements, it is visible as more advantage of the available beamforming space is taken. To this end, as the redundancy of the measurement frame reduces, *i.e.*, $\frac{TR}{M_T M_R} \rightarrow 1$, the gains in NMSE of the optimal incoherent-tight beamforming relative to the PTF-beamforming [126] fade away and the two methods come together. Nonetheless, under regular operation condition, with fast sweeps and normal and high redundancy, the proposed method outperforms the state of the art in terms of NMSE with up to 0.25 - 0.5 dB. Despite the fact that the gain value is not large, its normalized character stands to reason that this performance guarantees are typical and general between the two methods, and respectively, the optimal incoherent-tight approach leads in fact to a substantial better sparse recovery than the PTF approach [126, 138]. The more important value that the proposed optimal incoherent-tight training beamforming scheme brings, is its invariance of incoherence, tightness and equi-normality frame properties

under practical Kronecker factorization as exhibited in Subsection 4.3.4, leading to the realizable KD-beamforming.

The curves corresponding to the latter fall short in comparison to the jointly optimized beamforming schemes, but simultaneously provide a very good performance of about 2 dB NMSE on average over the randomized practically realizable equivalent. Furthermore, it is also visible that under the increase of the number of training beamforming vectors, the performance of the practically achievable schemes, *i.e.*, the KD-beamforming and the randomized beamforming, approaches the NMSE curves of their jointly optimized counterparts. Even so, the positive performance delta in favor of the KD method is maintained over the randomized one, which advocates for the general applicability of the frame-theoretic concepts utilized in the design of the former. Otherwise, all the curves exhibit the generally expected behavior of decreasing estimation performance with the increase in noise power levels.

4.5 Enhanced Sparse Solver for Channel Estimation

In the previous Subsection 4.4.2, the canonical ℓ_0 -norm minimization problem (4.18) has been suboptimally solved either greedily via OMP [123, 124], or approximated to an ℓ_1 -norm minimization via BPDN/ ℓ_1 -reweighted BPDN [145]. The ℓ_1 methods offer generally good performance under their approximation of the original ℓ_0 objective, but also are less robust in offering a sparse solution in noisy environments in comparison to the original problem (4.18). To challenge this conundrum, an ℓ_0 -norm alternatively approximated and transformed formulation of the canonical sparse recovery problem (4.18) is next derived and discussed, leading to a new iterative sparse mmWave channel estimation method.

Consider therefore the generic approximation of the ℓ_0 -norm introduced in [148] and applied to the vector $\mathbf{h} \in \mathbb{C}^{G_T G_R}$ as

$$\|\mathbf{h}\|_0 = \lim_{\alpha \rightarrow 0^+} \sum_{j=1}^{G_T G_R} \frac{|\mathbf{h}_j|}{|\mathbf{h}_j| + \alpha} = G_T G_R - \lim_{\alpha \rightarrow 0^+} \sum_{j=1}^{G_T G_R} \frac{\alpha}{|\mathbf{h}_j| + \alpha}. \quad (4.65)$$

The above expression is a parameterizable tight approximation of the ℓ_0 -norm whose tightness is steered via the parameter $\alpha > 0$. This leads actually for $0 < \alpha \ll 1$ to a much more accurate approximation than the ℓ_1 -norm regardless of the support size or the individual entries magnitudes \mathbf{h}_j of the vector \mathbf{h} . Applying the approximation (4.65) to the problem (4.18), the latter is transformed therefore to the optimization program

$$\begin{aligned} \underset{\mathbf{h} \in \mathbb{C}^{G_T G_R}}{\operatorname{argmin}} \quad & - \sum_{j=1}^{G_T G_R} \frac{\alpha}{|\mathbf{h}_j| + \alpha}, & (4.66a) \\ \text{subject to} \quad & \|\mathbf{y} - \Phi \Psi \mathbf{h}\|_2 \leq \delta, & (4.66b) \end{aligned}$$

where the hypersphere residual constraint follows similarly to (4.60b), and respectively, the non-negative approximation parameter is $\alpha \ll 1$.

The resulted problem (4.66) is categorized as part of standard family of non-convex optimiza-

tion programs. This is a result of the fact that the objective function (4.66a) is actually a sum of fractions containing concave-over-convex terms. This summation, also known as *concave-convex sum-ratio* in FPG, is in terms typically known as non-convex [149]. To tackle this issue, the authors in [149] have outlined that such non-convex functions expressed generally as sum-of-ratios of other non-convex terms can be transformed and convexized by the introduction of an additional parameter. This technique leads to a concrete partitioning of each individual term of the summation into difference of its numerator and denominator respectively. This transformation is realized by of the so-called *quadratic transform (QT)* technique in FPG terms [149]. Mathematically, a generic sum-of-ratios function can be transformed to a sum-of-quadratic functions given the QT equivalence

$$f(x) = \sum_{n=1}^N \frac{A_n(x)}{B_n(x)} = \sum_{n=1}^N 2\beta_n \sqrt{A_n(x)} - \beta_n^2 B_n(x) \quad (4.67)$$

where $A_n(x) : \mathbb{R} \rightarrow \mathbb{R}_+$ denotes an arbitrary nonnegative concave function, $B_m(x) : \mathbb{R} \rightarrow \mathbb{R}_{++}$ is an arbitrary strictly positive convex function, and x is a variable to be optimized over the constraint set \mathcal{X} .

The QT parameters $\beta_n, \forall n$, from the expression (4.67) above,

$$\beta_n \triangleq \frac{\sqrt{A_n(x)}}{B_n(x)}, \quad (4.68)$$

ensure that, for any fixed x , the original function $f(x)$ is in fact the equivalent of the QT resulted expression outlined in equation (4.67).

Nonetheless, the QT has been proposed only for continuously differentiable functions $A(x), B(x)$ [149]⁴. Upon this fact, the need of further manipulation to (4.66) is evident in order to be able to apply the QT and convexize the problem. To this end, the non-differentiable terms $|h_j|$ are replaced with slack variables t_j , forming the slack variable vector \mathbf{t} and resulting into the equivalent optimization program

$$\underset{\substack{\mathbf{h} \in \mathbb{C}^{G_T G_R}, \\ \mathbf{t} \in \mathbb{R}^{G_T G_R}}}{\text{argmin}} \quad - \sum_{j=1}^{G_T G_R} \frac{\alpha}{t_j + \alpha} \quad (4.69a)$$

$$\text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi} \mathbf{\Psi} \mathbf{h}\|_2 \leq \delta, \quad (4.69b)$$

$$|\mathbf{h}| \leq \mathbf{t}. \quad (4.69c)$$

It can be seen from the objective (4.69a) of the obtained problem (4.69) that the QT, [149], can be applied to convexize the original problem (4.66) given the continuously differentiable linearity of its denominator as a function of \mathbf{t} . Applying thus the QT to the objective function

⁴The reader interested in the QT and FPG is invited to take a mathematically detailed detour in [149] regarding these areas of Signal Processing and Optimization Theory.

of (4.69) above, the equivalent optimization problem

$$\begin{aligned} \underset{\substack{\mathbf{h} \in \mathbb{C}^{G_T G_R}; \\ \mathbf{t} \in \mathbb{R}^{G_T G_R}}}{\text{argmin}} \quad & - \sum_{j=1}^{G_T G_R} (2\beta_j \sqrt{\alpha} - \beta_j^2 (\mathbf{t}_j + \alpha)) \end{aligned} \quad (4.70a)$$

$$\text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi} \mathbf{\Psi} \mathbf{h}\|_2 \leq \delta, \quad (4.70b)$$

$$|\mathbf{h}| \leq \mathbf{t}, \quad (4.70c)$$

is obtained by fixing

$$\beta_j = \frac{\sqrt{\alpha}}{|\mathbf{h}_j| + \alpha} \quad \forall j, \quad (4.71)$$

following the original objective (4.66a) and (4.68).

Furthermore, given that the first term of the objective function described above is considered a constant in terms of the current search variables \mathbf{h} , and respectively, \mathbf{t} for all β_j , it can be dropped completely from the objective (4.70a). An analogous remark can be made regarding the terms $\beta_j^2 \alpha, \forall j$. So, in the light of the above, the final equivalent optimization program is resulted as

$$\underset{\substack{\mathbf{h} \in \mathbb{C}^{G_T G_R}; \\ \mathbf{t} \in \mathbb{R}^{G_T G_R}}}{\text{argmin}} \quad \mathbf{b}^T \mathbf{t} \quad (4.72a)$$

$$\text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi} \mathbf{\Psi} \mathbf{h}\|_2 \leq \delta, \quad (4.72b)$$

$$|\mathbf{h}| \leq \mathbf{t}, \quad (4.72c)$$

where implicitly the β_j^2 scalars were compacted by vectorization as

$$\mathbf{b} = [\beta_1^2, \beta_2^2, \dots, \beta_{G_T G_R}^2]^T \in \mathbb{R}^{G_T G_R}. \quad (4.73)$$

Algorithm 4 FPG MMWAVE CHANNEL ESTIMATION

Inputs: measurement matrix $\mathbf{\Phi} \triangleq (\mathbf{U}^T \otimes \mathbf{V}^H)$, dictionary matrix $\mathbf{\Psi} \triangleq (\mathbf{A}_T^* \otimes \mathbf{A}_R)$, residual upper bound δ , tightening parameter for ℓ_0 -norm approx. α , maximum number of iterations

I_{\max}

Outputs: sparse CSI estimate $\mathbf{h} \in \mathbb{C}^{G_T G_R}$

1: Set iteration counter $i = 0$

2: Set initial solution vector $\mathbf{h}^{(i)} = \mathbf{0}$

3: **repeat**

4: Update $\beta_j, \forall j$ as $\beta_j = \frac{\sqrt{\alpha}}{|\mathbf{h}_j^{(i)}| + \alpha}$ and form weight vector \mathbf{b} by (4.73)

5: Set $i = i + 1$

6: Obtain $\mathbf{h}^{(i)}$ by solving optimization problem (4.72)

7: **until** $\left(\frac{\|\mathbf{h}^{(i)} - \mathbf{h}^{(i-1)}\|_2^2}{G_T G_R} < 10^{-6} \text{ or } i > I_{\max} \right)$

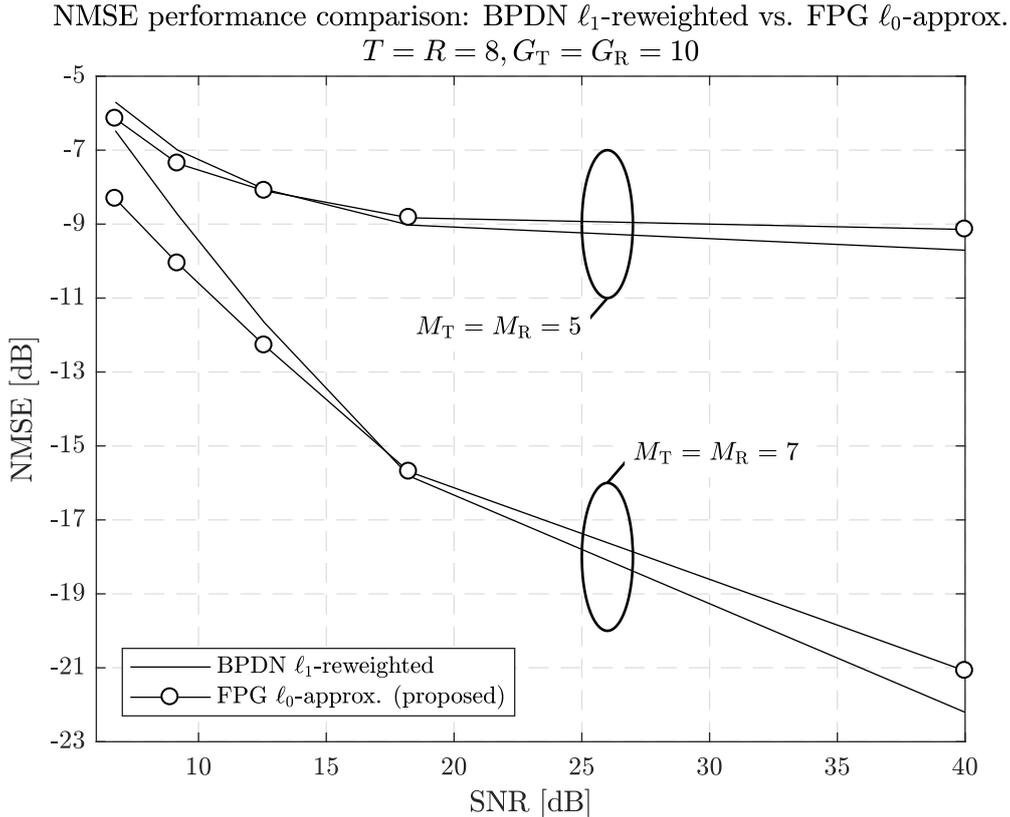


Figure 4.9: NMSE performance curves for BPDN ℓ_1 -reweighted and FPG ℓ_0 -approximated (proposed) estimators given systems with $M_T = M_R = 5$, and respectively, $M_T = M_R = 7$ training Tx/Rx beamforming/combining vectors. The training beamforming scheme applied is Φ_{OPT} across all realizations.

Remark that the final optimization problem (4.72) obtained upon successive equivalent transforms involving approximation of ℓ_0 -norm and FPG-based QT is as simple as a convex linear program, which represents the convexized epigraph form of the originally ℓ_0 -approximated problem (4.66). As a result, it can be efficiently solved via numerical convex solvers using interior point methods, such as for instance CVX [150].

What is more, given the concave-convex sum-ratio objective function in equation (4.69), it is guaranteed under the QT, [149, Th. 3], that the resulted equivalent optimization problem (4.72) converges actually to a stationary point by iteratively solving it, and subsequently updating \mathbf{b} according to (4.71) and (4.73) respectively. Upon this remark, the proposed sparse recovery method based on (4.72) is offered in full in Algorithm 4.

To highlight the expected upper bound performance gains of the proposed algorithm against the state-of-the-art ℓ_1 -reweighted BPDN, the same system scenarios as before were simulated. For the above goal however, the training beamforming strategy was selected to be the optimum, *i.e.*, the proposed joint optimal incoherent-tight approach. The NMSE results obtained are thus summarized by Figure 4.9. The performance of the FPG proposed method is on average better than the ℓ_1 -reweighted BPDN [145]. In particular the proposed algorithm provides improved performance over the state of the art in the low and medium SNR regions where the ℓ_0 -based

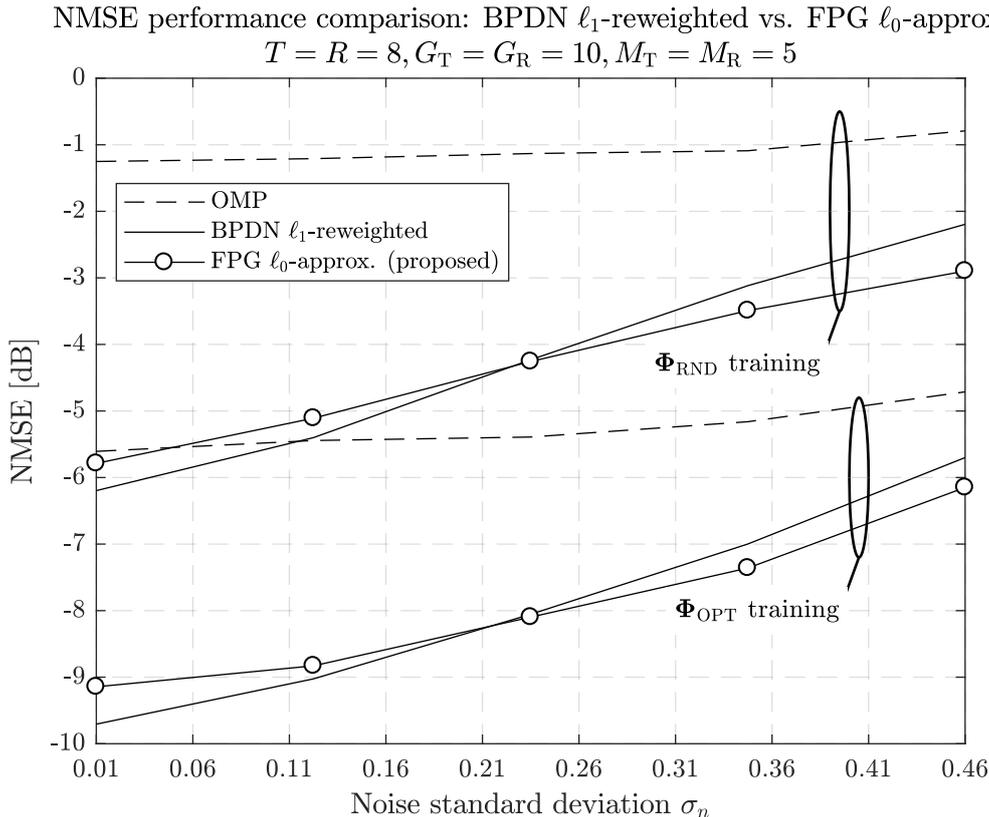


Figure 4.10: NMSE performance curves for BPDN ℓ_1 -reweighted and FPG ℓ_0 -approximated (proposed) estimators given systems with $M_T = M_R = 5$, and respectively, $M_T = M_R = 7$ training Tx/Rx beamforming/combining vectors. The training beamforming scheme applied is Φ_{OPT} across all realizations.

approximation (4.65) and its QT realization enforce better the physical sparsity constraint inherent in the recovered signal \hat{h} . The gains, however vary from less than 0.25 dB in the case of $M_T = M_R = 5$ to almost 2 dB for the realizations with $M_T = M_R = 7$. These are obtained for the low SNR areas where in fact sparse recovery is most problematic, thus arguing in favor of the reconstruction robustness the proposed ℓ_0 -centered method has over the ℓ_1 -focused basis pursuit approach [145]. Furthermore, the loss in performance in the high SNR regions is maintained almost constant to on average 0.5 dB regardless whether 5 or 7 training beamforming vectors were used at the Tx, respectively at the Rx.

Finally to add more evidence in support of the general claims made above related to the performance of the FPG proposed sparse solver in comparison to the ℓ_1 -reweighted BPDN, the Figure 4.10 was plotted. This includes both the optimal beamforming curves discussed above, but also the ones of the randomized beamforming measurement matrix, both plotted for the case $M_T = M_R = 5$. Therefore, Figure 4.10 outlines the fact that the general qualitative and quantitative behavior already discussed for Figure 4.9 are invariant and scale horizontally regardless of the chosen measurement matrices. To this point, the randomized beamforming given by Φ_{RND} was selected.

4.6 Conclusions

We proposed in this chapter a deep analysis of the current mmWave channel estimation problem from the perspective of Frame Theory. To this end, the state-of-the-art sparse formulation [107] of the mmWave CSI estimation was used. We highlighted the applicability of frame-theoretic concepts both at the level of virtual dictionary creation under the geometry of the antenna arrays of mmWave radios, but also in the design of mmWave optimized training beamformers for synchronization and channel estimation sequences employed in the context of IA. Therefore, the CS central problem to mmWave channel estimation was dissected and reviewed through the prism of Frame Theory, thus linking the latter and its CS applications in a closed-circle mathematical treatment.

As an outcome, the design of training beamforming sequences was summarized down to the design of measurement frames of CS linear systems. In conjunction with CS recovery guarantees the optimal design criteria for such frames were identified and elaborated as *measurement incoherence* and *sensing tightness*, respectively. As a result, two frame-theoretic beamforming schemes were proposed. The first one, *i.e.*, the *optimal incoherent-tight beamforming*, jointly optimized the Tx/Rx training beamforming vectors disregarding the physical separation of the Tx/Rx, outperforming the state-of-the-art equivalent of PTF-beamforming [126,138]. However, these schemes both were argued as being impractical given their disregard of separate Tx/Rx radio domains, and for this reason, the second method, the *KD-beamforming*, was proposed as a mean to fill in the practicality gap. The method proved in effect to consistently outperform the reference solution of randomized Gaussian beamforming [114]. The performance delta originated from the same general design principles of measurement incoherence and sensing tightness that were distributed to the Tx/Rx sides via the Kronecker product.

Lastly, we contributed to the robust sparse recovery mmWave estimation as well, by providing an alternative ℓ_0 -approximated convexized FPG solver capable of taking advantage of its ℓ_0 -norm based objective. The algorithm proved consistently on average better and more robust than the BPDN- ℓ_1 [145] state of the art, particularly in the medium and low SNR regimes, whereas the incurred NMSE performance losses for the high SNR cases were acceptably low.

Theoretic at its roots, the work on this subject of mmWave channel estimation developed throughout this chapter may be taken further by considering the practical hybrid beamforming realizations of the proposed channel measurement strategies and losses incurred in the process. Furthermore, massive MIMO mmWave extensions may be directly considered on top of the training beamforming methods, together with the accelerated realizations of the sparse solvers utilized.

Chapter 5

Non-orthogonal Multiple Access Framework via Frame Theory

As previously seen, given their redundant representational design, frames have an inherent capacity to both expand and compress signals depending on the input perspectives and dimensionality implied. Chapter 4 has detailed previously the link between CS and frame-theoretic properties and concepts, whereas the latter were used to optimize the sparse recovery synthesis of mmWave channels. Continuing the set trend, this chapter applies the same frame-driven analysis approach to the area of NOMA. This idea is motivated by the historical pattern of depleting communications resources experienced periodically by legacy communications systems, as detailed in Chapter 2.

The NOMA problem is thus decomposed under a generic frame-based treaty suited to 5G+ and next-generation wireless communications networks. The framework hereby introduced abstracts the core ideas of NOMA signal transmission/reception and captures their essence formulated as a frame design problem on the Tx side and as an overloaded multiuser detection task on the Rx side, respectively. Unsurprisingly to some extent, similar frame attributes as the ones essential to Chapter 4 beamforming designs are found mathematically to receive the best information-theoretic dividends under the generic NOMA disambiguation proposed.

The findings hereafter provide new techniques in terms of NOMA signal processing based on

Excerpts of this chapter are adapted and enhanced based on the selected articles:

R.-A. Stoica and G. T. Freitas de Abreu, “Massively Concurrent NOMA: A Frame-Theoretic Design for Non-Orthogonal Multiple Access,” in *IEEE 52nd Asilomar Conference on Signals, Systems, and Computers*, Oct. 2018, pp. 461-466. ©2018 IEEE

T. Hara, R.-A. Stoica, K. Ishibashi and G. T. Freitas de Abreu, “On the Sum-Rate Capacity and Spectral Efficiency Gains of Massively Concurrent NOMA Systems,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2019, to appear. ©2019 IEEE

R.-A. Stoica, G. T. Freitas de Abreu, T. Hara, K. Ishibashi, “Massively Concurrent Non-orthogonal Multiple Access for 5G Networks and Beyond,” in *IEEE Access – Special Section on Advances in Signal Processing for Non-orthogonal Multiple Access*, vol. 7, pp. 82080-82100, July 2019. ©2019 IEEE

R.-A. Stoica and G. T. Freitas de Abreu, “A Low-complexity Receiver for Massively Concurrent Non-orthogonal Multiple Access,” in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1-5. ©2019 IEEE

native signal space non-orthogonal representations and recovery methods applicable at scale. It is hoped therefore that in turn these advances provide the necessary toolset and features of solving at least partly some of the concerns and computationally restrictive points that have impeded so far the integration of NOMA communications within the current [9], and respectively, near-future [10] 5G standardization.

5.1 Core Idea and Preliminaries

The non-orthogonality idea in communications has always been a taboo subject. In practice, systems were designed over the years always along the premises of avoiding interference among signals and users, forcing therefore orthogonality at the signal space level as outlined over the course of Chapter 2. As seen though, the increasing number of users together with the fixed parametric orthogonal signaling employed jointly so far in conventional systems are not massively scalable as their communication signal space and resources are exploited gradually to depletion over time.

Under this high-level perspective, the idea of non-orthogonal user access schemes is well-founded albeit its introduced interference. A first glimpse of this strategy was on the brink of emerging actually given the turn of the century hype for CDMA communications, but eventually failed to do so upon the complexity of the multiuser detection which increased well beyond that of the envisioned matched filtering [4]. Essentially, CDMA slowly gave way to OFDM as detailed in Section 2.1.

Recognizing such shortfalls of orthogonality and identifying the upcoming fight over the scarce spectrum resources for increased communication rates and access density, NOMA was introduced as a standalone concept in [16]. At its core, the early NOMA concept relied on the simple yet effective idea of *power domain superposition* where power controlled multiplexing of two or more users was used [16, 17]. To highlight this core concept, a simple example is next provided with the gains incurred over a traditional orthogonal access system.

Example 5.1.1. *Consider without loss of generality the DL scenario where two users are served by a BS over a single carrier system with normalized bandwidth $W = 1\text{Hz}$ and total power constraint $P_t = 1$. Additionally, assume that user 1 experiences a worse channel than user 2 such that the channel ordering is $|h_1| \leq |h_2|$. The NOMA DL scheme transmits both user 1 and user 2 signals during one channel utilization over the available bandwidth taking advantage of this channel disparity. Concretely, since NOMA is enabled by successive interference cancellation (SIC), to allow the system to work and have both users capable of finally detecting their corresponding signals, the core idea of NOMA is to help the weak user, i.e., user 1, while still allowing user 2 to perform SIC. Thus, the BS amplifies the weak user signal s_1 by a power coefficient γ_1 , and equivalently it does so for user 2 by γ_2 , such that the total power constraint P_t is met. Hence, the NOMA Tx signal becomes*

$$x = \sqrt{\gamma_1}s_1 + \sqrt{\gamma_2}s_2 \quad (5.1)$$

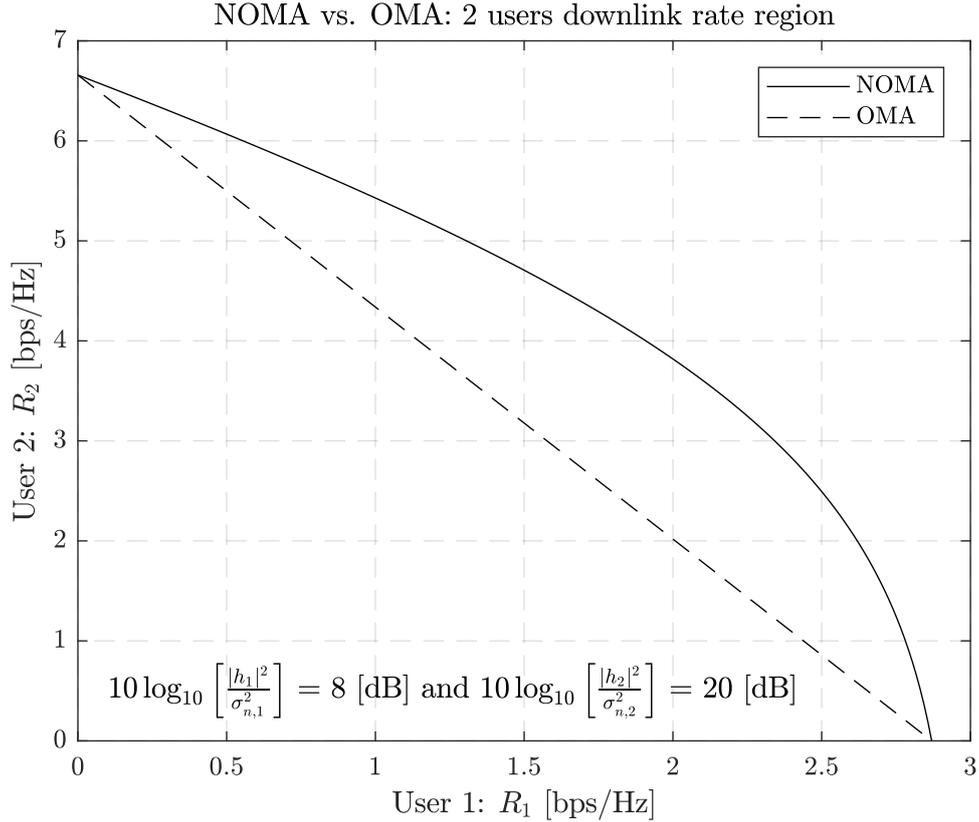


Figure 5.1: Illustration of the NOMA achievable rate region for a 2 user single carrier DL communication system in comparison to its OMA counterpart.

where the individual signals s_1, s_2 are considered to be normalized as $\mathbb{E}[|s_1|^2] = \mathbb{E}[|s_2|^2] = 1$.

Under the above ordering and provided that the detailed power control is enabled, the achievable rates of the two users are obtained as

$$R_1^{\text{NOMA}} \leq \log_2 \left(1 + \frac{\gamma_1 |h_1|^2}{\gamma_2 |h_1|^2 + \sigma_{n,1}^2} \right) \quad (5.2a)$$

$$R_2^{\text{NOMA}} \leq \log_2 \left(1 + \frac{\gamma_2 |h_2|^2}{\sigma_{n,2}^2} \right), \quad (5.2b)$$

where implicitly SIC detection was considered for the strong user, i.e., user 2.

For the latter to be always guaranteed, and respectively, to simultaneously also help user 1 decode its own signal treating user 2 as Gaussian noise, the NOMA power control idea implements the constraint $\sqrt{\gamma_1} \geq \sqrt{\gamma_2}$. As a result, user 1 decodes its message directly considering the signal of user 2 as noise, whereas user 2 will always be able to decode s_1 first, remove it and then detect its own signal as a consequence of the fact [18] that

$$\log_2 \left(1 + \frac{\gamma_1 |h_1|^2}{\gamma_2 |h_1|^2 + \sigma_{n,1}^2} \right) \leq \log_2 \left(1 + \frac{\gamma_1 |h_2|^2}{\gamma_2 |h_2|^2 + \sigma_{n,2}^2} \right). \quad (5.3)$$

On the contrary, orthogonal multiple access (OMA) excludes spectrum sharing by providing

exclusive access to the media resource. Given the single carrier nature of the considered system, thus TDMA is performed and orthogonal time sharing is in effect such that the achievable rates of the two users are obtained simply as

$$R_1^{\text{OMA}} \leq \log_2 \left(1 + \frac{|h_1|^2}{\sigma_{n,1}^2} \right) \quad (5.4a)$$

$$R_2^{\text{OMA}} \leq \log_2 \left(1 + \frac{|h_2|^2}{\sigma_{n,2}^2} \right), \quad (5.4b)$$

where given the exclusive transmissions the total power available was used for each transmission to maximize the individual rates.

The system achievable sum-rate is given therefore by

$$R \leq R_1 + R_2, \quad (5.5)$$

where R_1 , R_2 correspond to the achievable rates of either the NOMA or the OMA system realization respectively.

The theoretic rate region of NOMA compared to the one of conventional OMA is plotted in Figure 5.1 for the particular scenario where the CNR of the two users were set to $10 \log_{10} \left[\frac{|h_1|^2}{|\sigma_{n,1}|^2} \right] = 8$ dB for the user 1, and respectively, $10 \log_{10} \left[\frac{|h_2|^2}{|\sigma_{n,2}|^2} \right] = 20$ dB for user 2. It is important to note however that $R^{\text{NOMA}} \geq R^{\text{OMA}}$ [18, 151, 152] with equality met for $|h_1| = |h_2|$. As a consequence of this fact, NOMA is in fact using the channel gains diversity as a catalyst to encourage spectrum reusage and simultaneously increase communication rates.

The above NOMA core idea can also be applied to the UL scenario where SIC is performed on the BS to detect the multiple signals simultaneously received. At this point, it should be already clear that the term NOMA itself is overloaded. It is not used to denote alone the traditional information-theoretic *multiple access* meaning of UL, but interchangeably also to express the DL case. In fact, the multiple access in NOMA refers to *concurrent access* both in UL and DL. Under these premises, NOMA is fundamentally a multi-user simultaneous transmission (MUST) system limited by its self-induced interference. However, these designed distortions bring also advantages that make NOMA desirable. To this end, an instantiation of the Example 5.1.1 was also considered for inclusion as an advanced addendum to 4G LTE-Advanced for 2 users [153]. Once again the idea of non-orthogonality resurfaced thence towards the end cycle of development of 4G access techniques.

Along the lines of the above, it can be said that non-orthogonality as a core concept surfaced in both earlier cases, *i.e.* 3G and 4G, as limited late amendments to already matured wireless technologies via non-orthogonal CDMA [4, 154], and respectively, via MUST and carrier aggregation techniques [155]. However, the recent research works, nicely summarized for instance under [18, 152, 156], provide clear proofs of the NOMA gains over the conventional OMA. Among such positive deltas are in fact the increased system rates (for instance UL power-based NOMA can achieve the theoretic SIC-capacity of the MA channel [151]), the spectrum reusage

and sharing paradigm of scarce frequency resources, the improved user fairness, and lastly, the more performant QoS guarantees that NOMA offers opposite to OMA.

5.2 A Classification of Existing Solutions and Practical Limitations

To capitalize on the theoretic promised gains of NOMA and to achieve higher spectral efficiency than the one of current systems, non-orthogonality has thus been considered now as a core part of the future communication networks, rather than as an add-on as in the past [152]. Given the self-interference nature of NOMA systems, one key problem associated with their success is the design of controlled multi-user access patterns enabling the service of massively more users than the signal space dimensionality. As a result, over the last five years a multitude of NOMA schemes were proposed to this end [156]. However, they are all based on the core, very simple idea of *non-orthogonality* implemented over some signal space domain. A general, non-exhaustive, but largely accepted dichotomy [152,156] of NOMA systems is used to separate them under two categories: power-domain NOMA (PD-NOMA), and respectively, code-domain NOMA (CD-NOMA).

Obviously, Example 5.1.1 illustrated a concrete realization of PD-NOMA for a naive and simple 2 users DL case. Generally though, as seen, such systems are resumed to the principle of *superposition of multiple signals over the power domain*. To this extent, two or more users are linearly combined by weighted power coefficients. The associated power control needs therefore to be optimized given all the active users according to some common system-level objective and under the constraints subsequent to the inherent SIC detection scheme. As highlighted in Example 5.1.1, usually the latter involves prior knowledge of side information regarding the available CSI across the user pool and the BS, but also other additional individual constraints, such as power or required rates for instance [157]. As a consequence, the PD-NOMA realization is bounded to solving problems of resource allocation and user selection/clusterization [157, 158] seeking to maximize the system achievable sum-rate for instance. Provided that such optimization problems are indeed solved, then PD-NOMA is known to provide significant gains over equivalent OMA systems [158].

Although very simple at its origins, the idea of PD-NOMA leads therefore to associated problems which are inherently of combinatorial nature, *e.g.* optimum users clusterization and power allocation for maximizing sum-rate with QoS constraints [159,160]. Whereas solving such optimization problems for a small number of users is a computationally tractable task, this cannot be said anymore for the large scale systems NOMA is targeting [156,158]. Typical solutions to such problems involve new techniques such as *monotonic optimization* where the monotonically increasing objectives and constraints are bounded and thus approximated by polyhedra and polyblocks which are subsequently shrunk such that the globally optimum solutions are reached in a more tractable worst-case computational time [161]. For instance such problems were solved optimally for joint clusterization and resource allocation of differently constrained NOMA systems in [159,160,162]. Despite the fact that monotonic optimization reduces the

runtimes and provides computationally manageable paths to solving optimally or suboptimally such problems, it is actually quite heuristic in nature and its performance depends highly on the type of conditions imposed [161], which makes the method hard to scale universally. To top it off, PD-NOMA assumptions on the existence of side CSI partial, statistical or full information across all active users at the BS is yet again another hard to scale necessary systemic input, which highly increases the communication overhead.

The non-orthogonally multiplexing of CD-NOMA is constructed in contrast to PD-NOMA based on bit-, symbol- or code-level overloading, hence its nomenclature [152, 156]. This approach provides actually a relaxation of the massive resource allocation and user selection problems complexity of PD-NOMA. However, as nothing comes for free in reality, the complexity is, albeit reduced, moved to the receiver side where, instead of SIC, CD-NOMA systems necessitate more advanced algorithms and modern signal processing techniques for multiuser detection as a consequence of the code-level information superposition.

Despite the CD-NOMA paradigm shift in non-orthogonality with regards to PD-NOMA, the code-based approaches are capable of reaching similar performance as the latter. [152, 156]. As a consequence, the complexity-performance balance tilts on the side of CD-NOMA rather than PD-NOMA [18].

Furthermore, by moving the non-orthogonality multiplexing to bits or symbols level, CD-NOMA enabled more flexible and diverse designs in achieving high-performing NOMA systems. The effects of this design flexibility led therefore to a diversification of CD-NOMA methods over different realizations which may be categorized into four different subcategories according to the techniques used in fact to compress, and respectively, multiplex non-orthogonally the user streams of the active users over the physically available communication resources. Concretely, these subcategories may be differentiated as *scrambling*, *interleaving*, *spreading* and *coding* based CD-NOMA instances [156].

The scrambling CD-NOMA methods utilize scrambling signatures as a method to differentiate among the active users. An example is Qualcomm's resource spread multiple access (RSMA) system [163] which multiplexes users non-orthogonally by means of low-correlation long PN-sequences. The reciprocal descrambling operation has then the role to improve individual users' SNR and aid the multi-user detection (MUD) usually implemented as a straightforward minimum mean squared error (MMSE) with SIC or parallel interference cancellation (PIC) [164]. The potential errors of the detection are then further corrected by low-rate multi-user error correction codes [163].

One should maybe note that in fact by extrapolation of the concept of scrambling to using an entire *transparent* sequence of 1s, PD-NOMA itself can be seen prospectively as the simplest scrambling based NOMA scheme as well.

In a similar fashion, the interleaving and spreading CD-NOMA systems focus on lowering and controlling the inter-user interference by various methods involving bit/symbol interleaving/spreading techniques. These alter thus the original user information signals in a controlled way to aid the receivers cope with the system overloading interference. An instance of interleaving CD-NOMA is for instance Nokia's interleave division multiple access (IDMA) scheme

proposed in [165]. This utilizes different random patterns to interleave the overloaded number of active users. The obtained streams are then jointly combined over each RE and transmitted over the wireless channel. At the Rx side, the low-complexity elementary signal estimator (ESE) algorithm is used for detection, enhanced by extrinsic log-likelihood ratio (LLR) feedback of the channel forward-error correction decoder structure [165,166].

Another instance of interleaving CD-NOMA is Samsung’s interleave-grid multiple access (IGMA) [167]. This is in fact an enhanced variation of IDMA which uses interleaving both at the bit and symbol levels subsequently enhanced by sparse overlapping of the resulted interleaved symbols onto the available RE grid. The sparsity of the multiplexing is in fact inspired by low-density spreading (LDS) [168–170], making IGMA a hybrid of interleaving and spreading CD-NOMA.

From the CD-NOMA spreading perspective, some further examples are ZTE’s multi-user shared access (MUSA) [171], LG’s non-orthogonal coded multiple access (NCMA) [172], or respectively, Nokia’s non-orthogonal coded access (NOCA) [173] implementations.

Despite the various existent schemes of spreading NOMA approaches enumerated above, all are in fact relying on pseudo-randomly or carefully crafted spreading sequences which superimpose streams over all available RE, along the same core idea of CDMA. However, having dropped the orthogonality or pseudo-orthogonality constraint of CDMA in favor of system overloading, the design of the spreading sequences is required to be more complex and potentially lead to necessarily non-binary spreading codes. For instance, take the example of the quantized low-correlation random complex spreading sequences utilized in ZTE’s MUSA. These aim to design individual signatures for active users’ access and to provide the necessary support for grant-free communications [174], and thus, reduce signaling overhead and improve spurious UL communication links particularly of interest for instance in the context of IoT. To this end, tertiary 2D “chips” are used, in fact complex realizations of real and imaginary tertiary sequences $\{-1, 0, 1\}$, such that relatively low-complex MMSE-SIC (MMSE-SIC) or MMSE-PIC (MMSE-PIC) could still be used on the receiver side.

Lastly, code-based CD-NOMA methods rely on data streams mapped on top of specially designed sparse codebooks or linear symbol combinations describing their multiplexing over the different RE domains available for communication, *i.e.*, time, frequency or space. The two best and most popular code-centric CD-NOMA schemes are respectively, the pattern division multiple access (PDMA) [175–177] and Huawei’s sparse-coded multiple access (SCMA) [178, 179]. The receivers of these two access perform joint MUD via the message passing algorithm (MPA) [180] to that extent that the inherently designed sparsity is used to reduce the associated factor graph connectivity, [178], and thus keep the complexity of the MPA low as it scales exponentially with the size of the codebook.

In other words, *sparsity* is used as *a way to reduce the complexity of the MUD* on the Rx side *in the detriment of fully taking advantage of the signal space dimensionality* given the overloading factor of the system. Let us elaborate a bit more on the role the sparsity plays for these CD-NOMA designs by some examples. Consider to this end, the PDMA approach [175], where linear patterned operators are applied as precoders to users such that these are combined

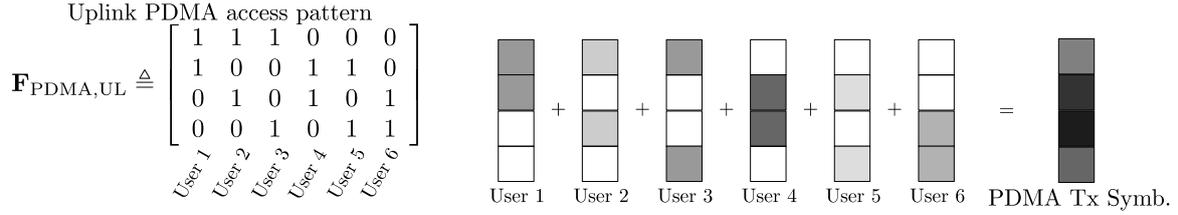


Figure 5.2: An illustration of the UL PDMA active user non-orthogonal multiplexing for $M = 4$ REs and $K = 6$ users.

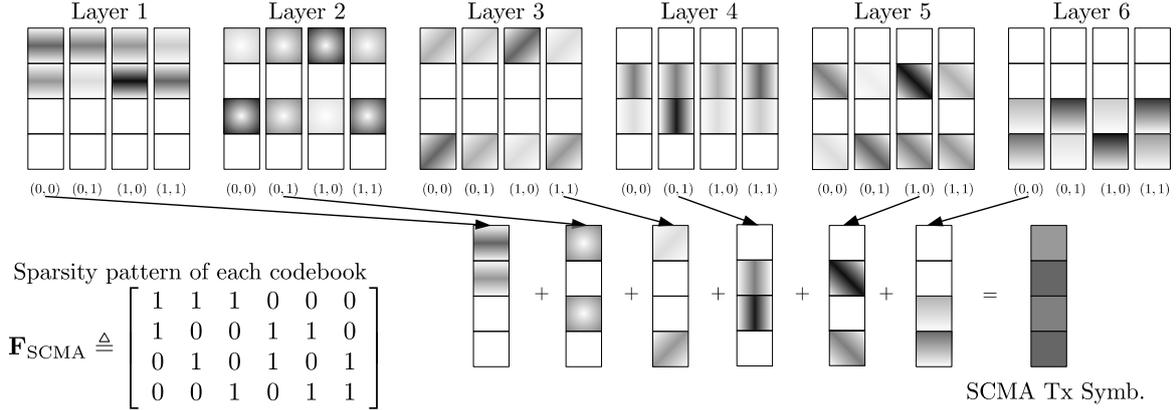


Figure 5.3: An illustration of the SCMA active user non-orthogonal multiplexing for $M = 4$ REs and $K = 6$ users.

with some sparsity on top of the RE. A pictorial depiction of the concept is portrayed in Figure 5.2 for the particular realization of UL overloading of a 4 REs system with 6 active users by means of the pattern matrix $\mathbf{F}_{\text{PDMA,UL}}$ [176]. As it can be seen from Figure 5.2, the symbols of the users are linearly combined such that each RE contains contributions of at most 3 users out of the total of 6 active ones. Essentially, these pattern matrices are generally linear binary operators mapping the active users on top of existing resources with or without additional power control. Upon knowing $\mathbf{F}_{\text{PDMA,UL}}$ at the Rx, it follows that the complexity of the search space has been reduced as the factor graph associated with the MUD problem is not fully connected. Analogously, PDMA behaves in DL [177]. Clearly, the practical performance of the MUD depends therefore on the pattern matrix realizations.

On the other hand, even though SCMA has a similar sparse allocation strategy, the user streams under consideration are abstracted and regarded as *layers*. Subsequently, SCMA designs for each layer an optimized sparse complex multi-dimensional codebooks of the same size as the constellation dimension of each user [179]. These multi-dimensional points undergo finally a constellation “shaping” by means of unitary rotations of their mother constellations such that the possible overlaps of collided constellation points are reduced, whereas codewords are sparsely represented and multiplexed to help the MUD simultaneously decrease complexity, but also maintain a good performance under controlled interference. The codebook sparsity is finally enforced upon a similar pattern matrix as for PDMA, where the codebook components are assigned to the non-zero entries of the sparse pattern essentially determining the factor graph

structure of the scheme. As a result, SCMA achieves a similar low density spreading to the idea of LDS [169, 178]. The structure of SCMA for the similar PDMA UL scenario of 4 available REs exploited by 6 active users simultaneously is presented in Figure 5.3. Based on the SCMA design, it follows that the codebook optimization and its subsequent constellation shaping are essential for the performance of the scheme. The codebook acts therefore as the core component in this regards, determining the high-level effects, whereas the constellation shaping provides additional gains, lowering even more the effects of interference, since in fact constellation shaping is known to provide improved performance over interference limited systems [181].

5.3 Frame-theoretic Representation of Non-orthogonal Access

As shortly outlined in the previous section, the research behind NOMA in general is very diverse. The rest of this chapter will however focus on the CD-NOMA localized context. Even so, the latter, as seen already, is quite eclectic, spanning over various schemes with multiple variations in realizing the wireless system non-orthogonal overloading. Nevertheless, some ideas are in fact recurrent, as seen already for instance with the reuse of special codes/scramblers meant to mix the active users' signals under controlled interference levels, somewhat similar to the overloaded CDMA performed at the bit level.

Moreover, among even the best performing CD-NOMA schemes [152, 156], *i.e.* PDMA [175, 176], and respectively, SCMA [178, 179], there can be identified some elements which suggest that optimal performance may yet not be attained. Concretely, this is the case of the sparsity constraints imposed allegedly to reduce the complexity of the decoding schemes. From a pure information-theoretic perspective, sparsity and structured constrained matrices can only decrease the total entropy available within a prescribed dimensionality given by the physical signal space.

In light of all of the above, we challenged the status quo of the CD-NOMA state of the art by asking and providing concrete answers to the following fundamental questions:

Q1 – *Can CD-NOMA be abstracted under a unified framework?*

Q2 – *What are the optimum design criteria for CD-NOMA best performance?*

Q3 – *Is the optimum performance practically realizable?*

Q4 – *How scalable is the optimum design?*

Let us first start with addressing Question Q1 throughout this section. Provided that essentially under an overloaded system scenario the active users' signals are multiplexed on less REs than the dimensionality of the joint virtual user space, an abstract *compressive mapping* between the collective active user signal space, and respectively, the available REs is necessary. Such a mapping, in fact a signal compression, is nothing but the joint linear allocator of REs and information precoder. Furthermore, as a pure mathematical operation, this is identified as the *synthesis* linear mapping of *Frame Theory* [19, 20], earlier discussed in Chapter 3.

Upon this remark, the CD-NOMA non-orthogonal multiplexing is thereby reduced to a *frame design* problem. The latter amounts to the assignment of optimized signature waveforms to users, thus *optimally spreading the active users over all resources simultaneously*.

Consider the communication system supposed to serve K active users constrained to having available only a group of $M < K$ orthogonal REs. Without loss of generality, hereafter the REs are envisioned to be carrier tones, such as for instance frequency subcarriers of a conventional OFDM system. This does not go against leveraging alternative singular orthogonal dimensions, *e.g.* time or antennas resources, but also joint domains, such as the space-time combination for instance.

Furthermore, the active K users are assumed, once more without loss of generality, to have the same source bit generation rate, and respectively, the same constellations \mathcal{S} of cardinality $|\mathcal{S}| = Q$. As a consequence,

$$B = K \log_2(Q) \quad (5.6)$$

bits need to be served simultaneously considering all the active users.

Upon the fact that $M < K$, radiating the entire information of the K users over the M REs would be without a doubt limited by inter-user interference, whereas the MUD would be distorted as well. However, for the transmission to be firstly realizable on the Tx side alone, one needs to account for the REs bottleneck and compress the joint users' information to the available signal space of dimensionality M . As a consequence, a *compressing mapping*, \mathcal{V} , must be used to multiplex *all information bits* through the overloaded system, *i.e.*

$$\mathcal{V} : \{0, 1\}^B \rightarrow \mathbb{C}^M, \quad (5.7)$$

with \mathcal{V} decomposable as $\mathcal{F} \circ \mathcal{B}$, such that

$$\mathcal{B} : \{0, 1\}^B \rightarrow \mathbb{C}^K, \quad (5.8a)$$

$$\mathcal{F} : \mathbb{C}^K \rightarrow \mathbb{C}^M. \quad (5.8b)$$

The latter decomposition model, $\mathcal{V} \triangleq \mathcal{B} \circ \mathcal{F}$, is used to abstractly describe the two-stage process of firstly modulating the information bits into a virtual *super-symbol* \mathbf{s} , and secondly compressing the information over the available RE space. In more detail, the virtual super-symbol \mathbf{s} is built by stacking together the individual constellation symbol mappings performed by each independent active user leading therefore to the explicit signal

$$\mathbf{s} \triangleq [s_1, s_2, \dots, s_K]^T \in \mathbb{C}^K. \quad (5.9)$$

The second stage, describes the linear compression operator \mathcal{F} , which is tasked to combining all the possible super-symbol realizations to the constrained physical signal space spanned under the REs domain, *i.e.*, \mathbb{C}^M , as detailed by the linear transform (5.8b).

Applying all of the above under a classical linearly modeled wireless communications system,

the concurrent access over an AWGN channel is resumed respectively to

$$\mathbf{y} = \mathbf{P}^{\frac{1}{2}} \mathbf{F} \mathbf{s} + \mathbf{n}, \quad (5.10)$$

with the radiated complex signal being formed of

$$\mathbf{x} \triangleq \mathbf{P}^{\frac{1}{2}} \mathbf{F} \mathbf{s}, \quad (5.11)$$

as a non-orthogonal superposition of *linearly compressed* active users symbols.

The matrix \mathbf{P} introduced in the model (5.10) describes the optional power control that could be applied to the Tx signal \mathbf{x} in (5.11). To this extent, \mathbf{P} is a diagonally structured power policy matrix. In addition, under the proposed modeling, prospective supplemental *bit-level operations*, *shaping* and *detection decorrelative techniques*, are applied independently of \mathcal{F} through the bits-to-symbols mapping \mathcal{B} from (5.8a). Finally, the *multiplexing / allocation / spreading* non-orthogonal concurrent access strategy over the physical RE is implemented by the linear mapping \mathbf{F} as \mathcal{F} from (5.8b).

Mathematically, given the generic overloaded concurrent access system (5.10), $\mathbf{F} \in \mathbb{C}^{M \times K}$ is a rank-deficient matrix upon $M < K$ that should exploit diversity fully and thus take advantage of the entire available signal space \mathbb{C}^M , *i.e.* $\text{rank}(\mathbf{F}) = M$. As seen in Chapter 3, this implies that \mathbf{F} is in fact the synthesis operator from \mathbb{C}^K to \mathbb{C}^M whose columns are formed by frame vectors over \mathbb{C}^M , *i.e.* by Definition 3.2.1, \mathbf{F} is a frame.

Having proposed the overloaded concurrent transmission model (5.10) for the AWGN channel above, let us now consider more realistic scenarios and outline its applicability. Consider therefore in the sequel the DL and UL practical realizations of CD-NOMA systems under fading wireless communications channels.

Downlink

The DL scenario is applicable to a multitude of upcoming applications in the context of URLLC and mMTC in future communication networks. In opposition to past practices, the DL case, and specifically the massive concurrent DL scenario, is becoming more and more important with the envisioned proliferation of new data-hungry applications that require a heavy DL component. Some concrete examples are for instance V2X communications, *e.g.*, semaphore-less automated traffic junctions, tele-operated valet parking, real-time sensor data exchange, high-definition dynamic maps and information exchange; massive data download in edge computing, *e.g.*, automated factory / logistics optimization operations applied at the edge networks; and high-throughput data streaming, *e.g.*, AR / VR empowered tours, large eGaming events, *etc.*

The assumption is therefore that an information source, for example a BS, downlinks data simultaneously by means of broadcast methods to a massive amount of users, larger in numbers than the traditional orthogonal resources that the system possesses. Furthermore, heterogeneous rates and QoS requirements may be expected by the downlink terminal nodes. Under these

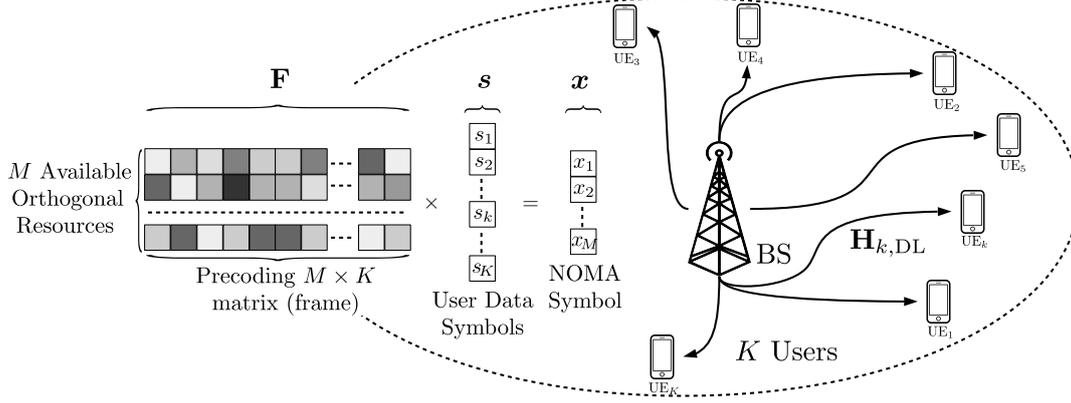


Figure 5.4: An illustration of a DL NOMA system realization according to (5.12).

conditions NOMA transmissions can help currently orthogonally limited systems. With these assumptions, and considering a single antenna BS, respectively UEs deployments, the DL receive signal at any user k is

$$\mathbf{y}_k = \mathbf{H}_{k,DL} \mathbf{F} \cdot \mathbf{s} + \mathbf{n}_k, \quad (5.12)$$

with $\mathbf{H}_{k,DL} \triangleq \text{diag}([h_{1,k}, h_{2,k}, \dots, h_{M,k}]^T) \in \mathbb{C}^{M \times M}$ the diagonal Rayleigh fading channel matrix between the BS and k -th UE, $\mathbf{s} \in \mathcal{S}^K \subset \mathbb{C}^K$ the super-symbol vector containing all users information, $\mathbf{n}_k \sim \mathcal{CN}(0, \sigma \mathbf{I}_M)$ the AWGN, and the *modulated signature waveforms* of each individual user as the columns of the complex *frame* $\mathbf{F} \in \mathbb{C}^{M \times K}$, respectively.

As before, without loss of generality, it is assumed that each user uses symbols out of the same Q -ary constellation \mathcal{S} . An illustration of the DL NOMA general scenario is given in Figure 5.4 for sake of completeness.

Uplink

The UL case can essentially be formulated as in a dual manner with respect to its DL counterpart, *e.g.* [169,171,178]. However, in this scenario, the K active users transmit simultaneously to a BS. As a consequence, the received BS signal is subsequently obtained as

$$\mathbf{y} = \sum_{k=1}^K \mathbf{H}_{k,UL} \mathbf{f}_k \cdot s_k + \mathbf{n}, \quad (5.13a)$$

$$= \sum_{k=1}^K (\text{diag}(\mathbf{H}_{k,UL}) \circ \mathbf{f}_k) \cdot s_k + \mathbf{n}, \quad (5.13b)$$

$$= \mathbf{H}_{eq,UL} \cdot \mathbf{s} + \mathbf{n}, \quad (5.13c)$$

where the matrix $\mathbf{H}_{eq,UL}$ has been used to compact the *equivalent frame-precoded multiple access channel*.

Explicitly, $\mathbf{H}_{eq,UL}$ is obtained as the grouped column vectors

$$\mathbf{H}_{eq,UL} \triangleq [\text{diag}(\mathbf{H}_{1,UL}) \circ \mathbf{f}_1, \dots, \text{diag}(\mathbf{H}_{K,UL}) \circ \mathbf{f}_K], \quad (5.14)$$

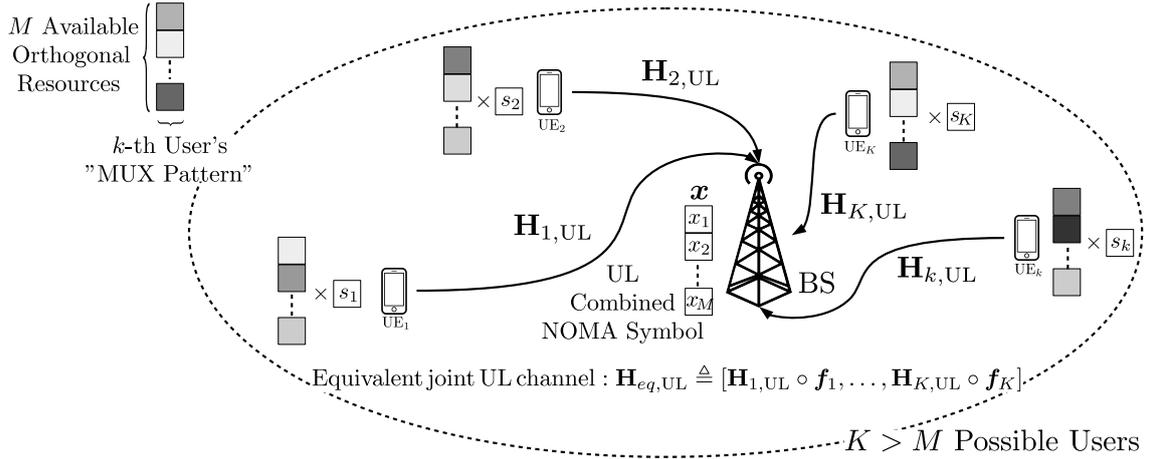


Figure 5.5: An example of an UL NOMA system.

obtained individually by the Hadamard product of each frame vector k and the diagonal entries of the diagonal matrix $\mathbf{H}_{k,\text{UL}}$ denoting the individual UL channel of the k -th user and the BS, where, similarly as before, single antenna BS and UE are considered.

An illustration of the UL use case is thus presented in Figure 5.5.

The DL-UL duality of CD-NOMA under the proposed simple modeling can be now remarked, since for the DL case one can define analogously for each user k an *equivalent frame-precoded DL channel matrix*

$$\begin{aligned} \mathbf{H}_{eq,k,\text{DL}} &\triangleq \mathbf{H}_{k,\text{DL}} \cdot \mathbf{F} \\ &= [\text{diag}(\mathbf{H}_{k,\text{DL}}) \circ \mathbf{f}_1, \dots, \text{diag}(\mathbf{H}_{k,\text{DL}}) \circ \mathbf{f}_K], \end{aligned} \quad (5.15)$$

such that the DL model is, from a pure mathematical viewpoint, a particular realization of the UL one, where essentially all frames precode the same channel $\mathbf{H}_{k,\text{DL}}$.

In the light of all of the above, Question Q1 is hereby considered to have been answered. In summary, the CD-NOMA access methods have been reduced to a compound two-step mapping generating a NOMA signal \mathbf{x} .

5.4 Massively Concurrent Non-orthogonal Multiple Access Transmission

Having previously answered Question Q1, and so, defined the DL (5.12), and respectively, UL (5.13) generic CD-NOMA models, the attention is turned next to Question Q2. With this objective in mind, upon (5.12) or (5.13), it is clear that the performance of the abstractly formulated NOMA system relies primarily on \mathbf{F} , since at the end of the day the symbols \mathbf{s} contain discretized entries of a constellation or of a codebook, respectively. Therefore, towards answering Question Q2, the objective is to find the optimum criteria of effectively designing \mathbf{F} . As a result, an information-theoretic motivated discourse about the choice and design of \mathbf{F} to

jointly benefit both schemes in DL and UL through a scalable design follows next.

5.4.1 Capacity Analysis and Performance Guarantees

The capacity and achievable rates are thus analyzed in order to extract the design conditions of the information multiplexing/spreading matrix \mathbf{F} . In doing so, it is inherently assumed that no CSIT is available. The latter assumption is imposed in order to eliminate the system overhead necessary to gather the CSI. In fact, such a lack of forward channel information is to be expected in the context of URLLC or mMTC where the links are required to be fast and scalable, to have low-latency, and respectively, to be energy-efficient.

The focus is placed therefore on the ergodic analysis of system sum-rates for both the DL and the UL cases. To this end, it is assumed that joint MUD is performed in the joint ML sense where each user is capable of detecting simultaneously all received symbols, not just its own. The multi-dimensional super-symbol \mathbf{s} containing all the information of all the K active users is furthermore considered without loss of generality to be constrained to unit average power such that $\mathbb{E}[\mathbf{s}^H \mathbf{s}] = 1$. In addition, the thermal noise at the receivers is universally assumed to follow the AWGN conventions with a flat spectral representation and an energy density of N_0 . Furthermore, the information-theoretic analysis is considering both continuous and discrete inputs over the wireless memoryless channels in order to highlight both the theoretical achievable rates, but also the practical achievable rates up to receive signal quantization.

Achievable Rates in Downlink without CSIT

Let us first start with the DL analysis and look at the ergodic sum-rates for the continuous-input-continuous-output memoryless channel (CCMC), and respectively, for the discrete-input-continuous-output memoryless channel (DCMC) scenarios.

Continuous Input Continuous Output Memoryless Channel

Considering all the previously formulated assumptions, in the continuous-input signal case \mathbf{s} becomes in fact a random variable drawn out of the maximum entropic distribution with known mean and variance, *i.e.*, the Gaussian distribution. As a consequence, $\mathbf{s} \sim \mathcal{CN}(0, \frac{1}{K} \mathbf{I}_K)$ since in fact \mathbf{s} is complex valued. The variance constraint follows in the light of the unit-power average constraint.

Upon this latter remark and the previously stated assumptions, the ergodic sum-rate capacity across all resources for any given k -th user is

$$C_{k,\text{DL}}^{\text{CCMC}} = \mathbb{E}_{\mathbf{H}_k} \left[\log_2 \det \left(\mathbf{I}_M + \underbrace{\frac{1}{KN_0}}_{\text{SNR}} \mathbf{H}_k \overbrace{\mathbf{F}\mathbf{F}^H}^{\text{frame operator } \mathbf{S}_F} \mathbf{H}_k^H \right) \right]. \quad (5.16)$$

Note that for simplicity of the notation we have implicitly replaced above in (5.16) the cumbersome notation of the channel $\mathbf{H}_{k,\text{DL}}$ as \mathbf{H}_k since the context of DL is well-delimited

under this section. Furthermore, given that \mathbf{F} is a fixed point, parametric system components, the ergodicity of the process is resumed above just over the realizations of the channel \mathbf{H}_k .

Note that the matrix $\mathbf{A} \triangleq \mathbf{I}_M + \frac{1}{KN_0} \mathbf{H}_k \mathbf{F} \mathbf{F}^H \mathbf{H}_k^H$ is Hermitian and positive semi-definite given its general form. As a consequence, it follows that Hadamard's determinant inequality [182, p. 233] may be invoked, *i.e.*

$$\det(\mathbf{A}) \leq \prod_{k=1}^K a_{kk}, \quad (5.17)$$

The upper bound in (5.17) is reached with equality only in the case where \mathbf{A} itself is a diagonal matrix. Reviewing the explicit form of \mathbf{A} from (5.16), since \mathbf{H}_k is diagonal it follows that $\mathbf{S}_{\mathbf{F}}$ should be diagonal as well in order to maximize $C_{k,\text{DL}}^{\text{CCMC}}$ for any user k . The latter condition is equivalent, as seen in Section 3.4, to \mathbf{F} being a *tight frame*. Additionally, considering the ergodicity across all RE, and the fact that CSIT is unavailable, it is optimum from a stochastic perspective for \mathbf{F} to be additionally *equi-normalized*. Note at this point that the initial assumption of no CSIT does not impede further optimizations alongside \mathbf{F} . Given for instance that such information is available, for the DL case these optimizations become a matter of power control via the optional matrix \mathbf{P} similarly to the problem of waterfilling, especially given that the individual frame vectors of \mathbf{F} are of the same norm.

Without loss of generality and from a normalized analysis perspective, let the frame \mathbf{F} be in fact unit-normalized such that $\|\mathbf{f}_k\|_2 = 1, \forall k$. As a result, the optimal frame \mathbf{F} is an UNTF, and in effect its frame operator $\mathbf{S}_{\mathbf{F}}$ satisfies Equation (3.58), thereby simplifying (5.16) to

$$C_{k,\text{DL}}^{\text{CCMC}} = \mathbb{E}_{\mathbf{H}_k} \left[\log_2 \det \left(\mathbf{I}_M + \frac{1}{MN_0} \mathbf{H}_k \mathbf{H}_k^H \right) \right]. \quad (5.18)$$

Moreover, considering the ergodic Rayleigh fading channel, the obtained $C_{k,\text{DL}}^{\text{CCMC}}$ can be computed in closed form upon Telatar's seminal work on the analysis of MIMO communications channels, *i.e.* concretely by [183, Th. 2] as

$$C_{k,\text{DL},\text{Rayl.}}^{\text{CCMC}} = \int_0^\infty \log_2 \left(1 + \frac{1}{MN_0} \lambda \right) \lambda^{K-M} \exp^{-\lambda} \cdot \sum_{m=0}^{M-1} \frac{m! \cdot [L_m^{K-M}(\lambda)]^2}{(m+K-M)!} d\lambda, \quad (5.19)$$

where λ represents a marginalized singular value dimension of the equivalent non-negative definite matrix $\mathbf{H}_k \mathbf{H}_k^H$ and L_m^j are the (m, j) -associated Laguerre polynomials [183].

Let us first remark the fact that given the abstract formulation from Subsection 5.3, the analysis from above yielding Equation (5.16) similarly applies to CD-NOMA schemes, as for example SCMA [178], PDMA [176] and MUSA [171]. To briefly detail, the matrix \mathbf{F} associated to SCMA becomes the factor graph matrix whereas the intricate optimized codebook design is bundled in \mathbf{s} ; for PDMA, the \mathbf{F} is simply the multiplexation pattern of the scheme; in the case of MUSA, the \mathbf{F} contains on its columns the waveform signatures or spreading signatures associated with each individual multiplexed user.

On the other hand, note however that the high-level criteria identified so far as desirable for the optimal design of \mathbf{F} , *i.e.* tightness and unit-normality, leading to the optimum rate expression in (5.18) are soft constraints on the structure of the matrix itself. In comparison, the \mathbf{F} realizations discussed above subsequent to SCMA, PDMA or MUSA impose hard constraints also on the structures of columns and rows of \mathbf{F} to achieve different levels of overloading. Consequently, for such schemes, the frame operator term $\mathbf{S}_{\mathbf{F}}$ in equation (5.16) corresponding may not generally reduce to the form from equation (3.58). This effects in term to a more complicated analysis of their maximum achievable sum-rates, while also implying the *loss of exploited degrees-of-freedom (DoF)*. Following this argument, it is therefore desirable, additionally to the previous constraints, that \mathbf{F} is a fully dense matrix capable of separating its users and proving a low mutual interference (MUI) not based on sparsity or structured constructs, but instead on *frame-theoretic low-coherent overloading*. The latter is not only optimal in terms of spreading the users across all REs and decreasing thus the inter-user correlation, but it also provides the highest degree of redundancy by dense, *massively concurrent overloading* whilst exploiting all DoF.

Given the desired frame properties of *tightness, unit-normality, incoherence*, and respectively, *massive concurrency* for an optimized design of complex valued \mathbf{F} matrices, we refer hereafter to the NOMA systems characterized by such a frame \mathbf{F} as MC-NOMA.

Factually, given Hadamard's inequality (5.17), it can be inferred that the CCMC capacity of (5.18) is in fact the optimum ergodic sum-rate across any DL NOMA single antenna based scheme without CSIT. Concretely, MC-NOMA has the theoretical potential to outperform over DL all other NOMA schemes in terms of maximum achievable ergodic sum-rate. Moreover, this gain over the state of the art can be achieved practically provided that two high-level conditions are simultaneously met:

- i)* $\forall (M, K)$, a frame \mathbf{F} can be designed with the properties detailed above;
- ii)* an optimal joint ML MUD decoder can be implemented.

Before going into the details of how the above two constraints can be met, let us firstly further analyze the practically feasible DCMC capacity in DL as well, and then secondly consider the same rate expressions for the UL scenario.

Discrete Input Continuous Output Memoryless Channel

We compute in this subsection the explicit DL ergodic sum-rate for any given user k in the DCMC case given the same previous general assumptions. Applying the Shannon-Hartley Theorem [Ch. 4] [184], the DL rate is given by the maximum of mutual information as

$$R_{k,\text{DL}}^{\text{DCMC}} \triangleq \max_{p(\mathbf{s}) \in \mathcal{S}^K} I(\mathbf{s}; \mathbf{y}_k) \quad (5.20a)$$

$$= \max_{p(\mathbf{s}) \in \mathcal{S}^K} (B - H(\mathbf{s}|\mathbf{y}_k)), \quad (5.20b)$$

where once for a tractable analysis it is assumed that the symbol of each k -th users in \mathbf{s} is

sampled out of the same Q -ary constellation \mathcal{S} , such that $\mathbf{s} \in \mathcal{S}^K$ and B is given by Equation (5.6).

Expanding the entropy term in (5.20), it is yielded that

$$R_{k,\text{DL}}^{\text{DCMC}} = B - \sum_{i=1}^{2^B} \int_{\mathbf{y}_k} p(\mathbf{s}_i, \mathbf{y}_k) \log_2 \left(\frac{1}{p(\mathbf{s}_i | \mathbf{y}_k)} \right) d\mathbf{y}_k \quad (5.21a)$$

$$= B - \sum_{i=1}^{2^B} \int_{\mathbf{y}_k} p(\mathbf{s}_i) p(\mathbf{y}_k | \mathbf{s}_i) \cdot \log_2 \left(\frac{p(\mathbf{y}_k)}{p(\mathbf{s}_i) p(\mathbf{y}_k | \mathbf{s}_i)} \right) d\mathbf{y}_k. \quad (5.21b)$$

Furthermore, the mutual information of (5.20) is known to be maximized under the maximum entropic discrete uniform input distribution [185], *i.e.*, for

$$p(\mathbf{s}) = \frac{1}{2^B}. \quad (5.22)$$

Now, considering the abstract noisy DL NOMA communication model from (5.12), it is obtained that

$$p(\mathbf{y}_k | \mathbf{s}_i) = p(\mathbf{H}_k \mathbf{F} \mathbf{s}_i + \mathbf{n}_k | \mathbf{H}_k \mathbf{F} \mathbf{s}_i) \quad (5.23a)$$

$$= (\pi N_0)^{-M} \exp \left(- \frac{\overbrace{\|\mathbf{y}_k - \mathbf{H}_k \mathbf{F} \mathbf{s}_i\|_2^2}^{\text{noise } \mathbf{n}_k}}{N_0} \right) \quad (5.23b)$$

$$= (\pi N_0)^{-M} \exp \left(- \frac{\|\mathbf{n}_k\|_2^2}{N_0} \right) = p(\mathbf{n}_k). \quad (5.23c)$$

Given therefore (5.23), it is trivial to remark that the conditional distribution of the output given the input over the fixed channel realization \mathbf{H}_k linearly precoded by the combining matrix \mathbf{F} is in fact reduced to the implicit noise distribution. In light of this fact, and by use of Bayes' rule followed by marginalization, the PDF of \mathbf{y} is calculated from (5.23) as

$$p(\mathbf{y}_k) = \sum_{\ell=1}^{2^B} p(\mathbf{s}_\ell) p(\mathbf{y}_k | \mathbf{s}_\ell) \quad (5.24a)$$

$$= \frac{1}{2^B} \sum_{\ell=1}^{2^B} \frac{1}{(\pi N_0)^M} \exp \left(- \frac{\|\mathbf{y}_k - \mathbf{H}_k \mathbf{F} \mathbf{s}_\ell\|_2^2}{N_0} \right) \quad (5.24b)$$

$$= \frac{1}{2^B} \sum_{\ell=1}^{2^B} p(\mathbf{H}_k \mathbf{F} \mathbf{s}_\ell + \mathbf{n}_k | \mathbf{H}_k \mathbf{F} \mathbf{s}_\ell). \quad (5.24c)$$

Applying (5.22) - (5.24) to (5.21), the explicit solution for the instantaneous rate in DL for

any k -th user is given by

$$R_{k,\text{DL}}^{\text{DCMC}} = B - \frac{1}{2^B} \sum_{i=1}^{2^B} \int_{\mathbf{n}_k} p(\mathbf{n}_k) \cdot \log_2 \left[\frac{\sum_{\ell=1}^{2^B} p(\mathbf{H}_k \mathbf{F} \mathbf{s}_\ell + \mathbf{n}_k | \mathbf{H}_k \mathbf{F} \mathbf{s}_i)}{p(\mathbf{n}_k)} \right] d\mathbf{y}_k \quad (5.25a)$$

$$= B - \frac{1}{2^B} \sum_{i=1}^{2^B} \mathbb{E}_{\mathbf{n}_k} \left[\log_2 \sum_{\ell=1}^{2^B} \exp(\eta_{k,\text{DL}}(\ell, i)) \right], \quad (5.25b)$$

where for convenience of the notation $\eta_{k,\text{DL}}(\ell, i), \forall \ell, i$ was implicitly defined as

$$\eta_{k,\text{DL}}(\ell, i) \triangleq \frac{\|\mathbf{n}_k\|^2 - \|\mathbf{H}_k \mathbf{F}(\mathbf{s}_\ell - \mathbf{s}_i) + \mathbf{n}_k\|^2}{N_0}. \quad (5.26)$$

Finally, to obtain the ergodic sum-rate for the DCMC case, the expectation over (5.25) is taken with respect to the random channel realizations of \mathbf{H}_k , and respectively, to the random noise realizations \mathbf{n}_k , yielding the final form

$$C_{k,\text{DL}}^{\text{DCMC}} = \mathbb{E}_{\mathbf{H}_k, \mathbf{n}_k} [R_{k,\text{DL}}^{\text{DCMC}}] \quad (5.27a)$$

$$= B - \frac{1}{2^B} \sum_{i=1}^{2^B} \mathbb{E}_{\mathbf{H}_k, \mathbf{n}_k} \left[\log_2 \sum_{\ell=1}^{2^B} \exp(\eta_{\text{DL}}(\ell, i)) \right]. \quad (5.27b)$$

Under the generically proposed model of (5.12), it can be observed given Equations (5.18) and (5.27) that actually the MC-NOMA systems are equivalent in information-theoretic terms to a MIMO overcomplete transmission scheme with increased Tx diversity in comparison to the Rx diversity [2, Chap. 10], [184]. In more details, the MIMO capacity derived by Telatar in [183] for Gaussian inputs and the CCMC rate for the k -th user derived for MC-NOMA in Equation (5.18) are similar, whereas the DCMC Equation (5.27) is analogous in its final form to the achievable ergodic sum-rate of MIMO systems with discrete inputs [186, 187]. Furthermore, as the model (5.12) is applicable to various CD-NOMA systems, the results and key remarks from above remain true for these as well. The main difference is however that the specialized form of \mathbf{F} as a MC-NOMA frame is not more valid. Explicitly therefore, the DL CCMC rate is “downgraded” to the generic Equation (5.16), while the DL DCMC remains expressed through the same Equation (5.27) with particularly defined \mathbf{F} matrices given the considered schemes.

To end this treatment of DL scenario, it is worth underlining the fact that the rate expressions of Equations (5.18) and (5.27) are the ergodic sum-rates of any k -th user across all M REs. As a conclusion, considering that all K users have similar SNRs, it follows therefore that in fact the DL system total achievable sum-rate is on average given for the CCMC and DCMC cases, respectively by

$$C_{\text{DL}}^{\text{CCMC}} = K \cdot C_{k,\text{DL}}^{\text{CCMC}}, \quad (5.28a)$$

$$C_{\text{DL}}^{\text{DCMC}} = K \cdot C_{k,\text{DL}}^{\text{DCMC}}. \quad (5.28b)$$

Achievable Uplink Sum-Rate without CSIT

In the sequel, we analyze the counterpart UL ergodic sum-rate system capacity under the proposed model (5.13). The signal processing chain assumes therefore that each k -th user utilizes an unique frame vector \mathbf{f}_k out of \mathbf{F} to spread non-orthogonally its own signal over all the available M REs.

Continuous Input Continuous Output Memoryless Channel

Recall the DL-UL duality outlined in Section 5.3 via Equations (5.15) and (5.14) respectively. Considering this DL-UL connection and leveraging the result of Equation (5.15), it follows directly that the ergodic sum-rate of the UL system under the assumption of the joint ML MUD scheme at the BS is obtained as

$$C_{\text{UL}}^{\text{CCMC}} = \mathbb{E}_{\mathbf{H}_{eq,\text{UL}}} \left[\log_2 \det \left(\mathbf{I}_M + \frac{1}{KN_0} \mathbf{H}_{eq,\text{UL}} \mathbf{H}_{eq,\text{UL}}^H \right) \right]. \quad (5.29)$$

As previously defined in (5.14), the $\mathbf{H}_{eq,\text{UL}} \in \mathbb{C}^{M \times K}$ is in fact an equivalent UL channel precoded by the matrix \mathbf{F} subsequent to the NOMA system considered. In effect, $\mathbf{H}_{eq,\text{UL}}$ is an equivalent \mathbf{F} as a fixed embedded component.

Unfortunately, in opposition to the Kronecker product which was earlier shown in Chapter 4 to preserve tightness, unit-norm and incoherence properties, the same cannot be said about the Hadamard product, especially in the context of the definition of the equivalent UL channel $\mathbf{H}_{eq,\text{UL}}$ as formulated by (5.14). Given this, more information about the structure or general properties of \mathbf{F} necessary or desirable to maximize the system spectral efficiency and communication rate aside from the ones already stated earlier for MC-NOMA cannot be further offered. This is the case since as there are simply no mathematical guarantees that attributes of \mathbf{F} would be propagated upon $\mathbf{H}_{eq,\text{UL}}$, especially given the assumption of no CSIT.

Nonetheless, it still makes sense from both system engineering and information-theoretic perspectives to require \mathbf{F} to be in essence the same between DL and UL, and so, force the frame to fulfill the MC-NOMA conditions from earlier on. On one hand, allowing users to have knowledge of just one matrix \mathbf{F} containing both the DL spreading codes, but also the UL signature waveforms would minimize the system overhead and network distribution effort coupled with disseminating the information about \mathbf{F} to all the active participants. Whereas on the other hand, the tightness, unit-normality, incoherence and massive concurrency (matrix density) in fact still preserve a lot of the salient signal information. Concretely, tightness and unit-normality assure the fact that any symbol $s_k \in \mathcal{S}$ is treated similarly from a power perspective, while similarly incoherence and dense realizations of \mathbf{F} spread the inherent signal information as best as possible from an information-theoretic perspective as they together generate as uncorrelated dense structures as possible. These qualitative remarks, albeit soft, are intuitive and applicable both to DL, as mathematically proven earlier, but also to UL from a pure entropic driven argumentation. Lastly, these choices can also be motivated under the lack of CSIT, as without any prior information spreading the signals and users as incoherently as

possible is in fact most desirable for the MUD operation such that potential symbol collisions are better avoided.

Discrete Input Continuous Output Memoryless Channel

To provide a complete circle, let us finally consider also the DCMC system sum-rate analysis in the UL case. Once more, applying the DL-UL duality reference by Equations (5.14) and (5.15), the ergodic achievable rate for CD-NOMA UL under the proposed generic model (5.13) can be derived in complete analogy to its DL counterpart (5.28b), yielding similarly

$$C_{\text{UL}}^{\text{DCMC}} = B - \frac{1}{2^B} \sum_{i=1}^{2^B} \mathbb{E}_{\mathbf{H}_{eq, \text{UL}}, \mathbf{n}} \left[\log_2 \sum_{\ell=1}^{2^B} \exp(\eta_{\text{UL}}(\ell, i)) \right], \quad (5.30)$$

where implicitly $\eta_{\text{UL}}(\ell, i)$ was yet again defined for convenience $\forall \ell, i$ as

$$\eta_{\text{UL}}(\ell, i) \triangleq \frac{\|\mathbf{n}\|^2 - \|\mathbf{H}_{eq, \text{UL}}(\mathbf{s}_\ell - \mathbf{s}_i) + \mathbf{n}\|^2}{N_0}. \quad (5.31)$$

The difference to the counterpart DL DCMC is yet again marked as before by the fact that in UL the channel matrix and the precoding frame \mathbf{F} are not simply separable as a matrix product anymore in the light of (5.14). However, as throughout this section, the result in (5.30) is valuable in its generality as it allows a common treaty of different NOMA schemes in a head to head comparison given the common UL generic model (5.13).

From the MC-NOMA perspective, as already argued above for the UL CCMC case, the same frame-theoretic prerogatives identified upon the DL MC-NOMA case are applicable universally here too, especially given imperfect or no CSIT. As a consequence, in the light of the results above and their accompanying argumentation, it thus follows that to implement the MC-NOMA scheme one needs to generate the frame \mathbf{F} as a *dense incoherent UNTF*. Additionally, to ensure scalability, such methods to design \mathbf{F} according to these criteria should be able to do so irrespective of the dimensionality constraints, *i.e.* for any $K > M$.

5.4.2 Transmitter Design via Incoherent Unit-norm Tight Frames

The frame-theoretic problem of designing *dense, tight, incoherent* and *equi-normalized* general frames has been actually already solved for complex-valued realizations in Chapter 4, Section 4.3.3 by means of Algorithm 3. The cores of this proposed algorithm are the CSIDCO optimization (4.50) leading to highly incoherent UNFs, and respectively, the polar decomposition (3.79) implementing Theorem 3.5.1 that solves the tight frame representation closest in terms of Frobenius norm to any given frame. As seen previously in Chapter 4, this method decorrelates the frame vectors up to UNTFs with very low mutual coherence such that the distributions of the non-diagonal absolute entries of the Gramian matrix are actually narrowly bell-shaped and closely centered around the WB. Since in terms of MC-NOMA schemes these values control in fact the inherent MUI upon non-orthogonal overloading, and should be there-

fore as low as possible to aid MUD, it follows that Algorithm 3 is directly applicable to the design of the MC-NOMA frames, especially given that all the requirements of the associated frame design problem of the scheme are jointly met.

Let us briefly discuss the matter of inter-user interference under the formulation of MC-NOMA and frame incoherence, respectively. Obviously, for a better detection, the correlation between the individual frame vectors (acting as waveform signatures or similarly spreading sequences for each individual active user) is preferred to be as low as possible to minimize inter-symbols interference and allow lower-complexities MUDs receivers. To this end, imported from CDMA literature, a lot of CD-NOMA schemes rely in fact on “hidden” frame-theoretic constructs that address frame vectors incoherence in one way or another.

For instance Ericsson’s Welch-bound spreading multiple access (WSMA) [156, 188], RSMA, PDMA, and MUSA all respectively use the sum of squared correlations (SSC) principle to minimize the inter-user correlation by minimizing the Frobenius norm of the signaling Gram matrix instead. As seen in Section 3.5 earlier, this is in fact given in frame-theoretic sense by the FP, with the global minimizer obtained by any tight frame. Of course, such a design criterion is still desirable, as when the minimization reaches the global optimum tight frames are realized such that DL sum-rate (5.16) is simultaneously also maximized for instance. However, the latter ideal case is not in general achievable as these schemes impose additional constraints on the contents of their corresponding frames such that these are either quantized, RSMA [163], WSMA [188], MUSA [171], or respectively sparse [175], for example. As a result, the SSC optimization achieves the minimization of the *average coherence* in the light of Equations (3.47) and (3.48). Whereas the proposed complex-valued MC-NOMA scheme, leverages all the available DoFs and under the realization of the proposed Algorithm 3 manages simultaneously to *minimize the worst-case coherence* under the CSIDCO framework [82], but also to *reduce the overall average coherence* via tightening. In this way both the optimality of coherence and of the achievable sum-rate are obtained universally for generic dimensionality constraints (M, K) . As a side note, recall that the existence of the actual globally optimum frames capable of satisfying both these conditions, *i.e.*, ETFs, is momentarily conjectured for generic (M, K) and restricted constructions exist only for particular dimensions, as discussed in detail in Chapter 3. In this sense, the relaxed optimum solutions, suboptimal in comparison to ETFs if these were to exist, are the MC-NOMA frames proposed and practically achievable for any (M, K) by Algorithm 3.

Picking up on the practicality and implemetability remark, notice that given the formulation (4.50), the CSIDCO problem is reduced to a convex-norm program, *i.e.*, the ∞ -norm, constrained by a quadratic inequality over the global search domain \mathbb{C}^M . This problem is already formulated by the authors of [82] following the disciplined convex programming conventions [189], and as a result is easily solvable via semidefinite convex programming via CVX [150] available in high-level scientific programming languages as MATLAB[®] [190] for instance. The authors of CVX lately also developed an automatic code generator for the low-level C programming language [191] implementations following the disciplined convex programming paradigm, but this generator is unfortunately limited to quadratic programs only [192]. To provide support to further implementations closer to the low-level hardware domain, we alleviated this

intractability of the original formulation of CSIDCO (4.50), and reformulated the problem into a simple convex quadratic program with box and quadratic constraints as outlined by (5.32) below.

$$\begin{aligned} & \underset{\mathbf{x} \triangleq [\mathbf{f}_{k,\mathcal{R}}; \mathbf{t}_{\mathcal{R}}; \mathbf{t}_{\mathcal{I}}] \in \mathbb{R}^{2M+2}}{\text{minimize}} && \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}, \end{aligned} \quad (5.32a)$$

$$\text{subject to} \quad \mathbf{A}_{\mathcal{R},1} \mathbf{x} \leq 0, \quad \mathbf{A}_{\mathcal{R},2} \mathbf{x} \leq 0, \quad (5.32b)$$

$$\mathbf{A}_{\mathcal{I},1} \mathbf{x} \leq 0, \quad \mathbf{A}_{\mathcal{I},2} \mathbf{x} \leq 0, \quad (5.32c)$$

$$\mathbf{x}^T \mathbf{B} \mathbf{x} - 2\mathbf{b}^T \mathbf{x} + 1 - T_k^2 \leq 0, \quad (5.32d)$$

where

$$\boldsymbol{\Sigma} \triangleq \begin{bmatrix} \mathbf{0}_{2M} & \mathbf{0}_{2M \times 2} \\ \mathbf{0}_{2 \times 2M} & \mathbf{I}_2 \end{bmatrix} \in \mathbb{R}^{(2M+2) \times (2M+2)}, \quad (5.32e)$$

$$\mathbf{A}_{\mathcal{R},1} \triangleq \begin{bmatrix} \tilde{\mathbf{F}}_{k,\mathcal{R}}^T & -\mathbf{1}_{(K-1) \times 1} & \mathbf{0}_{(K-1) \times 1} \end{bmatrix} \in \mathbb{R}^{(K-1) \times (2M+2)}, \quad (5.32f)$$

$$\mathbf{A}_{\mathcal{R},2} \triangleq \begin{bmatrix} -\tilde{\mathbf{F}}_{k,\mathcal{R}}^T & -\mathbf{1}_{(K-1) \times 1} & \mathbf{0}_{(K-1) \times 1} \end{bmatrix} \in \mathbb{R}^{(K-1) \times (2M+2)}, \quad (5.32g)$$

$$\mathbf{A}_{\mathcal{I},1} \triangleq \begin{bmatrix} \tilde{\mathbf{F}}_{k,\mathcal{R}}^T \mathbf{D}_M & \mathbf{0}_{(K-1) \times 1} & -\mathbf{1}_{(K-1) \times 1} \end{bmatrix} \in \mathbb{R}^{(K-1) \times (2M+2)}, \quad (5.32h)$$

$$\mathbf{A}_{\mathcal{I},2} \triangleq \begin{bmatrix} -\tilde{\mathbf{F}}_{k,\mathcal{R}}^T \mathbf{D}_M & \mathbf{0}_{(K-1) \times 1} & -\mathbf{1}_{(K-1) \times 1} \end{bmatrix} \in \mathbb{R}^{(K-1) \times (2M+2)}, \quad (5.32i)$$

$$\mathbf{D}_M \triangleq \mathbf{I}_M \otimes \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{(2M+2) \times (2M+2)} \quad (5.32j)$$

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{I}_{2M} & \mathbf{0}_{2M \times 2} \\ \mathbf{0}_{2 \times 2M} & \mathbf{0}_{2 \times 2} \end{bmatrix} \in \mathbb{R}^{(2M+2) \times (2M+2)}, \quad (5.32k)$$

$$\mathbf{b}^T \triangleq \begin{bmatrix} \tilde{\mathbf{f}}_{k,\mathcal{R}}^T & 0 & 0 \end{bmatrix} \in \mathbb{R}^{2M+2}. \quad (5.32l)$$

The alternative formulation of CSIDCO (4.50) proposed in (5.32) is referred to hereafter as the quadratic CSIDCO (QCSIDCO). Firstly, note that the quadratic program of QCSIDCO can easily be implemented in its given form as a disciplined convex program by means of CVX [150], but also using conventional interior point methods and search algorithms existent in generic convex optimization toolboxes for low-level programming languages.

Secondly, let us detail a bit more the core ideas of the reformulation of CSIDCO to QCSIDCO, (5.32). The proposed method provides the same input-output behavior as the CSIDCO by transforming the complex-domain associated problem of (4.50) to the slightly relaxed box-constrained real-valued problem (5.32). To this end, the complex values of the frame vectors have been expressed in an interleaved manner such that every complex-valued number $z \in \mathbb{C}$

became a column vector of two real values, *i.e.*,

$$z = z_{\mathcal{R}} + jz_{\mathcal{I}} \in \mathbb{C} \mapsto \begin{bmatrix} z_{\mathcal{R}} \\ z_{\mathcal{I}} \end{bmatrix} \in \mathbb{R}^2. \quad (5.33)$$

Effectively the size of each frame vector doubled as the frame $\tilde{\mathbf{F}}$ considered in CSIDCO got transformed to its real-valued equivalent $\tilde{\mathbf{F}}_{\mathcal{R}} \in \mathbb{R}^{2M \times K}$. In analogy to the CSIDCO notation, the frame obtained by pruning the k -th frame is hereby denoted by $\tilde{\mathbf{F}}_{k,\mathcal{R}}$. Thus, the task to decorrelate the K frame vectors became similarly to the CSIDCO idea to iteratively and sequentially solve the CSIDCO problem under the equivalent interlaced real-valued representation of the original complex frame vectors. To preserve the equivalence to the original problem, the complex-domain dot product was rewritten in terms of the proposed interlaced real representation. This was achieved essentially through a straightforward algebraic expansion of products of type

$$z_1 z_2^* = (z_{1,\mathcal{R}} + jz_{1,\mathcal{I}})(z_{2,\mathcal{R}} - jz_{2,\mathcal{I}}) \quad (5.34a)$$

$$= \underbrace{z_{1,\mathcal{R}}z_{2,\mathcal{R}} + z_{1,\mathcal{I}}z_{2,\mathcal{I}}}_{\text{real part } \Re[z_1 z_2^*]} + j \underbrace{(z_{2,\mathcal{R}}z_{1,\mathcal{I}} - z_{1,\mathcal{R}}z_{2,\mathcal{I}})}_{\text{imaginary part } \Im[z_1 z_2^*]} \quad (5.34b)$$

into the equivalent

$$\Re[z_1 z_2^*] = \left(\begin{bmatrix} z_{1,\mathcal{R}} \\ z_{1,\mathcal{I}} \end{bmatrix}^T \cdot \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{Z}_{\mathcal{R}}} \right) \cdot \begin{bmatrix} z_{2,\mathcal{R}} \\ z_{2,\mathcal{I}} \end{bmatrix}, \quad (5.35)$$

and respectively,

$$\Im[z_1 z_2^*] = \left(\begin{bmatrix} z_{1,\mathcal{R}} \\ z_{1,\mathcal{I}} \end{bmatrix}^T \cdot \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{Z}_{\mathcal{I}}} \right) \cdot \begin{bmatrix} z_{2,\mathcal{R}} \\ z_{2,\mathcal{I}} \end{bmatrix}. \quad (5.36)$$

Since the dot product over the complex domain is in fact a summation of the terms in (5.34), it follows that its rewriting upon the chosen real-valued representation (5.33) is given by a sum of real terms (5.35) and (5.36), respectively. Furthermore, the matrices $\mathbf{Z}_{\mathcal{R}}$, $\mathbf{Z}_{\mathcal{I}}$ are block-wise diagonally replicated up to the length of the individual left-hand-side complex-valued row vectors. This is for instance how matrix \mathbf{D}_M in (5.32j) was in fact obtained. In comparison, its counterpart yielding the real components of products of type (5.34) is just the simple identity matrix, and hence, no multiplications are required.

Another key transform from CSIDCO to QCSIDCO is given by the box constraints on the real and imaginary components of the inner products between the frame vector under optimization and all the other frame vectors, originally denoted under CSIDCO as $\tilde{\mathbf{F}}_k^H \mathbf{f}_k$. These

bounding boxes are applied via the slack variables $t_{\mathcal{R}}$ and $t_{\mathcal{I}}$ such that

$$\left| \Re \left[\tilde{\mathbf{F}}_k^H \mathbf{f}_k \right] \right| \leq t_{\mathcal{R}} \quad (5.37)$$

$$\left| \Im \left[\tilde{\mathbf{F}}_k^H \mathbf{f}_k \right] \right| \leq t_{\mathcal{I}}. \quad (5.38)$$

Furthermore, since the objective is still to decorrelate the current vector undergoing optimization \mathbf{f}_k relative to the others in the frame, these slack variables provide in fact a close upper bound to the original objective of CSIDCO, as

$$\|\tilde{\mathbf{F}}_k \mathbf{f}_k\|_{\infty} \leq \sqrt{t_{\mathcal{R}}^2 + t_{\mathcal{I}}^2} \quad (5.39)$$

trivially follows in the light of (5.37) and (5.38).

Considering this slight relaxation of the objective, the QCSIDCO program (5.32) follows, where the only constraint of CSIDCO, *i.e.*, (4.50b), was carried over with T_k determined as before, whereas calculated simply by applying the proposed complex-to-real representation on the frame vectors, together with the subsequent inner product equivalent transforms discussed above.

In summary, the proposed QCSIDCO is a real-valued slightly relaxed, but equivalently constrained quadratic program implementation of the complex-valued CSIDCO (4.50), [82]. To this end, the interchangeable usage of QCSIDCO with CSIDCO in Step 4 of Algorithm 3, is resumed to the following sequence.

1. Firstly, the same complex input frame $\tilde{\mathbf{F}}$ of CSIDCO is transformed into its equivalent real representation given (5.33) as

$$\tilde{\mathbf{F}}_{(0)} \in \mathbb{C}^{M \times K} \xrightarrow{(5.33)} \tilde{\mathbf{F}}_{(0),\mathcal{R}} \in \mathbb{R}^{2M \times K}. \quad (5.40)$$

2. Then the problem (5.32) is solved iteratively and sequentially for each real-transformed k -th frame vector of the frame $\tilde{\mathbf{F}}_{(i)}$ as long as $0 \leq i < I$ a number of maximum iterations.
3. The final real-valued solution $\mathbf{F}_{(I),\mathcal{R}}$ is then reconverted to the complex domain by reverting the mapping (5.33) as

$$\mathbf{F}_* = (\mathbf{I}_M \otimes [1 \quad 1j]) \cdot \mathbf{F}_{(I),\mathcal{R}} \quad (5.41)$$

Upon inequality (5.39) one may argue that the QCSIDCO degradation in yielding low-coherent frames given its usage in Algorithm (3) relative to the original CSIDCO may be significant. However this is not at all the case, and this is so primarily to the slack variables $t_{\mathcal{R}}$ and $t_{\mathcal{I}}$ in (5.32). These are in fact minimizing the mutual coherence of QCSIDCO by means of (5.39) and (5.32), but simultaneously also constraining and lowering the absolute values of the real and imaginary parts of the inner products between the optimized frame vector $\mathbf{f}_{k,\mathcal{R}}$ and all the other frame components $\tilde{\mathbf{F}}_{k,\mathcal{R}}$. In the light of the mathematically equivalent algebraic representation between the complex and real domains, *i.e.*, (5.33) and (5.41), it follows that the upper bound in (5.39) is actually tight.

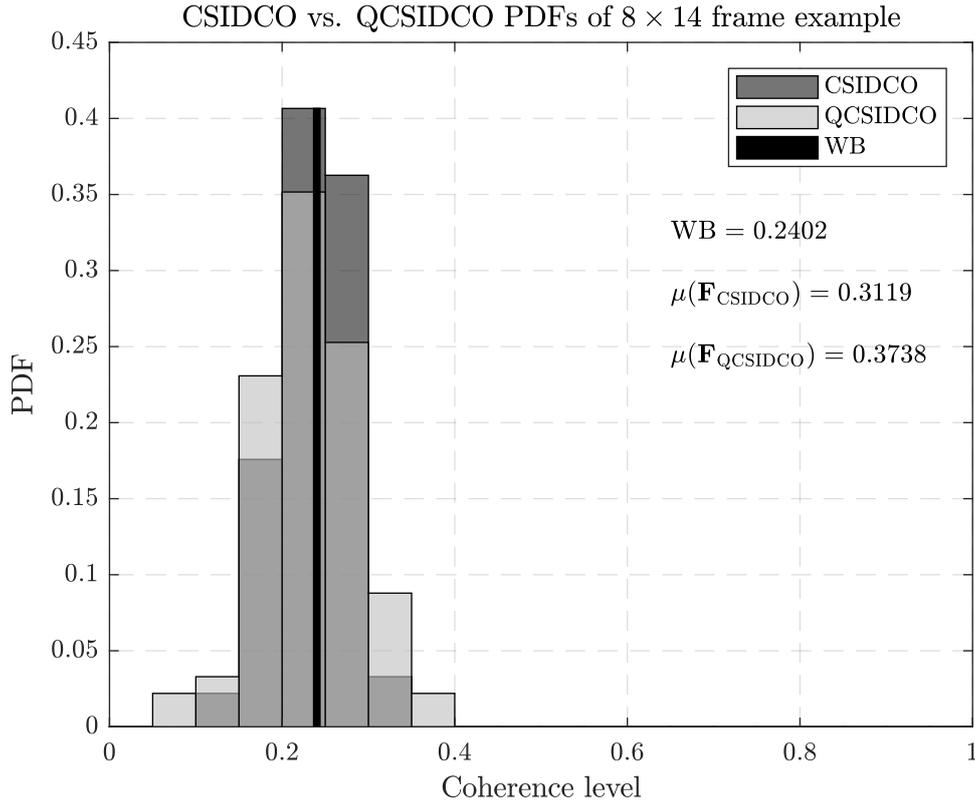


Figure 5.6: Comparison of the coherence PDF obtained between running the Algorithm 3 with CSIDCO, $\mathbf{F}_{\text{CSIDCO}}$, and QCSIDCO, $\mathbf{F}_{\text{QCSIDCO}}$, for $I = 10$ given the design of a frame $\mathbf{F} \in \mathbb{C}^{8 \times 14}$. The WB for this dimensions pair is $\text{WB} = 0.2402$ and marked as a bar plot.

To show this numerically as well, Algorithm 3 was implemented both with CSIDCO and QCSIDCO, and was used to generate complex dense, incoherent UNTF frames of sizes $(M, K) = \{(8, 14), (20, 32)\}$. The results of the obtained coherence PDFs for the same initialization frame $\mathbf{F}_{(0)}$ and a maximum of $I = 10$ iterations over each of the two methods are plotted in Figures 5.6 and 5.7, respectively.

As it can be seen, in both cases, the high-level bell-shaped WB-centered coherence distribution of the frames is preserved between the CSIDCO and QCSIDCO. Indeed there is a slight degradation for the QCSIDCO as a consequence of the relaxation (5.39) considered as its objective, but this is not critical, leading to a slight increase in the spread of the PDF's support in comparison to the CSIDCO case. This of course leads to slight increase in mutual coherence, whilst in fact the average coherence is similar to the CSIDCO method. As a result, QCSIDCO is in turn also yielding highly incoherent frames.

Furthermore, aside from the coherence perspective, it is worth to mention that under the Algorithm 3, the properties of frame density, tightness and unit-normality are invariant to the selection between CSIDCO and QCSIDCO. Hereafter, in the light of all of the above, the QCSIDCO is considered instead of CSIDCO in generating MC-NOMA precoding frames without loss of generality, especially given its slightly degraded coherence performance in comparison of CSIDCO.

To gain a better understanding of the MC-NOMA transmitter realized via dense, incoherent

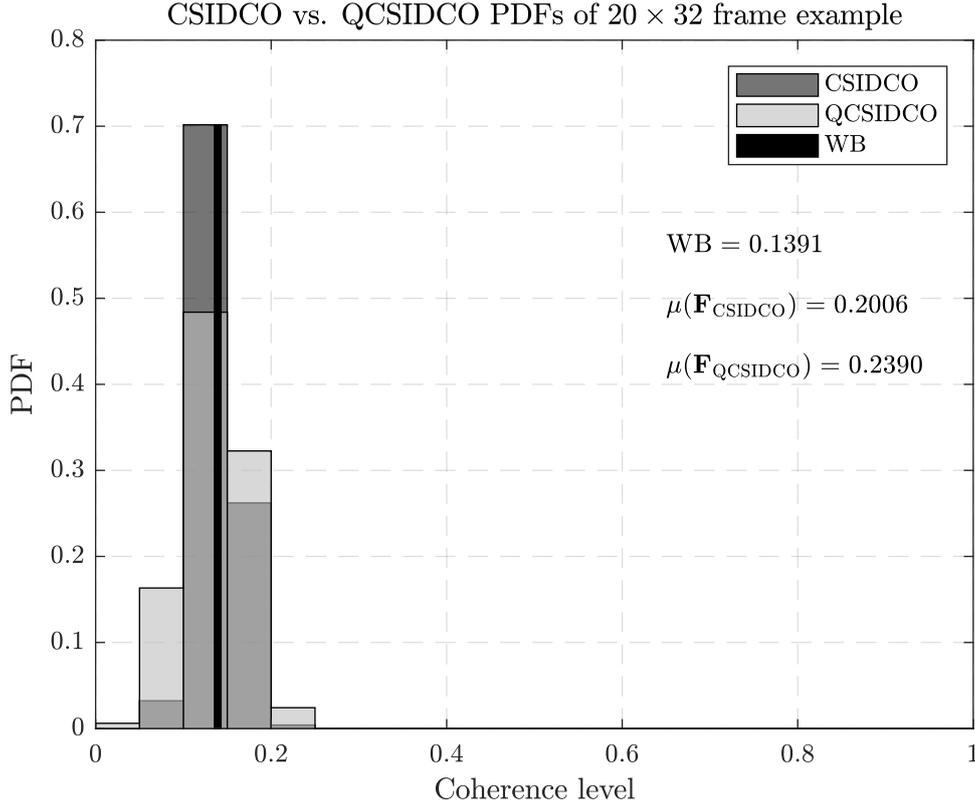


Figure 5.7: Comparison of the coherence PDF obtained between running the Algorithm 3 with CSIDCO, $\mathbf{F}_{\text{CSIDCO}}$, and QCSIDCO, $\mathbf{F}_{\text{QCSIDCO}}$, for $I = 10$ given the design of a frame $\mathbf{F} \in \mathbb{C}^{20 \times 32}$. The WB for this dimensions pair is $\text{WB} = 0.1391$ and marked as a bar plot.

UNTFs, designed as explained above, and to highlight further its embedded advantages, let us have a deeper look into the design concepts associated with the frame construction. Effort is firstly dedicated to outlining the structure of the equivalent DL, $\mathbf{H}_{eq,DL}$, and UL, $\mathbf{H}_{eq,UL}$, channels as perceived by the Rx, given the dense, incoherent UNTF MC-NOMA symbol precoding on the Tx side.

Consider thus a frame \mathbf{F} designed by the Algorithm 3. By means of incoherence and tightness alike, the MC-NOMA Tx frame is by design constrained to have low worst-case coherence, but also low average coherence respectively, being centered around the WB as seen earlier in Figures 5.6 and 5.7. In fact, one may regard the MC-NOMA Tx frames as approximated relaxations of WB optimal ETF discussed in Section 3.5, in the sense that the frames corresponding to MC-NOMA obtain almost a uniform coherence level given their tight distribution around the WB. This remark is clearly illustrated in Figure 5.8, showing an intensity map of the Gram matrix corresponding to a frame $\mathbf{F} \in \mathbb{C}^{8 \times 14}$. This design feature is a key attribute in maintaining the inter-user interference limited as it corresponds to the joint non-orthogonal spreading and differentiation of different users signaling for MC-NOMA.

In light of Figure 5.8, let us iterate and underscore once more the fact that MC-NOMA fulfills a similar design goal as aforementioned state-of-the-art CD-NOMA alternatives (*e.g.* LDS [169], SCMA [178], PDMA [176], RSMA [163]), *i.e.*, overloading the system while controlling and minimizing the MUI, *without structured sparsity*. In effect, MC-NOMA exploits all the available

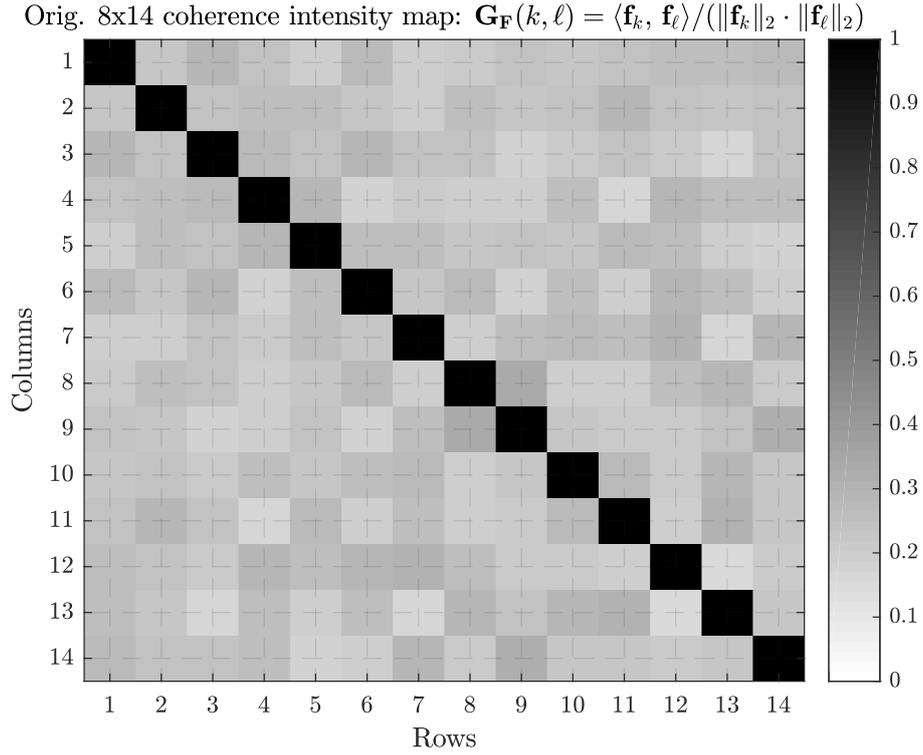


Figure 5.8: Normalized coherence intensity map of a dense, incoherent UNTF $\mathbf{F} \in \mathbb{C}^{8 \times 14}$ designed by Algorithm 3. ©2019 IEEE

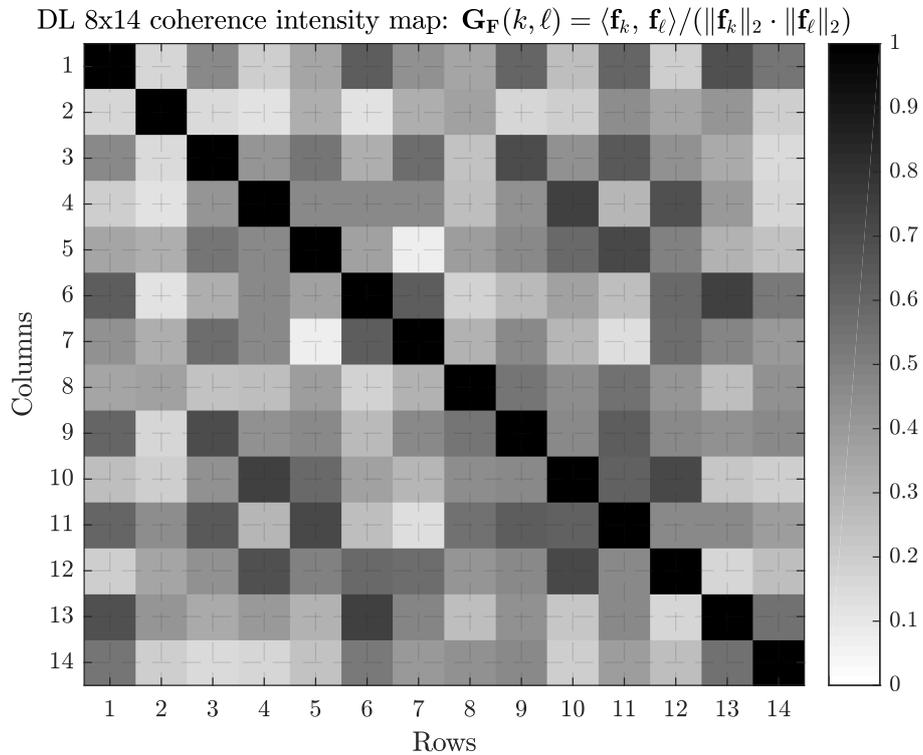


Figure 5.9: Normalized coherence intensity map of an equivalent DL MC-NOMA precoded channel $\mathbf{H}_{eq,DL}$ (5.15). Note that the same frame from Figure 5.8 was used. ©2019 IEEE

resources fully, in opposition to CD-NOMA technologies such as LDS [169], SCMA [178], PDMA [176], and this respectively, once again argues the superiority of MC-NOMA achievable rates illustrated in the next Subsection 5.4.3.

The effect of the unknown wireless channel onto the spreading performed by the optimized incoherent UNTF \mathbf{F} from Figure 5.8, and respectively, the DL-UL channel structural equivalence are next illustrated under Figures 5.9 and 5.10. Concretely, Figure 5.9 displays the intensity map of an equivalent DL channel (5.15), $\mathbf{H}_{eq,k,DL}$, corresponding to the frame \mathbf{F} in Figure 5.8 and one random Rayleigh fading channel realization. On the other hand, Figure 5.10 plots similarly the intensity map of the UL counterpart $\mathbf{H}_{eq,UL}$ as given by (5.14), under K independently distributed Rayleigh fading channels.

Given both Figures 5.9 and 5.10 it is thus seen that indeed the qualitative effects of the DL and UL channels are similar. Generally in fact, an unknown channel (both DL and UL) precoded with MC-NOMA incoherent UNTF leads to an increase in the overall coherence as well as worst-case coherence of the equivalently considered channel in comparison to the frame itself. However, a gradation and inherent diversity of interference power levels is obtained across various users, similar to what is achieved in other CD-NOMA schemes, such as SCMA [179] and PDMA [176] at the expense of larger design complexity and sparsity requirements. The latter are jointly avoided by the MC-NOMA scheme as argued so far.

All in all, Figures 5.8 - 5.10 outline that the incoherent UNTF MC-NOMA frame is able to strongly and non-orthogonally separate user signals, regardless whether the fading channels are

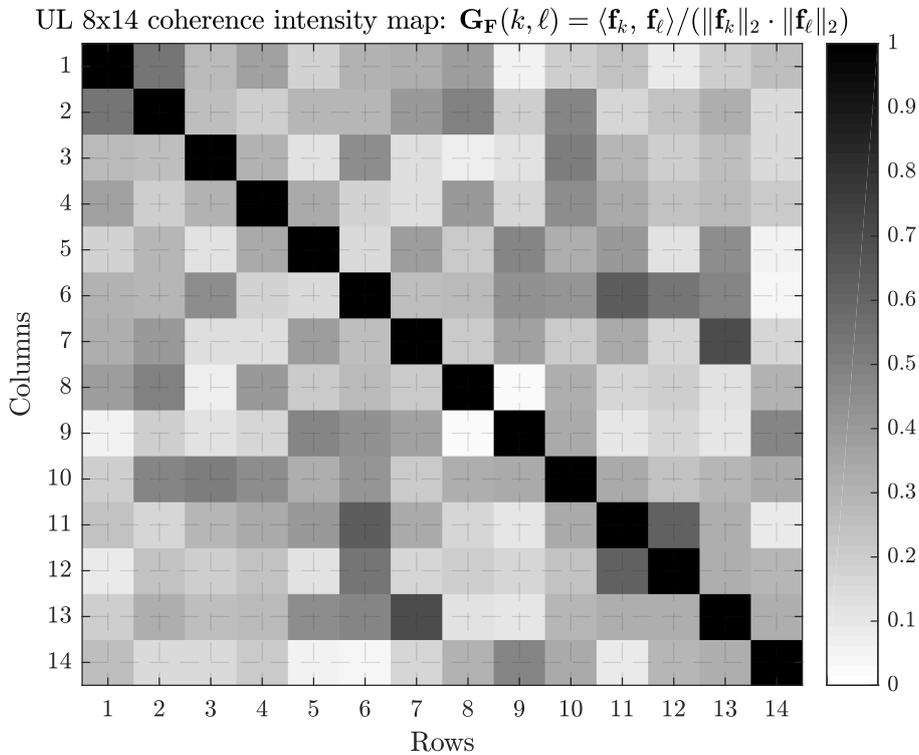


Figure 5.10: Normalized coherence intensity map of an equivalent UL MC-NOMA precoded channel $\mathbf{H}_{eq,UL}$ (5.15). Note that the same frame from Figure 5.8 was used. ©2019 IEEE

diagonal, as is the case for the DL (5.12), or fully rich, as for UL (5.13).

5.4.3 Achievable Sum-rate Performance

Having not only determined the optimal information-theoretic conditions towards the design of a prime CD-NOMA scheme, coined as MC-NOMA in Subsection 5.4.1, but also having established a method to implement it via the frame-theoretic construction of dense, incoherent UNTFs, as seen in the previous subsection, Question Q2 is therefore fully answered.

To add more context and numerical sense to the MC-NOMA answer to Question Q2, a qualitative and quantitative analysis is provided regarding the performance differences of MC-NOMA relative to other CD-NOMA schemes. For this purpose, the achievable DL and UL sum-rates of the proposed MC-NOMA system against those of SCMA [178] and PDMA [175] are therefore compared. These two benchmarks were selected given their well-known and proven performance gains over other CD-NOMA schemes [152, 156] in both idealized but also realistic scenarios.

For a meaningful comparison, the case of $(M, K) = (4, 6)$ was considered as a 150% over-

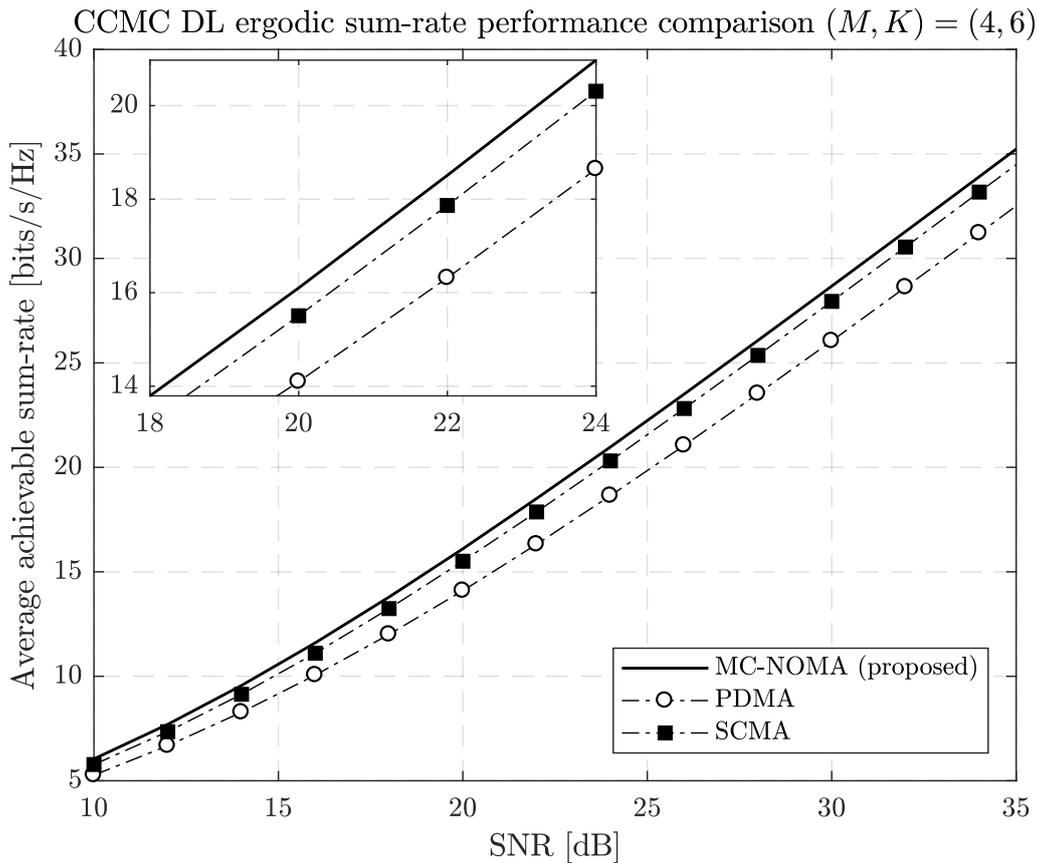


Figure 5.11: Theoretic spectrally normalized CCMC ergodic sum-rate for CD-NOMA DL systems according to Equation (5.16). The proposed MC-NOMA curve is given by (5.18) with the frame \mathbf{F} obtained by Algorithm 3, whereas the SCMA [178] is implemented on the basis of the optimized \mathbf{F}_{SCMA} from (5.42), and respectively, the PDMA [177] is implemented using the optimized allocation pattern $\mathbf{F}_{\text{PDMA,DL}}$ from (5.43). ©2019 IEEE

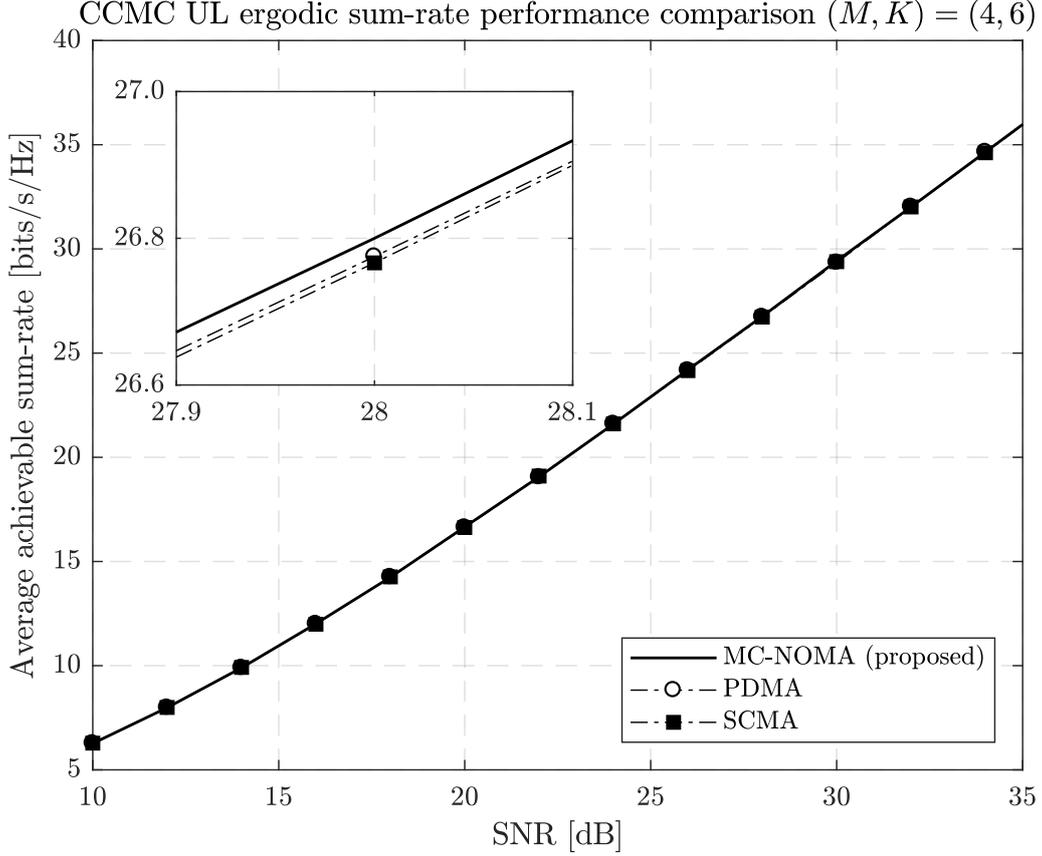


Figure 5.12: Theoretic spectrally normalized CCMC ergodic sum-rate for CD-NOMA UL systems, equation (5.29). The MC-NOMA curve is given by the proposed frame design \mathbf{F} in Algorithm 3, whereas the SCMA [178] is implemented on the basis of the optimized \mathbf{F}_{SCMA} from (5.42), and respectively, the PDMA [177] is implemented using the optimized allocation pattern $\mathbf{F}_{\text{PDMA,UL}}$ from (5.44). ©2019 IEEE

loaded system for both DL and UL. This particular dimensionality was selected upon the already highly optimized existent systems for SCMA [178, 179] and PDMA [176, 177].

Thus, the SCMA method considered the optimized factor graph [178, 179] corresponding to the sparse matrix

$$\mathbf{F}_{\text{SCMA}} \triangleq \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \quad (5.42)$$

On the other hand, the PDMA scheme was simulated based on the optimized pattern multiplexing matrix

$$\mathbf{F}_{\text{PDMA,DL}} \triangleq \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad (5.43)$$

for the DL case, [177], and respectively,

$$\mathbf{F}_{\text{PDMA,UL}} \triangleq \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}, \quad (5.44)$$

for the UL scenario [176].

Actually, the considered combining/allocation patterns from above provide also equivalent comparison conditions across MC-NOMA, SCMA and PDMA, as for the latter two schemes the multiplexing of non-orthogonal active users is implemented over a maximum overlap of 3 users per available RE.

Furthermore, given the SCMA paradigm, its characteristic optimized codebook design principles described throughout [179] were used over all the simulations for an enhanced baseline reference and a fair comparison relative to the state of the art. Since these optimization criteria are in fact developed for discrete constellation points, in the case of CCMC scenario these have been reduced to simple input symbol fixed relative rotations as described in [178, 179]. Lastly, to clarify the CCMC inputs, these values were drawn randomly in this case for all schemes from

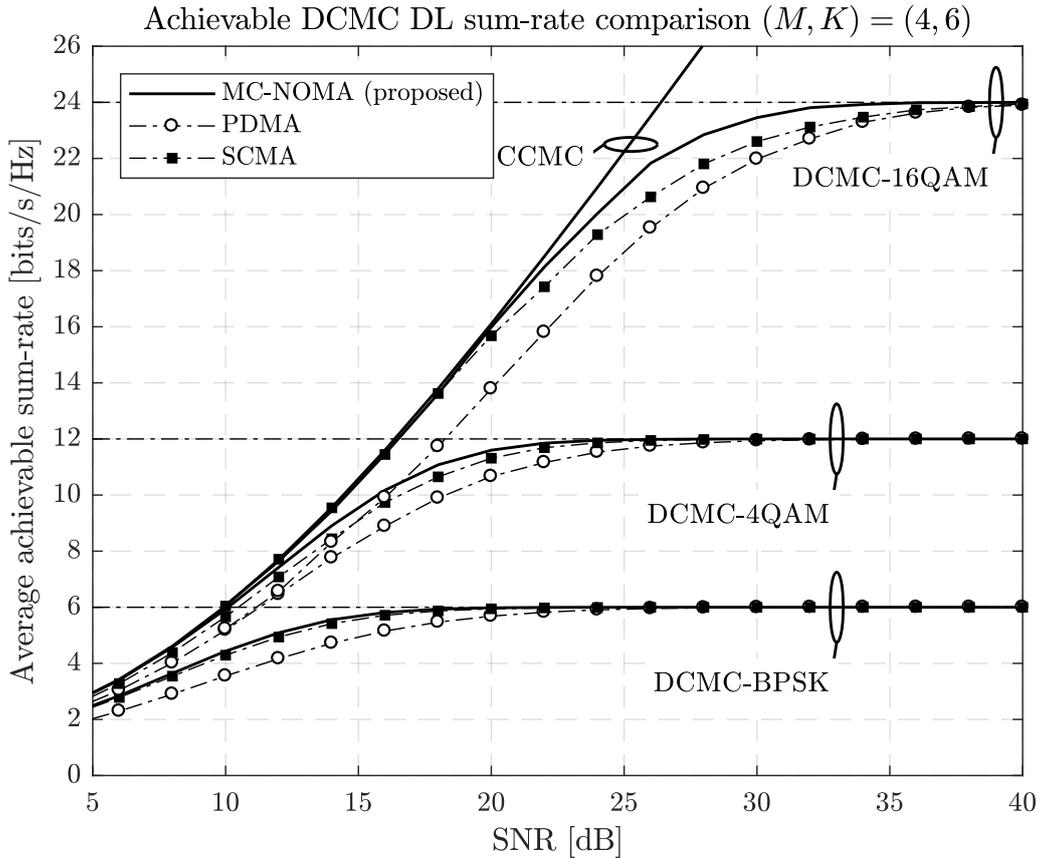


Figure 5.13: DL ergodic sum-rate over the DCMC case across MC-NOMA, SCMA [178], and PDMA [177]. The DL MC-NOMA CCMC sum-rate capacity is plotted as the upper bound of the discrete achievable ergodic sum-rates. ©2019 IEEE

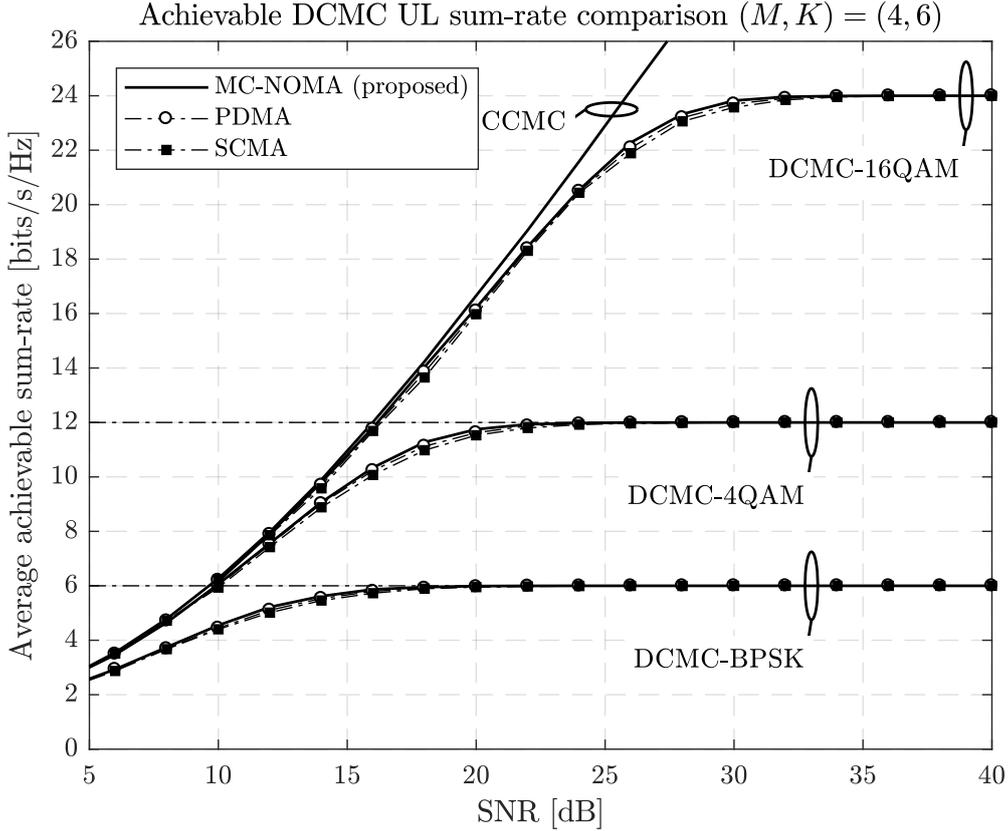


Figure 5.14: UL ergodic sum-rate over the DCMC case across MC-NOMA, SCMA [178], and PDMA [176]. The UL MC-NOMA CCMC sum-rate capacity is plotted as the upper bound of the discrete achievable ergodic sum-rates. ©2019 IEEE

the maximum entropic normalized zero-mean Gaussian distribution such that $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{K} \mathbf{I}_M)$. The CCMC results are displayed in Figures 5.11 and 5.12 given the DL and UL cases, respectively. It is remarked that indeed the proposed MC-NOMA system outperforms both SCMA and PDMA schemes over both DL and UL across the SNR range. However, the advantage of MC-NOMA over the state-of-the-art benchmarks for the UL scenario is marginal, mostly as consequence of the increased UL channel diversity which in turn improves the sum-rate of both SCMA and PDMA despite their inherent sparseness. These results are also a consequence of the fact that the MC-NOMA frame \mathbf{F} design criteria are optimized such that the DL sum-rate of (5.16) is maximized, while not harming the UL sum-rate as already detailed in Subsection 5.4.1 given (5.29). In other words, the proposed dense incoherent UNTF design is in fact better and more meaningful in shaping the overloaded multiplexing of active users in DL than it is in UL, and this is also visible through the prism of the equivalent UL and DL channels reduced respectively to (5.14) and (5.15).

Given the CCMC results, similar effects and deltas are expected for the DCMC cases realizable over practical systems with discrete-input symbols. And this expectation is confirmed actually by Figures 5.13 and 5.14, as the DCMC counterparts of the previous CCMC plots. The advantage of MC-NOMA over SCMA and PDMA is once again higher in the DL case rather than in UL one, but this behavior was to be expected in the light of Figure 5.12 and of the

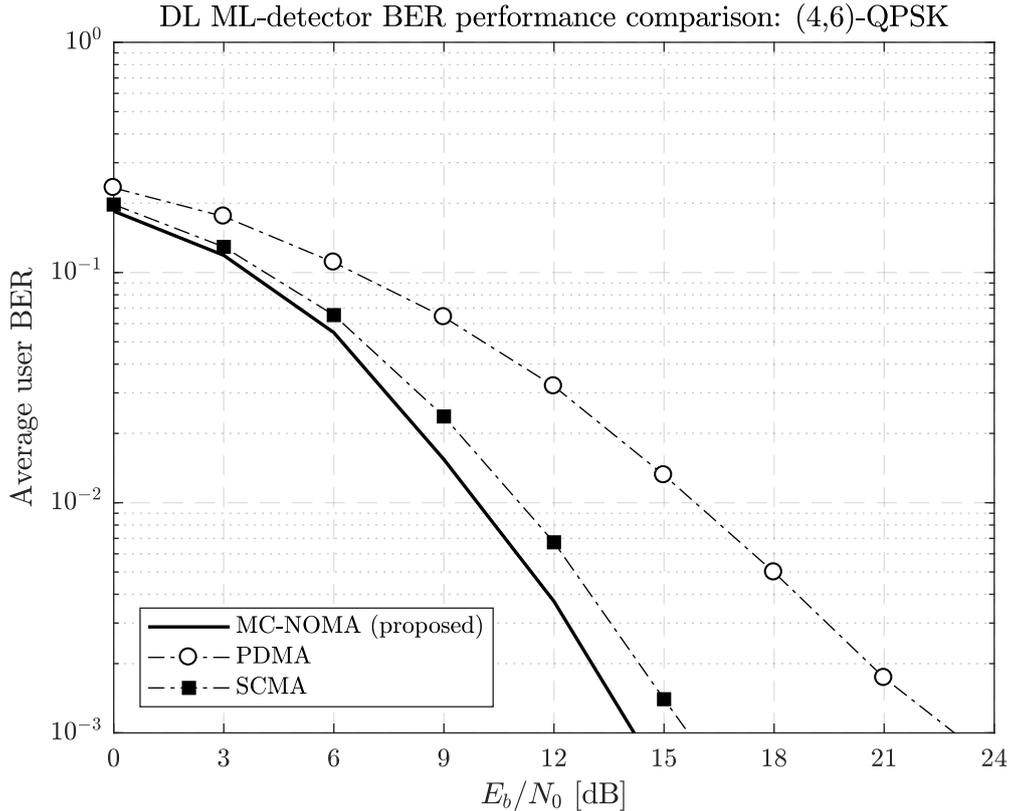


Figure 5.15: Uncoded DL average BER performance via optimum joint ML detection for QPSK. The MC-NOMA proposed scheme is compared against state-of-the-art SCMA [178] and PDMA [177] schemes. ©2019 IEEE

above argumentation, respectively. Moreover, the curves highlight the fact that over the knee SNR regions the gaps between the proposed MC-NOMA and the state-of-the-art benchmarks are larger between DCMC and CCMC across the constellations employed, but also increase upon the constellation order.

This last remark highlights the scalability of the MC-NOMA proposed scheme. This is capable to separate non-orthogonally users of overloaded communications systems by using complex-valued signature waveforms leading to complex spreading under limited and optimized interference levels. In effect, this provides a more efficient use of the available REs in comparison to the best CD-NOMA [156] schemes SCMA [178, 179] and PDMA [176, 177]. This is the case since on one hand PDMA relies on conventional bits-to-symbols mappings and subsequent channel coding gains in non-orthogonally multiplexing active users via sparse recoverable patterns, whereas on the other hand, SCMA uses optimized non-orthogonal codebooks to avoid collisions over multi-dimensional signaling and subsequent rotations [179]. These strategies, become both difficult to scale for larger systems.

Lastly, as a motivation for the section to come, but also to confirm the robustness and practical feasibility of the observed MC-NOMA rate gains over the state of the art, Figure 5.15 displays the average DL BER curves for uncoded joint ML MUD for the same (4, 6) system discussed above. The capacity performance exhibited by MC-NOMA over both SCMA

and PDMA is so further accompanied by increased MUD capability as a result of the frame-theoretic optimized inter-user interference management and full exploitation of the signal space DoFs proposed by MC-NOMA. This last figure coupled with the guaranteed theoretic and numeric sum-rate superior performance of MC-NOMA against state-of-the-art methods constitute therefore the motivation to develop high-performance low-complexity joint MUD in the next section.

5.5 Massively Concurrent Non-orthogonal Multiple Access Detection

In the light of all of the above, we proceed in answering Questions Q3 and Q4 by investigating the MUD at the Rx side. To highlight the optimum achievable performance, perfect CSI at the receiver (CSIR) is assumed and the MC-NOMA frame \mathbf{F} is considered to be known as well. The last assumption is however easily realizable in practice given the necessary dissemination of the waveform signature vectors across the communication terminals.

Furthermore, the uncoded joint MUD optimal capability is targeted in order to stress and outline the maximum detection performance of the proposed MC-NOMA scheme. This approach also provides the necessary detection structures in achieving the ergodic sum-rate results previously observed in Subsection 5.4.3. Additionally, given the high degree of overloading potentially incurred by MC-NOMA systems it is clear that simple detectors would not aid our quest, and traditional matched filter CDMA despreading structures are not sufficient either. As a result, the focus is placed firstly on the joint ML MUD detector capable of providing optimal detection performance with low-complexity. Secondly, the optimal detection complexity requirement is relaxed in favor of a fast and massive-scale joint MUD modern detection approach.

5.5.1 Maximum Likelihood Low-complexity Tree Search Decoding

Let us first turn our attention to the feasible low-complexity realization of a joint MUD scheme for the Rx MC-NOMA, targeting the optimal detection performance of the joint ML detector. Consider therefore at first, the joint ML receiver, given by

$$\hat{\mathbf{s}}_{\text{ML}} = \underset{\mathbf{s} \in \mathcal{S}^K}{\text{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2. \quad (5.45)$$

Note that for the sake of notation simplicity, \mathbf{H} has been used to denote above in Equation (5.45), and hereafter, the equivalent precoded channel in either DL or UL, given respectively by $\mathbf{H}_{k,\text{DL}}\mathbf{F}$ as (5.15) or $\mathbf{H}_{k,\text{UL}}\mathbf{F}$ as (5.14). Thus, the two communication models, *i.e.*, (5.12) and (5.13) respectively, are seen under the unified form of Equation (5.45) from the perspective of the MC-NOMA receivers. In this way, the previously discussed DL-UL equivalence is leveraged to provide practical detector reusability over the treatment of joint MUD in both DL and UL MC-NOMA contexts.

Additionally, recall that the optimum ML detector described in (5.45) is actually equivalent

to the joint maximum a posteriori (MAP) MUD, provided that input symbols are sampled independently and uniformly out of the same constellation \mathcal{S} [164]. This constitutes a divergent path compared to the CD-NOMA literature, which typically relaxes optimality of the detection in favor of lower-complexity achieved through receivers employing decoders like MMSE-SIC, MMSE-PIC, MPA, ESE as detailed in [152, 156] for instance.

As seen through the capacity analysis of Subsection 5.4.1, CD-NOMA in general, and respectively MC-NOMA in particular, may be observed from the perspective of a MIMO overloaded system given the non-orthogonal multiplexing of active users. This remark is even so clearer in the light of the ML problem formulation (5.45). Upon the later analogy, we regard the sphere decoding (SD) [193–196] as a strong candidate for joint MUD capable of reaching the optimal ML performance with much lower complexity. Thus, the initial optimization problem (5.45) is relaxed to its SD formulation

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathcal{S}^K; \mathcal{R}_{\mathbf{H}}^2 \leq \beta^2}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2, \quad (5.46)$$

given the search radius of the system $\mathcal{R}_{\mathbf{H}}$, defined as $\mathcal{R}_{\mathbf{H}} \triangleq \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2$, *i.e.*, the maximum residual of the objective for any prospective solution.

Comparing optimization problems (5.45) and (5.46), their equivalent is immediate, especially by considering the SD radius upper bounded by β as $\beta \rightarrow \infty$. However, in ML terms one would like to find the optimal solution, say \mathbf{s}_* , of the above two problems. The key idea of the search procedure behind solving the SD problem is given by the fact that solving the optimization problem (5.46) is equivalent to solving

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathcal{S}^K; \mathcal{R}_{\mathbf{H}}^2 \leq \|\mathbf{y} - \mathbf{H}\mathbf{s}_*\|_2^2}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2. \quad (5.47)$$

In other words, the solution to (5.46) is the same as the ML one as long as the search radius of the search space, β , is greater than the optimum residual, *i.e.*, $\beta^2 \geq \|\mathbf{y} - \mathbf{H}\mathbf{s}_*\|_2^2$. Based on this remark, the complexity of the search can be significantly reduced, firstly compared to the naive joint ML detector, by subsequently reducing the search radius β .

The approach to solve (5.46), and equivalently (5.47), involves the QR decomposition \mathbf{H} yielding equivalent problem

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathcal{S}^K; \mathcal{R}_{\mathbf{H}}^2 \leq \beta^2}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{Q}\mathbf{R}\mathbf{s}\|_2^2 \quad (5.48a)$$

$$\underset{\mathbf{s} \in \mathcal{S}^K; \mathcal{R}_{\mathbf{R}}^2 \leq \beta^2}{\operatorname{argmin}} \quad \underbrace{\|\mathbf{Q}^H \mathbf{y} - \mathbf{R}\mathbf{s}\|_2^2}_{\tilde{\mathbf{y}} \text{ proj.}} \quad (5.48b)$$

$$\underset{\mathbf{s} \in \mathcal{S}^K; \mathcal{R}_{\mathbf{R}}^2 \leq \beta^2}{\operatorname{argmin}} \quad \|\tilde{\mathbf{y}} - \mathbf{R}\mathbf{s}\|_2^2, \quad (5.48c)$$

given the fact that \mathbf{Q} is unitary, and in effect the projected residual metric $\mathcal{R}_{\mathbf{H}} = \mathcal{R}_{\mathbf{R}}$ is preserved, with $\mathcal{R}_{\mathbf{R}} \triangleq \|\tilde{\mathbf{y}} - \mathbf{R}\mathbf{s}\|_2$.

Under the assumption of system non-orthogonal overloading, the matrix \mathbf{R} becomes upper

trapezoidal, such that

$$\mathbf{R} = \left[\begin{array}{c|c} \underbrace{\mathbf{R}_1}_{\text{upper triangular} \in \mathbb{C}^{M \times M}} & \underbrace{\mathbf{R}_2}_{\text{dense} \in \mathbb{C}^{M \times (K-M)}} \end{array} \right], \quad (5.49)$$

typical for the case of generalized sphere decoding (GSD) applied to MIMO underdetermined linear systems [154, 197–203].

The conventional GSD search is firstly estimating the $K - M + 1$ symbols based on the last row of \mathbf{R} , and then subsequently proceeds in solving the traditional SD problem for the remaining $M - 1$ symbols given the remaining rows of \mathbf{R} . The search radius of the computed solution is then explicitly given by

$$\sum_{i=1}^M \left| \tilde{y}_i - \sum_{j=i}^K r_{ij} s_j \right|^2 \leq \beta^2. \quad (5.50)$$

as singular theoretic constraint imposed on the prospective solution $\mathbf{s} \in \mathcal{S}^K$.

Nevertheless already at the GSD level, to reduce complexity, the search is finely tuned on a symbol-by-symbol basis such that distinct radii are applied to each symbol in the light of the cumulative residual obtained upon each processed row of \mathbf{R} . Concretely, the search of the first $K - M + 1$ symbols corresponding to the M -th row of \mathbf{R} is performed over the local radius

$$\beta_M^2 \triangleq \left| \tilde{y}_M - \sum_{j=M}^K r_{Mj} s_j \right|^2 \leq \beta^2. \quad (5.51)$$

Having made a selection for the $K - M + 1$ symbols associated with the M -th system equation, the search is then resumed for each subsequent equation $\ell \in \{M - 1, M - 2, \dots, 1\}$ over the individual radii

$$\beta_\ell^2 \triangleq \left| \tilde{y}_\ell - \underbrace{r_{\ell\ell} s_\ell}_{\text{interference due to past choices}} - \underbrace{\sum_{j=\ell+1}^K r_{\ell j} s_j}_{\text{shrinking search radius at } \ell\text{-th row}} \right|^2 \leq \beta^2 - \underbrace{\sum_{i=\ell+1}^M \beta_i^2}_{\text{accumulated residual}}. \quad (5.52)$$

Given this general search strategy, in light of (5.52) it is ensured with higher granularity that each processed row of \mathbf{R} does not break through its candidate solution the overall search residual radius. In effect, whenever this is the case the search can be interrupted as any other subsequent solutions for lower rows become invalid themselves.

In essence, the specialized literature provides three alternative strategies to reduce the complexity of the GSD problem in case of the underdetermined systems such as the one under consideration. The first one is to break the GSD problem into subsequent determined SD sub-problems, [154, 197, 198], leading in fact to the search strategy covered above via Equations (5.51) and (5.52). The second approach relies on *Tikhonov regularized GSD* thus transforming the GSD problem of size (M, K) to a square trivially regularized SD problem of size (K, K) by

Tikhonov regularization, [199, 200]. The last strategy is called *slab-decoding* and uses a parametric slab planar constraint for the overcomplete dimensions of the system in a try to reduce their associated search computational complexity before employing the traditional SD for the remaining unknowns [202, 203].

Unfortunately, all of these methods have particular caveats of their own restricting their universal applicability to joint MUD of overloaded systems such as MC-NOMA. In short, the first strategy of GSD suffers from high complexity in detecting the overloaded users upon the fact that for the $K - M + 1$ users corresponding to M -th row of \mathbf{R} a search space of order $\mathcal{O}(Q^{K-M})$ needs to be explored [198].

On the other hand, the regularized GSD method appends trivially regularized equations to expand the system to a square upper triangular matrix \mathbf{R} . The effectiveness of this transform however is bounded on the input signals energy which are strictly required to have unit-power. Albeit, this is true for BPSK or 4-QAM alike, other constellations of bigger sizes or different shapes need to be transformed to these atomic forms [200]. Although this task is possible for regular rectangular quadrature amplitude modulation (QAM) constellations in a general sense [200], it may not be the case for other star-shaped or non-regular constellations that may be utilized. Furthermore, provided such a constellation decomposition exists, this also leads to substantial increases in complexity upon the input space decomposition and subsequent search process similar in complexity to the typical GSD in worst-case sense [199].

Lastly, the slab decoding fits a parametric planar slab over the space of prospective solutions for the $K - M + 1$ equations corresponding to the M -th row of the matrix \mathbf{R} [202]. The objective of this operation is to constrain these variables to a geometric search region able to reduce the exhaustive search space complexity of $\mathcal{O}(Q^{K-M})$ [203]. However, given the parameteric selection of the slab, potentially selected by the expected noise statistic, the scheme is susceptible to degradation and may fail to obtain optimal performance. Furthermore, although applicable to arbitrary constellations, this approach offers little complexity reduction in systems with high overloading ratios such as MC-NOMA, since effectively high-dimensional slabs become harder to design and optimize [203].

In the light of all the arguments above, it is clear that usual strategies to reduce the complexity of GSD may not be suitable to the generic requirements of MC-NOMA receivers. As a consequence, we propose our own variant of joint MUD detection in the sequel. For this purpose, the starting point is considered to be the standard GSD algorithm, given that it is a well-studied approach [154, 195, 197, 198], and that it can be easily optimized further without loss of generality. The core ideas behind this method have already been introduced above via equations (5.48)-(5.52), and so, the scheme is reduced in fact to a depth-first search (DFS) constrained by adapted SD radii [195, 198]. Despite that the worst-case complexity of such a DFS tree search is of order $\mathcal{O}(Q^K)$, [204], the authors of [195] have shown that on average the stochastic complexity associated with the SD is of polynomial order, with an exponent dependent on the SNR. Motivated by this remark, an enhanced GSD is proposed in the sequel, which was designed based on the goal to reduce the search complexity by the application of various *reduction*, *enumeration* and *pruning* techniques in the context of MC-NOMA receivers.

The Lattice-based Reduction

As mentioned earlier, a multitude of methods are available as tools to optimize the performance of generic SD and GSD. One of these, often employed in practice and also applicable to MC-NOMA systems, is the Lenstra-Lenstra-Lovász (LLL) lattice reduction [205, 206] over infinite lattices. Nevertheless, provided that standard QAM constellations are used, search recursions may also be accelerated upon lattice-reduction methods, such as those discussed in [154, 197–201, 207], spanning from LLL reduction, provided that the complex based values in equation (5.48) are mapped to their corresponding real-numbered equivalents. Fortunately, for any standard QAM finite constellation this is easily achievable by the mapping

$$\tilde{\mathbf{y}} \xrightarrow{\mathbb{R}} \tilde{\mathbf{y}} \triangleq \begin{bmatrix} \Re[\tilde{\mathbf{y}}] \\ \Im[\tilde{\mathbf{y}}] \end{bmatrix} \in \mathbb{R}^{2M}, \quad (5.53)$$

$$\mathbf{R} \xrightarrow{\mathbb{R}} \mathbf{R} \triangleq \left[\begin{array}{c|c} \Re[\mathbf{R}] & -\Im[\mathbf{R}] \\ \hline \Im[\mathbf{R}] & \Re[\mathbf{R}] \end{array} \right] \in \mathbb{R}^{2M \times 2K}, \quad (5.54)$$

$$\mathbf{s} \xrightarrow{\mathbb{R}} \mathbf{s} \triangleq \begin{bmatrix} \Re[\mathbf{s}] \\ \Im[\mathbf{s}] \end{bmatrix} \in \mathbb{R}^{2K}, \quad (5.55)$$

while also preserving the form of equation (5.48c) based on the new search space obtained under the projection of the individual \mathcal{S} constellations on the real and imaginary axes.

Clearly, this latter operation resumes in the case of QAM constellations to obtaining correspondingly discrete integer points search rulers across the real and imaginary axes, hereafter denoted as \mathcal{Z} . As a result, the original problem over the search space \mathcal{S}^K is thus reduced to the *underdetermined discrete least squares* problem [201] over the obtained search space \mathcal{Z}^{2K} .

Note as well that under the above complex-to-real transformation, the 2-norm of the residual metrics (5.50)-(5.52) remains invariant and change just their subsequent dimensionality given the doubling of the problem, *i.e.*, $(M, K) \mapsto (2M, 2K)$.

The Schnorr-Euchner Enumeration

Another typical structural optimization that is usually performed on top of the SD family of DFS tree search algorithms is given by the search space ordering. Clearly, proper sorting of the search space at each decision point in the tree affects the evolution of the decreasing radius [194]. In this way, “good” decisions sequentially decrease complexity as they generate lower residuals and decrease the detector’s search radius considerably upon early traversals. Effectively, this then leads in turn to pruning latter “bad” decisions [195].

An effective, well-studied and widely-applied enumeration technique is the Schnorr-Euchner (SE) enumeration proposed in [208]. This sorting, indexed non-decreasingly by the corresponding residuals generated by each prospective solution [208], leads actually to a zig-zag search over the integer line search solution space, being mostly referred to in conjunction with the SD directly as the Schnorr-Euchner sphere decoding (SESD) [209]. By means of SE enumeration, the

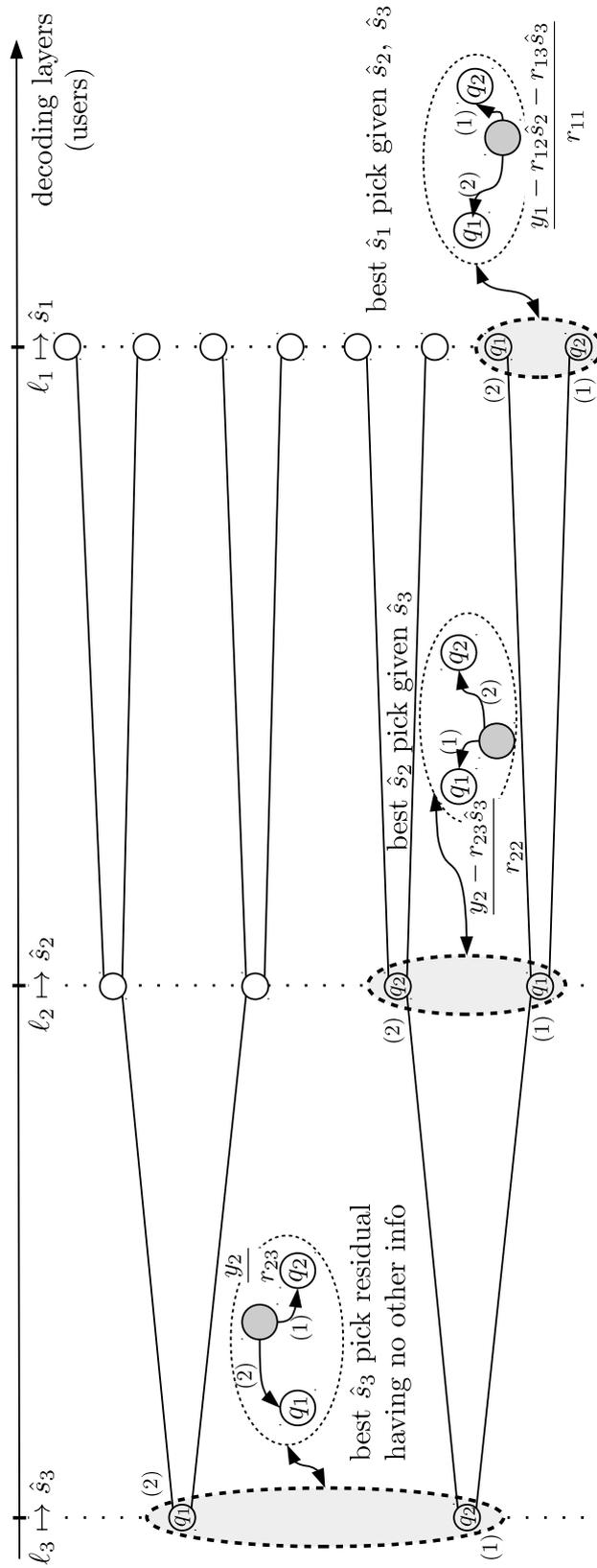


Figure 5.16: Illustration of SE ordering and tree search optimization for a (2, 3) demonstrative overcomplete access system serving 3 users over 2 resources with binary bipolar constellations. ©2019 IEEE

search tree itself is therefore sorted, arranging the solutions in accordance with non-decreasing residuals on the left-most branches, or alternatively the ones that are visited firstly, as depicted in Figure 5.16 for clarity. This approach combined with the SD conventional search radius update detailed earlier, decreases effectively the average search complexity down to polynomial times with varying exponents depending on the SNRs [195]. As a result, it can also help in the GSD case.

The Proposed Tree Pruning Strategies

Both the LLL-inspired reductions and SE enumeration optimizations previously summarized are unfortunately not sufficient to lower the complexity of MC-NOMA receivers, and generic overloaded systems for that matter. This is especially the case for large overloading cases, as the joint detection of the first $K - M + 1$ symbols is based upon just the last row of the trapezoidal matrix \mathbf{R} as depicted in Equation (5.49). As a consequence, the tree pruning vital to complexity reduction cannot be performed earlier than the M -th detection layer (user signal) in the tree, since just at that level the M -th system equation (M -th row of matrix \mathbf{R}) can be utilized together with its associated radius inequality (5.51). The net result is that both in average and worst-case terms complexity is governed by $\mathcal{O}(Q^{K-M})$ [200, 201].

To try and reduce this complexity and ensure practicality of the GSD-based decoding at least for smaller MC-NOMA systems, consider below the revisit of the residuals incurred in traversing user signals s_k , such that

$$\beta_k^2 \triangleq \left| \underbrace{\left(\tilde{y}_M - \sum_{j=k+1}^K r_{Mj}s_j \right)}_{\text{past residual}} - \underbrace{r_{Mk}s_k}_{k\text{-th pick}} \right|^2 \leq \beta^2 \quad (5.56a)$$

$$= |r_{Mk}|^2 \left| \left(\frac{\tilde{y}_M - \sum_{j=k+1}^K r_{Mj}s_j}{r_{Mk}} \right) - s_k \right|^2 \leq \beta^2 \quad (5.56b)$$

for each detection layer $k \in \{M, M + 1, \dots, K\}$.

In the light of the above Equation (5.56), some additional strong insights towards further search optimizations based on reduction and tree pruning can be derived, which are pursued in the sequel.

A. Algebraic Reordering

A key general remark is that in GSD problems, the DFS tree search associated with the general SD ideas is in fact a family of trees with each one rooted by a possible solution points at the top detection layer [200, 201]. As a consequence, it is really desirable to prune subtrees as early as possible to lower the MUD detection complexity. Given that the first opportunity to prune the tree is at the M -th layer, as previously noted, the focus is placed on the M -th row equation of the linear system (5.48c).

Observing the incremental inequalities (5.56), it can be further remarked that M -th row

entries of the matrix \mathbf{R} are in fact weights affecting the residual contributions some symbol selections may impact onto others. Thus, to perform an algebraic reduction of the search space given the matrix \mathbf{R} , it is of high interest to penalize bad choices earlier, *i.e.*, at higher layers than at lower layers. Mathematically this reduces to ordering the power of the weights non-decreasingly from layer M to layer K respectively, such that

$$|r_{MM}|^2 \leq \dots \leq |r_{Mk}|^2 \leq \dots \leq |r_{MK}|^2. \quad (5.57)$$

Based upon the same ideology it is also required to collect the most energy within the M -th row for each signal layer $k \in \{M, M+1, \dots, K\}$, *i.e.*, such that the $|r_{Mk}|^2$ entries have the highest possible magnitudes given the matrix \mathbf{R} . However, such a requirement leads to a combinatorial optimization problem that is practically hard to solve, and as a result, this condition is relaxed to a criterion of maximum sum-of-squares requirement for the last row of the trapezoidal matrix \mathbf{R} , as

$$\max_{\mathbf{R}} \sum_{k=M}^K |r_{Mk}|^2. \quad (5.58)$$

To implement the reduction conditions proposed in (5.57) and (5.58) complete column and row permutations of the receive signal model is necessary as follows

$$\underbrace{\mathbf{P}_r \mathbf{y}}_{\text{permuted } \mathbf{y}} = \underbrace{\mathbf{P}_r \mathbf{H} \mathbf{P}_c}_{\text{permuted } \mathbf{H}} \cdot \underbrace{\mathbf{P}_c^H \mathbf{s}}_{\text{permuted } \mathbf{s}}, \quad (5.59)$$

with \mathbf{P}_r and \mathbf{P}_c defined as corresponding row/column binary-valued permutation matrices.

In light of (5.59), the proposed algebraic reduction following conditions (5.57), and respectively, (5.59), is reduced to finding the optimized QR under the permutation matrices \mathbf{P}_r and \mathbf{P}_c . Let us next revisit the linear system in (5.59), reinterpreted in the light of such a permuted QR decomposition as

$$\mathbf{y}_P = \underbrace{\mathbf{P}_r \mathbf{H} \mathbf{P}_c}_{\mathbf{QR}} \cdot \mathbf{s}_P \quad (5.60a)$$

$$\mathbf{Q}^H \mathbf{y}_P = \mathbf{R} \mathbf{s}_P \quad (5.60b)$$

$$\tilde{\mathbf{y}}_P = \mathbf{R} \mathbf{s}_P, \quad (5.60c)$$

with the permuted received signal $\mathbf{y}_P \triangleq \mathbf{P}_r \mathbf{y}$, and respectively, the permuted input signal $\mathbf{s}_P \triangleq \mathbf{P}_c^H \mathbf{s}$.

Following (5.60), it results that as step preceding the tree traversal specific to the GSD family of algorithms, it is appropriate to carefully design the QR decomposition of (5.60), with the aim to maximize the objective in (5.58), while satisfying the chain of inequalities in (5.57).

The idea of QR column permutation and reordering as an enhancement of detection performance and complexity reduction is not new in the context of MIMO systems or space-time codes, see for instance related works of [201, 210–213]. Most of these works, recur in fact to ordering the signal space non-decreasingly such that either the users with the highest SNRs

Algorithm 5 RUBIK-QR DECOMPOSITION

Inputs: $\mathbf{H} \in \mathbb{C}^{M \times K}$, $M \leq K$

Outputs: unitary $\mathbf{Q} \in \mathbb{C}^{M \times M}$, trapezoidal $\mathbf{R} \in \mathbb{C}^{M \times K}$, rows permutation $\mathbf{P}_r \in \mathbb{Z}^{M \times M}$, columns permutation $\mathbf{P}_c \in \mathbb{Z}^{K \times K}$

- 1: Initialize $\mathbf{P}_r = \mathbf{I}_M$, $\mathbf{P}_c = \mathbf{I}_K$, $\mathbf{R} = \mathbf{H}$
 - 2: Sort columns of \mathbf{R} according to rule $\|\mathbf{R}_{(:,1)}\|_2 \leq \|\mathbf{R}_{(:,2)}\|_2 \leq \dots \leq \|\mathbf{R}_{(:,K)}\|_2$ and obtain columns permutation $\mathbf{p}_c \in \mathbb{Z}^K$
 - 3: Update $\mathbf{P}_c \mapsto \mathbf{P}_{c(:,\mathbf{p}_c)}$, $\mathbf{R} \mapsto \mathbf{R}\mathbf{P}_c$
 - 4: Sort rows of \mathbf{R} according to rule $\|\mathbf{R}_{(1,:)}\|_2 \leq \|\mathbf{R}_{(2,:)}\|_2 \leq \dots \leq \|\mathbf{R}_{(M,:)}\|_2$ and obtain rows permutation $\mathbf{p}_r \in \mathbb{Z}^M$
 - 5: Update $\mathbf{P}_r \mapsto \mathbf{P}_{r(\mathbf{p}_r,:)}$, $\mathbf{R} \mapsto \mathbf{P}_r\mathbf{R}$
 - 6: **for** $c = 1:M$ **do**
 - 7: Given $\mathbf{R}_{(:,c)}$, generate Householder reflector \mathbf{u} , [214]
 - 8: Cache column reflector $\mathbf{U}_{(c:M,c)} = \mathbf{u}$
 - 9: Apply Householder transform (HHT) to submatrix $\mathbf{R}_{(c:M,c:K)} \mapsto \text{HHT}(\mathbf{u}, \mathbf{R}_{(c:M,c:K)})$, [214]
 - 10: **end for**
 - 11: Sort last $K - M + 1$ columns by $|\mathbf{R}_{(M,M)}| \leq |\mathbf{R}_{(M,M+1)}| \leq \dots \leq |\mathbf{R}_{(M,K)}|$ and update permutation $\mathbf{p}_c \in \mathbb{Z}^{K-M+1}$
 - 12: Update $\mathbf{P}_{c(:,M:K)} \mapsto \mathbf{P}_{c(:,\mathbf{p}_c)}$
 - 13: Retrieve $\mathbf{Q} \mapsto \text{HHT}(\mathbf{U}, \mathbf{I}_M)$ from cached reflectors, [214]
-

or highest powers are detected first. In some sense the proposed idea is similar through its condition (5.57), but in conjunction with the sum-of-squares condition (5.58) what we achieve is actually maximizing the detection energy jointly over all users for the last row equation of \mathbf{R} . This strategy suits best our joint detection approach based on the GSD tree structure.

Since the QR decomposition preserves the column ordering of the input matrix [43] under unitary left transforms, it follows that column permutations can always be performed as long as the triangularity of the generated \mathbf{R} is not impeded. Upon this remark, firstly a full permutation of columns is done according so as to order ascendingly from left to right the column norms of the input matrix to be QR-decomposed. Next, to maximize the energy towards the lower-right corner of the input matrix and greedily solve the condition (5.58), a full rows permutation follows sorting rows according to their row norms, ascendingly from top to bottom. Having imposed this structure on the input matrix, then the QR decomposition is applied followed by a fine reordering of the last $K - M$ columns so as to satisfy criterion (5.57).

The method described above is implemented efficiently by algebraic reflections of Householder Transformation [214] to yield the desired permuted QR conditioned by (5.57) and (5.58) in polynomial time. In addition, the proposed permuted QR decomposition is hereafter referred to as the *Rubik-QR decomposition* upon the successive column and row permutations of \mathbf{R} that resemble block rotations necessary to solve a Rubik cube. Lastly, for completeness the

pseudo-code of the the Rubik-QR algorithm is listed¹ under Algorithm 5.

B. Probabilistic Tree Pruning

Despite all the efforts of the GSD optimization techniques discussed so far, the complexity of highly overloaded systems may well be still too complex to attack practically, given the impossibility of early tree pruning. To generally address this issue especially in order to cap the complexity, it is highly desirable to perform as early branch cuts as possible upon some “magic” threshold. This problem motivated vast amounts of work on various stochastic methods to identify such thresholds. For instance, the works [215–217] rely on noise statistics to define parameterized pruning thresholds. These approaches are generally named under the collective term *probabilistic tree pruning*.

In [217] a framework for probabilistic tree pruning was proposed based on different pruning strategies like K -best pruning, uniform pruning, pruning by thresholding or hybrids thereof. One of the central results of [217], *i.e.*, Proposition 2, relates the pruning strategy of hard thresholding to the residual noise probabilistic distribution under general signal space constraints. Despite the fact that [217] considered only the conventional full-rank system corresponding to the canonical SD and QR decomposition of square matrices, it is obvious however that the noise distribution of the residuals is preserved under the unitary transforms of Rubik-QR from Algorithm 5. In fact, the squared residuals $\beta_\ell^2, \ell = \{1, 2, \dots, M\}$, see Equations (5.51) and (5.52), are distributed according to the chi-squared distribution provided the assumption of AWGN system noise at the Rx. As a consequence, the same technique of hard thresholding derived in [217] on top of the chi-squared distribution of the squared noise residuals is applicable to our overloaded tree detector scheme as well.

Concretely, given the overloaded MC-NOMA detection, pruning via hard thresholding δ starting at the earliest possible M -th detection layer is obtained by [217, Eq. (14)], with δ respectively given as

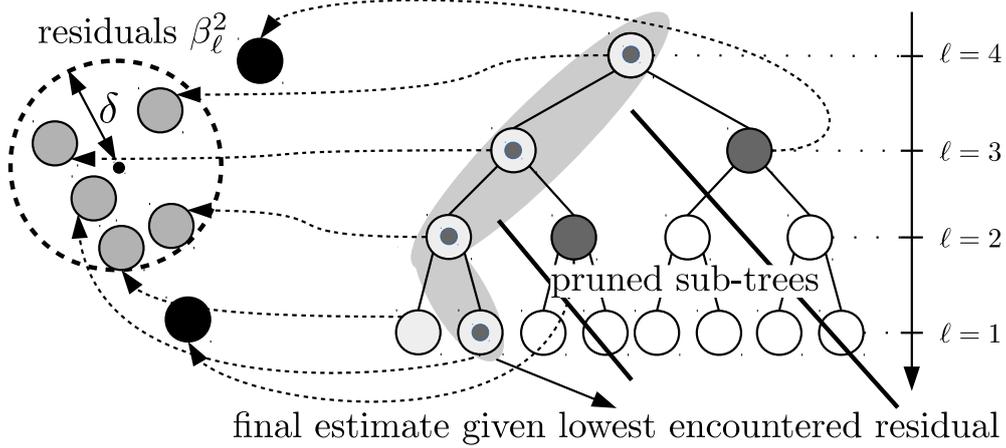
$$\delta = 2\sigma^2 \cdot \gamma^{-1} \left(1 - \left(1 + \frac{d_{min}^2}{4\sigma^2} \right)^{-M} \right), \quad (5.61)$$

In the expression (5.61) defining the threshold δ , d_{min} marks the minimum distance of the finite constellations \mathcal{S} radiated, and respectively, $\gamma(x)$ is the lower incomplete gamma function

$$\gamma(x) \triangleq \int_0^x \exp(-t) dt. \quad (5.62)$$

Thus, applying hard thresholding by δ from Equation (5.61) and given the general conditions

¹Note that Octave GNU and MATLAB[®] [190] notations for submatrix and subvector selection are used for simplicity.


 Figure 5.17: Probabilistic tree pruning by δ -thresholding. ©2019 IEEE

of SD, *i.e.*, (5.51) and (5.52), at any layer $\ell = \{1, 2, \dots, M\}$, only the leaves whose paths fulfill

$$\sum_{i=\ell}^M \beta_i^2 \leq \beta^2, \quad (5.63a)$$

$$\beta_\ell^2 \leq \delta, \quad (5.63b)$$

are further traversed, whereas the leaves violating these constraints and their corresponding subtrees are cut, as displayed in Figure 5.17.

In the light of all of the above optimization approaches towards the general GSD, we refer to the complete MC-NOMA MUD as the *rubik probabilistic tree search with thresholding (Rubik-PTST)*. A pseudo-code implementation of this MUD scheme summarizing its full logic is offered in Algorithm 6. Notice that internally the proposed Rubik-PTST uses an inner standard SESD decoder [207, Alg. 6], enhanced by the probabilistic tree pruning by thresholding $\text{SESD}_{\text{THRESH}}(\cdot)$, as described above. The algorithm [207, Alg. 6] is a streamlined and optimized implementation of a regular SESD improved via fast LLL-reduced recursions over finite QAM constellations, and hence, the problem it solves is in fact the real-valued one obtained via (5.53) - (5.55).

To summarize, the proposed Rubik-PTST listed as Algorithm 6 is an effective joint detector for overloaded systems that embeds the LLL lattice reduction, the SE enumeration, the proposed Rubik-QR algebraic reordering from Algorithm 5, and the probabilistic tree pruning by noise-optimized thresholding.

Performance Evaluation

The effectiveness and efficiency of the proposed Rubik-PTST method as a joint optimal low-complexity receiver for the introduced MC-NOMA scheme are evaluated in the sequel by means of numerical simulations. Motivated by the more stringent complexity restrictions of the UEs than those of the BS, the focus is thus heavily placed on the practical feasibility and decodability of DL MC-NOMA in detriment of UL MC-NOMA. However, in the light of the

Algorithm 6 RUBIK-PTST MULTI-USER DETECTOR

Inputs: $\mathbf{y} \in \mathbb{C}^M$, $\mathbf{H} \in \mathbb{C}^{M \times K}$, constellation \mathcal{S} , σ^2
Outputs: estimates $\mathbf{s} \in \mathcal{S}^K$

```

1: Set initial search radius  $\beta = +\infty$ 
2: Compute threshold  $\delta$  given (5.61)
3: Set layer  $\ell = K$ 
4: Allocate default candidate symbol sets  $\forall \ell$  layers,  $S_\ell$ 
5: Apply Rubik-QR reduction
   ( $\mathbf{Q}, \mathbf{R}, \mathbf{P}_r, \mathbf{P}_c$ ) = RUBIK-QR ( $\mathbf{H}$ )
   Update  $\mathbf{y} \mapsto \mathbf{Q}^H \mathbf{P}_r \mathbf{y}$ 
6: while (1) do
7:   if  $\ell \geq M$  then
8:     if EMPTY( $S_\ell$ ) then
9:       if  $\ell = K$  then
10:        break
11:      else
12:        Reset  $S_\ell$  to default ▷ clears visit
13:        Go up the tree  $\ell \mapsto \ell + 1$ 
14:        continue
15:      end if
16:    else
17:      if set  $S_\ell$  is not visited yet then
18:        Apply SE to  $S_\ell \mapsto \text{SE}(\mathcal{S}, \mathbf{R}, \mathbf{y}_M, \hat{\mathbf{s}}_{(\ell+1:K)})$ 
19:        if  $\ell = M$  then
20:          for all symbols  $S_{\ell,i} \in S_\ell$  do
21:            if not conditions (5.63) then
22:              REMOVE( $S_\ell, S_{\ell,i}$ ) ▷ prune  $S_{\ell,i}$  subtree
23:            end if
24:          end for
25:        end if
26:      end if
27:      Get layer  $\ell$  estimate  $\hat{\mathbf{s}}_{(\ell)} \mapsto \text{POP\_FRONT}(S_\ell)$ 
28:      Go down the tree  $\ell \mapsto \ell - 1$ 
29:    else
30:      Solve inner SD sub-problem
31:       $\hat{\mathbf{s}}_{(1:M-1)} \mapsto \text{SESDTHRESH}(\mathbf{y}_{(1:\ell)}, \mathbf{R}_{(1:\ell, 1:\ell)}, \delta, \mathcal{S})$ 
32:      Compute obtained residual  $\beta_{\text{err}}^2 = \|\mathbf{y} - \mathbf{R}\hat{\mathbf{s}}\|_2^2$ 
33:      if  $\beta_{\text{err}}^2 \leq \beta^2$  then
34:         $\beta \mapsto \beta_{\text{err}}$  ▷ Decrease search radius
35:      end if
36:      Update search layer  $\ell = M$ 
37:    end if
38:  end while
39: Revert Rubik-QR permutation:  $\mathbf{s} \mapsto \mathbf{P}_c \mathbf{s}$ 

```

DL-UL equivalence exhibited under the given model, this focused perspective is in fact irrelevant for the general behavior, achievable performance and complexity exhibited by the Rubik-PTST detector across both DL and UL MC-NOMA systems.

The simulation setup entailed therefore the software replication of end-to-end MC-NOMA communications links under the generic linear system model described in Section 5.3. The wireless channels were modeled across all realizations as i.i.d Rayleigh fading channels, and the assumptions of the system models (5.12) and (5.13) were followed for DL, and respectively, for UL cases. The MC-NOMA precoding frames were generated as dense, incoherent UNTFs as described in Subsection 5.4.2 using Algorithm 3. Moreover, each data point across all the performance curves, was generated upon averaging over at least 5000 stochastic noise and channel realizations to ensure sufficient statistical significance.

The optimized GSD based upon the ordered SESD accelerated by fast lattice-based recursions implemented by [207, Alg. 6] was considered as a reference benchmark to the proposed Rubik-PTST receiver. This choice is motivated by a two-fold argumentation. Firstly, it confers a method known to efficiently reach the ML-detection performance over finite search sets [196,207]. Secondly, it implements in a great part (except the additional pruning by hard thresholding) the same inner detection routine that the Rubik-PTST is based on. As a consequence of the latter, a direct benchmarking of the improvements brought by the two newly proposed tree op-

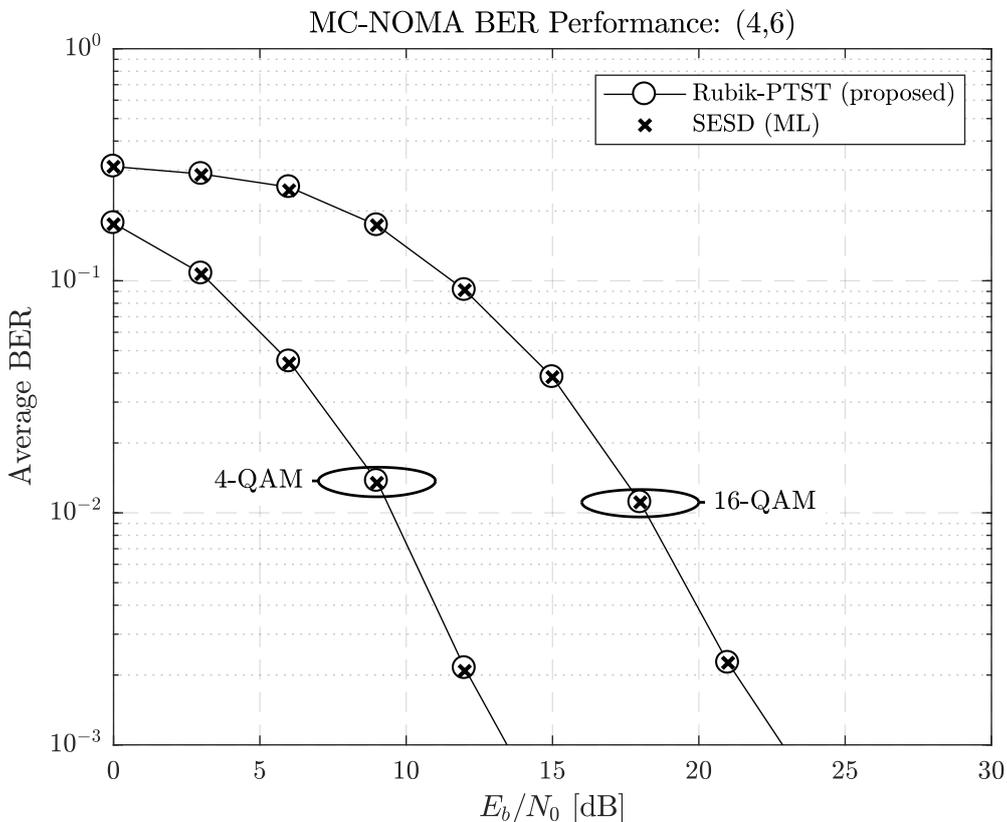


Figure 5.18: BER performance of MC-NOMA with Rubik-PTST and ML detection via SESD [207] optimized GSDs in a $(M, K) = (4, 6)$ system with 4-QAM and 16-QAM inputs. ©2019 IEEE

timization techniques, *i.e.*, algebraic reduction and tree search probabilistic pruning, is possible compared to a state-of-the-art equivalent.

Figure 5.18 displays the average BER performance of a DL system with 150% overloading given by $K = 6$ active users sharing $M = 4$ resources. It can be seen that indeed the proposed Rubik-PTST receiver achieves the joint ML detection performance. Based on this fact and given the capacity analysis performed under Subsection 5.4.1 upon the joint ML MUD assumption, it follows that the proposed MC-NOMA scheme with Rubik-PTST detection is able to practically achieve the capacity curves shown in Figures 5.13 and 5.14.

A natural question given the results of Figure 5.18 could be *what is the associated complexity with this performance*. The ML brute force search computational cost is obviously $\mathcal{O}(Q^K)$, whereas it is known that for determined systems the average SESD complexity is polynomial [195, 207, 217]. However, in the light of the overloading $K > M$, the latter tends to diverge from polynomial complexity and become dominated by $\mathcal{O}(Q^{K-M})$ [204] as $K - M$ increases. This is in fact a consequence of the DFS traversal technique solving the joint detection, and respectively, of the limitation of being able to prune the tree only at the M -th detected layer.

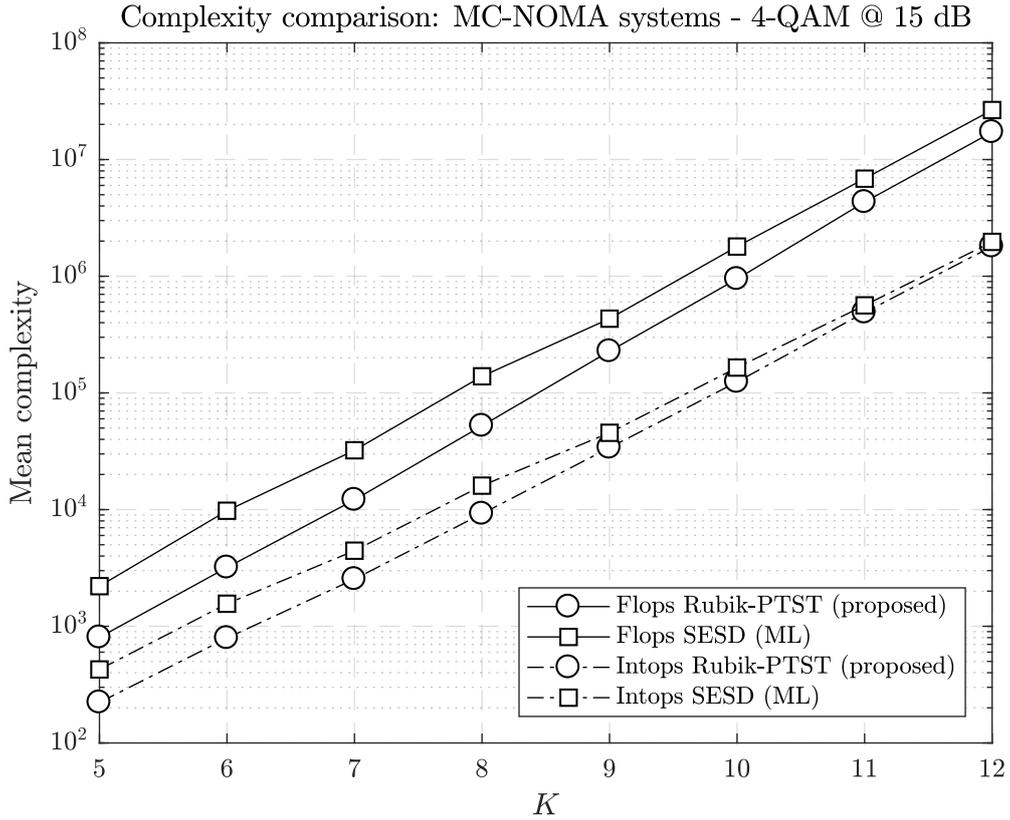
In essence then, a more in depth comparison between the complexity of the proposed Rubik-PTST and the GSD implementation via SESD is therefore necessary. This is offered by Figure 5.20, which shows both the absolute and the relative complexities of the two methods as a function of the number of users $K > M$, with the number of resources fixed to $M = 4$ under the SNR of 15dB and 4-QAM constellations.

Firstly, the absolute numbers counted in terms of floating point and integer point operations, *i.e.*, flops and intops, are displayed Figure 5.19a. This plot reveals in turn that the complexity of the proposed Rubik-PTST receiver measured in absolute flops and intops is not only lower than that of the state-of-the-art benchmark, but also follows the same trend. Under this remark, it is clear that unfortunately the shift from polynomial to exponential time complexity under increasing overloading is not avoided by the proposed Rubik-PTST. Nevertheless, the absolute values of the associated flops and intops necessary for Rubik-PTST detection in MC-NOMA indicate that even at 300% overloading, joint ML detection is achievable in milliseconds using standard modern computing² and even without requiring optimized or embedded computational environments, which speaks of the practical feasibility of MC-NOMA altogether. These numbers advocate for the practical feasibility of MC-NOMA altogether, as the Rubik-PTST could at any point be relaxed for computational efficiency in detriment of numeric performance by capping for instance the number of total visits to the last leafs in the tree [213].

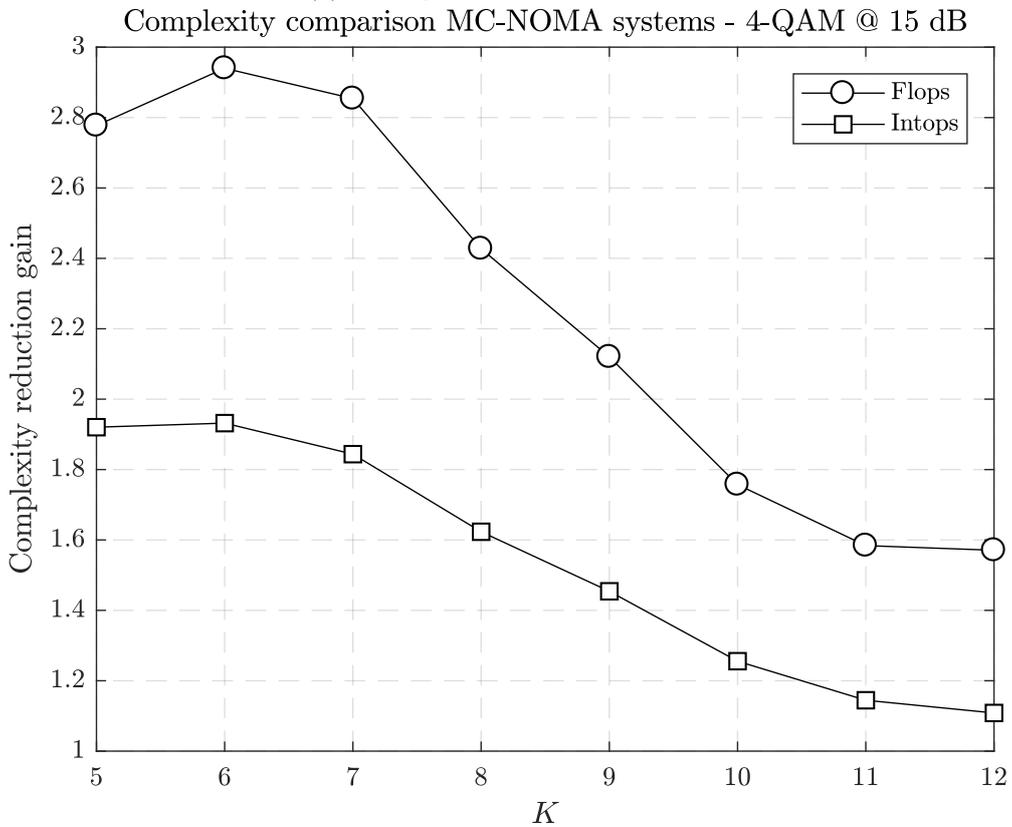
Subfigure 5.19b shows in the sequel the relative complexity reduction of the Rubik-PTST receiver over the optimized and accelerated GSD SESD implementation, defined as

$$G_R \triangleq \frac{\sum_{\mathbf{H}_i; i=1}^I \text{flops}_{\text{GSD}}}{\sum_{\mathbf{H}_i; i=1}^I \text{flops}_{\text{Rubik-PTST}}}, \quad (5.64)$$

²It is assumed that CPU's processing speeds are in the typical GHz range. For instance, the 300% (4, 12) MC-NOMA system is decoded at 15 dB on average by an unoptimized C implementation of Rubik-PTST running on an Intel i5 3rd Gen. quad-core processor @ 2.67GHz in 112ms



(a) Average number of operations.



(b) Complexity reduction gain.

 Figure 5.19: Complexity evaluation of MC-NOMA (M, K) systems with Rubik-PTST and SESD optimized GSD receivers over 4-QAM at SNR = 15dB, with $M = 4$ and $K = \{5, 6, \dots, 12\}$.

where the number of iterations considered, *i.e.*, I , was set to 1000.

The results reveal that while higher gains are achieved by the Rubik-PTST receiver at moderate overloading factors of up to 200%, the proposed decoder continues to outperform the state-of-the-art also for higher overloading scenarios as the two approach in complexity as $K - M$ increases.

Finally, in order to outline that the ML-achievable performance of the Rubik-PTST receiver is not affected given the proposed embedded algebraic reduction and tree pruning strategies, a comparison of the BER performance for the scenarios studied in Figure 5.19 is displayed in Figure 5.20 for the fixed SNR level of 15 dB. Clearly, the performance curve of Rubik-PTST matches the one of the SESD implementation, known to provide the optimum performance upon setting $\beta = \infty$, [207]. Along the same lines, the full curves over varying SNR of three different overloading levels, *i.e.*, (4, 6), (4, 8), (4, 12), are displayed in Figure 5.21. Once more it is seen that the Rubik-PTST joint detector is able to reach the optimum detection, whereas the curves exhibit otherwise the normal behavior presenting performance degradation upon the increase of the overloading. This comes as a consequence of increasing the mutual coherence of the MC-NOMA frame \mathbf{F} leading to higher inter-user interference, and thus subsequent detection impairments.

Having demonstrated the efficiency, efficacy and robustness of the proposed Rubik-PTST receiver for MC-NOMA overloaded systems, we lastly turn our attention to the scalability of

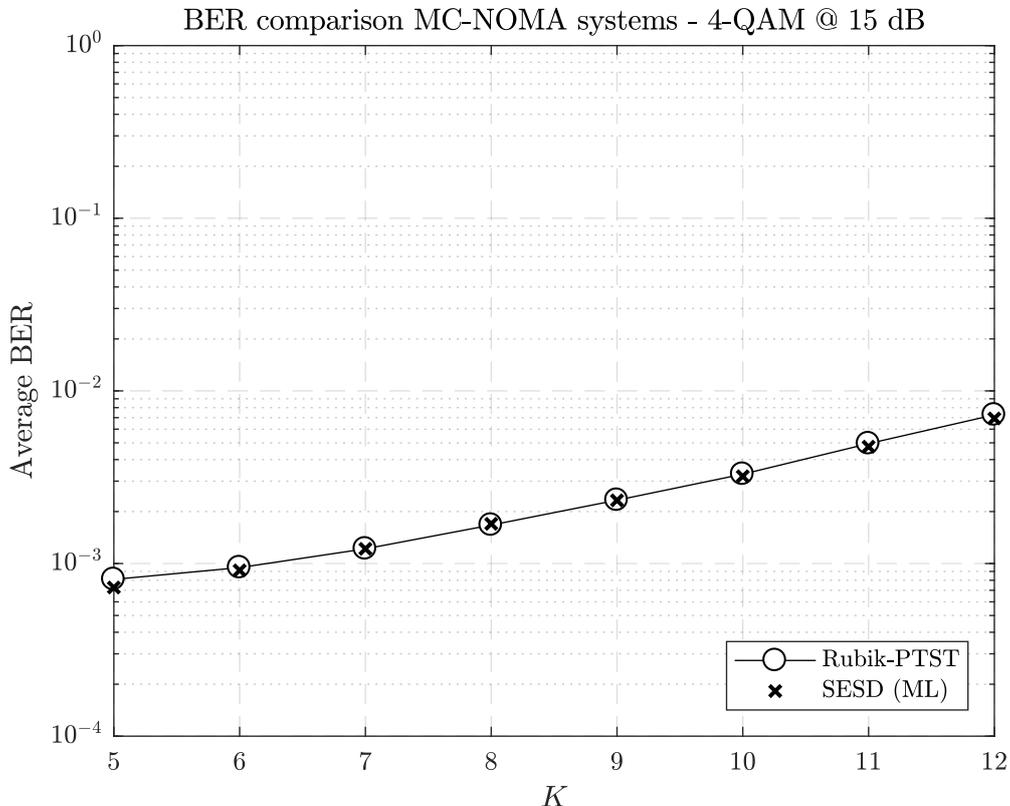


Figure 5.20: Average BER performance of MC-NOMA systems of size (M, K) , such that $M = 4$, and respectively, $K \in \{5, 6, 7, 8, 9, 10, 11, 12\}$. ©2019 IEEE

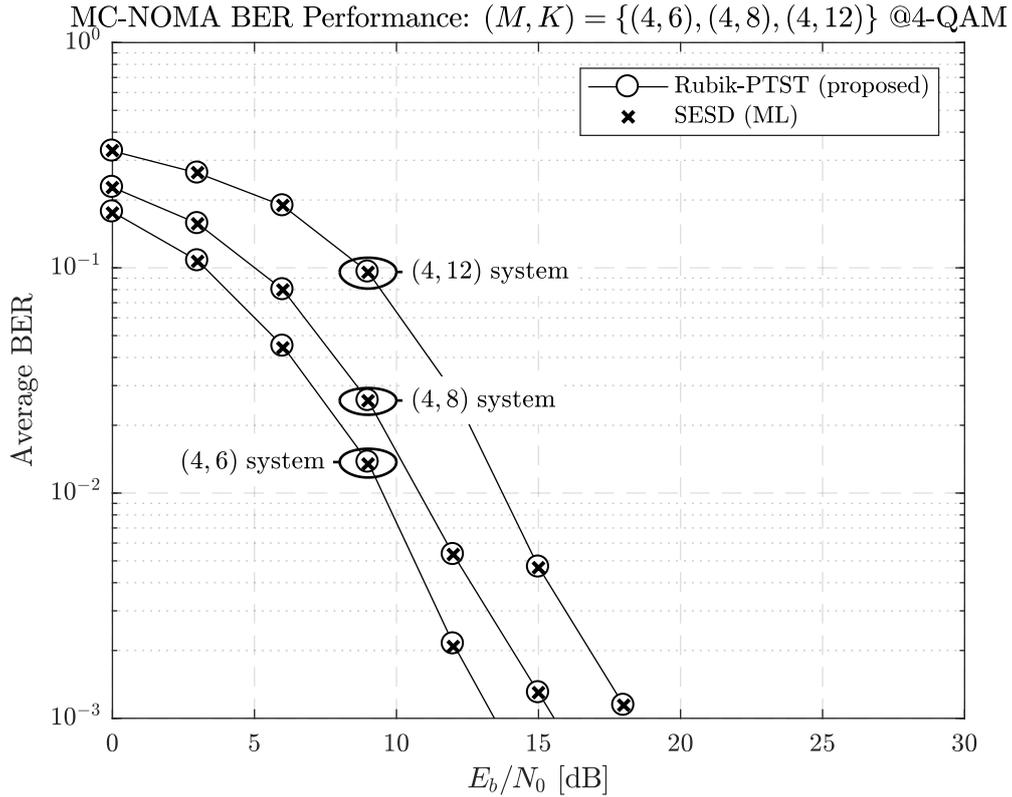
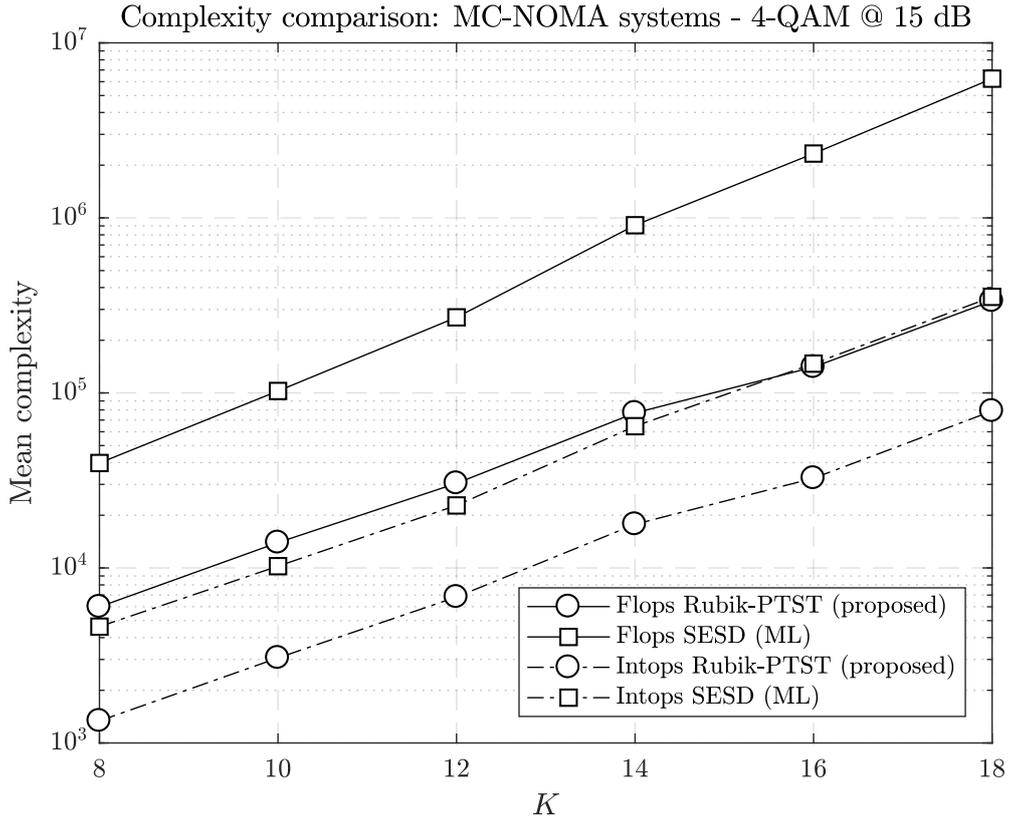


Figure 5.21: BER performance of MC-NOMA with Rubik-PTST and ML detection via SESD [207] optimized GSDs for systems of size $(M, K) = \{(4, 6), (4, 8), (4, 12)\}$ and 4-QAM. ©2019 IEEE

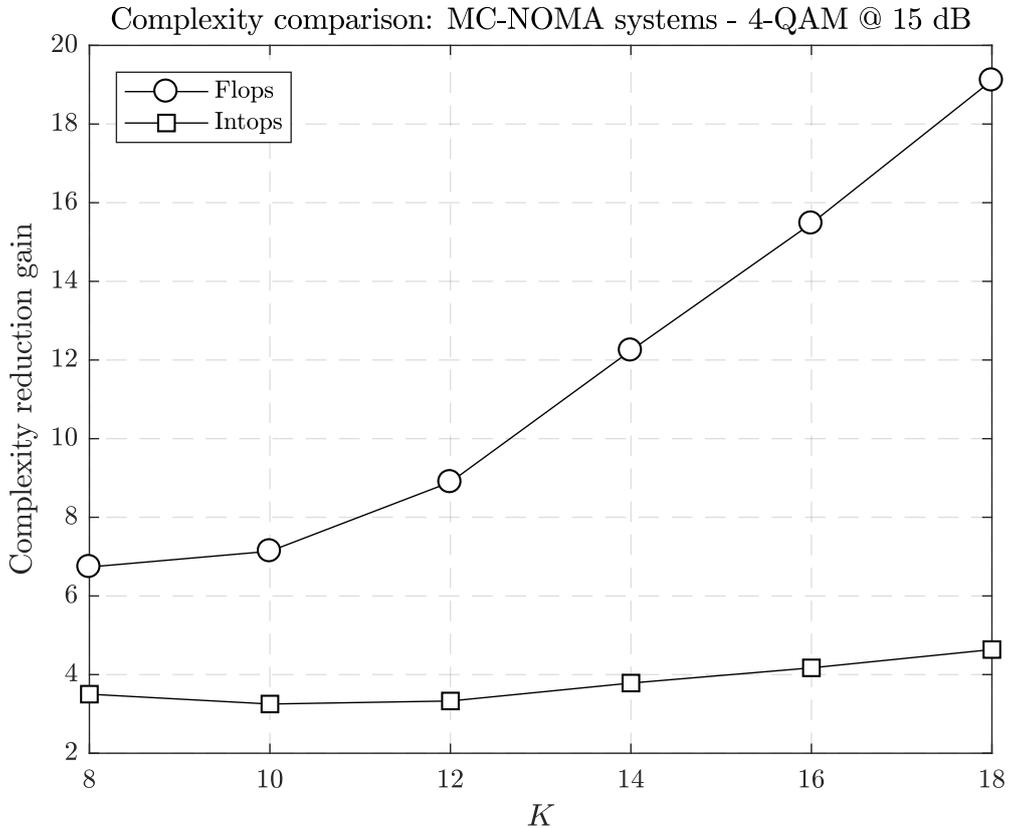
the scheme. Inherently, the vertical scalability of the overloading has been insofar addressed and actually outlined through Figures 5.19 - 5.21, as the overloading was increased in fine steps from 150% to 300%.

To provide a complete description of scalability, the performance of Rubik-PTST-decoded MC-NOMA systems of size $(M, M + 2)$ with $M \in \{6, 8, 10, 12, 14, 16\}$ was analyzed in terms of computational complexity. The results thus obtained, under a fixed SNR level of 15 dB, are displayed in Figure 5.22 to offer an analogous perspective to that of Figure 5.19 from the perspective of horizontal scaling, *i.e.*, the increase of the MC-NOMA system under a constant delta between users and REs, or alternatively, under a constant overloading ratio. In comparison to Figure 5.19, Figure 5.22 displays on one hand higher gains in the relative complexity reduction achieved by Rubik-PTST over the SESD optimized GSD, especially in the case of flops which are greatly reduced by the proposed optimizations for Rubik-PTST. On the other hand, this behavior is also visible in the absolute number of flops and intops, to the level that for instance a MC-NOMA system of $(M, K) = (16, 18)$ is jointly decodable within the order of sub-milliseconds given modern floating-point enabled processing units with typical GHz running frequencies.

These results of performance and complexity in terms of both vertical and horizontal scalability can be detailed as well through the prism of frame-theoretic constructs such as mutual coherence. Concretely, the vertical increase in overloading under keeping constant the number of REs leads to an increase in the worst-case coherence of the associated MC-NOMA precoding



(a) Average number of operations.



(b) Complexity reduction gain.

Figure 5.22: Complexity comparison of Rubik-PTST and SEDS optimized GSD receivers for MC-NOMA systems (M, K) , such that $M = \{6, 8, 10, 12, 14, 16\}$, and respectively, $K = M + 2$. ©2019 IEEE

frame just by means of the WB (3.68). Oppositely, given the same WB argumentation, increasing the absolute number of REs and maintaining the overloading ratio constant, or alternatively, preserving the delta $K - M$ leads to the decrease of mutual coherence of the spreading frame implementing the MC-NOMA scheme. By design, \mathbf{F} as a MC-NOMA frame exploits these mathematical via QCSIDCO/CSIDCO frame decorrelation upon Algorithm 3. Effectively, MC-NOMA minimizes therefore the maximum, but also average inter-user interference level as seen in Subsection 5.4.2, which in turn benefits the proposed tree-based detection upon its embedded search optimizations.

Thus, on one hand, the designed Rubik-PTST joint detector achieves low computational complexity (approaching the SD promised average polynomial times [195]) for larger systems with minimal overloading, as seen in Figure 5.22. On the other hand, it becomes however increasingly slower for highly overloaded systems with computational time diverging to $\mathcal{O}(Q^{K-M})$ complexity in the light of the search necessary for the overloaded user symbols as observed in Figure 5.19. However, despite the dimensionality, the Rubik-PTST detector maintains its ML-like performance decoding as witnessed above. These characteristics make this joint detection scheme together with the proposed MC-NOMA feasible compelling candidates for high performance URLLC, taking into consideration however the caveats detailed above regarding system dimensionality.

Finally, we can remark that the Rubik-PTST receiver solves Question Q3 to the extent that joint MUD ML performance is achievable even under practical time constraints for small systems with arbitrary overloading, and respectively, for moderate systems with reduced overloading. Hence, Question Q4 is partially addressed as well, but not yet fully closed, which is the objective of the sequel.

5.5.2 Sparsely-structured Massive Scale Joint Detector

Based upon the ending remarks of the previous subsection, a joint detector applicable to massive MC-NOMA would be highly desirable to solve practically the associated joint detection problem. The motivation behind this goal is in fact to fill the gap left by the Rubik-PTST joint detector which is more advantageous for low and moderate sized MC-NOMA scenarios upon the caveats discussed previously at the end of Subsection 5.5.1.

In the sequel, the joint detection problem is researched from a sparsely-structured perspective of joint signal recovery. This reinterpretation is solved effectively based on an ℓ_0 -convexized formulation via FPG, similar in some aspects to the mmWave channel estimation scheme detailed in Section 4.5.

General optimization problem from an ℓ_0 -norm perspective

Let us start by firstly reformulating the optimum ML problem formulation through the prism of sparsity inspired by the general idea presented in [218]. Consider for this purpose the

trivial inequality

$$\sum_{i=1}^Q \|\mathbf{s} - q_i \mathbf{1}_K\|_0 \geq KQ - K, \quad \forall \mathbf{s} \in \mathbb{C}^K, q_i \in \mathcal{S} \quad (5.65)$$

where by q_i each one of the Q discrete constellation points in \mathcal{S} is hereby denoted.

Given this inequality the search space of the ML problem (5.45) over a K -dimensional discrete complex-valued constellations can be expanded to the entire complex-valued domain \mathbb{C}^K . Consequently, the original ML problem (5.45) is revisited and hereby transformed as

$$\hat{\mathbf{s}}_{\text{ML}} = \underset{\mathbf{s} \in \mathbb{C}^K}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2, \quad (5.66a)$$

$$\text{subject to} \quad \sum_{i=1}^Q \|\mathbf{s} - q_i \mathbf{1}_K\|_0 = KQ - K, \quad q_i \in \mathcal{S}. \quad (5.66b)$$

To this end, it is important to note the fact that the equality in (5.65) is met just for the particular realizations of $\mathbf{s} \in \mathcal{S}^K$. Moreover, despite the fact that $KQ - K$ is the global minimum value of the left-hand-side of (5.65), this is obviously achieved non-uniquely for *all* $\mathbf{s} \in \mathcal{S}^K$, such that one could refer to the above inequality proposed as *the discrete search ℓ_0 inequality*. Leveraging these facts, the equivalent sparse formulation of the optimum ML detection for MC-NOMA systems is obtained in (5.66). This form, previously discussed within [218], offers therefore the basis of our next proposed detector.

Despite the present form (5.66) of the optimum detection problem, its ℓ_0 -norm constraint (5.66b) still makes it computationally intractable by preserving the exponential search complexity of the ML detection problem. To solve this conundrum, the system is firstly mapped to its real-valued equivalent by means of (5.53) - (5.55). As a direct consequence, the dimensionality of the involved signals is doubled for the real-valued case, whereas the original notation is kept for simplicity. Furthermore, as part of this complex-to-real transform, the constellation \mathcal{S} is projected onto its real and imaginary axes as $\mathcal{S}_{P,\mathcal{R}}$ and $\mathcal{S}_{P,\mathcal{I}}$, respectively. Hereafter, assume for sake of discussion and without loss of generality the original input constellations to be the same Q -sized QAM across all active users, such that upon their projections the real and imaginary components coincide. Then, $\mathcal{S}_{P,\mathcal{R}} = \mathcal{S}_{P,\mathcal{I}}$ is simply denoted as the $\frac{Q}{2}$ -sized \mathcal{S}_P equivalent projected constellation points corresponding the original Q -sized \mathcal{S} inputs available for each user. To eliminate any confusion following the complex-to-real transformation related to the problem various dimensions involved, problem (5.66) is rewritten from its real-valued, (5.53) - (5.55), equivalent perspective,

$$\hat{\mathbf{s}}_{\text{ML}} = \underset{\mathbf{s} \in \mathbb{R}^{2K}}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2, \quad (5.67a)$$

$$\text{subject to} \quad \sum_{i=1}^{\frac{Q}{2}} \|\mathbf{s} - q_i \mathbf{1}_{2K}\|_0 = KQ - 2K, \quad q_i \in \mathcal{S}_P. \quad (5.67b)$$

An equivalent sparse representation, \mathbf{x} , of the signal \mathbf{s} involved in the problem (5.67) can next be obtained based on the discrete search ℓ_0 inequality (5.65) via the 1-bit binary indicator

quantizer [218]

$$\mathbb{I}_{\mathcal{S}_P}(\mathbf{x}_k) \triangleq \begin{cases} 1 & \text{for } \mathbf{x}_k \in \mathcal{S}_P, \\ 0 & \text{for } \mathbf{x}_k \notin \mathcal{S}_P, \end{cases} \quad (5.68)$$

applied element-wise to \mathbf{x} , such that the original signal \mathbf{s} is yielded back by

$$\mathbf{s} = \mathbf{C}_{\mathcal{S}_P} \mathbf{x}, \quad \mathbf{x} \in \{0, 1\}^{KQ}. \quad (5.69)$$

The desparsifying synthesis matrix in the Equation (5.69) is defined as

$$\mathbf{C}_{\mathcal{S}_P} \triangleq \mathbf{I}_{2K} \otimes \mathbf{q}^T \quad (5.70)$$

as a band-diagonal matrix with a bandwidth of $\frac{Q}{2}$ containing the uniquely available constellation points from \mathcal{S}_P stacked as a column vector, *i.e.*, \mathbf{q} .

Remark that the *hard-quantization* (5.68) and binary representation \mathbf{x} of \mathbf{s} just offer a compact way to rewrite the constraint (5.66b), or equivalently (5.67b), and as a consequence, do not alleviate the exponential search characteristic of the considered problem. To mitigate this issue and provide computational tractability to the joint detection of massive MC-NOMA schemes, the relaxation of probabilistic *soft-quantization* of the hard method above is hereby proposed as fulfilling together

$$\mathbf{s} = \mathbf{C}_{\mathcal{S}_P} \mathbf{x}, \quad (5.71a)$$

$$\mathbf{1}_{2K} = \mathbf{C}_1 \mathbf{x} \quad (5.71b)$$

for real-valued box-constrained $\mathbf{x} \in [0, 1]^{KQ}$, and respectively, with $\mathbf{C}_1 \triangleq \mathbf{I}_{2K} \otimes \mathbf{1}_{Q/2}^T$.

The condition (5.71a) follows directly from the binary sparse description of \mathbf{s} from (5.69). Secondly, upon $\mathbf{x} \in [0, 1]^{KQ}$, the condition (5.71) becomes a probabilistic normalization ensuring that the soft values in $[0, 1]$ for the entries of \mathbf{x} corresponding to the same signal symbol in \mathbf{s} are sum-normalized to 1. Coherently, these two constraints yield therefore a \mathbf{x} -weighted and normalized synthesis of each signal symbol in \mathbf{s} based upon the discrete values of the available constellations \mathcal{S}_P .

Following these remarks, this probabilistic soft-quantization is applied to the problem (5.67) to yield the ML-equivalent reformulated optimization problem in terms of \mathbf{x} as

$$\underset{\mathbf{x} \in [0, 1]^{KQ}}{\operatorname{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{C}_{\mathcal{S}_P} \mathbf{x}\|_2^2 \quad (5.72a)$$

$$\text{subject to} \quad \|\mathbf{x}\|_0 = 2K, \quad (5.72b)$$

$$\mathbf{C}_1 \mathbf{x} = \mathbf{1}_{2K}. \quad (5.72c)$$

Note that the condition forcing the ML performance in program 5.72 is in fact constraint (5.72b) given that $\mathbf{x} \in [0, 1]^{KQ}$. Unfortunately, the involved ℓ_0 -norm is not convex and, as a result, the problem in (5.72) is not yet a convex program that could easily be optimized, relying in fact on search-based solvers as the Rubik-PTST joint detector previously proposed in

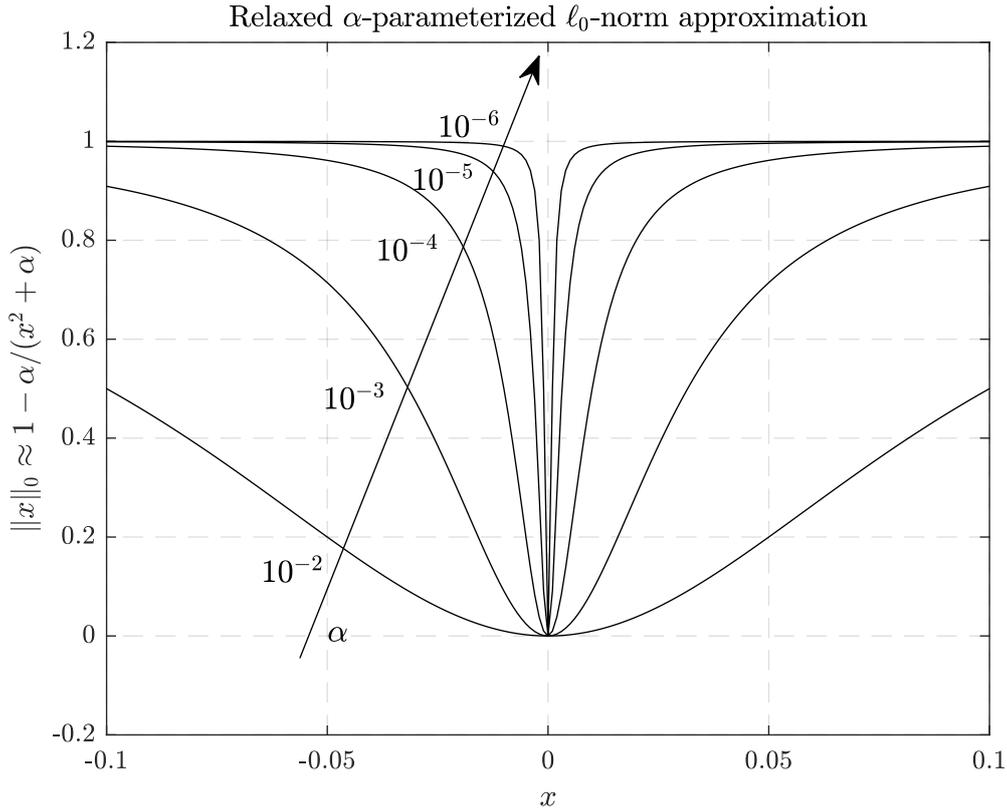


Figure 5.23: Smooth and differentiable α -parameterized approximation of ℓ_0 -norm.

Subsection 5.5.1. However, since the goal is to be able to jointly detect MC-NOMA schemes, let's break this deadlock and relax the problem (5.72), sacrificing some of the performance in favor of computational efficiency and formulation versatility. Thus, we aim to solve the regularized version of the problem (5.72), *i.e.*,

$$\underset{\mathbf{x} \in [0,1]^{KQ}}{\operatorname{argmin}} \quad \lambda \|\mathbf{x}\|_0 + \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{C}_{S_P} \mathbf{x}\|_2^2 \quad (5.73a)$$

$$\text{subject to} \quad \mathbf{C}_1 \mathbf{x} = \mathbf{1}_{2K} \quad (5.73b)$$

At this point we can remark the ℓ_0 -central formulation of the objective which we also similarly encountered in the case of mmWave channel estimation. As detailed there, the toolkit of FPG [149] is used here once more to solve (5.73) via the QT technique applied on top of an approximated ℓ_0 -norm. To this end, consider here instead the smooth and differentiable³ quadratic-based approximation

$$\|\mathbf{x}\|_0 = KQ - \lim_{\alpha \rightarrow 0^+} \sum_{i=1}^{KQ} \frac{\alpha}{x_i^2 + \alpha}, \quad \forall \mathbf{x} \in \mathbb{R}^{KQ}, \alpha \in \mathbb{R}_+, \quad (5.74)$$

illustrated for the sake of clarity in Figure 5.23 across an unidimensional planar section.

³It is important to note that albeit inspired by approximation (4.65) and [148], the approximation proposed here is different as it is using a smooth differentiable function of \mathbf{x} , *i.e.*, the squared of the individual entries, instead of the non-smooth and universally non-differentiable absolute value employed in Chapter 4.

Provided that the scalar parameter $0 < \alpha \ll 1$ and the denominators in the ratios above are convex smooth functions in x_i , the approximation in equation (5.74) can be rewritten using the QT of FPG [149] as

$$\|\mathbf{x}\|_0 \approx KQ - \left(\sum_{i=1}^{KQ} 2\beta_i \sqrt{\alpha} - \beta_i^2 (x_i^2 + \alpha) \right), \quad (5.75)$$

where β_i is defined in the light of the QT [149], following (4.68), as

$$\beta_i \triangleq \frac{\sqrt{\alpha}}{x_i^2 + \alpha}, \quad \forall i. \quad (5.76)$$

The approximation (5.75) can thus be easily vectorized under the quadratic form

$$\|\mathbf{x}\|_0 \approx KQ - 2\sqrt{\alpha} \text{trace}(\mathbf{B}) + \alpha \text{trace}(\mathbf{B}^T \mathbf{B}) + \mathbf{x}^T \mathbf{B}^T \mathbf{B} \mathbf{x}, \quad (5.77)$$

where the matrix $\mathbf{B} \triangleq \text{diag}(\beta_1, \beta_2, \dots, \beta_{KQ})$.

Given all of the above, the original NP-hard ℓ_0 -constrained detection problem (5.66) is subsequently relaxed by means of FPG and QT to its convex equivalent mixed-quadratic form

$$\underset{\mathbf{x} \in [0,1]^{KQ}}{\text{argmin}} \quad \lambda (\mathbf{x}^T \mathbf{B}^T \mathbf{B} \mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{H} \mathbf{C}_q \mathbf{x}\|_2^2, \quad (5.78a)$$

$$\text{subject to} \quad \mathbf{C}_1 \mathbf{x} = \mathbf{1}_{2K}, \quad (5.78b)$$

such that the constant term (in terms of \mathbf{x}) from (5.77), $KQ - 2\sqrt{\alpha} \text{trace}(\mathbf{B}) + \alpha \text{trace}(\mathbf{B}^T \mathbf{B})$, has been ignored as it bears no role on the optimization solution.

Furthermore, in the light of the QT associated with [149, Th. 3] and the fact that the ℓ_0 -approximation proposed entails a concave-convex sum-ratio, it follows that the convex program (5.78) converges iteratively to a stationary point. In other words, the problem is solved iteratively until convergence or a maximum number of iterations is reached.

The formulation of (5.78) is already cast into the CVX disciplined convex programming ruleset of [189], and so, the CVX [150] semi-definite programming associated solvers can be used directly. Nevertheless, this strategy albeit relatively computationally efficient for small-scale optimizations becomes very slow [150] for large problems such as the one that is considered, *e.g.* $20 \ll M < K$. To efficiently solve the latter scenarios, the ADMM implementation of (5.78) is proposed next.

Proximal splitting decomposition via ADMM

The ADMM toolbox [219] is a fast general solver recipe for problems of the type

$$\underset{\mathbf{x}, \mathbf{z}}{\text{minimize}} \quad f(\mathbf{x}) + g(\mathbf{z}) \quad (5.79a)$$

$$\text{subject to} \quad \mathbf{D}_x \mathbf{x} + \mathbf{D}_z \mathbf{z} - \mathbf{c} = \mathbf{0}, \quad (5.79b)$$

where the real-valued functions $f(\mathbf{x})$ and $g(\mathbf{z})$ are required to be closed, proper and convex [219].

Note that the above conditions do not restrict f, g to be finite-valued or differentiable, whereas no assumptions of $\mathbf{D}_x, \mathbf{D}_z$ are explicitly implied. Under these loose conditions, the ADMM ensures convergence to a possibly non-unique solution given the iterations [219]

$$\mathbf{x}^{(i+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \mathcal{L}_\rho(\mathbf{x}, \mathbf{z}^{(i)}, \mathbf{v}^{(i)}), \quad (5.80a)$$

$$\mathbf{z}^{(i+1)} = \underset{\mathbf{z}}{\operatorname{argmin}} \mathcal{L}_\rho(\mathbf{x}^{(i+1)}, \mathbf{z}, \mathbf{v}^{(i)}), \quad (5.80b)$$

$$\mathbf{v}^{(i+1)} = \mathbf{v}^{(i)} + \rho(\mathbf{D}_x \mathbf{x}^{(i+1)} + \mathbf{D}_z \mathbf{z}^{(i+1)} - \mathbf{c}), \quad (5.80c)$$

where \mathbf{v} denotes the dual variable associated with the augmented Lagrangian function

$$\mathcal{L}_\rho(\mathbf{x}, \mathbf{z}, \mathbf{v}) = f(\mathbf{x}) + g(\mathbf{z}) + \mathbf{v}^\top (\mathbf{D}_x \mathbf{x} + \mathbf{D}_z \mathbf{z} - \mathbf{c}) + \frac{\rho}{2} \|\mathbf{D}_x \mathbf{x} + \mathbf{D}_z \mathbf{z} - \mathbf{c}\|_2^2 \quad (5.81)$$

parameterized by $\rho > 0$.

Therefore, the optimization problem (5.78) can be directly translated into the canonical ADMM formulation as detailed next. Firstly, let us accommodate the constraint $\mathbf{x} \in [0, 1]^{KQ}$ by introducing the set indicator function

$$\iota_{[0,1]^{KQ}}(\mathbf{x}) = \begin{cases} +\infty & \text{for } \mathbf{x} \notin [0, 1]^{KQ}, \\ 0 & \text{for } \mathbf{x} \in [0, 1]^{KQ}. \end{cases} \quad (5.82)$$

Using the latter, and expanding the objective function (5.78a) to its quadratic form in \mathbf{x} , the functions f, g are defined secondly as

$$f(\mathbf{x}) \triangleq \mathbf{x}^\top \underbrace{\left(\lambda \mathbf{B}^\top \mathbf{B} + \frac{1}{2} \mathbf{C}_q^\top \mathbf{H}^\top \mathbf{H} \mathbf{C}_q \right)}_{\triangleq \mathbf{P}} \mathbf{x} - \mathbf{y}^\top \mathbf{H} \mathbf{C}_q \mathbf{x} + \frac{1}{2} \mathbf{y}^\top \mathbf{y} \quad (5.83a)$$

$$g(\mathbf{z}) \triangleq \iota_{[0,1]^{KQ}}(\mathbf{z}), \quad (5.83b)$$

such that the additional affine constraint $\mathbf{x} - \mathbf{z} = \mathbf{0}$ is necessary.

Given the above notation and logic, the optimization problem of (5.78) is thirdly transformed under the ADMM canonical form as the equivalent problem

$$\underset{\mathbf{x}, \mathbf{z} \in \mathbb{R}^{KQ}}{\operatorname{minimize}} \quad f(\mathbf{x}) + g(\mathbf{z}) \quad (5.84a)$$

$$\text{subject to} \quad \mathbf{x} - \mathbf{z} = \mathbf{0}, \quad (5.84b)$$

$$\mathbf{C}_1 \mathbf{x} - \mathbf{c} = \mathbf{0}, \quad (5.84c)$$

where for exposition's sake we implicitly denoted $\mathbf{c} \triangleq \mathbf{1}_{2K}$.

Lastly, the augmented Lagrangian function associated with the ADMM optimization prob-

lem formulated by (5.84) is calculated as

$$\begin{aligned}\mathcal{L}_\rho(\mathbf{x}, \mathbf{z}, \mathbf{v}) &= \mathbf{x}^\top \mathbf{P} \mathbf{x} - \mathbf{y}^\top \mathbf{H} \mathbf{C}_q \mathbf{x} + \frac{1}{2} \mathbf{y}^\top \mathbf{y} + \iota_{[0,1]^{KQ}}(\mathbf{z}) \\ &+ \mathbf{v}^\top (\mathbf{x} - \mathbf{z}) + \frac{\rho}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 \\ &+ \mathbf{u}^\top (\mathbf{C}_1 \mathbf{x} - \mathbf{c}) + \frac{\rho}{2} \|\mathbf{C}_1 \mathbf{x} - \mathbf{c}\|_2^2.\end{aligned}\quad (5.85)$$

Based upon the above reformulation, the optimization problem can be solved efficiently via ADMM iterations since the convergence conditions are met, [219]. Concretely, given $f(\mathbf{x}), g(\mathbf{z})$ defined in (5.83), the smoothly differentiable augmented Lagrangian (5.85) and following elementary multi-variate calculus, the solution \mathbf{x} is computed by the closed-form iterations

$$\mathbf{x}^{(i+1)} = (2\mathbf{P} + \rho(\mathbf{I}_{KQ} + \mathbf{C}_1^\top \mathbf{C}_1))^{-1} (\mathbf{C}_q^\top \mathbf{H}^\top \mathbf{y} + \rho(\mathbf{z}^{(i)} + \mathbf{C}_1^\top \mathbf{c}) - \mathbf{v}^{(i)} - \mathbf{C}_1^\top \mathbf{u}^{(i)}) \quad (5.86a)$$

$$\mathbf{z}^{(i+1)} = \frac{1}{\rho} (\rho \mathbf{x}^{(i+1)} + \mathbf{v}^{(i)}) \quad (5.86b)$$

$$\mathbf{v}^{(i+1)} = \mathbf{v}^{(i)} + \rho (\mathbf{x}^{(i+1)} - \mathbf{z}^{(i+1)}) \quad (5.86c)$$

$$\mathbf{u}^{(i+1)} = \mathbf{u}^{(i)} + \rho (\mathbf{C}_1 \mathbf{x}^{(i+1)} - \mathbf{c}). \quad (5.86d)$$

Algorithm 7 SPARSE-STRUCTURED ℓ_0 -FPG MASSIVE MUD VIA ADMM

Inputs: $\mathbf{H} \in \mathbb{R}^{2M \times 2K}$, $\mathbf{y} \in \mathbb{R}^{2M}$, $\rho, \lambda > 0$, $Q/2$ -sized real constell. \mathcal{S}_P , max. iterations O

Output: $\hat{\mathbf{s}} \in \mathbb{C}^K$

- 1: Set max. inner loops $I = 500$ and outer loops $O_{\max} = O$.
 - 2: Set $\mathbf{x}_s^{(0)} = \frac{2}{Q} \mathbf{1}_{KQ}$, outer loop counter $o = 0$, $\delta_o = +\infty$ and convergence threshold $\gamma = 10^{-6}$.
 - 3: **while** ($o < O_{\max}$ and $\delta_o > \gamma$) **do**
 - 4: Compute $\beta_i, \forall i$ and \mathbf{B} by (5.76) for $\mathbf{x}_s^{(o)}$.
 - 5: Compute \mathbf{P} by (5.83) and cache the inverse $(2\mathbf{P} + \rho(\mathbf{I}_{KQ} + \mathbf{C}_1^\top \mathbf{C}_1))^{-1}$.
 - 6: Generate randomly & uniformly $\mathbf{x}^{(0)}, \mathbf{z}^{(0)} \in [0, 1]^{KQ}$.
 - 7: Set $\mathbf{v}^{(0)} = \mathbf{0}$, $\mathbf{u}^{(0)} = \mathbf{0}$.
 - 8: Set inner loop counter $i = 1$, $\delta_i = +\infty$.
 - 9: **while** ($i < I$ and $\delta_i > \gamma$) **do**
 - 10: Compute $\mathbf{x}^{(i+1)}, \mathbf{z}^{(i+1)}, \mathbf{v}^{(i+1)}, \mathbf{u}^{(i+1)}$ given (5.86).
 - 11: Calculate $\delta_i \triangleq \|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\|_2$.
 - 12: Increase loop counter $i = i + 1$.
 - 13: **end while**
 - 14: Increase loop counter $o = o + 1$.
 - 15: Set estimate to current solution $\mathbf{x}_s^{(o)} = \mathbf{x}^{(i)}$.
 - 16: Calculate $\delta_o \triangleq \|\mathbf{x}_s^{(o)} - \mathbf{x}_s^{(o-1)}\|_2$.
 - 17: **end while**
 - 18: Recover $\mathbf{s} = \mathbf{C}_{\mathcal{S}_P} \mathbf{x}_s^{(o-1)}$.
 - 19: Remap to discrete symbols (slicing) $\mathbf{s} \xrightarrow{\mathcal{S}_P} \hat{\mathbf{s}}_P$.
 - 20: Recover complex domain signal estimate $\hat{\mathbf{s}}_P \xrightarrow{(5.55)^{-1}} \hat{\mathbf{s}}$.
-

For convenience, the complete implementation of the detection problem described by (5.78) obtained efficiently via the ADMM from (5.84) is summarized as Algorithm 7.

To simulate the expected uncoded BER performance of the proposed detection scheme detailed above, DL MC-NOMA systems were simulated similarly to the case of performance evaluations from the end of Subsection 5.5.1. Throughout the simulations, the ρ, λ algorithm parameters were fixed and set to be $\rho = 0.25$, and $\lambda = 1$, respectively. The BER results for MC-NOMA systems of sizes $(M, K) = (32, 48)$ and $(M, K) = (80, 120)$, both corresponding to an overloading ratio of 150%, are displayed in Figures 5.24, and respectively, 5.25.

For the sake of comparison and benchmarking, the conventional naive reference (given the massive overloading) of MMSE decoder [164] was on one hand considered, while on the other hand the modern convex ℓ_1 -optimization method of sum of absolute value (SOAV) [222] was implemented as a state-of-the-art alternative. In short, the SOAV approach has proved to offer very good performance in the context of MIMO overloaded systems [220, 221], by essentially solving the ℓ_1 -relaxation of the problem, [220],

$$\underset{\mathbf{s} \in \mathbb{R}^{2K}; q_i \in \mathcal{S}_P}{\operatorname{argmin}} \quad \frac{2}{Q} \sum_{i=1}^{\frac{Q}{2}} \|\mathbf{s} - q_i \mathbf{1}_{2K}\|_1, \quad (5.87a)$$

$$\text{subject to} \quad \mathbf{y} = \mathbf{H}\mathbf{s}. \quad (5.87b)$$

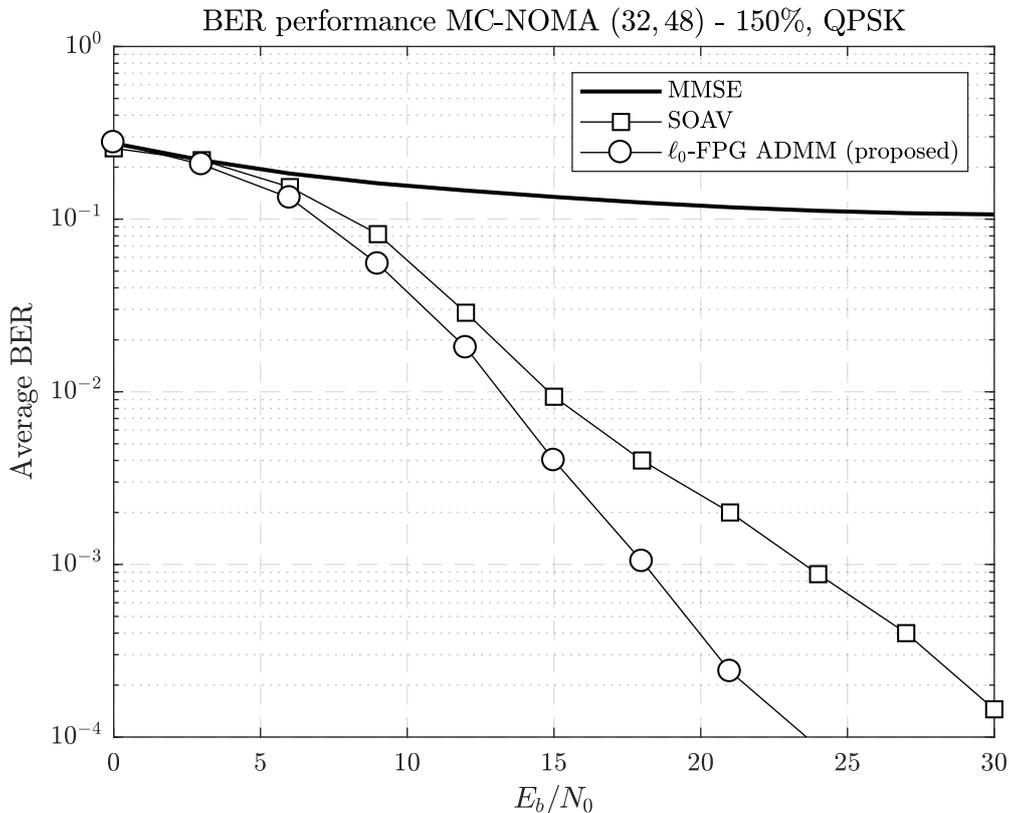


Figure 5.24: Average BER performance of a MC-NOMA system serving $K = 48$ active users over $M = 32$ available resources. Detection is performed via MMSE, SOAV [220, 221] and the proposed ℓ_0 -FPG MUD detector implemented via ADMM as given in Algorithm 7.

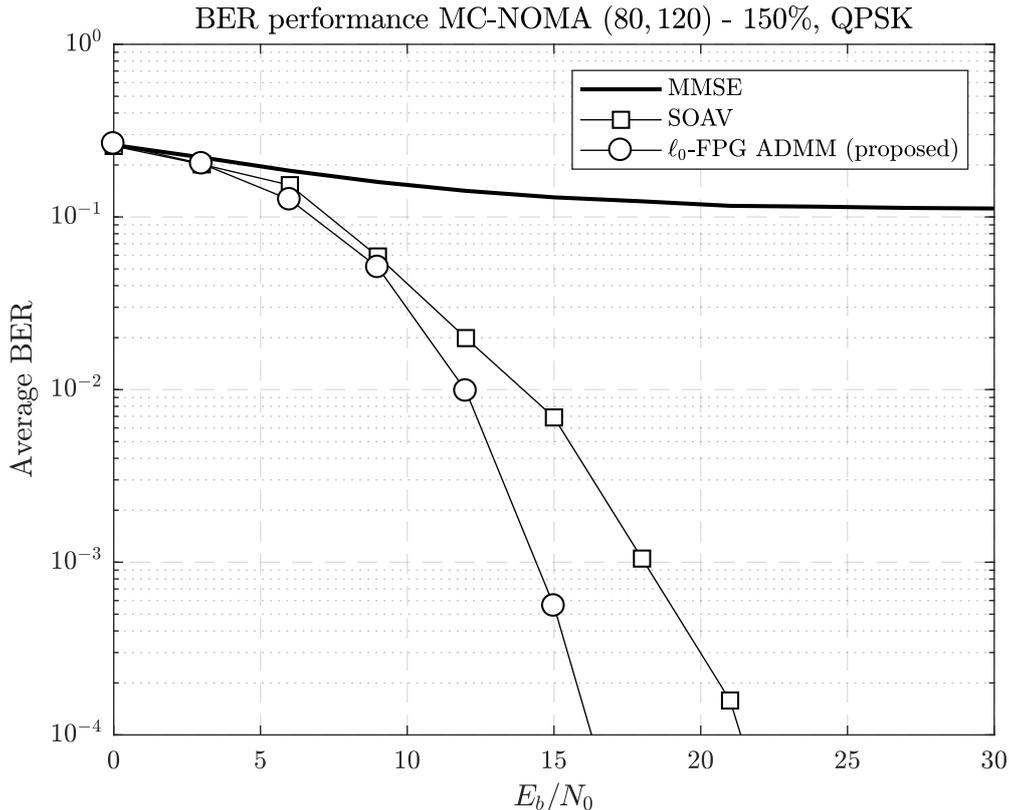


Figure 5.25: Average BER performance of a MC-NOMA system serving $K = 120$ active users over $M = 80$ available resources. Detection is performed via MMSE, SOAV [220, 221] and the proposed ℓ_0 -FPG MUD detector implemented via ADMM as given in Algorithm 7.

For efficient computational complexity, the SOAV and its iteratively weighted variations use the method of proximal splitting as a generalization of ADMM [219], and as a result they leverage the Douglas-Rachford splitting, optimized and derived in fact just for the case of the real constellation points BPSK or 4-QAM [220, 221].

It can be seen firstly that the proposed MC-NOMA multi-user detector manages to successfully recover the user information despite the overloading-induced inter-user interference. Furthermore, the state-of-the-art SOAV decoder is consistently but also significantly outperformed by the introduced method. Additionally, note that between the two Figures, 5.24 and 5.25, the detection performance of both, the ADMM proposed detector and the SOAV detector, increases with the system increase, scaling with the MC-NOMA dimensionality given fixed overloading level. As mentioned earlier in the end of Subsection 5.5.1, this is a feature that benefits MC-NOMA given its incoherent frame-theoretic inherent attributes minimizing the active users' mutual coherence. To this end, the method proposed in Algorithm 7 follows the trend of the state-of-the-art and leverages therefore these traits of MC-NOMA to their fullest, such as BER performance better than the SOAV reference is achieved.

One of the merits of the proposed MUD receiver is its embedded probabilistic nature based upon the introduced soft-quantization of the objective being minimized, see for instance (5.71), and respectively, (5.73). This attribute leads therefore to additional soft-outputs that could be

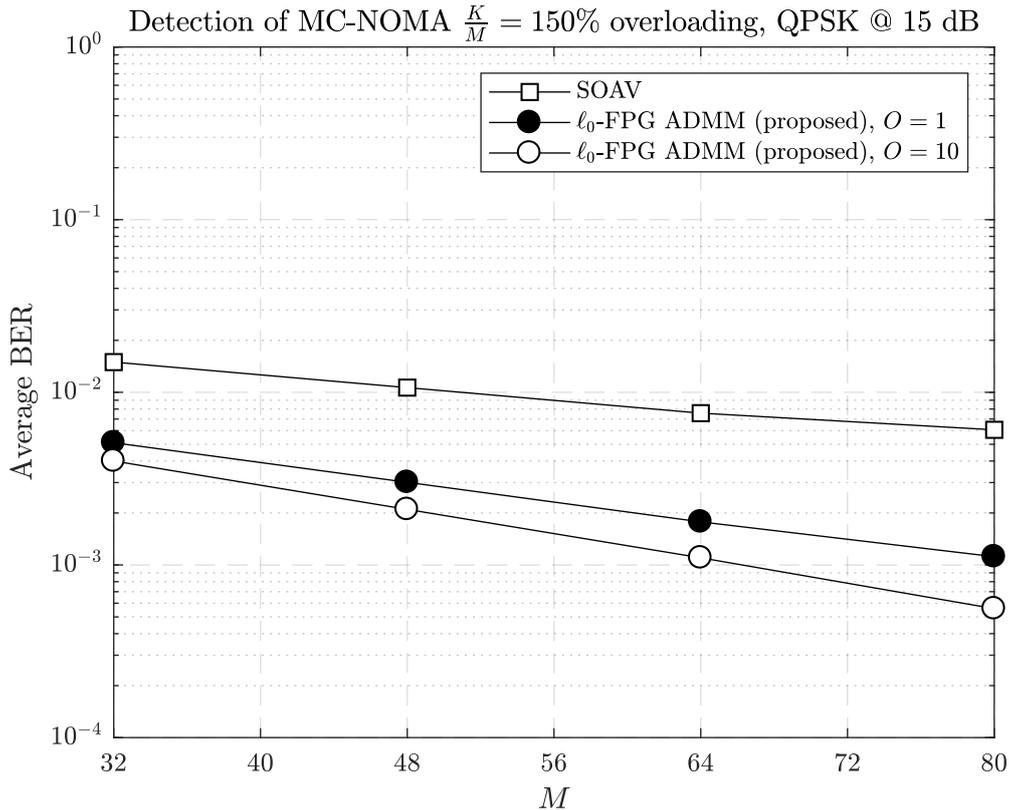


Figure 5.26: Multi-user detection performance for systems with overloading factor of 150% and $M = \{32, 48, 64, 80\}$ available resources. The reference receiver is considered to be the SOAV [220, 221] detector against which the proposed ℓ_0 -FPG detector from Algorithm 7 is compared in its configuration with $O = 1$ or $O = 10$ iterations.

combined with subsequent decoding blocks and feedback mechanisms if necessary.

Yet another key component is represented by the ℓ_0 -norm approximation and its FPG and QT [149] convexization with respect to the original ML problem. These methods not only improve versatility against parameter selections, but also lead to increased recovery robustness [148] in comparison to ℓ_1 methods, such as the SOAV, since they are able to better cover and represent the essential sparsity within the system. This is visible in the case of BER deltas between the Algorithm 7 and the SOAV which increase with larger SNRs.

Furthermore, the relaxed ℓ_0 -norm regularization via FPG and QT is also a sparsity enhancement technique, similar in its iterative solution to the BPDN ℓ_1 -reweighted scheme [145]. The sparsity level is modulated further through the regularization parameter λ , which controls the weight placed on forcing the solution to reach the desired $2K$ -sparse structure of the ML estimation problem. This is achieved iteratively via the outer loop of the Algorithm 7 which acts therefore as a sparsity-based performance enhancer.

This last remark can thus be used in practical deployments leveraging also the soft-outputs available within the receiver to reduce the outer number of iterations for applications where latency is valued more than detection resolution accuracy. Moreover, it can be also linked to the MC-NOMA large-scale reduced inter-user interference argument to advocate further the scalability of the proposed art. To illustrate this point, Figure 5.26 was rendered outlining the

performance decay between running Algorithm 7 with $O = 10$, and respectively, $O = 1$ across increasing MC-NOMA systems with fixed overloading.

To comment on the complexity of the detector, let us have a look at one of the outer loops of Algorithm 7. It is clear that most complex operation is in fact the square matrix inversion $(2\mathbf{P} + \rho(\mathbf{I}_{KQ} + \mathbf{C}_1^T \mathbf{C}_1))^{-1}$ of dimensionality $KQ \times KQ$. For this reason, the inversion has been moreover cached prior to the inner ADMM loop to avoid multiplicative effects on performance. This is generally possible as the matrices $\mathbf{P}, \mathbf{B}, \mathbf{C}_1$ are constant over the scope of the outer loop of the Algorithm 7. In turn, the complexity of the inner ADMM loop is governed just by simple algebraic multiplications of matrices with vectors. Consequently, it follows that the computational complexity of the proposed joint detector in its ADMM implementation from Algorithm 7 is of order $\mathcal{O}(O \cdot (KQ)^3)$, similar in fact to the SOAV complexity [220].

All in all, the proposed sparsely-structured ℓ_0 -FPG multi-user receiver provides a scalable solution for massive MC-NOMA systems via its ADMM implementation from Algorithm 7.

5.6 Conclusions

The problem of non-orthogonal user multiplexing has been addressed generically for both DL and UL cases under an unified frame-theoretic interpretation of the central NOMA information transmission. This abstract formulation led to the identification of the NOMA active user multiplexing through the prism of a precoding NOMA frame and its subsequent design. In turn, the optimum design was analyzed in terms of achievable system sum-rates by means of ergodic capacity expressions for both DL and UL general CD-NOMA scenarios. Having formulated these expressions analytically and having analyzed their maximization potentials, the main design criteria for the precoding frame have been derived, and given these, an optimal frame design mechanism has been proposed as MC-NOMA.

The MC-NOMA philosophy is based therefore on spreading all the active users over all the available REs, thus reaching fully dense precoding and waveform signatures over each user. Thus, instead of sparse access patterns or low-density codes used to manage inter-user interference, MC-NOMA advocates in turn to densely optimized user incoherence requiring less attention and effort for the design of sparsity and sparse codebooks. In fact, the design of MC-NOMA may be summarized to the generation of dense, incoherent UNTFs, which we solved by successive decorrelation via QCSIDCO and tightening of arbitrary random structures.

Provided all of these, the superiority of the proposed scheme was demonstrated not only theoretically, but also numerically in comparison to two of the best state-of-the-art CD-NOMA schemes, *i.e.*, SCMA and PD-NOMA, highlighting its performance gain in terms of ergodic sum-rates for both continuous and discrete channel inputs.

Having completed the design of the ideal CD-NOMA Tx as MC-NOMA, attention was turned to the practical feasibility of detection and information decoding. To support the promised achievable sum-rates and gains firstly the implementation of a joint ML-capable MUD was considered. This led to the introduction of the Rubik-PTST receiver as an optimized tree search low-complexity ML-reaching joint detector. For reducing the computational effort of

Rubik-PTST, two strategies were used: the algebraic search space reduction, and respectively, the probabilistic tree pruning by hard thresholding based on the noise statistics. These proved efficient in speeding up the joint optimum detection for MC-NOMA, making it in fact practically feasible for small systems with high overloading, but also to medium-sized systems with low-levels of overloading.

Finally, to outline the scalability of MC-NOMA the problem of massive joint detection was studied. To this end, an alternative probabilistic sparsely-structured ℓ_0 -norm formulation of the ML joint detection problem was introduced. This was subsequently relaxed by means of smooth and differentiable ℓ_0 -approximation and FPG techniques to a simple quadratic program efficiently solved via ADMM up to tractable cubic worst-time complexity. The proposed sparsely-structured ℓ_0 -FPG joint detector via ADMM outlined good and scalable performance with the size of the MC-NOMA systems considered.

Chapter 6

The 6G Horizon – Sense the Waveforms

Over the last couple of chapters frame-theoretic designs for systems pertaining to emerging technologies across 5G and 5G+ have been discussed in mathematical details. Furthermore, as briefly mentioned in the introductory Section 1.3, these may become as well candidates for future iterations of wireless communications collapsed under what may be defined as 6G.

Of course, one may not know yet what will the sixth generation of wireless system entail. Nevertheless, as taught by the short but rich history of cellular and wireless communications explored succinctly in Chapter 1, one may anticipate disruptive trends and try to predict as an educated guess what would be the directions to be pursued next.

This chapter is in these regards a “soft” *humble personal depiction*, by no means exhaustive, of such disruptions that will transform 5G into a more stable, and established successor, *i.e.* 6G. In this sense, it becomes a daring “letter to the future,” serving as hard comparison between the current expectations, and respectively, the reality achieved in 20-30 years time in the realm of wireless communications.

6.1 Trends and Future Directions

During the course of last years, 5G technologies and standardizations have been carried by both academia and industry alike, converging to a, so called, new radio (NR) specification currently to be released under Release 16 [10]. This effort would be practically concluded by the release of large-scale rollouts of 5G networks expected for late 2019. Obviously, such deployments will in fact be limited to a subset of the fully envisioned functionality of the 5G standard, being subsequently advanced by updates and evolutions of new features which are

Excerpts of this chapter are adapted and enhanced based on the article:

R.-A. Stoica, G. T. Freitas de Abreu, “6G: the Wireless Communications Network for Collaborative and AI Applications,” *arXiv preprint arXiv:1904.03413*, Apr. 2019.

addressed already in specialized circles as 5G+.

A great merit of the fifth iteration of wireless networks, beyond the hyped promised gains in communications rates, access and reliability, is actually represented by the architectural paradigm shift with respect to its predecessor, which in fact enabled all these gains. To this end, the 5G development has been designed from a ground-up perspective with a focus on solving existent vertical applications or empowering new ones.

One key design pattern at the core of 5G was the rethinking of the traditional hierarchical layered network architecture into a more flexible format, capable of functional virtualization and network aggregation [9]. Another major deviation was the resolution to extend the spectrum above the canonical sub-6 GHz bands utilized by the previous systems and complement these by mmWave and higher frequencies, leading to a tenfold increase of available bandwidth. From a connectivity density perspective, 5G is based on massive MIMO and more advanced MIMO architectures than those featured in 4G which increase Tx/Rx diversity and simultaneously offer an enhanced directional communications potential [223].

Moreover, these points are complemented in terms of access by carrier aggregation and MUST techniques imported from 4G LTE [153], expectations being that more decisive non-orthogonal concurrent access schemes for both UL and DL would eventually find their way into the later stages of standardization as enhancements. All of these *soft* changes may also be complemented by the *hard* strides made on the RF hardware front in going beyond MIMO via wireless full-duplex capabilities [224–227].

Thus, analyzing these technologies and their current integration within 5G, one may naturally ask in effect: “*what else could we want and why even bother thinking about 6G at all ?*”. The simple answer to these otherwise well-anchored questions is that in fact *revolution never comes from within, but is rather imposed by radical changes in exterior conditions*. Such a radical change, preparing to emerge at a large scale is given by the raise of AI and its propagation in different business verticals. As briefly introduced in Section 1.2, we expect this to cause a major disruption towards the way wireless communications are currently addressed by 5G, acting as a catalyst to future advances which we categorize as central to 6G.

In Chapter 1, the idea that 5G and 6G address together the wireless networking towards the *organic Internet access* for both conventional people users and machines/intelligent agents alike was introduced, see right section of Figure 1.2. For this purpose and in the light of the proliferation of various AI applications, the trend of diversified services with highly heterogeneous requirements of access, rate, latency or energy consumption will certainly only increase.

6.2 Candidate Technologies and Expectations

AI is currently establishing itself both through scientific advancements [11], but also via publicized and marketed hype as one of the most important emerging technologies during the last decades. Undoubtedly, computational intelligence bears definitive prospects in unlocking problems that have proved so far difficult to solve. It is true that most of the time, AI and its embedded component of machine learning are very useful in fact to generate solutions where

human understanding and collective knowledge of research have failed to propose viable highly defined models. In addition to extracting such models based on big training data sets, AI is further useful in evolving existent designs or states based on adapting to experienced events or ingested data which gives it a scalable applicability in practice over abstract problems essentially dispersed across a large palette of fields.

Similarly, AI has the potential of developing high-performance collective applications beyond what expert-based strategies and protocols could achieve. A recent example of *collaborative intelligence* is in fact the use of five AI entities by OpenAI Five to beat a team of world-class players in the strategic complex multi-player video game of Dota[®] 2 [228]. One key aspect of this recent “success”, besides the advances in deep reinforcement learning [229], relies in the ability of multiple neural network artificial agents to cooperate in real-time and to outperform the human state-of-the-art.

The key point in the examples above, the cooperation, was achieved of course at local level by just observing similar inputs aggregated over the states of the game, and respectively, the positions over the game map. However, even so, this example shows how important *collaborative AI* could be in reaching complex common objectives to benefit all rather than just some. Moreover, given the nature of the highly-paced mobile society of the 21st century, it is clear that real-world collaboration at scale may only be achieved by supplementing the sensing component with communication links. For practicality reasons, most of these links will be in fact driven by wireless communications.

Additionally, the proliferation of increasingly more variate and accurate sensors in modern day devices will lead undoubtedly to *advanced context-awareness* that may be exploited to reach collaboratively common goals. To clarify, consider for instance a mass of completely autonomous vehicles (potentially the future moving BSs) driving collaboratively through a crowded urban canopy, freed of traffic lights or signs. The collaboration of intelligent agents (the autonomous vehicles) will be mostly locally-oriented as the links amongst close-by vehicles approaching a common spatial area take precedence. However, the ego-centric sensing of each vehicle will be augmented with outer layers of information coming from the proximity traffic “partners”, such as for instance the planned routes of each vehicle, the compressively sensed perspective of surrounding vehicles, the prospective proximity of emergency vehicles or highly priority nodes (*e.g.* people, bicyclists) *etc.* All of these ego-generated and collected data points will then contribute distributively in navigating all the approaching road users through the busy roads safely and with reduced overall congestion.

In the light of the above, it is clear that vast amounts of connections will be necessary to solve such large-scale distributed problems where massive access, large data volumes, ultra low-latency, but also sensing services beyond those that are to be offered by 5G networks will be required. In fact, such applications sparked by collaborative intelligence will constitute in end effect dynamic opportunistic distributed systems in need of ad-hoc local networks generated bottom-up to serve them. In other words, these ensembles would constitute *collaborative quilt networks* relying on distributed rather than centralized logic enabled mostly by sidelinks instead of traditional cellular UL and DL.

Furthermore, intelligent agents may also be deployed to BSs as well. To this end, AI technology may be utilized to optimize not only wireless access and network tasks, a direction which 5G network slicing and function virtualization is already aiming to solve, but also to provide business and application intelligence for users. One should note that this does not mean that AI will replace wireless engineers and system architects. On the contrary, it should act as a support for the latter in creating more efficient protocols to enable scalable massive access models under different constraints of latency, bandwidth or power utilization. In some way this may lead to *soft and adaptive protocols* involving more distributed consensus among communicating nodes than the hard rules implemented today by expert-based rules.

From the above arguments, we identify some potential qualitative requirements of a prospective future 6G standard as:

- embedded *wireless environment sensing*
- pervasive *enhanced context-awareness*
- nodes/network *contextual self-reconfiguration*
- functional/situational/positional network *self-aggregation*
- *real-time and opportunistic* latency, rate and access *provisioning*

Given the above list and previously detailed applications, it is clear that a lot of network research is necessary to extend the current 5G *policy-driven* approaches of network slicing and virtualization to more organic and flexible constructs capable to adaptively and scalably address heterogeneity beyond what the current targets are. Thus, *context-aware adaptive device self-reconfiguration* and *network self-aggregation* will require vast amounts of research and development to empower at scale and in real-time collaborative AI applications as previously described.

However, one would probably be mistaken to consider that 6G shall focus mostly on network related aspects, and that as a result the PHY layer research will become limited or play a secondary role. History tells us that a similar situation daunted upon the research community at yet another cross-road, at the end of 4G research phase, when it was asked whether “*the PHY layer was dead*” [230]. The main concern at the time was that PHY research would actually slowly decrease and limit to fulfilling short-term and derivative needs of the telecommunications industry. Time has shown that the concern was in fact ephemeral as the interest in PHY layer research was actually rekindled by both hardware related and signal processing advances. To name a few, consider just the recently emerged innovations of NOMA, full-duplex (FD) radio and hybrid mmWave RF technologies.

Having learned from this history lesson, we believe that in fact PHY interest will not become extinct. Instead the research focus will be subsequently transformed by advances and developments to come, both in the hardware, but also in the software areas. Concretely, we believe that over the course of the next decades the PHY will be segregated into two components: the *low-PHY* and the *high-PHY*, respectively.

On the one hand, the “low-PHY” shall address more signal processing oriented problems, closer to the RF hardware under the embedded signal impairments, acting in fact as an *advanced wireless sensor* under the dynamics of 6G operation. To this end, the “low-PHY” should provide the handling of flexible waveforms modulation/demodulation, precoding & combining, beamforming & interference management, given all hardware impairments, such as for instance, compacting of massive MIMO antennas, FD operation, mmWave RF distortions, CSI imperfections *etc.*. Furthermore, alongside these canonical functions, we envision that this layer will also focus in providing processed wireless information beyond what is available today in wireless transceivers, enabling accurate estimation of side information such as ranging, angle of arrival/angle of departure, detection and categorization of transmitter, transmitter tracking, radar imaging and so on. These communication metadata could then be leveraged by upper layers or by the business logic beyond what is de facto possible today and thus help the envisioned emerging top-level collaborative AI applications. As a result, this “low-PHY” will develop strongly towards enabling software defined radios (SDRs) architectures, a dream many times attempted before, but which was never really fully realized at scale.

On the other hand, the “high-PHY” will aim to complete the current virtualization trend started by 5G, and so, will focus on becoming the *software-oriented driver* of the “low-PHY”. For this purpose, we envision therefore research on “high-PHY” to focus on application-based functional post-filtering of the information coming from the “low-PHY”, and respectively, on the interface with the upper layers and intelligence core. In summary, “high-PHY” shall thus develop the code-domain technologies required to control, interact and exploit the capabilities of the “low-PHY”.

To illustrate the last points regarding the PHY layer, consider for instance the MC-NOMA method proposed in Chapter 5 as non-orthogonal access scheme, and respectively, the last requirement of opportunism cited previously and envisioned as one qualitative requirements of 6G. Although 5G by design is enabled to offer a multitude of differentiated services, including for instance URLLC and mMTC, it still relies on the traditional model of grant-based communications [9, 10]. In contrast, a truly opportunistic system, should rely on the premises of grant-free communications, where users could uplink for instance their information whenever that is available without going thorough the entire handshaking procedure of creating an established communication link. Grant-free access has therefore the potential to offer both efficient and power-optimized massive access.

The proposed MC-NOMA scheme could be leveraged to this end in its UL scenario for IoT/industrial IoT (IIoT) use cases where the signature waveforms of the MC-NOMA precoding matrices, *i.e.*, the columns of the MC-NOMA precoding frame, are leased towards individual users which are sporadically active. The lease management and bookkeeping of these signatures would then be handled by the high-PHY together with other higher-level functions, whereas the precoding/detection would be performed by the low-PHY, so as to decouple the processing and allow for further virtualization at the PHY level.

6.3 Concluding Remarks

To summarize, the development prospects and future directions of 5G and 5G+ are of paramount interest provided the recent advances and potential applications of massively distributed collaborative AI. As previously detailed, the collaborative intelligence could become a catalyst breaking the technologies addenda to 5G and its evolutionary successors into a standalone 6G new paradigm.

The definitive fact is in end effect whether the “predicted” causes and effects detailed within this chapter find somehow their way in the future wireless communications ecosystem, either as (an unimaginatively named) 5G+ or as 6G, to finally provide an organic massive Internet access for people and intelligent machines alike. To get the real-world answer to this dilemma one just has thus to wait and observe...

Chapter 7

Epilogue

The manuscript is concluded in the sequel with the presentation of the main points of the research, key results and future work aspects worth pursuing as derivatives of this work.

7.1 Main Conclusions

A new design approach to modern and future wireless communications has been introduced by this thesis via the concepts of Frame Theory. Motivated by the access, accuracy and performance limitations observed in the case of orthogonal communications systems along the short but very rich history of radio communications, we have resorted to the alternative treaty of wireless communication systems via non-orthogonal and redundant signal processing. From this perspective, mathematical frame-theoretic constructs were formalized and interpreted in engineering contexts to facilitate the understanding of the use cases of such overcomplete signal representations for wireless communication systems. To this end, the compressive synthesis and expansive analysis of signal vectors were discussed and clarified in Chapter 3, together with the special linear operators, *i.e.* frames, rendering them.

Upon this applied analytical approach, two general frameworks were introduced and projected directly onto two different popular research fields within future wireless communications, namely mmWave, and respectively, NOMA systems.

Punctually, the mmWave contribution is linked to the non-trivial sparse recovery of the communication channels, which in terms is in fact a particular problem of CS. Thus, we clarified analytically the links between CS and Frame Theory and defined the design criteria of the embedded measurement matrices central to the mmWave problem to be incoherence, tightness and equi-normality. The optimum design was thus found to be an incoherent UNTF, constituting therefore the training joint beamforming vectors at the Tx and Rx. Since such a matrix represented an idealized realization infeasible in practice given the physical separation between Tx and Rx, we have derived two general mathematical results outlining the fact that the attributes

of such designs are in fact invariant to the Kronecker product describing the combined measurement matrix previously optimized. As a result, a practical approach to separate the idealized solution into feasible Tx/Rx training beamformers as incoherent UNTFs has been found based on the same method used to generate the ideal solution, the generation of incoherent UNTFs. The latter was achieved by leveraging iteratively successive convex optimization decorrelations of the frame vectors given the CSIDCO framework [82], followed respectively by subsequent tightening under the polar decomposition, Theorem 3.5.1.

The beamformers thus obtained, both in their idealized, and respectively, practical realizations, have generally provided an increased sparse recovery robustness of the mmWave channels regardless of the recovery algorithms utilized. Thus the proposed designs are highly desirable in practical deployments for mmWave beamforming management, in particular for the cases where no prior information about the communication channels are available, *i.e.* for the IA scenarios. Furthermore, since the problem of sparse recovery is central to CS, the designs obtained in Chapter 4 constitute in fact also optimum measurement matrices for general sparse recovery problems, beyond the one of mmWave channel estimation. This is a consequence of the fact that the design procedure utilized for such measurement matrices, Algorithm 3, yields optimal frames with properties of tightness, incoherence and uniform sensing energy that universally maximize the guarantees of sparse recovery.

Lastly, to provide even further gains beyond the state of the art, the sparse recovery of the mmWave signal was resumed to an approximated ℓ_0 -norm formulation leading to a FPG problem solved by QT [149] to a linear program with box constraints. The latter formulation allowed upon the utilization of the ℓ_0 -norm a better recovery of the sparse mmWave channel in low and medium SNR areas than the equivalent ℓ_1 -based methods of BPDN and ℓ_1 -reweighted BPDN.

On the other hand, the NOMA problem was firstly abstracted and formulated in terms of a generic linear system where the non-orthogonal multiplexing of K active users onto M RE such that $K > M$ was resumed to a finite frame operator. This generic formulation compressed in turn the multitude of CD-NOMA schemes under a simplified linear model that allowed further exploitation and comparison of the methods in terms of the sum-rates for both DL and UL. Thus, an ergodic achievable sum-rate capacity analysis over the generic model of both DL and UL under joint ML detection for the cases of CCMC, and respectively, DCMC revealed the optimum conditions of designing the multiplexing matrix corresponding to CD-NOMA. These criteria identified the ideal non-orthogonal multiplexing matrix to be a dense, incoherent UNTF. Given the similarity to the measurement matrices constructed for the case of mmWave beamforming, the same method was used in this case to generate the dense, incoherent UNTFs.

However, to aid practical implementations, the CSIDCO formulation [82] was slightly relaxed to an equivalent box- and sphere-constrained quadratic program coined as QCSIDCO. It was shown moreover that the proposed alternative was on par in terms of frame generation to CSIDCO, managing to generate similar dense, incoherent and UNTFs. In general, the frames obtained via Algorithm 3 are in fact approximations of ideal low-coherence frames, also known as optimal Grassmannian frames. Upon this fact, the inter-user interference thus controlled via

such frames is minimized aiding the MUD without any requirement of embedded sparsity which is a recurrent design principle across the majority of other CD-NOMA schemes. As a result, the NOMA multiplexing implemented by dense, incoherent UNTFs was named MC-NOMA, as it is capable of providing massive access while spreading users by individual signature waveforms across all the available resources at once.

In the light of the MC-NOMA optimized design, it was shown furthermore that this was indeed optimal in terms of ergodic sum-rates across both DL and UL, and respectively, CCMC and DCMC scenarios, in comparison with the best performing CD-NOMA schemes [152, 156], *i.e.* SCMA [178] and PDMA [176].

To complete the circle of the MC-NOMA system and simultaneously advocate that indeed the optimum ergodic sum-rate is achievable a low-complexity effective and efficient MUD was then proposed in the form of Rubik-PTST. This is a DFS tree search algorithm enhanced on one hand by SD classical search ordering optimizations, *i.e.* SE and constellation reduction, but also on the other hand by additionally proposed complexity reduction mechanisms such as the Rubik-QR decomposition, and respectively, tree pruning with hard thresholding.

The first optimization is an algebraic reduction strategy rearranging the search space by increasing the search energy associated with the overloaded equation to first decode. This is achieved by means of subsequent column- and row-permuted QR (QR) ordering and decomposition. The second optimization however is stochastically based on the expected distribution of the squared residual under the assumption of AWGN receiver noise. The idea is very simply reducing the search space by tree pruning of solutions and corresponding subtrees whose subsequent squared residuals fall outside of a statistically chosen confidence interval radius given the expected chi-squared distribution of the latter.

The results showed that such a search based strategy is able generally to retrieve the joint ML solution within feasible detection times for modern day processors. However, the proposed optimizations and detector structure cannot generally lower the expected worst-time exponential complexity of such search structures. As an effect, the detector is successfully and feasibly applicable for low-scale MC-NOMA systems with high overloading, but also to medium-scale MC-NOMA systems with low overloading.

Finally, to address this problem of joint detection at massive scale, we introduced a new sparsely-structured receiver. This was formulated based on FPG, and respectively, on the QT of an ℓ_0 -approximated smooth and differentiable formulation of the search space. As a result, a probabilistic-constrained quadratic program was obtained and solved efficiently via ADMM. This joint detector proved to surpass the state of the art effectively across various system dimensions, scaling well with the system size.

Lastly, our expected directions of 6G were provided given the goal of achieving organic Internet access for both human and machine users. To this end, collaborative AI was identified as the main disruptor triggering the standardization evolution and the emergence of new paradigms, like for instance the low-PHY and high-PHY separation.

7.2 Outlook and Extensions

The work covered throughout this manuscript does not try and cannot be by any means an exhaustive treatment of all the touched topics. In some cases, particular simplifying assumptions are made, whereas in some other idealized scenarios are assumed. Despite these facts, caveats are made clear and discussed inline within the main body of the work.

We summarize here however some valuable working points to complementing or to extending this research:

- Despite of formalizing practically achievable fully digital designs for the training Tx/Rx beamformers, Chapter 4 does not treat the practical realization of the latter in terms of common hybrid RF architectures of mmWave systems. This problem could constitute an interesting extension of the mmWave section where advantage of the structure of the Tx/Rx training beamformers may be taken towards their realization with hybrid architectures.
- Chapter 4 does not treat the case of massive mmWave systems, but instead considers low-cost transceivers with small numbers of Tx/Rx antennas, *i.e.* $T = R = 8$. An extension of the numerical simulations and analysis could be done to systems with more number of antennas and higher number of training beamforming/combining vectors, which may be more common in practice.
- The MC-NOMA receivers assume perfect CSIR, which is an idealized assumption, providing an upper bound of achievable system performance. To study real-world feasibility, the decoders should be tested under the imperfect CSIR assumption and the performance degradation relative to the ideal case should be measured.
- The MC-NOMA simulated channels are mainly specular stochastic Rayleigh fading, but it could be worth it to investigate more realistic channels instead, with and without perfect CSIR.
- The near-far effect specific to CDMA may be apparent to MC-NOMA as well. Despite the expectations that this phenomenon should be mitigated by the joint receivers proposed, an analysis in this direction may be of practical interest.
- The extension and applicability of MC-NOMA towards truly grant-free access could be another interesting extension of future work.

Appendix A

Normed, Inner Product and Hilbert Spaces

NOTE: This appendix is not meant to provide a full-depth clarification of linear algebra or functional analysis topics. Instead, it lists sufficient review material to provide a convenient reference and help the reader follow the work more easily. For a more in-depth refresher of such topics, the reader is directed to literature such as [20, 72, 231].

Firstly, recall that a *vector space* is the closed set V under finite vector addition, *i.e.*

$$+ : V \times V \rightarrow V, \tag{A.1}$$

and scalar multiplication, *i.e.*

$$\cdot : \mathcal{F} \times V \rightarrow V, \tag{A.2}$$

defined over the scalar field \mathcal{F} [231], such that the following eight axioms hold for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and scalars $a, b \in \mathcal{F}$:

- commutativity: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- identity under addition: $\mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}$
- identity under scalar multiplication: $1 \cdot \mathbf{x} = \mathbf{x}$
- existence of additive inverse: $\forall \mathbf{x} \in V, \exists -\mathbf{x}$ such that $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$
- associativity of vector addition: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- associativity of scalar multiplication: $a(b\mathbf{x}) = (ab)\mathbf{x}$
- distributivity of scalar sum: $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$
- distributivity of vector sum: $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$

Under the concept of bases, widely covered in Chapter 3, vector spaces are fully characterized from an algebraic perspective solely by their dimensionality, *i.e.* their size. However, from a functional analysis perspective, the algebraic definition of vector spaces does not provide the necessary structure to determine convergence relations for sequences of functions to another function, given that these are defined over the vector space in question. In response, the concept of *normed* or *inner product vector spaces* was introduced [72] to address the “measuring” necessary to determine the convergence of such functions and sequences thereof. Concretely, these two reduce to the following definitions [72].

Definition A.0.1. Consider the vector space V . The function $\|\cdot\| : V \rightarrow [0, \infty) \cup \infty$ satisfying the conditions

- i). $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
- ii). $\|a\mathbf{x}\| = |a| \|\mathbf{x}\|, \quad \forall \mathbf{x} \in V, a \in \mathcal{F}$
- iii). $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in V.$

is a norm on V .

Definition A.0.2. Consider the vector space V . The function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathcal{F}$ for which the properties

- i). $\langle \mathbf{x}, \mathbf{y} \rangle = (\langle \mathbf{y}, \mathbf{x} \rangle)^*, \quad \forall \mathbf{x}, \mathbf{y} \in V$
- ii). $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle, \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, a, b \in \mathcal{F}$
- iii). $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \quad \forall \mathbf{x}, \mathbf{y} \in V$
- iv). $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

hold, is an inner product on V .

Thus, if V is equipped with a norm in the sense of Definition A.0.1, V is a *normed vector space*, whereas if the vector space V is equipped with an inner product as detailed by Definition A.0.2, then V is a *inner product vector space*. Moreover, notice in the light of the two definitions from above that if V is an inner product vector space, then it can be equipped with a norm defined as

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}, \tag{A.3}$$

and as a result, V is also a normed space.

In addition, the inner product spaces define also the property of orthogonality, such that two elements $\mathbf{x}, \mathbf{y} \in V$ are orthogonal if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, where the orthogonal complement of any subspace $U \subset V$ is further defined respectively as

$$U^\perp \triangleq \{\mathbf{x} \in V \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0, \mathbf{y} \in U\}. \tag{A.4}$$

Finally, given all of the above a Hilbert space is defined as follows [72].

Definition A.0.3. *An inner product space H with the norm defined by Equation A.3 is a Hilbert space if its norm is a complete metric space.*

In other words, an inner product space H is a Hilbert space if its norm metric satisfies convergence over all Cauchy sequences [72], *i.e.*

$$\lim_{\min(m,n) \rightarrow \infty} \|\mathbf{x}_m - \mathbf{x}_n\| = 0, \quad (\text{A.5})$$

for every sequence $\mathbf{x}_1, \mathbf{x}_2, \dots$ of elements in H .

As examples of finite-dimensional Hilbert spaces [72] that are also explored in this work are:

1. The real numbers vector space \mathbb{R}^M with the inner product $\langle \mathbf{x}, \mathbf{x} \rangle$ defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \mathbf{x}^T \mathbf{y}, \quad (\text{A.6})$$

or equivalently,

$$\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \mathbf{y}^T \mathbf{x}. \quad (\text{A.7})$$

2. The complex numbers vector space \mathbb{C}^M with the inner product $\langle \mathbf{x}, \mathbf{x} \rangle$ defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \mathbf{y}^H \mathbf{x}. \quad (\text{A.8})$$

Appendix B

Own Publications

Journal Papers:

[J1] R.-A. Stoica, S. Severi and G. T. Freitas de Abreu, “A Self-Organizing Frequency Approach to 802.11p Channel Estimation,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1930-1942, July 2017.

[J2] R.-A. Stoica, G. T. Freitas de Abreu, T. Hara and K. Ishibashi, “Massively Concurrent Non-Orthogonal Multiple Access for 5G Networks and Beyond,” in *IEEE Access – Special Section on Advances in Signal Processing for Non-orthogonal Multiple Access*, vol. 7, pp. 82080-82100, July 2019.

[J3] R.-A. Stoica, H. Iimori and G. T. Freitas de Abreu, K. Ishibashi, “Frame Theory and Fractional Programming for Sparse Recovery-based mmWave Channel Estimation,” in *IEEE Access*, submitted (in review).

Conference and ArXiv Papers:

[C1] H. Iimori, R.-A. Stoica and G. T. Freitas de Abreu, “Constellation shaping for rate maximization in AWGN channels with non-linear distortion,” *IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curaçao, Dutch Antilles, 2017, pp. 1-5.

[C2] R.-A. Stoica, G. T. Freitas de Abreu and H. Iimori, “A Frame-Theoretic Scheme for Robust Millimeter Wave Channel Estimation,” *IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, IL, USA, 2018, pp. 1-6.

[C3] R.-A. Stoica and G. T. Freitas de Abreu, “Massively Concurrent NOMA: A Frame-Theoretic Design for Non-Orthogonal Multiple Access,” *52nd Asilomar Conference on Signals, Systems,*

and Computers, Pacific Grove, CA, USA, 2018, pp. 461-466.

[C4] R.-A. Stoica and G. T. Freitas de Abreu, “6G: the Wireless Communications Network for Collaborative and AI Applications,” *arXiv preprint*, arXiv:1904.03413.

[C5] T. Hara, R.-A. Stoica, K. Ishibashi and G. T. Freitas de Abreu, “On the Sum-Rate Capacity and Spectral Efficiency Gains of Massively Concurrent NOMA Systems,” *IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakech, Marroco, 2019, pp. to appear.

[C6] R.-A. Stoica and G. T. Freitas de Abreu, “Frame-theoretic Precoding and Beamforming Design for Robust mmWave Channel Estimation,” *IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakech, Marroco, 2019, pp. to appear.

[C7] R.-A. Stoica and G. T. Freitas de Abreu, “A Low-complexity Receiver for Massively Concurrent Non-orthogonal Multiple Access”, *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cannes, France, 2019, pp. 1-5.

[C8] R.-A. Stoica, H. Iimori and G. T. Freitas de Abreu, “Sparsely-structured Multiuser Detection for Large Massively Concurrent NOMA Systems,” *53rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2019, accepted (to appear).

[C9] R.-A. Stoica, H. Iimori and G. T. Freitas de Abreu, “Multiuser Detection for Large Massively Concurrent NOMA Systems via Fractional Programming,” *IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Guadalupe, West Indies, 2019, accepted (to appear).

Bibliography

- [1] V. Garg, *Wireless Communications & Networking*. Elsevier, 2010.
- [2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, 1st ed. Cambridge University Press, 2005.
- [3] B. P. Lathi and Z. Ding, *Modern Digital and Analog Communication Systems*, 4th ed. Oxford University Press, 2009.
- [4] J. Schiller, *Mobile Communications*, 2nd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2003.
- [5] Y. G. Li and G. L. Stuber, *Orthogonal frequency division multiplexing for wireless communications*. Springer Science & Business Media, 2006.
- [6] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [7] International Telecommunication Union (ITU), “Recommendation ITU-R M.2083-0: IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond,” Sep 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf
- [8] —, “Recommendation ITU-R M.2012-3: Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-Advanced (IMT-Advanced),” Jan 2018. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2012-3-201801-I!!PDF-E.pdf
- [9] 3rd Generation Partnership Project (3GPP), “TR 21.915 V1.0.0 (2019-03): Release description; Release 15,” Mar 2019. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/21_series/21.915/21915-100.zip
- [10] —, “Release 16,” Public Progress Report. [Online]. Available: <https://www.3gpp.org/release-16>
- [11] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ: Pearson Education Limited, 2016.
- [12] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, “Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!” *IEEE Access*, vol. 1, pp. 335–349, 2013.

-
- [13] 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.
- [14] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, “An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems,” *IEEE Journal on Selected Topics on Signal Processing*, vol. 10, no. 3, pp. 436 – 453, 2016.
- [15] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, “MIMO for Millimeter-wave Wireless Communications: Beamforming, Spatial Multiplexing, or Both?” *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, Dec. 2014.
- [16] Y. Saito and Y. Kishiyama and A. Benjebbour and T. Nakamura and A. Li and K. Higuchi, “Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access,” in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013, pp. 1–5.
- [17] Y. Saito and A. Benjebbour and Y. Kishiyama and T. Nakamura, “System-level Performance Evaluation of Downlink Non-orthogonal Multiple Access (NOMA),” in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 611–615.
- [18] Z. Ding and X. Lei and G. K. Karagiannidis and R. Schober and J. Yuan and V. K. Bhargava, “A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, Oct 2017.
- [19] P. G. Casazza and G. Kutyniok, *Finite Frames: Theory and Applications*, 1st ed. Birkhäuser Basel, 2012.
- [20] O. Christensen, *An Introduction to Frames and Riesz Bases*, 1st ed. Springer, 2003.
- [21] J. Cahill, P. G. Casazza, and G. Kutyniok, “Operators and frames,” *Journal of Operator Theory*, vol. 70, no. 1, pp. 145–164, 2013.
- [22] J. Kovačević and A. Chebira, “An Introduction to Frames,” *Foundations and Trends in Signal Processing*, vol. 2, no. 1, pp. 1–94, 2008.
- [23] V. I. Morgenshtern and H. Bölcskei, “A Short Course on Frame Theory,” *arXiv preprint arXiv:1104.4300*, 2011.
- [24] *IEEE Guide for Wireless Access in Vehicular Environments (WAVE) – Architecture*, Std., March 2014.
- [25] *IEEE Standard for Information technology – Telecommunications and information exchange between systems Local and metropolitan area networks — Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Std., March 2012.

- [26] ETSI, *ETSI ES 202 663 V1.1.0: Intelligent Transport Systems (ITS); European Profile Standard for the Physical and Medium Access Control Layer of Intelligent Transport Systems Operating in the 5 GHz Frequency Band*, Std., 2009.
- [27] —, *ETSI EN 302 663 V1.2.1: Intelligent Transport Systems (ITS); Access Layer Specification for Intelligent Transportation Systems Operating in the 5GHz Frequency Band*, Std., 2013.
- [28] —, *ETSI EN 302 665 V1.1.1: Intelligent Transport Systems (ITS); Communications Architecture*, Std., 2010.
- [29] *IEEE Standard for Information Technology – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments*, Std., July 2010.
- [30] N. Abramson, “THE ALOHA SYSTEM: Another Alternative for Computer Communications,” in *Proceedings of the Fall Joint Computer Conference*, ser. AFIPS ’70 (Fall). New York, NY, USA: ACM, Nov 1970, pp. 281–285. [Online]. Available: <http://doi.acm.org/10.1145/1478462.1478502>
- [31] D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data Networks*, 2nd ed. Prentice-Hall International New Jersey, 1992.
- [32] R. B. Cooper, *Introduction to Queueing Theory*, 2nd ed. Elsevier North Holland, 1981.
- [33] A. V. Oppenheim and R. W. Schaffer, *Discrete-time Signal Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.
- [34] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*, 1st ed. McGraw-Hill, 1994.
- [35] S. Verdu and S. Shamai, “Spectral efficiency of CDMA with random spreading,” *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, March 1999.
- [36] J. W. Cooley and J. W. Tukey, “An algorithm for the machine calculation of complex Fourier series,” *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.
- [37] Y. Rahmatallah and S. Mohan, “Peak-To-Average Power Ratio Reduction in OFDM Systems: A Survey And Taxonomy,” *IEEE Communications Surveys Tutorials*, vol. 15, no. 4, pp. 1567–1592, Dec 2013.
- [38] M. Pätzgold, *Mobile Fading Channels*, 1st ed. John Wiley & Sons, 2002.
- [39] F. Hlawatsch and M. G., *Wireless Communications over Rapidly Time-Varying Channels*, 1st ed. Elsevier, 2011.
- [40] G. Matz, “On non-WSSUS wireless fading channels,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2465 – 2478, 2005.

-
- [41] G. Acosta-Marum and M. A. Ingram, “Six Time-and Frequency-selective empirical channel models for vehicular wireless LANs,” *Vehicular Technology Magazine, IEEE*, vol. 2, no. 4, pp. 4–11, 2007.
- [42] G. Acosta-Marum, “Measurement, modeling, and OFDM synchronization for the wide-band mobile-to-mobile channel,” Ph.D. dissertation, 2007.
- [43] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. JHU press, 2013.
- [44] *IEEE Standard for Telecommunications and Information Exchange Between Systems - LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: High Speed Physical Layer in the 5 GHz band*, Std., Dec 1999.
- [45] C. F. Mecklenbräuker, A. F. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernadó, T. Zemen, O. Klemp, and N. Czink, “Vehicular channel characterization and its implications for wireless system design and performance,” *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1189–1212, 2011.
- [46] P. Alexander, D. Haley, and A. Grant, “Outdoor Mobile Broadband Access with 802.11,” *IEEE Communications Magazine*, vol. 45, no. 11, pp. 108–114, November 2007.
- [47] —, “Cooperative Intelligent Transport Systems: 5.9-GHz Field Trials,” *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1213–1235, July 2011.
- [48] A. F. Molisch, F. Tufvesson, J. Karedal, and C. F. Mecklenbräuker, “A survey on vehicle-to-vehicle propagation channels,” *IEEE Wireless Communications*, vol. 16, no. 6, pp. 12–22, December 2009.
- [49] M. Boban, J. Barros, and O. Tonguz, “Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation,” *Vehicular Technology, IEEE Transactions on*, vol. 63, no. 9, pp. 4146–4164, 2014.
- [50] W. Viriyasitavat, M. Boban, H.-M. Tsai, and A. Vasilakos, “Vehicular communications: Survey and challenges of channel and propagation models,” *Vehicular Technology Magazine, IEEE*, vol. 10, no. 2, pp. 55–66, 2015.
- [51] A. Paier, T. Zemen, L. Bernadó, G. Matz, J. Karedal, N. Czink, C. Dumard, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, “Non-WSSUS vehicular channel characterization in highway and urban scenarios at 5.2GHz using the local scattering function,” in *Smart Antennas, 2008. WSA 2008. International ITG Workshop on*, Feb 2008, pp. 9–15.
- [52] L. Bernadó, T. Zemen, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, “The (in-) validity of the WSSUS assumption in vehicular radio channels,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, Sept 2012, pp. 1757–1762.

- [53] J. Heiskala and J. Terry Ph D, *OFDM wireless LANs: A theoretical and practical guide*. Sams, 2001.
- [54] *IEEE Standard for Information Technology – Local and Metropolitan Area Networks – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments*, Std., July 2010.
- [55] E. G. Ström, “On Medium Access and Physical Layer Standards for Cooperative Intelligent Transport Systems in Europe,” *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1183–1188, 2011.
- [56] R.-A. Stoica, S. Severi, and G. T. Freitas de Abreu, “Learning the Vehicular Channel Through the Self-Organization of Frequencies,” in *2015 IEEE Vehicular Networking Conference (VNC)*, Dec 2015, pp. 68–75.
- [57] —, “A Self-Organizing Frequency Approach to 802.11p Channel Estimation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. TBD — to appear, DOI:10.1109/TITS.2016.2612886.
- [58] C. Douillard, M. Jézéquel, C. Berrou, D. Electronique, A. Picart, P. Didier, and A. Glavieux, “Iterative correction of intersymbol interference: Turbo-equalization,” *European transactions on telecommunications*, vol. 6, no. 5, pp. 507–511, 1995.
- [59] G. Ren, H. Zhang, and Y. Chang, “SNR estimation algorithm based on the preamble for OFDM systems in frequency selective channels,” *IEEE Transactions on Communications*, vol. 57, no. 8, pp. 2230–2234, Aug 2009.
- [60] J. A. Fernandez, K. Borries, L. Cheng, B. Vijaya Kumar, D. D. Stancil, and F. Bai, “Performance of the 802.11p physical layer in vehicle-to-vehicle environments,” *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 1, pp. 3–14, 2012.
- [61] Z. Zhao, M. Wen, X. Cheng, C.-X. Wang, and B. Jiao, “A novel effective channel estimation scheme applicable to IEEE 802.11p,” in *ITS Telecommunications (ITST), 2012 12th International Conference on*, Nov 2012, pp. 88–92.
- [62] Z. Zhao, X. Cheng, M. Wen, L. Yang, and B. Jiao, “Constructed Data Pilot-Assisted Channel Estimators for Mobile Environments,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 947–957, April 2015.
- [63] R. Koetter, A. C. Singer, and M. Tuchler, “Turbo equalization,” *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 67–80, Jan 2004.
- [64] J. A. Fessler and A. O. Hero, “Space-Alternating Generalized Expectation-Maximization Algorithm,” *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, Oct 1994.

-
- [65] C. Smith, *The Car Hacker's Handbook: a Guide for the Penetration Tester*. No Starch Press, 2016.
- [66] J. Misener, S. Andrews, P. Cannistra, T. Adams, J. Collins, D. Garcia, A. Waite, R. Walsh, B. Sheppard, and B. A. Hamilton, "Communications data delivery system analysis task 2 report: high-level options for secure communications data delivery systems." United States. Joint Program Office for Intelligent Transportation Systems, Tech. Rep., 2012. [Online]. Available: <https://rosap.ntl.bts.gov/view/dot/3359>
- [67] E. J. Candès and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [68] I. Daubechies, A. Grossmann, and Y. Meyer, "Painless nonorthogonal expansions," *Journal of Mathematical Physics*, vol. 27, no. 5, pp. 1271–1283, 1986.
- [69] D. Gabor, "Theory of communication. Part 1: The analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.
- [70] —, "Theory of communication. Part 2: The analysis of hearing," *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 442–445, Nov 1946.
- [71] —, "Theory of communication. Part 3: Frequency compression and expansion," *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 445–457, Nov 1946.
- [72] N. Young, *An introduction to Hilbert space*. Cambridge University Press, 1988.
- [73] A. Ben-Israel and T. N. Greville, *Generalized Inverses: Theory and Applications*. Springer, 2003, vol. 15.
- [74] K. Grochenig, "Acceleration of the Frame Algorithm," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3331–3340, Dec 1993.
- [75] J. J. Benedetto and M. Fickus, "Finite Normalized Tight Frames," *Advances in Computational Mathematics*, vol. 18, no. 2, pp. 357–385, Feb 2003.
- [76] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [77] Bodmann, Bernhard G and Casazza, Peter G and Kutyniok, Gitta, "A quantitative notion of redundancy for finite frames," *Applied and Computational Harmonic Analysis*, vol. 30, no. 3, pp. 348–362, 2011.
- [78] T. Strohmer and R. W. Heath Jr, "Grassmannian Frames with Applications to Coding and Communication," *Applied and computational harmonic analysis*, vol. 14, no. 3, pp. 257–275, 2003.

- [79] L. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Transactions on Information Theory*, vol. 13, pp. 619–621, 1967.
- [80] M. Rosenfeld, "In praise of the Gram matrix," in *The Mathematics of Paul Erdős II*. Springer, 1997, pp. 318–323.
- [81] P. Delsarte, J. Goethals, and J. Seidel, "Bounds for systems of lines, and Jacobi polynomials," *Philips Research Reports*, vol. 30, pp. 91–105, 1975.
- [82] C. Rusu, N. G. Prelcic, and R. W. Heath, "Algorithms for the Construction of Incoherent Frames under Various Design Constraints," *Signal Processing*, vol. 152, pp. 363–372, 2018.
- [83] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.
- [84] B. Garrett, "Polar Decomposition of a Matrix," *Course Lecture Notes*, Apr 2014. [Online]. Available: <http://buzzard.ups.edu/courses/2014spring/420projects/math420-UPS-spring-2014-buffington-polar-decomposition.pdf>
- [85] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
- [86] M. Elad, "Optimized projections for compressed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, Dec 2007.
- [87] E. V. Tsiligiani, L. P. Kondi, and A. K. Katsaggelos, "Construction of incoherent unit norm tight frames with application to compressed sensing," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2319–2330, Apr 2014.
- [88] C. Rusu and N. González-Prelcic, "Optimized compressed sensing via incoherent frames designed by convex optimization," *arXiv preprint arXiv:1507.02454*, 2015.
- [89] C. Rusu and N. González-Prelcic, "Designing incoherent frames through convex techniques for optimized compressed sensing," *IEEE Transactions on Signal Processing*, vol. 64, no. 9, pp. 2334–2344, May 2016.
- [90] R. B. Holmes and V. I. Paulsen, "Optimal frames for erasures," *Linear Algebra and its Applications*, no. 277, pp. 31–51, Jan 2004.
- [91] P. G. Casazza, O. Christensen, and A. J. Janssen, *Contemporary Mathematics*, no. 247, pp. 131–148, Jan 1999.
- [92] T. Strohmer and S. Beaver, "Optimal ofdm design for time-frequency dispersive channels," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1111–1122, July 2003.
- [93] R. Hadani, S. Rakib, S. Kons, M. Tsatsanis, A. Monk, C. Ibars, J. Delfeld, Y. Hebron, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," *arXiv preprint arXiv:1808.00519*, 2018.

-
- [94] P. Raviteja, K. T. Phan, Q. Jin, Y. Hong, and E. Viterbo, “Low-complexity iterative detection for orthogonal time frequency space modulation,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr 2018, pp. 1–6.
- [95] I. Daubechies, “The Wavelet Transform, Time-Frequency Localization and Signal Analysis,” *IEEE Transactions on Information Theory*, vol. 36, p. 5, Sep 1990.
- [96] —, *Ten Lectures on Wavelets*, 1st ed. SIAM, 1992.
- [97] D. L. Donoho, “Compressed Sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [98] Z. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Kuang, W. Chen, L. Li, and S. Li, “A survey on terahertz communications,” *China Commun.*, vol. 16, no. 2, pp. 1–35, Feb. 2019.
- [99] D. Reinsel, J. Gantz, and J. Rydning, “The digitization of the world from edge to core,” *IDC White Paper*, pp. 1–28, Nov. 2018.
- [100] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, “Overview of millimeter wave communications for fifth-generation (5G) wireless networks—with a focus on propagation models,” *IEEE Trans. Ant. Propagation*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.
- [101] S. Mumtaz, J. M. Jornet, J. Aulin, W. H. Gerstacker, X. Dong, and B. Ai, “Terahertz communication for vehicular networks,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5617–5625, Jul. 2017.
- [102] H. T. Friis, “A note on a simple transmission formula,” in *Proc. I.R.E. and Waves and Electrons*, May. 1946, pp. 254–256.
- [103] E. Björnson, L. V. der Perre, S. Buzzi, and E. G. Larsson, “Massive MIMO in sub-6GHz and mmwave: Physical, practical, and use-case differences,” *IEEE Trans. Wireless Commun.*, 2019.
- [104] V. Raghavan, A. Partyka, A. Sampath, S. Subramanian, O. H. Koymen, K. Ravid, J. Cezanne, J. Cezanne, K. Mukkavilli, and J. Li, “Millimeter-wave MIMO prototype: Measurements and experimental results,” *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 202–209, Jan. 2018.
- [105] 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.
- [106] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C.-L. I, and A. Ghosh, “Millimeter wave communications for future mobile networks,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sept. 2017.

- [107] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Sig. Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [108] F. Sohrabi 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, April 2016.
- [109] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Sig. Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
- [110] X. Yu, J. Zhang, and K. B. Letaief, "Hybrid precoding in millimeter wave systems: How many phase shifters are needed?" in *Proc. IEEE Globecom*, Singapore, Singapore, 2017, pp. 1–7.
- [111] C. Rusu, R. Mèndez-Rial, N. González-Prelcic, and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8380–8393, Dec 2016.
- [112] D. Jagyasi and P. Ubaidulla, "Low-complexity transceiver design for multi-user millimeter wave communication systems under imperfect CSI," in *Proc. IEEE VTC-Fall*, Montreal, Canada, Sept. 2016, pp. 1–5.
- [113] C. Xu, R. Ye, Y. Huang, S. He, and C. Zhang, "Hybrid precoding for broadband millimeter-wave communication systems with partial CSI," *IEEE Access*, vol. 6, pp. 50 891–50 900, 2018.
- [114] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial access in millimeter wave cellular systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 7926–7940, Dec 2016.
- [115] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems With Large Antenna Arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
- [116] 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.
- [117] H. Ghauch, M. Bengtsson, T. Kim, and M. Skoglund, "Subspace Estimation and Decomposition for Hybrid Analog-digital Millimetre-wave MIMO Systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 395–399.
- [118] S. Montagner, N. Benvenuto, and P. Baracca, "Channel Estimation Using a 2D DFT for Millimeter-Wave Systems," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

-
- [119] J. Lee, G. T. Gil, and Y. H. Lee, “Exploiting Spatial Sparsity for Estimating Channels of Hybrid MIMO Systems in Millimeter Wave Communications,” in *IEEE Global Communications Conference*, Dec. 2014, pp. 3326–3331.
- [120] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, “Channel Estimation and Hybrid Combining for mmWave: Phase Shifters or Switches?” in *Information Theory and Applications Workshop (ITA), 2015*, Feb. 2015, pp. 90–97.
- [121] A. A. M. Saleh and R. Valenzuela, “A statistical model for indoor multipath propagation,” *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, February 1987.
- [122] H. Yan and D. Cabric, “Compressive initial access and beamforming training for millimeter-wave cellular systems,” *arXiv preprint arXiv:1901.11220*, 2019.
- [123] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, Nov 1993, pp. 40–44 vol.1.
- [124] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [125] S. Chen and D. Donoho, “Basis Pursuit,” in *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, vol. 1, Oct. 1994, pp. 41–44 vol.1.
- [126] S. Malla and G. T. Freitas de Abreu, “Channel Estimation in Millimeter Wave MIMO Systems: Sparsity Enhancement via Reweighting,” in *Proceedings of IEEE ISWCS*, Poznan, Poland, Sept. 2016.
- [127] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, “Design and optimization on training sequence for mmwave communications: A new approach for sparse channel estimation in Massive MIMO,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1486–1497, July 2017.
- [128] X. Li, J. Fang, H. Li, and P. Wang, “Millimeter wave channel estimation via exploiting joint sparse and low-rank structures,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1123–1133, Feb 2018.
- [129] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, “Massive mimo channel estimation for millimeter wave systems via matrix completion,” *IEEE Signal Processing Letters*, vol. 25, no. 11, pp. 1675–1679, Nov 2018.
- [130] —, “Hybrid beamforming with random analog sampling for wideband channel estimation in millimeter wave massive MIMO systems,” in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1–5.

- [131] E. Vlachos, G. Alexandropoulos, and J. S. Thompson, “Wideband MIMO channel estimation for hybrid beamforming millimeter wave systems via random spatial sampling,” *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–15, Aug 2019.
- [132] J. W. Choi and B. Shim and Y. Ding and B. Rao and D. I. Kim, “Compressed Sensing for Wireless Communications: Useful Tips and Tricks,” *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1527–1550, July-Sept. 2017.
- [133] D. L. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003. [Online]. Available: <https://www.pnas.org/content/100/5/2197>
- [134] E. Candès and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse problems*, vol. 23, no. 3, p. 969, 2007.
- [135] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, *Introduction to compressed sensing*. Cambridge University Press, 2012, p. 1–64.
- [136] E. J. Candès and T. Tao, “Decoding by Linear Programming,” *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec 2005.
- [137] W. Chen, M. R. D. Rodrigues and I. J. Wassell, “On the Use of Unit-Norm Tight Frames to Improve the Average MSE Performance in Compressive Sensing Applications,” *IEEE Signal Processing Letters*, vol. 19, no. 1, pp. 8–11, Jan. 2012.
- [138] —, “Projection Design for Statistical Compressive Sensing: A Tight Frame Based Approach,” *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2016–2029, Apr. 2013.
- [139] M. Rossi, A. M. Haimovich, and Y. C. Eldar, “Spatial Compressive Sensing for MIMO Radar,” *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 419–430, Jan. 2014.
- [140] H. Yan and D. Cabria, “Compressive sensing based initial beamforming training for massive mimo millimeter-wave systems,” in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec 2016, pp. 620–624.
- [141] A. M. Sayeed, “Deconstructing multiantenna fading channels,” *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct 2002.
- [142] A. Sayeed and N. Behdad, “Continuous aperture phased mimo: Basic theory and applications,” in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2010, pp. 1196–1203.
- [143] M. Thill and B. Hassibi, “Group Frames With Few Distinct Inner Products and Low Coherence,” *IEEE Transactions on Signal Processing*, vol. 63, no. 19, Oct. 2015.

-
- [144] M. Thill and B. Hassibi, “Low-Coherence Frames From Group Fourier Matrices,” *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3386–3404, June 2017.
- [145] E. J. Candès, M. B. Wakin, and S. P. Boyd, “Enhancing Sparsity by Reweighted ℓ_1 Minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, Dec. 2008.
- [146] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Img. Sci.*, vol. 2, no. 1, pp. 183–202, Mar 2009. [Online]. Available: <http://dx.doi.org/10.1137/080716542>
- [147] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009.
- [148] M. Medra, A. W. Eckford, and R. Adve, “Using Fractional Programming for Zero-Norm Approximation,” *arXiv e-prints*, p. arXiv:1810.11725, Oct 2018.
- [149] K. Shen and W. Yu, “Fractional programming for communication systems – Part I: Power control and beamforming,” *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.
- [150] M. Grant and S. Boyd. (2017) CVX: Matlab software for disciplined convex programming. [Online]. Available: <http://cvxr.com/cvx/>
- [151] K. Higuchi and A. Benjebbour, “Non-orthogonal multiple access (noma) with successive interference cancellation for future radio access,” *IEICE Transactions on Communications*, vol. 98, no. 3, pp. 403–414, 2015.
- [152] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, “A survey of non-orthogonal multiple access for 5g,” *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2294–2323, thirdquarter 2018.
- [153] 3rd Generation Partnership Project (3GPP), “Study on downlink multiuser superposition transmission for LTE,” Mar 2015. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/36_series/36.859/36859-020.zip
- [154] P. Dayal and M. K. Varanasi, “A fast generalized sphere decoder for optimum decoding of under-determined MIMO systems,” in *Proc. of Allerton Conf. on Commun., Control, and Computing*, 2003, pp. 1216–1225.
- [155] M. Iwamura, K. Etemad, M. H. Fong, and R. Nory, “Carrier aggregation framework in 3GPP LTE-Advanced,” vol. 48, no. 8, pp. 60–67, Aug. 2010.
- [156] Z. Wu, K. Lu, C. Jiang, and X. Shao, “Comprehensive study and comparison on 5G NOMA schemes,” *IEEE Access*, vol. 6, pp. 18 511–18 519, 2018.

- [157] S. M. R. Islam, M. Zeng, O. A. Dobre, and K. Kwak, "Resource allocation for downlink NOMA systems: Key techniques and open issues," vol. 25, no. 2, pp. 40–47, Apr. 2018.
- [158] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," vol. 65, no. 8, pp. 6010–6023, Aug. 2016.
- [159] Y. Sun, D. W. K. Ng, and R. Schober, "Optimal resource allocation for multicarrier MISO-NOMA systems," in *IEEE Int. Conf. on Commun. (ICC)*, May 2017, pp. 1–7.
- [160] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for MC-NOMA systems," in *IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [161] Y. J. A. Zhang, L. Qian, J. Huang *et al.*, "Monotonic optimization in communication and networking systems," *Foundations and Trends® in Networking*, vol. 7, no. 1, pp. 1–75, 2013.
- [162] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," vol. 65, no. 3, pp. 1077–1091, Mar. 2017.
- [163] Y. Cao, H. Sun, J. Soriaga, and T. Ji, "Resource spread multiple access - a novel transmission scheme for 5G uplink," in *IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.
- [164] S. Verdú, *Multiuser Detection*. Cambridge University Press, 1998.
- [165] L. Ping, L. Liu, K. Y. Wu, and L. WK, "On interleave-division multiple-access," in *IEEE Int. Conf. on Commun. (ICC)*, vol. 5, Jun. 2004, pp. 2869–2873.
- [166] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," vol. 5, no. 4, pp. 938–947, Apr. 2006.
- [167] Q. Xiong, C. Qian, B. Yu, and C. Sun, "Advanced NOMA scheme for 5G cellular network: Interleave-grid multiple access," in *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–5.
- [168] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," vol. 56, no. 4, pp. 1616–1626, Apr. 2008.
- [169] J. van de Beek and B. M. Popovic, "Multiple access with low-density signatures," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.
- [170] R. Hoshyar, R. Razavi, and M. Al-Imari, "LDS-OFDM an efficient multiple access technique," in *IEEE 71st Veh. Technol. Conf. (VTC-Spring)*, May 2010, pp. 1–5.
- [171] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-user shared access for Internet of Things," in *IEEE 83rd Veh. Technol. Conf. (VTC-Spring)*, May 2016, pp. 1–5.

- [172] LG Electronics, “Considerations on DL/UL multiple access for NR,” in *3GPP Meeting RAN1#84b, Document R1-162517*, May 2016, retrieved Jan. 2019, pp. 1–4. [Online]. Available: http://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_84b/Docs/R1-162517.zip
- [173] Nokia, “Non-orthogonal multiple access for new radio,” in *3GPP Meeting RAN1#85, Document R1-165019*, May 2016, retrieved Jan. 2019, pp. 1–6. [Online]. Available: http://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_85/Docs/R1-165019.zip
- [174] Z. Yuan, C. Yan, Y. Yuan, and W. Li, “Blind multiple user detection for grant-free msa without reference signal,” in *IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.
- [175] J. Zeng, B. Li, X. Su, L. Rong, and R. Xing, “Pattern division multiple access (PDMA) for cellular future radio access,” in *Int. Conf. on Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.
- [176] B. Ren, Y. Wang, X. Dai, K. Niu, and W. Tang, “Pattern matrix design of PDMA for 5G UL applications,” *China Commun.*, vol. 13, no. 2, pp. 159–173, Nov. 2016.
- [177] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, “Pattern division multiple access — a novel nonorthogonal multiple access for fifth-generation radio networks,” vol. 66, no. 4, pp. 3185–3196, Apr. 2017.
- [178] H. Nikopour and H. Baligh, “Sparse code multiple access,” in *IEEE 24th Annu. Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 332–336.
- [179] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, “SCMA codebook design,” in *IEEE 80th Veh. Tech. Conf. (VTC-Fall)*, Sep. 2014, pp. 1–5.
- [180] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor graphs and the sum-product algorithm,” vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [181] H. Iimori, R.-A. Stoica, and G. T. Freitas de Abreu, “Constellation shaping for rate maximization in AWGN channels with non-linear distortion,” in *IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec 2017, pp. 1–5.
- [182] D. J. Garling, *Inequalities: A Journey into Linear Analysis*. Cambridge University Press, 2007.
- [183] E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
- [184] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE J. on Sel. Areas in Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

- [185] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. New York, NY, USA: Wiley-Interscience, 2006.
- [186] C. Xiao, Y. R. Zheng, and Z. Ding, “Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels,” vol. 59, no. 7, pp. 3301–3314, Jul. 2011.
- [187] S. X. Ng and L. Hanzo, “On the MIMO channel capacity of multidimensional signal sets,” vol. 55, no. 2, pp. 528–536, Mar. 2006.
- [188] J. Zeng, T. Lv, R. P. Liu, X. Su, M. Peng, C. Wang, and J. Mei, “Investigation on evolving single-carrier noma into multi-carrier noma in 5g,” *IEEE Access*, vol. 6, pp. 48 268–48 288, 2018.
- [189] M. Grant, S. Boyd, and Y. Ye, “Disciplined convex programming,” in *Global optimization*. Springer, 2006, pp. 155–210.
- [190] MATLAB, *Version 9.20.0 (R2017a)*. Natick, Massachusetts: The MathWorks Inc., 2017.
- [191] B. W. Kernighan and D. M. Ritchie, *The C programming language*, 2nd ed. Prentice-Hall, 2000.
- [192] J. Mattingley and S. Boyd, “CVXGEN: A code generator for embedded convex optimization,” *Optimization and Engineering*, vol. 13, no. 1, pp. 1–27, 2012.
- [193] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Mathematics of Computation*, vol. 44, no. 170, pp. 463–471, 1985.
- [194] H. Vikalo, “Sphere decoding algorithms for digital communications,” Ph.D. dissertation, Stanford University, 2003.
- [195] B. Hassibi and H. Vikalo, “On the sphere-decoding algorithm I. Expected complexity,” vol. 53, no. 8, pp. 2806–2818, Aug. 2005.
- [196] H. Vikalo and B. Hassibi, “On the sphere-decoding algorithm II. Generalizations, second-order statistics, and applications to communications,” vol. 53, no. 8, pp. 2819–2834, Aug. 2005.
- [197] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, “Generalised sphere decoder for asymmetrical space-time communication architecture,” *Electron. Lett.*, vol. 36, no. 2, pp. 166–167, Jan. 2000.
- [198] Z. Yang, C. Liu, and J. He, “A new approach for fast generalized sphere decoding in MIMO systems,” vol. 12, no. 1, pp. 41–44, Jan. 2005.
- [199] T. Cui and C. Tellambura, “An efficient generalized sphere decoder for rank-deficient MIMO systems,” vol. 9, no. 5, pp. 423–425, May 2005.

-
- [200] X.-W. Chang and X. Yang, “An efficient regularization approach for underdetermined MIMO system decoding,” in *Proc. of the Int. Conf. on Wireless Commun. and Mobile Computing (IWCMC)*. ACM, Aug 2007, pp. 349–353.
- [201] X. Chang and X. Yang, “An efficient tree search decoder with column reordering for underdetermined MIMO systems,” in *IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2007, pp. 4375–4379.
- [202] K. K.-. Wong and A. Paulraj, “Near maximum-likelihood detection with reduced-complexity for multiple-input single-output antenna systems,” in *Proc. of 38th Asilomar Conf. on Signals, Syst. and Comput.*, vol. 1, Nov. 2004, pp. 1158–1162.
- [203] K.-K. Wong, A. Paulraj, and R. D. March, “Slab-sphere decoding: efficient maximum-likelihood detection for asymmetric MIMO antenna systems,” in *IEEE 61st Veh. Technol. Conf. (VTC-Spring)*, vol. 1, May 2005, pp. 692–696.
- [204] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*, 3rd ed. MIT press, 2009.
- [205] A. K. Lenstra, H. W. Lenstra, and L. Lovász, “Factoring polynomials with rational coefficients,” *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, Dec. 1982.
- [206] X. Chang, J. Wen, and X. Xie, “Effects of the LLL reduction on the success probability of the Babai Point and on the complexity of sphere decoding,” vol. 59, no. 8, pp. 4915–4926, Aug. 2013.
- [207] A. Ghasemmehdi and E. Agrell, “Faster recursions in sphere decoding,” vol. 57, no. 6, pp. 3530–3536, Jun. 2011.
- [208] C. P. Schnorr and M. Euchner, “Lattice basis reduction: Improved practical algorithms and solving subset sum problems,” *Math. Programming*, vol. 66, no. 1, pp. 181–199, Aug. 1994.
- [209] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, “Closest point search in lattices,” vol. 48, no. 8, pp. 2201–2214, Aug. 2002.
- [210] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,” *IEEE J. on Select. Areas in Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.
- [211] D. Wubben, R. Bohnke, J. Rinas, V. Kuhn, and K. D. Kammeyer, “Efficient algorithm for decoding layered space-time codes,” *Electron. Lett.*, vol. 37, no. 22, pp. 1348–1350, Oct. 2001.
- [212] D. L. Milliner, E. Zimmermann, J. R. Barry, and G. Fettweis, “A fixed-complexity smart candidate adding algorithm for soft-output MIMO detection,” vol. 3, no. 6, pp. 1016–1025, Dec. 2009.

- [213] R. Y. Chang and W. Chung, “Best-first tree search with probabilistic node ordering for MIMO detection: Generalization and performance-complexity tradeoff,” vol. 11, no. 2, pp. 780–789, Feb. 2012.
- [214] A. S. Householder, “Unitary triangularization of a nonsymmetric matrix,” *J. of the ACM*, vol. 5, no. 4, pp. 339–342, Oct. 1958.
- [215] R. Gowaikar and B. Hassibi, “Statistical pruning for near-maximum likelihood decoding,” vol. 55, no. 6, pp. 2661–2675, Jun. 2007.
- [216] B. Shim and I. Kang, “Sphere decoding with a probabilistic tree pruning,” vol. 56, no. 10, pp. 4867–4878, Oct. 2008.
- [217] T. Cui, S. Han, and C. Tellambura, “Probability-distribution-based node pruning for sphere decoding,” vol. 62, no. 4, pp. 1586–1596, May 2013.
- [218] A. Aïssa-El-Bey, D. Pastor, S. M. A. Sbaï, and Y. Fadlallah, “Sparsity-based recovery of finite alphabet solutions to underdetermined linear systems,” *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 2008–2018, April 2015.
- [219] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [220] R. Hayakawa, K. Hayashi, H. Sasahara, and M. Nagahara, “Massive overloaded MIMO signal detection via convex optimization with proximal splitting,” in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 1383–1387.
- [221] R. Hayakawa and K. Hayashi, “Convex Optimization-Based Signal Detection for Massive Overloaded MIMO Systems,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7080–7091, Nov. 2017.
- [222] M. Nagahara, “Discrete signal reconstruction by sum of absolute values,” *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1575–1579, Oct 2015.
- [223] E. Björnson, J. Hoydis, L. Sanguinetti *et al.*, “Massive mimo networks: Spectral, energy, and hardware efficiency,” *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [224] D. Bharadia, E. McMilin, and S. Katti, “Full duplex radios,” in *ACM SIGCOMM computer communication review*, vol. 43, no. 4. ACM, 2013, pp. 375–386.
- [225] D. Bharadia and S. Katti, “Full duplex MIMO radios,” in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI) ’14*, 2014, pp. 359–372.
- [226] G. C. Alexandropoulos and M. Duarte, “Low complexity full duplex mimo: Novel analog cancellation architectures and transceiver design,” *arXiv preprint arXiv:1809.09474*, 2018.

- [227] H. Iimori, G. T. Freitas de Abreu, and G. C. Alexandropoulos, “MIMO beamforming schemes for hybrid SIC FD radios with imperfect hardware and CSI,” *IEEE Transactions on Wireless Communications*, pp. 1–15, July 2019.
- [228] OpenAI, “OpenAI Five,” <https://blog.openai.com/openai-five/>, 2018, accessed June 29th, 2019.
- [229] Y. Li, “Deep reinforcement learning,” *arXiv preprint arXiv:1810.06339*, 2018.
- [230] M. Dohler, R. W. Heath, A. Lozano, C. B. Papadias, and R. A. Valenzuela, “Is the PHY Layer Dead?” *IEEE Communications Magazine*, vol. 49, no. 4, pp. 159–165, Apr. 2011.
- [231] G. Strang, *Introduction to linear algebra*, 3rd ed. Wellesley-Cambridge Press Wellesley, MA, 1993.