Asymmetric Signaling: A New Dimension of Interference Management in Hardware Impaired Communication Systems

Dissertation by

Sidrah Javed

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology

Thuwal, Kingdom of Saudi Arabia

October, 2021

EXAMINATION COMMITTEE PAGE

The thesis of Sidrah Javed is approved by the examination committee.

Committee Chairperson:	Prof. Mohamed-Slim Alouini
External Committee Member:	Prof. Octavia A. Dobre
Committee Members :	Prof. Basem Shihada
	Prof. Ahmed El Tawil
	Prof. Taous-Meriem Laleg-Kirati
	Dr. Osama Amin

©October, 2021

Sidrah Javed

All Rights Reserved

ABSTRACT

Asymmetric Signaling: A New Dimension of Interference Management in Hardware Impaired Communication Systems

Sidrah Javed

Hardware impairments (HWIs) impose a huge challenge on modern wireless communication systems owing to the characteristics like compactness, least complexity, cost effectiveness and high energy efficiency. Numerous techniques are implemented to minimize the detrimental effects of these HWIs ,however, the residual HWIs may still appear as an additive distortion, multiplicative interference, or an aggregate of both. Numerous studies have commenced efforts to model one or the other forms of hardware impairments in the radio frequency (RF) transceivers. Many presented the widely linear model for in-phase and quadrature imbalance (IQI) but failed to recognize the impropriety induced in the system because of the self-interfering signals. Therefore, we have presented not only a rigorous aggregate impairment model along with its complete impropriety statistical characterization but also the appropriate performance analysis to quantify their degradation effects. Latest advances have endorsed the superiority of incorporating more generalized impropriety phenomenon as opposed to conventional propriety.

In this backdrop, we propose the improper Gaussian signaling (IGS) to mitigate the drastic impact of HWIs and improve the system performance in terms of achievable rate and outage probability. Recent contributions have advocated the employment of IGS over traditional proper Gaussian signaling (PGS) in various interference limited scenarios even in the absence of any improper noise/interference. It is pertaining to the additional degree of freedom (DoF) offered by IGS, which can be optimized to reap maximum benefits. This reduced-entropy signaling is the preferred choice to pose minimal interference to a

legitimate network yielding another mechanism to tackle undesired interference. Evidently, the incorporation of both inherent and induced impropriety characteristics is critical for effective utilization.

Most of the recent research revolves around the theoretical analysis and advantages of improper signaling with minimal focus on its practical realization. We bridge this gap by adopting and optimizing asymmetric signaling (AS) which is the finite discrete implementation of the improper signaling. We propose the design of both structural and stochastic shaping to realize AS. Structural shaping involves geometric shaping (GS) of the symbol constellation using some rotation and translation matrices. Whereas, stochastic shaping assigns non-uniform prior probabilities to the symbols. Furthermore, hybrid shaping (HS) is also proposed to reap the gains of both geometric and probabilistic shaping. AS is proven superior to the conventional M-ary symmetric signaling in all of its forms. To this end, probabilistic shaping (PS) demonstrates the best trade-off between the performance enhancement and added complexity.

This research motivates further investigation for the utilization of impropriety concepts in the upcoming generations of wireless communications. It opens new paradigms in interference management and another dimension in the signal space. Besides communications, the impropriety characterization has also revealed numerous applications in the fields of medicine, acoustics, geology, oceanography, economics, bioinformatics, forensics, image processing, computer vision, and power grids.

ACKNOWLEDGEMENTS

First of all, thanks to Almighty ALLAH for giving me strength, ability to learn and accomplish my academic goals. Next, I would like to recognize incisive review from committee members.

I would like to express my sincerest gratitude for my PhD advisor Professor Mohamed-Slim Alouini for his endless, unwavering and instrumental support. I am highly indebted for his continuous support and guidance throughout my stay in KAUST well beyond the level of my expectation. His compassion and empathy has indeed enabled me to complete this doctorate given a number of personal and professional challenges. He has always remained a beacon of inspiration and encouragement. Indeed, it will always remain an honor and a matter of immense privilege to be a part of his team. I would also like to acknowledge my co-supervisor Professor Basem Shihada for proposing interesting research problems and valuable feedback. I am grateful to Dr Osama Amin, Research Scientist at KAUST, for his expert opinion and passionate participation in my research work. I am also thankful to Dr Ahmed Elzanaty, Postdoctoral Fellow at KAUST, for his guidance and support during the research. I would like to acknowledge all the faculty members from the CSPN group - specially Dr Tareq Al-Nafouri and Dr. Ahmed Sultan Salem, for their utmost dedication and passion in course lectures. I am indebted to KAUST for giving me the opportunity to study in a world-class university. I express my profound gratitude to my friends and colleagues. Especially Annie, Konpal, and Basma. I will always cherish their friendship, guidance, support and celebrations.

Finally, my sincerest thanks to my family (Mama and Adnan) for their love and endless support. My greatest gratitude goes to my husband, Ali Anjum, for motivating and encouraging me in every possible way to pursue and achieve my goals. It is to my beloved sons Affan Ali and Rayyan Ali, I dedicate this thesis. May they prosper in life!

TABLE OF CONTENTS

Examination Committee Page **2**

Copyright **3**

Abstract **4**

Acknowledgements **6**

List of Abbreviations **13**

List of Figures **16**

List of Tables **18**

1 Introduction **19**

1.1 Overview 19
1.2 Background and Limitations 20
1.3 Sources of Impropriety 22
1.3.1 Communications 22
1.3.2 Power Systems 24
1.3.3 Medicine 25
1.3.4 Optics and Acoustics 25
1.3.5 Oceanography and Geophysics 26
1.4 Motivation 27
1.4.1 Impairment Modeling and Characterization 27
1.4.2 Improper Characterization 29
1.4.3 Adaptive Asymmetric Signaling 30
1.4.4 Remarks 31
1.5 Literature Review 32
1.5.1 Impairments Model Incorporation 32
1.5.2 Error Probability Analysis of the Improper Systems 34
1.5.3 Improper Characterization 35

1.5.4 Adaptive Asymmetric Signaling 37
1.5.5 Deductions 39
1.6 Objectives and Contributions 39
1.7 Organization and Road map 43
1.8 Notations 43

2 Technical Framework 45
2.1 Stochastic Data Modeling 46
2.1.1 Complex Random Vectors 46
2.1.2 Quaternion Random Vectors 50
2.1.3 Summary and Insights 53
2.2 Propriety versus Circularity 53
2.2.1 Complex Random Vectors 54
2.2.2 Quaternion Random Vectors 62
2.2.3 Summary and Insights 65
2.3 Transformations and Operations 66
2.3.1 Complex Random Vectors 66
2.3.2 Quaternion Random Vectors 69
2.3.3 Summary and Insights 70
2.4 Entropy and Probability Density Functions 70
2.4.1 Complex Random Vectors 70
2.4.2 Quaternion Random Vector 72
2.4.3 Summary and Insights 74
2.5 Testing for Impropriety 74
2.5.1 GLRT for Complex Random Vectors 76
2.5.2 GLRT for Quaternion Random Vectors 80
2.5.3 Summary and Insights 81

3 Hardware Impairment Modeling 82
3.1 Transceiver Hardware Distortion Model 83
3.2 Transceiver I/Q Imbalance Model 85

4 Hardware Impaired Multihop Full-Duplex Relaying Systems 88
4.1 Significance and Contributions 89
4.2 FDR under HWD System Model 90
4.3 Achievable Rates 92
4.4 HWD- and RSI-Aware Signaling Design 93

4.4.1 Joint Optimization 94
4.4.2 Distributed Optimization 99
4.4.3 Complexity Analysis 105
4.5 Numerical and Simulation Results 106
4.5.1 Effect of RSI and HWD 107
4.5.2 Number of Participating Relays 109
4.5.3 Impact of Cluster size and Placement in Distributed Algorithms . . 110
4.6 Conclusion 112

5 Asymmetric Signaling for HWI Systems 114
5.1 Main Contributions 115
5.2 System Description 115
5.3 Receiver Design 117
5.3.1 Optimal Maximum Likelihood Receiver 117
5.3.2 Minimum Euclidean Distance Receiver 119
5.3.3 Linear Minimum Mean Square Error Receiver 119
5.4 Error Probability Analysis 120
5.4.1 Symbol Error Probability 120
5.4.2 Asymptotic Analysis 122
5.5 Average Probability of Error 123
5.5.1 System with Transmitter and Receiver Hardware Impairments . . . 123
5.5.2 Zero-Distortion Transmitter 127
5.5.3 System with Negligible Transmitter I/Q Imbalance 129
5.5.4 System with Negligible Receiver I/Q Imbalance 130
5.6 Transmit Signaling Design 132
5.6.1 Special Cases 134
5.6.2 Maximum Pairwise Error Probability Minimization 134
5.6.3 Maximum Symbol Error Rate Minimization 135
5.7 Numerical Results 136
5.7.1 Optimal Receiver 137
5.7.2 Bounds and Approximations 139
5.7.3 Asymmetric Transmission 141
5.8 Conclusion 143

6 When Probabilistic Shaping Realizes Improper Signaling for Hardware Distortion Mitigation 145
6.1 Significance and Contributions 145

6.2 System Description 146
6.2.1 Transceiver Hardware Distortion Model 146
6.2.2 Optimal Receiver 148
6.3 Error Probability Analysis 149
6.4 Proposed Probabilistic Signaling Design 151
6.4.1 Problem Formulation 152
6.4.2 Optimization Framework 153
6.4.3 Toy Examples 155
6.5 Hybrid Shaping where Conventional meets State-of -the-Art 157
6.5.1 Hybrid Shaping Parameterization 157
6.5.2 Optimal Receiver 158
6.5.3 Error Probability 159
6.5.4 Problem Formulation 160
6.5.5 Proposed Algorithm 162
6.5.6 Illustrative Example 164
6.6 Numerical Results 164
6.7 Conclusion 171

7 Theoretical Analysis and Performance Limits 172
7.1 Achievable Rate 172
7.1.1 Average Achievable Rate Limits 175
7.1.2 Rate Region Analysis 178
7.2 Outage Probability 186
7.2.1 Rate Outage Probability 186
7.2.2 Secrecy Outage Probability 187
7.3 Power Efficiency 188
7.4 Degrees of Freedom 190
7.5 IGS Signaling Design 190
7.5.1 Closed Form Solutions 191
7.5.2 Convex Optimization 192
7.5.3 Non-Convex Optimization 193
7.5.4 Intractable Optimization Framework 198
7.5.5 Summary and Insights 199
7.6 IGS Detection and Estimation 200
7.6.1 Detection 200
7.6.2 Estimation 203

7.6.3 Source Separation 206

8 Practical Realization and Implementation 209

8.1 Asymmetric Signal Design 209

8.1.1 Probabilistic Shaping 211

8.1.2 Geometric Shaping 212

8.1.3 Orthogonal/Non-Orthogonal Sharing 216

8.1.4 Hybrid Signaling 217

8.1.5 Deep Learning 218

8.2 Asymmetric Signal Recovery 218

8.2.1 Equalization 219

8.2.2 Estimation 219

8.2.3 Filtering 220

8.2.4 Detection 221

8.3 Error Probability Analysis 222

8.3.1 Multiuser Direct Sequence Multiple Access Systems 222

8.3.2 Coded Communication Systems 223

8.3.3 Multiple Antenna Systems 224

8.3.4 Multi-carrier and Single-carrier Systems 226

8.3.5 Systems with Frequency Selective Channels 227

8.3.6 Systems with Improper Noise 227

8.3.7 Hardware Impaired Systems 228

8.3.8 Multiuser Interference Channel 229

9 Applications 232

9.1 Data Analysis 232

9.1.1 Independent Component Analysis 232

9.1.2 Principal Component Analysis 234

9.1.3 Complex Least Mean Square Analysis 236

9.1.4 Miscellaneous 237

9.2 Signal Processing 238

9.2.1 Array Processing 238

9.2.2 System Identification and Feature Extraction 238

9.2.3 Pattern Recognition and Image Processing 239

9.3 Communication Systems 239

9.3.1 Cellular Networks 239

9.3.2 Cognitive Radios 239

9.3.3 Full Duplex Systems 240
9.3.4 Hardware Impairments Mitigation 241
9.3.5 Phase Estimation 242
9.3.6 Interference Channels 242
9.3.7 Noisy Channels 243
9.3.8 Trelli's Coding 243

10 Conclusion **245**
10.1 Summary 245
10.2 Inferences 247
10.3 Lessons Learned 249
10.3.1 Data Representations 249
10.3.2 Inherent Impropriety 250
10.3.3 Imposed Impropriety 250
10.3.4 Design, Detection, and Estimation 251
10.3.5 Practical Realization 252
10.4 Challenges and Way Forward 253
10.4.1 Impairment Modeling and Incorporation 253
10.4.2 System Identification 253
10.4.3 Transmission Design 254
10.4.4 Parameter Optimization 255
10.4.5 Joint and Disjoint Design 255
10.4.6 Performance Analysis 256
10.4.7 Time-Sharing 257
10.4.8 Realization 257
10.4.9 Implementation Cost 258
10.4.10 Recommendations 259
10.5 Concluding Remarks 260

References **261**

Appendices **298**

LIST OF ABBREVIATIONS

AS	asymmetric signaling
AWGN	additive white Gaussian noise
BC	broadcast channel
BER	bit error rate
CDF	cumulative distribution function
CLT	complex linear transformation
CSI	channel state information
DC	difference of convex
DF	decode-and-forward
DoF	degree of freedom
DoI	degree of impropriety
DS-CDMA	direct-sequence code division multiple access
ECG	electrocardiogram
EEG	electroencephalogram
EOG	electroculogram
EVD	eigen-vlaue decomposition
FD	full duplex
FDR	full duplex relaying
fMRI	functional magnetic resonance imagine
FSO	free space optics
GLRT	generalized likelihood ratio test
GS	geometric shaping
HD	half duplex
HS	hybrid shaping
HWD	hardware distortions
HWI	hardware impaired
HWIs	hardware impairments
IBC	interference broadcast channel
IC	interference channel
IGS	improper Gaussian signaling

IQI	in-phase and quadrature imbalance
LED	light-emitting diode
LMMSE	linear minimum mean square error
LRT	likelihood ratio test
LT	linear transformation
MAP	maximum a posteriori
MED	minimum Euclidean distance
MH	multihop
MH-DF-FDR	multihop decode-and-forward FDR
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
ML	maximum likelihood
NC	non-circular
NNUB	nearest neighbor union bound
OFC	optical fiber communication
OFDM	orthogonal frequency division multiplexing
PAM	pulse amplitude modulation
PDF	probability density function
PEP	pairwise error probability
PGS	proper Gaussian signaling
PoD	probability of detection
PS	probabilistic shaping
PSD	positive semi-definite
PSK	phase shift keying
QAM	quadrature amplitude modulation
QCLP	quadratic contraint linear programming
QoS	quality of service
RF	radio frequency
RLT	real linear transformation
RSI	residual self-interference
RV	random vector
RX	receiver
SCP	sequential convex programming
SEP	symbol error probability
SER	symbol error rate
SI	self-interference

SIMO	single-input multiple-output
SISO	single-input single-output
SNR	signal-to-noise ratio
SOS	second-order statistics
TX	transmitter
WLT	widely linear transformation
WUT	widely unitary transformation
Z-IC	one-sided interference channel

LIST OF FIGURES

1.1 Interdisciplinary Sources of Improperness 23

2.1 Extents of Circularity and their Implications 56
2.2 The Interplay between Propriety and Circularity 58
2.3 Extents of Propriety and their Implications 59

3.1 Impropriety Characterization 85
3.2 Transmitter I/Q imbalance 86

4.1 FDR System under HWD and RSI. 91
4.2 Different Sized Clusters with Their Two Possible Placements 100
4.3 Impact of RSI and HWD on System Performance 107
4.4 Average Achievable Rate Performance. 110
4.5 Average achievable rate versus self-interference for various cluster sizes. . . 111

5.1 HWI-Aware Receivers 137
5.2 Bounds and Approx on $\bar{P}_{\mathrm{b}}$ 138
5.3 Asymmetric Transmission Schemes 140

6.1 8-QAM probability distribution at Rate = 2 bits/symbol. 156
6.2 8-QAM probability distribution at Rate = 2 bits/symbol. 157
6.3 Different asymmetric signaling designs for proper and improper aggregate interference 163
6.4 BER Performance for a range of EbNo in AWGN channel. 166
6.5 Throughput Performance in AWGN channel. 167
6.6 HWD mitigation for various AS in different channels. 168
6.7 BER Performance for a range of EbNo 169

7.1 Overview and Configuration of Theoretical Analysis 173
7.2 Percentage Reduction in Outage Probability with IGS 188
7.3 Conceptual Classification of IGS Detection, Estimation and Identification . 201

8.1 Practical Implementation of IS: From Realization to Recovery and Analysis 210

8.2 Celebrated 16-QAM Lattice Structures 213
8.3 Geometrical Packing/Mask Structures 214

9.1 Applications of Impropriety in Data Analysis, Signal Processing and Communication Networks 233

10.1 Lessons Learned and Way Forward 249

LIST OF TABLES

2.1 Consequences of Vanishing Pseudo-Variances 63
2.2 Impropriety Evaluation based on various Tests involving Different Data Representations 75
2.3 Test Thresholds for $\mathbb{C}-$GLR Hypothesis Tests 78
2.4 Quaternion GLR Hypothesis Tests and Decision Criterion 81

4.1 Complexity Analysis of the proposed algorithms 106
4.2 Rayleigh Fading and Free Space Distance Pathloss Model 109

5.1 Rotation angles for maximal symbol separation in one-dimension 133
5.2 Percentage Error of Gaussian Quadrature Legendre Polynomial Approximation 141

7.1 Achievable Rate Improvement by IGS over PGS in Different Settings . . . 179
7.2 Outage Probability Reduction with IGS Relative to PGS 187

8.1 Error Probability Reduction with Appropriate Modeling (Signaling, Filtering, Estimation or Detection) 230

K.1 First Order Necessary KKT Conditions 306

Chapter 1

Introduction

Exponentially rising demands of high data rates and reliable communications given the limited power and bandwidth resources impose enormous challenges on the next generations of wireless communication systems [1–4]. It has compelled researchers to think beyond the traditional approaches and techniques. Various research contributions propose new configurations and novel techniques to address these challenges [5, 6]. Nonetheless, the performance of such systems can be highly degraded by the hardware imperfections in RF transceivers [7–9].

Hardware impairments (HWIs) impose a huge challenge on next-generation network planning and deployment especially at high-frequency [3,7–9]. HWIs emerge in various RF stages including imperfections in analog-to-digital/digital-to-analog converters, non-linear high power amplifier/low noise amplifier, mismatched local oscillator and phase shifter, etc. [8–14]. HWIs are considered as the biggest limitation while considering the deployment of massive amount of circuitry operating at relatively higher frequencies [13, 15]. Therefore, RF impairments form a key design challenge for developing new techniques/configurations in modern wireless communications, as these imperfections dominate the performance of the overall system [8].

1.1 Overview

We use signals either to quantify real-world physical quantities or to exchange information. Thus, we require appropriate signals and models for accurate performance analysis and de-

sign of the underlying systems. Most of the studies adopt circularly symmetric (proper) complex distribution for analysis rendering simplified computations owing to it's uncorrelated and equal power real and imaginary components. However, a more generalized approach is to deal with circularly asymmetric (improper) signals which may have correlated/unequal components and contains the proper signals as a special case.

Communication systems conventionally employ proper signals pertaining to their mathematical tractability and optimal performance in P2P interference-free system setup. However, with the every increasing demands of wireless connectivity, most of the communication systems are limited by the interference/noise. This demands the adoption of the more generalized improper/asymmetric signaling to improve the system performance. In particular, improper Gaussian signaling has recently gained popularity for it's superiority in various interference-limited system configurations. The theoretical limits achieved by IGS can be practically realized using adaptive asymmetrical discrete constellations. Asymmetric signaling can be implemented using the structural and/or stochastic shaping of the symmetric constellation.

The research work studies the theoretical impact of improper Gaussian signaling on communication networks based on fundamental information theory studies, and realizes practical systems using asymmetric discrete constellation based on appropriate signal processing techniques.

1.2 Background and Limitations

In the backdrop of the aforementioned overview, the effectiveness of numerous research contributions is limited due to certain deficiencies.

- Many applications in applied sciences employ complex analysis to model the real-world data in complex domain. The complex analysis brings advantages like comprehensiveness, compactness, computational economy, extra dimension, elegant analysis and much more. Paradoxically, researchers resort to the more tedious two-

dimensional real modeling of the complex variables for their statistical investigation [16] and the complex theory lost most of its beauty, elegance, and interest [17].

- Majority of the contributions assumed proper signal model for the underlying complex phenomenon rendering simplified computations, which is in contrast with most of the real-world scenarios [18]. The improper nature of complex models has been proven for various entities in numerous diverse fields. For instance, the accumulative additive thermal noise model in communication systems [19, 20], complex envelop of the scalar optical fields [21], empirical speech model [22], complex traces of the seismic signals [23], ocean-current spectra [24], wind fields [25–27] and fluid dynamics [28], complex valued model of unbalanced three-phase voltage in power systems [29, 30] and neural activity in brain/spinal cord as measured by functional magnetic resonance imaging (fMRI) [31–33].

- The ideal assumption of the perfect transceiver hardware or the effective impairment mitigation allows researchers to resort to the conventional simplified system models. However, inaccurate parameters estimation, unreliable distortion modeling, and unsophisticated compensation algorithms with limited capabilities lead to a certain residual deterioration effect which is not neutralized by these strategies [34]. Additionally, HWIs like transceiver in-phase and quadrature-phase (I/Q) imbalance (IQI) result not only in phase/amplitude errors and raised noise floor but also in an inevitable mixing of the desired and image signals. This transforms proper signal to improper signal as well as proper thermal and distortion noise to improper accumulative noise in the wireless communication systems [35]. It is worthy to emphasize that the existing performance analysis literature does not investigate whether these hardware impairments exhibit symmetric or asymmetric characteristics. On the other hand, statistical signal processing research highlighted the asymmetric characteristics of baseband communication signals due to the I/Q imbalance [31]. Residual

improper HWIs can highly degrade the effective signal-to-interference plus noise ratio (SINR), achievable rate, system outage probability, system error probability, and throughput causing overall performance degradation.

These defficiencies and superficial assumptions not only lead to inaccurate modeling, misleading analysis, and incorrect results but also deprives us from seeking the maximum benefit of the additional design freedom [36, 37]. Thus, motivating researchers to develop accurate models and propose effective compensation methods to meet the expected performance.

1.3 Sources of Impropriety

The most common speculation in the complex analysis assumes propriety and/or circularity of the r.v. under investigation. This phenomenon assumes uncorrelated real and imaginary components of the complex entities with identical variance which is generally not true. This section reveals numerous examples from real-world problems in diverse fields which render improper complex signatures e.g., in medicine [31–33, 38–42], oceanography [24, 43, 44], geology [23, 25–27, 45–48], optics [21, 49, 50], acoustics [22, 51, 52], power systems [29, 30, 53, 54]. and communication systems [20, 31, 36, 55–72].

The two main sources of impropriety include asymmetric signals and asymmetric data. Asymmetric signals may occur naturally or result from some transforming phenomenon whereas empirical data is generally asymmetric. The improper/asymmetric signatures in numerous fields are highlighted in various technical contributions as shown in Fig. 1.1.

1.3.1 Communications

A necessary, yet insufficient condition for impropriety is that the real-valued random process be at least nonstationary [73]. This implies that the complex baseband representations of wide-sense stationary thermal noise and nonstationary transmitted data signals are appropriately modelled as proper and improper (may reduce to proper) [63]. Thus, the po-

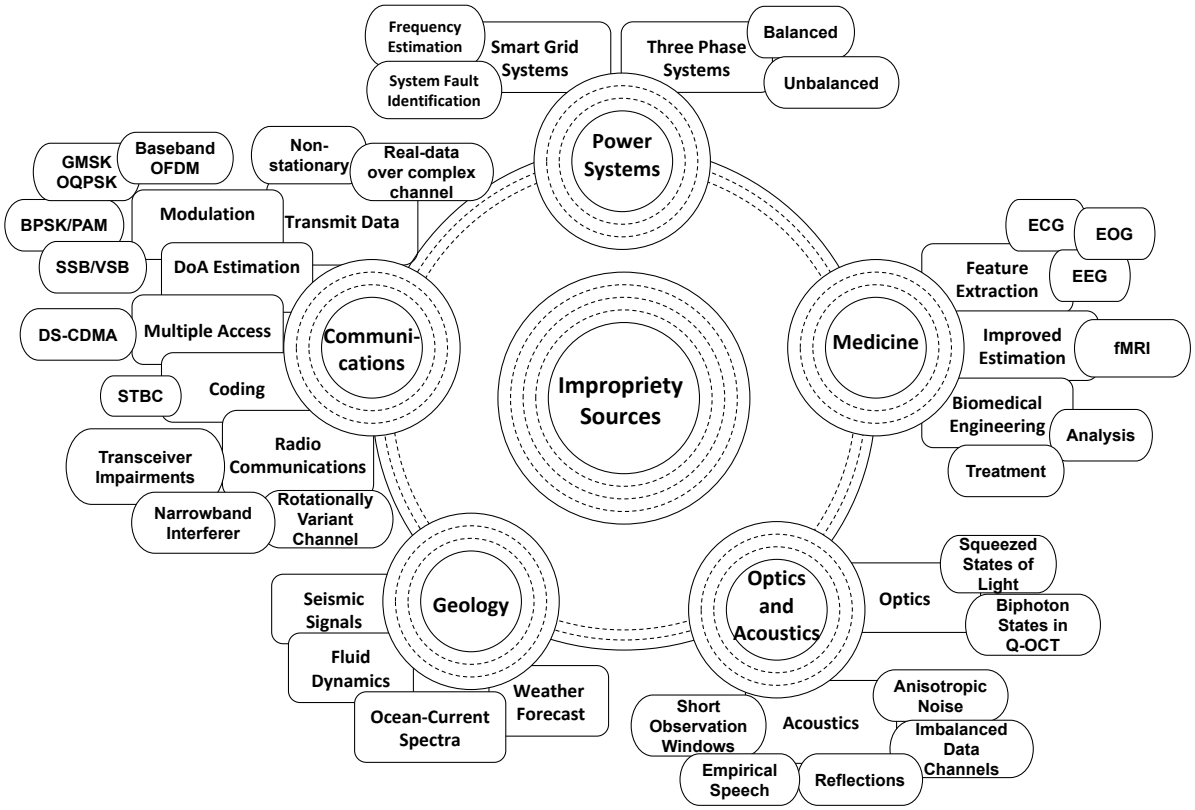


Figure 1.1: Interdisciplinary Sources of Improperness

tentially improper nature of signals must be taken into account when designing detection algorithms.

Improperness may originate from various modulation, coding or access schemes. Important examples of such digital modulation schemes include BPSK [55, 56], PAM [57], GMSK [58, 59], OQPSK or staggered QPSK [60], SSB, VSB [61] and baseband (but not passband) orthogonal frequency division multiplexing (OFDM) [74]. On the other hand, coding schemes like STBC also result in improper signals [62]. Interestingly, the real-valued data transmission over a complex-valued channel also results in non-zero pseudo-covariance of the received observations or non-zero cross pseudo-covariance between the received observations and the desired variable, thus resulting in improper received signal. Such a scenario arises in the GSM, binary CPM, offset QAM [65] and general simplex signals. In many sensor applications, DoA estimation with electronically steerable antenna

arrays is improved by employing NC signal constellations [36, 63, 64]. In addition, multiple access scheme like direct-sequence code division multiple access (DS-CDMA) [75] with narrowband interferer results in NC interference [65]. Consequently, such modulation, coding and multiple access procedures can corrupt sensor observations with NC complex noise. This necessitates appropriate observation models and distributed complex filters to tackle them [66].

In areas such as radio communications, beamforming and spectral sensing, the amplitude and phase imbalances between its I/Q components may also introduce non-circularity in the received signals [67]. Additionally, hardware impairments (HWIs) like non-linearity, IQI and phase noise rise in various RF and baseband blocks such as analog-to-digital convertors, high power and low noise amplifiers, low-pass and band-pass filters etc. [8, 9]. These impairments accumulate at both the ttransmitter (TX) and receiver (RX) RF chains and yield undesired effects. For instance, non-linearities result in additive Gaussian distortion noise [12, 14, 76], whereas IQI not only induces phase and amplitude errors but also mixes the desired and image signals [8,69,77,78]. This can be caused by transceiver imperfections [20,68–70], communication channels that are not rotationally invariant [31] or NC interference from other sources [57]. Besides wireless communications, other noise models such as underwater propeller noise from maritime ship also demonstrate an improper nature [42].

1.3.2 Power Systems

Accurate real-time estimation of system frequency is a major technical challenge for future smart grids with dynamically updating generation and loading topology. Recently, the complex signal retrieved after the Clarke's ($\alpha\beta$) transformation of unbalanced three-phase voltages is shown to be second-order NC [29, 30]. Therefore, the augmented statistics and the corresponding WL models are exploited to manipulate varying degrees of non-circularity relative to different frequency variation sources. Hence supporting the estima-

tion of instantaneous frequency in a three-phase system under both balanced and unbalanced conditions [53, 54]. The complete second-order information offers next-generation solutions for accurate, adaptive and robust frequency estimation as well as system fault identification [29].

1.3.3 Medicine

In medical data analysis, non-circularity is exploited for features extraction and improved estimations in electrocardiogram (ECG) [38] and functional magnetic resonance imagine (fMRI) [39, 40], respectively. The fMRI measured neural activities in the brain or spinal cord are recently recognized as improper signals [31–33]. Thus, the NC probability distributions of the sources enable real-time extraction of eye muscle activity: electroculogram (EOG) from the electroencephalogram (EEG) recordings using some blind source extraction algorithms [41]. Rhythms of brain waves are non-stationary signals [42] and thus potentially improper because the temporal information arises from correlativity in the frequency domain. Hence, NC characteristics are a useful resource for suitable analysis and treatment in biomedical engineering.

1.3.4 Optics and Acoustics

Analogous to stochastic complex fields, arbitrary second moments of the complex envelopes of scalar optical fields are also completely characterized by its phase-insensitive and phase-sensitive correlation functions [21]. Phase-insensitive correlation function is analogous to the covariance matrix whereas phase-sensitive correlation function is equivalent to the pseudo-covariance matrix in the optics. Thus, the non-zero phase-sensitive correlation function marks improper optical field. Conventional light sources like sunlight, light-emitting diode (LED) and lasers have trivial phase-sensitive correlation whereas advanced non-linear optical sources like squeezed states of light possess phase sensitive correlation [49]. The biphoton states in quantum optical coherence tomography (Q-OCT)

depict entanglement properties and offer various benefits owing to the phase-sensitive correlation between its two photons [50]. Optical coherence theory should accommodate phase-sensitive fluctuations for complete second order characterization pertaining to the significant propagation differences between phase-insensitive and phase-sensitive optical sources.

It is quite natural to assume real-world acoustic signals as NC owing to the anisotropic noises, unequal powers of data channels and reflections, and short observation windows. Thus, augmented statistical framework is required to accurately model the acoustic sources for speech recognition [51, 52]. It is based on the fact that speech can be empirically improper in the frequency domain [22].

1.3.5 Oceanography and Geophysics

Complex trace analysis of seismic signals [23], interpretation of ocean-current spectra [24] and the analysis of wind fields [25–27] exploit the improperness of their respective complex representations. The impropriety concept is well-known and appreciated by meteorologists and oceanographers since the early 1970s [43, 44]. Recently, the improper nature of the real-time wind data along with the tracking of degree of circularity is identified for renewable energy applications [46,47]. Similarly, adaptive filtering of the real-world wind signals is based on its NC and nonstationary observations [48]. Others employ impropriety concepts to further physical applications with asymmetric/anisotropic data structures e.g., in fluid dynamics [28] and seismic data signals [45] to overcome modeling challenges. Most of the above mentioned contributions are backed by [37], which claims the improper nature of signals involved in various real problems.

This section highlights the existence of the impropriety concepts in diverse science and engineering domains. Therefore, such real-world problems require rigorous characterization for appropriate analysis and treatment.

1.4 Motivation

Besides covering the broader aspects of various disciplines, this thesis is focused on the communication systems with intentional/unintentional improper signatures. Next we highlight some important deriving forces which motivated our research in impairment modeling, impropriety characterization, and improper/asymmetric signaling.

1.4.1 Impairment Modeling and Characterization

Inevitable imperfections/impairments can render drastic effects on the system performance. Therefore, it is imperative to inculcate them in the system model and design problems. In addition, the effective compensation schemes require an accurate statistical model of these imperfections including hardware impairments and undesired self-interference whenever they arise.

Hardware Impairments

Hardware impairments (HWIs) impose a huge challenge on next-generation network planning and deployment especially at high-frequency [3, 7–9]. HWIs emerge in various RF stages including imperfections in analog-to-digital/digital-to-analog converters, non-linear high power amplifier/low noise amplifier, mismatched local oscillator and phase shifter, etc. [8, 9]. These hardware imperfections result not only in phase/amplitude errors and raised noise floor but also in an inevitable mixing of the desired and image signals. Especially, the performance of multihop systems can be severely affected by the hardware distortions (HWD) due to the accumulated effect of impaired transceivers at the participating nodes.

Hardware distortion is caused by different circuit impairments such as the phase noise and non-linear distortion [8]. They render deteriorating effects on the wireless communication systems performance. On the other hand, the widely linear model of IQI is backed

by some solid contributions [78–80]. Thus, according to statistical signal processing studies, widely linear precoders/transformations can efficiently map symmetric information-bearing signals to asymmetric signals [31, 81]. However, only few works have acknowledged the asymmetric characteristics induced by this WL structure of IQI [20, 68, 80, 82]. The WLT of IQI at the transmitter transforms PGS to IGS, which can be verified by finding non-zero pseudo-variance of the resulting signal [79]. In contrast, the receiver IQI is mainly responsible for the improper Gaussian additive noise [71,72]. Therefore, transceiver impairments specifically IQI play a vital role in transforming the propriety of the information signals and undesired noise/interference [20, 70, 83–85]. This motivates an accurate modeling and thorough investigation of the multiple HWIs.

Self Interference in Full Duplex Relaying Systems

Several studies have been carried out to investigate extreme node densification and collaborative radio technologies to improve the spectral efficiency and meet the exponentially growing wireless data traffic demands [86, 87]. Relaying technology has gained much interest as it can be used in different network topologies and applications, to improve the quality-of-service, such as unmanned autonomous vehicles and self-driving robots [88]. Particularly, multi-hop relaying can significantly extend the coverage and improve energy efficiency [89, 90]. However, spectral efficiency decreases with the increase of number of relays, where the frequency (or time) is shared between the nodes [91]. FDR is a promising technology that can compensate spectral efficiency loss by allowing each node to transmit and receive simultaneously [92]. However, it drastically suffers from self-interference (SI) that can limit it's operation. The SI signal is relatively larger than the desired signal of interest, which increases the dynamic range span of the low-noise amplifier (LNA) and the analog-to-digital converter (ADC) at the receiver side. As a result, both the undesired interference and the hardware noise levels increase, which can greatly suppress or even destroy the information bearing signal [93]. Despite the multiple analog and digital cancellation

stages, the FDR performance is still limited by the residual self-interference (RSI). Hence, it needs to be taken into account for effective compensation.

1.4.2 Improper Characterization

Divergence from norms/conventions can sometimes beat the traditional and long-practiced assumptions. Same is the case with long assumed circularly symmetric complex (CSC) signal conjecture followed before the advent of last three decades. The advancement from the real stochastic domain to the complex stochastic domain came with the naive assumption of equal energy and uncorrelated real and imaginary components of a complex random entity, later named as proper complex random variables (r.v.) which is generally not true. This section presents the motivation behind the impropriety specifications and implementation. The theoretical implementation can be seen as IGS whereas the practical implementation can be realized as asymmetric discrete signaling.

Impropriety and Circularity

A proper complex r.v. is uncorrelated with its complex conjugate whereas any correlation between the two results in improper complex r.v. The investigation of the presence and absence of this property is coined as propriety and impropriety, respectively. Another related yet distinct phenomenon is the concept of CSC or circular r.v., demonstrating rotationally invariant probability distribution in the complex plane [17, 31, 36, 55, 57, 94, 95]. The absence of this property renders non-circular complex r.v. and the evaluation study to determine the circular and non-circular nature of complex entities is termed as circularity and non-circularity, respectively.

Circularity is an assumption that was originally introduced for the definition of the probability distribution function [17]. For instance, the PDF of a complex Gaussian random vector (RV) assumes the anticipated and familiar natural form only for proper RVs [55, 94]. Moreover, the definitions of independence and/or correlation are inadequate without

complete second order statistical (SOS) characterization. For example, it is shown that contrary to the real case, uncorrelated complex normal r.vs. are not generally independent [16]. Similarly, the propriety characterization is inevitable for the appropriate treatment and accurate entropy quantification of the complex signals and complex impulse responses of the equivalent base-band channels.

Improper Gaussian Signaling

IGS is a generalized complex signaling scheme that relaxes the symmetric characteristics of PGS scheme allowing a correlation between the signal components and/or unequal power of each component, as opposed to proper Gaussian signaling [94]. Hence, IGS incorporation can offer an additional degree of design, pertinent to its circularly asymmetric nature characterized by the circularity coefficient [57].

1.4.3 Adaptive Asymmetric Signaling

Despite the overwhelming benefits of IGS, it is practically infeasible owing to the high detection complexity and unbounded peak-to-average power ratio [2, 96]. This motivates the researchers to design some equivalent finite and discrete asymmetric signaling (AS) schemes for practical implementation. Such AS (improper discrete constellation) entails redesigning the symmetric discrete signal constellation to convert it into an asymmetric signal [2]. This can be achieved through the structural or/and stochastic shaping of the finite discrete symmetric constellations such as M-ary QAM.

Structural Shaping

Several studies focused on geometric shaping (GS) to induce structural asymmetry in order to improve the system performance. GS transforms equally spaced symbols to unequally spaced symbols (due to correlated and/or unequal power distribution between quadrature components of the symbols) in a distinct geometric envelop such as ellipse [97], parallelo-

gram [96, 98] or some irregular envelop [99]. Thus, structural shaping demands incorporation and further investigation for practical implementation.

Stochastic Shaping

The asymmetric discrete family of constellations is practical, but they exhibit two types of loss, i.e., shaping loss and packing loss in approaching IGS theoretical limits [96]. Most of the efforts to close the gap between AS and ideal IGS are concentrated around GS with a limited focus on stochastic/probabilistic shaping as another way to implement AS. Given a fixed number of symbols and the symbol locations, an asymmetric constellation can be obtained by adjusting the symbol probabilities [100]. PS maps equally distributed input bits into constellation symbols with non-uniform prior probabilities [101]. This can be achieved using DM for rate adaptation such as constant composition DM [102], adaptive arithmetic DM [103], syndrome DM [104, 105] DM-based compressed sensing [106, 107]. The idea is to employ a higher-order constellation (with non-uniform probabilities) as opposed to the uniform constellation, while targeting a minimum transmission rate. This offers additional DoF and adaptive rates.

1.4.4 Remarks

The detailed and thorough research on IGS and AS schemes encouraged us to put-forth a comprehensive tutorial encompassing all the technical preliminaries. It also surveys all the contributions in the journey from theoretical limits to practical realization pertaining to the impropriety concepts. IGS enables improved performance limits in terms of achievable rate, outage probability, power efficiency, and degree of freedom. On the other hand, the practical implementation of these impropriety concepts using asymmetric signaling improve the quality of service and reliability of the interference-limited systems.

1.5 Literature Review

This section presents numerous contributions in the adopted research streams such as impairments incorporation and compensation, impropriety characterization and implementation, rigorous and extensive performance analysis, and adaptive asymmetric signaling. Followed by the concluding remarks enclosed in the deductions referring to the deficiencies in these contributions and way forward.

1.5.1 Impairments Model Incorporation

The taxonomy of the literature emphasizing the notable hardware impairments and compensation schemes in different system setups is presented in this subsection.

Hardware Distortions

Numerous efforts have been carried out to accurately model various forms of HWIs such as HWD and IQI. Many studies focused on the statistical modeling of additive hardware distortions at the transmitter and the receiver [10–13, 76, 108, 109]. However, few contributions emphasized the distinct improper behavior of these HWDs [19, 83, 84, 110]. Such characterization requires dedicated compensation techniques to meet the performance demands.

I/Q Imbalance

Numerous contributions present the models of in-phase and quadrature-phase (I/Q) imbalance where the self interference (SI) signal induces the amplitude and rotational imbalance besides the receiver thermal noise. The effects of IQI imbalance were investigated for different communication systems in [12, 14, 72, 72, 78, 111–116]. Various solutions have been proposed to improve overall system performance in terms of achievable rate and outage probability in the presence of I/Q imbalance at tranceivers. For example, [114, 115] dis-

cusses an opportunistic relaying scheme to compensate IQI and loop-back self-interference in orthogonal frequency-division multiplexing FDR system to reduce outage probability. Additionally, Boulogeorgos *et al.* proposes an I/Q imbalance self-interference coordination scheme to improve system diversity and quality of the received signal but at the expense of reduced transmission rate [111]. Similarly, Mokhtar *et al.* discusses an opportunistic relaying scheme to compensate IQI and loop-back self-interference in orthogonal frequency-division multiplexing FDR system [115].

Joint Hardware Impairments

Few studies analyzed multiple RF front-end impairments and their individual baseband equivalent error models [8, 117]. Schenk studied the modeling procedure, impact of non-ideal hardware on the system performance, and digital compensation schemes of various RF imperfections in high data-rate wireless systems [8]. Similarly, Boulogeorgos *et al.* studied the impact of various HWIs on the energy detection spectrum sensing in cognitive radio systems in [117].

RSI and HWI in FDR

Various studies have been carried out to address the deteriorating impact of RSI or HWD in full-duplex communication systems. For example, the effects of HWIs were investigated for different full-duplex communication systems in [12, 14, 71, 72, 78, 112, 114, 115] and particularly for relay systems in [12, 14, 71, 72, 114, 115]. Of all these contributions, [112] proposes a digital cancellation scheme to supplement RF/analog cancellation techniques for self-interference mitigation in single-channel FD wireless communication. On the other hand, [12] focuses on a low complexity hardware impairments aware transceiver scheme to mitigate distortions in the transmitter and the receiver. Likewise, [71, 72, 114] analyze the system performance under IQI at relay and destination with ideal transmitter considering a half-duplex amplify-and-forward (AF) relay, orthogonal frequency-division mul-

tiplexing (OFDM), and IQI at all nodes in OFDM dual-hop opportunistic AF relaying. Similarly, other contributions [118–122] presented power control mechanisms to improve system throughput or outage performance in the presence of RSI and inter-relay interference assuming ideal transceivers.

Very few studies focus on both RSI and HWD in a full-duplex operation mode. For example, [14] carries out the performance analysis of dual-hop proactive DF relaying networks with best relay selection under hardware impairment and co-channel interference. Similarly, [115] analyzes the outage probability of dual-hop DF FDR for an OFDM system in the presence of IQI and loop-back SI. Besides carrying out the performance analysis, few works proposed some compensation schemes to improve the system performance. For example, [78] proposes a novel widely linear digital cancellation processing to mutually mitigate SI and practical hardware imperfections in direct-conversion FD transceiver. In [123] and [124], authors proposed compensation schemes for various HWIs and RSI at the AF relay(s), considering dual-hop HWI-FD-AF-relay system with ideal source and destination. This motivates us to propose a mitigation signaling design to concurrently combat RSI and transceiver distortions at the source and destination besides multiple HWI-FD-relays. Furthermore, we have focused on the DF relaying strategy in-place of AF relaying scheme, to support the communication in an interference limited environment, in order to meet next generation traffic demands.

1.5.2 Error Probability Analysis of the Improper Systems

The error probability performance analysis of systems subjected to HWIs has recently received few attentions [125–128]. Windisch and Fettweis quantified the impact of the receiver I/Q imbalance in terms of closed-form error probability [125],. Similarly, Qi and Aissa studied the bounds on average symbol error probability (SEP) for the receiver I/Q imbalanced MIMO system [126]. Moreover, Krishnan *et al.* derived the SEP for systems suffering from Gaussian phase error [127]. Additionally, Bouhlel *et al.* analyzed the pair-

wise error probability (PEP) over correlated Rayleigh and Ricean fading channel and incorporated both imperfect channel state information and additive hardware distortions [128]. To the best of author's knowledge, none of these works incorporated the improper characteristics of underlying HWIs while carrying out the error probability analysis and deriving the optimal detectors. Therefore, the accuracy of the analysis and derived results is questionable and demands appropriate treatment.

1.5.3 Improper Characterization

Interestingly, the propriety concept is not unheard-of and vast majority of contributions were put forward in one domain or the other. This concept is quite popular in the statistics,signal processing, and information theory community but yet to find it's standing in the communication theory.

Impropriety and Circularity

Improper processes and models have seen growing interest in the statistics community [63, 129, 130]. Similarly, the pioneering works in signal processing community which emphasized the significance of detailed propriety characterization include [16,17,36,55,57, 131]. Moreover, the significant contributions from the information theory group comprises of [94, 132–138]. Furthermore, few early contributions from the communication circle include [56, 65, 75, 139–141].

Various studies discussed the significance of a complete characterization and appropriate treatment of the systems involving improper and non-circular signals. These studies deal with impropriety concepts that are focused on Gaussianity deviants [142], interference mitigation [143], filtering [37], detection [144], estimation [145–147], source identification [148] and separation [39]. Furthermore, a comprehensive treatment of theoretical fundamentals of improper and non-circular signals along with their diverse applications is carried out in [31]. Despite of all these contributions, an exhaustive yet comprehensive sur-

vey is required which not only encompasses all these domains but also furnishes complete evolution from the theoretical aspects to the realization ones.

Improper Gaussian Signaling

Improper Gaussian signaling has caught significant attention in the last two decades. It was expected to outperform it's counterpart proper Gaussian signaling under improper noise/interference. Surprisingly, it has emerged as a strong competitor in many interference-limited scenarios even in the absence of any improper noise/interference. This section marks valuable contributions, which quantified the superiority of IGS over traditional PGS in both cases.

Improper Interference/Noise Links Improper/asymmetric signaling dominates traditional signaling when the system is contaminated by the improper complex interference [149–152], self-interference [78, 153, 154], improper noise [19, 75, 155, 156], asymmetric noise/distortions [83, 108, 133] and non-circular hardware imperfections [20, 68–70, 77, 82]. IGS is proven really effective in dampening improper noise and distortion effects in multi-antenna or multi-nodal system settings [83, 110, 157–159]. The ergodic rate maximization and outage probability minimization based on a generalized error model for accumulative hardware impairments in SIMO and MIMO systems are studied in [20, 108] Moreover, IGS benefits can also be reaped in various full-duplex/half-duplex relay settings by effectively compensating the residual self-interference, inter-relay interference and/or HWD [79, 83, 153, 158–160].

General Interference Links Recent studies have demonstrated the perks of improper/asymmetric signaling in the general interference-limited scenarios even in the absence of improper contamination. Notable edge attained by improper transmission over proper transmission in various interference-limited scenarios include interference broadcast channel (IBC) [161, 162], broadcast channel (BC) [163–165], multiple access channel (MAC)

[166, 167], cross-interference channel (X-IC or IC) [81, 168–174], one-sided interference channel (Z-IC) [175–178], relaying systems [158, 179–182] multi-antenna systems [183, 184], multi-cell systems [185], and multi-tier networks [186], etc. The reduced-entropy IGS can also be beneficial in unlicensed spectrum-sharing techniques with minimal interference to the legitimate users in underlay [187–191], overlay [192] and interweave [193] cognitive radio setups.

1.5.4 Adaptive Asymmetric Signaling

After realizing the infeasibile practical implementation of the celebrated IGS, some researchers started exploring the asymmetric signaling as a way to adopt improper signaling in the discrete finite space. Wireless communications community focused on Geometric shaping whereas optical fiber and free-space optics adopted probabilistic shaping to realize asymmetric signaling.

Structural/Geometric Shaping

A family of improper discrete constellations generated by widely linear processing of a square M-ary QAM depict parallelogram envelop [96]. Similarly, GS based on optimal translation and rotation also yields parallelogram envelop [98]. However, conditioned on high SNR and higher order QAM, the optimal constellation is the intersection of the hexagonal lattice/packing with an ellipse where the eccentricity determines the circularity coefficient [97]. GS has emerged as a competent player to reduce shaping loss and improve reception at lower signal-to-noise ratios in terrestrial broadcast systems [194, 195]. GS parameters can be designed for diverse objectives such as capacity maximization [96], BER reduction [98], and symbol error probability minimization [97].

Stochastic/Probabilistic Shaping

PS-based schemes have been employed to enhance the system performance in optical fiber communication (OFC) and free space optics (FSO). In OFC, multiple transformations are presented to approach Gaussian channel capacity using PS including prefix codes [196, 197], many-to-one mappings combined with a turbo code [198], distribution matching [199] and cut-and-paste method [200]. Furthermore, multidimensional coded modulation format with hybrid probabilistic and geometric constellation shaping can effectively compensate non-linearity and approach Shannon limits in OFC [201]. Coded modulation scheme with PS aims to solve the shaping gap and coarse mode granularity problems [202]. Interested reader can read the classic work [132] for the design guidelines of AS in the coherent Gaussian channel with equal signal energies and unequal a priori probabilities. Probabilistic amplitude shaping is another concept that can only be used for symmetric constellation with coherent modulation, which greatly limits its application [203]. For FSO, a practical and capacity achieving PS scheme with adaptive coding modulation is proposed with intensity modulation/direct detection [204].

The concept of PS is widely employed in the OFC and FSO systems. However, it is quite not well investigated in wireless communication systems and only a few studies have contributed in this domain [205, 206]. For example, enumerative amplitude shaping is proposed as a constellation shaping scheme for IEEE 802.11 which renders Gaussian distribution on the constituent constellation [205]. Moreover, PS has been proposed to maximize the mutual information between transmit and receive signals for non-linear distortion effects in AWGN channels [206]. To the best of authors' knowledge, PS has not been used to enhance the error performance or to realize the IGS for wireless communication systems with HWD.

1.5.5 Deductions

Numerous contributions have come forward in these broader streams to address the aforementioned limitations. However, they lack in one or multiple aspects which seeks further investigation. For instance, various studies incorporate one or the other form of HWIs but very few focus on the accumulative impact of these HWIs. The rare contributions which deal with the accumulative effects have failed to recognize the improper signatures caused by the WLT of IQI. Interestingly, the improper characterization demands meticulous characterization and generalized analysis which is certainly lacking in most of the literature. Similarly, the impact of IGS has been abundantly studied in interference management but none of the work recognized its effectiveness in mitigating the HWIs. Although IGS has emerged as a promising candidate in improving system performance in various interference-limited scenarios, it is practically unfeasible. A handful of researches signified this aspect and proposed GS as the practical implementation of IGS. Nonetheless, we are the pioneers in suggesting PS as another way to induce impropriety which is far superior to GS. We also presented HS accumulating the perks of both GS and PS to approach the theoretical limits offered by IGS. The details of our objectives and contributions to address the aforementioned deficiencies are furnished in the next section.

1.6 Objectives and Contributions

The main objective of this dissertation is to encompass the journey from the theoretical limits of IGS to practical implementation of asymmetric signaling with a special emphasis on HWI systems. This necessitates accurate modeling and characterization of the aggregate hardware impairments along with the efficient signaling design and compensation algorithms to mitigate the aggregate degrading effect of improper impairments and interference. Unlike traditional proper Gaussian signaling scheme, we adopt improper Gaussian transmission signaling scheme to efficiently combat all forms of impairments and interference.

Next, we take a step further towards practical feasibility and adopt geometric, probabilistic, and hybrid shaping. The main contributions of thesis are enumerated below:

- We consider the combined effect of various HWIs including both transmitter and receiver I/Q mismatch as well as accumulative additive distortions at the transmitter as well as the receiver [20].The aggregate model is advantageous for estimating and modeling the combined impact for joint mitigation. Amplitude and rotational errors from all impairment sources are modeled as multiplicative error whereas the distortion noises are incorporated as additive errors, which result in a noise cloud around the detected symbols.

- This research investigates full characterization of the self-interfering (SI) information signals and improper Gaussian noise components inspired by the statistical signal processing studies. Rigorous and accurate statistical model of aggregate hardware imperfections captured their asymmetric characteristics pertaining to the widely linear transformation of the transmitted and received signal under I/Q imbalance in the up- and down- conversion stage [31]. We extensively studied, meticulously modeled, and precisely quantify the improper characteristics of the aggregate HWIs from various impairment sources.

- We advocate adopting IGS signaling scheme as opposed to the traditional PGS scheme for interference-limited scenarios. Our objective is to quantify the gain obtained by IGS over PGS and to evaluate if the gain is significant enough to adopt IGS optimization framework as the optimal IGS solution can sometimes reduce to PGS. IGS benefits are analyzed in improving the achievable rate performance while efficiently mitigating the improper HWDs.

- We explore the utilization of improper signaling scheme instead of the proper signaling scheme to combat both the RSI and HWD in MH-DF-FDR systems. This work studies the effect of HWDs and RSI on the achievable rate performance of the

MH-FDR system in the absence of direct link. Then, employs IGS to compensate the degradation on the achievable rate performance due to both intrusions. We developed a rigorous joint optimization framework to design the signal characteristics by tuning the signal symmetry degree in terms of the pseudo-variance in order to maximize the end-to-end achievable rate of the MH-FDR system. We further proposed a distributed optimization framework that suits practical implementation of the proposed transmission scheme offering reduced round-trip delays, computational complexity and communication overhead.

- We analyzed the error probability performance of communication systems suffering from HWIs induced by IQI and additive distortions at both transceivers. We derive the optimal maximum likelihood (ML) detector and the suboptimal linear minimum mean square error (LMMSE) receiver for the improper self-interfering information signals [207]. We analyzed instantaneous and then average error probability based on the pairwise error probability followed by the bounds and approximations for various HWI system configurations.

- Motivated by the theoretical limits results, which demonstrate the benefits of employing improper Gaussian signaling to improve the performance of hardware impaired systems, we adopt asymmetric signaling scheme to minimize the error probability performance. Asymmetric signaling is inevitable for practical realization of improper information-bearing signals. Therefore, we proposed and designed structural/geometric shaping of finite discrete constellations such as M-QAM using the optimal rotation and translation matrices as opposed to the widely linear precoders. Such transformation enabled us to significantly reduce the average symbol error rate.

- The issues associated with GS, such as high shaping gap and coarse granularity, motivated us to adopt stochastic/probabilistic shaping to realize the IGS scheme and combat HWD to assure reliable wireless communication. Our proposed PS is based

on the optimal maximum a posteriori (MAP) detector and the design is governed by the power and rate constraints for hardware distorted system. We further suggest a hybrid shaped asymmetric signaling scheme that reaps benefits of both PS and GS and present an adaptive algorithm that tunes both signal probability and shaping parameters. Finally, we verify our proposed techniques i.e., PS, GS, and HS in terms of BER and throughput performance in both AWGN and Rayleigh fading channels.

- Our survey bridges the interdisciplinary gap between the fields of information theory and signal processing for the wireless communication applications pertaining to signal characterization. We emphasized the significance of propriety characterization and elucidated the main differences between the intermingled terms of impropriety and circularity. We carried out the taxonomy of literature to feature naturally occurring sources of impropriety and their consequences. We further elaborated the performance comparison of the theoretical limits achieved by IGS as compared to the conventional PGS e.g., achievable rate, outage probability, power/energy efficiency and DoF. We presented various design guidelines covering suitable optimization tools for the IGS design in addition to the relevant impropriety detection, estimation, filtering and separation procedures. This survey encompassed the journey from theoretical IGS to practical asymmetric discrete signaling and corresponding asymmetric signal recovery methodologies namely equalization, estimation, filtering and detection. Moreover, error probability (EP) analysis demonstrated the maximum reported percentage decrease and the corresponding SNR gains to attain a certain error rate with asymmetric characterization relative to symmetric characterization in various system configurations. Comprehensive survey of the applications, in data analysis, signal processing and communication theory domains, reaping benefits by exploiting or incorporating impropriety concepts is also included. In the end, we summarized the lesson learned throughout this study while pointing out the main challenges and way forward.

The proposed study not only encloses the existing contributions but also serves as an introductory and motivational guideline for the beginners in this domain. It further elaborates various tools and techniques for appropriate improper/asymmetric signaling to reap the maximum benefits.

1.7 Organization and Road map

The rest of the thesis is organized as follows. In Chapter 2, we provide the technical framework to understand the preliminary impropriety concepts regarding complex and quaternion random vectors. Next we present the hardware impairment aggregate model along with its statistical characteristics in Chapter 3. A case study of IGS performance in improving the end-to-end achievable rate of a MH-DF-FDR system suffering from HWD and RSI is tackled in Chapter 4. In the journey from theoretical IGS to practical AS, Chapter 5 analyzes the error probability performance and design the GS for a wireless communication link suffering from numerous HWIs. In Chapter 6, we design and examine probabilistic shaping as another way to implement AS and improve reliability and throughput of a HWD system. Additionally, we employ hybrid of PS and GS to seek the maximum benefits from the added DoF. Chapter 7 and 8 present the theoretical and practical aspects of improper signaling for communications, respectively. Chapter 9 highlights the application of impropriety characterization in data analysis, signal processing, and communication systems. Lastly, we summarize and conclude the dissertation followed by the challenges, lessons learned, and way forward in Chapter 10.

1.8 Notations

In this report, scalars are denoted by lower-case italic letters, while vectors and matrices are denoted by boldfaced lower- and upper-case letters, respectively. For a complex scalar x, the conjugate and absolute value of x are represented by x^* and $|x|$, respectively. On the other hand, for a given vector $\mathbf{x}$, the L2-norm, complex-conjugate, transpose and conjugate-

transpose of $\mathbf{x}$ are represented by $\|\mathbf{x}\|_2, \mathbf{x}^*$, $\mathbf{x}^{\mathrm{T}}$ and $\mathbf{x}^{\mathrm{H}}$, respectively. The complex augmented random vector $\underline{\mathbf{x}}$ is defined as $\underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x}^{\mathrm{T}} & \mathbf{x}^{\mathrm{H}} \end{bmatrix}^{\mathrm{T}}, \mathbf{x} \in \mathbb{C}^{\mathrm{N}}$. Identity matrix with N dimension is presented by I_N. For a square matrix $\mathbf{A}$, the complex-conjugate, transpose, conjugate-transpose, and square-root matrix of $\mathbf{A}$ are represented by $\mathbf{A}^*$, $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$, $\mathbf{A}^{-1}$ and $\mathbf{A}^{1/2}$, respectively. As for the trace and determinant of $\mathbf{A}$, $\mathrm{Tr}\left(\mathbf{A}\right)$ and $|\mathbf{A}|$ are used, respectively. Moreover, $\mathbf{A} \succeq \mathbf{0}$ and $\mathbf{A} \succ \mathbf{0}$ denote that $\mathbf{A}$ is a semi-definite positive matrix and positive definite matrix, respectively. $\mathbb{C}^{M \times N}$ and $\mathbb{R}^{M \times N}$ describe a complex-valued and real-valued matrix with dimensions $M \times N$, respectively. $\mathbb{C}_*^{2N}$ represents the set of augmented vectors $\underline{\mathbf{x}}$. The expected value operator is given by $\mathrm{E}\left[.\right]$ and the probability of occurrence of an event Ω is expressed as $\mathrm{Pr}\left\{\Omega\right\}$. Imaginary number i has it's conventional representation and is defined by $i^2 = -1$.

Chapter 2

Technical Framework

Many applications in applied sciences employ complex analysis to model the real-world data in complex domain. The complex analysis brings advantages like comprehensiveness, computational economy, extra dimension, elegant analysis and much more. For instance, multi-component alternating current circuits assume complex impedance, in place of real resistance and reactance, for refined circuit analysis. The resistance and reactance are treated as the real and imaginary components, respectively, of a complex impedance [208]. Analogously, in digital communications, complex numbers enabled us to deal with quadrature amplitude modulation (QAM) in a tractable mathematical way. Moreover, one square-QAM signal compactly carries the information of two pulse amplitude modulation (PAM) signals in its inphase and quadrature phase components [209]. The data from some physical systems should be analyzed as complex-valued signals because the data represent motion on the complex plane (e.g., tidal analysis in oceanography and two-component observations in meteorology [210]). Furthermore, directional processes (radar, sonar, Doppler ultrasound, vector fields, bearings only estimation), where both the "intensity" (amplitude) and "direction" (phase) components carry the information, are also most conveniently analyzed as complex valued processes [129, 211, 212].

But why stop at complex numbers? In fact, there are several applications which require higher dimensional representation such as quaternions, which is convenient to represent the rotations of three-dimensional space. Quaternions are used to characterize data of several systems/applications including aerospace [213], computer graphics [214], signal array processing [215], Fourier transforms of images [216], design of orthogonal polarized

STBC [217], wave separation [218], wind forecasting [219], nonlinear estimation [220], adaptive filtering [221, 222], and vector sensors [223]. One compelling application is the unified treatment of the relative position and orientation in hand-eye calibration of a robot [224].

In a nutshell, complex numbers and their extensions are immensely used in countless real-world applications. Some of these applications treat variables, signals or images as deterministic quantities. However, many applications require a stochastic modeling of the underlying phenomena such as electromagnetic waves carrying random codes, polarized magnetic disturbances, and noise in image processing etc. [225]. Therefore, complex and quaternion r.v. require appropriate and complete characterization to be fully understood and applied.

This section explains the basic technical framework to understand the interplay of proper/improper and circular/non-circular RVs at length. It begins with the introduction of various stochastic data representations and distinguishes intermingled terms of propriety and circularity. It further elaborates the appropriate transformations, operations, expressions and testing for the improper characterizations.

2.1 Stochastic Data Modeling

Various data presentation techniques are proposed in order to connect data with its modeling domain. In the remaining section, we focus on complex and quaternion forms pertaining to their vast applications [36, 129, 211–215].

2.1.1 Complex Random Vectors

Complex analysis of data can be carried out based on three data representations known as complex, real composite and complex augmented representation of the complex random vector (RV)s [31, 36]. Based on these representations, we will introduce the first-order and second-order statistics (SOS) characteristics of the RVs and their intuitive meanings.

Complex Representation

Consider an N-dimensional complex RV $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ and $\mathbf{z} \in \mathbb{C}^N$. The first order statistic given by the statistical mean) of $\mathbf{z}$ is

$$\mu_{\mathbf{z}} = \mathrm{E}\{\mathbf{z}\} = \mathrm{E}\{\mathbf{x}\} + i\mathrm{E}\{\mathbf{y}\} = \mu_{\mathbf{x}} + i\mu_{\mathbf{y}}, \tag{2.1}$$

where, $\mu_{\mathbf{x}}, \mu_{\mathbf{y}} \in \mathbb{R}^N$ and $\mu_{\mathbf{z}} \in \mathbb{C}^N$. The conventional SOS characterization of $\mathbf{z}$ is perceived by the covariance matrix, which is defined as

$$\mathbf{R}_{\mathbf{zz}} = \mathrm{E}\{(\mathbf{z} - \mu_{\mathbf{z}})(\mathbf{z} - \mu_{\mathbf{z}})^{\mathrm{H}}\} = \mathbf{R}_{\mathbf{xx}} + \mathbf{R}_{\mathbf{yy}} + i(\mathbf{R}_{\mathbf{xy}}^{\mathrm{T}} - \mathbf{R}_{\mathbf{xy}}), \tag{2.2}$$

where $\mathbf{R}_{\mathbf{xx}}$ and $\mathbf{R}_{\mathbf{yy}}$ are the auto covariance matrices of the real and imaginary components, respectively and $\mathbf{R}_{\mathbf{xy}}$ is the cross covariance matrix between them. Furthermore, $\mathbf{R}_{\mathbf{xx}}, \mathbf{R}_{\mathbf{yy}}, \mathbf{R}_{\mathbf{xy}} \in \mathbb{R}^{N\mathrm{x}N}$ and $\mathbf{R}_{\mathbf{zz}} \in \mathbb{C}^{N\mathrm{x}N}$. The covariance matrices in (2.2) are found from

$$\mathbf{R}_{\mathbf{uv}} = \mathrm{E}\{(\mathbf{u} - \mu_{\mathbf{u}})(\mathbf{v} - \mu_{\mathbf{v}})^{\mathrm{T}}\}, \tag{2.3}$$

where $\mathbf{u}$ and $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}\}$. However, the complete SOS depiction involves another matrix, $\tilde{\mathbf{R}}_{\mathbf{zz}} \in \mathbb{C}^{N\mathrm{x}N}$, named as *pseudo-covariance matrix* [57, 94].

$$\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathrm{E}\{(\mathbf{z} - \mu_{\mathbf{z}})(\mathbf{z} - \mu_{\mathbf{z}})^{\mathrm{T}}\} = \mathbf{R}_{\mathbf{xx}} - \mathbf{R}_{\mathbf{yy}} + i(\mathbf{R}_{\mathbf{xy}}^{\mathrm{T}} + \mathbf{R}_{\mathbf{xy}}). \tag{2.4}$$

Complete SOS in the sense that $\tilde{\mathbf{R}}_{\mathbf{zz}}$ accounts for correlation and unequal power distribution of the quadrature components of a complex RV in addition to covariance matrix which assumes equal power and uncorrelated real and imaginary components. For non-singular $\mathbf{R}_{\mathbf{zz}}$, the following three conditions are necessary and sufficient for $\mathbf{R}_{\mathbf{zz}}$ and $\tilde{\mathbf{R}}_{\mathbf{zz}}$ to be the covariance and pseudo-covariance matrices of $\mathbf{z}$ [16]

- $\mathbf{R}_{\mathbf{zz}}$ is Hermitian and positive semi-definite (PSD).

- $\tilde{\mathbf{R}}_{\mathbf{zz}}$ is Symmetric
- the Schur complement $\mathbf{R}_{\mathbf{zz}} - \tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-*}\tilde{\mathbf{R}}_{\mathbf{zz}}^{*}$ is PSD.

SOS are sufficient for the widely known Gaussian RVs as they can completely characterize their distribution, characteristic function, and higher order moments [16].

Definition 2.1a. *A multivariate complex Gaussian RV* $\mathbf{z}$ *can be fully characterized using complete SOS properties as* $\mathcal{CN}\left(\mu_{\mathbf{z}}, \mathbf{R}_{\mathbf{zz}}, \tilde{\mathbf{R}}_{\mathbf{zz}}\right)$ *[20].*

Consequently, a scalar Gaussian r.v. $z = x+iy$ can be fully described as $\mathcal{CN}\left(\mu_z, \sigma_z^2, \tilde{\sigma}_z^2\right)$ [57, 226], where $\mu_z = \mathrm{E}\{z\}$, $\sigma_z^2 = \mathrm{E}\{|z-\mu_z|^2\}$ and $\tilde{\sigma}_z^2 = \mathrm{E}\{(z-\mu_z)^2\}$ represent the statistical mean, variance, and pseudo-variance of z, respectively.

Definition 2.2a. *Two complex RVs* $\mathbf{z}_1$ *and* $\mathbf{z}_2 \in \mathbb{C}^N$ *are uncorrelated iff both* $\mathbf{R}_{\mathbf{z}_1\mathbf{z}_2}$ *and* $\tilde{\mathbf{R}}_{\mathbf{z}_1\mathbf{z}_2}$ *matrices vanish, where cross covariance* $\mathbf{R}_{\mathbf{z}_1\mathbf{z}_2} = \mathrm{E}\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_2 - \mu_{\mathbf{z}_2})^{\mathrm{H}}\}$ *and cross pseudo-covariance* $\tilde{\mathbf{R}}_{\mathbf{z}_1\mathbf{z}_2} = \mathrm{E}\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_2 - \mu_{\mathbf{z}_2})^{\mathrm{T}}\}$ *[94].*

Real Composite Representation

The complex RV $\mathbf{z}$ can be alternately represented as the real composite RV $\mathbf{u} = [\mathbf{x}^{\mathrm{T}}\ \mathbf{y}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2N}$. The first- and second-order statistical characteristics of this representation are described by $\mu_{\mathbf{u}} \in \mathbb{R}^{2N}$ and $\mathbf{R}_{\mathbf{uu}} \in \mathbb{R}^{2N\mathrm{x}2N}$, respectively [16, 36, 55].

$$\mu_{\mathbf{u}} = \mathrm{E}\{\mathbf{u}\} = \begin{bmatrix} \mu_{\mathbf{x}} \\ \mu_{\mathbf{y}} \end{bmatrix}, \ \mathbf{R}_{\mathbf{uu}} = \begin{bmatrix} \mathbf{R}_{\mathbf{xx}} & \mathbf{R}_{\mathbf{xy}} \\ \mathbf{R}_{\mathbf{xy}}^{\mathrm{T}} & \mathbf{R}_{\mathbf{yy}}. \end{bmatrix} \tag{2.5}$$

Definition 2.1b. *A complex Gaussian RV* $\mathbf{z}$ *with alternate representation* $\mathbf{u}$ *can be fully described as* $\mathcal{N}\left(\mu_{\mathbf{u}}, \mathbf{R}_{\mathbf{uu}}\right)$ *[166].*

Definition 2.2b. *Two complex RVs* $\mathbf{z}_1$ *and* $\mathbf{z}_2$ *with real representations* $\mathbf{u}_1 = [\mathbf{x}_1^{\mathrm{T}}\ \mathbf{y}_1^{\mathrm{T}}]^{\mathrm{T}}$ *and* $\mathbf{u}_2 = [\mathbf{x}_2^{\mathrm{T}}\ \mathbf{y}_2^{\mathrm{T}}]^{\mathrm{T}}$ *are uncorrelated iff all four cross covariance matrices* $\mathbf{R}_{\mathbf{x}_1\mathbf{x}_2}$, $\mathbf{R}_{\mathbf{x}_1\mathbf{y}_2}$, $\mathbf{R}_{\mathbf{y}_1\mathbf{x}_2}$ *and* $\mathbf{R}_{\mathbf{y}_1\mathbf{y}_2}$ *vanish [94].*

Augmented Representation

The complex RV $\mathbf{z}$ is sometimes represented as an augmented complex vector $\underline{\mathbf{z}} = [\mathbf{z}^{\mathrm{T}}\ \mathbf{z}^{\mathrm{H}}]^{\mathrm{T}} \in \mathbb{C}^{2N\mathrm{x}2N}$ for convenience. The presented representations $\mathbf{u}$ and $\underline{\mathbf{z}}$ are interchangeable as $\underline{\mathbf{z}} = \sqrt{2}\mathbf{T}\mathbf{u}$ and $\mathbf{u} = \frac{1}{\sqrt{2}}\mathbf{T}^{\mathrm{H}}\underline{\mathbf{z}}$. where $\mathbf{T}$ is unitary transformation matrix defined as [16, 36, 55]

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_N & i\mathbf{I}_N \\ \mathbf{I}_N & -i\mathbf{I}_N \end{bmatrix}. \tag{2.6}$$

Analogous to other representations, the complete first and second order characterization is given by $\mu_{\underline{\mathbf{z}}} \in \mathbb{C}^{2N}$ and the augmented covariance matrix $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}} \in \mathbb{C}^{2N\mathrm{x}2N}$, respectively, which are expressed as [16]

$$\mu_{\underline{\mathbf{z}}} = \mathrm{E}\{\underline{\mathbf{z}}\} = \begin{bmatrix} \mu_{\mathbf{x}} + i\mu_{\mathbf{y}} \\ \mu_{\mathbf{x}} - i\mu_{\mathbf{y}} \end{bmatrix},\ \mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{zz}} & \tilde{\mathbf{R}}_{\mathbf{zz}} \\ \tilde{\mathbf{R}}^*_{\mathbf{zz}} & \mathbf{R}^*_{\mathbf{zz}} \end{bmatrix}. \tag{2.7}$$

The matrix $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$, which was first used in [134], has the following features [31]

- Block pattern structure ($\begin{smallmatrix} \square & \blacksquare \\ \blacksquare^* & \square^* \end{smallmatrix}$)
- Hermitian and PSD

Owing to the special structure of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$, the matrix factorization like eigen-vlaue decomposition (EVD), singular value decomposition, or Cholesky factorization work differently to the regular matrices in the sense that all decomposed factors must follow the similar block pattern structure [55]. $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ can be connected with $\mathbf{R}_{\mathbf{uu}}$ as

$$\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}} = \mathrm{E}\{\underline{\mathbf{z}}\underline{\mathbf{z}}^{\mathrm{H}}\} = 2\mathrm{E}\{\mathbf{T}\mathbf{u}\mathbf{u}^{\mathrm{H}}\mathbf{T}^{\mathrm{H}}\} = 2\mathbf{T}\mathbf{R}_{\mathbf{uu}}\mathbf{T}^{\mathrm{H}}. \tag{2.8}$$

Definition 2.1c. *A complex Gaussian RV* $\mathbf{z}$ *can alternately be fully characterized as* $\mathcal{CN}\left(\mu_{\underline{\mathbf{z}}}, \mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}\right)$.

Definition 2.2c. *Two complex RVs* $\mathbf{z}_1$ *and* $\mathbf{z}_2$ *with complex augmented representations* $\underline{\mathbf{z}}_1$ *and* $\underline{\mathbf{z}}_2$ *are uncorrelated iff* $\mathbf{R}_{\underline{\mathbf{z}}_1\underline{\mathbf{z}}_2} = \mathrm{E}\{(\underline{\mathbf{z}}_1 - \mu_{\underline{\mathbf{z}}_1})(\underline{\mathbf{z}}_2 - \mu_{\underline{\mathbf{z}}_2})^{\mathrm{H}}\}$ *vanishes.*

2.1.2 Quaternion Random Vectors

The algebra of quaternions was invented by Sir W. R. Hamilton in 1844 [227] while multiplying triplets of real numbers. Fortunately, he failed and defined an elegant way to multiply quadruplets of numbers by giving up one familiar feature of ordinary multiplication: commutativity. Quaternions are non-commutative extension of complex numbers to hyper-complex numbers [228]. Analogous to the case of complex numbers, we present three different representations of the quaternion RV.

Complex Representation

Consider an N-dimensional quaternion RV $\mathbf{q} \in \mathbb{H}^N$, $\mathbf{q} = \mathbf{r}_1 + i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k$ where $\mathbf{r}_1, \mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k \in \mathbb{R}^N$ and the basis elements i,j and k satisfy $i^2 = j^2 = k^2 = ijk = -1$, $ij = -ji = k$, $jk = -kj = i$ and $ki = -ik = j$ (Non-commutative multiplication). The quaternion conjugate is given by $\mathbf{q}^* = \Re(\mathbf{q}) - \Im(\mathbf{q}) = \mathbf{r}_1 - i\mathbf{r}_i - j\mathbf{r}_j - k\mathbf{r}_k$ and the involution of $\mathbf{q}$ over a pure unit quaternion $\alpha \in \{i, j, k\}$ is defined as $\mathbf{q}^{(\alpha)} = -\alpha\mathbf{q}\alpha$ [138, 229]. For instance, the involution $\mathbf{q}^{(i)} = \mathbf{r}_1 + i\mathbf{r}_i - j\mathbf{r}_j - k\mathbf{r}_k$ inverts the sign of $\mathbf{r}_i$ in $\mathbf{q}^*$. Based on these descriptions, the complete SOS characterization of a zero-mean $\mathbf{q}$ requires the covariance matrix $\mathbf{R}_{\mathbf{qq}} = \mathrm{E}\{\mathbf{q}\mathbf{q}^{\mathrm{H}}\}$ and three pseudo-covariance matrices $\mathbf{R}_{\mathbf{qq}^{(i)}}, \mathbf{R}_{\mathbf{qq}^{(j)}}$ and $\mathbf{R}_{\mathbf{qq}^{(k)}}$. These pseudo-covariance matrices quantify the correlation between $\mathbf{q}$ and its involutions $\mathbf{q}^{(\alpha)}$ over three pure unit quaternions [230]

$$\mathbf{R}_{\mathbf{qq}^{(\alpha)}} = \mathrm{E}\{\mathbf{q}\mathbf{q}^{(\alpha)\mathrm{H}}\} = -\mathrm{E}\{\mathbf{q}\alpha\mathbf{q}^{\mathrm{H}}\alpha\}. \tag{2.9}$$

Definition 2.3a. *A zero-mean quaternion Gaussian RV* $\mathbf{q}$ *is completely characterized using SOS of* $\mathbf{q}$ *and its involutions* $\mathbf{q}^{(i)}, \mathbf{q}^{(j)}$ *and* $\mathbf{q}^{(k)}$ *as* $\mathcal{QN}\left(\mathbf{R}_{\mathbf{qq}}, \mathbf{R}_{\mathbf{qq}^{(i)}}, \mathbf{R}_{\mathbf{qq}^{(j)}}, \mathbf{R}_{\mathbf{qq}^{(k)}}\right)$.

Real Composite Representation

The zero-mean quaternion RV $\mathbf{q}$ can be alternately represented as the quadrivariate real composite RV $\mathbf{v} \in \mathbb{R}^{4N}$, $\mathbf{v} = [\mathbf{r}_1^{\mathrm{T}}\ \mathbf{r}_i^{\mathrm{T}}\ \mathbf{r}_j^{\mathrm{T}}\ \mathbf{r}_k^{\mathrm{T}}]^{\mathrm{T}}$. Similar to complex case, the SOS properties of $\mathbf{v}$ are given by the $\mathbb{R}^{4N\mathrm{x}4N}$ covariance matrix [230]

$$\mathbf{R}_{\mathbf{vv}} = \begin{bmatrix} \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_k} \\ \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} \\ \mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_k} \\ \mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_k} \end{bmatrix}, \tag{2.10}$$

where, $\mathbf{R}_{\mathbf{r}_\gamma\mathbf{r}_\zeta} = \mathrm{E}\{\mathbf{r}_\gamma\mathbf{r}_\zeta^{\mathrm{T}}\}$ and $\mathbf{R}_{\mathbf{r}_\gamma\mathbf{r}_\zeta} = \mathbf{R}_{\mathbf{r}_\zeta\mathbf{r}_\gamma}^{\mathrm{T}}$ with $\gamma, \zeta \in \{1, i, j, k\}$.

Definition 2.3b. *A zero-mean quaternion Gaussian RV* $\mathbf{q}$ *with quadrivariate real composite representation* $\mathbf{v}$ *is completely characterized using the symmetric correlation matrix* $\mathbf{R}_{\mathbf{vv}}$.

Augmented Representation

Let $\underline{\mathbf{q}}$ denotes the augmented representation of $\mathbf{q}$ and its involutions as

$$\underline{\mathbf{q}} = \begin{bmatrix} \mathbf{q} \\ \mathbf{q}^{(i)} \\ \mathbf{q}^{(j)} \\ \mathbf{q}^{(k)} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 + i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k \\ \mathbf{r}_1 + i\mathbf{r}_i - j\mathbf{r}_j - k\mathbf{r}_k \\ \mathbf{r}_1 - i\mathbf{r}_i + j\mathbf{r}_j - k\mathbf{r}_k \\ \mathbf{r}_1 - i\mathbf{r}_i - j\mathbf{r}_j + k\mathbf{r}_k \end{bmatrix}. \tag{2.11}$$

Analogous to complex case, the alternate representations of $\mathbf{q}$ are interchangeable using $\underline{\mathbf{q}} = 2\mathbf{A}_N\mathbf{v}$ and $\mathbf{v} = \frac{1}{2}\mathbf{A}_N^{\mathrm{H}}\underline{\mathbf{q}}$, where $\mathbf{A}_N$ is a unitary transformation matrix defined as [231]

$$\mathbf{A}_N = \frac{1}{2}\begin{bmatrix} \mathbf{I}_N & i\mathbf{I}_N & j\mathbf{I}_N & k\mathbf{I}_N \\ \mathbf{I}_N & i\mathbf{I}_N & -j\mathbf{I}_N & -k\mathbf{I}_N \\ \mathbf{I}_N & -i\mathbf{I}_N & j\mathbf{I}_N & -k\mathbf{I}_N \\ \mathbf{I}_N & -i\mathbf{I}_N & -j\mathbf{I}_N & k\mathbf{I}_N \end{bmatrix}. \tag{2.12}$$

The SOS characterization of $\underline{\mathbf{q}}$ is given by the following augmented covariance matrix [232]

$$\mathbf{R}_{\underline{\mathbf{q}}\underline{\mathbf{q}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{qq}} & \mathbf{R}_{\mathbf{qq}^{(i)}} & \mathbf{R}_{\mathbf{qq}^{(j)}} & \mathbf{R}_{\mathbf{qq}^{(k)}} \\ \mathbf{R}^{(i)}_{\mathbf{qq}^{(i)}} & \mathbf{R}^{(i)}_{\mathbf{qq}} & \mathbf{R}^{(i)}_{\mathbf{qq}^{(k)}} & \mathbf{R}^{(i)}_{\mathbf{qq}^{(j)}} \\ \mathbf{R}^{(j)}_{\mathbf{qq}^{(j)}} & \mathbf{R}^{(j)}_{\mathbf{qq}^{(k)}} & \mathbf{R}^{(j)}_{\mathbf{qq}} & \mathbf{R}^{(j)}_{\mathbf{qq}^{(i)}} \\ \mathbf{R}^{(k)}_{\mathbf{qq}^{(k)}} & \mathbf{R}^{(k)}_{\mathbf{qq}^{(j)}} & \mathbf{R}^{(k)}_{\mathbf{qq}^{(i)}} & \mathbf{R}^{(k)}_{\mathbf{qq}} \end{bmatrix}, \tag{2.13}$$

where $\mathbf{R}^{(\beta)}_{\mathbf{qq}^{(\alpha)}} = -\beta\mathbf{R}_{\mathbf{qq}^{(\alpha)}}\beta$ with $\beta \in \{i, j, k\}$. The correlation matrix $\mathbf{R}^{(\alpha)}_{\mathbf{qq}^{(\alpha)}} = \mathbf{R}^{(\alpha)\mathrm{H}}_{\mathbf{qq}^{(\alpha)}}$ is i-Hermitian, j-Hermitian and k-Hermitian for $\alpha = i, j$ and k, respectively. Thus, proving a non-trivial extension of augmented complex statistics to its quaternion counterpart [231]. $\mathbf{R}_{\underline{\mathbf{q}}\underline{\mathbf{q}}}$ is linked with $\mathbf{R}_{\mathbf{vv}}$ as

$$\mathbf{R}_{\underline{\mathbf{q}}\underline{\mathbf{q}}} = \mathrm{E}\{\underline{\mathbf{q}}\underline{\mathbf{q}}^{\mathrm{H}}\} = 4\mathrm{E}\{\mathbf{A}_N\mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{A}_N^{\mathrm{H}}\} = 4\mathbf{A}_N\mathbf{R}_{\mathbf{vv}}\mathbf{A}_N^{\mathrm{H}}. \tag{2.14}$$

Definition 2.3c. *A zero-mean quaternion Gaussian RV* $\mathbf{q}$ *with augmented representation* $\underline{\mathbf{q}}$ *is completely characterized using the augmented correlation matrix* $\mathbf{R}_{\underline{\mathbf{q}}\underline{\mathbf{q}}}$.

The uncorrelation between two quaternions $\mathbf{q}_1$ and $\mathbf{q}_2$ with augmented representations $\underline{\mathbf{q}}_1$ and $\underline{\mathbf{q}}_2$, respectively, require the regular cross covariance $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2} = \mathrm{E}\{\mathbf{q}_1\mathbf{q}_2^{\mathrm{H}}\}$ and three involutional cross covariance matrices $\mathbf{R}^{(\alpha)}_{\mathbf{q}_1\mathbf{q}_2} = -\alpha\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2}\alpha$ along with all possible combinations of cross pseudo-covariance matrices $\mathrm{E}\{\mathbf{q}_1^{(\gamma)}\mathbf{q}_2{}^{(\zeta)\mathrm{H}}\}$ contained in $\mathbf{R}_{\underline{\mathbf{q}}_1\underline{\mathbf{q}}_2} = \mathrm{E}\{\underline{\mathbf{q}}_1\underline{\mathbf{q}}_2^{\mathrm{H}}\}$ to be zero.

2.1.3 Summary and Insights

In a nutshell, this subsection covers three well-known data presentation techniques i.e., complex, real composite, and complex augmented representation for both complex and quaternion RVs. However, why do we need multiple equivalent representations of the same phenomenon? To address this concern, consider the tedious analysis of real composite representation which can be significantly simplified using complex representation [16]. Besides this intuitive reasoning, there are other limitations that will be discussed in next subsection. Another curiosity, that arises, is why do we need redundant complex augmented representation of the RVs. The response to this query is three-folds. 1) The uncorrelation of real $\mathbf{x}_1$ and $\mathbf{x}_2$ is completely defined as $\mathbf{R}_{\mathbf{x}_1\mathbf{x}_2} = \mathbf{0}$. However, the uncorrelation of the complex $\mathbf{z}_1$ and $\mathbf{z}_2$ is not completely defined as $\mathbf{R}_{\mathbf{z}_1\mathbf{z}_2} = \mathbf{0}$. Thus, we require augmented $\underline{\mathbf{z}}_1$ and $\underline{\mathbf{z}}_2$ to completely characterize the uncorrelation as $\mathbf{R}_{\underline{\mathbf{z}}_1\underline{\mathbf{z}}_2} = \mathbf{0}$ [31]. 2) Pertaining to the block structure of the $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$, interestingly it can be invertible even when $\mathbf{R}_{\mathbf{zz}}$ and $\tilde{\mathbf{R}}_{\mathbf{zz}}$ are not [55]. 3) It is a powerful tool for WL transformations (refer to Section 2.3) [36]. On the other hand, vector representations i.e., real composite and augmented formulation allow easier geometrical interpretations in high-dimensional space relative to the complex representation [225]. In conclusion, we prefer complex representation for comprehensive analysis, real composite representation for easy geometrical interpretations and augmented formulation for complete modeling, characterization, and operations (e.g., inversion and transformation) [231].

2.2 Propriety versus Circularity

Properness evaluation i.e., the identification of correlation between the complex valued RVs and their complex conjugates is a popular subject in signal processing [17, 55, 57] and information theory [94]. It has also been extended to quaternion valued r.v. [225] and vectors [138, 231] for various applications. This subsection illustrates extensive definitions

of impropriety and the measures for degree of impropriety (DoI) for complex as well as quaternion RVs based on different representations.

2.2.1 Complex Random Vectors

The full statistical characterization of complex RVs involves the analysis of the moments and probability distributions (if exist). A complex RV is designated as proper/improper and circular/non-circular based on these characteristics.

Propriety

Based on the real and complex representation of complex RVs, a *proper complex* RV is described using the following definition

Definition 2.4 (Proper RV)**.** *A proper complex RV is composed of real and imaginary vectors with identical auto covariance matrices* $\mathbf{R_{xx}} = \mathbf{R_{yy}}$ *and skew-symmetric cross covariance matrix* $\mathbf{R_{xy}} = -\mathbf{R_{xy}}^{\mathrm{T}}$*. Alternately, a proper complex RV* $\mathbf{z}$ *renders zero pseudo-covariance matrix* $\tilde{\mathbf{R}}_{\mathbf{zz}}$ *[94]. Properness may also be seen as the uncorrelation of* z *with its complex conjugate* z^* *[95, 233]. These equivalent propriety definitions are sometimes referred as Strict Propriety [55].*

This implies that a proper RV $\mathbf{z}$ has $\mathbf{R_{xy}}$ with zero main diagonal elements rendering uncorrelated real and imaginary components of each element z_n in $\mathbf{z}$. However, the off-diagonal elements can be non-zero yielding correlated $\Re\{z_k\}$ and $\Im\{z_l\}$ for $k \neq l$ [94]. For a zero-mean scalar complex r.v. $z = x + iy$, $\sigma_{xy} = \mathrm{E}\{xy\} = 0$ is necessary for propriety.

With slight abuse of terminology, we term $\underline{\mathbf{z}}$ proper if $\mathbf{z}$ is proper where proper $\mathbf{z}$ implies vanishing pseudo-variance $\tilde{\mathbf{R}}_{\mathbf{zz}}$. However, proper $\underline{\mathbf{z}}$ does not imply vanishing $\tilde{\mathbf{R}}_{\underline{\mathbf{zz}}}$ rather it implies a block diagonal structure of $\mathbf{R}_{\underline{\mathbf{zz}}} = \textbf{Block-Diag}\,(\mathbf{R}_{\mathbf{zz}}, \mathbf{R}^*_{\mathbf{zz}})$. This demonstrates the equivalence of the propriety statements for $\mathbf{z}$ and $\underline{\mathbf{z}}$ yet different implications. It is evident that for proper $\mathbf{z}$, all eigenvalues of $\mathbf{R}_{\underline{\mathbf{zz}}}$ are real (same as that of Hermitian $\mathbf{R}_{\mathbf{zz}}$) and

have even multiplicity (paired because of diagonal blocks $\mathbf{R}_{\mathbf{zz}}$ and $\mathbf{R}^*_{\mathbf{zz}}$). Thus, strict propriety implies even multiplicity of the eigenvalues, however, it is not a sufficient condition. Therefore, Schreier *et al.* established the following generalization based on the augmented representation.

Definition 2.5 (Generalized Propriety)**.** *A complex RV is termed as generalized proper if all eigenvalues of* $\mathbf{R}_{\underline{\mathbf{zz}}}$ *have even multiplicity [55].*

The EVD of $\mathbf{R}_{\underline{\mathbf{zz}}}$ takes on the form $\mathbf{R}_{\underline{\mathbf{zz}}} = \mathbf{V}\left(\mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{\mathrm{H}}\right)\mathbf{V}^{\mathrm{H}}$ with diagonal $\mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{\mathrm{H}}$ iff $\mathbf{z}$ is generalized proper, otherwise block matrix $\mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{\mathrm{H}}$ with diagonal blocks [55]. Moreover, the complex RVs $\mathbf{z}_1$ and $\mathbf{z}_2$ are *cross proper* iff the cross pseudo-covariance matrix $\tilde{\mathbf{R}}_{\mathbf{z_1 z_2}}$ vanishes. Notably, they are *jointly proper* if the composite RV having $\mathbf{z}_1$ and $\mathbf{z}_2$ as subvectors is proper [94] or if they are proper and cross proper [138].

Definition 2.6 (Improper RV)**.** *A complex RV* $\mathbf{z}=\mathbf{x}+i\mathbf{y}$ *is called improper if any of the following statements holds [234]*

- *Non identically distributed* $\mathbf{x}$ *and* $\mathbf{y}$ *i.e.,* $\mathbf{R}_{\mathbf{xx}} \neq \mathbf{R}_{\mathbf{yy}}$
- *Correlated* $\mathbf{x}$ *and* $\mathbf{y}$ *i.e.,* $\mathbf{R}_{\mathbf{xy}} \neq -\mathbf{R}_{\mathbf{xy}}{}^{\mathrm{T}}$
- *Non-zero* $\tilde{\mathbf{R}}_{\mathbf{zz}}$
- *Correlated* $\mathbf{z}$ *and* $\mathbf{z}^*$
- *Lack of block diagonal structure in* $\mathbf{R}_{\underline{\mathbf{zz}}}$
- *Lack of even multiplicity of eigenvalues in* $\mathbf{R}_{\underline{\mathbf{zz}}}$

Absence of properness of a RV is termed as *improperness* and the extent of improperness is argued by the eigenvalue spread of $\mathbf{R}_{\underline{\mathbf{zz}}}$ [234]. For instance, a RV $\mathbf{z}_1$ is *less improper* than another RV $\mathbf{z}_2$ if the eigenvalues of $\mathbf{R}_{\underline{\mathbf{z}}_1\underline{\mathbf{z}}_1}$ are majorized by (less spread out) those of $\mathbf{R}_{\underline{\mathbf{z}}_2\underline{\mathbf{z}}_2}$ [63].

Figure 2.1: Extents of Circularity and their Implications

Definition 2.7 (Maximally Improper RV)**.** *For a given* $\mathbf{R}_{\mathbf{zz}}$*, the vector whose* $\mathbf{R}_{\underline{\mathbf{zz}}}$ *has least eigenvalue spread must be proper* $\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$ *whereas the vector whose* $\mathbf{R}_{\underline{\mathbf{zz}}}$ *has maximum possible eigenvalue spread must be maximally improper [31].*

Circularity

A stronger version of propriety considering the probability distribution of a RV is named *Circularity* and is defined as

Definition 2.8 (Circular RV)**.** *A complex RV* $\mathbf{z}$ *is said to be circular (having CSC distribution about the origin), iff its distribution remains invariant under multiplication by any (complex) number on the unit complex circle i.e.,* $\mathbf{z}$ *and* $\hat{\mathbf{z}} = \mathbf{z}e^{j\alpha}$ *have the same distribution for any real* α *[36, 95].*

For a scalar z, is equivalent to spherical symmetry of the corresponding real composite vector $u = [x\ y]^{\mathrm{T}}$ [235]. On the other hand, the magnitude of circularity of a complex RV $\mathbf{z}$ is further classified based on the behavior of the underlying random variables [17, 47]:

- *Marginal Circular:* If its components are complex circular r.v.
- *Weakly Circular:* If $\mathbf{z}$ and $e^{j\alpha}\mathbf{z}$ have the same probability distribution for any real α

- *Strongly Circular:* If $\mathbf{z}$ and $e^{j\mathbf{a}} \odot \mathbf{z}$ have the same probability distribution for any real vector $\mathbf{a}$

- *Total Circular:* If its components are independent and circular. The circularity characteristics are related to each other as shown in Figure 2.1.

Relation between Propriety and Circularity

Often the terms proper and circular are used interchangeably. However, as a matter of fact, they are quite related yet distinct phenomenon.

Corollary 1. *Propriety and circularity are related as:*
Circularity $\underset{\not\Leftarrow}{\Rightarrow}$ *Propriety, Impropriety* $\underset{\not\Leftarrow}{\Rightarrow}$ *Non-circularity*

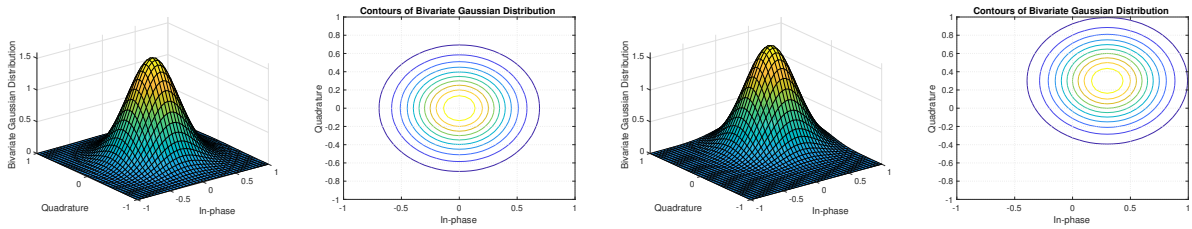
- For a zero-mean circular RV $\mathbf{z}$, the respective pseudo-covariance matrices $\tilde{\mathbf{R}}_{\mathbf{zz}}$ and $\tilde{\mathbf{R}}_{\hat{\mathbf{z}}\hat{\mathbf{z}}}$ are related as:

$$\tilde{\mathbf{R}}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} = \mathrm{E}\left[\hat{\mathbf{z}}\hat{\mathbf{z}}^{\mathrm{T}}\right] = e^{j2\alpha}\tilde{\mathbf{R}}_{\mathbf{zz}} \quad \forall\alpha. \tag{2.15}$$

  Thus, (2.15) implies $\tilde{\mathbf{R}}_{\mathbf{zz}}$ should be a zero-matrix for any given α to satisfy circularly symmetry. Thus, circularity implies properness whereas the converse is not true in general.

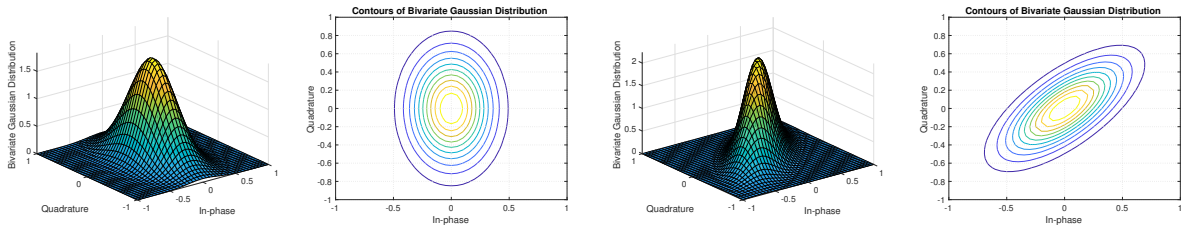
- Propriety requires the second-order moments to be rotationally invariant. However, circularity requires that the probability density function (PDF) and thus all existing moments to be rotationally invariant. Thus, circularity implies zero mean and propriety, but not vice versa [36].

- In the light of the aforementioned arguments, an improper RV with rotationally variant second moment is essentially non-circular (NC). However, a NC RV with non-zero mean but zero pseudo-covariance is still proper, proving the fact that non-circularity does not imply improperness in general.



(a) Proper and Circular Gaussian RV (Centered)

(b) Proper and NC Gaussian RV (Non-Centered)



(c) Improper Gaussian RV (Unequal Power)

(d) Improper Gaussian RV (Correlated)

Figure 2.2: The Interplay between Propriety and Circularity

Example. Propriety and circularity are equivalent for a zero-mean Gaussian r.v. having uncorrelated real and imaginary components with equal power distribution as depicted in Figure 2.2a, whereas the same distribution with non-zero mean is proper but not circular in Figure 2.2b. Furthermore, the deviation from equal power distribution and uncorrelation result in improper as well as NC complex r.v. as shown in Figure 2.2c and Figure 2.2d, respectively. The quadrature component has more power/variance than the in-phase component in Figure 2.2c and non-perpendicular distribution contours to x- or y-axis depict the correlation between I/Q components in Figure 2.2d.

For other zero-mean RVs propriety and circularity are related as shown in Figure 2.3. The rotational invariance of all existing moments of circular RV certainly implies the rotational invariance of second-order moments and thus it is equivalent to designating it as a proper RV. Second-order circular or strictly proper RV demonstrates even multiplicity of eigenvalues owing to the block-diagonal structure of $\mathbf{R}_{\underline{\mathbf{zz}}} = \textbf{Block-Diag}\left(\mathbf{R}_{\mathbf{zz}}, \mathbf{R}^{*}_{\mathbf{zz}}\right)$ rendering generalized proper RV. Consequently. a circular RV is always proper/generalized proper. On the other hand, lack of any propriety condition as elaborated in Definition 2.6

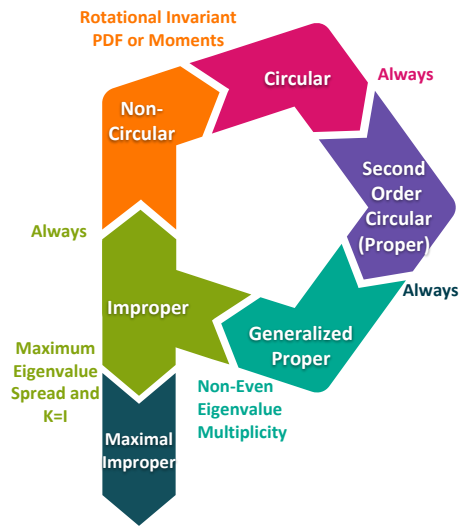


Figure 2.3: Extents of Propriety and their Implications

leads to improper RV. An improper RV with maximum eigenvalue spread of $\mathbf{R}_{\underline{\mathbf{zz}}}$ is termed as maximal improper RV. Moreover, improperness always implies non-circularly as it nullifies the condition of rotational invariant second-moment.

Degree of Improperness

Impropriety implies non-circularity, thus the rotational variance of the distribution of a complex entity is characterized by the DoI. For a complex scalar r.v. z, with finite variance σ_z^2 and pseudo-variance $\tilde{\sigma}_z^2$, the measure of correlation between z and z^* is given by,

Definition 2.9. *Circularity quotient ρ_z is defined as the fraction between the pseudo-variance and the variance*

$$\rho_z = \frac{\tilde{\sigma}_z^2}{\sigma_z^2} = ke^{i\phi}, \tag{2.16}$$

where, DoI is measured by the circularity coefficient $k = |\tilde{\sigma}_z^2|/\sigma_z^2$ and the circularity angle is given by ϕ [57].

The term *circularity coefficient* for k is originated from [135] while the terms *non-circularity rate* and *non-circularity phase* are also used for k and ϕ, respectively [236]. In reality, the circularity coefficient is the *canonical correlation* between z and z^* [237].

The geometric interpretation for the circularity quotient is given as $\rho_z \in \Omega$, where $\Omega = \{z \in \mathbb{C} : |z| \leq 1\}$ is the unit circle [57]. Thus, the circularity coefficient lies in the range $k \in [0, 1]$ providing an interesting result $0 \leq |\tilde{\sigma}_z^2| \leq \tilde{\sigma}_z^2$. The circularity coefficient measures the degree of circularity, as in $k = 0$ means z is second-order circular whereas $k = 1$ means z is maximally NC (i.e., x or y is constant, equal to zero, or x is a linear function of y) [238]. This quantifies DoI for a r.v., next we present DoI of a RV as:

Definition 2.10. *For a complex RV* $\mathbf{z}$*, the coherence matrix* $\mathbf{C} = \mathbf{R}_{\mathbf{zz}}^{-1/2}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-\mathrm{T}/2}$ *with Takagi factorization* $\mathbf{C} = \mathbf{FKF}^{\mathrm{T}}$ *contains the canonical correlations or circularity coefficients* k_i *in* $\mathbf{K} = \boldsymbol{Diag}(k_1, k_2, \ldots, k_N)$ *and* $\mathbf{F}$ *is a complex unitary matrix. It is important to highlight that* k_i *is the circularity coefficient of the* i^{th} *variable in* $\mathbf{z}$*. The DoI can be defined as a function of canonical correlations in following ways [234]*

$$\rho_1 = 1 - \prod_{i=1}^{r} \left(1 - k_i^2\right) \overset{(r=N)}{=} 1 - |\mathbf{R}_{\underline{\mathbf{zz}}}| \, |\mathbf{R}_{\mathbf{zz}}|^{-2} \,, \tag{2.17}$$

$$\rho_2 = \prod_{i=1}^{r} k_i^2 \overset{(r=N)}{=} \left|\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-*}\tilde{\mathbf{R}}_{\mathbf{zz}}^{*}\right| |\mathbf{R}_{\mathbf{zz}}|^{-1} \,, \tag{2.18}$$

$$\rho_3 = \frac{1}{N}\sum_{i=1}^{r} k_i^2 \overset{(r=N)}{=} \frac{1}{N}\mathrm{Tr}\left(\mathbf{R}_{\mathbf{zz}}^{-1}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-*}\tilde{\mathbf{R}}_{\mathbf{zz}}^{*}\right) , \tag{2.19}$$

where $r = \mathrm{rank}(\mathbf{R}_{\underline{\mathbf{zz}}})$. The set of canonical correlations $\{k_i\}_{i=1}^{N}$ is also referred to as *circularity spectrum* [135]. DoI must satisfy $0 \leq \rho_i \leq 1$, ranging from proper signal ($\rho_i = 0$) to maximally improper signal ($\rho_i = 1$) [57]. The intuitive meaning of the aforementioned measures of improperness is summarized as:

- ρ_1 is 1 if at least one $k_i = 1$ i.e., any one maximal improper variable in RV $\mathbf{z}$ will result in maximum DoI. Additionally, it is 0 if all k_i are 0. Thus, it helps in identifying if any one element of the RV is maximally improper or if all entries in a RV are proper. However, it fails to discriminate if all or any subset of the entries are maximally improper. Similarly, it is insufficient to identify any subset of proper entries. Nevertheless, ρ_1 gives the entropy loss due to improperness of a RV [36] (refer to

Section 2.4) and it is also used for impropriety likelihood-ratio testing [237] (refer to Section 2.5) as it measures the linear dependence between $\mathbf{z}$ and $\mathbf{z}^*$.

- ρ_2 is 0 if at least one $k_i = 0$ i.e., any one proper variable in RV $\mathbf{z}$ will result in minimum DoI. Moreover, ρ_2 is 1 if all k_i are 1. Thus, it helps in identifying if any one element of the RV is proper or if all entries in a RV are improper. However, it fails to apprehend if all or any subset of the entries are proper. Likewise, it cannot assess any subset of maximally improper entries.

- ρ_3 attains maximum value when all entries in $\mathbf{z}$ are maximally improper and attains minimum value when all entries in $\mathbf{z}$ are proper. Nonetheless, it cannot identify subset of proper or maximally improper entries in a RV.

In short, ρ_1 and ρ_2 complement each other and provide the missing information, whereas ρ_3 is the preferred choice for joint assessment. Intuitively, $\mathbf{K} = \mathbf{I}$ portrays all elements of the RV $\mathbf{z}$ to be maximal improper. Nevertheless, it is a necessary but not sufficient condition for maximal improperness of a RV On the other hand, maximum eigenvalue spread is the sufficient condition for such maximal impropriety.

Discussion

The concepts of propriety and circularity and the absence of these phenomena i.e., impropriety and non-circularity along with the extent of improperness (namely DoI) are discussed at length. Numerous degrees of propriety and circularity from strictly proper to generalized proper and from marginally circular to total circular, respectively, are also distinctly stated. Eventually, the interplay between propriety and circularity can be concluded as: circularity is a subset of propriety with more restrictions, whereas impropriety is a subset of non-circularity.

2.2.2 Quaternion Random Vectors

Impropriety characterization is not limited to the complex domain. For instance, Vakhania studied the concept of properness in quaternion domain [239]; however, his definition of $\mathcal{Q}$-properness is restricted to the invariance of the PDF under some specific rotations around angle of $\pi/2$. Amblard *et al.* further relaxed the conditions of $\mathcal{Q}$-properness to $\mathcal{C}^{\alpha}$-properness with an arbitrary axis and angle of rotation φ, $\mathbf{q} \triangleq e^{\alpha\varphi}\mathbf{q}$ for any pure unit quaternion α [225]. The evolved and refined definitions of quaternions propriety rely on the vanishing properties of the pseudo-covariance matrices as:

Definition 2.11 ($\mathcal{R}^{\alpha}$-properness)**.** *A quaternion RV* $\mathbf{q}$ *is* $\mathcal{R}^{\alpha}$*-proper iff the pseudo-covariance matrix* $\mathbf{R}_{\mathbf{q}\mathbf{q}^{(\alpha)}}$ *vanishes [138].*

Definition 2.12a ($\mathcal{C}^{\alpha}$-properness)**.** *A quaternion RV* $\mathbf{q}$ *is* $\mathcal{C}^{\alpha}$*-proper if it is correlated with* $\mathbf{q}^{(\alpha)}$ *and uncorrelated with the rest i.e.,* $\mathbf{q}^{(\bar{\alpha})}$ *[231].*

Example A quaternion RV $\mathbf{q}$ is $\mathcal{C}^{j}$-proper iff the pseudo-covariance matrices $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}}$ and $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(k)}}$ vanish [138]. Equivalently, $\mathcal{C}^{j}$-proper quaternion exhibits the following $\mathbf{R}_{\mathbf{v}\mathbf{v}}$ structure [230]

$$\mathbf{R}_{\mathbf{v}\mathbf{v}} = \begin{bmatrix} \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i} \end{bmatrix}, \tag{2.20}$$

Proof. Evidently, (2.20) is obtained from (2.10) by exploiting the definitions of vanishing $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}}$ and $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(k)}}$, which are expressed as

$$\begin{aligned} \mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}} &= -\mathrm{E}\{(\mathbf{r}_1+i\mathbf{r}_i+j\mathbf{r}_j+k\mathbf{r}_k)(\mathbf{r}_1^{\mathrm{T}}-i\mathbf{r}_i^{\mathrm{T}}+j\mathbf{r}_j^{\mathrm{T}}+k\mathbf{r}_k^{\mathrm{T}})\}. \\ \mathbf{R}_{\mathbf{q},\mathbf{q}^{(k)}} &= -\mathrm{E}\{(\mathbf{r}_1+i\mathbf{r}_i+j\mathbf{r}_j+k\mathbf{r}_k)(\mathbf{r}_1^{\mathrm{T}}+i\mathbf{r}_i^{\mathrm{T}}+j\mathbf{r}_j^{\mathrm{T}}-k\mathbf{r}_k^{\mathrm{T}})\}. \end{aligned} \tag{2.21}$$

Equating these pseudo-variances to zero render the set of four real-valued equations as

Table 2.1: Consequences of Vanishing Pseudo-Variances

$\mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}} = \mathbf{0}$	$\mathbf{R}_{\mathbf{q},\mathbf{q}^{(j)}} = \mathbf{0}$	$\mathbf{R}_{\mathbf{q},\mathbf{q}^{(k)}} = \mathbf{0}$
$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_1}+\mathbf{R}_{\mathbf{r}_i\mathbf{r}_i}=\mathbf{R}_{\mathbf{r}_j\mathbf{r}_j}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_k}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_1}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_j}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_i}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_k}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_1}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_k}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_i}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_j}$
$\mathbf{R}_{\mathbf{r}_i\mathbf{r}_1}-\mathbf{R}_{\mathbf{r}_1\mathbf{r}_i}=\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j}-\mathbf{R}_{\mathbf{r}_j\mathbf{r}_k}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_i}+\mathbf{R}_{\mathbf{r}_i\mathbf{r}_1}=-\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j}-\mathbf{R}_{\mathbf{r}_j\mathbf{r}_k}$	$\mathbf{R}_{\mathbf{r}_i\mathbf{r}_1}+\mathbf{R}_{\mathbf{r}_1\mathbf{r}_i}=\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_k}$
$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_j}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i}$	$\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1}-\mathbf{R}_{\mathbf{r}_1\mathbf{r}_j}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k}-\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_j}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1}=-\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k}-\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i}$
$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_k}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1}=-\mathbf{R}_{\mathbf{r}_i\mathbf{r}_j}-\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_k}+\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_j}+\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i}$	$\mathbf{R}_{\mathbf{r}_1\mathbf{r}_k}-\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1}=\mathbf{R}_{\mathbf{r}_i\mathbf{r}_j}-\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i}$

shown in Table 2.1. In case of $\mathcal{C}^j$-proper quaternion, the four set of equations each from $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}}$ and $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(k)}}$ are simultaneously solved. The solution implies $\mathbf{R}_{\mathbf{r}_j\mathbf{r}_j} = \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1}$, $\mathbf{R}_{\mathbf{r}_k\mathbf{r}_k} = \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i}$, $\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} = \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1}$, $\mathbf{R}_{\mathbf{r}_j\mathbf{r}_k} = \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i}$, $\mathbf{R}_{\mathbf{r}_1\mathbf{r}_j} = -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1}$, $\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} = -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i}$, $\mathbf{R}_{\mathbf{r}_1\mathbf{r}_k} = -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i}$, and $\mathbf{R}_{\mathbf{r}_i\mathbf{r}_j} = -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1}$. Consequently, substituting these relations in (2.10) yields $\mathbf{R}_{\mathbf{vv}}$ in (2.20). □

Definition 2.12b. *Alternately, a $\mathcal{C}^\alpha$-proper quaternion is defined to exhibit a distribution that is invariant by left Clifford translation, i.e., $\mathbf{q} \triangleq e^{\alpha\varphi}\mathbf{q}$ $\forall\varphi$, for one and only one imaginary unit α [225].*

Definition 2.13 ($\mathcal{Q}$-properness)**.** *A quaternion RV $\mathbf{q}$ is $\mathcal{Q}$-proper iff all three pseudo-covariance matrices $\mathbf{R}_{\mathbf{qq}^{(i)}}$, $\mathbf{R}_{\mathbf{qq}^{(j)}}$ and $\mathbf{R}_{\mathbf{qq}^{(k)}}$ vanish [138]. It also implies that $\mathbf{q}$ is uncorrelated with its three vector involutions. Moreover, the corresponding $\mathbf{R}_{\underline{\mathbf{qq}}}$ is real-valued, positive definite, and symmetric [231].*

$\mathcal{Q}$-properness is also referred as $\mathcal{H}$-properness and is equivalently reported as the distribution invariance of axis α and φ, i.e., $\mathbf{q} \triangleq e^{\alpha\varphi}\mathbf{q}$ $\forall\varphi$, for any imaginary unit α [225]. Analogous to the $\mathcal{C}^j$-proper case, the equivalent $\mathbf{R}_{\mathbf{vv}}$ for $\mathcal{Q}$-proper quaternion is obtained by simultaneously solving the 12 set of equations (as given in Table 2.1) obtained by setting

$\mathbf{R}_{\mathbf{q}\mathbf{q}^{(i)}}$, $\mathbf{R}_{\mathbf{q}\mathbf{q}^{(j)}}$ and $\mathbf{R}_{\mathbf{q}\mathbf{q}^{(k)}}$ to zero [230].

$$\mathbf{R}_{\mathbf{v}\mathbf{v}} = \begin{bmatrix} \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} \\ \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} \\ \mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} \\ \mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} \end{bmatrix}. \tag{2.22}$$

Intuitively, a $\mathcal{Q}$-proper quaternion is $\mathcal{C}^\alpha$-proper for all pure unit quaternions $\alpha = i, j,$ and k. Additionally, $\mathcal{R}^\alpha$ and $\mathcal{C}^\alpha$-properness are complementary and together they result in $\mathcal{Q}$-properness. As a special case, the propriety of scalar quaternion $q = r_1 + ir_i + jr_j + kr_k$, is equivalent to sphericity of v, i.e., it is called proper iff r_1, r_i, r_j and r_k are independent and identically distributed (i.i.d) [230]. Analogous to the complex case, two quaternion RVs $\mathbf{q}_1$ and $\mathbf{q}_2$ are *cross proper* vectors as

- Cross $\mathcal{R}^\alpha$-proper iff $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(\alpha)}} = \mathbf{0}$ with $\alpha \in \{i, j, k\}$.
- Cross $\mathcal{C}^\alpha$-proper iff $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(\alpha')}} = \mathbf{0}$ and $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(\alpha'')}} = \mathbf{0}$ with arbitrary one-to-one mapping between $\{\alpha, \alpha', \alpha''\}$ and $\{i, j, k\}$.
- Cross $\mathcal{Q}$-proper iff all $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(i)}} = \mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(j)}} = \mathbf{R}_{\mathbf{q}_1\mathbf{q}_2^{(k)}} = \mathbf{0}$.

Similarly, two quaternion RVs $\mathbf{q}_1$ and $\mathbf{q}_2$ are *jointly-proper* iff they are $\mathcal{R}^\alpha$-, $\mathcal{C}^\alpha$- or $\mathcal{Q}$-proper and respectively cross proper [138].

Degree of Improperness

The degree of Π-properness of a quaternion RV $\mathbf{q}$ with $\Pi \in (\mathcal{R}^\alpha, \mathcal{C}^\alpha, \mathcal{Q})$ and augmented covariance matrix $\mathbf{R}_{\underline{\mathbf{q}}\mathbf{q}}$ is evaluated as

$$\mathcal{P}_\Pi = \min_{\hat{\mathbf{R}}_{\underline{\mathbf{q}}\mathbf{q}} \in \mathfrak{R}_\Pi} D\left(\mathbf{R}_{\underline{\mathbf{q}}\mathbf{q}} \parallel \hat{\mathbf{R}}_{\underline{\mathbf{q}}\mathbf{q}}\right), \tag{2.23}$$

where $\mathfrak{R}_{\Pi}$ represents the set of proper augmented covariance matrices with Π-properness e.g.,

$$\mathfrak{R}_{\mathcal{Q}} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(i)}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(j)}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(k)}} = \mathbf{0}\}, \tag{2.24}$$

$$\mathfrak{R}_{\mathcal{C}^{\alpha}} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha')}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha'')}} = \mathbf{0}\}, \tag{2.25}$$

$$\mathfrak{R}_{\mathcal{R}^{\alpha}} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha)}} = \mathbf{0}\}. \tag{2.26}$$

Moreover, $D\left(\mathbf{R}_{\underline{\mathbf{qq}}} \parallel \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}\right)$ is the Kullback-Leibler divergence [240] between two zero-mean quaternion Gaussian distributions with $\mathbf{R}_{\underline{\mathbf{qq}}}$ and $\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}$ [138, Table III]. Interestingly, the Pythagorean theorem for exponential families of PDF's render $\mathcal{P}_{\mathcal{Q}} = \mathcal{P}_{\mathcal{R}^{\alpha}} + \mathcal{P}_{\mathcal{C}^{\alpha}}$ with $\mathcal{P}_{\mathcal{Q}}, \mathcal{P}_{\mathcal{R}^{\alpha}},$ and $\mathcal{P}_{\mathcal{C}^{\alpha}}$ as defined in Table 2.4 [138].

Discussion

Quaternion propriety definitions are restricted to the SOS which completes the analysis for Gaussian quaternions, e.g., the impropriety measure $\mathcal{P}_{\Pi}$ is also the non-circularity measure for a zero-mean Gaussian quaternion RV based on its complete characterization using SOS. Nevertheless, further investigation of the nth-order properness is required to handle other distributions. As an example, the higher-order discrete rotational invariance analysis is required to tackle 4D constellations in communication systems [241].

2.2.3 Summary and Insights

The extent of complex and quaternion RV properness is classified in two (strict and generalized proper) and three ($\mathcal{R}^{\alpha}, \mathcal{C}^{\alpha}, \mathcal{Q}$-proper) categories, respectively. However, the strongest versions of properness in complex and quaternion RVs have different implications i.e., a strictly proper complex RV $\mathbf{z}$ may contain correlated $\Re\{z_k\}$ and $\Im\{z_l\}$ for $k \neq l$ [94] whereas a strictly proper quaternion ($\mathcal{Q}$-Proper) RV $\mathbf{q}$ cannot contain correlated $\Re\{q_k\}$ and $\Im\{q_l\}$ with $\mathrm{E}\{\Re\{q_k\}\Im\{q_l\}\} = 0 \; \forall \, (k \neq l)$. Furthermore, the correlation between two RVs can be assessed using cross and joint properness in both complex and quaternion RVs. Ad-

ditionally, the respective DoI not only provides the entropy loss due to the improperness but also helps in impropriety testing to identify the underlying type of properness for appropriate transformation and processing [232]. Moreover, the complex domain DoI is also useful in blind source separation [31] and the quaternion domain DoI can provide the error exponent of Neyman-Pearson detector for binary hypothesis testing. Nonetheless, these perks are only obtained with the usage of complex or augmented representations. The real composite representation can only identify the improperness of a complex RV based on the structure of the real covariance matrix, yet it fails to provide a measure for the DoI.

2.3 Transformations and Operations

Impropriety classification helps to identify the simplified form of processing and transformation in terms of computational complexity [219]. This subsection highlights suitable processing models for complex as well as quaternion RVs based on their propriety characterization.

2.3.1 Complex Random Vectors

Various transformations and operations of the complex RVs depend on their propriety characterization. The appropriate processing enables us to exploit the additional design freedom offered by the RVs' improperness and extract the information embedded in them.

Transformations

Three forms of transformations, i.e., real linear transformation (RLT) or widely linear transformation (WLT), complex linear transformation (CLT) and widely unitary transformation (WUT) are reviewed as:

- RLT $\mathbf{L} \in \mathbb{R}^{2M\text{x}2N}$ on the real composite vector $\mathbf{u} \in \mathbb{R}^{2N}$ yields another real composite vector $\mathbf{u}_{\text{RLT}} \in \mathbb{R}^{2M}$ as

$$\mathbf{u}_{\text{RLT}} = \begin{bmatrix} \mathbf{x}_{\text{RLT}} \\ \mathbf{y}_{\text{RLT}} \end{bmatrix} = \mathbf{L}\mathbf{u} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \tag{2.27}$$

Equally, the augmented complex model of the transformation $\mathbf{u}_{\text{RLT}} = \mathbf{L}\mathbf{u}$ is given as

$$\underline{\mathbf{w}} = \begin{bmatrix} \mathbf{w} \\ \mathbf{w}^* \end{bmatrix} = \sqrt{2}\mathbf{T}\mathbf{u}_{\text{RLT}} = \sqrt{2}\mathbf{T}\mathbf{L}\mathbf{u} = \{\mathbf{T}\mathbf{L}\mathbf{T}^{\text{H}}\}\{\sqrt{2}\mathbf{T}\mathbf{u}\} = \underline{\mathbf{N}}\mathbf{z}, \tag{2.28}$$

where,

$$\underline{\mathbf{N}} = \begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_2 \\ \mathbf{N}_2^* & \mathbf{N}_1^* \end{bmatrix}. \tag{2.29}$$

is the augmented description of the WLT or *linear-conjugate-linear transformation* preserving the block pattern structure. Thus, $\mathbf{w} = \mathbf{z}_{\text{WLT}} = \mathbf{N}_1\mathbf{z} + \mathbf{N}_2\mathbf{z}^*$ with

$$\mathbf{N}_1 = \frac{1}{2}\left[\mathbf{L}_{11} + \mathbf{L}_{22} + i\left(\mathbf{L}_{21} - \mathbf{L}_{12}\right)\right], \tag{2.30}$$

$$\mathbf{N}_2 = \frac{1}{2}\left[\mathbf{L}_{11} - \mathbf{L}_{22} + i\left(\mathbf{L}_{21} + \mathbf{L}_{12}\right)\right]. \tag{2.31}$$

- A CLT or *strictly linear* (SL) *transformation* (SLT) is a special case of the WLT when $\mathbf{z}_{\text{CLT}} = \mathbf{N}_1\mathbf{z}$ with $\mathbf{N}_2 = \mathbf{0}$ and the corresponding RLT is given as

$$\mathbf{z}_{\text{CLT}} = \begin{bmatrix} \mathbf{x}_{\text{CLT}} \\ \mathbf{y}_{\text{CLT}} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ -\mathbf{L}_{12} & \mathbf{L}_{11} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \tag{2.32}$$

To summarize, RLT on $\mathbb{R}^{2N}$ are linear on $\mathbb{C}^N$ if they have the matrix structure as in (2.32), otherwise they are WLT. Surprisingly, WLT in $\mathbb{C}$-domain provides more insight than the RLT in $\mathbb{R}$-domain but offers computationally expensive hardware

implementation [31].

- Interestingly, the EVD of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ is different from EVD of ordinary matrices and takes the unfamiliar form $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}} = \underline{\mathbf{U}}\Lambda\underline{\mathbf{U}}^{\mathrm{H}}$ with WUT $\underline{\mathbf{U}}$, satisfying $\underline{\mathbf{U}}^{\mathrm{H}}\underline{\mathbf{U}} = \underline{\mathbf{U}}\underline{\mathbf{U}}^{\mathrm{H}} = \mathbf{I}$ and the following structure [55, 234]

$$\underline{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_2^* & \mathbf{U}_1^* \end{bmatrix}. \tag{2.33}$$

  Moreover, the eigen-valued matrix $\underline{\Lambda}$ is given as

$$\underline{\Lambda} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{\mathrm{H}} = \frac{1}{2}\begin{bmatrix} \Lambda^{(1)} + \Lambda^{(2)} & \Lambda^{(1)} - \Lambda^{(2)} \\ \Lambda^{(1)} - \Lambda^{(2)} & \Lambda^{(1)} + \Lambda^{(2)} \end{bmatrix}, \tag{2.34}$$

  where, $\Lambda^{(1)} = \mathbf{Diag}\left(\lambda_1, \lambda_3, \ldots, \lambda_{2n-1}\right)$ and $\Lambda^{(2)} = \mathbf{Diag}\left(\lambda_2, \lambda_2, \ldots, \lambda_{2n}\right)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{2n}$ being the ordered eigenvalues of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$. This simplifies to a diagonal matrix $\Lambda = \mathbf{Diag}\left[\Lambda^{(1)}, \Lambda^{(2)}\right]$ for the special case of proper RVs.

Definition 2.14. *Any RV obtained from a complex proper RV* $\mathbf{z}$ *by a linear or affine transformation, i.e.,* $\hat{\mathbf{z}} = \hat{\mathbf{A}}\mathbf{z} + \hat{\mathbf{b}}$*, where* $\hat{\mathbf{A}} \in \mathbb{C}^{M\mathrm{x}N}$ *and* $\hat{\mathbf{b}} \in \mathbb{C}^M$ *are constant, is also proper. Thus, properness is preserved under affine transformations [94]. On contrary, the generalized propriety is also preserved by unitary WLT and arbitrary SLT [55]*

Operations

Inner products and quadratic forms of complex RVs are distinctly defined for different data representations.

- The inner product of two N-dimensional complex RVs $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{C}^N$ with real composite representation $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^{2N}$ and augmented representation $\underline{\mathbf{z}}_1, \underline{\mathbf{z}}_2 \in \mathbb{C}_*^{2N}$ is defined as

$$\mathbf{u}_1^{\mathrm{T}}\mathbf{u}_2 = \frac{1}{2}\underline{\mathbf{z}}_1^{\mathrm{H}}\mathbf{T}\mathbf{T}^{\mathrm{H}}\underline{\mathbf{z}}_2 = \frac{1}{2}\underline{\mathbf{z}}_1^{\mathrm{H}}\underline{\mathbf{z}}_2 = \Re\{\mathbf{z}_1^{\mathrm{H}}\mathbf{z}_2\}. \tag{2.35}$$

- Moreover, the real-valued quadratic operation can be described using different representations as [16, 36].

$$\mathbf{u}^{\mathrm{T}}\mathbf{L}\mathbf{u} = \frac{1}{2}\underline{\mathbf{z}}^{\mathrm{H}}\underline{\mathbf{N}\mathbf{z}} = \Re\{\mathbf{z}^{\mathrm{H}}\mathbf{N}_1\mathbf{z} + \mathbf{z}^{\mathrm{H}}\mathbf{N}_2\mathbf{z}^*\}. \tag{2.36}$$

Discussion

Remarkably, the maximal invariants of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ depend on the underlying transformations. Meaning thereby, any function of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ that is invariant under particular transformation must be a function of these maximal invariants only. For example, the eigenvalues of $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ constitute a maximal invariant for $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ under WUT whereas the circularity coefficients constitute a maximal invariant for $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$ under non-singular SLT [31].

2.3.2 Quaternion Random Vectors

The type of quaternion properness defines several types of linear processing [232]

- *Full-widely* linear processing is the general optimal linear processing with the simultaneous operation on the quaternion and its involutions as

$$\mathbf{q}_{\mathrm{f}} = \mathbf{F}_1^{\mathrm{H}}\mathbf{q} + \mathbf{F}_i^{\mathrm{H}}\mathbf{q}^{(i)} + \mathbf{F}_j^{\mathrm{H}}\mathbf{q}^{(j)} + \mathbf{F}_k^{\mathrm{H}}\mathbf{q}^{(k)}, \tag{2.37}$$

  where, $\mathbf{q}_{\mathrm{f}} \in \mathbb{H}^{P\mathrm{x}1}$ and $\mathbf{F}_1^{\mathrm{H}}, \mathbf{F}_i^{\mathrm{H}}, \mathbf{F}_j^{\mathrm{H}}, \mathbf{F}_k^{\mathrm{H}} \in \mathbb{H}^{P\mathrm{x}N}$. Full-WLT for augmented representation takes the form $\underline{\mathbf{q}}_{\mathrm{f}} = \mathbf{F}_{\underline{\mathbf{q}}}^{\mathrm{H}}\underline{\mathbf{q}}$ with $\mathbf{F}_{\underline{\mathbf{q}}} = \left[\mathbf{F}_1^{\mathrm{T}}\ \mathbf{F}_i^{\mathrm{T}}\ \mathbf{F}_j^{\mathrm{T}}\ \mathbf{F}_k^{\mathrm{T}}\right]^{\mathrm{T}}$, whereas for the real composite representation it becomes $\mathbf{v}_{\mathrm{f}} = \mathbf{F}_{\mathbf{v}}^{\mathrm{T}}\mathbf{v}$ with $\mathbf{F}_{\mathbf{v}} = \mathbf{A}_N^{\mathrm{H}}\mathbf{F}_{\underline{\mathbf{q}}}\mathbf{A}_P$.

- *Semi-widely* linear processing is for the $\mathcal{C}^{\alpha}$-proper vectors and takes the form $\mathbf{q}_{\mathrm{s}} = \mathbf{F}_1^{\mathrm{H}}\mathbf{q} + \mathbf{F}_{\alpha}^{\mathrm{H}}\mathbf{q}^{(\alpha)}$.

- The *conventional* linear processing of the $\mathcal{Q}$-proper vector take the simplified form $\mathbf{q}_{\mathrm{c}} = \mathbf{F}_1^{\mathrm{H}}\mathbf{q}$.

Notably, $\mathcal{R}^{\alpha}$-proper quaternion RVs do not result in a simplified linear processing model [232]. A $\mathcal{Q}$-proper RV is invariant under affine transformation [231] whereas the $\mathcal{C}^{\alpha}$-properness is invariant under semi-WLT [138].

2.3.3 Summary and Insights

In conclusion, all the transformations and operations can be equivalently carried out in all three data formulations. Of all transformations, the properness is only preserved under affine or SLT i.e., the resultant vector after transformation will carry the same propriety value as that of the original vector. Propriety characterization is especially significant to apply the most simplified form of processing while carrying out transformations and operations. Furthermore, exploiting the WL model of the RV provides vast benefits, e.g., it can provide four times faster convergence of the quaternion least-mean square adaptive filtering relative to its quadrivariate counterpart [219, 221].

2.4 Entropy and Probability Density Functions

Information entropy quantifies the average ambiguity i.e., the amount of information or uncertainty in an event. Thus, impropriety incorporation is crucial to evaluate the information theoretic entropy losses as well as the probability distributions of the generalized Gaussian RVs. This section presents the generalized framework of these analysis for both complex and quaternion RVs.

2.4.1 Complex Random Vectors

This part contains the definitions of differential entropy and probability distributions for a general complex Gaussian RV.

Differential entropy

The entropy is broadly categorized as discrete and differential entropy to measure the amount of surprise in discrete and continuous RV, respectively. Focusing on the differential entropy of a complex RV $\mathbf{z}$ with composite real representation $\mathbf{u} = [\mathbf{x}^{\mathrm{T}}\ \mathbf{y}^{\mathrm{T}}]^{\mathrm{T}}$ when $\mathbf{x}$ and $\mathbf{y}$ are both Gaussian distributed, gives [55]

$$H(\mathbf{u}) = \frac{1}{2}\log\left[(2\pi e)^{2N}\left|\mathbf{R}_{\mathbf{uu}}\right|\right]. \tag{2.38}$$

Equivalently, the differential entropy of $\mathbf{z}$ in terms of the augmented covariance matrix $\mathbf{R}_{\underline{\mathbf{zz}}}$ is [36]

$$H(\mathbf{z}) = \frac{1}{2}\log\left[(\pi e)^{2N}\left|\mathbf{R}_{\underline{\mathbf{zz}}}\right|\right]. \tag{2.39}$$

Definition 2.15. *The differential entropy of a complex RV* $\mathbf{z}$ *with a fixed correlation matrix is maximum, iff the RV is zero-mean Gaussian and proper [94], where it is given by*

$$H_P(\mathbf{z}) = \log\left[(\pi e)^{N}\left|\mathbf{R}_{\mathbf{zz}}\right|\right]. \tag{2.40}$$

Remark For a scalar, the differential entropy is appropriately defined as the joint differential entropy of its real and imaginary parts [94]. The difference is the differential entropies of a proper and improper complex Gaussian RV $\mathbf{z}$ is given by the mutual information $\mathcal{I}(\mathbf{z};\mathbf{z}^*)$ between $\mathbf{z}$ and $\mathbf{z}^*$ as [36]

$$H_I(\mathbf{z}) = H_P(\mathbf{z}) - \mathcal{I}(\mathbf{z};\mathbf{z}^*), \tag{2.41}$$

where, $\mathcal{I}(\mathbf{z};\mathbf{z}^*) = -\frac{1}{2}\log\prod_{i=1}^{N}(1-k_i^2)$ is a function of ρ_1 (2.17).

Complex-valued Gaussian Distribution

Van Den Bos was the pioneer to formulate a general multivariate Gaussian distribution for improper complex processes. He demonstrated that the conventional definition of the complex Gaussian distribution (based on the covariance matrix) is only a special case and thus, applicable to proper processes only [134]. The PDF of a general complex Gaussian RV $\mathbf{z}$ with $\mathbf{u} = [\mathbf{x}^{\mathrm{T}}\ \mathbf{y}^{\mathrm{T}}]^{\mathrm{T}}$ when $\mathbf{x}$ and $\mathbf{y}$ are both Gaussian distributed, is given as

$$p\left(\mathbf{u}\right) = \frac{1}{\sqrt{\left(2\pi\right)^{2N}\left|\mathbf{R}_{\mathbf{uu}}\right|}} \exp\left\{-\frac{1}{2}\left(\mathbf{u}-\mu_{\mathbf{u}}\right)^{\mathrm{T}} R_{\mathbf{uu}}^{-1}\left(\mathbf{u}-\mu_{\mathbf{u}}\right)\right\}. \tag{2.42}$$

Equivalently, the PDF of a complex RV $\mathbf{z}$ with augmented covariance matrix $\mathbf{R}_{\underline{\mathbf{zz}}}$ is written as [134]

$$p\left(\mathbf{z}\right) = \frac{1}{\sqrt{\pi^{2N}\left|\mathbf{R}_{\underline{\mathbf{zz}}}\right|}} \exp\left\{-\frac{1}{2}\left(\underline{\mathbf{z}}-\mu_{\underline{\mathbf{z}}}\right)^{\mathrm{H}} \mathbf{R}_{\underline{\mathbf{zz}}}^{-1}\left(\underline{\mathbf{z}}-\mu_{\underline{\mathbf{z}}}\right)\right\}. \tag{2.43}$$

2.4.2 Quaternion Random Vector

This subsection contains the definitions of differential entropy and probability distributions for a general multivariate quaternion RV.

Differential entropy

The differential entropy of a quaternion RV $\mathbf{q}$ with real representation $\mathbf{v}$ and the respective covariance matrix $\mathbf{R}_{\mathbf{vv}}$ is given by [138]

$$H_{\mathbf{q}}\left(\mathbf{R}_{\mathbf{vv}}\right) = 2N\ln\left(2\pi e\right) + \frac{1}{2}\ln\left|\mathbf{R}_{\mathbf{vv}}\right|. \tag{2.44}$$

Equivalently, the augmented representation renders [231]

$$H_{\mathbf{q}}\left(\mathbf{R}_{\underline{\mathbf{qq}}}\right) = 2N\ln\left(\frac{\pi e}{2}\right) + \frac{1}{2}\ln\left|\mathbf{R}_{\underline{\mathbf{qq}}}\right|. \tag{2.45}$$

Definition 2.16. *The differential entropy of any quaternion RV is upper bounded by the differential entropy of a centered $\mathcal{Q}$-proper Gaussian RV with $\mathbf{R}_{\mathbf{vv}} = \sigma^2 \mathbf{I}_{4N}$ and is given by [231]*

$$H_{\mathbf{q}} \leqslant H_{\mathcal{Q}-\text{proper}} = 2N \ln \left(2\pi e \sigma^2\right). \tag{2.46}$$

Beyond the mutual information between two complex RVs, its generalization to higher dimensions is termed as *interaction information* [242]. Took *et al.* presents the interaction information between quaternion-valued Gaussian RVs $\mathbf{q}, \mathbf{q}_i, \mathbf{q}_j$ and $\mathbf{q}_k$ in [231, eq. 43]. Unlike mutual information, interaction information can be negative depicting a decrease in the degree of association between the variates in a multivariate quantity, when one variable is dealt as a constant. The impropriety measure $\mathcal{P}_{\mathcal{Q}}$ measures the interaction information between $\mathbf{q}$ and its involutions. It represents the entropy loss due to improperness as $\mathcal{P}_{\mathcal{Q}} = H_{\mathbf{q}}\left(\mathbf{R}_{\mathbf{qq}}\right) - H_{\mathbf{q}}\left(\mathbf{R}_{\underline{\mathbf{qq}}}\right)$. Similarly, $\mathcal{P}_{\mathcal{C}^\alpha}$ and $\mathcal{P}_{\mathcal{R}^\alpha}$ presents the entropy losses due to $\mathcal{C}^\alpha$- and $\mathcal{R}^\alpha$-improperness of $\mathbf{q}$, respectively [138]. In other words, $\mathcal{P}_{\mathcal{R}^\alpha}$ measures of the entropy loss due to the $\mathcal{Q}$-improperness of the $\mathcal{C}^\alpha$-proper quaternion vector.

Quaternion-valued Gaussian Distribution

The PDF of a zero-mean Gaussian RV $\mathbf{q}$ with real representation $\mathbf{R}_{\mathbf{vv}}$ is given by [138]

$$p\left(\mathbf{v}\right) = \frac{\exp\left(-\frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{v}\right)}{\left(2\pi\right)^{2N}\left|\mathbf{R}_{\mathbf{vv}}\right|^{1/2}}. \tag{2.47}$$

Equivalently, it can be expressed using the augmented representation as [231]

$$p\left(\underline{\mathbf{q}}\right) = \frac{\exp\left(-\frac{1}{2}\underline{\mathbf{q}}^{\mathrm{H}}\mathbf{R}_{\underline{\mathbf{qq}}}^{-1}\underline{\mathbf{q}}\right)}{\left(\pi/2\right)^{2N}\left|\mathbf{R}_{\underline{\mathbf{qq}}}\right|^{1/2}}. \tag{2.48}$$

In the Gaussian case, the distribution of a $\mathcal{H}$-proper variable is contained in a 4D hypersphere [225].

2.4.3 Summary and Insights

Impropriety i.e., correlation between a RV and it's conjugate results in loss of entropy and this loss can be quantified in terms of mutual information and interaction information in case of complex and quaternion RVs, respectively. Geometrically, it is the loss in capacity by enclosing the codewords in a multidimensional ellipsoid instead of a hypersphere [243]. The differential entropy is maximum for a zero-mean proper Gaussian RV. Any deviation from this trio i.e., zero-mean, properness, or Gaussianity will result in a loss of entropy. Conventionally, researchers adhered to this trio to maximize entropy in a communication system, however later divergence from this property was exploited to gain benefits in some interference-limited applications.

2.5 Testing for Impropriety

Impropriety testing is a procedure to characterize the impropriety features of the random signals. In particular, it involves the identification of a RV as proper or improper RV based on its random samples. Moreover, it sometimes includes the quantification of the extent of non-circularity of the potential improper RV. Impropriety testing is an important consideration in order to exploit the significant performance gains offered by the improper signals through appropriate processing. Impropriety defining pseudo-covariance matrix is practically estimated from the available data. However, the general non-zero estimate does not indicate that the source is actually improper.

Therefore, various studies have proposed strategies to test for the impropriety of the signal from its observations based on different assumptions as highlighted in Table 2.2. They include hypothesis testing based on likelihood ratio test (LRT) [129, 131, 156, 230, 232, 237, 238, 244–248, 251], maximum likelihood (ML) estimation of circularity coefficients [250], Wald's type detector [226, 249] and invariant testing [129] etc. For instance, Schreier *et al.* [237] and Olilla *et al.* [131] independently proposed a binary hypothesis test for im-

Table 2.2: Impropriety Evaluation based on various Tests involving Different Data Representations

Testing Criteria	**Analysis**	**Variables**	**Assumptions**
GLRT with augmented representation	Conventional	Complex RV	i.i.d samples of Gaussian RV and simulated test threshold [131] [237]
			independent but non-identical samples [244]
			i.i.d complex elliptically symmetric distributions [238]
			i.i.d samples of Gaussian RV and analytical test threshold [245]
	Frequency	Complex vector time-series	Stationary Gaussian vector sequence with spectral identification [246]
		Complex vector sequences	Stationary non-Gaussian vector sequence [247]
			Stationary non-Gaussian signal with improper/colored noise [156]
	Spectral Image	2D-spectrum	Stationary non-Gaussian image with improper Gaussian noise [248]
	Conventional	Quaternion RV	i.i.d sample of quaternion Gaussian RV [232]
LRT with real representation	Conventional	Quaternion RV	i.i.d sample of quaternion Gaussian RV [230]
GLRT with real representation		Complex RV	i.i.d samples of Gaussian RV and analytical test threshold [129]
Wald's type detector			i.i.d CES distributions [249] [226]
Circularity coefficients estimation			Asymptotic analysis for arbitrary distribution [250]
Invariant testing			i.i.d samples of Gaussian RV and analytical test threshold [129]

propriety of the complex Gaussian data based on generalized likelihood ratio test (GLRT) and simulated test thresholds. Whereas, [129] proposed analytical test threshold based on the theoretical analysis of the null asymptotic distribution of the test statistic. In [252], a real-valued formulation based on block-skew circulant matrices is considered for GLRT. Interestingly, these studies deal with the i.i.d random samples taken from the Gaussian distribution. The case of independent but not necessarily identical or non-Gaussian distribution of RVs is discussed in [244]. A robust Walds type ML (WTML)-detector for propriety is also presented, robust to deviations from the Gaussianity assumption, based on multiple i.i.d samples in the broad class of complex elliptically symmetric (CES) distributions [226, 249].

Previous parametric impropriety tests are limited to a sequence of independent RV. To address this concern, [246] and [247] extended the results to stationary Gaussian and non-Gaussian sequences, respectively, using power spectral representations. Interestingly, [246] specifies the frequencies causing improperness when propriety is invalid, whereas the approach in [247] provides only a decision on propriety. The impropriety testing is not only limited to vectors and sequences rather it has also been extended to quaternions [230, 232] and spectral images [248]. A compelling practical application is the detection of random transmit waveforms from the Noise Radars (electromagnetic systems that use random signals for detecting and locating reflecting objects) based on the circularity tests [251].

2.5.1 GLRT for Complex Random Vectors

Of all impropriety tests, hypothesis testing based on GLRT and its variants are the most popular owing to the simple derivation and intuitive interpretation of the detection rules. Although the procedure is not generally optimal in the Neyman-Pearson sense, it is still practical and reliable [129, 131, 253]. Moreover, GLR is well-known for its invariance

characteristics as the hypothesis test (like propriety) must be invariant to but not WLT [237, 254].

Let $\mathbf{z} \in \mathbb{C}^N$ be a complex Gaussian RV with the PDF given by (2.49). Now consider M i.i.d random samples $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_M$ taken from the Gaussian distribution with augmented mean $\mu_{\underline{\mathbf{z}}}$ and augmented covariance $\mathbf{R}_{\underline{\mathbf{zz}}}$. Let $\underline{\mathbf{Z}} = [\underline{\mathbf{z}}_1, \underline{\mathbf{z}}_2, \ldots, \underline{\mathbf{z}}_M]$ denotes the augmented sample matrix, where $\underline{\mathbf{z}}_m = [\mathbf{z}_m^{\mathrm{T}}\ \mathbf{z}_m^{\mathrm{H}}]^{\mathrm{T}}$ is the augmented sample vector. Then, the joint PDF of these samples is given by

$$p\left(\mathbf{Z}\right) = \pi^{-MN}\left|\mathbf{R}_{\underline{\mathbf{zz}}}\right|^{-M/2} \exp\left\{-\frac{1}{2}\sum_{m=1}^{M}\left(\underline{\mathbf{z}}_m - \mu_{\underline{\mathbf{z}}}\right)^{\mathrm{H}} \mathbf{R}_{\underline{\mathbf{zz}}}^{-1}\left(\underline{\mathbf{z}}_m - \mu_{\underline{\mathbf{z}}}\right)\right\}, \tag{2.49}$$

$$= \pi^{-MN}\left|\mathbf{R}_{\underline{\mathbf{zz}}}\right|^{-M/2} \exp\left\{-\frac{M}{2}\mathrm{Tr}(\mathbf{R}_{\underline{\mathbf{zz}}}^{-1}\hat{\mathbf{R}}_{\underline{\mathbf{zz}}})\right\}, \tag{2.50}$$

where $\hat{\mathbf{R}}_{\underline{\mathbf{zz}}}$ is the sample augmented covariance matrix

$$\hat{\mathbf{R}}_{\underline{\mathbf{zz}}} = \frac{1}{M}\sum_{m=1}^{M}(\underline{\mathbf{z}}_m - \hat{\mu}_{\underline{\mathbf{z}}})(\underline{\mathbf{z}}_m - \hat{\mu}_{\underline{\mathbf{z}}})^{\mathrm{H}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{zz}} & \hat{\tilde{\mathbf{R}}}_{\mathbf{zz}} \\ \hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}^{*} & \hat{\mathbf{R}}_{\mathbf{zz}}^{*}, \end{bmatrix} \tag{2.51}$$

where $\hat{\mu}_{\underline{\mathbf{z}}} = \frac{1}{M}\sum_{m=1}^{M}\underline{\mathbf{z}}_m$ is the sample augmented mean vector. The aim is to distinguish signals based on the binary hypothesis tests presented in Table 2.3. GLRT is the ratio of likelihood with $\mathbf{R}_{\underline{\mathbf{zz}}}$ constrained to have zero off-diagonal blocks i.e., $\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$, to likelihood with unconstrained $\mathbf{R}_{\underline{\mathbf{zz}}}$. Thus, GLRT is testing the block-diagonal structure of the augmented covariance matrix as [131, 237]

$$\lambda = \frac{\max\limits_{\mathbf{R}_{\underline{\mathbf{zz}}}, \tilde{\mathbf{R}}_{\mathbf{zz}}=\mathbf{0}} p\left(\mathbf{Z}\right)}{\min\limits_{\mathbf{R}_{\underline{\mathbf{zz}}}} p\left(\mathbf{Z}\right)}. \tag{2.52}$$

In a GLR, the unknown parameters are replaced by ML estimates [253]. It is well known that the unconstrained and constrained ($\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$) ML estimate of $\mathbf{R}_{\underline{\mathbf{zz}}}$ is the sample covariance matrix $\hat{\mathbf{R}}_{\underline{\mathbf{zz}}}$ and $\hat{\mathbf{R}}_{\mathbf{0}}$, respectively, where $\hat{\mathbf{R}}_{\mathbf{0}} = \textbf{Block-Diag}\left(\hat{\mathbf{R}}_{\mathbf{zz}}, \hat{\mathbf{R}}_{\mathbf{zz}}^{*}\right)$. Therefore,

Table 2.3: Test Thresholds for $\mathbb{C}-$GLR Hypothesis Tests

Null Hypothesis (Proper)	**Alternate Hypothesis** (Improper)	**Test Threshold**	**Ref**
$\mathsf{H}_0: \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$	$\mathsf{H}_1: \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \tilde{\mathbf{R}}^{\dagger}_{\mathbf{zz}}$	Absent	[131]
$\mathsf{H}_0: \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$	$\mathsf{H}_1: \ \tilde{\mathbf{R}}_{\mathbf{zz}} \neq \mathbf{0}$	Simulated	[237]
$\mathsf{H}_0: \ \text{all } \delta_n = 0$	$\mathsf{H}_1: \ \text{all } \delta_n \neq 0$	Approximated	[129]
$\mathsf{H}_0: \ \rho = 0$	$\mathsf{H}_1: \ \rho \neq 0$	Asymptotic	[226]

GLR can be expressed as

$$\begin{aligned} l = \lambda^{\frac{2}{M}} &= \left|\hat{\mathbf{R}}_{\mathbf{0}}^{-1}\hat{\mathbf{R}}_{\underline{\mathbf{zz}}}\right| \left(\exp\left\{-\frac{M}{2}\mathrm{Tr}(\hat{\mathbf{R}}_{\mathbf{0}}^{-1}\hat{\mathbf{R}}_{\underline{\mathbf{zz}}} - \mathbf{I})\right\}\right)^{\frac{2}{M}} \\ &= \frac{\left|\hat{\mathbf{R}}_{\underline{\mathbf{zz}}}\right|}{\left|\hat{\mathbf{R}}_{\mathbf{zz}}\right|^2} = \frac{\left|\hat{\mathbf{R}}_{\mathbf{zz}} - \hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}\hat{\mathbf{R}}_{\mathbf{zz}}^{-*}\hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}^{*}\right|}{\left|\hat{\mathbf{R}}_{\mathbf{zz}}\right|}. \end{aligned} \tag{2.53}$$

Thus, GLR test statistic is the quotient between the determinant of the Schur complement of $\hat{\mathbf{R}}_{\underline{\mathbf{zz}}}$ and the determinant of $\hat{\mathbf{R}}_{\mathbf{zz}}$. This test statistic is defined in [131] and the hypothesis testing involves the test for null $\tilde{\mathbf{R}}_{\mathbf{zz}}$ signifying proper RV $\mathbf{z}$ or known $\tilde{\mathbf{R}}^{\dagger}_{\mathbf{zz}}$ signifying improper RV $\mathbf{z}$. On the other hand, [237] presents slightly different hypothesis test as given in Table 2.3 along with the simulated thresholds. We can also employ the estimated canonical correlation matrix $\hat{\mathbf{K}}$ using $\hat{\mathbf{R}}_{\mathbf{zz}}$ and $\hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}$ as discussed in Definition 2.10. We then have

$$l = \left|\mathbf{I} - \hat{\mathbf{K}}\hat{\mathbf{K}}^{\mathrm{H}}\right| = \prod_{n=1}^{N}\left(1 - \hat{k}_n^2\right). \tag{2.54}$$

This explains that the GLR is a function of the squared canonical correlations which make up a complete/maximal set of invariants under linear transformation (LT). Contributions like [250] suggest to directly find the ML estimates of circularity coefficients which are the singular values of the empirical coherence matrix $\hat{\mathbf{C}}$, estimated from $\hat{\mathbf{R}}_{\mathbf{zz}}$ and $\hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}$. Moreover, two invariant tests in [129] not only rely on these maximal invariants but also suggest

the critical regions to minimize false alarms based on the Box approximation. These two invariant tests based on the real composite representation state null-hypothesis as all $\delta_k = 0$ where Takagi Factorization of $\mathbf{R_{uu}} = \mathbf{G}\Delta_1\mathbf{G}^{\mathrm{T}}$ with $\Delta_1 = \mathbf{Diag}\left(\left(\mathbf{I}+\Delta_2\right), \left(\mathbf{I}-\Delta_2\right)\right)$ and $\Delta_2 = \mathbf{Diag}\left(\delta_1, \delta_2, \ldots, \delta_N\right)$. They accept H_0 iff

$$T_1 \equiv \prod_{n=1}^{N}\left(1-\hat{\delta}_n^2\right) \geq c_1\left(\eta_1, M, N\right), \tag{2.55}$$

$$T_2 \equiv \sum_{n=1}^{N}\left(\hat{\delta}_n^2\right) \leq c_2\left(\eta_1, M, N\right), \tag{2.56}$$

where $c_1\left(\eta_1, M, N\right)$ and $c_2\left(\eta_1, M, N\right)$ are constants based on the probability of false alarm (PFA) and η_1 is the specific size of the test [129]. Interestingly, T_2 is the locally most powerful (LMP) test as it has as high a power as possible for alternatives H_1 i.e., those for which all k_n are small. However, no uniformly most powerful (UMP) invariant test for impropriety exists for the problem $N \geq 2$ [129]. Based on the standard ML theory, the statistic l possesses an asymptotic chi-squared distribution with $F = N(N+1)$ DoF under the null hypothesis [129, 238, 244]. Thus, the statistical analysis allows us to accept H_0 iff $-(M-N)\log T_1 < \chi_F^2(1-\eta_1)$ where $\chi_F^2(1-\eta_1)$ is the $100(1-\eta_1)\%$ point of the chi-square distribution with F DoF. Additionally, the Wald's type detector [226] for the uni-variate case suggests to reject null hypothesis based on the ML estimate of circularity quotient $\hat{\rho}$ of the underlying CES distribution with finite fourth-order moment if $M|\hat{\rho}|^2/\hat{\varsigma}_0 \geq \chi^2_{2;1-\eta_2}$, where $\chi^2_{2;1-\eta_2}$ denotes $(1-\eta_2)^{\mathrm{th}}$ quantile of the chi-square distribution with 2 DoF, η_2 PFA and $\hat{\varsigma}_0$ as given by [226, (20)]. An interesting extension is to look for the extent of non-circularity when the null-hypothesis is rejected e.g., [238] tests the hypothesis that a particular number of circularity coefficients vanish.

To conclude, the impropriety testing can be implemented with great computational efficiency using the block-skew circulant matrices for the composite real formulation of the GLRT [252]. One of the proposed efficient implementation involves the following test

statistic

$$r = \frac{2^{2N}\left|\hat{\mathbf{R}}_{\mathbf{uu}}\right|}{\left|\hat{\mathbf{R}}_{\mathbf{zz}}\right|^2}, \tag{2.57}$$

where $\hat{\mathbf{R}}_{\mathbf{uu}} = \frac{1}{M}\sum_{m=1}^{M}(\mathbf{u}_m - \hat{\mu}_{\mathbf{u}})(\mathbf{u}_m - \hat{\mu}_{\mathbf{u}})^{\mathrm{T}}$ is the sample covariance matrix based on the real representation in Definition 2.1a. Hellings *et al.* claim that this implementation reduces the computational complexity roughly by a factor of four when compared to the implementation of the augmented complex version in (2.53), owing to its inherent redundancy.

2.5.2 GLRT for Quaternion Random Vectors

Like complex RV, propriety testing has also been expanded for the quaternion RV based on its real composite [230] as well as augmented representation [232]. For instance, Via *et al.* suggest three GLRTs to identify two main kinds of quaternion properness. Based on the ML estimates of $\mathbf{R}_{\underline{\mathbf{qq}}}$ under three distinct hypotheses namely $\mathcal{Q}$-proper, $\mathcal{C}^{\alpha}$-proper and possibly improper vectors, the GLRTs binary hypothesis tests are given in Table 2.4. The hypothesis testing is based on the test whether $\mathbf{R}_{\underline{\mathbf{qq}}}$ belongs to the convex set of $\mathcal{Q}$-proper and $\mathcal{C}^{\alpha}$-proper augmented covariance matrices, given as $\mathfrak{R}_{\mathcal{Q}}$ (2.24) and $\mathfrak{R}_{\mathcal{C}^{\alpha}}$ (2.25), respectively, or not. Moreover, the GLRT test statistics are based on the ML estimate of $\mathbf{R}_{\underline{\mathbf{qq}}}$ i.e., $\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}$ and its constraint formulations $\hat{\mathbf{D}}_{\mathcal{Q}}$ and $\hat{\mathbf{D}}_{\mathcal{C}^{\alpha}}$ under the hypothesis $\mathsf{H}_{\mathcal{Q}}$ and $\mathsf{H}_{\mathcal{C}^{\alpha}}$, respectively [232]

$$\hat{\mathbf{D}}_{\mathcal{Q}} = \textbf{Block-Diag}\left(\hat{\mathbf{R}}_{\mathbf{qq}}, \hat{\mathbf{R}}^{(i)}_{\mathbf{qq}}, \hat{\mathbf{R}}^{(j)}_{\mathbf{qq}}, \hat{\mathbf{R}}^{(k)}_{\mathbf{qq}}\right), \tag{2.58}$$

$$\hat{\mathbf{D}}_{\mathcal{C}^{\alpha}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{qq}} & \hat{\mathbf{R}}_{\mathbf{qq}^{(i)}} & \mathbf{0}_N & \mathbf{0}_N \\ \hat{\mathbf{R}}^{(i)}_{\mathbf{qq}^{(i)}} & \hat{\mathbf{R}}^{(i)}_{\mathbf{qq}} & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \hat{\mathbf{R}}^{(j)}_{\mathbf{qq}} & \hat{\mathbf{R}}^{(j)}_{\mathbf{qq}^{(i)}} \\ \mathbf{0}_N & \mathbf{0}_N & \hat{\mathbf{R}}^{(k)}_{\mathbf{qq}^{(i)}} & \hat{\mathbf{R}}^{(k)}_{\mathbf{qq}} \end{bmatrix}. \tag{2.59}$$

GLRT comparisons for $\mathcal{Q}$-properness, $\mathcal{C}^{\alpha}$-properness and improperness are established

Table 2.4: Quaternion GLR Hypothesis Tests and Decision Criterion

	Null and Alternate Hypothesis	GLRT Statistic	GLRT Comparison
$\mathcal{Q}$-properness	$\mathsf{H}_{\mathcal{Q}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \in \mathfrak{R}_{\mathcal{Q}}$ $\mathsf{H}_{\mathcal{I}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \notin \mathfrak{R}_{\mathcal{Q}}$	$\mathcal{P}_{\mathcal{Q}}=-\frac{1}{2}\ln\left\lvert\hat{\mathbf{D}}_{\mathcal{Q}}^{-1/2}\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}\hat{\mathbf{D}}_{\mathcal{Q}}^{-1/2}\right\rvert$	$\mathcal{P}_{\mathcal{Q}} \underset{\mathsf{H}_{\mathcal{Q}}}{\overset{\mathsf{H}_{\mathcal{I}}}{\lessgtr}} \gamma_{\mathcal{Q}}$
$\mathcal{C}^{\alpha}$-properness	$\mathsf{H}_{\mathcal{C}^{\alpha}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \in \mathfrak{R}_{\mathcal{C}^{\alpha}}$ $\mathsf{H}_{\mathcal{I}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \notin \mathfrak{R}_{\mathcal{C}^{\alpha}}$	$\mathcal{P}_{\mathcal{C}^{\alpha}}=-\frac{1}{2}\ln\left\lvert\hat{\mathbf{D}}_{\mathcal{C}^{\alpha}}^{-1/2}\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}\hat{\mathbf{D}}_{\mathcal{C}^{\alpha}}^{-1/2}\right\rvert$	$\mathcal{P}_{\mathcal{C}^{\alpha}} \underset{\mathsf{H}_{\mathcal{C}^{\alpha}}}{\overset{\mathsf{H}_{\mathcal{I}}}{\lessgtr}} \gamma_{\mathcal{C}^{\alpha}}$
$\mathcal{R}^{\alpha}$-improperness	$\mathsf{H}_{\mathcal{Q}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \in \mathfrak{R}_{\mathcal{Q}}$ $\mathsf{H}_{\mathcal{C}^{\alpha}}$: $\mathbf{R}_{\underline{\mathbf{qq}}} \in \mathfrak{R}_{\mathcal{C}^{\alpha}}$	$\mathcal{P}_{\mathcal{R}^{\alpha}}=-\frac{1}{2}\ln\left\lvert\hat{\mathbf{D}}_{\mathcal{Q}}^{-1/2}\hat{\mathbf{D}}_{\mathcal{C}^{\alpha}}\hat{\mathbf{D}}_{\mathcal{Q}}^{-1/2}\right\rvert$	$\mathcal{P}_{\mathcal{R}^{\alpha}} \underset{\mathsf{H}_{\mathcal{Q}}}{\overset{\mathsf{H}_{\mathcal{C}^{\alpha}}}{\lessgtr}} \gamma_{\mathcal{R}^{\alpha}}$

with predefined fixed thresholds $\gamma_{\mathcal{Q}}, \gamma_{\mathcal{C}^{\alpha}}$ and $\gamma_{\mathcal{R}^{\alpha}}$, respectively.

2.5.3 Summary and Insights

Various studies present different test statistic and threshold levels for the GLR hypothesis tests. Nonetheless, no invariant test for impropriety is uniformly most powerful, they all are inclined and thus locally robust tests [129]. Most of the contributions describe the GLRT of circularity assuming complex normal data [131, 237], which is further adjusted to accommodate a broad class of CES distributions. If the propriety test is invalid, it is then useful to detect the number of latent NC signals in a complex Gaussian RV based on multiple hypothesis tests [255]. Similarly, GLRT for quaternions is focused on the identification of $\mathcal{Q}$-proper, $\mathcal{C}^{\alpha}$-proper or possibly improper RV.

The technical framework provides a comprehensive understanding of the impropriety concepts and identification. We will apply this knowledge to appropriately characterize different hardware impairments as presented in the next chapter.

Chapter 3

Hardware Impairment Modeling

HWIs impose a huge challenge on next-generation network planning and deployment especially at high-frequency [3, 7–9]. HWIs emerge in various RF stages including imperfections in analog-to-digital/digital-to-analog converters, non-linear high power amplifier/low noise amplifier, mismatched local oscillator and phase shifter, etc. [8, 9]. These hardware imperfections result not only in phase/amplitude errors and raised noise floor but also in an inevitable mixing of the desired and image signals. This distinct behavior motivates researchers to develop accurate models and propose effective compensation methods to meet the expected performance.

Numerous efforts have been carried out to accurately model various forms of HWIs. As an example, many studies focused on the statistical modeling of additive hardware distortions at the transmitter and the receiver [10–13, 76, 108, 109]. Other studies focused on the modeling of in-phase and quadrature phase IQI where the self-interference (SI) signal incorporates the amplitude and rotational imbalance besides the receiver thermal noise [71, 72, 113, 114, 116, 256]. Moreover, a few studies like [8] and [117] analyzed multiple RF front-end impairments and their individual baseband equivalent error models. In [8], Schenk studied the modeling procedure, impact of non-ideal hardware on the system performance, and digital compensation schemes of various RF imperfections in high data-rate wireless systems. Similarly, Boulogeorgos *et al.* studied the impact of various HWIs on the energy detection spectrum sensing in cognitive radio systems in [117]. However, we considered the combined effect of various HWIs including both transmitter and receiver I/Q mismatch as well as accumulative additive distortions at the transmitter as well as the

receiver [20]. This research investigates full characterization of the self-interfering (SI) information signals and improper Gaussian noise components inspired by the statistical signal processing, which highlighted the asymmetric characteristics of baseband communication signals due to the IQI [31].

3.1 Transceiver Hardware Distortion Model

Consider a single-link wireless communication system suffering from various hardware impairments. The non-linear transfer functions of various transmitter RF stages, such as digital-to-analog converter, band-pass filter and high power amplifier result in accumulative additive distortion noise $\eta_{\rm t} \sim \mathcal{CN}(0, \kappa_{\rm t}, \tilde{\kappa}_{\rm t})$, where $|\tilde{\kappa}_{\rm t}| \leq \kappa_{\rm t}$ [8, 76]. These distortions raise the noise floor of the transmitted signal $x_{\rm tx} = x + \eta_{\rm t}$, where x is the transmission signal. The transmitted signal further undergoes a slowly varying flat Rayleigh fading channel $h \sim \mathcal{CN}(0, \lambda, 0)$. Moreover, the receiver further induces an additive distortion $\eta_{\rm r}$, resulting from the non-linear transfer function of low noise amplifier, band-pass filters, image rejection low pass filter, analog-to-digital converter. It is important to highlight that the receiver distortions are in addition to the conventional thermal noise at the receiver.

$$y = \sqrt{p} h \left(x + \eta_{\rm t}\right) + \eta_{\rm r} + z, \tag{3.1}$$

where p is the transmitted power and z is the additive white Gaussian noise (AWGN) with variance σ_z^2. Furthermore, $\eta_{\rm t}$ and $\eta_{\rm r}$ are the respective additive impairment distortions at the transmitter and the receiver. Various theoretical investigations and measurement results indicate that the Gaussian model accurately describes the aggregate of all residual RF impairments when compensation algorithms are applied to mitigate hardware impairments ([10, 12–14, 76, 109, 117, 257–261] and references therein). This can also be motivated analytically by the central limit theorem. In addition, $\eta_{\rm t}$ and $\eta_{\rm r}$ are generalized as asymmetric signals pertaining to the transformation caused by some hardware impairments such as

HWD with wide linear transformation characteristics [31, 81]. Thus, the aggregate HWDs at the transmitter and the receiver are random variables with $\eta_{\rm t} \sim \mathcal{CN}(0, \kappa_{\rm t}, \tilde{\kappa}_{\rm t})$ where $|\tilde{\kappa}_{\rm t}| \leq \kappa_{\rm t} \leq \sigma_x^2$ and $\eta_{\rm r} \sim \mathcal{CN}(0, p|h|^2\kappa_{\rm r}, ph^2\tilde{\kappa}_{\rm r})$, where $|\tilde{\kappa}_{\rm r}| \leq \kappa_{\rm r} \leq \sigma_x^2$. Clearly, the proposed model reduces to the well-known system model $y = \sqrt{p}hx + z$ in the absence of transceiver distortions i.e. $\eta_{\rm t} = \eta_{\rm r} = 0$.

Lemma 1. *The equivalent generalized aggregate model of the HWD is given by*

$$y = \sqrt{p}h\left(x + \eta\right) + z, \tag{3.2}$$

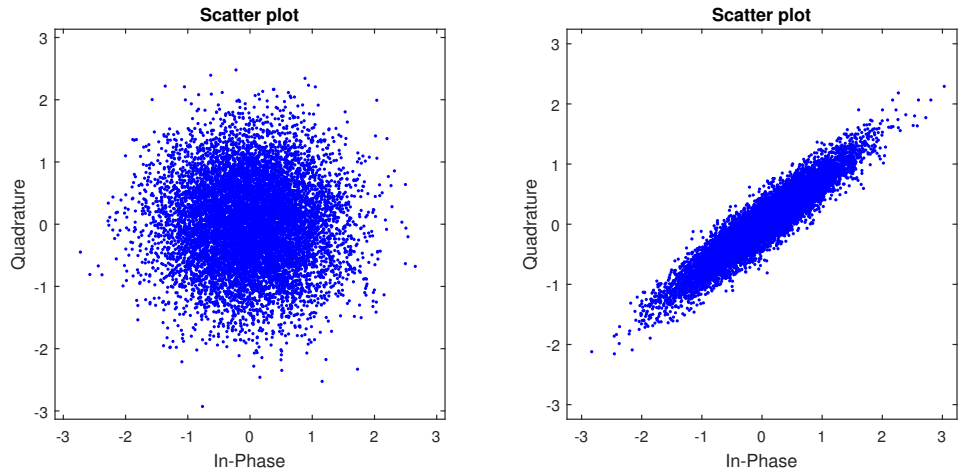
with $\eta = \eta_{\rm t} + \eta_{\rm r}$ distributed as $\mathcal{CN}(0, \kappa, \tilde{\kappa})$. Also, $\kappa = \kappa_{\rm t} + \kappa_{\rm r}$ and $\tilde{\kappa} = \tilde{\kappa}_{\rm t} + \tilde{\kappa}_{\rm r}$ capture the aggregate HWD at both the transmitter and the receiver along with impact of fading channels.

Proof. We assumed a general asymmetric model for the additive distortion with $\eta_{\rm t} \sim \mathcal{CN}(0, \kappa_{\rm t}, \tilde{\kappa}_{\rm t})$ and $\eta_{\rm r} \sim \mathcal{CN}(0, p|h|^2\kappa_{\rm r}, ph^2\tilde{\kappa}_{\rm r})$. Different from the existing literature that assumes symmetric HWD, we assume the general asymmetric scenario where having symmetric distortion at both in-phase and quadrature components is not the only possible scenario. Furthermore, the symmetric distortion can be transformed into asymmetric one after passing through WLT [31]. Illustrative demonstrated shows that the symmetric distribution of the proper additive distortions noise in Figure 3.1a transforms to the improper aggregated noise in Figure 3.1b pertaining to the WLT induced by the I/Q imbalance (refer to Section 3.2 for details). For a given fading channel, the variance and pseudo-variance of the aggregated impairments in (3.1) are expressed, respectively, as follows

$$\mathbb{E}\left[\left|\sqrt{p}h\eta_{\rm t} + \eta_{\rm r}\right|^2\right] = p|h|^2\left(\kappa_{\rm t} + \kappa_{\rm r}\right), \tag{3.3}$$

$$\mathbb{E}\left[\left(\sqrt{p}h\eta_{\rm t} + \eta_{\rm r}\right)^2\right] = ph^2\left(\tilde{\kappa}_{\rm t} + \tilde{\kappa}_{\rm r}\right). \tag{3.4}$$

The variance and pseudo-variance of the aggregated asymmetric distortions in (3.2) are



(a) Proper Distortion Noise

(b) Improper Distortion Noise

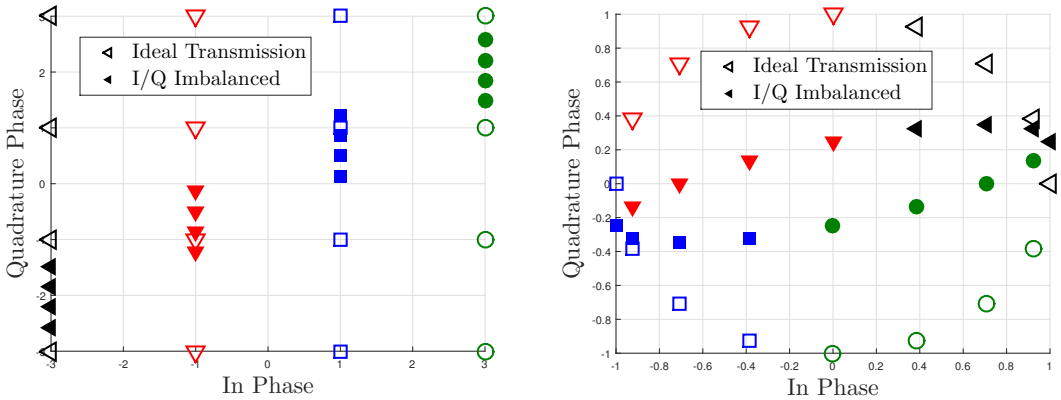
Figure 3.1: Impropriety Characterization

$p|h|^2\kappa$ and $ph^2\tilde{\kappa}$, respectively. Thus, the signal model in (3.1) can be equivalently modeled as in (3.2) when $\kappa = \kappa_{\rm t} + \kappa_{\rm r}$ and $\tilde{\kappa} = \tilde{\kappa}_{\rm t} + \tilde{\kappa}_{\rm r}$. □

It is important to note that in case of ideal hardware i.e. $\eta = 0$ is imposed by $\kappa = 0$ and $\tilde{\kappa} = 0$. Where, $\kappa = 0$ is dictated by the negligible transmitter and receiver distortion variances, $\kappa_{\rm t} = 0$ and $\kappa_{\rm r} = 0$, respectively. Also, $\tilde{\kappa} = 0$ follows from Definition 2.9.

3.2 Transceiver I/Q Imbalance Model

Consider a wireless communication system employing synchrodyne architecture transceivers in the RF front-end [262]. At the transmitter, the digital baseband modulated signal undergoes pulse shaping and up-conversion to the desired carrier frequency. During this homodyne up-conversion stage, the imperfect local oscillator and phase shifter introduce I/Q imbalance. Transceivers with I/Q imbalance not only contribute to the amplitude and phase errors but also induce self-interference (SI) signal $x_{\rm IQI}$, due to the limited image-rejection capability [5]. Based on the practically validated I/Q mismatch model presented in [8, 72, 256, 263–267] and references therein, the equivalent quadrature imbalanced transmit signal is expressed as $x_{\rm IQI} = \nu_1 x + \nu_2 x^*$, where x is a band-pass modulated signal taken



(a) QAM Modulation (b) PSK Modulation

Figure 3.2: Transmitter I/Q imbalance

from M-QAM, M-PSK or M-PAM constellation, ν_1 and ν_2 account for the amplitude (a_{t}) and phase (θ_{t}) errors at the transmitter I/Q mixer stage and are modeled as

$$\nu_1 = \left(1 + a_{\mathrm{t}} e^{j\theta_{\mathrm{t}}}\right)/2;\ \nu_2 = 1 - \nu_1^* = \left(1 - a_{\mathrm{t}} e^{-j\theta_{\mathrm{t}}}\right)/2. \tag{3.5}$$

The image rejection ratio (IRR) of transmitter I/Q mixer stages is defined as $\Upsilon_{\mathrm{t}} = |\nu_1|^2/|\nu_2|^2$, resides between 20 dB and 40 dB [117, 256]. The baseband equivalent impact of transmitter I/Q imbalance on 16-QAM and 16-PSK modulations is evident in Figure 3.2a and 3.2b, respectively. I/Q imbalance transmits the filled symbols from a skewed and rotated signal constellation in-place of the ideally modulated hollow symbols. Note that the ideal transmitter I/Q mixer stage exhibits a_{t}=1 and θ_{t}=0, yielding $x_{\mathrm{IQI}} = x$ and $\Upsilon_{\mathrm{t}} = \infty$. In addition, non-linear transfer functions of various transmitter RF stages result in accumulative additive distortion noise as modelled in Section 3.1.

$$x_{\mathrm{tx}} = x_{\mathrm{IQI}} + \eta_{\mathrm{t}} = \nu_1 x_m + \nu_2 x_m^* + \eta_{\mathrm{t}}. \tag{3.6}$$

The transmitted signal undergoes a slowly varying flat Rayleigh fading channel $h \sim \mathcal{CN}(0, \lambda, 0)$ and receiver additive distortion η_{r}. The received signal is expressed in terms

of the average transmit power p as $y_{\rm rx} = \sqrt{p}hx_{\rm tx} + \eta_{\rm r}$, where it again experiences I/Q imbalance at the down-conversion stage yielding $y = \mu_1 y_{\rm rx} + \mu_2 y_{\rm rx}^*$ [8, eq. 5.28] and [72, eq. 3]. The parameters $\mu_1 = \left(1 + a_{\rm r}e^{j\theta_{\rm r}}\right)/2$ and $\mu_2 = 1 - \mu_1^* = \left(1 - a_{\rm r}e^{-j\theta_{\rm r}}\right)/2$ capture the receiver's amplitude $(a_{\rm r})$ and rotational errors $(\theta_{\rm r})$. Thus, the receiver's IRR in the down-conversion process is $\Upsilon_{\rm r} = |\mu_1|^2/|\mu_2|^2$, where the perfect receiver I/Q balance occurs when $\mu_1 = 1$ and $\mu_2 = 0$, which follows from $a_{\rm r} = 1$ and $\theta_{\rm r} = 0$. Consequently, the received signal y under the aggregate HWIs model is given as

$$y = \sqrt{p}\tilde{h}_1 x + \sqrt{p}\tilde{h}_2 x^* + z, \tag{3.7}$$

where $\tilde{h}_1 = \mu_1\nu_1 h + \mu_2\nu_2^* h^*$ and $\tilde{h}_2 = \mu_1\nu_2 h + \mu_2\nu_1^* h^*$ are modified channel gain parameters accommodating fading and I/Q imbalance characteristics. Moreover, the aggregated noise is given by

$$z = \mu_1\left(\sqrt{p}h\eta_{\rm t} + \eta_{\rm r}\right) + \mu_2\left(\sqrt{p}h^*\eta_{\rm t}^* + \eta_{\rm r}^*\right). \tag{3.8}$$

Theorem 1. *Transceiver I/Q imbalance transforms the symmetric transmitted signal to asymmetric received signal and the proper Gaussian interference to improper Gaussian interference [20].*

Proof. This follows from the non-zero pseudo-variance of the received signal $\tilde{\sigma}_y^2 = 2\tilde{h}_1\tilde{h}_2 + \tilde{\sigma}_z^2$ and the accumulative interference term under HWI $\tilde{\sigma}_z^2 = 2\mu_1\mu_2\left(p|h|^2\sigma_{\rm t}^2 + \sigma_{\rm r}^2\right)$. □

The appropriate detection and precise modeling of HWIs is essential for accurate analysis and efficient mitigation. My research aims at the performance analysis, effective transmission design and optimal detection based on the presented impairment models. Next, we present different case studies to study the drastic impact of different HWIs on various system settings and the appropriate coping mechanisms.

Chapter 4

Hardware Impaired Multihop Full-Duplex Relaying Systems

In this chapter, we analyze the performance degradation of a multihop decode-and-forward FDR (MH-DF-FDR) system caused by the residual self-interference (RSI) and hardware distortions (HWD) imposed by the full duplex relaying (FDR) operation and imperfect hardware, respectively. In addition, we study the benefits of employing improper Gaussian signaling (IGS) in the MH-FDR system. Different from the traditional symmetric signaling scheme, i.e., proper Gaussian signaling (PGS), IGS has non-zero pseudo-variance that can limit the impact of RSI and HWD in the MH-FDR system. To evaluate the system performance gain using IGS, first we express the end-to-end achievable rate of the MH system as the minimum rate supported by all participating links. Then, we optimize the pseudo-variance of all participating transmitters including source and relays to compensate the interference impact and improve the end-to-end achievable rate. We propose two network optimization schemes based on the system characteristics i.e. joint optimization framework and distributed optimization scenario. Interestingly, IGS-based scheme outperforms its counterpart PGS-based scheme, especially at higher interference-to-noise ratio. Our findings reveal that using IGS in single-user detection systems that suffer from both RSI and HWD can effectively mitigate the degradation in the achievable rate performance.

The rest of the chapter is organized in the following fashion. Section 4.1 highlights the significance and contributions of this chapter. Section 4.2 studies the statistical model for the MH-DF-FDR system under asymmetric self-interference and transceiver distortions. Section 4.3 focuses on the information-theoretic achievable rate in the presence of improper Gaussian interference and signal. Section 4.4 focuses on designing the transmission

parameters based on the RSI and HWD characteristics. It deals with the joint as well as distributed optimization framework to fine-tune the statistical IGS parameters to achieve optimum system performance. Section 4.5 numerically analyzes the system performance with/without IGS assuming different system parameters followed by the insightful conclusion in Section 4.6.

4.1 Significance and Contributions

Analyzing the impact of hardware imperfections and RSI on the system performance and evaluating different compensation schemes require an accurate statistical model of these imperfections. RSI and HWD are modelled by widely linear transformations as discussed in [78] and [79], respectively. Thus, according to statistical signal processing studies, widely linear precoders/transformations can efficiently map symmetric information-bearing signals to asymmetric signals at each transmitter [31,81]. Therefore, this work models both RSI and HWD as asymmetric signals. Furthermore, the proposed research can employ IGS for signal transmission to jointly mitigate the deterrent effect of both RSI and HWD. IGS scheme has already been proven to evidently improve system performance in various system configurations such as multiple-input multiple output systems [108, 110], cognitive radio systems [191,192], full-duplex relaying [79,179,268], and alternating relaying [159]. The IGS transmission scheme is expected to outperform PGS in the presence of RSI and HWD in a MH-DF-HWD-FDR system. Our objective is to quantify the gain obtained by optimal IGS over PGS and to evaluate if the gain is significant enough to adopt IGS optimization framework as the optimal IGS solution can sometimes reduce to PGS. In order to maximize the end-to-end rate, the IGS transmit signaling characteristics of the source and relay have already been optimized as part of the MS dissertation for dual-hop DF-FDR system [153].

In this chapter, we study the utilization of asymmetric signaling scheme instead of the symmetric signaling scheme to combat both the RSI and HWD in MH-DF-FDR systems.

Symmetric signaling or PGS is the traditional signaling scheme that assumes independent signal components with equal power, which is described by its variance. On the other hand, asymmetric signaling or IGS relaxes the PGS characteristics and can have dependent signal component with/without equal power. Therefore, an IGS needs an additional statistical quantity to be accurately characterized, which is called the pseudo-variance [94]. We should note that IGS can be practically implemented using widely linear precoders, which efficiently maps symmetric information-bearing signals to asymmetric signals at each transmitter [31, 81]. The main contributions of this chapter are summarized as follows:

- Studying the effect of HWDs and RSI on the achievable rate performance of the MH-FDR system in the absence of direct link.
- Employing IGS to compensate the degradation on the achievable rate performance due to both HWDs and RSI in MH-FDR systems.
- Developing a rigorous joint optimization framework to design the signal characteristics by tuning the signal symmetry degree in terms of the pseudo-variance in order to maximize the end-to-end achievable rate of the MH-FDR system.
- Developing a distributed optimization framework that suits practical implementation of the proposed transmission scheme offering reduced round-trip delays, computational complexity and communication overhead.

4.2 FDR under HWD System Model

Consider a MH relaying system, where a source (R_0) intends to communicate with a destination (R_{k+1}) as shown in Figure 4.1. Both the high shadowing and the severe path loss effect are responsible for the absence of the direct link between R_0 and R_{k+1}. As such,

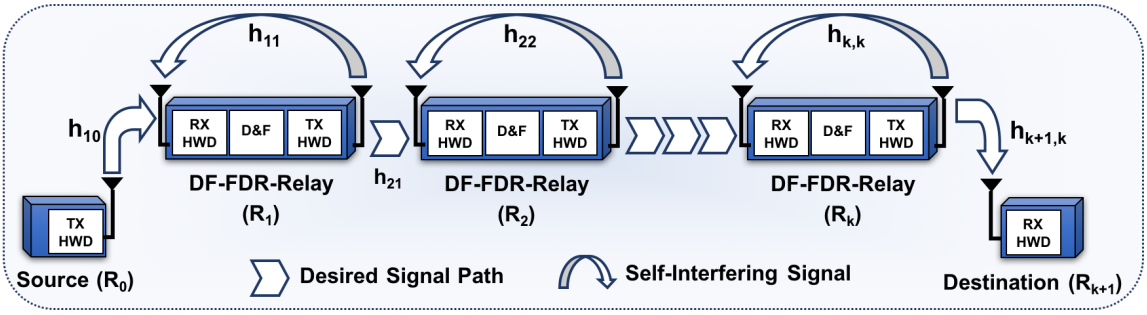


Figure 4.1: FDR System under HWD and RSI.

multiple FDRs ($R_1 - R_k$) operate as intermediate nodes to facilitate the end-to-end communication by expanding the coverage area with a full transmission rate. The limited power budget of the participating relays renders negligible inter-relay interference. All relays operate in a DF relaying strategy and the SI at the relays can only be canceled partially, thus resulting in RSI as depicted in Figure 4.1. Furthermore, various HWDs at the transmitter and receiver RF branches can drastically degrade the overall system performance. We model the MH-DF-FDR system suffering from transceiver HWDs.

In the proposed MH-FDR with the DF relaying strategy under HWD, the mth node R_m transmits an IGS signal $x_m \sim \mathcal{CN}(0, \sigma_m^2, \tilde{\sigma}_m^2)$ to the nth node R_n in one hop. Pertaining to the FDR operation, the received signal at the relay node R_n suffers from RSI h_{nn} in addition to the aggregate effect of transceiver distortions for both m-n link, η_{nm}, and n-n link, η_{nn}. The aggregate distortion $\eta_{nm} = \eta_{\text{tx}}^m + \eta_{\text{rx}}^n$ includes the transmit distortions from the mth node and the receiver distortions from the nth node in x_m, which is transmitted from the mth node to n^{th} node. Similarly, $\eta_{nn} = \eta_{\text{tx}}^n + \eta_{\text{rx}}^n$ includes the transmit distortions as well as the receiver distortions from the nth node in x_n, which is transmitted from and looped-back to n^{th} node owing to its full-duplex operation. The aggregate transceiver distortions exhibit the same statistical characteristics as detailed in (3.2). The signal is transmitted from R_0 to R_{k+1} in a sequential order. Therefore, the generalized received signal at n^{th} receiver is given by

$$y_n = \sqrt{p_m} h_{nm} \left(x_m + \eta_{nm}\right) + \sqrt{p_n} h_{nn} \left(x_n + \eta_{nn}\right) + z_n, \tag{4.1}$$

where $n = 1, 2, \ldots, k+1$ denotes the receiver nodes and $m = n - 1$ represents the sequential transmission nodes, thus $m = 0, 1, \ldots, k$. In addition, p_m is the transmit power of the mth node taken from a limited power budget, h_{nm} is the flat-fading channel of the m-n link[1] and $z_n \sim \mathcal{CN}\left(0, \sigma_z^2, 0\right)$ is independent identically distributed (iid) AWGN at the nth receiver node. The receiver at the relay node decodes the transmitted signal using single user decoder, then encodes it from IGS code-book as $x_n \sim \mathcal{CN}(0, \sigma_n^2, \tilde{\sigma}_n^2)$ for further transmission. The same transmitted signal causes self-interference in the FDR transmission mode through the h_{nn} link. Measurement-driven experimental studies [269–271] have shown that, after undergoing all possible isolation/cancellation techniques and assuming the perfect cancellation of slowly-varying line-of-sight path, the residual interference can be well characterized as a flat-fading channel [115, 179, 272–275] and the RSI channel h_{nn} can be modeled as a zero mean symmetric complex Gaussian random variable ([276–280], and references therein).[2] It is worthy noting that the self-interference link does not exist at the destination node R_{k+1} as there is no further transmission and self-interference. Additionally, the transceiver HWD of the m-n link, i.e., η_{nm}, is assumed to show the statistical characteristics: $\eta_{nm} \sim \mathcal{CN}(0, \kappa_{nm}, \tilde{\kappa}_{nm})$.

4.3 Achievable Rates

The overall end-to-end achievable rate of the MH-DF-FDR system, R_T, is given as

$$R_T = \min_n \left\{R_{nm}\right\}; \quad m = n - 1, \tag{4.2}$$

[1]It is important to mention that the narrow band assumption has been adopted to simplify the presentation of the mathematical modeling and optimization analysis. However, the same contribution can be straightforwardly extended to the multipath channel scenario such as OFDM for each sub-channel / sub-band after incorporating the effective inter-carrier interference. The extension to OFDM does not affect the optimization framework as the various transmission streams do not share the common resource budget.

[2]Note that the flat-fading assumption and Gaussian RSI model do not compromise the insights of the analysis that follows. The same conclusions can be reached if more complex RSI models are used.

where R_{nm} is the achievable rate of the m-n link in bits/sec. In our work, we deal with the RSI and HWD as interference terms, thus R_{nm} considering the IGS transmission scheme and asymmetric HWD terms can be obtained as [81]

$$R_{nm} = \frac{B_{nm}}{2}\log_2 \frac{\sigma_{y_n}^4 - \left|\tilde{\sigma}_{y_n}^2\right|^2}{\sigma_{I_n}^4 - \left|\tilde{\sigma}_{I_n}^2\right|^2}, \tag{4.3}$$

where B_{nm} is the bandwidth of the m-n link. In addition, $\sigma_{y_n}^2$ and $\tilde{\sigma}_{y_n}^2$ are the variance and the pseudo-variance of the received signal at the nth node, respectively. Also, $\sigma_{I_n}^2$ and $\tilde{\sigma}_{I_n}^2$ are the variance and the pseudo-variance of the self-interference signal plus noise at the receiver end. Therefore, the achievable rate of the link between the mth transmit node and the nth receiver node, R_{nm} can be expressed as

$$R_{nm} = \frac{B_{nm}}{2}\log_2 \frac{\alpha_{nm} - \left|p_m h_{nm}^2 \left(\tilde{\sigma}_m^2 + \tilde{\kappa}_{nm}\right) + p_n h_{nn}^2 (\tilde{\sigma}_n^2 + \tilde{\kappa}_{nn})\right|^2}{\beta_{nm} - \left|p_m h_{nm}^2 \tilde{\kappa}_{nm} + p_n h_{nn}^2 \left(\tilde{\sigma}_n^2 + \tilde{\kappa}_{nn}\right)\right|^2}, \tag{4.4}$$

where β_{nm} and α_{nm} are defined, respectively as,

$$\beta_{nm} = \left(p_m |h_{nm}|^2 \kappa_{nm} + p_n |h_{nn}|^2 \left(\sigma_n^2 + \kappa_{nn}\right) + \sigma_z^2\right)^2, \tag{4.5}$$

$$\alpha_{nm} = \left(p_m |h_{nm}|^2 \left(\sigma_m^2 + \kappa_{nm}\right) + p_n |h_{nn}|^2 \left(\sigma_n^2 + \kappa_{nn}\right) + \sigma_z^2\right)^2. \tag{4.6}$$

According to (4.4), R_{nm} is a function of the pseudo-variance $\tilde{\sigma}_m^2$ of the transmitted signal and the pseudo-variance $\tilde{\sigma}_n^2$ of the self-interfering signal, which provides additional degrees of freedom to mitigate the asymmetric interference caused by the HWD as well as RSI. However, the achievable rate of the last hop between node R_k and R_{k+1} is only a function of $\tilde{\sigma}_k^2$ due to the absence of the self-interfering link.

4.4 HWD- and RSI-Aware Signaling Design

In this section, we design the transmit signals for all transmitting nodes, to maximize R_T under HWD and RSI for the adopted MH-DF-FDR system. The main goal of the system

design is to optimize the statistical asymmetric characteristics of the transmitted signals to maximize R_T in (4.2) as follows

$$\begin{aligned} \textbf{4-P1}: \quad & \text{maximize} \quad \min_n \{R_{nm}(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2)\} \\ & \text{subject to} \quad 0 \leq |\tilde{\sigma}_m^2| \leq \sigma_m^2, \ \forall m, \end{aligned} \tag{4.7}$$

where the constraint is adopted to confirm the bounds on circularity coefficient, for the pseudo-variances of all transmitting signals from nodes R_0 to R_k. Throughout the rest of this work, we solve **4-P1** by proposing two optimization frameworks with different implementation, complexity and performance. Firstly, we propose an efficient joint optimization scheme which requires a centralized processing framework. Later, we present a distributed framework which optimizes the transmission parameters for a cluster of nodes.

4.4.1 Joint Optimization

In the joint optimization setup, we assume having a central node that gathers the channel state information (CSI), HWD and RSI levels from all the participating nodes to jointly optimize their signal parameters in order to maximize the overall end-to-end achievable rate. After processing the gathered information, the central node distributes the optimal transmission parameters for data transmission.

To solve **4-P1**, we write all optimization parameters (transmit pseudo-variances) in a vector form. The total number of the optimization parameters depend on the number of participating relays and the corresponding communication hops. Considering k intermediate relays between the source and the destination results in $k+1$ hops and hence $k+1$ transmitting nodes. Therefore, we need to optimize $k+1$ complex pseudo-variance variables or $2(k+1)$ real transmit pseudo-variance variables. Thus, we define a real vector $\mathbf{s}$ that captures the real and imaginary variables as

$$\mathbf{s} = \begin{bmatrix} \Re\{\tilde{\sigma}_0^2\} & \Im\{\tilde{\sigma}_0^2\} & \Re\{\tilde{\sigma}_1^2\} & \Im\{\tilde{\sigma}_1^2\} \ldots \Re\{\tilde{\sigma}_k^2\} & \Im\{\tilde{\sigma}_k^2\} \end{bmatrix}^T \tag{4.8}$$

Then, we express the link rate between nodes $R_m - R_n$ in (4.4) as a function of the vector $\mathbf{s}$ carrying optimization variables:

$$R_{nm}(\mathbf{s}) = \frac{B_{nm}}{2}\log_2\frac{\alpha_{nm} - \left|\mathbf{s}^T\mathbf{u}_{nm} + v_{nm}\right|^2}{\beta_{nm} - \left|\mathbf{s}^T\mathbf{w}_{nm} + v_{nm}\right|^2}, \tag{4.9}$$

where $\mathbf{u}_{nm}$ and $\mathbf{w}_{nm}$ are defined, respectively, as follows:

$$\mathbf{u}_{nm} = \langle\langle\left[p_m h_{\mathrm{nm}}^2 \;\; jp_m h_{\mathrm{nm}}^2 \;\; p_n h_{\mathrm{nn}}^2 \;\; jp_n h_{\mathrm{nn}}^2\right]^T\rangle_{(\mathrm{Az},2k-2)}\rangle_{(\mathrm{Ds},2m)}, \tag{4.10}$$

$$\mathbf{w}_{nm} = \langle\langle\left[p_n h_{\mathrm{nn}}^2 \;\; jp_n h_{\mathrm{nn}}^2\right]^T\rangle_{(\mathrm{Az},2k)}\rangle_{(\mathrm{Ds},2n)}. \tag{4.11}$$

Furthermore, the complex scalar v_{nm} is defined as a function of transmit power, CSI and HWD as,

$$v_{nm} = p_m h_{nm}^2 \tilde{\kappa}_{nm} + p_n h_{nn}^2 \tilde{\kappa}_{nn}. \tag{4.12}$$

Let us denote the numerator and denominator of the fraction in (4.9) as $\mathcal{N}_{nm}(\mathbf{s})$ and $\mathcal{D}_{nm}(\mathbf{s})$ respectively. Thus, $\mathcal{N}_{nm}(\mathbf{s})$ can be written in a simplified form as

$$\mathcal{N}_{nm}(\mathbf{s}) = \alpha_{nm} - \mathbf{s}^T\mathbf{U}_{nm}\mathbf{s} - \mathbf{s}^T\mathbf{c}_{nm} - |v_{nm}|^2, \tag{4.13}$$

where $\mathbf{U}_{nm}$ is the outer product of $\mathbf{u}_{nm}$, i.e., $\mathbf{U}_{nm} = \mathbf{u}_{nm}\mathbf{u}_{nm}^H$ and $\mathbf{c}_{nm} = v_{nm}^*\mathbf{u}_{nm} + v_{nm}\mathbf{u}_{nm}^*$. The positive semi-definite characteristic of $\mathbf{U}_{nm}$ renders the concavity characteristic of $\mathcal{N}_{nm}$ in $\mathbf{s}$. Analogously, $\mathcal{D}_{nm}(\mathbf{s})$ can be rewritten as

$$\mathcal{D}_{nm}(\mathbf{s}) = \beta_{nm} - \mathbf{s}^T\mathbf{W}_{nm}\mathbf{s} - \mathbf{s}^T\mathbf{d}_{nm} - |v_{nm}|^2, \tag{4.14}$$

with $\mathbf{d}_{nm} = v_{nm}^*\mathbf{w}_{nm} + v_{nm}\mathbf{w}_{nm}^*$ and $\mathbf{W}_{nm} = \mathbf{w}_{nm}\mathbf{w}_{nm}^H$, which implies the concavity of $\mathcal{D}_{nm}$ in $\mathbf{s}$ thanks to the positive semi-definite properties of $\mathbf{W}_{nm}$. Therefore, using the

aforementioned representations, optimization problem $\mathbf{P1}$ can be equivalently written as

$$\textbf{4-P2}: \quad \text{maximize} \quad \min_{n} \left\{ \frac{B_{nm}}{2} \log_2 \frac{\mathcal{N}_{nm}(\mathbf{s})}{\mathcal{D}_{nm}(\mathbf{s})} \right\} \tag{4.15}$$
$$\text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \quad \forall m,$$

where $\mathbf{P}_m = \text{diag}\left\{ \langle \langle [1 \;\; 1]^T \rangle_{(\text{Az},2k)} \rangle_{(\text{Ds},2m)} \right\}$. The max-min fractional problem in **4-P2** can be efficiently solved by exploiting the properties of the logarithmic function. In addition, we can transform the secondary minimization problem into the following maximization problem

$$\textbf{4-P3}: \text{ maximize } -\max_{n} \left\{ \frac{B_{nm}}{2} \log_2 \mathcal{N}_{nm}(\mathbf{s}) + \frac{B_{nm}}{2} \log_2 \mathcal{D}_{nm}(\mathbf{s}) \right\} \tag{4.16}$$
$$\text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2 \quad , \forall m.$$

Evidently, $\log_2 \mathcal{N}_{nm}(\mathbf{s})$ and $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ are concave functions owing to the the positive-concave nature of $\mathcal{N}_{nm}(\mathbf{s})$ and $\mathcal{D}_{nm}(\mathbf{s})$, respectively. Therefore, the subtractive form of the objective function in **4-P3** is universally known as difference of convex (DC)-programming (difference of concave) and cannot be handled straightforwardly. Thus, we employ sequential convex programming (SCP) to efficiently transform **4-P3** into iterative convex problems which can be optimally solved in each iteration [281]. In this approach, we use the affine Taylor series approximation of the function $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ to yield a convex objective function. The first-order Taylor series expansion of the function $g(x)$ at point $x^{(k)}$ is given by

$$\hat{g}(x, x^{(k)}) = g(x^{(k)}) + \nabla g(x^{(k)})^T \left(x - x^{(k)} \right). \tag{4.17}$$

Thus, by employing the same expansion for the concave function $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ gives an

affine approximation at $\mathbf{s}^{(i)}$ as

$$\breve{\mathcal{D}}_{nm}\left(\mathbf{s},\mathbf{s}^{(i)}\right)=\frac{B_{nm}}{2\ln 2}\left(\ln\mathcal{D}_{nm}\left(\mathbf{s}^{(i)}\right)+\frac{\nabla^T\mathcal{D}_{nm}\left(\mathbf{s}^{(i)}\right)}{\mathcal{D}_{nm}\left(\mathbf{s}^{(i)}\right)}\left(\mathbf{s}-\mathbf{s}^{(i)}\right)\right), \tag{4.18}$$

where $\nabla\mathcal{D}_{nm}\left(\mathbf{s}^{(i)}\right)=-(\mathbf{W}_{nm}+\mathbf{W}_{nm}^T)\mathbf{s}^{(i)}-\mathbf{d}_{nm}$ is the gradient of $\mathcal{D}_{nm}(\mathbf{s})$ evaluated at $\mathbf{s}^{(i)}$. It is important to note that no trust region is required as $\breve{\mathcal{D}}_{nm}(\mathbf{s},\mathbf{s}^{(i)})\leq\mathcal{D}_{nm}(\mathbf{s}^{(i)})$ [281]. Thus, **4-P3** can be convexified using the aforementioned procedure giving the following problem that needs to be solved successively while updating $\mathbf{s}^{(i)}$,

$$\begin{aligned}\textbf{4-P4}:\ &\text{maximize}\ -\max_n\left\{-\frac{B_{nm}}{2}\log_2\mathcal{N}_{nm}\left(\mathbf{s}\right)+\breve{\mathcal{D}}_{nm}(\mathbf{s},\mathbf{s}^{(i)})\right\}\\ &\text{subject to}\quad 0\leq\mathbf{s}^T\mathbf{P}_m\mathbf{s}\leq\sigma_m^2\quad,\forall m.\end{aligned} \tag{4.19}$$

Given, the convex objective function and convexity preservation by the point-wise maximization, the primary maximization problem can be equivalently written as follows [282]

$$\begin{aligned}\textbf{4-P5}:\ &\text{minimize}\ \max_n\left\{-\frac{B_{nm}}{2}\log_2\mathcal{N}_{nm}\left(\mathbf{s}\right)+\breve{\mathcal{D}}_{nm}(\mathbf{s},\mathbf{s}^{(i)})\right\}\\ &\text{subject to}\quad 0\leq\mathbf{s}^T\mathbf{P}_m\mathbf{s}\leq\sigma_m^2\quad,\forall m.\end{aligned} \tag{4.20}$$

The formulated problem **4-P5** yields an optimal solution for a given $\mathbf{s}^{(i)}$ as it is the minimization of a convex function pertaining to the convexity preservation by point-wise maximization. One way to solve **4-P5** is by introducing an auxiliary variable τ in order to capture the point-wise maximization problem as follows

$$\begin{aligned}\textbf{4-P6}:\ &\text{minimize}\ \tau\\ &\text{subject to}\quad -\frac{B_{nm}}{2}\log_2\mathcal{N}_{nm}\left(\mathbf{s}\right)+\breve{\mathcal{D}}_{nm}(\mathbf{s},\mathbf{s}^{(i)})\leq\tau,\ \forall m\backslash\{k\}\\ &\qquad -\frac{B_{nm}}{2}\log_2\mathcal{N}_{nm}\left(\mathbf{s}\right)+\frac{B_{nm}}{2}\log_2\mathcal{D}_{nm}(\mathbf{s})\leq\tau,\ m=k\\ &\qquad 0\leq\mathbf{s}^T\mathbf{P}_m\mathbf{s}\leq\sigma_m^2\quad,\forall m.\end{aligned} \tag{4.21}$$

Therefore, the solution of the proposed optimization problem reduces to solving **4-P6** iteratively using the SCP as discussed in the following subsection.

Centralized Joint Algorithm

To solve **4-P1** or equivalently **4-P2**, we use the SCP that deals with a convexified version of the difference of convex/concave problem **4-P3** in each iteration developing Algorithm 1. The proposed algorithm starts with a feasible starting point $\mathbf{s}^{(i)}$ to find the affine approximation of $\frac{1}{2}\log_2\mathcal{D}_{nm}(\mathbf{s})$ for all transmitters, i.e., R_m and $m < k$. The last node does not suffer from self-interference, which renders constant $\frac{1}{2}\log_2\mathcal{D}_{nm}(\mathbf{s})$ at $m = k$. Then, the affine approximation $\breve{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)})$ is used to optimally solve a quadratic contraint linear programming (QCLP) problem defined in **4-P6** using any available convex optimization solvers such as CVX-MATLAB employing interior-point method. Next, the solution of the QCLP problem is used to update $\mathbf{s}$ obtaining $\mathbf{s}^{(i+1)}$, which is the starting point for next iteration. The algorithm solves successive convex QCLPs and updates the solution values in each iteration until the desired stopping condition is met. The stopping convergence criterion is when the absolute difference between two successive solutions is less than a predefined threshold δ.

Algorithm 1 Sequential Convex Programming

Initialize $i \leftarrow 0$, $\epsilon \leftarrow \infty$ and **Set** tolerance δ
Choose feasible starting point $\mathbf{s}^{(i)}$
while $\epsilon \geq \delta$ **do**
    Evaluate $\breve{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)})\ \forall m \backslash \{k\}$
    Solve **4-P6** and obtain $\mathbf{s}$ using $\mathbf{s}^{(i)}$
    $\mathbf{s}^{(i+1)} \leftarrow \mathbf{s}$
    Update $\epsilon \leftarrow |\mathbf{s}^{i+1} - \mathbf{s}^{i}|$
    $i \leftarrow i + 1$
end while
$\mathbf{s}^{*} \leftarrow \mathbf{s}^{i+1}$
$R_{\mathrm{T}} = \min_{n} \{R_{nm}(\mathbf{s}^{*})\}$

The obtained solution vector $\mathbf{s}$ contains the real and imaginary components of the trans-

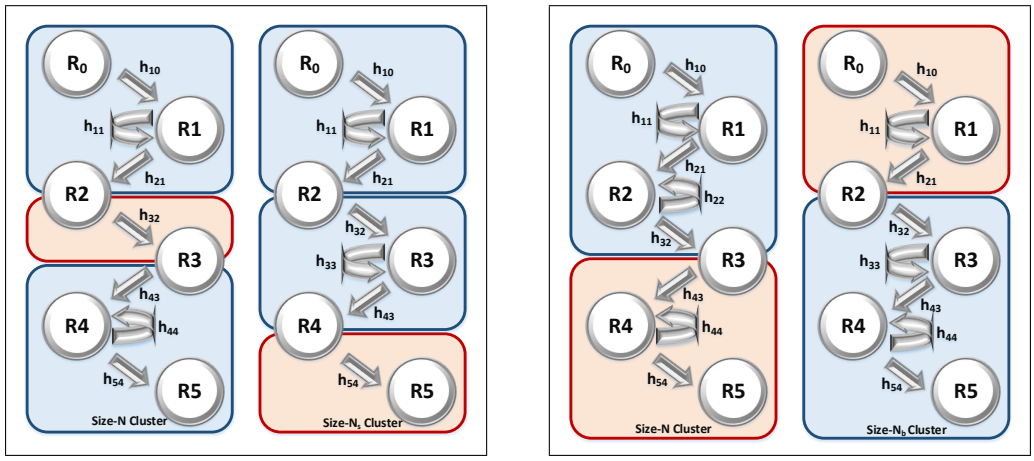
mit pseudo-variances for all transmitting nodes. Thus, the maximized end-to-end achievable rate R_{T} can be computed using (4.2) and (4.9). Despite of the efficacy of the joint-optimization, it requires a centralized network realization, where all nodes share their information with a central node prior to the transmission. Then, the signal design is carried out at the central node in order to update the relaying nodes with the optimized signal parameters.

4.4.2 Distributed Optimization

The implementation of the joint optimization requires a centralized network realization. However, the centralized realization may not be suitable in different scenarios due to:

- **Time Delays:** In the proposed MH-FDR system, we employ relay(s) to establish a communication link between any two distant nodes in the absence of a direct link. The same communication links are utilized to transmit CSI from each node to the central node that performs the joint optimization of the transmit parameters. Therefore, exchanging CSI results in notable delay depending on number of intermediate nodes. Moreover, the CSI data is subjected to error propagation. Then, the central node processes the received information, performs joint optimization and sends the optimized parameters back to each node. Furthermore, the processing/computational delay at the central node also increases with higher number of intermediate nodes. Also, for large networks, the round trip time and computational delay may exceed channel coherence time, which results in degraded performance. On the other hand, the round trip in each cluster decreases significantly compared with the centralized configuration due to the reduced cluster size.

- **Communication Overhead:** For a network involving large number of relays to establish a communication link between two distant nodes, pilot signals are sent between intermediate nodes to estimate the channels, then the CSI is sent to the central



(a) C+1 Clusters with One Size-N_s Cluster (b) C Clusters with One Size-N_b Cluster

Figure 4.2: Different Sized Clusters with Their Two Possible Placements

nodes via the relays. Thus, the increase in number of relays will increase the cost of communication overhead. However, it decreases in the distributed realization as the entire network is divided into smaller clusters with fewer number of relays where each cluster locally performs the optimization process. This local optimization procedure can be carried out at any node within cluster. However, the centrally located node and the computationally competent node are two preferred choices in order to reduce round trip delays and computing time, respectively. The selected node only requires the CSI of the links within that cluster for local optimization and this CSI can be communicated through control channels to save the communication overhead on the information channels [283].

Therefore, to address the practical limitations of the centralized realization, reduce the communication/processing overhead on one central node and avoid the excessive time delays, the distributed optimization framework is proposed.

Distributed Algorithm

In the distributed framework, we aim at grouping the neighboring nodes into clusters of equal sizes, having N nodes and $N-1$ hops, in order to equally distribute the computational

load among all clusters. Each cluster acts as a whole system and runs the joint-optimization algorithm locally to maximize the end-to-end achievable rate within that cluster. However, the common scenario arises when the total number of nodes is not an integer multiple of the desired cluster size, thus they cannot be equally grouped into the clusters of same size. As such, consider a system with k-relays between the source and the destination. Then, we propose to divide this system into C equal clusters of size N, where each cluster has N nodes, where C is found to be expressed as

$$C = \left\lfloor \frac{k+1}{N-1} \right\rfloor . \tag{4.22}$$

As a result, this leads to N_{s} unassigned nodes where N_{s} is smaller than N and is found to be

$$N_{\mathrm{s}} = k - \left\lfloor \frac{k+1}{N-1} \right\rfloor (N-1) + 2. \tag{4.23}$$

In the following, we propose two grouping configurations to deal with the N_{s} nodes:

- Group the N_{s} nodes to form an additional small-cluster. Thus, the system will have a total of $C+1$ clusters.
- Group the remaining N_{s} nodes with a pre-defined cluster. Thus, the number of the nodes in the new re-defined cluster becomes N_{b} and is found to be

$$N_{\mathrm{b}} = N + N_{\mathrm{s}} - 1. \tag{4.24}$$

The performance of the distributed realization approach can be improved by smartly choosing the cluster size and their placement in the network. A heuristic approach is to locate this odd sized cluster either at the transmitter side involving source node or at the receiver end involving destination node. However, it is important to note that the cluster formulation yields sub-optimal solution owing to the neglected RSI at the end nodes. Therefore, the big odd cluster is expected to outperform the small odd cluster and even the regular

cluster pertaining to the inclusion of RSI of intermediate nodes. Thus, a good suggestion is to place the big odd cluster at the weakest channel location as it will contain the bottleneck link dictating the end-to-end achievable rate. On the other hand, the grouping scenario with small cluster is preferred to be placed at the strongest channel location so that its compromised transmission parameters are not detrimental for the overall end-to-end achievable rate. The proposed suggestions have been supported and validated using simulation results in Section 4.5.

For comprehensive illustration, consider the distributed network scheme presented in Figure 4.3. We aim to divide this system having $k = 4$ relays into $N = 3$ sized clusters, which is clearly not possible to have equally sized groups. Therefore, one approach is to group the remaining $N_s = 2$ nodes in another cluster which can either be placed in the middle or at the destination as shown in Figure 4.2a. Another possibility is to place this at the transmitting end. As discussed, we prefer to formulate this cluster at the $\max\left(|h_{10}|, |h_{32}|, |h_{54}|\right)$. So that the compromised transmission parameters due to neglected RSI do not dominate the system performance, which is dictated by the minimum rate link. Another approach is to group the remaining $N_s = 2$ nodes with another cluster formulating a $N_b = 4$ sized-cluster which can either be placed at the start or at the end as shown in Figure 4.2b. Again, the preferred approach is to place this big cluster around transmitter if the $\min\left(|h_{10}|, |h_{21}|\right) \leq \min\left(|h_{43}|, |h_{54}|\right)$, else place it at the end to accommodate destination node. Thus, bigger cluster will host the bottleneck link and the conforming transmission parameters, owing to the inclusion of RSI at larger number of intermediate nodes, will dictate overall system throughput.

It is important to highlight that the above mentioned approach requires a rough idea of the strong and/or weak channel gains throughout the span of the system to practically locate the odd cluster. The small cluster and big cluster scenarios are separately dealt in Algorithm 2 and 3 respectively. As mentioned earlier, the small cluster size leads to inefficient solution, therefore we place this small cluster at a location which offers relatively

higher channel gain. This is advantageous as the cluster with better channel conditions does not dictate the overall system achievable rate. In the same way, we place big cluster at a weaker channel location for more accurate solution parameters as it will dictate the overall system rate. The following two algorithms are enumerated to practically implement the distributed networking approach.

Algorithm 2 Distributed Algorithm with $C+1$ clusters

Choose Cluster size N
Compute Smaller cluster size N_{s} using (4.23).
Formulate the small cluster at strongest channel location.
Initialize Cluster counter $j \leftarrow 1$
while $j \leq C+1$ **do**
    if Before N_{s} **then**
        Solve **4-P2** for $m = (j-1)(N-1), \ldots, (j)(N-1)-1$.
    end if
    if At N_{s} **then**
        Solve **4-P2** for $m = (j-1)(N-1), \ldots, (j)(N-1)-(N-N_{\mathrm{s}})-1$.
    end if
    if Beyond N_{s} **then**
        Solve **4-P2** for $m = (j-1)(N-1)-(N-N_{\mathrm{s}}), \ldots, (j-1)(N-1)+N_{\mathrm{s}}-2$.
    end if
    Distribute the locally optimized transmit parameters within cluster.
    $j \leftarrow j+1$
end while

In Algorithm 2, we begin by choosing a cluster size N and divide the entire system into C clusters each having N nodes. The remaining nodes are grouped into a small cluster of size N_{s} and are chosen to have relatively high channel gain. Then, Algorithm 2 groups the clusters into three main groups with a total of $C+1$ clusters. The first group consists of all clusters before the N_{s} cluster. The second group comprises only the N_{s} cluster. Finally, the last group comprises of all the clusters after the N_{s} cluster. The physical intuition behind the three mentioned groups in the algorithm is to provide the appropriate indexing of the involved transmitting nodes in each cluster for their transmission parameter optimization by solving **4-P2**. The transmit parameters are locally designed in an individual cluster by solving optimization problem **4-P2** for the corresponding range of m-transmitters. It is

worth mentioning that the problem **4-P2** is solved by iteratively solving the problem **4-P6** using Algorithm 1. The optimized variables are then distributed to each node within the cluster and the process continues.

Algorithm 3 Distributed Algorithm with C clusters

Choose Cluster size N
Compute bigger cluster size N_{b} using (4.24).
Formulate the big cluster at the weakest channel location.
Initialize Cluster counter $j \leftarrow 1$
while $j \leq C$ **do**
    if Before N_{b} **then**
        Solve **4-P2** for $m = (j-1)(N-1), \ldots, (j)(N-1)-1$.
    end if
    if At N_{b} **then**
        Solve **4-P2** for $m = (j-1)(N-1), \ldots, (j-1)(N-1)+N_{\mathrm{b}}-2$.
    end if
    if Beyond N_{b} **then**
        Solve **4-P2** for $m = (j-1)(N-1)+N_{\mathrm{b}}-N, \ldots, (j-1)(N-1)+N_{\mathrm{b}}-2$.
    end if
    Distribute the locally optimized transmit parameters within cluster.
    $j \leftarrow j+1$
end while

Algorithm 3 realizes the C cluster scenario, where the N_{s} cluster is merged with a regular cluster to formulate N_{b} sized cluster. Similar to Algorithm 2, Algorithm 3 deals with three groups of clusters by scanning the C participating clusters. In the absence of CSI, the remaining cluster can be randomly placed anywhere in the network. Although the division into clusters reduces the round trip delays, communication overhead, and computational complexity, it does so at the expense of deviated solution parameters. The distributed solution ignores the drastic effects of RSI at the destination node in each cluster. Thus, the solution of the distributed setup is expected to deviate from the joint optimization one depending on the cluster size. The bigger the cluster size, the closer the distributed optimization solution to the joint optimization one, as it accommodates RSI of the intermediate nodes and vice versa. In addition, the distributed optimization focuses on the minimum rate performance of local cluster without considering other links, which may cause devi-

ation from the joint optimization solution. This portrays a trade off between distributed computational load and optimal performance.

4.4.3 Complexity Analysis

The computational complexity analysis of the proposed algorithms is carried out in the sequel. The proposed joint and distributed transmit optimization algorithms depict a trade-off between performance and the computational complexity assuming negligible round-trip delays and the high end computational capabilities of the central node. The computational complexity of these strategies can be expressed as follows:

- Joint Approach: $O\left(I_1^{\mathrm{SCP}}\left(N_1^2 k + \alpha \max\left(N_1^3, N_1^2 M_1, F_1\right)\right)\right)$
- Distributed Approach: $O\left(I_2^{\mathrm{SCP}}\left(N_2^2\left(\Gamma - 2\right) + \alpha \max\left(N_2^3, N_2^2 M_2, F_2\right)\right)\right)$

where I_x^{SCP} is the number of SCP iterations before convergence ($x = 1$ for joint algorithm, $x = 2$ for distributed algorithm with regular clusters of size N, $x = 2a$ for distributed algorithm with small cluster of size N_s and $x = 2b$ for distributed algorithm with big cluster of size N_b). In addition, the total number of optimization variables in joint and distributed optimization algorithms are given as $N_1 = 2k + 2$ and $N_2 = 2\Gamma - 2$, respectively. Where k is the total number of relays and Γ is the number of nodes in a given distributed cluster. As for α, it is assumed to be between 10 and 100 for the interior point method [282]. Also, M_1 and M_2 are the number of inequalities representing constraints of joint and distributed optimization problems, respectively, defined as $M_1 = 2(k + 1)$ and $M_2 = 2(\Gamma - 1)$, respectively.

Furthermore, $F_1 = N_1(1 + M_1) + N_1^2 M_1$ and $F_2 = N_2(1 + M_2) + N_2^2 M_2$ are the costs of evaluating the first and second derivatives of the objective and constraint functions in joint and distributed algorithms, respectively. The computational complexities of joint and distributed algorithms are further simplified in Appendix A and Appendix B, respectively. The simplified complexity analysis is presented in Table 4.1 for joint-algorithm with k

Table 4.1: Complexity Analysis of the proposed algorithms

Algorithm	Size	Complexity
Joint Algorithm	k-relays	$O\left(\alpha I_1^{\mathrm{SCP}} k^3\right)$
Distributed Algorithm	N-sized Clusters	$O\left(\alpha I_2^{\mathrm{SCP}} N^3\right)$
	N_s-sized Cluster	$O\left(\alpha I_{2a}^{\mathrm{SCP}} {N_s}^3\right)$
	N_b-sized Cluster	$O\left(\alpha I_{2b}^{\mathrm{SCP}} {N_b}^3\right)$

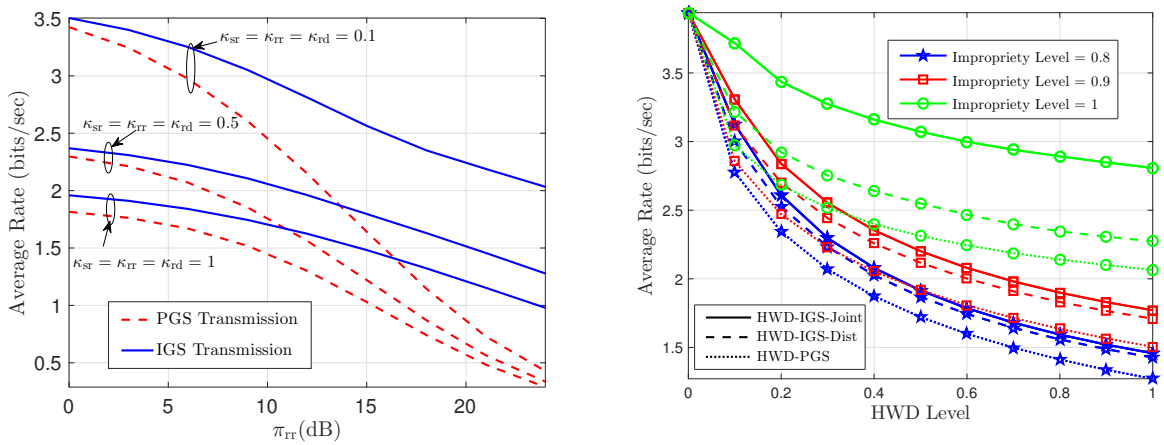
relays and distributed-algorithm with $\Gamma = N$ sized regular cluster, $\Gamma = N_s$ sized small cluster, and $\Gamma = N_b$ sized big cluster.

Evidently, the trend $1 \leq N_s \leq N \leq N_b \leq k$ follows from (4.23) and (4.24), depicts the least complexity for the N_s-cluster distributed algorithm relative to the N-cluster distributed algorithm. Moreover, N_b-cluster distributed algorithm depicts the most complexity of all distributed algorithms. However, the relation $\Gamma = (k + C + 1)/C$ demonstrates that $\Gamma << k$. Thus, validating that the computational complexity of the distributed algorithm in a cluster of size Γ is far less than that of the joint optimization algorithm involving k relays.

4.5 Numerical and Simulation Results

In this section, we quantify the gain reaped by employing IGS transmission scheme in place of PGS scheme for the adopted MH-DF-FDR system suffering from RSI and HWD. Besides studying the degradation effect caused by these interferences, we also investigate the impact of multiple relays and the cluster size in distributed optimization approach on the overall system performance. In addition, we compare the performance of the two proposed algorithm for distributed approach with N_{b} and N_{s} cluster sizes and their respective placement in the system network.

As for the simulation parameters, we assume unity transmit power from all transmitters, 1 Hz bandwidth of all involved links, noise variance $\sigma_z^2 = 1$ and HWD level κ_{nm} = 1 along with the impropriety level $|\tilde{\kappa}_{nm}| = 0.9$ at all participating nodes. Moreover, h_{nm} of the m-n link is modeled as a slowly-varying Rayleigh flat-fading channel with $h_{nm} \sim$



(a) Average rate vs. RSI $\pi_{\rm rr}$ for PGS and IGS schemes.

(b) Average rate vs. HWD level.

Figure 4.3: Impact of RSI and HWD on System Performance

$\mathcal{CN}(0, \pi_{nm}, 0)$ [3]. Furthermore, we assume 25dB SNR for m-n link and the RSI of 10dB unless otherwise specified.

4.5.1 Effect of RSI and HWD

First, we study the performance degradation caused by the RSI and HWD in Figure 4.3a and 4.3b respectively. In the first simulation example, we study the advantage of employing IGS in suppressing the RSI effect on the average achievable end-to-end rate for a dual-hop FDR system as shown in Figure 4.3a. Average rate is observed at various RSI gains $\pi_{\rm rr}$ ranging from 0dB to 25dB for three different HWD levels. For simplicity, we assume equal HWD and impropriety levels at source, relay and destination as presented by $\kappa_{\rm sr} = \kappa_{\rm rr} = \kappa_{\rm rd} = 0.1, 0.5 \,\&\, 1$ and $|\tilde{\kappa}_{\rm sr}| = |\tilde{\kappa}_{\rm rr}| = |\tilde{\kappa}_{\rm rd}| = 0.9$, respectively. Evidently, the increasing self-interference severely degrades the achievable rate performance. In addition, increasing HWD $\kappa_{\rm nm}$ from 0.1 to 1 also deteriorates the system performance.

s

Interestingly, the proposed IGS scheme is capable of providing significant performance

[3] It is important to highlight that the presented technical contribution holds true for any form of fading including Rayleigh, Ricean or Nakagami. It is because of the fact that the derived framework does not depend on the statistical characteristics of the given channel model.

enhancement at lower-residual HWD levels, assuming effective joint compensation of HWD and SI, for the entire range of RSI levels as shown in Figure 4.3a. Similar results have been demonstrated in [179] to emphasize the significance of employing IGS transmission in an attempt to alleviate the RSI adverse effect in FDR system considering ideal transceivers. It is important to highlight that the PGS scheme undergoes saturation at higher RSI levels irrespective of the HWD level. Thus, it can be safely concluded that the RSI dominates in degrading the rate performance for PGS. On the other hand, IGS scheme efficiently mitigates the RSI impact and reduces the rate degradation. The best relative improvement is achieved at high RSI and low HWD levels.

Secondly, we study the degradation effect of HWD on the average achievable rate for various impropriety levels and the relative performance gains obtained by the proposed joint IGS scheme and a sub-optimal less-complex distributed IGS approach over the conventional PGS scheme in Figure 4.3b. We assume a MH system incorporating 3 relays between the source and destination with favorable channel gains of h_{nm}= 30dB at each participating link. Clearly, the rate performance drastically deteriorates with increasing HWD variance, from 0 indicating the ideal hardware to 1 indicating the maximal impairment level[4] even at very good channels gain. Moreover, joint IGS optimally mitigates the HWD impact relative to PGS scheme at various impropriety levels. We aim to quantify the practical impropriety levels through exact mapping of HWDs aggregating from I/Q mixers, PA, and LNAs in future measurement phase. In addition, distributed approach with cluster size $N = 2$ locally optimizes the transmission parameters reducing complexity, undesired communication overhead, and delays. The joint IGS scheme outperforms the sub-optimal distributed approach, which exceeds traditional PGS in achievable rate performance. Interestingly, the distributed approach performs close to optimal joint IGS for lower impropriety levels.

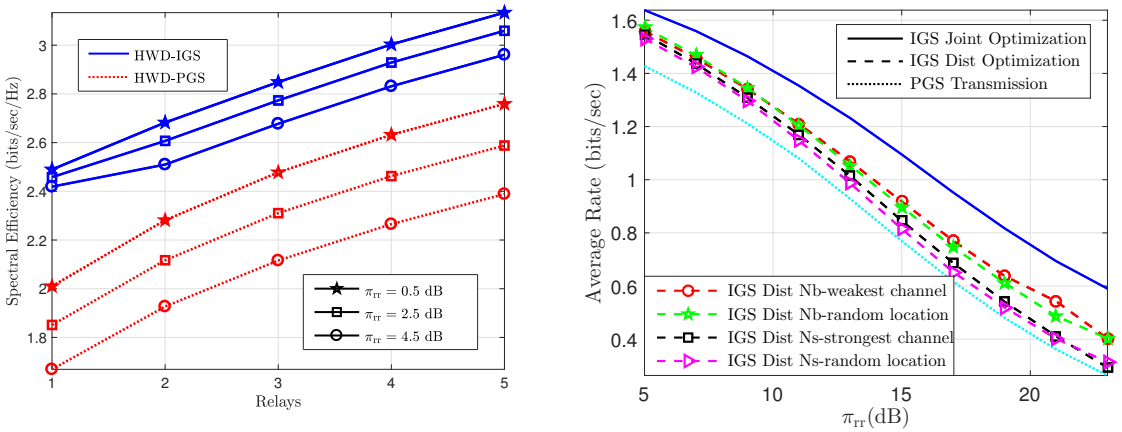
[4]Maximally impaired hardware is attained when the one-dimensional additive distortion power/variance becomes equal to the transmitted signal power in the absence of any other mitigation strategy. Alternatively, it is occurred when the in-phase and quadrature-phase distortion components are fully correlated for a given impropriety level.

Table 4.2: Rayleigh Fading and Free Space Distance Pathloss Model

System Parameters	Parameter Values				
Number of relays	1	2	3	4	5
Transmitting nodes (Source and Relays)	2	3	4	5	6
Distance (d) between adjacent nodes (km)	0.30	0.20	0.15	0.12	0.10
Transmit power per node (mW)	5.0	3.3	2.5	2.0	1.7

4.5.2 Number of Participating Relays

Next, we investigate the effect of increasing number of relays on the system spectral efficiency in the absence of a direct link between a given source and destination located 600m apart. We assume a limited power budget of 10mW for all participating transmitters including source and relays. Thus, increasing the number of relays decreases the power transmitted by one transmitter yielding negligible inter-relay interference. We have adopted a Rayleigh fading and free-space distance path-loss model (Pathloss(dB) $= 92.45 + 20\log_{10} d_{km} + 20\log_{10} f_{GHz}$) [284]. We assume increasing number of relays ranging from 1 to 5 which decreases inter-node distances and path losses. We also choose uniform transmit power distribution among all transmitting nodes as shown in Table 4.2. We further assume 2.1GHz carrier frequency, 20MHz channel bandwidth and three levels of RSI i.e., 0.5dB, 2.5dB and 4.0dB as shown in Figure 4.4a. For simplicity, we assume equal HWDs and impropriety levels at source, relays and destination as presented by $\kappa_{\mathrm{nm}} = 0.5$ and $|\tilde{\kappa}_{\mathrm{nm}}| = 0.9$, respectively. Intuitively, the absence of inter-relay interference and better link strength with increasing number of relays guarantees an increase in the spectral efficiency. However, the presence of HWDs and limited power budget limits this performance gain. Moreover, increasing RSI drastically degrades the average rate performance especially for PGS and higher number of participating relays. It is evident from Figure 4.4a that IGS transmission can improve the spectral efficiency up to 58% as compared to the existing PGS transmission in the presence of residual interferences. Conclusively, the joint IGS



(a) Spectral efficiency vs number of relays.

(b) Average rate vs RSI for IGS-Dist.

Figure 4.4: Average Achievable Rate Performance.

scheme efficiently mitigates the interference effects for any given number of relays and RSI levels.

4.5.3 Impact of Cluster size and Placement in Distributed Algorithms

Another simulation example in Figure 4.4b illustrates the system achievable rate attained by various forms of distributive algorithms keeping the joint-IGS scheme as a benchmark. A multihop (MH) system using 4-relays between the source and destination is optimized using four different forms of $N = 3$ distributive framework employing algorithm 2 and 3. IGS distributed algorithm with N_{b} cluster divides the 5-link network into a cluster of size $N = 3$ and groups the remaining nodes in a bigger cluster of size $N_{\mathrm{b}} = 4$. Further division is based on the placement of this N_{b} cluster, it is evident from Figure 4.4b that the odd cluster placement at weak channel location (IGS Dist Nb) outperforms the random placement in the absence of CSI. Identically, IGS distributed algorithm with N_{s} cluster divides the 5-link network into 2 clusters of size $N = 3$ and the remaining link is isolated. Again, the placement of the segregated link at the strongest channel location (IGS Dist Ns) performs better than the random placement at any other location. It is worth noting that the N_{b} cluster formation outperforms the N_{s} cluster formation irrespective of the cluster placement.

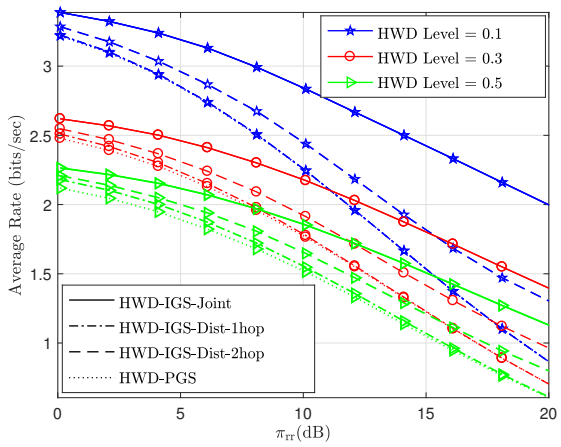


Figure 4.5: Average achievable rate versus self-interference for various cluster sizes.

Therefore, N_b distributed algorithm is the preferred choice when feasible. Evidently, all the proposed IGS schemes provide significant rate compensation/improvement at different interference levels as compared to its counterpart PGS scheme.

Finally, the impact of cluster size is analyzed on the end-to-end achievable rate for three different residual HWD levels with 0.9 impropriety in Figure 4.5. 3-relay FDR system under HWD and RSI is locally optimized using 3-different cluster sizes $N = 2$ or 1-hop, $N = 3$ or 2-hop, and $N = 5$ or 4-hop or joint IGS optimization. We study the simulation results for a range of RSI from 0dB to 20dB. Interestingly, the proposed IGS schemes outperform PGS scheme especially for the significant HWD and RSI levels. It is apparent from Figure 4.5 that the bigger the cluster size, the closer rate performance of the distributed to the joint optimization one. This proves the previously discussed intuition in Section 4.4.2. However, there is a trade-off between the cluster size and the system complexity, round-trip delays and communication overhead. Therefore, cluster size can be decided as per system configuration.

Numerical and simulation results clearly advocate the benefits of deploying various forms of the IGS scheme over the PGS scheme to improve the system performance in terms of end-to-end achievable rate for various levels of RSI and HWD in the MH-FDR system. Moreover, the distributed approach performs closer to the joint optimization approach for

big cluster sizes and lesser number of clusters as shown in Figure 4.4b and Figure 4.5, respectively. It is owing to the accommodation of the RSI effects at all intermediate nodes.

The theoretical limits attained by IGS transmission for the achievable rate of a MH-DF-FDR system suffering from HWDs establish the performance limits that can be achieved through future communication standards. The existing techniques can achieve the performance demonstrated by the PGS scheme. However, this work motivates the future research to propose appropriate adaptive coding and modulation which can achieve the IGS performance. As an example, one of our recent work quantifies the error performance improvement obtained by transforming traditional symmetric M-QAM to asymmetric M-QAM transmission to mitigate the performance degradation due to hardware impairments and realize the benefits of asymmetric (improper) signaling [285].

4.6 Conclusion

In this chapter, we analyzed the effectiveness of using improper Gaussian signaling (IGS) scheme in multi-hop decode-and-forward full-duplex relaying (MH-DF-FDR) systems under residual self-interference (RSI) and hardware distortions (HWD). To this end, we expressed the achievable rate for the underlying system and tuned the IGS pseudo-variance to maximize the end-to-end achievable rate. We presented, analyzed and illustrated two realization schemes named as joint and distributed optimization schemes. Distributed-IGS is further categorized as per the cluster size as well as its relative position in the system network. Distinct forms of IGS-scheme can be adopted for suitable system configurations. For a small system configuration with fewer hops joint-IGS is the preferred choice. However, for a larger-hops system, the joint-IGS renders sub-optimal results pertaining to the inevitable processing and round-trip delays back and forth from the central-node at the cost of increased system complexity and communication overhead. Therefore, distributed-IGS is the preferred approach for large system configurations. Furthermore, distributed-IGS with bigger odd-cluster along with the optimal cluster placement is the preferred choice

as per the acquired simulation results. In a nutshell, all forms of IGS are proven to be promising candidates for next generation networks that can significantly improve the overall achievable rate under various HWD and RSI levels, which have asymmetric signatures on the useful signal.

Chapter 5

Asymmetric Signaling for HWI Systems

Error probability study of hardware impaired (HWI) systems highly depends on the adopted model. Considering the distinct improper Gaussian features of HWI systems, captured by recent models, HWI-aware receivers are designed. An optimal ML receiver serves as a performance benchmark, and a sub-optimal linear minimum mean square error (LMMSE) receiver introduces a reduced-complexity implementation. Whereas, the conventional HWI-unaware minimum Euclidean distance (MED) receiver, based on the proper noise assumption, exhibits substandard performance. Next, the average error probability of the proposed optimal ML-receiver is analyzed, where several tight bounds and approximations are derived for various HWI systems. Motivated by the benefit of improper Gaussian signaling in mitigating HWI, which is proven in recent studies, asymmetric modulation is adopted and optimized for transmission. The numerical results demonstrate a bit error rate (BER) reduction up to 70% of the proposed HWI-aware receivers over HWI-unaware receivers. Moreover, the asymmetric modulation is shown to reduce the BER by 93%. These results signify the importance of incorporating accurate HWI models, designing appropriate receivers and optimizing signal transmission for BER performance compensation.

This chapter is organized as: Section 5.1 describes the significance and main contributions of this work. Section 5.2 describes the adopted system along with its complete characterization. In section 5.3, we propose an optimal receiver and a sub-optimal linear receiver for the system under mentioned HWIs. Next, we present error probability analysis based on the pairwise error probability (PEP) for generalized M-ary modulation scheme in section 5.4. Later, average error probability bounds and approximations are computed

for three different adopted systems in section 5.5. Section 5.6 illustrates various forms of transmit signal designs to reduce error probability followed by the numerical results in Section 5.7 and the conclusion in Section 5.8.

5.1 Main Contributions

In this chapter, we analyze the error probability performance of communication systems suffering from HWIs represented in IQI and additive distortions at both transmitter and receiver. Moreover, motivated by the theoretical limits results in [20], which demonstrate the benefits of employing improper Gaussian signaling to improve the performance of hardware impaired systems, we adopt asymmetric signaling scheme to minimize the error probability performance. In the following, we summarize the main contributions as:

- Studying and accurately quantifying the asymmetric characteristics of the aggregate HWIs from various impairment sources.
- Deriving the optimal maximum likelihood (ML) detector and the suboptimal linear minimum mean square error (LMMSE) receiver based on the improper interference, SI information signals, and asymmetric signaling transmission.
- Analyzing PEP based on the accurate model and optimal ML detector.
- Analyzing the average error probability and deriving bounds and approximations for various HWI system configurations.
- Designing an asymmetric modulation signal to minimize the error probability.

5.2 System Description

Consider a wireless communication system with RF transceiver impairments including hardware distortions and IQI. These impairments can be modeled as detailed in Chapter

3. Therefore, the received signal with aggregate impairments is given in (3.7) as

$$y = \sqrt{p}\tilde{h}_1 x + \sqrt{p}\tilde{h}_2 x^* + z, \tag{5.1}$$

where x is a band-pass modulated signal taken from M-QAM, M-PSK or M-PAM constellation. Moreover, $\tilde{h}_1 = \mu_1\nu_1 h + \mu_2\nu_2^* h^*$ and $\tilde{h}_2 = \mu_1\nu_2 h + \mu_2\nu_1^* h^*$ are modified channel gain parameters accommodating fading h and IQI $(\nu_1, \nu_2, \mu_1, \mu_2)$ characteristics . The aggregated noise z is given by (3.8). Thus, the transceiver IQI transforms the symmetric transmitted signal to asymmetric received signal and the proper Gaussian interference to improper Gaussian interference [20].

The useful signal component in (3.7) comprises of both the actual signal x and the SI signal $x*$. Thus, treating the SI term as mere interference will result in a loss of useful information. For a given channel and I/Q parameters, the instantaneous signal-to-noise ratio (SNR) averaged over additive distortions and M signal constellation with $\mathrm{E}\{|x_m|^2\} = \sigma_x^2$ and $\mathrm{E}\{x_m^2\} = \tilde{\sigma}_x^2$, $m = 1, 2, \ldots, M$, is given as

$$\Lambda = \frac{p\left(\left|\tilde{h}_1\right|^2\sigma_x^2 + \left|\tilde{h}_2\right|^2\sigma_x^2 + \tilde{h}_1\tilde{h}_2^*\tilde{\sigma}_x^2 + \tilde{h}_2\tilde{h}_1^*\tilde{\sigma}_x^{2*}\right)}{\left(p|h|^2\sigma_t^2 + \sigma_r^2\right)\left(|\mu_1|^2 + |\mu_2|^2\right)}. \tag{5.2}$$

We relax the transmit distortion obtaining reduced complex system and express p in terms of the average SNR $\bar{\Lambda} = \mathbb{E}_h\left[\Lambda\right]$ as

$$p = \frac{\sigma_r^2}{\lambda}\left[\frac{\bar{\Lambda}}{|\nu_1|^2\sigma_x^2 + |\nu_2|^2\sigma_x^2 + 2\Re\left\{\nu_1\nu_2^*\tilde{\sigma}_x^2\right\}}\right]. \tag{5.3}$$

The transceiver IQI and the improper additive noise change the system characteristics and render the traditional MED receiver as a sub-optimal detection solution. MED can neither support the SI structure induced by the IQI nor the asymmetric characteristics of the additive noise.

5.3 Receiver Design

In this section, we propose an optimal ML receiver for the adopted HWI system model. We further propose a least complex LMMSE receiver and use the conventional minimum Euclidean distance receiver for performance comparison.

5.3.1 Optimal Maximum Likelihood Receiver

Given the equiprobable symbols and improper interference, the conventional MED receiver for Gaussian interference performs sub-optimally as it fails to accommodate the dependent and non-identical real and imaginary components of the improper noise. Therefore, we propose an optimal ML receiver for the presented hardware impaired system model. The aggregated received signal in (3.7) is equivalent to

$$y = \sqrt{p}\chi^m + z; \quad m = 1, 2, \ldots, M \tag{5.4}$$

where

$$\zeta_d^m = \frac{\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m\right)^2}{\sigma_{\mathrm{I}}^2} + \frac{\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m\right)^2}{\sigma_{\mathrm{Q}}^2} - \frac{2\rho\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m\right)\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} \tag{5.5}$$

and $\chi^m = \tilde{h}_1 x_m + \tilde{h}_2 x_m^*$ depend on the transmitted symbol m and z is the accumulated noise component (3.8). For a given channel, IQI estimates [286] and transmitted signal x_m, the real component y_r and imaginary component y_i of the received signal y are jointly Gaussian with PDF as given in (5.6),

$$f_{y_{\mathrm{r}}y_{\mathrm{i}}|\chi^m}\left(y_{\mathrm{r}}, y_{\mathrm{i}}|\chi^m\right) = \frac{1}{2\pi\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}\sqrt{1-\rho^2}} \exp\frac{-1}{2(1-\rho^2)}\left(\zeta_d^m\right), \tag{5.6}$$

where the useful in-phase signal component χ_{r}^m and quadrature component χ_{i}^m are

given, respectively, as

$$\chi_{\mathrm{r}}^{m} = \left(\tilde{h}_{1\mathrm{r}} + \tilde{h}_{2\mathrm{r}}\right) x_{\mathrm{r}}^{m} + \left(\tilde{h}_{2\mathrm{i}} - \tilde{h}_{1\mathrm{i}}\right) x_{\mathrm{i}}^{m},$$
$$\chi_{\mathrm{i}}^{m} = \left(\tilde{h}_{1\mathrm{i}} + \tilde{h}_{2\mathrm{i}}\right) x_{\mathrm{r}}^{m} + \left(\tilde{h}_{1\mathrm{r}} - \tilde{h}_{2\mathrm{r}}\right) x_{\mathrm{i}}^{m}, \tag{5.7}$$

where $\tilde{h}_{k\mathrm{r}}$ and $\tilde{h}_{k\mathrm{i}}$ are the real and imaginary components of $\tilde{h}_k$ respectively. In addition, x_{r}^{m} and x_{i}^{m} are the real and imaginary components of the transmitted symbol x_m, respectively. Moreover, the in-phase z_{r} and quadrature component z_{i} of z are zero-mean Gaussian RV with variances σ_{I}^2 and σ_{Q}^2, respectively. $\sigma_{\mathrm{Q}}^2 = a_{\mathrm{r}}^2 \sigma_{\mathrm{I}}^2$ and σ_{I}^2 is written as

$$\sigma_{\mathrm{I}}^2 = \frac{\sigma_{\mathrm{r}}^2}{2} + \frac{p\sigma_{\mathrm{t}}^2}{2}|h|^2. \tag{5.8}$$

Interestingly, the noise variance depends not only on the IQI parameter and the distortion variance but also on the transmitted power and channel gain. Therefore, an increase in the transmitted power marks a proportional increase in the noise variance owing to the presence of transmit distortions. Using (3.8), the correlation coefficient ρ between z_{r} and z_{i} simplifies to

$$\rho = \frac{\mathbb{E}\left(z_{\mathrm{r}} z_{\mathrm{i}}\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} = \sin\theta_{\mathrm{r}}. \tag{5.9}$$

Note that ρ merely depends on the amount of rotational error induced by the receiver IQI. Considering the improper Gaussian interference, the optimal ML detection is based on maximizing the following conditional probability

$$\hat{x}_m = \underset{m=1,2,\ldots,M}{\arg\max}\, f_{y_{\mathrm{r}} y_{\mathrm{i}}|\chi^m}\left(y_{\mathrm{r}}, y_{\mathrm{i}}|\chi^m\right) \tag{5.10}$$

Using (5.6), the ML receiver in (5.10) reduces to the minimization of the argument of $f_{y_{\mathrm{r}} y_{\mathrm{i}}|\chi^m}\left(y_{\mathrm{r}}, y_{\mathrm{i}}|\chi^m\right)$, i.e., $\hat{x}_m = \underset{m=1,2,\ldots,M}{\arg\min}\left\{\zeta_d^m\right\}$. Using (5.8) and (5.9), the optimal ML

receiver is simplified as

$$\hat{x}_m = \underset{m=1,2,\ldots,M}{\arg\min} \left\{ (y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m)^2 + \frac{1}{a_{\mathrm{R}}^2}(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m)^2 - \frac{2\sin\theta_{\mathrm{R}}}{a_{\mathrm{R}}}\,(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m)\,(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m) \right\}. \tag{5.11}$$

5.3.2 Minimum Euclidean Distance Receiver

Considering the traditional MED receiver with the assumption of circularly symmetric complex Gaussian interference

$$\hat{x}_m = \underset{m=1,2,\ldots,M}{\arg\min} |y - \sqrt{p}\chi^m|^2, \tag{5.12}$$

where $\chi^m = hx_m$ under ideal hardware assumption. The disagreement between (5.11) and (5.12) is owing to the presence of amplitude translation a_{r} and rotational error θ_{r} caused by the receiver IQI. The amplitude error is responsible for the non-identical real and imaginary components and the rotational error is accountable for the mutual correlation.

5.3.3 Linear Minimum Mean Square Error Receiver

Linear receivers are popular for the ease of implementation and least complex receiver design. Therefore, the LMMSE receiver is derived for the adopted HWI system model. Consider the linear estimator $\hat{x} = c^* y$, we aim to design the estimator in order to minimize the mean square error $\mathrm{E}\left\{\|\hat{x} - x\|^2\right\} = \mathrm{E}\left\{\|c^* y - x\|^2\right\}$. Given the convex nature of norm minimization problem, second order sufficient condition renders the stationary point as the global optimal solution. Thus, the optimal c is given as $\bar{c} = R_{yy}^{-1} R_{yx}$, where R_{yy} is the covariance of the received signal y given in (3.7) and R_{yx} is the cross-covariance of the received signal y and the transmitted signal x. Considering the asymmetric transmission,

the $\bar{c}$ is derived as:

$$\bar{c} = \frac{\sqrt{p}h_1\sigma_x^2 + \sqrt{p}h_2\tilde{\sigma}_x^{2*}}{p\left(|h_1|^2 + |h_2|^2\right)\sigma_x^2 + 2p\Re\{h_1 h_2^* \tilde{\sigma}_x^2\} + \sigma_z^2}, \tag{5.13}$$

where the aggregate noise variance $\sigma_z^2 = (p|h|^2\sigma_t^2 + \sigma_r^2)(|\mu_1|^2 + |\mu_2|^2)$ from (3.8). Note that the obtained $\hat{x}$ is LMMSE estimate of the transmitted symbol x_m. Thus, the ultimate detection relies on finding the closest possible symbol in a reference constellation using slicers.

The optimal ML receiver serves as the performance benchmark for the conventional MED receiver, which is considered optimal under the ideal hardware and AWGN assumption. Moreover, the performance analysis of LMMSE receiver in Section 5.7 justifies its suitability for reduced complexity practical implementation.

5.4 Error Probability Analysis

Symbol error probability analysis has been carried out based on the proposed optimal ML receiver presented in section 5.3. This work focuses on the derivation of pairwise error probabilities in higher order modulation schemes for the simplified and tractable analysis of the adopted system model under HWI. We further analyze the asymptotic behavior of the derived error probability expression with increasing SNR.

5.4.1 Symbol Error Probability

Symbol error probability (P_{s}) is defined as the probability of detecting χ^n given χ^m was transmitted with prior probability $\Pr(\chi^m)$ for all $m \neq n$. Thus, P_{s} is expressed as

$$P_{\mathrm{s}} \leq \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} \Pr(\chi^m \to \chi^n | \chi^m) \Pr(\chi^m) \tag{5.14}$$

where the term $\Pr\left(\chi^m \to \chi^n | \chi^m\right)$ is the well-known PEP expression. Under the assumption of equal likely transmitted symbols of M^{-1}, (5.14) reduces to

$$P_{\mathrm{s}} \leq \frac{1}{M} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} \Pr\left(\chi^m \to \chi^n | \chi^m\right). \tag{5.15}$$

At high SNR, most error events occur with neighboring symbols, where the probability of making error with neighboring symbol is higher than that with a farther point which is not located at minimum distance ($d_{\min}$) from the transmitted symbol. Thus, we can upper bound P_{s} by the PEP between two closest possibilities as $P_{\mathrm{s}} \leq (M-1) \Pr\left(\chi^m \to \chi^n | \chi^m\right)$, where x_n is assumed to be the neighbor located at $d_{\min}$ from x_m. This upper bound is the loose union bound as it assumes all neighbors to be located at $d_{\min}$, which is only valid for orthogonal modulation schemes and invalid for other modulation schemes. Therefore, we employ the nearest neighbor union bound (NNUB) for M-ary modulation.

$$P_{\mathrm{s}} \leq \psi \Pr\left(\chi^m \to \chi^n | \chi^m\right) \tag{5.16}$$

where ψ represents the average number of neighbors located at $d_{\min}$ distance to each other. For example, for M-phase shift keying (PSK) $\psi = 2$ and for M-QAM square constellation, it is found to be

$$\psi = \frac{1}{M}\left(4\mathrm{x}2 + 4\left(\sqrt{M}-2\right)\mathrm{x}3 + \left(\sqrt{M}-2\right)^2 \mathrm{x}4\right). \tag{5.17}$$

Based on the optimal ML receiver in (5.11), PEP is defined by the conditional probability expressed in (5.18).

$$\Pr(\chi^m \to \chi^n | \chi^m) = \Pr\left\{ \begin{matrix} \frac{\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m\right)^2}{\sigma_{\mathrm{I}}^2} + \frac{\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m\right)^2}{\sigma_{\mathrm{Q}}^2} - \frac{2\rho\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^m\right)\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^m\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} > \\ \frac{\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^n\right)^2}{\sigma_{\mathrm{I}}^2} + \frac{\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^n\right)^2}{\sigma_{\mathrm{Q}}^2} - \frac{2\rho\left(y_{\mathrm{r}} - \sqrt{p}\chi_{\mathrm{r}}^n\right)\left(y_{\mathrm{i}} - \sqrt{p}\chi_{\mathrm{i}}^n\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} \end{matrix} \right\}. \tag{5.18}$$

Given that x_m was transmitted, it is justified to replace all occurrences of y_{r} with $\sqrt{p}\chi_{\mathrm{r}}^m + z_{\mathrm{r}}$ and y_{i} with $\sqrt{p}\chi_{\mathrm{i}}^m + z_{\mathrm{i}}$. Thus, we obtain the following simplified PEP expression

$$\Pr\left(\chi^m \to \chi^n | \chi^m\right) = \Pr\left\{\eta > \frac{p\zeta_{\mathrm{r}}^2}{\sigma_{\mathrm{I}}^2} + \frac{p\zeta_{\mathrm{i}}^2}{\sigma_{\mathrm{Q}}^2} - \frac{2p\rho\zeta_{\mathrm{r}}\zeta_{\mathrm{i}}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}\right\}, \tag{5.19}$$

where η is obtained by the superposition of z_{r} and z_{i}

$$\eta = \left(\frac{2\rho\zeta_{\mathrm{i}}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} - \frac{2\zeta_{\mathrm{r}}}{\sigma_{\mathrm{I}}^2}\right)\sqrt{p}z_{\mathrm{r}} + \left(\frac{2\rho\zeta_{\mathrm{r}}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} - \frac{2\zeta_{\mathrm{i}}}{\sigma_{\mathrm{Q}}^2}\right)\sqrt{p}z_{\mathrm{i}}, \tag{5.20}$$

One can show that η is Gaussian RV with zero mean and a variance σ_η^2 that is expressed as

$$\sigma_\eta^2 = 4\left(1-\rho^2\right)\left[\frac{p\zeta_{\mathrm{r}}^2}{\sigma_{\mathrm{I}}^2} + \frac{p\zeta_{\mathrm{i}}^2}{\sigma_{\mathrm{Q}}^2} - \frac{2p\rho\zeta_{\mathrm{r}}\zeta_{\mathrm{i}}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}\right]. \tag{5.21}$$

Moreover, $\zeta_{\mathrm{r}} = {\chi_{\mathrm{r}}}^m - {\chi_{\mathrm{r}}}^n$ and $\zeta_{\mathrm{i}} = {\chi_{\mathrm{i}}}^m - {\chi_{\mathrm{i}}}^n$ are the distances between real and imaginary components of the transmitted and received useful signal component in an error event, respectively. Consequently, the probability in (5.19), is the cumulative distribution function (CDF) of the Gaussian RV η and is given by

$$\Pr\left(\chi^m \to \chi^n | \chi^m\right) = Q\left(\sqrt{\frac{1}{4\left(1-\rho^2\right)}\left(\frac{p{\zeta_{\mathrm{r}}}^2}{\sigma_{\mathrm{I}}^2} + \frac{p{\zeta_{\mathrm{i}}}^2}{\sigma_{\mathrm{Q}}^2} - \frac{2p\rho\zeta_{\mathrm{r}}\zeta_{\mathrm{i}}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}\right)}\right). \tag{5.22}$$

The instantaneous PEP expression depends on the distance between the transmitted and received erroneous signal component under Rayleigh fading, IQI and transceiver additive distortions. It further relies on the average transmit power and statistical characteristics of the non-identical and dependent I/Q phase improper interference components.

5.4.2 Asymptotic Analysis

Throughout the following discussion, we investigate the system performance of HWIs systems under high SNR assumption. The dependence of the statistical characteristics of the

improper interference components (5.8) on the average SNR motivates us to analyze the asymptotic behavior of the PEP expression with respect to the parameter p. Substituting (5.8) and (5.9) in (5.22) yields

$$\mathrm{PEP} = Q\left(\sqrt{\frac{a_{\mathrm{r}}^2 p\zeta_{\mathrm{r}}^2 + p\zeta_{\mathrm{i}}^2 - 2a_{\mathrm{r}} p \sin\theta_{\mathrm{r}}\zeta_{\mathrm{r}}\zeta_{\mathrm{i}}}{2a_{\mathrm{r}}^2\cos^2\theta_{\mathrm{r}}\left(\sigma_{\mathrm{r}}^2 + p\sigma_{\mathrm{t}}^2|h|^2\right)}}\right). \tag{5.23}$$

The asymptotic behavior of the instantaneous PEP depicts an irreducible error floor, which is found from the following limit,

$$\lim_{p\to\infty} \mathrm{PEP} = Q\left(\sqrt{\frac{a_{\mathrm{r}}^2\zeta_{\mathrm{r}}^2 + \zeta_{\mathrm{i}}^2 - 2a_{\mathrm{r}}\sin\theta_{\mathrm{r}}\zeta_{\mathrm{r}}\zeta_{\mathrm{i}}}{2a_{\mathrm{r}}^2\sigma_{\mathrm{t}}^2\cos^2\theta_{\mathrm{r}}|h|^2}}\right). \tag{5.24}$$

It is important to note that out of all the aforementioned impairments, transmitter distortions are mainly responsible for the error floor. Intuitively, increasing SNR also increases the transmitter distortions, thus rendering it ineffective to reduce the error probabilities. From (5.23), it is evident that assuming $\sigma_{\mathrm{t}}^2 = 0$ will result in $\lim_{p\to\infty} \mathrm{PEP} = 0$.

5.5 Average Probability of Error

In this section, we analyze the average symbol error probability (SEP) for three different scenarios; system under transmitter and receiver HWIs, system with only receiver HWIs and system with only transmitter HWIs.

5.5.1 System with Transmitter and Receiver Hardware Impairments

Consider the generalized system model under transmitter and receiver IQI as well as the transmitter and receiver additive distortions. The PEP of the underlying system is given by (5.23). The average SEP ($\bar{P}_{\mathrm{s}}$) of the generalized model is investigated assuming small θ_{r}, where it has been shown experimentally that $\theta_{\mathrm{r}} \leq 5^0$ is a valid assumption for the IQI rotational errors in [8]. As a result, we can assume $\sin\theta_{\mathrm{r}} \approx 0$, which reduces the PEP

expression to

$$\mathrm{PEP} = Q\left(\sqrt{\frac{\alpha h_{\mathrm{r}}^2 + \beta h_{\mathrm{i}}^2}{1+\gamma\left(h_{\mathrm{r}}^2 + h_{\mathrm{i}}^2\right)}}\right), \tag{5.25}$$

where $\alpha = p\left(a_{\mathrm{r}}^2\varsigma_1^2 + \varsigma_2^2\right)/(2\sigma_{\mathrm{r}}^2 a_{\mathrm{r}}^2 \cos^2\theta_{\mathrm{r}})$ and $\beta = p\left(a_{\mathrm{r}}^2\varsigma_3^2 + \varsigma_4^2\right)/(2\sigma_{\mathrm{r}}^2 a_{\mathrm{r}}^2 \cos^2\theta_{\mathrm{r}})$. In addition, γ is the scaled ratio of transmitter distortion variance to receiver distortion variance, i.e., $\gamma = p\sigma_{\mathrm{t}}^2/\sigma_{\mathrm{r}}^2$. Moreover, $\varsigma_s, s \in \{1,2,3,4\}$ are obtained using IQI parameters and the respective modulation scheme characteristics as $\varsigma_1 = \xi_{\mathrm{r}},\ \varsigma_2 = a_{\mathrm{t}} a_{\mathrm{r}} \cos\theta_{\mathrm{t}} \cos\theta_{\mathrm{r}} \xi_{\mathrm{i}},\ \varsigma_3 = a_{\mathrm{t}} cos\theta_{\mathrm{t}} \xi_{\mathrm{i}}$ and $\varsigma_4 = a_{\mathrm{r}} \cos\theta_{\mathrm{r}} \xi_{\mathrm{r}}$. They particularly depend on the symbol separation between two closest possibilities with $\xi_{\mathrm{r}} = (x_{\mathrm{r}}^m - x_{\mathrm{r}}^n)$ being the separation between the real components and $\xi_{\mathrm{i}} = (x_{\mathrm{i}}^m - x_{\mathrm{i}}^n)$ being the separation between the corresponding imaginary components. Let us denote the argument of the Q-function in (5.25) as

$$\vartheta = \frac{\alpha h_{\mathrm{r}}^2 + \beta h_{\mathrm{i}}^2}{1+\gamma\left(h_{\mathrm{r}}^2 + h_{\mathrm{i}}^2\right)}. \tag{5.26}$$

Under the assumption of Rayleigh fading channels, the RV ϑ for a given signal constellation and IQI parameters has a CDF of $F_\vartheta(\vartheta)$ that is derived in Appendix C as

$$F_\vartheta(\vartheta) = \frac{1}{k_1(\vartheta)} \sum_{m=0}^{\infty} \frac{2^{-2m} k_3(\vartheta)^{2m}}{[m!]^2 k_2(\vartheta)^{2m+1}} \left[2m! - \Gamma\left(2m+1, \vartheta k_2(\vartheta)\right)\right]. \tag{5.27}$$

Using (5.16), $\bar{P}_{\mathrm{s}}$ is obtained by averaging the PEP expression in (5.25) with respect to ϑ as

$$\bar{P}_{\mathrm{s}} \leq \psi \mathrm{E}_\vartheta\left[Q\left(\sqrt{\vartheta}\right)\right] = \psi \int_0^\delta Q\left(\sqrt{\vartheta}\right) f_\vartheta(\vartheta) d\vartheta. \tag{5.28}$$

Average PEP Bound

Using the Chernoff bound $Q(x) \leq 0.5\exp(-x^2/2)$, yields

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2} \int_0^\delta e^{-\frac{\vartheta}{2}} f_\vartheta(\vartheta) d\vartheta = \frac{\psi}{2} e^{-\frac{\delta}{2}} + \frac{\psi}{4} \int_0^\delta e^{-\frac{\vartheta}{2}} F_\vartheta(\vartheta) d\vartheta. \tag{5.29}$$

The equality in (5.29) is obtained after applying integration by parts, which enables us to represent the presented bound in terms of $F_{\vartheta}(\vartheta)$. Average PEP bound can be accurately evaluated using numerical integration techniques such as Gaussian quadrature. Existence of the definite integral with finite limits motivates us to employ the popular $n^{\text{th}}-$ order Legendre orthogonal polynomial $P_N(t)$ [287, (25.4.29-30)]. We obtain the required form by following parametrization,

$$\bar{P}_{\text{s}} \leq \frac{\psi}{2} e^{-\frac{\delta}{2}} + \frac{\delta\psi}{8} \int_{-1}^{+1} s\left(0.5\delta\left(t+1\right)\right) dt. \tag{5.30}$$

where $s(\vartheta) = e^{-\frac{\vartheta}{2}} F_{\vartheta}(\vartheta)$. Gaussian quadrature enables us to numerically approximate the integral in (5.30) with weights ω_n and $s(\vartheta)$ evaluation at the corresponding instances.

$$\bar{P}_{\text{s}} \leq \frac{\psi}{2} e^{-\frac{\delta}{2}} + \frac{\delta\psi}{8} \left[\sum_{n=1}^{N} \omega_n(t_n) s\left(0.5\delta\left(t_n+1\right)\right) + R_N \right], \tag{5.31}$$

where t_n is the nth zero of the $N^{\text{th}}-$order Legendre polynomial $P_N(t)$ and the corresponding weights are obtained using [287, (25.4.29)].

$$\omega_n(t_n) = \frac{2}{\left(1-t_n^2\right)\left[\mathrm{P}_N{}^{(1)}\left(t_n\right)\right]^2}. \tag{5.32}$$

The approximation error is given by residual term R_N [287, (25.4.30)], which decreases significantly with the increasing order of the Legendre polynomial and expressed by

$$R_N = \frac{\delta^{2N+1}(N!)^4}{(2N+1)\left[(2N)!\right]^3} f^{(2N)}(\phi) \quad ; -1 < \phi < 1. \tag{5.33}$$

This numerical integration technique accurately evaluates the intractable integral in (5.29) as (5.31).

Average PEP Approximation

The Craig's formula representation of the Q-function [288, (4.2)] can be employed to approximate the average probability of error as

$$\bar{P}_{\mathrm{s}} \leq \psi \int_0^{\delta} \left(\frac{1}{\pi} \int_0^{\pi/2} e^{-\frac{\vartheta}{2\sin^2\varphi}} d\varphi \right) f_{\vartheta}(\vartheta) d\vartheta. \tag{5.34}$$

Based on uniform convergence, swapping the integrals and employing integration by parts yields the $\bar{P}_{\mathrm{s}}$ as a function of $F_{\vartheta}(\vartheta)$ as

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{\pi} \int_0^{\pi/2} e^{-\frac{\delta}{2\sin^2\varphi}} d\varphi + \frac{\psi}{2\pi} \int_0^{\pi/2} \int_0^{\delta} \frac{e^{-\vartheta/\left(2\sin^2\varphi\right)}}{2\sin^2\varphi} F_{\vartheta}(\vartheta)\, d\vartheta d\varphi. \tag{5.35}$$

The first integral in (5.35) reduces to the complementary error function and the second term involving double integration leads to an intractable analysis. Therefore, we employ Gaussian Quadrature to numerically integrate this definite integration as

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2} \mathrm{erfc}\left(\sqrt{\frac{\delta}{2}}\right) + \frac{\psi}{2\pi} \int_0^{\pi/2} \int_0^{\delta} g(\vartheta, \varphi) d\vartheta d\varphi. \tag{5.36}$$

Choosing the following parametrization for numerical integration enables us to employ $n^{\mathrm{th}}-$ order Legendre orthogonal polynomial in the square interval [-1, 1] as

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2} \mathrm{erfc}\left(\sqrt{\frac{\delta}{2}}\right) + \frac{\delta\psi}{8} \int_{-1}^{1} \int_{-1}^{1} g\left(\frac{\delta}{2}(y+1), \frac{\pi}{4}(x+1)\right) dy\, dx, \tag{5.37}$$

where $y = \frac{2}{\delta}\vartheta - 1$ and $x = \frac{4}{\pi}\varphi - 1$. Eventually the numerical approximation of double integrals is obtained using $K^{\mathrm{th}}-$order and $L^{\mathrm{th}}-$order Legendre polynomial approximations for integration over dy and dx respectively as

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2} \mathrm{erfc}\left(\sqrt{\frac{\delta}{2}}\right) + \frac{\delta\psi}{8} \sum_{l=1}^{L} \sum_{k=1}^{K} \omega_l \omega_k g\left(\frac{\delta}{2}(y_k+1), \frac{\pi}{4}(x_l+1)\right), \tag{5.38}$$

where the weights $\omega_k(y_k)$ and $\omega_l(x_l)$ are obtained using (5.32). Moreover, y_k and x_l are the kth and l^{th} zeros of the $K^{\text{th}}-$order Legendre polynomial $\mathrm{P}_K(y)$ and $L^{\text{th}}-$order Legendre polynomial $\mathrm{P}_L(x)$ respectively. Higher order K and L yield better approximation with minimal residual errors at the cost of computational overhead.

5.5.2 Zero-Distortion Transmitter

In this subsection, we propose a closed form bound and approximation for the average PEP expression of the adopted system model under transmitter and receiver IQI along with the thermal noise at receiver but negligible transmitter distortion. This simplifies the PEP expression in (5.25) to $\mathrm{PEP} = Q\left(\sqrt{\varrho}\right)$ where $\varrho = \alpha h_{\mathrm{r}}^2 + \beta h_{\mathrm{i}}^2$. Following similar steps as in Appendix C, the PDF of $f_\varrho(\varrho)$ for the Rayleigh fading channel is given as

$$f_\varrho(\varrho) = \frac{1}{\lambda\sqrt{\alpha\beta}} e^{-\frac{\varrho}{2\lambda}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)} \mathcal{I}_0\left\{\frac{\varrho}{2\lambda}\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)\right\};\ \varrho \geq 0. \tag{5.39}$$

Thus the average PEP can be evaluated by solving $\bar{P}_{\mathrm{s}} \leq \psi \mathbb{E}_\varrho\left[Q\left(\sqrt{\varrho}\right)\right]$.

Average PEP Bound

Employing the Chernoff bound on the Q-function, we obtain

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2}\int_0^\infty e^{-\frac{\varrho}{2}} f_\varrho(\varrho) d\varrho. \tag{5.40}$$

The uniform convergence and the interchangeable integral and summation enable us to evaluate the closed form expression for the $\bar{P}_{\mathrm{s}}$ bound as shown in Appendix D obtaining

$$\bar{P}_{\mathrm{s}} \leq \frac{\sqrt{\alpha\beta}\psi}{|\beta-\alpha|}\sum_{m=0}^{\infty}\frac{(2m)!}{[m!]^2 2^{2m}}\left[\frac{|\beta-\alpha|}{(\alpha+\beta+\alpha\beta\lambda)}\right]^{2m+1}. \tag{5.41}$$

This yields a closed form upperbound on the average SEP using (5.15) or (5.16).

Exact Average PEP using Craig Representation

The Craig representation of Q-function yields exact average PEP expression and a close approximation for $\bar{P}_{\mathrm{s}}$ using NNUB as

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{\pi} \int_{0}^{\pi/2} \int_{0}^{\infty} e^{-\frac{\varrho}{2\sin^2\varphi}} f_{\varrho}\left(\varrho\right) d\varrho d\varphi. \tag{5.42}$$

Using similar steps to the presented in Appendix D, we solve the integral over ϱ obtaining

$$\bar{P}_{\mathrm{s}} \leq \frac{2\sqrt{\alpha\beta}\psi}{\pi\left|\beta-\alpha\right|} \sum_{m=0}^{\infty} \frac{(2m)!}{(m!)^2 2^{2m}} \int_{0}^{\pi/2} \left(\frac{\left|\beta-\alpha\right| \sin^2\varphi}{\alpha\beta\lambda+(\alpha+\beta)\sin^2\varphi} \right)^{(2m+1)} d\varphi. \tag{5.43}$$

Furthermore, the Wolfram Mathematica integrator solves the complex integral in (5.43) as

$$\bar{P}_{\mathrm{s}} \leq \frac{2\sqrt{\alpha\beta}\psi}{\pi|\beta-\alpha|} \sum_{m=0}^{\infty} \frac{(2m)!}{(m!)^2 2^{2m}(3+4m)} \left(\frac{|\beta-\alpha|}{\alpha\beta\lambda}\right)^{(2m+1)} F_1\left(2m+\tfrac{3}{2}; \tfrac{1}{2}, 2m+1; 2m+\tfrac{5}{2}; 1, \tfrac{-(\alpha+\beta)}{\alpha\beta\lambda}\right) \tag{5.44}$$

Equivalently, $\bar{P}_{\mathrm{s}}$ can be represented with Gauss hypergeometric function for easy implementation as given in (5.45).

$$\bar{P}_{\mathrm{s}} \leq \frac{2\psi}{\pi\lambda\sqrt{\alpha\beta}} \sum_{m=0}^{\infty} \frac{(2m)!}{(m!)^2(3+4m)} \left(\frac{|\beta-\alpha|}{2\alpha\beta\lambda}\right)^{2m} \Gamma\left(\tfrac{1}{2}\right) \frac{\Gamma\left(2m+\frac{5}{2}\right)}{\Gamma(2m+2)} {}_2F_1\left(\left[2m+\tfrac{3}{2}, 2m+1\right], 2m+2, \tfrac{-(\alpha+\beta)}{\alpha\beta\lambda}\right) \tag{5.45}$$

Exact Average PEP using Moment Generating Function (MGF)

The lack of correlation between the real and imaginary channel coefficients in $\varrho = \alpha h_{\mathrm{r}}^2 + \beta h_{\mathrm{i}}^2$ enables us to compute approximately $\bar{P}_{\mathrm{s}}$ using MGF approach. The average PEP is given in terms of MGF of ϱ as

$$\mathbb{E}_{\varrho}\left[Q\left(\sqrt{\varrho}\right)\right] = \frac{1}{\pi} \int_{0}^{\pi/2} M_{\varrho}\left[\frac{-1}{2\sin^2\varphi}\right] d\phi. \tag{5.46}$$

where $M_{\varrho}(s)$ is expressed as, see Appendix E for derivation,

$$M_{\varrho}(s) = \frac{1}{\sqrt{1-\alpha s}\sqrt{1-\beta s}}. \tag{5.47}$$

Therefore, the bound on $\bar{P}_s$ is given as

$$\bar{P}_s \leq \frac{\psi}{\pi}\int_0^{\pi/2} \frac{1}{\sqrt{1+\frac{\alpha}{2\sin^2\varphi}}\sqrt{1+\frac{\beta}{2\sin^2\varphi}}} d\phi. \tag{5.48}$$

The Chernoff, Craig and MGF bounds provide tight bound and approximation on the average PEP, respectively, which can then be employed to evaluate NNUB on average SEP as well as average BER for the adopted HWI system.

5.5.3 System with Negligible Transmitter I/Q Imbalance

This section deals with the receiver IQI under the assumption of minimal transmitter IQI. An example of such scenario is the cellular downlink case when the BTS employs sophisticated signal processing techniques rendering minimal IQI. However, the mobile station employs homodyne RF front-end architecture to assist compactness and energy efficiency. In such scenarios, the system only suffers from the receiver IQI.

In the presence of the thermal noise, we further consider two scenarios with and without transmitter distortion to evaluate average PEP. In the first case, system suffers from receiver IQI as well as non-trivial transmitter and receiver distortions. In this case the average PEP has the same bound as given in (5.29),-(5.31) and a close approximation as given in (5.37),(5.38) with $F_{\vartheta}(\vartheta)$ and the corresponding parameters are given in (C.6)-(C.7). However, the parameters α and β are now defined as $\alpha = p(\xi_{\mathrm{r}}^2 + \cos^2\theta_{\mathrm{r}}\xi_{\mathrm{i}}^2)/(2\sigma_{\mathrm{r}}^2\cos^2\theta_{\mathrm{r}})$, and $\beta = p(\xi_{\mathrm{i}}^2 + \cos^2\theta_{\mathrm{r}}\xi_{\mathrm{r}}^2)/(2\sigma_{\mathrm{r}}^2\cos^2\theta_{\mathrm{r}})$. In the second scenario, when the transmitter exhibits negligible transmit distortion, the receiver I/Q imbalanced system follows the same closed form bound as given in (5.41) and exact average PEP in (5.45), with the corresponding

parameters α and β.

5.5.4 System with Negligible Receiver I/Q Imbalance

After neglecting IQI at the receiver, the PEP expression in (5.25) reduces to

$$\mathrm{PEP} = Q\left(\sqrt{\frac{\alpha|h|^2}{1+\gamma|h|^2}}\right) = Q\left(\sqrt{\Omega}\right), \tag{5.49}$$

where $\alpha = \frac{p}{2\sigma_{\mathrm{r}}^2}\left(\tau_1^2+\tau_2^2\right)$, $\tau_1 = \xi_{\mathrm{r}}$ and $\tau_2 = \left(a_{\mathrm{t}}\sin\theta_{\mathrm{t}}\right)\xi_{\mathrm{r}} + \left(a_{\mathrm{t}}\cos\theta_{\mathrm{t}}\right)\xi_{\mathrm{i}}$. To evaluate the average PEP, we first need to investigate the PDF of Ω that is derived in Appendix F as

$$f_{\Omega}\left(\Omega\right) = \frac{\alpha\lambda}{\left(\alpha-\gamma\Omega\right)^2} e^{\frac{-\lambda\Omega}{\alpha-\gamma\Omega}}; \quad 0 \leq \Omega \leq \alpha/\gamma. \tag{5.50}$$

Average PEP Bound

The average PEP is upper bounded using the Chernoff bound as

$$\bar{P}_{\mathrm{s}} \leq \frac{\alpha\lambda\psi}{2}\int_0^{\alpha/\gamma} \frac{e^{\frac{\gamma\Omega^2-(\alpha+2\lambda)\Omega}{2(\alpha-\gamma\Omega)}}}{\left(\alpha-\gamma\Omega\right)^2} d\Omega. \tag{5.51}$$

Again, we employ Gaussian quadrature numerical integration to evaluate the presented upper bound using Legendre polynomial, which simplifies the average PEP as

$$\bar{P}_{\mathrm{s}} \leq \frac{\lambda\psi}{\gamma}\int_{-1}^{+1} g\left(t\right)dt = \frac{\lambda\psi}{\gamma}\sum_{k=1}^{n}\omega_k g\left(t_k\right) + R_n, \tag{5.52}$$

where the nodes t_k are the zeros of $n^{\mathrm{th}}-$order Legendre polynomial and the weights ω_k and the residual term R_n are defined in (5.32) and (5.33) respectively. In addition, the function $g(t)$ is derived from the integrand in (5.51) by choosing the parametrization $t =$

$-1+2\Omega\gamma/\alpha$ as

$$g\left(t\right)=\frac{1}{\left(t-1\right)^{2}}\exp\left(\frac{-\alpha\left(t+1\right)}{4\gamma}+\frac{\lambda\left(t+1\right)}{\gamma\left(t-1\right)}\right). \tag{5.53}$$

Average PEP Approximation

Using the Craig representation of the Q-function, the average error probability of the transmitter only IQI reduces to

$$\bar{P}_{\mathrm{s}}\leq\frac{\psi}{\pi}\int_{0}^{\alpha/\gamma}\int_{0}^{\pi/2}e^{-\frac{\Omega}{2\sin^{2}\phi}}\frac{\alpha\lambda}{\left(\alpha-\gamma\Omega\right)^{2}}e^{\frac{-\lambda\Omega}{\alpha-\gamma\Omega}}d\phi d\Omega. \tag{5.54}$$

which cannot be evaluated in a tractable way. However, by incorporating similar procedure to approximate the $\bar{P}_{\mathrm{s}}$ bound in (5.54) using numerical integration Gaussian Quadrature technique by choosing the following parameterization yields

$$\bar{P}_{\mathrm{s}}\leq\frac{\alpha^{2}\lambda\psi}{8\gamma}\int_{-1}^{1}\int_{-1}^{1}h\left(\frac{\alpha}{2\gamma}\left(u+1\right),\frac{\pi}{4}\left(v+1\right)\right)du\,dv, \tag{5.55}$$

where $u=\frac{2\gamma}{\alpha}\Omega-1$ and $v=\frac{4}{\pi}\varphi-1$. The numerical approximation of double integrals in the square region is obtained using $K^{\mathrm{th}}-$ order and $L^{\mathrm{th}}-$ order Legendre polynomial approximations for integration over du and dv, respectively.

$$\bar{P}_{\mathrm{s}}\leq\frac{\alpha^{2}\lambda\psi}{8\gamma}\sum_{l=1}^{L}\sum_{k=1}^{K}\omega_{l}\omega_{k}h\left(\frac{\alpha}{2\gamma}\left(u_{k}+1\right),\frac{\pi}{4}\left(v_{l}+1\right)\right), \tag{5.56}$$

where the weights $\omega_k(u_k)$ and $\omega_l(v_l)$ are obtained using (5.32). Moreover, u_k and v_l are the kth and l^{th} zeros of the $K^{\mathrm{th}}-$ order and $L^{\mathrm{th}}-$ order Legendre polynomials, respectively.

The average BER can be derived as $\bar{P}_{\mathrm{b}}\leq\bar{P}_{\mathrm{s}}/\log_{2}\mathrm{M}$ assuming gray coding in the high SNR regime. The average PEP bounds and approximations are given in (5.31) and (5.38) for the system under dual IQI with transmit distortion respectively. Similarly, the average PEP bounds and approximations for the dual IQI without transmit distortion are

given in (5.41) and (5.45) respectively. The same results hold for the system with negligible transmitter IQI with and without transmit distortion given modified parameters α and β. Similarly, the average PEP bounds and approximations for the system with negligible receiver IQI are proposed in (5.52) and (5.56), respectively.

5.6 Transmit Signaling Design

Transmit signaling can be optimized to reduce the probability of error by effectively mitigating HWI effects. Practically, the signal constellation can be modified within power constraints to achieve a lower probability of error. The rotation and translation in-variance of error probability is a well-known fact. However, this is only true in the presence of proper additive noise. We propose an asymmetric modulation scheme where effective rotation and optimal scaling of a symmetric signal constellation is performed to mitigate the HWIs impact in distorting signals and deforming the aggregate noise to improper Gaussian one. The PEP expression in (5.22) can be presented in a matrix form as follows

$$\Pr\left(\mathbf{x}^m \to \mathbf{x}^n | \mathbf{x}^m\right) = Q\left(\sqrt{\frac{p}{4\left(1-\rho^2\right)}\mathbf{x}_{mn}^{\mathrm{T}}\tilde{\mathbf{H}}\mathbf{Y}\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{x}_{mn}}\right), \tag{5.57}$$

where $\mathbf{x}^m = [\Re\{x_m\} \;\; \Im\{x_m\}]^{\mathrm{T}}$, $\tilde{\mathbf{H}}$ and $\mathbf{Y}$ are expressed as

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{h}_{1r} + \tilde{h}_{2r} & \tilde{h}_{2i} - \tilde{h}_{1i} \\ \tilde{h}_{1i} + \tilde{h}_{2i} & \tilde{h}_{1r} - \tilde{h}_{2r} \end{bmatrix}, \; \mathbf{Y} = \begin{bmatrix} 1/\sigma_{\mathrm{I}}^2 & -\rho/\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}} \\ -\rho/\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}} & 1/\sigma_{\mathrm{Q}}^2 \end{bmatrix} \tag{5.58}$$

and $\mathbf{x}_{mn}$ represent the distance vector between any two information symbols x_m and x_n and is written as

$$\mathbf{x}_{mn} = \begin{bmatrix} \Re\{x_m\} - \Re\{x_n\} \\ \Im\{x_m\} - \Im\{x_n\} \end{bmatrix}. \tag{5.59}$$

In order to design the rotation and scaling, consider the information bearing signal $s = s_{\mathrm{r}} + is_{\mathrm{i}}$ and the vector representation of mth symbol as $\mathbf{s}^m = [\Re\{s_m\} \;\; \Im\{s_m\}]^{\mathrm{T}}$. Thus,

Table 5.1: Rotation angles for maximal symbol separation in one-dimension

Modulation Scheme	θ	Modulation Scheme	θ
4-QAM	26.56^0	16-QAM	14.0375^0
8-QAM	26.56^0	64-QAM	7.0^0

the suggested rotation and translation transforms the symmetric modulated symbols $\mathbf{s}^m$ to asymmetric modulated symbols $\mathbf{x}^m$ as

$$\mathbf{x}^m = \mathbf{A}\left(\eta\right)\mathbf{R}\left(\theta\right)\mathbf{s}^m \tag{5.60}$$

where, the rotation matrix $\mathbf{R}$ with rotation angle θ is:

$$\mathbf{R}\left(\theta\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{5.61}$$

Moreover, the translation matrix $\mathbf{A}$ is given in (5.62). It is important to highlight that $\eta \in [0\ 1]$. It can be easily derived using Definition 3; where $\mathcal{C}_x$=$|\tilde{\sigma}_x^2|/\sigma_x^2$, and $\mathcal{C}_x \in [0\ 1]$ with $|\tilde{\sigma}_x^2| = \eta\sigma_x^2$ given translation $\mathbf{x} = \mathbf{A}\left(\eta\right)\mathbf{s}$,

$$\mathbf{A}\left(\eta\right) = \begin{bmatrix} \sqrt{1+\eta} & 0 \\ 0 & \sqrt{1-\eta} \end{bmatrix} \tag{5.62}$$

We aim to design the rotation angle θ and scaling η to dampen error probability by effectively mitigating various HWI effects. Table 5.1 contains the intuitive θ for various modulation schemes to attain maximum separation between any two adjacent symbols in one-dimension and can be introduced as $\mathbf{x} = \mathbf{R}\mathbf{s}$. Next, we propose several design schemes for the translation matrix $\mathbf{A}$.

5.6.1 Special Cases

The existing symmetric signaling schemes induce $\theta = 0$ and $\eta = 0$ which render identity rotational and scaling matrices. The symmetric signaling scheme fails to attain lower error probability in the presence of improper impairments. Therefore, a sub-optimal approach for the asymmetric transmission i.e, maximal asymmetric transmission with $\eta = 1$ can be employed. This η renders the following scaling matrix

$$\mathbf{A} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \tag{5.63}$$

This scheme reduces computational complexity to solve any optimization problem and outperforms the existing symmetric signaling scheme in some scenarios as discussed in Section 5.7. Next, we propose two optimization schemes for the design of optimal translation matrix.

5.6.2 Maximum Pairwise Error Probability Minimization

In this subsection, we design the scaling matrix $\mathbf{A}$ for asymmetric signal transmission to minimize the maximum PEP. The PEP in (5.57) after the requisite transformation in (5.60) can be presented as

$$\Pr(\mathbf{s}^m \rightarrow \mathbf{s}^n | \mathbf{s}^m) = Q\left(\sqrt{\frac{p}{4\left(1-\rho^2\right)} \mathbf{s}_{mn}^{\mathrm{T}} \mathbf{R}^{\mathrm{T}} \mathbf{A}^{\mathrm{T}} \widetilde{\mathbf{H}} \mathbf{Y} \widetilde{\mathbf{H}}^{\mathrm{T}} \mathbf{A} \mathbf{R} \mathbf{s}_{mn}}\right) \tag{5.64}$$

The following optimization problem, **5-P1**, minimize the maximum PEP based on Definition 2.9 for all possible error events between any two different symbols in a given signal

constellation:

$$\begin{aligned}
\textbf{5-P1} : \underset{\eta,\kappa}{\text{minimize}} \quad & \kappa \\
\text{subject to} \quad & \Pr\left(\mathbf{s}^m \to \mathbf{s}^n | \mathbf{s}^m\right) \leq \kappa, \ \forall s^m \neq s^n \ \forall \ m, n \\
& 0 \leq \eta \leq 1.
\end{aligned}$$

The minimization problem **5-P1** can be equivalently reformulated to the following maximization problem by disregarding the Q-function in (5.64) for the ease of implementation,

$$\begin{aligned}
\textbf{5-P2} : \underset{\eta,\tau}{\text{maximize}} \quad & \tau \\
\text{subject to} \quad & \mathbf{s}_{mn}^{\mathrm{T}} \mathbf{R}^{\mathrm{T}} \mathbf{A}^{\mathrm{T}} \widetilde{\mathbf{H}} \mathbf{Y} \widetilde{\mathbf{H}}^{\mathrm{T}} \mathbf{A} \mathbf{R} \mathbf{s}_{mn} \geq \tau, \ \forall \mathbf{s}^m \neq \mathbf{s}^n, \ \forall \ m, n \\
& 0 \leq \eta \leq 1.
\end{aligned}$$

5-P2 is prove to be a concave problem as shown in Appendix G, thus it can be optimally solved using solvers like CVX.

5.6.3 Maximum Symbol Error Rate Minimization

The asymmetric modulation can also be designed based on another optimization criteria, where the maximum SEP has to be minimized. To this end, we employ the SEP bound in (5.15) for the information bearing symbols $\mathbf{s}^m, \ m \in \{1, 2, \ldots, M\}$ as

$$P_{\mathrm{s}} \leq \frac{1}{M} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} \Pr\left(\mathbf{s}^m \to \mathbf{s}^n | \mathbf{s}^m\right). \tag{5.65}$$

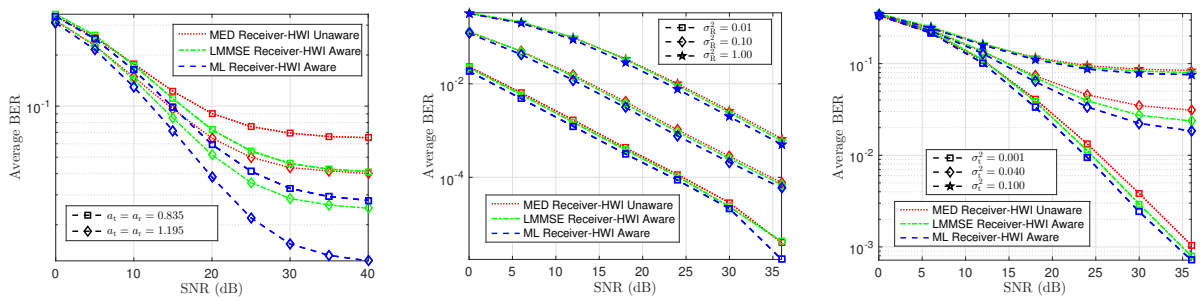
The minimization problem can be formulated as in **5-P3** along with the box constraint on the design parameter η based on Definition 2.9.

$$\textbf{5-P3}: \underset{\eta,\tau}{\text{minimize}} \quad \tau$$
$$\text{subject to} \quad \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{n=1\\n\neq m}}^{M} Q\left(\sqrt{\frac{p}{4\left(1-\rho^2\right)}}\mathbf{s}_{mn}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\tilde{\mathbf{H}}\mathbf{Y}\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{A}\mathbf{R}\mathbf{s}_{mn}\right) \leq \tau \quad \forall \mathbf{s}^m \neq \mathbf{s}^n$$
$$0 \leq \eta \leq 1$$

The problem **5-P3** is proven to be convex in Appendix H and thus it can be optimally solved using any available solver which supports Q-function such as fmincon in MATLAB. The optimal η renders an optimal scaling matrix $\mathbf{A}$ using (5.62). Therefore, symmetric signaling based on the traditional modulation schemes is transformed to asymmetric transmission using (5.60) to effectively combat HWI effects. The proposed asymmetric transmission schemes are compared in the following section while keeping the existing symmetric transmission scheme as a benchmark and an upper-bound on the average error probability of HWIs.

5.7 Numerical Results

In this section, we numerically investigate the BER or symbol error rate (SER) performance of the proposed HWIs system model where IQI results in self-interference. Furthermore, we analyze the performance of the two proposed receivers relative to the conventional MED receiver in the presence of improper additive noise. Moreover, we investigate the tightness of the designed bounds and approximations on the average probability of error. Eventually, we compare various suggested asymmetric transmission schemes to effectively combat HWIs in order to attain lower error probability. The numerical investigations are carried out for the HWI systems under dual IQI with or without transmitter distortion as well as the individual IQI at the transmitter or receiver. Moreover, the results for average



(a) Dual I/Q imbalance (b) Receiver I/Q imbalance. (c) Transmitter I/Q imbalance.

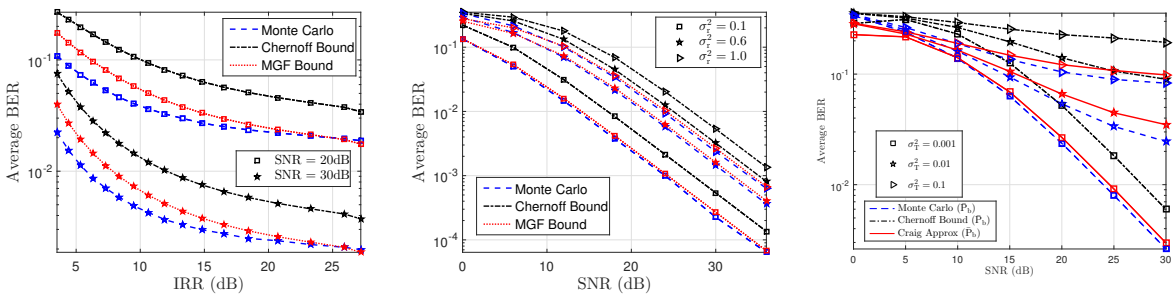
Figure 5.1: HWI-Aware Receivers

error probability are validated only for gray coded QAM modulation scheme for brevity. However, same results can be extended to other modulation schemes.

5.7.1 Optimal Receiver

In this subsection, we analyze the performance gain obtained by employing the proposed HWI-aware optimal ML and sub-optimal LMMSE receivers in-place of the conventional HWI-unaware MED receiver for various system configurations. As for the system parameters, we assume normalized 16-QAM modulation with HWI parameters $a_{\mathrm{t}}{=}a_{\mathrm{r}}{=}\ 0.835$, $\theta_{\mathrm{t}}{=}\theta_{\mathrm{r}}{=}\ 5^0$, $\sigma_{\mathrm{r}}^2 = 1$, $\sigma_{\mathrm{t}}^2 = 0.05$, $\lambda = 1$, the average SNR in 0-36dB range and $\psi = 3$, unless otherwise stated.

Firstly, we analyze the average BER of the HWIs system with dual IQI as well as additive distortions at both transmitter and receiver as shown in Figure 5.1a. We investigate average BER versus average SNR for two different symmetric amplitude IQI estimates $a_{\mathrm{t}}{=}a_{\mathrm{r}}{=}\ 0.835\,\&\,1.195$. Evidently, the impaired system depicts a high error probability especially for the lower value of I/Q amplitude scaling. The increasing SNR helps to improve the average BER performance but eventually undergoes saturation. It is important to highlight that this error floor is due to the presence transmit distortions as discussed in (5.24). The HWI-aware ML and LMMSE receivers clearly outperform the HWI-unaware MED receiver, which was the optimal choice for negligible HWIs. The performance gain is



(a) Dual IQI versus $\Upsilon_{\mathrm{t}} = \Upsilon_{\mathrm{r}}$. (b) Dual IQI for σ_{r}^2. (c) Transmitter IQI for σ_{t}^2.

Figure 5.2: Bounds and Approx on $\bar{P}_{\mathrm{b}}$

particularly significant for higher amplitude errors and higher order SNR. Although the optimal ML receiver demonstrates superior performance relative to the LMMSE receiver for any given SNR and I/Q amplitude error but the LMMSE receiver is a fairly good candidate for least complex practical receiver design.

Secondly, the performance gain of the optimal receiver is observed for a receiver impaired system under receiver IQI and additive distortions at the receiver only. Figure 5.1b represents the average BER versus average SNR for three different receiver distortion levels $\sigma_{\mathrm{r}}^2 = 0.2, 0.6 \ \& \ 1$. Evidently, the drastic effect of receiver HWIs significantly decreases with increasing SNR. Moreover, the HWI-aware optimal ML receiver outperforms suboptimal HWI-aware LMMSE receiver and HWI-unaware MED receiver at all distortion levels, especially at higher SNR. Similarly, the average BER performance of the proposed receivers of the transmitter impaired system is observed for three different transmit distortion levels $\sigma_{\mathrm{t}}^2 = 0.001, 0.04 \ \& \ 0.1$ in Figure 5.1c. The average BER depicts an error floor for significant transmit distortion levels but decreases with increasing SNR for negligible σ_{t}^2.

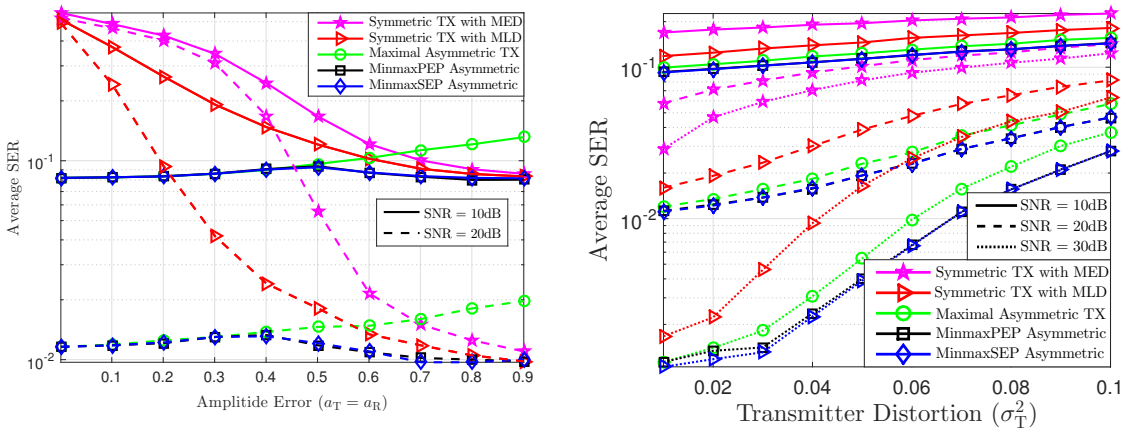
Conclusively, the HWI-aware receivers outperform HWI-unaware receiver in all presented scenarios especially at higher amplitude errors, lower transmitter distortions, and higher SNR. This signifies the importance of incorporating HWIs in the system model and receiver design. Moreover, the optimal ML receiver sets the performance benchmark

for sub-optimal LMMSE receiver. Although the superior performance of ML receiver is unprecedented, but the LMMSE receiver is suitable for practical implementation with subsidiary performance. Additionally, the transmitter distortions are mainly responsible for the error floor at higher SNR levels. Increasing SNR amplifies transmit distortion as well and thus rendering it ineffective to reduce error probability. An important observation is the quantification of the deteriorating effects of various HWIs. Figure 5.1a demonstrates that the dual I/Q impairments and distortions are more drastic than the individual impaired systems in Figure 5.1b and 5.1c. Moreover, the transmitter impairments are far more degrading than receiver impairments.

5.7.2 Bounds and Approximations

This subsection deals with the evaluation of Chernoff bounds and numerical integration approximations on the average probability of error for two system configurations. They include dual I/Q imbalanced system with negligible transmit distortions and transmitter impaired system for brevity. We assume the following system parameters, unless otherwise stated, a_{t}=a_{r}= 0.835, θ_{t}=θ_{r}= 5^0, $\sigma_{\mathrm{r}}^2 = 1$, $\sigma_{\mathrm{t}}^2 = 0$, $\lambda = 1$, $\psi = 3$ for NNUB, normalized 16-QAM modulation scheme.

Firstly, the average BER is analyzed for IRR ranging from 3.5-27dB ($a_{\mathrm{t}} = a_{\mathrm{r}} = 0.2 - 1.0$) in a dual I/Q imbalanced system with receiver distortion but negligible transmitter distortion. The bounds and approximations are observed for two different SNR conditions (20dB and 30dB) in Figure 5.2a. The derived closed form Chernoff bound and the MGF bound are in great agreement with the average BER trend and become tighter with increasing IRR. The MGF bound is remarkably tight in the region of interest when IRR$\geq$20dB. Similarly, Figure 5.2b investigates the similar system with ($a_{\mathrm{t}} = a_{\mathrm{r}} = 0.835$) for three different receiver distortion levels $\sigma_{\mathrm{r}}^2 = 0.1, 0.6 \ \& \ 1.0$. Evidently, lower receiver distortion guarantees lower probability of error. In addition, we observe fairly tight bounds for all distortion and SNR levels. Moreover, the NNUB bound on average BER using



(a) Dual impaired system: SER vs $a_{I/Q}$.

(b) Dual impaired system: SER vs σ_t^2.

Figure 5.3: Asymmetric Transmission Schemes

closed form $\bar{P}_s$ approximation (5.45) closely follows the simulated average BER for IRR = 20dB. Interestingly, the closed form Chernoff bound and Craig's approximation given by (5.41) and (5.44), respectively, can be accurately computed by summing few initial terms exhibiting negligible percentage errors, as given in Table 5.2(a).

Lastly, the derived bounds and approximations are investigated for a transmitter impaired system under transmitter IQI and transmitter additive distortions with thermal noise at the receiver and negligible receiver IQI in Figure 5.2c. The results are analyzed for three different transmit distortion levels $\sigma_t^2 = 0.001, 0.01 \ \& \ 0.1$. The Chernoff bound (5.51) is examined using the proposed Gaussian quadrature numerical integration employing $30 - order$ Legendre polynomial. Analysis depicts that employing lower order $P_N(t)$ exhibits promising results for lower SNR ranges. However, for higher SNR ranges, we need to incorporate a higher order Legendre polynomial to achieve a close bound. Figure 5.2c depicts that employing $30 - order$ $P_N(t)$ follows the same trend as the simulated average probability of error. In order to achieve a close approximation, $30 - order$ $P_K(u)$ and $32 - order$ $P_L(v)$ is employed to numerically integrate the $\bar{P}_s$ approximation intractable double integral in (5.54). The approximation shows promising results especially for medium to high SNR and higher impairment levels.

Table 5.2: Percentage Error of Gaussian Quadrature Legendre Polynomial Approximation

(a) Dual I/Q: Zero-Distortion Transmitter

Sum Terms	Chernoff Bound Equation (5.41)	Craig's Approx Equation (5.45)
1	2.84898e-16	3.41016e-17
2	6.54392e-24	3.28161e-25
3	1.57825e-31	3.44822e-33
4	3.86638e-39	4.97314e-40

(b) Transmitter Impaired System

Order (N, L, K)	Chernoff Bound Equation (5.51)	Craig's Approx Equation (5.54)
1	4.66701e-01	9.51256 e-01
5	1.33278e-03	2.16892e-02
10	8.65744e-05	2.28766e-03
20	1.95911e-06	3.32592e-04

In order to evaluate the accuracy of the proposed Gaussian quadrature numerical integration method for intractable Chernoff bound and Craig's approximation on average SEP, we derived the percentage error for varying ordered Legendre Polynomials as shown in Table 5.2 (b). We assumed the system with negligible receiver IQI as adopted in Figure 5.2c with 25dB SNR, 0.01 transmitter distortion variance and varying Legendre polynomial orders. It can be safely concluded that 10-20 ordered Legendre Polynomials can accurately approximate the intractable integrals with negligible percentage error.

5.7.3 Asymmetric Transmission

The drastic effects of HWIs can be partially mitigated by employing optimal detection scheme. However, we can attain better system performance using asymmetric transmission scheme to effectively dampen the improper interference. This subsection studies the effectiveness of various proposed transmit optimization schemes relative to the traditional

symmetric transmission scheme with or without optimal receiver for various impairments. We assume medium impairment levels of $\theta_{\rm t}$=$\theta_{\rm r}$= 5^0 , $a_{\rm I/Q} = 0.5$, $\sigma_{\rm r}^2 = 1$, $\sigma_{\rm t}^2 = 0.01$, $\lambda = 1$ and normalized 4-QAM modulation scheme, unless otherwise specified, in a dual impaired transmission system.

Firstly, we analyze the average SER for the range of symmetric amplitude errors $a_{\rm I/Q}(a_{\rm t}$= $a_{\rm r}) = 0 - 0.9$ for two given SNR levels; 10 dB and 20 dB in Figure 5.3a. We compare traditional symmetric scheme with or without optimal receiver with the three proposed asymmetric transmission schemes to quantify the error performance gain. Evidently, symmetric transmission with suboptimal receiver performs far worst than the symmetric transmission with the optimal receiver, especially for higher SNR and lower amplitude scaling (higher impairment level). Moreover, the error performance of the adopted system under symmetric transmission improves with increasing $a_{\rm I/Q}$. On contrary, the asymmetric transmission schemes depict a significantly lower error probability even for lower $a_{\rm I/Q}$. The suboptimal asymmetric transmission i.e. maximal asymmetric performs equally good as any other optimized asymmetric transmission for higher impairment levels but renders higher error probability than symmetric transmission for $a_{\rm I/Q} \geq 0.6$. However, the optimal asymmetric transmission schemes MinmaxPEP and MinmaxSEP equally outperform all other transmission schemes. Thus based on this analysis, one can conclude that the application of mere optimal receiver with symmetric transmission does not reduce the error probability to a greater extent but a better performance can be achieved by employing asymmetric transmission schemes along with the optimal receiver. Moreover, Maximal asymmetric transmission is a fairly good choice for highly impaired systems as it performs equally good as any other optimized asymmetric scheme without inducing additional computational overhead.

Furthermore, we investigated the average SER for a viable range of transmitter distortions $\sigma_{\rm t}^2 = 0.01 - 0.1$ for three different SNR values 10 dB, 20 dB and 30 dB in Figure 5.3b. We observe a similar pattern between symmetric signaling with the suboptimal re-

ceiver and symmetric signaling with the optimal receiver for all distortion levels and SNR values. However, a different trend is observed for the maximal asymmetric transmission scheme. It clearly outperforms symmetric signaling scheme at all distortion levels and renders suboptimal performance as compared to the optimized asymmetric transmission schemes (MinmaxPEP and MinmaxSEP). Thus, it depicts a trade-off between the error probability performance gain and computational expenses of the optimization problems.

5.8 Conclusion

In this paper, we demonstrated the significance of incorporating HWIs in accurate system modeling and analysis. We mainly focused on the detailed modeling of IQI and several distortion noises at the transceiver. We also proposed an optimal and a sub-optimal linear receiver which incorporates improper interference characteristics. Both transmitter and receiver IQI render SI information bearing signal whereas only receiver IQI is responsible for transforming AWGN to improper Gaussian noise. The transmitter distortion is subject to channel fading while the receiver added impairment is not. Analytical results are validated using simulations where the derived Chernoff bounds and numerical integration approximations are in close agreement with the simulated average BER trend and are significantly tight for lower impairments levels. Further performance improvement can be achieved using the proposed asymmetric modulation schemes, which outperform the existing symmetric signaling with or without the optimal receiver. Moreover, the maximal asymmetric scheme can be a fairly good candidate to achieve better performance without rendering any optimization expenses for highly impaired systems.

A possible extension to this paper would be to consider the impact of aggregated HWIs on the massive MIMO systems, which is a key concept to attain higher area throughput in future wireless networks. Interestingly, the studies have shown that the huge degrees of freedom offered by the massive densification provide robustness to only some of the impairments. For example, [289] proved that the concentrated antennas deployment offers im-

munity to the hardware distortions but not the phase drifts through closed-form achievable rate performance analysis. Similarly, [290] demonstrated that the effects of impairments and noise at the massive-antenna fusion center vanish while the sensor impairment dominates the achievable distributed detection performance, in the limit of an infinite number of antennas and infinite sensors reporting power budget.

Chapter 6

When Probabilistic Shaping Realizes Improper Signaling for Hardware Distortion Mitigation

HWD render drastic effects on the performance of communication systems. They are recently proven to bear asymmetric signatures; and hence can be efficiently mitigated using IGS, thanks to its additional design degrees of freedom. Discrete AS can practically realize the IGS by shaping the signals' geometry or probability. In this paper, we adopt the PS instead of uniform symbols to mitigate the impact of HWD and derive the optimal maximum a posterior detector. Then, we design the symbols' probabilities to minimize the error rate performance while accommodating the improper nature of HWD. Although the design problem is a non-convex optimization problem, we simplified it using successive convex programming and propose an iterative algorithm. We further present a HS design to gain the combined benefits of both PS and GS. Finally, extensive numerical results and Monte-Carlo simulations highlight the superiority of the proposed PS over conventional uniform constellation and GS. Both PS and HS achieve substantial improvements over the traditional uniform constellation and GS with up to one order magnitude in error probability and throughput.

6.1 Significance and Contributions

In this chapter, we propose PS as a method to realize improper signaling, which is beneficial in mitigating the impact of HWD on the BER performance. Motivated by IGS's theoretical results in various scenarios [2] and the issues associated with GS, such as high shaping gap

and coarse granularity, we adopt PS to realize the IGS scheme and combat HWD to assure reliable communications. In the following, we summarize the main contributions as:

- We derive the optimal MAP detector for a discrete sAS and carry out BER analysis for the adopted HWD communication system.
- We design the probabilistic shaped AS under power and rate constraints for hardware distorted system and propose adaptive algorithm that tune the symbol probabilities for PS to minimize the BER performance.
- We further suggest a hybrid shaped AS scheme that reaps benefits of both PS and GS and present an adaptive algorithm that tune both signal probability and shaping parameters.
- We present two algorithms to find the optimal parameters, i.e., symbol probabilities for PS and additionally shaping parameters for HS design.
- Finally, we present numerical Monte-Carlo simulations to validate the performance of the proposed techniques and compare the BER and throughput performance of PS, GS, and HS in AWGN and Rayleigh fading channels.

6.2 System Description

Impropriety incorporation is crucial for the systems dealing with improper signals, noise, or interference. Such characterization helps in meticulous system modeling, accurate performance analysis, and optimum signaling design. We begin by presenting the transceiver HWD model and the optimal receiver.

6.2.1 Transceiver Hardware Distortion Model

Consider a single-link wireless communication system suffering from various hardware impairments. The non-linear transfer functions of various transceiver RF stages result in

accumulative additive distortion noise as modeled in Chapter 3. We adopt the aggregate HWD model as presented in (3.2).

$$y = \sqrt{\alpha} g x_m + z; \quad m \in \{1, 2, \ldots, M\}, \tag{6.1}$$

where α is the transmitted power, $g \sim \mathcal{CN}(0, \lambda, 0)$ is the slowly varying flat Rayleigh fading channel, x_m is the single-carrier band-pass modulated signal taken from M-ary QAM, M-ary PSK, or M-ary PAM constellation with a probability mass function $p_m \triangleq p_X(x_m)$ rendering the transmission probability of symbol x_m, and $\mathbf{p} \triangleq [p_1, p_2, \cdots, p_M]$. Let us define the set that includes all possible symbol distributions as

$$\mathbb{S} = \left\{ \mathbf{p} : \mathbf{p} = [p_1, p_2, \cdots, p_M], \sum_{j=1}^{M} p_j = 1, p_j \geq 0, \forall j \in \{1, 2, \cdots, M\} \right\}. \tag{6.2}$$

Also, $z \triangleq \sqrt{\alpha} g \eta + w$ is modeled as improper noise, i.e., $z \sim \mathcal{CN}\left(0, \alpha |g|^2 \kappa + \sigma_w^2, \alpha g^2 \tilde{\kappa}\right)$. Moreover, the variance of z_{I} and z_{Q} are given in (6.3) and (6.4), respectively, as

$$\sigma_{\mathrm{I}}^2 = \frac{\alpha |g|^2 \kappa + \sigma_w^2 + \alpha \Re\left(g^2 \tilde{\kappa}\right)}{2}, \tag{6.3}$$

$$\sigma_{\mathrm{Q}}^2 = \frac{\alpha |g|^2 \kappa + \sigma_w^2 - \alpha \Re\left(g^2 \tilde{\kappa}\right)}{2}. \tag{6.4}$$

Furthermore, the non-zero pseudo-variance $\tilde{\sigma}_z^2$ motivates us to evaluate the correlation between z_{I} and z_{Q} using the correlation coefficient ρ_z as

$$\rho_z = \frac{\alpha \Im\left(g^2 \tilde{\kappa}\right)}{\sqrt{\left(\alpha |g|^2 \kappa + \sigma_w^2\right)^2 - \left(\alpha \Re\left(g^2 \tilde{\kappa}\right)\right)^2}}. \tag{6.5}$$

Proof of (6.3)-(6.5) is presented in Appendix I. HWD can leave drastic effects on the system performance as they raise the noise floor. Although, the entropy loss of improper noise is less than the proper noise but it is difficult to tackle. It requires some meticulously

designed improper signaling like IGS for effective mitigation. However, IGS is difficult to implement because of the unbounded peak-to-average power ratio and high detection complexity [2,96]. Therefore, researchers resort to the finite discrete AS schemes obtained by GS.

We propose PS as another way to realize AS in order to effectively dampen the deteriorating effects of improper HWD. PS aims to design non-uniform symbol probabilities for a higher order QAM to minimize BER offering more degrees of freedom and adaptive rates. In the following section, we carry out the error probability analysis of the adopted system which lays foundation for the proposed PS design.

6.2.2 Optimal Receiver

Conventional systems with Gaussian interference employ least-complex receivers with either minimum Euclidean or maximum likelihood detectors. However, such receivers cannot accommodate the unequal symbol probabilities and improper noise. Therefore, the optimal detection in the presented scenario can only be achieved by the MAP detector at the expense of increased receiver complexity. Considering the improper Gaussian HWD and the non-uniform priors of the constellation symbols, the optimal MAP detection is given by

$$\hat{m}_{\mathrm{PS}} = \underset{1\leq m\leq M}{\arg\max} \quad p_X(x_m) f_{Y_{\mathrm{I}},Y_{\mathrm{Q}}|X,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|x_m, g\right), \tag{6.6}$$

$$f_{Y_{\mathrm{I}},Y_{\mathrm{Q}}|X,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|x_m, g\right) = \frac{1}{2\pi\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}\sqrt{1-\rho_z^2}} e^{\left\{ \frac{-1}{2(1-\rho_z^2)} \left[\begin{array}{l} \frac{\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_m)\right)^2}{\sigma_{\mathrm{I}}^2} + \frac{\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_m)\right)^2}{\sigma_{\mathrm{Q}}^2} + \\ -\frac{2\rho_z\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_m)\right)\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_m)\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} \end{array} \right] \right\}}. \tag{6.7}$$

where $f_{Y_{\mathrm{I}},Y_{\mathrm{Q}}|X,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|x_m, g\right)$ is the conditional Gaussian PDF of y representing ML function given x_m and g, as expressed in (6.7) at the top of next page.

6.3 Error Probability Analysis

Considering the non-uniform priors and improper noise, the error probability analysis is carried out based on the optimal MAP detector presented in Section 6.2. Symbol error probability P_{s} is the accumulated error probability of all symbols with respect to their prior probabilities and is given as

$$P_{\mathrm{s}} = \sum_{m=1}^{M} p_m \Pr\left(e|x_m\right), \tag{6.8}$$

where $\Pr\left(e|x_m\right)$ is the probability of an error event given symbol x_m was transmitted. In order to yield a tractable and simplified analysis especially for higher order modulation schemes, P_{s} can be upper bounded as

$$P_{\mathrm{s}} \leq \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n\neq m}}^{M} p_m P_{mn}, \tag{6.9}$$

where, P_{mn} is the pairwise error probability (PEP), which represents the probability of deciding x_n given x_m was transmitted, ignoring all the other symbols in the constellation [209]. The PEP can be evaluated using the MAP rule in (6.6) as

$$P_{mn} = \Pr\left\{p_m\, f_{Y_{\mathrm{I}},Y_{\mathrm{Q}}|X,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|x_m, g\right) \leq p_n f_{Y_{\mathrm{I}},Y_{\mathrm{Q}}|X,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|x_n, g\right)\right\}. \tag{6.10}$$

By substituting the conditional probability from (6.7) in (6.10) and after some mathematical simplifications, the PEP can be written as in (6.11).

$$P_{mn}=\Pr\left\{2\left(1-\rho_z^2\right)\ln\left(\frac{p_m}{p_n}\right)\leq\left[\begin{array}{l}\frac{\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_m)\right)^2-\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_n)\right)^2}{\sigma_{\mathrm{I}}^2}+\frac{\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_m)\right)^2-\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_n)\right)^2}{\sigma_{\mathrm{Q}}^2} \\ +\frac{2\rho_z\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_n)\right)\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_n)\right)-2\rho_z\left(y_{\mathrm{I}}-\sqrt{\alpha}\Re(gx_m)\right)\left(y_{\mathrm{Q}}-\sqrt{\alpha}\Im(gx_m)\right)}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}\end{array}\right]\right\}. \tag{6.11}$$

Now, we find the in-phase and quadrature-phase components of the received signal y for a given transmitted symbol x_m as follows

$$y_{\mathrm{I}} = \sqrt{\alpha}\Re\left(gx_m\right) + z_I, \tag{6.12}$$

and

$$y_{\mathrm{Q}} = \sqrt{\alpha}\Im\left(gx_m\right) + z_Q, \tag{6.13}$$

respectively. Then, we substitute y_{I} and y_{Q} in (6.11), which can be further simplified obtaining,

$$P_{mn} = \Pr\left\{\psi \geq 2\left(1-\rho_z^2\right)\ln\left(\frac{p_m}{p_n}\right) + \alpha\gamma_{mn}\right\}, \tag{6.14}$$

where

$$\gamma_{mn} \triangleq \frac{{\xi_{mn}}_I^2}{\sigma_{\mathrm{I}}^2} + \frac{{\xi_{mn}}_Q^2}{\sigma_{\mathrm{Q}}^2} - \frac{2\rho_z\xi_{mnI}\xi_{mnQ}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}, \tag{6.15}$$

with $\xi_{mn} = g\,d_{mn} = g\left(x_m - x_n\right)$ representing the distance between m^{th} and n^{th} symbol with channel coefficient g, and ψ is obtained by the superposition of z_I and z_Q as

$$\psi{=}2\sqrt{\alpha}\rho_z\left[\left(\frac{\xi_{mnQ}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}{-}\frac{\xi_{mnI}}{\rho_z\sigma_{\mathrm{I}}^2}\right)z_I{+}\left(\frac{\xi_{mnI}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}{-}\frac{\xi_{mnQ}}{\rho_z\sigma_{\mathrm{Q}}^2}\right)z_Q\right]. \tag{6.16}$$

Clearly, ψ is another zero mean Gaussian random variable with variance σ_ψ^2 expressed as

$$\sigma_\psi^2 = 4\left(1-\rho_z^2\right)\alpha\gamma_{mn}. \tag{6.17}$$

Conclusively, P_{mn} is the complementary cumulative distribution function of ψ and is given as

$$P_{mn} = \mathcal{Q}\left(\frac{2\left(1-\rho_z^2\right)\ln\left(\frac{p_m}{p_n}\right) + \alpha\gamma_{mn}}{2\sqrt{\left(1-\rho_z^2\right)\alpha\gamma_{mn}}}\right). \tag{6.18}$$

Substituting the PEP derived in (6.18) to (6.9) along with the gray mapping assumption yields the following bound on BER

$$P_{\mathrm{b}} \leq \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}} \triangleq \frac{1}{\log_2(M)} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} p_m \mathcal{Q}\left(\beta_{mn} \ln\left(\frac{p_m}{p_n}\right) + \frac{1}{2\beta_{mn}}\right), \tag{6.19}$$

where $\beta_{mn} \triangleq \sqrt{1-\rho_z^2}/\sqrt{\alpha\gamma_{mn}}$. The BER expression depends on the size of the constellation, prior probabilities of all the symbols, power budget, mutual distances between the transmitted and received erroneous symbols under Rayleigh fading, and HWD statistical characteristics.

In contrast to the monotonically decreasing BER for the ideal systems, the BER saturates after a specific SNR in the hardware-distorted transceivers. In this regard, we carry out the asymptotic analysis of the bit error probability to quantify the error floor as high SNR. Let us set

$$\Upsilon \triangleq 1 - \frac{\left(\Im\left(g^2\tilde{\kappa}\right)\right)^2}{\left(|g|^4\kappa^2\right) - \left(\Re\left(g^2\tilde{\kappa}\right)\right)^2}, \tag{6.20}$$

the error floor can be upper bounded from (6.19) as in (6.21). We can see that the error floor depends on the adopted M-ary constellation, channel coefficient, HWD statistical characteristics, and symbol probabilities.

$$\lim_{\alpha\to\infty} P_{\mathrm{b}} \leq \frac{1}{\log_2(M)} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} p_m \mathcal{Q}\left(\frac{2\Upsilon \ln\left(\frac{p_m}{p_n}\right) + \left(\frac{2\Re(gd_{mn})^2}{|g|^2\kappa\Re(g^2\tilde{\kappa})} + \frac{2\Im(gd_{mn})^2}{|g|^2\kappa\Re(g^2\tilde{\kappa})} - 2\frac{\Re(gd_{mn})\Im(gd_{mn})\Im\left(g^2\tilde{\kappa}\right)}{\left(|g|^2\kappa\right)^2 - (\Re(g^2\tilde{\kappa}))^2}\right)}{\sqrt{4\Upsilon\left(\frac{2\Re(gd_{mn})^2}{|g|^2\kappa\Re(g^2\tilde{\kappa})} + \frac{2\Im(gd_{mn})^2}{|g|^2\kappa\Re(g^2\tilde{\kappa})} - 2\frac{\Re(gd_{mn})\Im(gd_{mn})\Im(g^2\tilde{\kappa})}{\left(|g|^2\kappa\right)^2 - (\Re(g^2\tilde{\kappa}))^2}\right)}}\right). \tag{6.21}$$

6.4 Proposed Probabilistic Signaling Design

We aim to design the non-uniform symbol probabilities, which minimize the BER of the adopted system suffering from HWD. The optimization is carried out given power and rate

constraints. The rate of the conventional QAM with uniform symbol probabilities and modulation order M_{u} is fixed, i.e., $R = \log_2(M_{\mathrm{u}})$. However, we seek the maximum benefits of PS by allowing a higher-order modulation with $M_{\mathrm{nu}} > M_{\mathrm{u}}$, where M_{nu} is the modulation order of the constellation with non-uniform probabilities $\mathbf{p}$. Thus, the rate of this scheme can be designed such that $R \triangleq \mathrm{H}(\mathbf{p}) \geq \log_2(M_{\mathrm{u}})$, rendering more design flexibility and hence is capable of reducing the BER. PS is capable of changing the transmission rate by changing the symbol distribution for a fixed modulation order, unlike uniform signaling, which needs to change the modulation scheme's order to change the rate for uncoded communications.

After designing the symbol probabilities, we can implement PS by using distribution matching at the transmitter to map uniformly distributed input bits to M_{nu}-QAM/PSK symbols [102,103,106]. Moreover, they can be detected using the proposed MAP detector (6.6) at the receiver that incorporates the prior symbol distribution. In the following, we formulate the PS design problem and propose an algorithm to obtain the non-uniform symbol probabilities followed by some toy examples.

6.4.1 Problem Formulation

The probability vector $\mathbf{p} \triangleq [p_1, p_2, \ldots, p_{M_{\mathrm{nu}}}]$, containing probabilities of the symmetric $M_{\mathrm{nu}}-$QAM/PSK modulated symbols with $M_{\mathrm{nu}} > M_{\mathrm{u}}$ [1], is designed to minimize the upper bound on the BER derived in (6.19). In particular, we formulate the problem as

$$\textbf{6-P1}: \underset{\mathbf{p}\in\mathbb{S}}{\text{minimize}} \quad \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}(\mathbf{p}) \tag{6.22a}$$

$$\text{subject to} \quad \sum_{m=1}^{M_{\mathrm{nu}}} |x_m|^2 p_m \leq 1, \tag{6.22b}$$

$$\mathrm{H}(\mathbf{p}) \geq \log_2(M_{\mathrm{u}}), \tag{6.22c}$$

[1] For $M_{\mathrm{nu}} = M_{\mathrm{u}}$, the distribution should be uniform to satisfy the rate constraint because uniform signaling has the largest entropy.

where (6.22b) and (6.22c) represent the average power and rate constraints, respectively, and $\mathrm{H}(\mathbf{p})$ is the source entropy, which represents the transmitted rate in terms of bits per symbol per channel use and is defined as

$$\mathrm{H}(\mathbf{p}) \triangleq \sum_{m=1}^{M_{\mathrm{nu}}} -p_m \log_2 (p_m) . \tag{6.23}$$

The concave nature of information entropy in (6.22c) renders a convex constraint in $\mathbf{p}$ and the rate fairness is justified based on the trade off between BER minimization and rate maximization, while satisfying a minimum rate. Therefore, the idea is to employ a higher order non-uniformly distributed $M_{\mathrm{nu}}-$QAM/PSK as compared to a lower order uniformly distributed $M_{\mathrm{u}}-$QAM/PSK with same energy and at least the same rate to minimize BER.

6.4.2 Optimization Framework

The optimization problem **6-P1** (6.22) is a non-convex optimization problem owing to the non-convex objective function even though all the constraints are convex. Therefore, we propose successive convex approximation approach to tackle it. We begin by approximating $\mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}(\mathbf{p})$ with its first order Taylor series approximation. First order Taylor series approximation of a function $f(x)$ around a point $x^{(k)}$ is given as

$$\tilde{f}\left(x, x^{(k)}\right) \approx f\left(x^{(k)}\right) + \nabla_x f\left(x^{(k)}\right)\left(x - x^{(k)}\right) . \tag{6.24}$$

Thus, we need to compute $\nabla_{\mathbf{p}} \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}$ and evaluate it at $\mathbf{p}^{(k)}$ to compute $\tilde{\mathrm{P}}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}, \mathbf{p}^{(k)}\right)$.

$$\nabla_{\mathbf{p}} \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}} = \left[\frac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial p_1} \quad \frac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial p_2} \quad \cdots \quad \frac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial p_{M_{\mathrm{nu}}}}\right] . \tag{6.25}$$

Algorithm 4 Successive Convex Programming

Initialize $i \leftarrow 0$, $\epsilon \leftarrow \infty$ and **Set** tolerance δ
Choose feasible starting point $\mathbf{p}^{(i)}$
while $\epsilon \geq \delta$ **do**
    Evaluate $\tilde{\mathrm{P}}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}, \mathbf{p}^{(i)}\right)$
    Solve **6-P1a** and obtain $\mathbf{p}$ using $\mathbf{p}^{(i)}$
    $\mathbf{p}^{(i+1)} \leftarrow \mathbf{p}$
    Update $\epsilon \leftarrow \left\|\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)}\right\|$
    $i \leftarrow i + 1$
end while
$\mathbf{p}^{*} \leftarrow \mathbf{p}^{i+1}$
$\mathrm{P}_{\mathrm{b}}^{*} \leq \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathcal{P}^{*}\right)$

In order to compute $\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}} / \partial p_t$, we rewrite (6.19) as

$$\mathrm{P}_{\mathrm{b}}^{\mathrm{UB}} = \frac{1}{\log_2\left(M_{\mathrm{nu}}\right)} \sum_{m=1}^{M_{\mathrm{nu}}} \sum_{\substack{n=1 \\ n \neq m}}^{M_{\mathrm{nu}}} p_m \int_{\Omega_{mn}}^{\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du, \tag{6.26}$$

where

$$\Omega_{mn} = \beta_{mn} \ln\left(\frac{p_m}{p_n}\right) + \frac{1}{2\beta_{mn}}. \tag{6.27}$$

From (6.26) and by applying the Leibniz integral rule, we get

$$\frac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial p_t} \leq \frac{1}{\log_2\left(M_{\mathrm{nu}}\right)} \sum_{\substack{n=1, \\ n \neq t, \\ m=t}}^{M_{\mathrm{nu}}} \left(\mathcal{Q}\left(\Omega_{mn}\right) - \frac{\beta_{mn}}{\sqrt{2\pi}} e^{-\frac{\Omega_{mn}^2}{2}}\right) + \frac{1}{\log_2\left(M_{\mathrm{nu}}\right)} \sum_{\substack{m=1, \\ m \neq t, \\ n=t}}^{M_{\mathrm{nu}}} \frac{\beta_{mn} p_m}{\sqrt{2\pi} p_n} e^{-\frac{\Omega_{mn}^2}{2}}. \tag{6.28}$$

Now, $\mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}$ can be approximated from (6.24), (6.25), and (6.28) using first order Taylor series expansion around an initial probability vector $\mathbf{p}^{(k)}$ as

$$\tilde{\mathrm{P}}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}, \mathbf{p}^{(k)}\right) \triangleq \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}^{(k)}\right) + \nabla_{\mathbf{p}} \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}^{(k)}\right)\left(\mathbf{p} - \mathbf{p}^{(k)}\right). \tag{6.29}$$

Successive convex programming minimizes **6-P1** by iteratively solving its convex approx-

imation **6-P1a** as presented in Algorithm 4.

$$\textbf{6-P1a}: \underset{\mathbf{p}\in\mathbb{S}}{\text{minimize}} \quad \tilde{\mathrm{P}}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}, \mathbf{p}^{(k)}\right) \tag{6.30a}$$

$$\text{subject to} \quad \sum_{m=1}^{M_{\mathrm{nu}}} |x_m|^2 p_m \leq 1, \tag{6.30b}$$

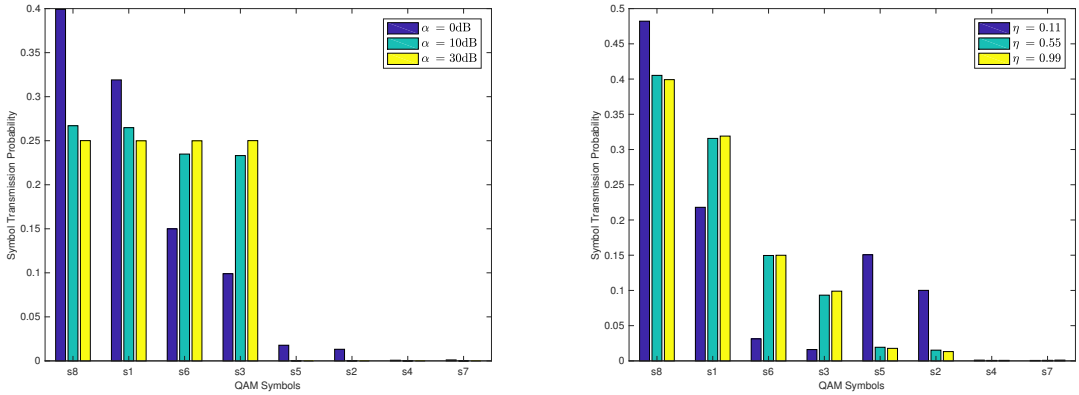
$$\mathrm{H}(\mathbf{p}) \geq \log_2\left(M_{\mathrm{u}}\right), \tag{6.30c}$$

It begins with the initiation of counter i, stopping criteria ϵ and the stopping threshold δ. Secondly, we choose some feasible PMF set $\mathbf{p}^{(i)} \in \mathbb{S}$ which satisfies the constraints (6.22b) and (6.22c). The while loop starts by evaluating the approximation $\tilde{\mathrm{P}}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathbf{p}, \mathbf{p}^{(i)}\right)$ around $\mathbf{p}^{(i)}$.

The convex problem **6-P1a** is solved using the Karush Kuhn Tucker (KKT) conditions derived in Appendix K to obtain the optimal probabilities for **6-P1a** [282]. The solution obtained in this iteration is updated as $\mathbf{p}^{(i+1)}$ and is used to evaluate the stopping criteria $\epsilon \leftarrow \left\|\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)}\right\|$ as shown in Algorithm 4. The loop ends when the change in two subsequent solution parameters in terms of the ℓ_2 norm is less than a predefined threshold δ. Once the stopping criteria is attained, the solution parameters $\mathbf{p}^{(*)}$ are guaranteed to render a BER $\mathrm{P}_{\mathrm{b}}^{*}$ which will be lower than the bound $\mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}\left(\mathcal{P}^{*}\right)$.

6.4.3 Toy Examples

A comprehensive illustration of probabilistically shaped $M_{\mathrm{nu}} = 8$-QAM with a 2 bits/symbol rate constraint, corresponding to $M_{\mathrm{u}} = 4$, is presented in Fig. 6.1a and Fig. 6.1b. The relation between prior probabilities and different SNR values is presented in Fig. 6.1a. Clearly, the probability distribution is quite random for lower SNR level such as $\alpha = 0$ dB. However, it starts adopting uniform distribution of 0.25 for four of it's symbols, i.e., s1, s3, s6, and s8 while zero probabilities for the rest four symbols. This technique provides lower BER while maintaining 2 bits/symbol rate for a fair comparison with traditional 4-QAM.



(a) Three SNR levels at $\eta = 0.99$

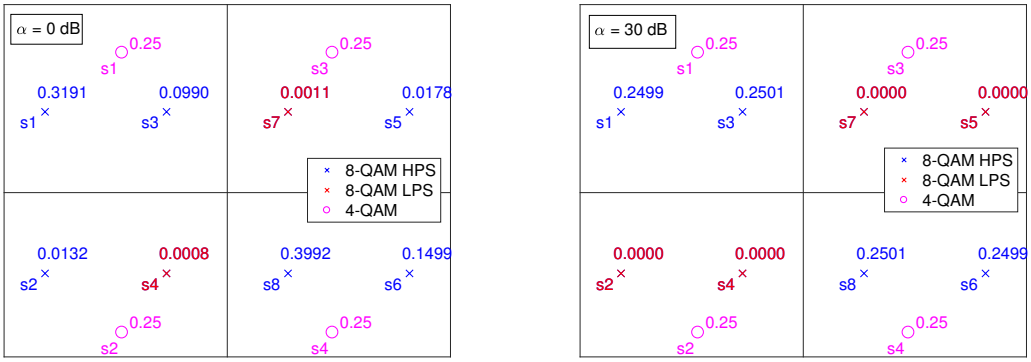
(b) Three HWD levels at $\alpha = 0$ dB

Figure 6.1: 8-QAM probability distribution at Rate = 2 bits/symbol.

Interestingly, it achieves a lower BER by transmitting half of the symbols which are not the nearest neighbors. It is important to highlight that the proposed approach achieves this performance with the same power budget and transmission rate.

Another example illustrates the trend of probabilitic shaping for 8-QAM constellation at lower SNR level (keeping in mind that it assigns the uniform probabilities to four symbols at high SNR levels). The trend for lower HWD level such as $\eta = 0.11$ is quite random. However, it follows a decreasing probability trend for middle to higher HWD levels. Intuitively, it assigns higher probabilities to the symbols with least power and lower probabilities to the symbols with higher powers. This trend decreases the BER while maintaining the average power constraint.

It is interesting to visualize the corresponding symbol constellations for both $\alpha = 0$ dB and $\alpha = 30$ dB. For $\alpha = 0$ dB, probabilistic shaped 8-QAM designates six symbols with significant transmission probabilities as highly probable symbol (HPS) whereas renders two symbols as least probable symbols (LPS) as depicted in Fig. 6.2a. On the other hand, PS at $\alpha = 30$ dB only resorts to transmitting four of it's symbols, i.e., s1, s3, s6, and s8 (HPS) and discards the rest as depicted in Fig. 6.2b. Notably, this technique assigns lowest probabilities to the symbols which are mostly affected by the highly improper noise in first



(a) PS 8-QAM versus NS 4-QAM at $\alpha = 0$ dB

(b) PS 8-QAM versus NS 4-QAM at $\alpha = 30$ dB

Figure 6.2: 8-QAM probability distribution at Rate = 2 bits/symbol.

and third quadrant. It is important to emphasize that the distortion power is proportional to the transmit power. This strengthens distortions at high SNR and leads to the negligible transmission probabilities for the highly affected symbols. Hence, it is capable of achieving lower BER while maintaining 2 bits/symbol rate for a fair comparison with traditional 4-QAM.

6.5 Hybrid Shaping where Conventional meets State-of -the-Art

In this section, we increase the AS design flexibility by allowing joint GS and PS, which we call it here HS, to improve the underlying communication system performance further. Throughout the design procedure, HS transforms the equally spaced uniformly distributed QAM/PSK symbols to unequally spaced symbols in a geometric envelope with non-uniform prior distribution. Thus, HS aims to optimize the symbol probabilities (i.e., PS) and some spatial shaping parameters for the constellation (i.e., GS).

6.5.1 Hybrid Shaping Parameterization

Apart from the non-uniform priors, consider the asymmetric transmit symbol $\mathbf{v}_m = [v_{mI} \quad v_{mQ}]^{\mathrm{T}}$ resulting from the GS on the conventional baseband symmetric M-QAM/M-PSK symbol

$\mathbf{x}_m = [x_{mI} \quad x_{mQ}]^{\mathrm{T}}$ i.e., $\mathbf{v}_m = \mathbf{A}\mathbf{R}\mathbf{x}_m$ as discussed in Chapter 5. The rotation matrix $\mathbf{R}(\theta)$ and translation matrix $\mathbf{A}(\zeta)$ are given in (5.61) and (5.62), respectively, with translation parameter $\zeta \in (0,1)$ and rotation angle $\theta \in (0, \mu\,\pi/2)$ for some constant μ. Uniformly distributed symmetric M-QAM constellation has a rotation symmetry of $n\,\pi/2, n \in \mathcal{Z}^+$ rendering $\mu = n$ to be good choice for GS. However, non-uniformly distributed M-QAM constellation can only be rotationally symmetric after $2n\pi$, thus $\mu = 4n$ is suitable for HS. This technique renders non-uniformly spaced symbols in a parallelogram envelop. It is important to highlight that this transformation preserves the power requirement. Power invariance of the rotation is a well known fact in the literature [209]. However, the wisdom behind the structure of $\mathbf{A}(\zeta)$ is unfolded in the following theorem.

Remark 1. *GS parameterization using translation matrix $\mathbf{A}(\zeta)$ preserves the power invariance of a complex random variable and inculcates asymmetry/improperness with the circularity coefficient ζ.*

Proof. The proof is presented in Appendix J. Furthermore, the generalization of the same concept to the symmetric discrete constellations such as M-QAM and M-PSK is also described in Appendix J. □

6.5.2 Optimal Receiver

The optimal receiver for hybrid shaped AS is also a MAP detector as derived in (6.6), but with a modified reference constellation v_m in place of x_m for all $m \in \{1, 2, \cdots, M_{\mathrm{nu}}\}$. More precisely, the detected symbol, $\hat{m}_{\mathrm{HS}}$, is the one that maximizes the posterior distribution, i.e.,

$$\hat{m}_{\mathrm{HS}} = \underset{1 \le m \le M_{\mathrm{nu}}}{\arg\max} \quad p_V(v_m) f_{Y_{\mathrm{I}}, Y_{\mathrm{Q}}|V,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|v_m, g\right), \tag{6.31}$$

where, $f_{Y_{\mathrm{I}}, Y_{\mathrm{Q}}|V,g}\left(y_{\mathrm{I}}, y_{\mathrm{Q}}|v_m, g\right)$ is similar to (6.7) by replacing all appearances if x_m with v_m for all $m \in \{1, 2, \cdots, M_{\mathrm{nu}}\}$. It is worth noting that non-uniform prior probabilities are inculcated in the detection process using MAP detector in place of ML detector. Moreover,

the geometrically shaped symbols are taken from a modified symbol constellation. Hence, this requires updating the reference constellation for appropriate detection.

6.5.3 Error Probability

HS follows the same BER bound as derived in (6.19) but with modified γ_{mn}. It can now be written using the following quadratic formulation as a function of ζ and θ.

$$\gamma_{mn}\left(\zeta,\theta\right)=\mathbf{x}_{mn}^{\mathrm{T}}\mathbf{R}\left(\theta\right)^{\mathrm{T}}\mathbf{A}\left(\zeta\right)^{\mathrm{T}}\mathbf{G}\mathbf{A}\left(\zeta\right)\mathbf{R}\left(\theta\right)\mathbf{x}_{mn}, \tag{6.32}$$

where $\mathbf{x}_{mn}$ is the real composite vector form of $\xi_{mn}=gd_{mn}$ given as $\mathbf{x}_{mn}=\begin{bmatrix}\xi_{mn\mathrm{I}} & \xi_{mn\mathrm{Q}}\end{bmatrix}^{T}$ and $\mathbf{G}$ contains the statistical characteristics of the aggregate noise including in-phase noise variance, quadrature-phase noise variance, and the correlation between these components.

$$\mathbf{G}=\begin{bmatrix}\frac{1}{\sigma_I^2} & \frac{-\rho_z}{\sigma_I\sigma_Q}\\ \frac{-\rho_z}{\sigma_I\sigma_Q} & \frac{1}{\sigma_Q^2}\end{bmatrix}. \tag{6.33}$$

Thus, the BER of HS can be upper bounded as

$$\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}\left(\mathbf{p},\zeta,\theta\right)=\frac{1}{\log_2\left(M_{\mathrm{nu}}\right)}\sum_{m=1}^{M_{\mathrm{nu}}}\sum_{\substack{n=1\\ n\neq m}}^{M_{\mathrm{nu}}}p_m\mathcal{Q}\left(\frac{\sqrt{1-\rho_z^2}}{\sqrt{\alpha\gamma_{mn}\left(\zeta,\theta\right)}}\ln\left(\frac{p_m}{p_n}\right)+\frac{\sqrt{\alpha\gamma_{mn}\left(\zeta,\theta\right)}}{2\sqrt{1-\rho_z^2}}\right). \tag{6.34}$$

6.5.4 Problem Formulation

HS targets the joint design of PS PMF $\mathbf{p}$ and GS parameters involving translation ζ and rotation θ parameter to minimize the BER bound given in (6.34).

$$\textbf{6-P2}: \underset{\substack{\mathbf{p}\in\mathbb{S}, 0\leq\zeta\leq 1,\\ 0\leq\theta\leq 2\pi}}{\text{minimize}} \quad \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}\left(\mathbf{p},\zeta,\theta\right) \tag{6.35a}$$

$$\text{subject to} \quad \sum_{m=1}^{M_{\mathrm{nu}}} |v_m|^2 p_m \leq 1, \tag{6.35b}$$

$$\mathrm{H}(\mathbf{p}) \geq \log_2\left(M_{\mathrm{u}}\right), \tag{6.35c}$$

where the average power constraint (6.22b) is updated as (6.35b) to account for the possible change in the power of the symbols by geometrically shaping the constellation. However, the proposed rate constraint (6.35c) remains intact. Additionally, there are some boundary constraints on ζ and θ, respectively.

Intuitively, it is quite difficult to tackle this non-convex multimodal joint optimization problem. Therefore, we resort to the alternate optimization of PS parameters ($\mathbf{p}$) and GS parameters (ζ,θ) using sub-problems **6-P2a** and **6-P2b**, respectively. Problem **6-P2a** designs the PS parameters for some given ζ and θ. It is quite similar to the problem **6-P1** and thus, can be solved using Algorithm 4.

$$\textbf{6-P2a}: \underset{\mathbf{p}\in\mathbb{S}}{\text{minimize}} \quad \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}\left(\mathbf{p},\zeta,\theta\right) \tag{6.36a}$$

$$\text{subject to} \quad (6.35\text{b}), (6.35\text{c}). \tag{6.36b}$$

On the other hand, the GS optimization problem designs ζ and θ for fixed symbol probabilities $\mathbf{p}$, given as

$$\textbf{6-P2b}: \underset{\substack{0\leq\zeta\leq 1,\\ 0\leq\theta\leq 2\pi}}{\text{minimize}} \ \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}\left(\mathbf{p},\zeta,\theta\right). \tag{6.37}$$

The optimization problem **6-P2b** is a multimodal non-convex problem which is hard to tackled even by the SCP approach as employed in Section 6.4. The difficulty arises due to the absence of any constraints which restrict the feasibility region. The feasibility space enclosed by the boundary constraints is highly insufficient to serve our purpose. Therefore, we can approximate the solution using any of the following two methods

- Trust region reflective method: This method defines a trust region around a specific initial point and then approximate the function within that region. The convex approximation is the first order Taylor series approximation using the gradient. It begins by minimizing convex approximation of the function to obtain a solution. This solution is the perturbation in the initial point rendering a new point which should minimize the original function. Otherwise, we need to shrunk the trust region and repeat the process. Reflections are used to increase the step size while satisfying box constraints. After each iteration, we receive a new point which renders a lower objective function than the initial point. This iterative approach leads us to a local minimum and stops when some specified stopping criterion are met [291, 292].

- Gradient descent: This method is a relatively faster approach to tackle the problem at hand. It is owing to the fact that it does not involve any approximation and underlying optimization. It begins with an initial point and keeps updating the point in the descent direction using the gradients and a step size until it reaches a local solution or satisfies some stopping criterion [282].

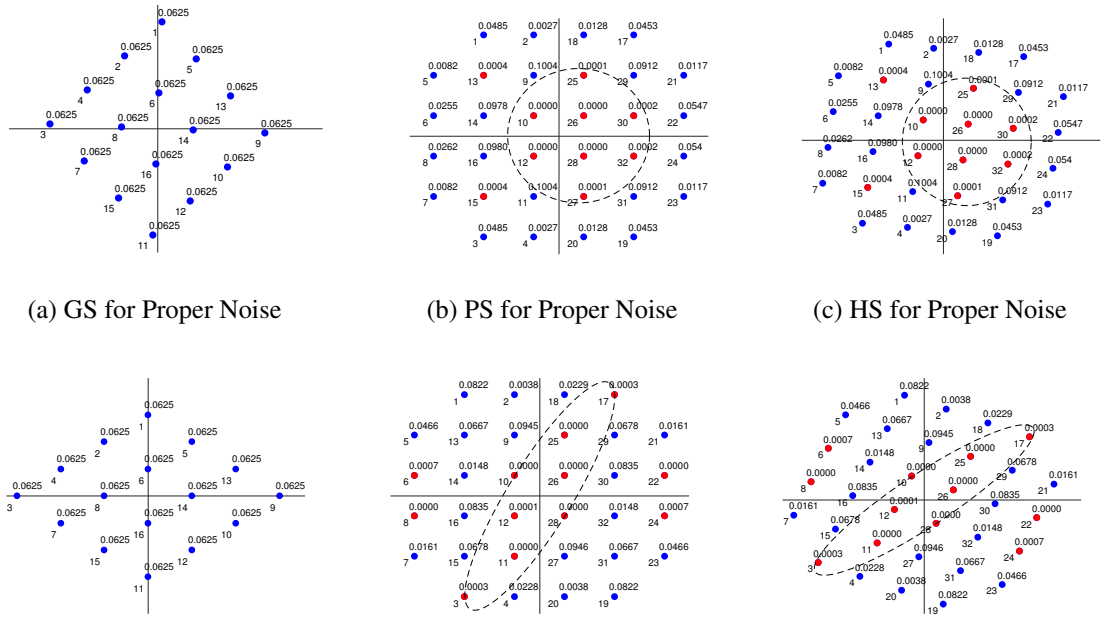
Interestingly, both of these methods require the gradients of $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}(\mathbf{p}, \zeta, \theta)$ with respect to ζ and θ. Gradients are used either to approximate the function with it's first order Taylor series approximation within a trust region or to find the next point in the descent direction. The gradients are evaluated and presented in Appendix L.

Algorithm 5 Alternate Optimization

1: **Initialize** $j \leftarrow 0$, $\epsilon \leftarrow \infty$ and **Set** tolerance δ
2: **Choose** feasible starting points $\mathbf{p}^{(j)}$, $\zeta^{(j)}$, and $\theta^{(j)}$.
3: **Evaluate** $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j)}\left(\mathbf{p}^{(j)}, \zeta^{(j)}, \theta^{(j)}\right)$.
4: **while** $\epsilon \geq \delta$ **do**
5: Solve **6-P2a** using Algorithm 4 with starting point $\mathbf{p}^{(j)}$ and given $\zeta^{(j)}$, $\theta^{(j)}$ to obtain $\mathbf{p}^{(j*)}$
6: Solve **6-P2b** with starting points $\zeta^{(j)}$,$\theta^{(j)}$ and given $\mathbf{p}^{(j*)}$ to obtain $\zeta^{(j*)}$, $\theta^{(j*)}$
7: $\mathbf{p}^{(j+1)} \leftarrow \mathbf{p}^{(j*)}$, $\zeta^{(j+1)} \leftarrow \zeta^{(j*)}$, and $\theta^{(j+1)} \leftarrow \theta^{(j*)}$
8: Evaluate $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j+1)}\left(\mathbf{p}^{(j+1)}, \zeta^{(j+1)}, \theta^{(j+1)}\right)$.
9: Update $\epsilon \leftarrow \left\|\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j+1)} - \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j)}\right\|$
10: $j \leftarrow j + 1$
11: **end while**
12: Solution parameters: $\mathbf{p}^* \leftarrow \mathbf{p}^{j+1}$, $\zeta^* \leftarrow \zeta^{j+1}$, $\theta^* \leftarrow \theta^{j+1}$
13: Objective function: $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}*} \leftarrow \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j+1)}$
14: Consequence: $\mathrm{P}_{\mathrm{b,HS}}^{*} \leq \mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}*}$

6.5.5 Proposed Algorithm

The joint optimization problem **6-P2** can be tackled using the alternate optimization algorithm as presented in Algorithm 5. It solves the sub problems **6-P2a** and **6-P2b** alternately and iteratively. It begins with some starting feasible points $\mathbf{p}^{(j)}$, $\zeta^{(j)}$, and $\theta^{(j)}$ and evaluates $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j)}\left(\mathbf{p}^{(j)}, \zeta^{(j)}, \theta^{(j)}\right)$ as a benchmark. The alternate optimization begins by solving **6-P2a** to minimize $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}}$ with respect to $\mathbf{p}$ given a pair of ζ and θ. It is achieved by replacing all entries of x_m with $v_m = \mathbf{AR}x_m \ \forall m$. $\mathbf{p}^{(j*)}$ is obtained using the framework provided in Algorithm 4 which solves **6-P1a** iteratively. Then, the optimum $\mathbf{p}^{(j*)}$ is used as a given PMF to obtain the pair $\zeta^{(j*)}$ and $\theta^{(j*)}$ by solving **6-P2b**. These optimum parameter values are updated to attain next initial points. Moreover, $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}(j+1)}\left(\mathbf{p}^{(j+1)}, \zeta^{(j+1)}, \theta^{(j+1)}\right)$ is also evaluated to compare the decrease in objective function. The norm of this difference is stored in ϵ and the process is repeated until this value drops below a preset threshold δ. Eventually, the solution parameters are updated in $(\mathbf{p}^*, \zeta^*, \theta^*)$ which yield the minimized BER upper bound $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}*}$ using HS. Therefore, these HS parameters are capable of rendering a BER $\mathrm{P}_{\mathrm{b,HS}}^{*}$ lower than the bound $\mathrm{P}_{\mathrm{b,HS}}^{\mathrm{UB}*}$.



(a) GS for Proper Noise (b) PS for Proper Noise (c) HS for Proper Noise

(d) GS for Improper Noise (e) PS for Improper Noise (f) HS for Improper Noise

Figure 6.3: Different asymmetric signaling designs for proper and improper aggregate interference

Numerical evaluations reveal that the stopping criteria is mostly met in just one iteration. Interestingly, Step 5 and 6 in Algorithm 5 are interchangeable and need to be chosen carefully. For instance, PS demonstrates better performance at higher HWD levels so it is intuitive to design the HS by first PS and then GS in order to attain further gain over PS. Whereas, GS depicts lower BER at lower HWD levels so it is recommended to design HS by first GS and then PS in order to achieve better performance than GS using the added DoF offered by PS.

HS can be implemented by choosing the transmit symbols for the translated and rotated signal constellation, i.e., $v_m = \mathbf{A}\left(\zeta^*\right)\mathbf{R}\left(\theta^*\right)x_m$. Furthermore, the symbols are transmitted according to the optimized $\mathbf{p}^*$ where ζ^*, θ^* and $\mathbf{p}^*$ are designed using Algorithm 5. Upon reception, they are detected using the MAP detector as presented in (6.31).

6.5.6 Illustrative Example

We present a comprehensive example to highlight the design of various distinct asymmetric constellations for a fixed rate of 4 bits/symbol. The red color visualizes constellation symbols with quite low probabilities. Fig. 6.3 represents different AS schemes assuming either proper $C_\eta = 0$ or highly improper $C_\eta = 0.9$ HWDs, respectively. We use 16-QAM for GS whereas 32-QAM for both PS and HS. The shaping parameters are designed/optimized for a system suffering from high HWD, i.e., $\kappa = 0.99$ at 30dB SNR. Fig. 6.3a and 6.3d illustrate equally prior geometrically shaped constellation symbols in the presence of proper and maximally improper noise, respectively. Fig. 6.3a is a mere rotation of the original 16-QAM in the presence of proper HWDs whereas Fig. 6.3d also inculcates the translation rendering a squeezed parallelogram envelop in vertical axis. Next, probabilistic shaped constellations are presented in Fig. 6.3b and 6.3e for proper and maximally improper distortions, respectively. Evidently, the formation of red symbols around the origin transforms from a symmetric circle in Fig. 6.3b to an ellipse in 1st and 3rd quadrant in Fig. 6.3e corresponding to the respective symmetric and asymmetric noise. This reveals the reason behind superior performance of PS as it is capable of assigning negligible transmission probabilities to the symbols which are mostly affected by the aggregate noise as per it's proper/improper characteristics. Furthermore, this probabilistic shaped constellation undergoes GS to obtain hybrid shaped QAM constellation as shown in Fig. 6.3c and 6.3f under proper and maximally improper noise, respectively. This transformation allows the constellation to align itself as per the underlying noise characteristics and further improves the system performance.

6.6 Numerical Results

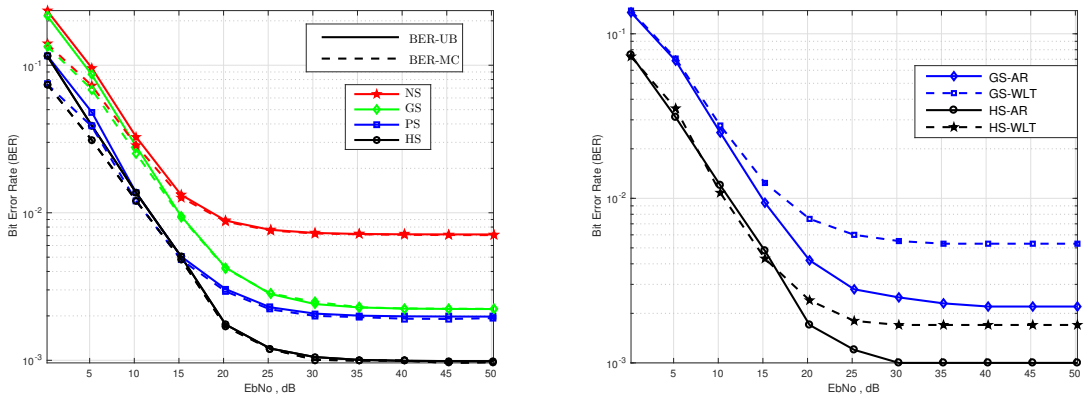
Numerical evaluations of the adopted HWD system are carried out to study the drastic effects of hardware imperfections and the effectiveness of the mitigation strategies. The

performance of the proposed asymmetric transmission schemes PS and HS is quantified as opposed to the benchmark NS and conventional GS, with varying EbNo and HWD levels. EbNo is obtained by normalizing SNR with the transmission rate. Moreover, GS can be implemented by transmitting symbols from a reshaped constellation $\mathbf{v}_m = \mathbf{A}\left(\zeta^*\right)\mathbf{R}\left(\theta^*\right)\mathbf{x}_m$, where ζ^* and θ^* can be obtained by solving **P2a** given uniform prior distribution. Upon reception, they are detected using the ML detector which is the simplified form of optimal MAP detector (6.31) given uniform prior probabilities. This ML detector considers the reshaped constellation symbols $\mathbf{v}_m$ as the reference to detect the received symbols.

For most of the numerical evaluations we assume Gray coded square QAM constellations of order $M_{\mathrm{u}} = 8$, i.e., $R = \log_2(M_{\mathrm{u}})$, for NS and GS as benchmarks. For PS and HS we employ $M_{\mathrm{nu}} = 32$-QAM with rate at least as high as that of GS, i.e., $R \geq \log_2(M_{\mathrm{u}})$. Moreover, we consider practical HWD values for the transmitter $\kappa_t = 0.01$ and receiver $\kappa_r = 0.12$. The pseudo-variances are derived from the $\tilde{\kappa}_{\mathrm{tI}} = \kappa_t/4$, $\tilde{\kappa}_{\mathrm{rI}} = \kappa_r/4$, and correlation coefficient $\rho_\eta = 0.9$. Intuitively, AWGN channel assumes $g = 1$ and circularly symmetric Rayleigh fading channel is generated using $\lambda = 1$. Furthermore, the transmission EbNo is taken as 30 dB. The aforementioned values of the parameters are used throughout the numerical results, unless specified otherwise.

First, we evaluate the performance of various AS schemes for a range of EbNo from 0 dB to 50 dB in an AWGN channel as shown in Fig. 6.4a. We employ M_{u}-QAM for NS and GS whereas M_{nu}-QAM for PS and HS. The BER upper bound (BER-UB) of PS and HS are given by (6.19) and (6.34), respectively, whereas the BER-UB of NS and GS are derived from (6.19) and (6.34) by assuming uniform distribution, respectively.

The BER performance improves with increasing EbNo till 30 dB and then undergoes saturation owing to the presence of HWD. Further increase in bit energy also results in an increase in the distortion variance, as the system experiences an error floor which can be deduced from (6.21). Evidently, the proper/symmetric QAM is suboptimal and the BER performance is significantly improved using AS. Conventional GS is not beneficial at lower



(a) BER performance of AS schemes

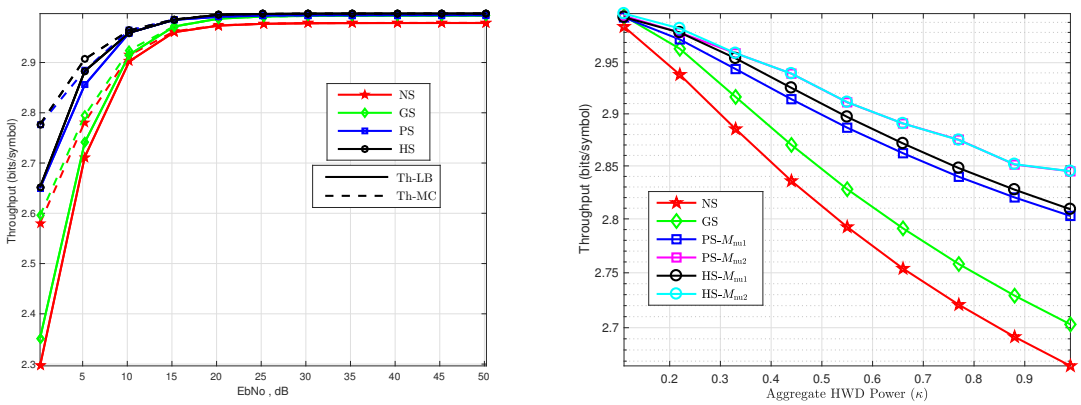
(b) Comparison of two GS Schemes

Figure 6.4: BER Performance for a range of EbNo in AWGN channel.

EbNo values, but it significantly improves the performance for higher EbNo values pertaining to the increased symbol space [98]. On the other hand, the proposed PS is capable of minimizing the BER for the entire range of EbNo. Substantial gains can be achieved by taking another step forward and employing HS. Therefore, we can safely conclude that the best performance can be achieved using PS for EbNo ≤ 15 dB and HS for EbNo ≥ 15 dB. At 20 dB, the BER reductions for GS, PS, and HS schemes with respect to unshaped constellation are approximately 52.22%, 66.67%, 80%, respectively. The numerical results in Fig. 6.4a depict close accordance between the derived BER-UBs and the corresponding MC performance of the various transmission schemes.

For the same system settings, we compare two different parameterization techniques to achieve asymmetric GS, which is a building block of HS, in Fig. 6.4b. The GS-AR scheme represents our proposed GS scheme based on the optimal translation A and rotation R. This scheme induces a power imbalance between in-phase and quadrature components instead of their mutual correlation [98]. We compare this GS scheme with the well known WLT scheme referred as GS-WLT. We use the similar parameterization as adopted in [96] for our BER minimization problem and numerically solve the resultant non-convex optimization problem [2]. The comparison of GS schemes has been extended to hybrid shaping: where

[2]We omit the derivation and implementation details of GS-WLT and HS-WLT due to the limited space.



(a) Throughput versus EbNo

(b) Throughput versus HWD levels

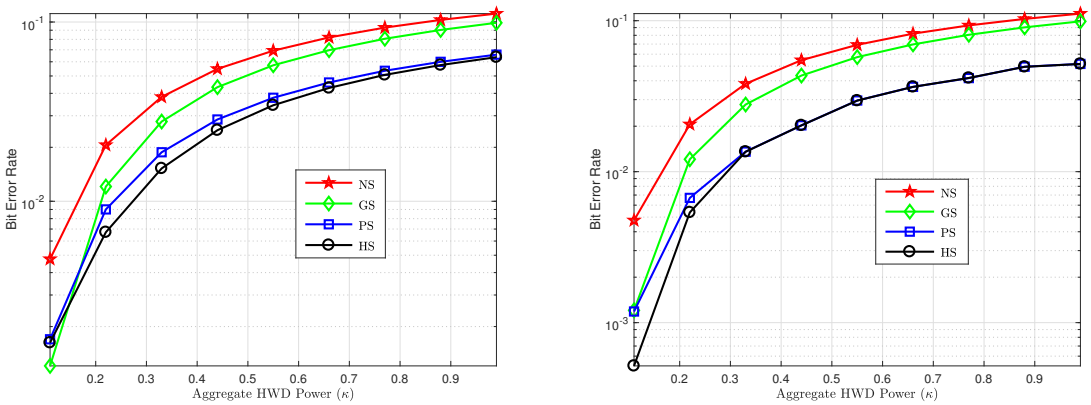
Figure 6.5: Throughput Performance in AWGN channel.

HS-AR and HS-WLT apply the proposed PS scheme to determine non-uniform probability distribution but respective GS techniques. Evidently, the candidate schemes perform equally good at low EbNo values but our proposed AR scheme outperforms WLT scheme in both GS and HS for relatively higher EbNo values.

Given the simulation settings as in Fig. 6.4a, we analyze system throughput (correctly received bits/symbol) for a range of EbNo values where the lower bound on system throughput can be obtained as

$$\mathcal{T}^{\mathrm{LB}}(\mathbf{p}) = \left[1 - \mathrm{P}_b^{\mathrm{UB}}(\mathbf{p})\right] \mathrm{H}(\mathbf{p}) \tag{6.38}$$

The throughput lower-bound (Th-LB) of all the transmission schemes can be calculated using their respective BER-UBs in (6.38). Fig. 6.5a validates the derived Th-LBs using MC simulations. It further depicts negligible throughput gain of GS over NS but noticeable throughput improvement using PS or HS. For instance, 1.5%, 6% and 7% percentage increase in throughput can be observed using GS, PS, and HS over NS at EbNo $= 5$ dB. At very low SNR, all the schemes depict unsatisfactory performance, as the required transmission rate can be higher than the maximum achievable rate which is related to the channel capacity. For moderate SNR, the throughput gain of the proposed schemes is quite sub-



(a) Using 16-QAM

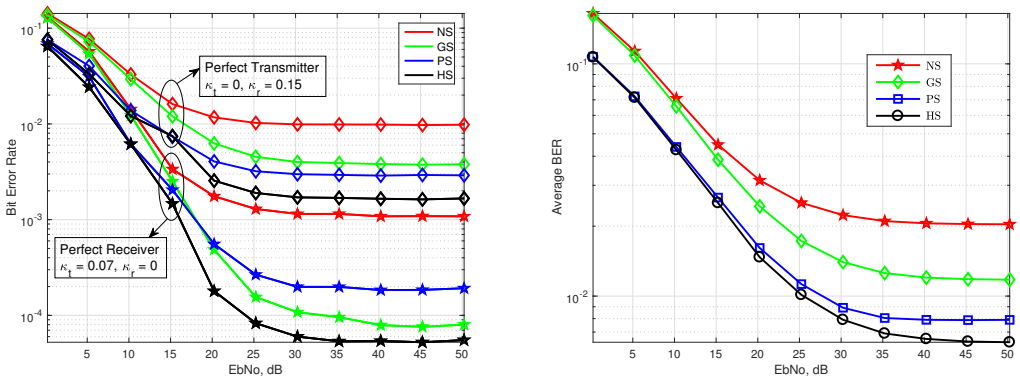
(b) Using 32-QAM

Figure 6.6: HWD mitigation for various AS in different channels.

stantial; nevertheless, it undergoes saturation when EbNo ≥ 20 dB. Interestingly, PS/HS saturates at 3 bits/symbol following rate fairness constraint with negligible BER whereas other schemes saturate below 3 bits/symbol depicting significant BER even though the entropy of 8-QAM with uniform distribution is $\log_2(8) = 3$.

Later, we analyze the behavior of various AS schemes with increasing distortion levels and their impact on the system throughput at EbNo $= 30$ dB. Fig. 6.5b compares the throughput performance of M_{u}-QAM NS and GS with $M_{\mathrm{nu1}} = 16$-QAM PS and HS as well as with $M_{\mathrm{nu2}} = 32$-QAM PS and HS. System throughput decreases almost linearly with increasing HWD for all forms of signaling but with different slopes. NS demonstrates the steepest slope with increasing HWD and all the other AS schemes render gradual slopes. Quantitative analysis shows the slopes of -0.55, -0.41, -0.28, and -0.24 using NS, GS, 16-QAM PS/HS, and 32-QAM PS/HS, respectively, with increasing HWD. Therefore, PS and HS present the most favorable results as compared to the GS. Their performance can be even improved by increasing the modulation order. Another important observation is the overlapping response of PS and HS especially for higher ordered QAM, which suffices PS and revokes the need of HS to perform even better.

A similar analysis is undertaken to study the impact of increasing HWD on the system



(a) BER for Perfect Receiver & Transmitter

(b) Average BER in Rayleigh fading.

Figure 6.7: BER Performance for a range of EbNo

BER performance in an AWGN channel. We assume 8-QAM for NS and GS whereas 16-QAM for PS and HS as depicted in Fig. 6.6a. Expectedly, the BER increases with increasing HWD levels and AS based systems achieve lower BER by efficiently mitigating the drastic HWD effects. Undoubtedly, the NS scheme suffers the most, but GS helps to decrease the BER to some extent. Further compensation can be achieved using the proposed PS and HS. Surprisingly, GS outperforms PS and HS at the lowest HWD values, e.g., $\kappa = 0.11$, in Fig. 6.6a but PS/HS maintain their superiority for $\kappa \geq 0.17$. Interestingly, PS/HS are still capable of outperforming GS even for the lowest HWD levels pertaining to their rate adaptation capability and added DoF using 32-QAM as highlighted in Fig. 6.6b. We can observe enhanced mitigation offered by the 32-QAM PS/HS as compared to the 16-QAM PS/HS due to the added DoF. Evidently, there is a trade-off between increased complexity and performance gain, which must be taken into account while choosing M_{nu} as per the system capability. For instance, we observe BER compensation of 66% and 77.5% using 32-QAM PS and HS, respectively, whereas BER compensation of 55% and 65% using 16-QAM PS and HS, respectively, at $\kappa = 0.22$ HWD level.

Another simulation example depicts the performance of the discussed AS schemes over a range of EbNo for two distinct scenarios of perfect receiver and perfect transmitter as presented in Fig. 6.7a. Perfect receiver system as the name specifies includes ideal zero-

distortion receiver but imperfect transmitter with $\kappa_t = 0.07$ whereas perfect transmitter system involves ideal zero-distortion transmitter but imperfect receiver with $\kappa_r = 0.15$. Note that the lower value of κ_t relative to κ_r is due to the fact that transmitters employ sensitive equipment to exhibit low distortions because the transmitter distortions are far more drastic than the receiver distortions. Interestingly, GS outperforms PS at EbNo $\geq$ 15 dB for the perfect receiver case as opposed to EbNo ≤ 15 dB where PS is still a better choice. HS outperforms both of them irrespective of the EbNo range classification. At such low HWD level, the BER percentage reduction of 81.82%, 90.91%, 94.55% is observed using PS, GS, and HS at 30 dB EbNo. Regarding the perfect transmitter scenario, GS and PS reverse the trend for higher EbNo level. Now the PS clearly outperforms GS for the entire range of EbNo and the HS marks its superiority over both of these schemes. At 0.15 HWD level, the EbNo gain of 8 dB, 12 dB, and 13 dB are estimated using GS, PS, and HS to attain the BER of 10^{-2}.

Finally, the average (ergodic) BER performance of the adopted system with $\kappa = 0.22$ HWD level is evaluated over a Rayleigh fading channel for a range of EbNo values as given in Fig. 6.7b. Evidently, the AS schemes preserve their BER trends and order. Clearly, average BER decreases with increasing EbNo and then undergoes saturation yielding an error floor. The derived BER bounds are also validated using MC simulations rendering a tighter bound for higher EbNo values. GS improves the average BER as compared to the NS scenario but PS and HS maintain their superior performance. Signaling schemes of GS, PS, and HS offer a percentage reduction of 54.55%, 63.64%, and 70.45%, respectively, in the average BER performance at 40 dB EbNo.

In a nutshell, we can conclude that the GS offers significant BER reduction at higher SNR values as opposed to the PS which offers universal gains. Moreover, the perks of HS are also prominent for higher SNR and higher M-ary modulation but depicts PS comparable performance at lower SNR values. Therefore, we recommend to employ HS given high SNR but resort to PS for lower SNR values to save additional computational expense.

Additionally, GS is a better choice for slightly distorted systems whereas PS/HS are the optimal choice for moderate to severely distorted systems. Furthermore, we can achieve improved performance by employing higher-order QAM constellations for PS/HS given adequate resources. On the other hand, the throughput gains are eminent at considerably lower SNR values and higher distortion values.

6.7 Conclusion

This work proposes probabilistic and hybrid shaping to realize asymmetric signaling in digital wireless communication systems suffering from improper HWD. Instinctively, all forms of asymmetric shaping are capable of decreasing the BER, and this performance gain improves with increasing SNR and/or increasing HWD levels with respect to NS. However, PS outperforms GS and performs equally well as HS. We can achieve more than 50% BER reduction with PS/HS over traditional GS. The perks of PS come at the cost of increased complexity in the design and decoding process. The HS scheme is capable of improving the system performance in terms of the BER as well as throughput. However, for less HWD levels and low EbNo, the benefits of HS over PS are limited while requiring additional complications in optimization, modulation, and detection procedures. Therefore, PS emerges as the best choice in the trade-off between enhanced performance and added complexity.

Chapter 7

Theoretical Analysis and Performance Limits

General analysis of a wireless communication system with improper transmission, interference, or noise should at least accommodate the complete SOS for valid theoretical analysis and performance limits. Improper Gaussian noise or interference are inevitable, however, the employment of proper or improper Gaussian signaling is questionable. Conventional information theoretic studies advocated the capacity achieving PGS in the point-to-point (P2P), BC and MAC [240]. However, recent studies have demonstrated some scenarios and conditions when PGS gains are inferior to that obtained from the IGS owing to the additional design freedom [137]. The performance gains are theoretically evaluated by metrics of interest like achievable rate, outage probability, power efficiency, and DoF. This section highlights the performance limits obtained by both conventional PGS and appealing IGS in various system settings. Followed by the design guidelines to fine tune the IGS transmission parameters in order to exploit the possible performance gains as detailed in Figure 7.1.

7.1 Achievable Rate

The performance of a wireless communication system can be theoretically quantified as achievable rate. A rate R is achievable if there exists a sequence of codes such that the maximal probability of error tends to 0 for sufficiently large block lengths. Recent contributions have highlighted a more general framework to quantify the achievable rate in order to accommodate the improper nature of the participating signals arising from

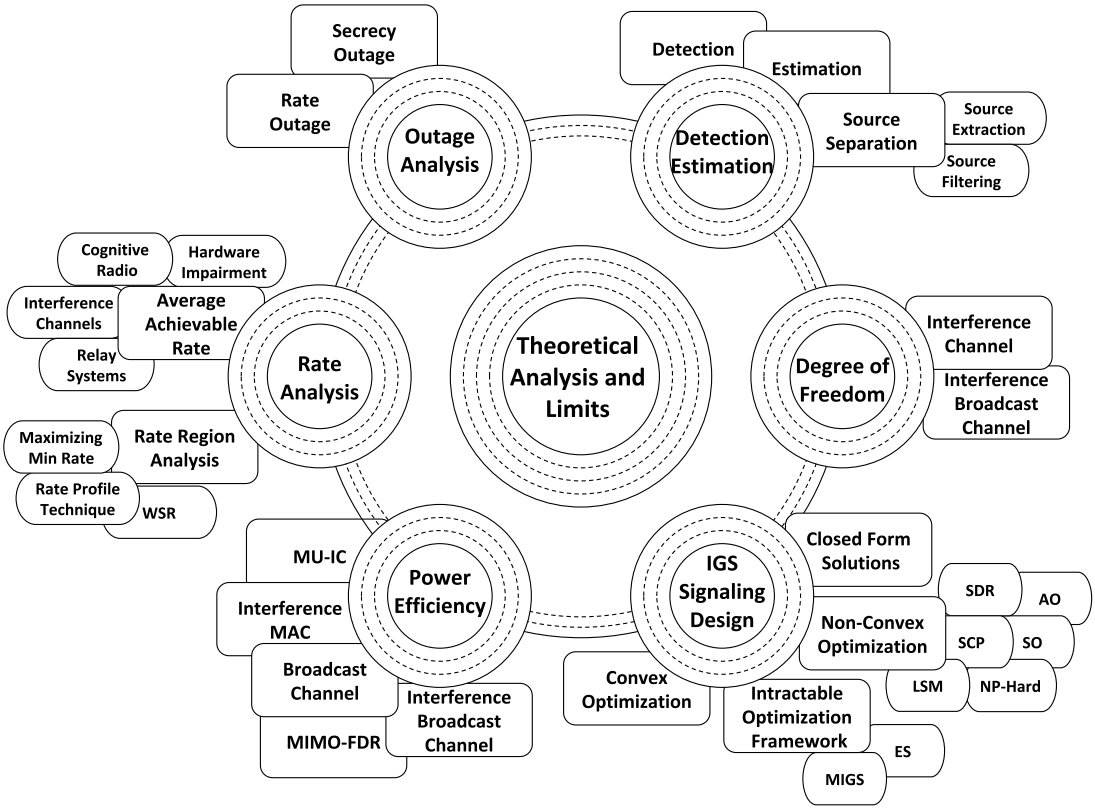


Figure 7.1: Overview and Configuration of Theoretical Analysis

IGS [20, 83, 159, 172, 176, 178, 188, 190, 192, 193, 293–295], WL precoding of PGS [77, 81, 149, 162, 163, 170, 174, 177, 184], or complex beamforming resulting in dependent real and imaginary parts [158, 296].

As an illustrative example, consider the most general case of K-user multiple-input multiple-output (MIMO)-interference channel (IC), where each user intends to communicate with its own RX but results in an interference to other $K-1$ RXs. Each user transmits general IGS $\mathbf{x}_k \in \mathbb{C}^M$ with the distribution $\mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{\mathbf{x}_k\mathbf{x}_k}, \tilde{\mathbf{R}}_{\mathbf{x}_k\mathbf{x}_k}\right)$ $\forall k = 1, \ldots, K$ from M transmitting antennas. Thus, the accumulated received signal vector at user k with N receive antennas is

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{x}_k + \sum_{j \neq k, j=1}^{K} \mathbf{H}_{kj}\mathbf{x}_j + \mathbf{n}_k, \tag{7.1}$$

where, $\mathbf{n}_k \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}_N, \mathbf{0}\right)$ is the CSC Gaussian thermal noise and $\mathbf{H}_{kj} \in \mathbb{C}^{N\text{x}M}$ is the quasi-static fading channel from TX-j to RX-k. Based on the augmented representations

and Schur complement, the differential entropy of $\mathbf{y}_k \in \mathbb{C}^N$ and interference plus noise term $\mathbf{w}_k = \sum \mathbf{H}_{kj}\mathbf{x}_j + \mathbf{n}_k$ are given as $H(\mathbf{z})$ for $\mathbf{z} = \mathbf{y}_k$ and $\mathbf{z} = \mathbf{w}_k$, respectively [31].

$$H(\mathbf{z}) = \log\left((\pi e)^N \left|\mathbf{R}_{\mathbf{zz}}\right|\right) + \frac{1}{2}\log\left(\left|\mathbf{I} - \mathbf{R}_{\mathbf{zz}}^{-1}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-\mathrm{T}}\tilde{\mathbf{R}}_{\mathbf{zz}}^{\mathrm{H}}\right|\right).$$

Shannons capacity formula assumes additive Gaussian noise and coded transmission with codewords drawn from a Gaussian codebook. Therefore, the instantaneous achievable rate per unit bandwidth (spectral efficiency) at RX-k with IGS can be obtained as [81]

$$\begin{aligned} R_k =& \mathcal{I}\left(\mathbf{y}_k; \mathbf{x}_k\right) = H(\mathbf{y}_k) - H(\mathbf{y}_k|\mathbf{x}_k) = H(\mathbf{y}_k) - H(\mathbf{w}_k) \\ =& \frac{1}{2}\log\frac{\mathbf{R}_{\underline{\mathbf{y}}_k\underline{\mathbf{y}}_k}}{\mathbf{R}_{\underline{\mathbf{w}}_k\underline{\mathbf{w}}_k}} = \underbrace{\log\frac{\left|\sigma^2\mathbf{I} + \sum\limits_{j=1}^{K}\mathbf{H}_{kj}\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}\mathbf{H}_{kj}^{\mathrm{H}}\right|}{\left|\sigma^2\mathbf{I} + \sum\limits_{j\neq k, j=1}^{K}\mathbf{H}_{kj}\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}\mathbf{H}_{kj}^{\mathrm{H}}\right|}}_{\triangleq R_{k,\text{proper}}\left(\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}\right)} \\ &+ \frac{1}{2}\log\frac{\left|\mathbf{I} - \mathbf{R}_{\mathbf{y}_k\mathbf{y}_k}^{-1}\tilde{\mathbf{R}}_{\mathbf{y}_k\mathbf{y}_k}\mathbf{R}_{\mathbf{y}_k\mathbf{y}_k}^{-\mathrm{T}}\tilde{\mathbf{R}}_{\mathbf{y}_k\mathbf{y}_k}^{\mathrm{H}}\right|}{\left|\mathbf{I} - \mathbf{R}_{\mathbf{w}_k\mathbf{w}_k}^{-1}\tilde{\mathbf{R}}_{\mathbf{w}_k\mathbf{w}_k}\mathbf{R}_{\mathbf{w}_k\mathbf{w}_k}^{-\mathrm{T}}\tilde{\mathbf{R}}_{\mathbf{w}_k\mathbf{w}_k}^{\mathrm{H}}\right|}. \end{aligned} \tag{7.2}$$

The achievable rate R_k at RX-k has two components, the first term $R_{k,\text{proper}}\left(\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}\right)$ is the conventional achievable rate attainable by traditional PGS scheme as it only depends on covariance $\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}$. Whereas, the additional term is the consequence of employing IGS as it is a function of both $\mathbf{R}_{\mathbf{x}_j\mathbf{x}_j}$ and $\tilde{\mathbf{R}}_{\mathbf{x}_j\mathbf{x}_j}$. Therefore, IGS yields another privilege to appropriately choose/optimize $\tilde{\mathbf{R}}_{\mathbf{x}_j\mathbf{x}_j}$ under the power constraints to yield the second term in (7.2) strictly positive. Interestingly, this property cannot be exploited in the equivalent real composite domain analysis [81].

The rate expression obtained for multiuser (MU) MIMO-IC can be easily reduced to MU single-input single-output (SISO)-IC [293], SISO one-sided interference channel (Z-IC) [178],single-input multiple-output (SIMO) systems with hardware distortions (HWDs) [108], MU multiple-input single-output (MISO)-IC [295], MU MIMO-IC [174], MU MIMO-

broadcast channel (BC) [163], MIMO-interference broadcast channel (IBC) [162], cognitive radio settings [188, 190, 192, 294], point-to-point (P2P) systems with asymmetric HWIs [20], relaying systems [159], full-duplex systems [83] and other similar interference-limited scenarios where interference can be treated as Gaussian noise. Various studies argue that IGS can offer significant achievable rate improvement relative to the PGS. Optimizing the covariance and pseudo-covariance matrices allow us to optimize the system performance to maximize the average achievable rate or achievable sum rate, minimize achievable rate, or to decide achievable Pareto rate region boundaries [81].

7.1.1 Average Achievable Rate Limits

Average achievable rate or ergodic rate is an insightful performance metric to quantify the capability of a network to transmit the number of bits per second per Hertz. This section summarizes the average achievable rate performance gains of IGS over PGS in various interference limited scenarios arising in cognitive radio schemes [190, 192, 294], multi-antenna or MU systems [20, 149, 177, 184], and relay systems [83, 158, 159] etc.

In a cognitive radio setting for dynamic spectrum access, the unlicensed secondary user (SU) coexists with a licensed primary user (PU) and opportunistically utilize its spectrum resources to improve the overall spectral efficiency [61, 297]. However, this approach renders undesirable interference on the legit primary network. It is noteworthy that the primary TX sticks to PGS as it is using its own spectrum being unaware of the secondary IC [190, 191]. Interestingly, the least entropy loss due to IGS makes it a suitable transmission scheme for SU as the improper interference on PU is far less detrimental compared to proper interference. Thus, IGS is the preferred choice for SU to maximize its achievable rate while maintaining PU rate or quality-of-service (QoS) constraints. The average achievable rate performance gain of SU with IGS over PGS are analyzed in three different cognitive setups i.e., underlay, overlay and interweave.

- In an underlay cognitive system, IGS is only beneficial when the ratio of the squared

modulus between the SU-PU interference link and the SU direct link exceeds a given threshold. Upon meeting this criteria, the SU adopts IGS as optimal signaling and demonstrates magnified relative performance gains up to 256% over PGS especially when PU is not heavily loaded [152, 190, 298].

- In an overlay cognitive system, where SU broadcasts a mixture of PGS and IGS to aid the primary message transmission and minimize the interference effect of SU on PU respectively. The optimized IGS offers 33.33% performance gain for 30dB SNR with partial CSI, while meeting PU quality of service (QoS). This gain is conditional and improves with decreasing primary network direct-link gain [192, 299].

- In an interweave cognitive system, the achievable rates of both the PU and the SU depend on the activity of the PU and the detection ability of the SU. Employed IGS yield percentage increase up to 8.26% over PGS with 50% probability of detection (PoD) and 10% PFA. The gain is significant especially at low sensing and detection capabilities of the SU, lower PU direct link and higher SU interference on the PU side [294].

Inspired by the perks of IGS in interference limited environments, various contributions have reaped IGS benefits in full duplex (FD)/half duplex (HD) multi-hop decode-and-forward (DF) relay systems by effectively compensating residual self-interference (RSI), inter-relay interference (IRI), and/or HWDs [83, 158, 159]. For instance, IGS is proposed in single antenna DF-FD relay channels to eliminate the self-interference (SI) and increase the throughput. This scheme with SI-elimination provides significant improvement of 107.14% over the conventional symmetric signaling without SI-elimination [158]. IGS is further employed to compensate not only RSI but also improper HWDs in multi-hop DF-FDR system. IGS parameters can be centrally optimized at one node or distributively optimized at multiple nodes to maximize the end-to-end (E2E) achievable rate. Centralized and distributed optimization offer relative performance gains of 166.67% and 80%, respectively. Gain can

rise up to 355.56% with centralized approach for higher interference-to-noise ratio with increased communication overhead and system complexity [83]. Moreover, the potential benefits of IGS are also exploited in two-hop alternate relaying system to relieve IRI and maximize total achievable rate in the absence of CSI at source. Two detection schemes are employed based on the IRI level relative to the desired signal level i.e., 1) for low IRI: single-user decoding (SUD) treating IRI as noise and 2) for strong IRI: successive decoding (SD) that first decodes IRI and subtracts it from received signal before detection. Average rate percentage improvement up to 161.54% and 66.67% is achieved with SUD and adaptive scheme that switches between SUD and SD, respectively. This gain is especially significant when the source-relay channel is a bottleneck [159].

IGS superiority is also proven in multi-antenna setup to mitigate various HWIs [20, 108]. Interestingly, aggregate impairments specially IQI transform symmetric noise and transmitted signal to asymmetric noise and received signal, respectively, motivating the employment of IGS to combat them. Javed *et al.* proposed optimal IGS and maximal IGS schemes for tractable and intractable optimization of SIMO and MIMO systems, respectively. Percentage average achievable rate improvement of 7.76% and 3% is depicted in highly impaired SIMO systems with optimal IGS and adaptive maximal IGS-PGS, respectively. Similarly, the employment of maximal IGS yields up to 10% increase in average achievable rate of MIMO relative to PGS even in low SNR regime [20].

Other contributions highlight the perks of IGS in single user [177] and MU MIMO-ICs [149, 184]. Firstly, IGS transmission with uniform- (UPA) or optimal- (OPA) power allocation strategies is exploited in MIMO P2P channel with interference (P2P-I) to improve the achievable rate. Relative rate gains of OPA-UPA of IGS over PGS are 102-115% for 1x1, 46-37% for 2 x2, 24-20% for 4x4, and 17-13% for 8x8 in MIMO P2P-I [177]. Secondly, for 2-user (2U) MIMO-IC, transmit covariance matrix is designed based on two maximal IGS schemes i.e., Improper-LB that provides the SU with the minimal rate, and Improper-UB that provides the SU with the maximum rate [184]. Percentage relative gain

of user-1 and user-2 achievable rate is 195.45% increase and 8.7% decrease with Improper-LB relative to PGS. However, the percentage relative gain of user-1 and user-2 achievable rate is 10.53% decrease and 209.52% increase with Improper-UB relative to PGS. Interestingly, in an interference limited scenario, if the IGS scheme is beneficial for interferer, it will be detrimental for the sufferer and vice versa. Lastly, for MU MIMO-IC transmission rates increase with IGS and transmit coordination and IGS instead of conventional interference management (IM) with PGS. Lagen *et al.* claims IGS performance gains over PGS in terms of mean user throughput are 10.64%, 13.95% and 22.92% with centralized-IM, decentralized-IM, and no-IM, respectively [149].

7.1.2 Rate Region Analysis

Pareto boundary of the achievable rate region comprises of all the Pareto optimal points, which are defined as:

Definition 17 (Pareto-Optimal)**.** *The rate pair (R_1, R_2) is called Pareto-optimal if ($\bar{R}_1, R_2$) and ($R_1, \bar{R}_2$), with $\bar{R}_1 > R_1$ and $\bar{R}_2 > R_2$, are not achievable [178].*

Interestingly, IGS and/or WL processing significantly improves the achievable rate region in MU interference setup. Various studies evaluated such improved rate regions whereas many others focused on the boundary of an achievable rate region, called the Pareto boundary, based on sum rate analysis, rate profile technique (RPT), or minimum weighted rate maximization [300].

Sum Rate Analysis

Pareto boundary of the rate region can be acquired by maximizing the sum rate when dealing with multiple nodes. Impropriety characterization renders substantial increase in the pareto boundaries as observed in interference MAC [166,167], Z-IC [176], and X-IC [172]. Apart from the complete boundary characterization, some studies focused on the weighted

Table 7.1: Achievable Rate Improvement by IGS over PGS in Different Settings

System	Type	IGS-TX	Metric	Procedure	Improvement	Ref
Cognitive Radio	Underlay P2P	SU	Max SU-Rate with PU-Rate Constraint	Maximal IGS	up to 256%	[190]
	Underlay MAC	S-MAC Users	Rate Region Boundary Maximize Sum Rate	Algorithm: Closed-Form Expressions	R1:133%, R2:117% up to 200%	[188]
	Overlay	SU	Max SU-Rate with PU-QoS Constraint	Piecewise Closed-Form Solution	up to 191.7%	[192]
	Interweave	SU			up to 8.3%	[294]
MIMO	Massive MU DL	BS	Maximize Sum Rate	Numerical Evaluation	ZF/MMSE:75% BD:97%	[77]
	2U X-IC	SU	Maximize SU-Rate	Maximal IGS	R1:195.5% increase R2:8.7% decrease	[184]
	MU X-IC	All	Minimize MSE	AO	cent-IM:10.6% decent-IM:13.9% no-IM:22.9%	[149]
	LP/WLP MU X-IC	WL Users	Maximize WSR ($\equiv$ Minimize WMSE)	BCD with AO	WLP:133%, HetTX(1):84.2% HetTX(K/2):28.9%	[174]
	Z-IC	Interferer	Average Sum Rate Minimum Rate	Maximal IGS	Sum rate:55.2% Min rate:640%	[177]
	Multicell BC N/w	BS	Maximize WSR	BCD with AO	up to 11.8%	[163]
	HCRAN IBC	Femto BSs		WMMSE and ADMM for AO	WMMSE-IGS:12.5% WMMSE-PGS:21.7%	[162]
	MU-IBC	BS	Maximize min rate	Path following Algo	30dBm: 37.93%	[301]
	P2P-I	MIMO TX	Maximize Rate Minimize MSE	Majorization Theory Tools	Rate Gap:1.55 MSE GAP:0.2 Rate Gain:115%	[177]
MISO	MU P2P-IC	Users	Pareto Rate Region Max-Min Rate	SDR with SOCP and GR	R1:70%, R2:25% up to 42.8%	[295]
	2U-BC	BS	Pareto Rate Region	SO with Bisection Search	R1:150%, R2:100%	[295]
	3U-BC	BS	Pareto Rate Region	Gradient based Rate Balancing	R1:100%, R2/R3:42.86%	[155]
	Multicell NOMA	BS	Max-min Fairness	LMI based Path following Algorithm	NOMA: 87.50% OMA:121.05%	[302]

System	Type	IGS-TX	Metric	Procedure	Improvement	Ref
SIMO	P2P HWI	TX	Maximize Achievable Rate	Convex QCQP using IPM	up to 10%	[20]
SISO	2U X-IC	1U	Rate Region	Closed-form	R1:133.7%,R2:86.6%	[303]
		All users	Max-Min Fairness Proportional Fairness	Closed-Form Solution	up to 83% up to 633%	[172]
			Max-Min Rate	SDR	up to 228.57%	[295]
			Pareto Rate Region Max-Min Rate Sum Rate	SDR for QCQP	R1:357%, R2:389% JO:58.6%, SO:54.8% JO:21.5%, SO:18.8%	[81]
	4U X-IC		Maximize Sum Rate	Alternating Minimization	up to 35.2%	[170]
	MU X-IC		Max-Min Weighted Rate	SDR and GR	2U:221.4%, 3U:304.3%	[293]
	P2P Z-IC		Maximize Sum Rate	Closed-Form Solution	up to 30.8%	[176]
			Pareto Rate Region		R1:83%, R2:150%	[178]
					R1:30.9%,R2:36.2%	[303]
	2U BC NOMA	BS	Max Sum Rate	KKT conditions	$\Delta P = 0.4\mathrm{P_t}$: 23.60%	[150]
	MU IBC 3-Cells	Each BS	Maximize Sum Rate	Exhaustive Search	K=5: 13.7% K=10:17.6% K=20:20.6%	[296]
DF Relays	Multihop FD	Source and Relays	Maximize E2E Rate	SCP for QCLP	Centralized: up to 355.6% Distributed: up to 80%	[83]
	Dualhop FD		Throughput	Line Search Method	up to 107%	[158]
	Alternate HD Relay	Relays only	Maximize Total Achievable Rate	Piecewise Closed-Form Solution	SUD:161.5% SUD-SD:66.7%	[159]

sum-rate (WSR) maximization achieved by IGS in various system settings like cognitive radio [188], massive MIMO with IQI [77], X-IC [81, 170, 174, 303], Z-IC [176, 177], BC [150, 163], and IBC [162, 296].

In underlay cognitive radio scheme, IGS has emerged as a promising candidate to improve the sum rates of the MAC users. Firstly, for primary-MAC (P-MAC) interfered by a P2P channel, IGS and symbol extensions can improve the achievable rates up to three times based on the interferer strength [166, 167]. Secondly, underlay 2U secondary-MAC (S-MAC) exploited IGS to improve average sum rates up to 200% [188]. Moving on to the large-scale MIMO systems with TX-IQI, Zhang *et al.* analyzed WLP algorithms based on ZF (WL-ZF), matched filter (WL-MF), WL-MMSE, and block-diagonalization (WL-BD) for the downlink (DL) scenario [77]. They argued the achievability of same multiplexing gains and WSR with WL-ZF and WL-BD in the presence of IQI as their counterparts ZF and BD in the absence of IQI. However, this performance is attained at the expense of minor power loss owing to the increased system scale. Interestingly, the WSR analysis of IQI system with WL-ZF/WL-MMSE and WL-BD depicted percentage increase up to 97.67% and 75% over ZF/MMSE and BD for single- and multi-antenna users, respectively [77].

Rate region improvement of Z-IC has also been studied extensively [176]. This contribution employs real-composite representation for easy optimization and characterizes only one point of rate region [176]. Improved sum rates offered by IGS with effective IM in Z-IC are also widely investigated in 2U SISO Z-IC. PGS is preferred for weak interference whereas optimal IGS can provide WSR improvements up to 30.8% in strong interference regime [176].

Sum rate analysis is also extended to X-IC in various multi-user and multi-antenna setups. For instance, Ho *et al.* explore the Pareto region for the 2U SISO-IC with cooperative (IGS) as well as non-cooperative (PGS) transmission strategy. They prioritize improper rank one signals over full-rank signals because of their simplicity, easy implementation

and close to optimal sum rate [172]. The study focuses on improving the system efficiency in terms of max-min fairness and proportional fairness while carrying out the rate region analysis. Evidently, at the max-min fairness, both users share the same maximum possible rate in Pareto region however at proportional fairness, the aim is to maximize the product of improvement over the Nash equilibrium. IGS provides remarkable percentage max-min fairness and proportional fairness improvements up to 83.33% and 633.33%, respectively, in the medium SNR regime. The rate region improvement is more substantial for decreasing SNR and asymmetric channel i.e., one IC is stronger than the other [172]. On the other hand, the Pareto region attained with minimum mean square error (MMSE) scheme is larger than and contains the corresponding region with ZF scheme. On the other hand, for 2U SISO-IC, Zeng *et al.* propose a joint (JO) and separate (SO) IGS optimization framework which achieves 21.52% and 18.85% WSR improvement, respectively, relative to PGS scheme [81]. Similarly, Soleymani *et al.* report average sum rate increase up to 150% in 2U SISO-IC even with imperfect CSI [303]. Moreover, Lameiro *et al.* illustrate WSR improvement up to 35.14% at 60dB SNR with IGS and linear interference alignment (IA) for 4U SISO-IC [170]. The WSR analysis is also extended from MU SISO-IC to MU MIMO-IC in a transitional heterogeneous (HetTX) setting where some legacy linear transceivers i.e., linear precoding and linear estimation (LP-LE) coexist with other WL transceivers i.e., WL precoding and WL estimation (WLP-WLE) [174]. This work addresses WLT filter design to maximize WSR and presents iterative procedure to solve equivalent minimum weighted-mean square error (W-MSE) problem. Transition from LP-LE to WLP-WLE, HetTX(1)-WLE (1 LP,$K-1$ WLP), and HetTX($K/2$)-WLE ($K/2$ LP,$K/2$ WLP) achieve the percentage improvements up to 133.33%, 84.21%, and 28.95%, respectively. Interestingly, this performance gain increases with increasing number of users, increasing aggregate interference levels or decreasing number of antennas. Surprisingly, the WL transceivers with no interference coordination performed worse than the linear transceivers with full coordination among users [174].

Apart from IC, WSR maximization problem is also extensively studied for BC and IBC [150, 162, 163, 296]. IGS offers significant sum-rate maximization for the downlink non-orthogonal multiple access i.e., non-OMA (NOMA) with imperfect successive interference cancellation (SIC). In a 2U SISO-BC NOMA setup, IGS optimization based on Karush-Kuhn-Tucker (KKT) conditions renders 23.60% and 18.71% sum-rate improvement when the power difference between two users is 40% and 20% of the total power ($\mathrm{P_t}$), respectively [150]. Similarly, increasing number of users, receiver antennas and transmission power of the base-station (BS) offer considerable WSR gain of 11.84% with WL design in MU MIMO-BC [163]. In contrast to the IC and BC, few studies focused on analyzing the potential benefits of IGS in the combined IBC [162, 296]. For instance, Shin *et al.* propose a new IA strategy based on IGS and MU diversity (MUD) for 3-cell SISO-IBC where each BS covers K users per cell. The percentage WSR improvement of 13.68%, 17.65%, and 20.56% are achieved with the proposed strategy relative to conventional IA strategies for $K = 5$, $K = 10$ and $K = 20$ users, respectively, at 20dB SNR [296]. WSR maximization problem in MIMO-IBC (e.g., heterogeneous cloud radio access network (H-CRAN)) is a non-trivial extension of MIMO-IC [174] and MIMO-BC [165]. Thus, Lin *et al.* propose a distributed beamforming algorithm for separate optimization. This algorithm outperforms existing WMMSE with IGS and PGS in terms of WSR by 12.5% and 21.74%, respectively [162].

Rate Profile Technique

Rate region boundary can also be established using RPT instead of maximizing the sum rate. This technique also advocates impropriety incorporation to render improved pareto boundaries in numerous interference-limited setups like underlay MAC [188], Z-IC [178, 303], X-IC [81, 295, 303, 304], and BC [155, 295].

Lameiro *et al.* extended their work in underlay P2P cognitive system [190] to underlay MAC setup in order to study the improved rate region by IGS. The IGS transmitting

unlicensed S-MAC coexists with PGS transmitting licensed primary link. The numerical results for 2U S-MAC with zero-forcing (ZF) decoding present the rate improvements up to 117% and 133% for SU-2 and SU-1, respectively. IGS guarantees rate improvement if the sum of IC gains is above a certain threshold and surprisingly, the relative gain increases with increasing number of users [188].

Rate region improvement of Z-IC has also been studied extensively [178]. The contribution [178] employs augmented complex representation for more insightful analysis and characterizes entire rate region boundary. Lameiro *et al.* emphasize the conditional optimality of IGS in SISO Z-IC that attains 83.33% and 150% percentage increase in the rates of user-1 (R1) and user-2 (R2), respectively, at the maximum sum rate point on Pareto boundary [178]. Likewise, Soleymani *et al.* claim 30.95% and 36.25% increase in R1 and R2, respectively, in a 2U SISO Z-IC with IGS transmission under imperfect CSI [303].

The perks of IGS are not only limited to Z-IC but also extend to X-IC. By far, 2U SISO-IC is the mostly studied X-IC for the employment of IGS with some substantial results [81, 295, 303, 304]. For instance, Soleymani *et al.* claim 86.67% and 133.77% increase in R2 and R1, respectively, with IGS in a 2U SISO X-IC under imperfect CSI of the interfering links [303]. Moreover, they also propose a practical IGS scheme i.e., maximal IGS (which does not require any optimization) for 2U SISO-IC with Rayleigh fading. They again present substantial increase in Ergodic rate region as well as sum-rate with IGS under strong interference [304]. Similarly, Zeng *et al.* present significant increase in the Pareto rate region of 2U SISO-IC as IGS improves up to 388.89% and 357.14% rates for user-2 and user-1, respectively [81]. Additionally, for a given the optimal rank-1 transmit covariance matrices, rank-1 pseudo-covariance matrices are proven optimal for achievable rate region in MU MISO-IC. Percentage improvements of R1:25% and R2:70% are attained for 2U MISO-IC [295].

IGS can render effective interference suppression in BCs when interference is treated as noise. For instance, 2U MISO-BC can achieve up to R1:100% and R2:150% rate im-

provement [295]. Similarly, for a 3U MISO-BC, IGS offers R2/R3:42.86% and R1:100% increase relative to PGS [155].

Minimum Achievable Rate

IGS transmission can also be optimized in order to maximize the minimum achievable rate of the system. Such contributions analyzed MU SISO-IC [81, 293], MU MISO-IBC [302], MU MISO-IC [295], MU MIMO Z-IC [177], MU MIMO-IC [300], and MU MIMO-IBC [301].

Firstly, average max-min rate improvement of 228.57% and 42.86% is reported in 2U SISO-IC with joint optimal and 3U MISO-IC with suboptimal optimization, respectively [295]. Similarly, percentage improvements up to 58.62% and 54.83% in average max-min rate based on the JO and SO, respectively, are observed in 2U SISO-IC [81]. Furthermore, numerical results revealed up to 221.43% and 304.26% improvement in 2U and 3U SISO-IC, respectively. The performance gains further improved with increasing SNR and number of users [293].

Secondly, IGS under OMA or NOMA is reported to exhibit almost two-fold gain in users minimum throughput, combating both intra-cell and inter-cell interference, in a MU multi-cell broadcast network. The proposed IGS design algorithms based on linear matrix inequality (LMI) optimization result upto 87.50% and 121.05% improvement in worst user rate relative to PGS in 4U, 3-cell MISO-IBC under NOMA and OMA schemes, respectively [302].

Lastly, the IGS benefits to maximize the minimum user rate are also reaped in MU MIMO ICs. For instance, IGS offers up to 640% improvement in min-rate performance of 2U MIMO Z-IC [177]. Moreover, a hardware impaired MU MIMO-IC depicts more than 80% improvement in fairness rate with IGS over PGS in a 10U setup [300]. Similarly, a path following IGS optimization algorithm yields up to 37.93% max-min rate improvement at 30dBm in a MIMO-IBC with 3 cells and 6U/cell [301].

In a nutshell, various contributions presented rigorous analysis and substantial results to demonstrate significant achievable rate improvements for different interference limited scenarios as summarized in Table 7.1.

7.2 Outage Probability

Several researchers resort to more conceptual analysis such as outage probability, i.e., the probability of the event when the system performance falls below a pre-defined threshold. Such abstracted analysis sacrifices the model depth for simplicity leading to simple expressions that characterize high-level network behavior, highlight general trade-offs, and facilitate network design. Outage probability can be quantified in terms of rate outage, SNR outage, secrecy outage and error outage etc. Of all these, rate and secrecy outage are focused as they suitably reflect the IGS system operation quality.

7.2.1 Rate Outage Probability

ROP is generally defined as the probability of the event $\Pr\{\text{Rate} < \text{Threshold}\}$. Various contributions have demonstrated substantial decrease in the Rate Outage Probability (ROP) when IGS transmission is adopted as opposed to PGS. These studies analyzed various systems including but not limited to cognitive radio [187, 191], multi-antenna [108] and relay systems [179] suffering from external interference or internal HWIs.

The utility of IGS at SU in underlay cognitive radio setting is supported by the claim that it decreases SU ROP by 77.5% while meeting PU QoS and SU power constraints. The performance gain of IGS increases as the interference-to-noise ratio (INR) of the SU to the PU increase for a certain SU target rate [191]. In another spectrum sharing scenario with coexisting FD PU and HD SU, IGS offers up to 91.43% reduction in the ROP when the allowable INR at the PUs exceed a certain threshold and the SU follows a maximum allowable target rate [187]. Gaafar *et al.* extended their work to mitigate RSI in FDR using IGS under Nakagami-m fading. The observed percentage decrease in ROP of IGS relative

Table 7.2: Outage Probability Reduction with IGS Relative to PGS

System	Tech	Desc	IGS TX	Metric	Procedure	Improvement	Ref
Multiple Antennas	SIMO	HWDs	Single TX	Minimize Rate Outage	Closed-Form Solution	SISO:85% SIMO:100%	[108]
DF Relays	FDR	Dual-hop	Relays	Minimize E2E Outage	Coordinate Descent	MPGS: 90% OPGS: 68.18%	[179]
Cognitive Radio Setting	Underlay	HD-PU	SU only	Minimize Rate Outage	Closed-Form Solution	up to 77.5%	[191]
		FD-PU			Algorithm Design	up to 91.43%	[187]
		Eaves-dropper		Minimize Secrecy Outage	Numerically	up to 95%	[189]

to PGS with maximal power allocation (MPGS) and PGS with optimal power allocation (OPGS) is 90% and 68.18%, respectively, with 5W power budget at relay. Unlike PGS, IGS maintained a fixed performance even with increasing RSI [179].

Furthermore, Javed *et al.* demonstrated the effectiveness of IGS transmission to efficiently combat the drastic effects of asymmetric HWDs on SIMO systems. IGS parameters were optimized to maximize instantaneous achievable rate and consequently reduce ROP up to 100% and 85% for the adopted SIMO [108] and SISO [79] system, respectively. The ROP gain in imperfect hardware system is especially significant with increased distortion levels and more receiver streams N_{R} for any given threshold rate.

7.2.2 Secrecy Outage Probability

A secrecy outage is the probability of event when the mutual information of the desired link (A→B) is lower/equal to that of the undesired link (A→E) i.e., $\Pr\{\mathcal{I}_{\mathrm{AB}} - \mathcal{I}_{\mathrm{AE}} < \text{Threshold}\}$. Consider the underlay cognitive setup with primary network link source-destination (S→D) and a secondary link Alice-Bob (A→B) in the presence of an eavesdropper Eve (E). Alice adopts IGS scheme in order to reduce the Secrecy Outage Probability (SOP). It is evident from the fact that IGS demonstrate lower differential entropy

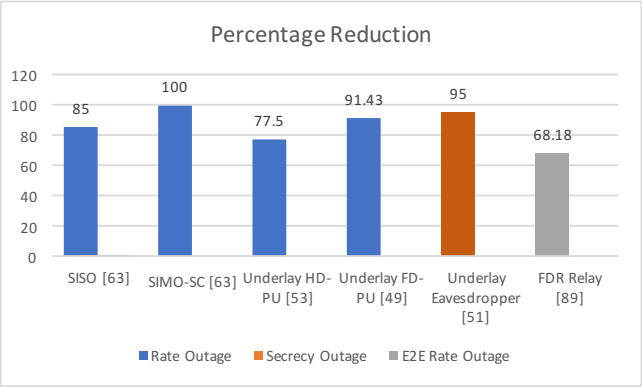


Figure 7.2: Percentage Reduction in Outage Probability with IGS

than PGS. Thus, an improper interference will be less harmful on the achievable rates relative to a regular proper interference when interference is treated as noise. Therefore, this allows Alice to transmit with higher power and reap higher achievable rates without violating S→D licensed link. Nonetheless, SOP reduction is a trade-off between impropriety and power of the transmitted signals. Oliveira *et al.* claim to be the first to establish IGS superiority in physical layer security of cognitive setup. They present closed formulation for the SOP and revealed up to 95% lower SU-SOP with IGS [189].

All of these contribution support the advantageous trend of IGS in reducing outage probability of the adopted systems suffering from interference or improper noise as highlighted in Table 7.2 and Figure 7.2.

7.3 Power Efficiency

Interestingly, IGS allows achieving the desired QoS while spending less power at the TXs owing to the additional tunable parameter. This leads to power efficient solutions for various MU, multi-antenna, and relay setups. IGS transmission with and without symbol extensions for MAC interfered by a P2P channel renders up to 32% and 40% power saving ratios, respectively [166]. Likewise, IGS optimization offers up to 19.38%, 8.47% and 3.5% energy efficiency with 4, 5 and 6 transmitting BS antennas in a MU MIMO-IBC with 3 cells and 6Us per cell [301].

IGS is further incorporated for joint rate-energy optimization in multi-antenna heterogeneous two-tier networks where users are subject to TX noise, MU interference and RSI. Although these impairments are detrimental for achievable rate, they are beneficial for energy harvesting because they carry RF energy [186]. Energy harvesting of 4μwatts/channel-use is achieved with 0.5bits/channel-use achievable rate in FD D2D 2U pairs coexisting with HD MU-BC in a cellular network. Similarly, for a single-carrier MISO BC the decrease in transmit power requirements by IGS over optimal PGS is 1dB, 3dB and 12dB for 2,4 and 6 users, respectively [155].

IGS is beneficial to improve the power efficiency of one-way DF-FD MIMO relay [182] and two-way AF-FD MIMO relay [181] for MU interference networks. MU network with MIMO-FDR is not only prone to inter-user interference and transceiver noise but also self-interference due to MIMO-FDR relay operation. Optimal IGS at the sources and relay requires 0.95dB minimum sum-power requirement relative to 1.2dB, 1.6dB and 1.75dB for OPGS, PGS-ZF and PGS-MRC, respectively, under QoS demand of 1 bit/channel-use by two-pair MIMO relay network. Power efficiency attained by IGS scheme improves with the increasing interference levels and TX noise for the entire range of rate demands. Interestingly, only IGS scheme is capable of meeting high rate demands by single real streaming whereas PGS fails to do so even with infinite transmission powers [182]. Additionally, AF FD MIMO relay system enhances power efficiency through SNR balancing or transmit power minimization using WL transmit strategies. By fine-tuning non-circularity of WL complex transmitted signals significant performance gains up to 150% percentage increase in average minimum SINR and 5dBW less average minimum relay transmit power relative to linear precoding schemes [181]. The WL gain increases with increasing number of pairs but decreases with increasing number of relay antennas. Although IGS offers significant power/energy efficiency gains over PGS, the relative energy efficiency benefits are less than rate benefits [300].

7.4 Degrees of Freedom

Apart from the discussed metrics, IGS is also beneficial to improve the achievable DoF in a given interference scenario including but not limited to X-IC [169, 170], BC [164]and IBC [296]. For instance, Lameiro *et al.* illustrated improved DoF $\approx 4/3$ (requiring minimal symbol extensions) with IGS and linear IA for 4U SISO-IC [170]. Similarly, Yang *et al.* demonstrated achievable DoF of 0.5 more than the outer bound of DoF of MIMO X-IC when spatial IA and ZF framework are employed along with the IGS [169]. Moreover, in MIMO-BC system with WL transceivers IGS per-user transmit signals outperform PGS counterpart under QoS constraints without time-sharing [164]. IGS guarantees faster convergence of effective DoF to its upper bound relative to PGS for a given rate requirement [164]. Extension to 3-cell SISO-IBC new IA strategy based on IGS and MUD revealed 1.5 DoF (1 by IGS + 0.5 by MUD) whereas higher cells required symbol extensions to achieve proportional DoFs [296].

7.5 IGS Signaling Design

Theoretical analysis of the IGS revealed tremendous payoffs in various MU scenarios such as the X-IC [81, 172, 174, 305], Z-IC [175–178], BC [173, 306], cognitive radio networks [190–192, 294], HWI systems [12, 35, 84, 108] and relay channels [83, 159, 179]. The vast majority of them assumed Gaussian codebooks for more efficient IM [96]. IGS scheme requires optimization of the following parameters based on the underlying presentation.

- Transmit covariance and pseudo-covariance matrices with complex representation which reduces to power and circularity coefficient for small-scale systems.
- Augmented covariance matrix with complex augmented representation
- Composite covariance matrix with real-composite representation

These parameters are optimized in order to maximize some performance metric like ergodic

rate, minimum rate, sum rate, or Pareto rate region. They can also be fine-tuned to minimize ROP, SOP, or MSE etc. This section includes the guidelines and insights of the popular design problems with celebrated optimization techniques.

7.5.1 Closed Form Solutions

The convexity of the IGS parameter optimization problem depends on the objective function as well as constraints. Few objective functions like complete characterization of Pareto boundary in SISO Z-IC and SISO-IC (by restriction to rank-1 transmit covariance matrices) [172, 178] and maximizing sum-rate in SISO Z-IC [176] yield closed form solutions. Alternately, other optimization problems produce piece-wise closed form solutions. Consider an overlay cognitive radio [192] or interweave cognitive radio [294] setting which involve licensed primary and unlicensed secondary link. The conventional PU is transmitting PGS whereas the SU employs optimized IGS to maximize its achievable rate R_s while meeting PU QoS requirements $R_p\left(p_s, \mathcal{C}_x\right) \geq R_{\text{min}}$ (R_p is the achievable rate of the PU). The optimization problem is similar to **7-P1** and tunes IGS transmission parameters of SU i.e., transmit power p_s and circularity coefficient $\mathcal{C}_x$.

$$
\begin{aligned}
\textbf{7-P1}: \max_{p_s, \mathcal{C}_x} \quad & R_s\left(p_s, \mathcal{C}_x\right) \\
\text{s. t.} \quad & R_p\left(p_s, \mathcal{C}_x\right) \geq R_{\text{min}}, \\
& 0 \leq p_s \leq p_{s,\text{max}}, \\
& 0 \leq \mathcal{C}_x \leq 1.
\end{aligned}
$$

SU achievable rate is maximized under PU QoS constraint such that the interference from SU can be limited. This ensures a minimum achievable rate for the PU to maintain QoS. Moreover, p_s is constrained under the transmission power budget $p_{s,\text{max}}$ and $\mathcal{C}_x$ range decides the transmission to be anywhere between proper and maximally improper. Likewise, maximizing total achievable rate in alternate HD-DF relay system [159] is equivalent to

7-P1 with maximum power transmission eliminating any primary rate constraint and can also be solved as piece-wise closed form solution.

Another form of **7-P1** is the ROP minimization problem **7-P2** in cognitive radio setup. The SU optimizes IGS transmission parameters p_s and $\mathcal{C}_x$ to minimize its ROP i.e., $P_{\text{out},s}\left(p_s, \mathcal{C}_x\right) = \Pr\{R_s\left(p_s, \mathcal{C}_x\right) \leq R_{\text{min}}\}$ while maintaining PU QoS $P_{\text{out},p}\left(p_s, \mathcal{C}_x\right) \leq P_{\text{out,th}}$ [191].

$$
\begin{aligned}
\textbf{7-P2}: \ \min_{p_s, \mathcal{C}_x} \quad & P_{\text{out},s}\left(p_s, \mathcal{C}_x\right) \\
\text{s. t.} \quad & P_{\text{out},p}\left(p_s, \mathcal{C}_x\right) \leq P_{\text{out,th}}, \\
& 0 \leq p_s \leq p_{s,\text{max}}, \\
& 0 \leq \mathcal{C}_x \leq 1.
\end{aligned}
$$

Interestingly, problems similar to **7-P2** can also be solved in closed form by investigating the monotonic trend of the objective function with respect to the optimization variables.

7.5.2 Convex Optimization

The closed-form solutions of the convex optimization problems are attractive but not always achievable. For instance, the achievable rate maximization problem **7-P3** of SIMO P2P system with transceiver HWIs reformulates as a quadratic-constraint quadratic programming (QCQP) problem. The IGS transmission parameters enclosed in $\mathbf{s} = [\Re\{\tilde{\sigma}_x^2\} \ \Im\{\tilde{\sigma}_x^2\} \ \sigma_x^2 \ ]^T$ are efficiently optimized using interior point method (IPM) pertaining to the convex quadratic constraints [20].

$$
\begin{aligned}
\textbf{7-P3}: \ \max_{\mathbf{s}} \quad & \text{R}_{\text{SIMO-IGS}}\left(\mathbf{s}\right) \\
\text{s. t.} \quad & \mathbf{A}_1\mathbf{s} \leqslant \mathbf{b}, \\
& \mathbf{s}^{\text{T}}\mathbf{A}_2\mathbf{s} \leqslant \mathbf{0},
\end{aligned}
$$

where $\mathbf{A}_1 = \text{diag}\,[0\ 0\ 1]^{\mathrm{T}}$ and $\mathbf{b} = [0\ 0\ \mathrm{P_T}]^{\mathrm{T}}$ signify the transmission power constraint whereas $\mathbf{A}_2 = \text{diag}\,[1\ 1\ -1]^{\mathrm{T}}$ tracks the magnitude of transmit pseudo-covariance. Thankfully, the complexity of convex optimization problems is polynomial in the problem dimension. However, such fancy convex optimization problems which yield elegant solutions are occasional.

7.5.3 Non-Convex Optimization

Most of the IGS design problems are non-convex in nature and require exponential efforts. Interestingly, all non-convex problems are not hard but lack convexity owing to their inappropriate formulation. As a matter of fact, many non-convex optimization problems admit a convex reformulation using relaxation approaches like semi-definite relaxation (SDR) and sequential convex programming (SCP). Moreover, for separable problems, alternate optimization is preferred over separate optimization if the underlying sub-problems are convex. Furthermore, line-search methods with gradient descent or Newton method converge to a local solution for unconstrained optimization. However, the NP-hard class of non-convex problems requires a different treatment.

Semi-Definite Programming

The joint optimization of covariance $\mathbf{R}_{x_k x_k}$ and pseudo-covariance $\tilde{\mathbf{R}}_{x_k x_k}$ matrices to achieve the Pareto-optimal rates emerges is a non-convex problem like **7-P4** in case of 2U Gaussian SISO-IC [81].

$$
\begin{aligned}
\textbf{7-P4}: \quad & \max_{\mathbf{R}_{x_k x_k}, \tilde{\mathbf{R}}_{x_k x_k}, R} \quad R \\
& \text{s. t.} \quad R_k \geq \alpha_k R,\ \forall k \\
& \qquad\quad 0 \leq \text{Tr}\left(\mathbf{R}_{x_k x_k}\right) \leq P_k,\ \forall k \\
& \qquad\quad \mathbf{R}_{\underline{x}_k \underline{x}_k} \succeq \mathbf{0},\ \text{or}\ 0 \leq |\tilde{\mathcal{C}}_{x_k}|^2 \leq \mathcal{C}_{x_k}^2\ \forall k
\end{aligned}
$$

where α_k is the target ratio between user k's achievable rate R_k and the sum-rate of all users R. Zeng *et al.* suggests SDR to transform non-convex QCQP (as in **7-P4**) to quasi-convex semi-definite programming problem which can either be solved using bisection search or Gaussian Randomization (GR) procedure based on the achievable rank-1 constrained solution [81]. The extension of this problem to MU SISO-IC is a sequence of non-convex minimum weighted rate maximization problems which cannot be solved optimally. Thus, SDR technique along with GR provides an efficient approximation to jointly optimize the transmission parameters [293].

The non-convex Pareto rate-region boundary characterization problem of interfering MAC setup can also be transformed from quadratic to linear form using SDP. Moreover, SDR of rank constraints convexifies the problem at hand which can then be solved using IPM. This problem is equivalent to **7-P4** except that the objective is to maximize R_Σ in place of auxiliary variable R and the first constraint is replaced by the individual link capacity constraint i.e., $\alpha_q R_\Sigma \leq L_q \forall q$. The solution of the relaxed problem can then be projected into the feasible set of the original problem using GR [166].

Another important problem is the sum-power minimization in MIMO FDR for MU interference networks with QoS demands of the communicating pairs. IGS design to tackle RSI and IQI with minimal power requirement leads to an SDP optimization problem with non-convex constraint set. Therefore, linearization of the second concave function in difference of concave (DC) programming problem using Fenchels inequality [282] is suggested [182].

Sequential Convex Programming

Transmit parameters of the source and participating relays can be optimized in a Multi-hop DF-FDR system to improve E2E achievable rate by efficiently mitigating RSI and HWDs. Javed *et al.* propose maximum allowable power transmission $\sigma_m^2 = P_t$ along with the optimized pseudo-variances $\tilde{\sigma}_m^2$ for all transmitting nodes to promote fairness [83].

Considering the achievable link rate between nodes m and n as $R_{nm}\left(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2\right)$, **7-P5** aims at maximizing the minimum link rate, i.e.,

$$\begin{aligned} \textbf{7-P5}: \quad & \text{maximize} \quad \min_n \left\{ R_{nm}\left(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2\right) \right\} \\ & \text{subject to} \quad 0 \leq \left|\tilde{\sigma}_m^2\right| \leq \sigma_m^2, \ \forall m. \end{aligned}$$

The joint optimization **7-P5** turned out to be a max-min fractional programming problem which can be solved in two different ways.

1. Generalized Dinkelbach algorithm (GDA) [307] can transform non-linear fractional programming to non-linear parametric programming and SCP can transform the resultant DC problem to QCQP convex optimization problem [153]

2. Alternately, the logarithmic properties can transform fractional programming problem to DC problem and eliminate GDA step. However, SCP is inevitable to solve further [83].

SCP for DC programming approximates second concave to its first-order Taylor series expansion and solves the resultant convex problem iteratively.

Alternate Optimization (AO)

Joint optimization of transmission parameters is not always manageable therefore some researchers suggest AO method to iteratively improve the approximate solution of maximizing WSR problems. For instance, Lagen *et al.* propose block coordinate descent (BCD) algorithm for AO in MU MIMO-IC with heterogeneous (some LT and WLT) deployment [174] and decentralized processing [149]. They emphasize WLP of the information symbols $\mathbf{b}_k$ from k^{th}-user i.e., $\mathbf{x}_k = \mathbf{T}_{1,k}\mathbf{b}_k + \mathbf{T}_{2,k}\mathbf{b}_k^*$ and WLE from the received vector $\mathbf{y}_k$ at k^{th}-user i.e., $\hat{\mathbf{b}}_k = \mathbf{R}_{1,k}^{\mathrm{H}}\mathbf{y}_k + \mathbf{R}_{2,k}^{\mathrm{H}}\mathbf{y}_k^*$. The goal is to design these precoding $\mathbf{T}_{1,k}, \mathbf{T}_{2,k}$ and estimation matrices $\mathbf{R}_{1,k}, \mathbf{R}_{2,k}$ for all users to maximize WSR by equiva-

lently minimizing W-MSE with weights μ_k, i.e.,

$$\textbf{7-P6}: \quad \min_{\substack{\mathbf{T}_{1,k},\mathbf{T}_{2,k}\\ \mathbf{R}_{1,k},\mathbf{R}_{2,k}}} \sum_{k\in\mathcal{K}} \frac{\mu_k}{2} \log_2 |\mathbf{E}_k\mathbf{F}_k^*|$$
$$\text{s.t.} \quad \text{Tr}\left(\mathbf{T}_{1,k}\mathbf{T}_{1,k}^{\text{H}} + \mathbf{T}_{2,k}\mathbf{T}_{2,k}^{\text{H}}\right) \leq P_k \ \forall k,$$

where, $\mathbf{E}_k$ is the MSE matrix and $\mathbf{F}_k = \mathbf{E}_k - \tilde{\mathbf{E}}_k\mathbf{E}_k^{-*}\tilde{\mathbf{E}}_k^*$. Similarly, WLP and WLE design to maximize WSR for the MU MIMO-BC was attained by equivalently minimizing W-MSE **7-P6** with BCD and AO [163]. Intuitively, this approach renders suboptimal stationary point solutions of precoding and estimation matrices through iterative computation. Additionally, the transmit characteristics of DF-FDR are also optimized using coordinate descent algorithm with AO owing to the monotonic objective function in the individual optimization variables [179]. Furthermore, Lameiro *et al.* also relied on alternating minimization algorithm to design IA precoders and decoders with IGS to provide achievable DoF bounds in 4U SISO-IC [170].

Separate Optimization (SO)

Alternating optimization is the preferred choice with iterative convergence especially if the sub-problems are convex for the subset of optimization variables by treating the remaining variables as constants. Otherwise, we resort to SO. The separate tuning of transmit covariances $\mathbf{R}_{x_kx_k}$ and pseudo-covariances $\tilde{\mathbf{R}}_{x_kx_k}$ for $k = 1, 2$ can be carried out in two ways [81].

1. Exclusive optimization
2. Optimizing $\mathbf{R}_{x_kx_k}$ assuming zero $\tilde{\mathbf{R}}_{x_kx_k}$ and then obtaining $\tilde{\mathbf{R}}_{x_kx_k}$ with given $\mathbf{R}_{x_kx_k}$.

As an illustration, **7-P4** for 2U SISO-IC can also be dealt using SO. The covariance optimization problem emerges as a linear feasibility problem necessitating bisection algorithm for its efficient solution. However, pseudo-covariance optimization with fixed $\mathbf{R}_{x_kx_k}$ is a

set of feasibility problems and thus can be solved as a finite number of second-order cone programming (SOCP) problems [81].

Similarly, the joint optimization of non-convex **7-P4** for MU Gaussian MISO-IC and MISO-BC does not yield a global optimal solution. Therefore, Zeng *et al.* propose SO of $\mathbf{R}_{x_k x_k}$ and $\tilde{\mathbf{R}}_{x_k x_k}$. $\mathbf{R}_{x_k x_k}$ by solving feasibility problem with $\tilde{\mathbf{R}}_{x_k x_k} = \mathbf{0}$ using bisection algorithm. Whereas, $\tilde{\mathbf{R}}_{x_k x_k}$ is obtained with fixed $\mathbf{R}_{x_k x_k}$ by solving equivalent minimum weighted sum-rate maximization (WSR-Max) [295].

Line Search Methods (LSM)

Dualhop DF-FDR systems can adopt improper signaling by finding the optimum weights to maximize the minimum SNR between the two hops under the perfect SI nulling constraint. The parameterization of the adopted problem renders one-dimensional optimization problem which can be efficiently solved using LSM [158].

Algorithms for NP-Hard Optimization

Unfortunately, all non-convex problems cannot be relaxed or convexified, rendering a class of NP-hard optimization problems. In computational complexity theory, these problems are informally "at least as hard as the hardest problems in NP". Generally, WSRMax problems are proven to be NP-hard [308]. Surprisingly, some subclasses of the general NP-hard problem can still be solved in polynomial time [309]. Whereas, others are solved using sub-optimal/approximation algorithms e.g., game-theory based algorithm [310], interference pricing based algorithm [311], gradient descent algorithm [312] with line search methods for unconstrained optimization [158], iterative weighted MMSE based algorithm [313], monotonic optimization frameworks [314, 315], graph theory for combinatorial optimization [316] and SDR for solving non-convex QCQPs with GR by restricting to rank-1 solutions [295]. To summarize, some NP-hard problems can be efficiently solved by combining multiple techniques with certain restrictions or by breaking the problem into sub-problems

and then employing suitable optimization technique to solve each sub-problem.

Here we present an example of such a scenario employing separate optimization and then WMMSE algorithm and alternating direction method of multipliers (ADMM) are used to solve two subproblems, respectively. The non-convex NP-hard WSR maximization in MIMO-IBC poses a huge challenge and cannot be straightforwardly solved using any of the aforementioned techniques. Such problem can be dealt by the separate optimization of transmission parameters where covariance matrices are designed (assuming zero pseudo-covariance) using WMMSE algorithm. Next, target is to design the pseudo-covariance matrices using the pre-designed covariance matrices. This non-convex quadratic programming problem can neither be solved using SDR nor with SCP (as they do not warrant a unique and globally optimal solution). Thus, ADMM-based multi-agent distributed algorithm is suggested to solve an AO sub-problem [162]. However, the global optimal solution of AO sub-problem only guarantees the convergence to a stationary solution of the overall problem. This problem is a classical example of employing both separate and alternate optimization to solve a NP-hard problem.

7.5.4 Intractable Optimization Framework

Unfortunately, some complicated system configurations result in the intractable optimization problems with no definite framework. Therefore, we have to resort to the brute-force attack with exhaustive search. Such exhaustive search comes at the cost of factorial time complexity and may not be desirable. Thus, a fairly simple but suboptimal procedure is to adopt Maximal IGS conditional to IGS superiority.

Exhaustive Search (ES)

IGS parameters can be fine tuned using ES in the feasible domain. For example, the parameters of WLP in downlink [77] and WL RX for uplink [82] in multi-cell massive MIMO systems can be chosen using ES. Similarly, exhaustive user scheduling algorithms com-

bined with IGS in SISO IBC can improve its sum-rate performance [296]. Moreover, the Pareto-optimal transmit covariance matrices for SINR balancing to improve the worst-user rate in the 2U SISO-IC case are also obtained exhaustively [185]. Although, ES leads to a near-optimal solution but it comes at the expense of factorial time complexity [317–319]. Therefore, a rule of thumb is to employ ES when nothing else works.

Maximal IGS (MIGS)

MIGS is usually adopted either to ease optimization overhead or to overcome intractable optimization issues. For instance, Javed *et al.* propose adaptive scheme which switches between PGS and maximal IGS in multi-antenna systems under HWIs based on some switching criterion [20]. However, MIGS is rarely the optimal improper signaling choice e.g., Lameiro *et al.* argues that maximal IGS is the optimal signaling for SU transmission while operating in IGS favorable domain in an underlay cognitive radio network [190]. They further advocated MIGS in underlay MIMO cognitive radio networks in order to restrict the SU interference to protect interference temperature constraint of the PU [320]. MIGS dominance over conventional PGS can be evaluated using majorization theory tools (MTT). Extension of MIMO P2P-I to two-tier HCN (with multiple MIMO Z-ICs) [177] and multi-antenna systems (specifically 2U MIMO-IC) [184] exploit MTT to demonstrate the superiority of MIGS [321]. Furthermore, MTT also help to demonstrate that the eigenvalue spread of augmented covariance matrix is greater for improper signals and becomes maximum for maximal improper signals [36].

7.5.5 Summary and Insights

The applicability and effectiveness of the most popular optimization frameworks is presented in this subsection. However, the theoretical analysis of IGS in more complicated systems and scenarios opens the research areas for other efficient, optimal, and fast-converging optimization techniques.

7.6 IGS Detection and Estimation

In the engineering sciences, the three main branches of statistical signal processing are estimation, detection, and signal analysis [31]. Therefore, various contributions have addressed these issues related to IS as enumerated in Fig. 7.3.

7.6.1 Detection

The problem of detecting the presence of improper complex random signal $s(t)$ from the observed complex signal $r(t)$ under additive noise $n(t)$ is carried out by a simple hypothesis test

$$\mathsf{H}_0 : \; r(t) = n(t), \qquad \mathsf{H}_1 : \; r(t) = s(t) + n(t)$$

Improper Signal in Proper Noise

The detection of an improper signal is based on a finite-dimensional log-likelihood ratio which can be designated as a cascade of an estimator and a correlator. For a zero-mean complex Gaussian signal $s(t)$, Schreier *et al.* propose the detection based on improper version of Karhunen-Loève (K-L) expansion [63]. The performance metric in terms of deflection yields double performance gain when pseudo-covariance is taken into account, generalizing the 3-dB gain of coherent processing over non-coherent processing [63]. On the other hand, for improper complex second-order cyclostationary random signal $s(t)$, Yeo *et al.* suggest properizing frequency shift vectorizer to exploit periodic and symmetric correlations of the complex envelope in the frequency domain. The probability of miss is significantly reduced by the joint utilization of cyclostationarity and impropriety [61].

General Possible Improper Signal in Improper and Colored Noise

Another interesting scenario is the detection of possible improper complex-valued signal common among two or more sensors, in the presence of possible improper and colored

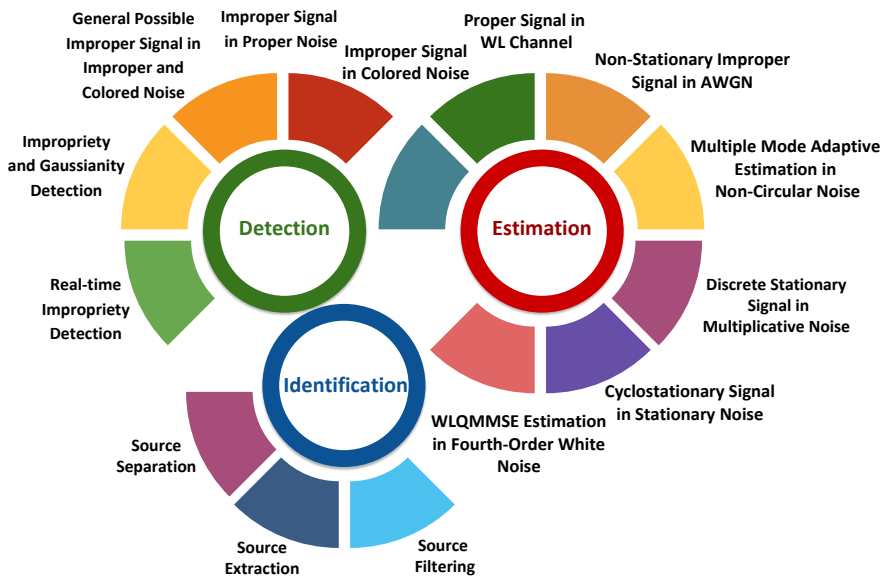


Figure 7.3: Conceptual Classification of IGS Detection, Estimation and Identification

noise. Tugnait *et al.* propose GLRT using asymptotic distribution of a frequency-domain sufficient statistic, based on the discrete Fourier transform of an augmented measurement sequence. Interestingly, they present 133.33% and 70.91% increase in the PoD of improper signals in improper noise relative to that in proper noise at -10dB and -7.5dB SNR, respectively, while achieving 0.1 PFA [156].

Impropriety and Gaussianity Detection

Let us now consider the case when the presence of complex random signal is known but we would like to evaluate if it is improper and/or Gaussian signal or not. Novey *et al.* address this problem of detecting possible improper Gaussian signal $z = x + iy$ using GLRT based on complex generalized Gaussian distribution (CGGD) [322] i.e.,

$$p\left(\underline{\mathbf{z}}, c, \mathbf{R}_{\underline{\mathbf{zz}}}\right) = \frac{\beta(c)}{\sqrt{\left|\mathbf{R}_{\underline{\mathbf{zz}}}\right|}} \exp\left\{-\alpha\left(c\right)\left(\mathbf{z}^{\mathrm{H}}\mathbf{R}_{\underline{\mathbf{zz}}}^{-1}\mathbf{z}\right)\right\}^{c}, \tag{7.3}$$

where, $\beta\left(c\right) = \Gamma\left(2/c\right)/\pi\Gamma\left(1/c\right)^{2}$ and $\alpha\left(c\right) = c\Gamma\left(2/c\right)/2\Gamma\left(1/c\right)$ with Gamma function $\Gamma\left(.\right)$ and shape parameter c. This problem involves two detection mechanisms, 1) Non-

circularity 2) Non-Gaussianity detection based on the following two binary hypothesis tests [323]

Non-circularity	Non-Gaussianity
$\mathsf{H}_0 : \mathbf{R}_{\underline{\mathbf{zz}}} = \sigma_z^2 \mathbf{I},$	$\mathsf{H}_0 : c = 1, \mathbf{R}_{\underline{\mathbf{zz}}} = \sigma_z^2 \mathbf{I},$
$\mathsf{H}_1 : \mathbf{R}_{\underline{\mathbf{zz}}} \neq \sigma_z^2 \mathbf{I}.$	$\mathsf{H}_1 : c \neq 1, \mathbf{R}_{\underline{\mathbf{zz}}} \neq \sigma_z^2 \mathbf{I}.$

The adjusted-GLRT detector [324] performs fairly good as CGGD [323] for Gaussian data ($c = 1$) but inferior to CGGD for sub-Gaussian ($c = 1.5$) and super-Gaussian data ($c = 0.25$). For instance, PoD of 0.75 with CGGD is reported relative to 0.3 with adjusted-GLRT for super-Gaussian data [323].

Real-time Impropriety Detection

Sometimes the static detection of improper signals is inadequate as the underlying applications may require real-time identification of improperness. Thus, Jelfs *et al.* propose collaborative adaptive filters trained by the complex least mean square (LMS) algorithms to detect and track improperness in real-time unlike competing static detectors [46].

Discussion

The detection process of improper signals varies under the presence of proper, improper, or colored noise. Once detected, the interest may reside in the evaluation of DoI as well as the underlying distribution (dictated by the shape parameter). Apart from this static detection, contributions have successfully dealt with the problem of impropriety detection in real-time applications.

7.6.2 Estimation

While detection problems merely identify the presence or absence of improper complex signal, estimation problems include estimating the value of a parameter, or vector of parameters, from a sequence of measurements. Several engineering applications require complex-valued estimations, such as training neural networks [37], passive radar tracking [325], target tracking [326], power systems frequency estimation [29] and fault diagnosis [327] etc. From communication theory perspective, extraction of the transmitted information signal $\mathbf{x}$ from the received observations $\mathbf{y}$ after undergoing a system, namely channel, is sometimes carried out using estimation based on the likelihood $p(\mathbf{y}/\mathbf{x})$. This conditional probability relies on the prior probabilities as per *Baysian* approach whereas other estimation techniques may not require priors e.g., *Frequentists* approach treats $\mathbf{x}$ as a vector of unknown constants [31].

Considering the problem of estimating y from complex observations vector $\mathbf{x}$, linear estimation aims to design $\mathbf{u}$ such that $\hat{y} = \mathbf{u}^{\mathrm{H}}\mathbf{x}$ minimizes linear MSE. Alternately, WLE aims to design $\mathbf{v}$ and $\mathbf{w}$ such that $\hat{y} = \mathbf{v}^{\mathrm{H}}\mathbf{x} + \mathbf{w}^{\mathrm{H}}\mathbf{x}^{*}$ minimizes WL-MSE [39]. MSE of the real data can be accurately carried out by linear estimators whereas WL estimators are generally optimum for complex data. Few relaxations from this rule, which prefer one estimator over the other, are highlighted for a broader picture:

- For jointly circular observations and trivial correlation between observations and estimandum, WL estimators reduce to SL estimators [31]

- For joint circularity between observations however correlated observations and estimandum, WL-MSE still offers better estimates than linear MSE [328]

- For NC observations, it is possible to design $\mathbf{v}$ and $\mathbf{w}$ such that y is uncorrelated with $\mathbf{x}$. This implies zero estimation with the SL procedure and perfect estimation with the WL procedure even when the MSE is zero [328]

- For maximally improper i.e., $\mathbf{x} = \varphi \mathbf{x}^*$ with probability 1, WLE is unnecessary as $\mathbf{x}$ and $\mathbf{x}^*$ carry the same information regarding proper/improper y [31]

- The WL-MMSE estimate of a real signal from a complex signal is always real whereas the LMMSE estimate is generally complex [329]

This section further characterizes estimation problems of different signals with different types of noise using SL and WL estimators based on MMSE criterion.

Discrete Stationary Signal in Multiplicative Noise

Estimation problem of discrete second-order stationary signals can be efficiently solved using WL recursive algorithms. Interestingly, WL predictor proposed by Navarro *et al.* offers significant performance gains. These benefits increase with the increasing improperness of observations and stabilize at a certain value [330].

Cyclostationary Signal in Stationary Noise

Blind estimation is a key concept to facilitate spectral efficiency by eliminating the need of pilot transmission. Napolitano *et al.* propose a blind algorithm to estimate amplitude, phase, relative time delay, and frequency shift of each user transmitting NC signals in a multiple access system. The presented algorithm, based on the cyclostationarity properties, not only provides mean-square consistent estimates of the unknown parameters but is also robust to stationary noise and non-stationary narrowband interference [331].

Non-stationary Improper Signal in additive white Gaussian noise (AWGN)

Next is the estimation of non-stationary improper complex zero-mean random signal in AWGN. Schreier *et al.* suggest WL-MMSE estimator using improper version of K-L expansion to address this problem. Interestingly, this procedure yields perfect estimates rendering arbitrarily large performance gain over SL estimator in the presence of improper

noise. Moreover, WL-MMSE estimation of an improper complex signal in uncorrelated noise can render twice performance advantage over LMMSE estimation at diminishing noise levels [31].

Multiple Mode Adaptive Estimation in NC Noise

The improper stochastic hybrid system with discrete and continuous states can be estimated using WL multiple model adaptive estimation (MMAE) algorithms. These algorithms are based on augmented Kalman filters which are matched to different modes of the hybrid system. WL-MMAE utilizing pseudo-covariance not only converges faster but also offers up to 30% less MSE than their counterparts [130]. Mohammadi *et al.* also extended their work to distributed estimation using diffusion strategies, when a system is observed distributively using an agent/sensor network [66].

Improper Signal in Colored Noise

Estimation of a random improper signal in the presence of colored noise having an additive white part is carried out with Hilbert space theory yielding 10% less MSE with WL estimator as compared to SL estimator [136].

Proper Signal in WL Channel

Underlying channel also impacts the performance of the MSE estimators, irrespective of the correlations between data. For instance, Trampitsch demonstrates the superiority of WL-MMSE over LMMSE for white Gaussian data with complex AWGN. This lead is observed either due to the WL characteristics of the underlying channel even in the absence of correlations between the data or for highly correlated data in a SL channel [233].

WL Quadratic Estimation in Fourth-order White Noise

For complex NC case, a scalar complex y can also be estimated from a measurement $\mathbf{x}$ using the WL quadratic estimator as $\hat{y} = c + \mathbf{g}^{\mathrm{H}}\underline{\mathbf{x}} + \underline{\mathbf{x}}^{\mathrm{H}}\mathbf{H}\underline{\mathbf{x}}$, where c is chosen as $-\mathrm{Tr}\left(\mathbf{R}_{\underline{yy}}\mathbf{H}\right)$ to ensure zero-mean $\hat{y}$ if the observations are also zero-mean. Interestingly, the WL part of WLQ-MMSE is not the same as WL-MMSE. Moreover, the better estimation obtained by WLQ-MMSE relies on the complete statistical information up to fourth order [31].

Discussion

The performance of various estimators can be distinguished using complementary MSE analysis which quantifies the DoI of the SL and WL estimation errors [67]. In a nutshell, the main difficulty in the state estimation comes from structural uncertainty which arises from the lack of knowledge of the true behavior of observations and noise in the underlying system. Therefore, a generalized approach in terms of WL estimators is preferred to accommodate all possibilities and uncertainties. However, this may come at the cost of over-fitting and slower convergence owing to the increased dimensions [39].

7.6.3 Source Separation

Separating one source from a mixture of noisy sources can be carried out using source extraction or source filtering.

Source Extraction

Source extraction aims to recover the original sources from their linear (or non-linear) mixtures in both noisy and noise-free environments. Moreover, blind source separation (BSS) does so with neither explicit knowledge of the sources nor the linear mixing process. Such source extraction is crucial in diverse areas like biomedical engineering, communications, radar, and sonar etc.

Only few contributions cater for the signals with NC PDFs in the source extraction process. For instance, Javidi *et al.* propose second-order complex domain blind source extraction algorithms permitting normalized MSE prediction, at the output of a WL predictor, to be the extraction criterion. Interestingly, the presented framework is suitable for both circular or NC sources with possible improper noise. An important application of BSS is the removal of useful artifacts from the corrupted EEG signals [41]. Another extension of BSS i.e., independent component analysis (ICA), which assumes the statistical independence of the underlying unknown source signals, is the most popular way to separate sources. Some important separability/identification results from the complex ICA based on the circularity coefficients include [39]:

- For the real-valued case, identification is only possible in the absence of two Gaussian sources having proportional covariance matrices in the mixture. Moreover, the knowledge of sample correlation also allows the segregation of Gaussian sources.
- For the complex case, a mixture of improper Gaussian sources with distinct circularity coefficients can be separated using the strong uncorrelating transform even without sample correlation.
- Two Gaussian sources are non-identifiable in a mixture if both of their covariance and pseudo-covariance matrices are proportional to each other.
- A unique maximally improper source in the mixture of sources can be perfectly separated.

Source Filtering

As opposed to extraction, another important phenomenon is the filtering of NC complex signals which can be achieved using various adaptive algorithms based on the WL modeling. Few of these algorithms are:

- Augmented affine projection algorithm [18]
- WL least mean squares to minimize MSE [39]
- WL least stochastic entropy algorithm [39]
- Complex augmented least-mean kurtosis algorithm [332]
- Incremental augmented complex LMS [25]
- Complex-valued Gaussian sum filter [333]

Discussion

Source extraction, especially BSS, is mainly dependent on the structure/distribution of the mixing sources. For instance, two Gaussian sources with proportional augmented covariance matrices or two improper Gaussian sources with similar circularity coefficients cannot be separated. Coming over to the filtering process, the performance of source filters can be quantified using different metrics such as mean square deviation, mean square error, prediction gain, and convergence rate, etc. which mainly depends on their adaptive or batch-wise implementation.

Chapter 8

Practical Realization and Implementation

Overwhelming theoretical performance advantages of IGS in various interference-limited system configurations motivated the researchers to propose the practical improper discrete realization i.e., asymmetric signaling. Finite discrete constellations are the preferred choice for implementation over Gaussian signals owing to their robustness, reduced detection complexity and bounded peak-to-average power ratio [96]. Therefore, the design of new family of asymmetric constellations is imperative for practical realization which is the counterpart of the standard proper discrete constellations i.e., symmetric signaling. Such asymmetric signaling along with the appropriate signal recovery mechanism can significantly reduce the error probability as discussed in the studies presented in Fig. 8.1.

8.1 Asymmetric Signal Design

Apart from the inherently asymmetric signaling schemes like M-PAM, OQPSK and GMSK, what are the possible ways to induce asymmetry in a symmetric discrete constellation like M-QAM and M-PSK? What should be the design objective? How to optimize the design parameters such as transmit power and circularity quotient to meet our objectives? How to generate optimal asymmetric signaling given a certain power constraint and circularity coefficient? Will the generated signaling be capable of achieving superior performance as demonstrated by the theoretical bounds? This section intends to address all of these concerns and highlights the major contributions in this regard. The asymmetric signaling is designed to achieve improper Gaussian capacity except for the 1.53dB shaping loss

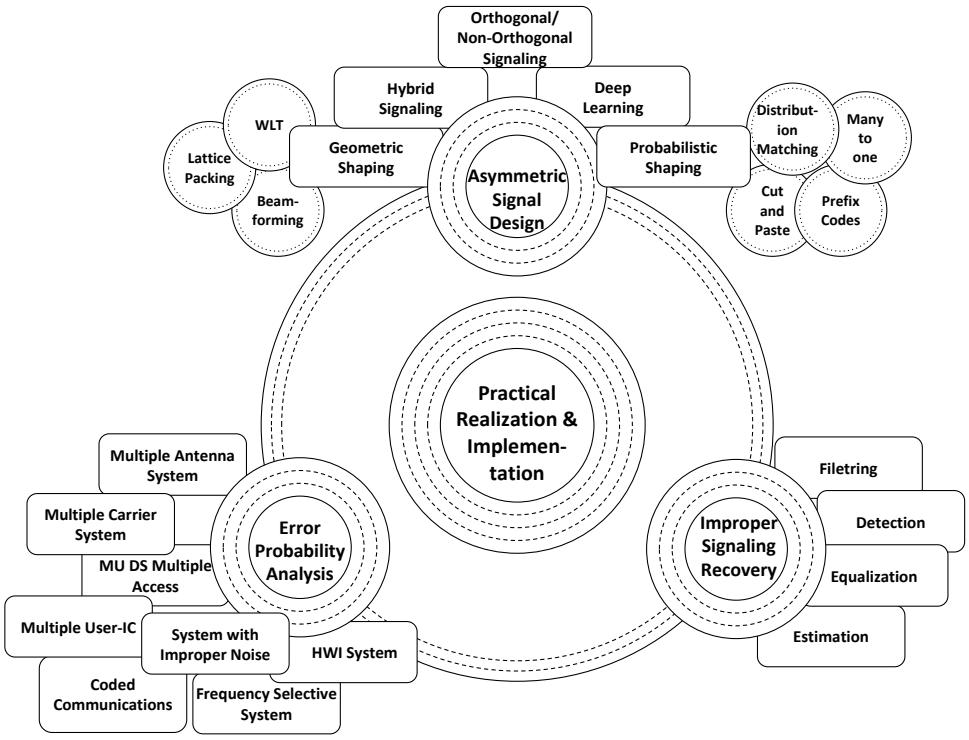


Figure 8.1: Practical Implementation of IS: From Realization to Recovery and Analysis

between IGS capacity and the envelop of capacity curves with sufficiently large M-ary constellations [96]. Moreover, they are optimized to reduce the EP of the constrained systems. Asymmetric signaling may arise from conventional symmetric signaling in the following ways:

1. Probabilistic Shaping (PS): Non-uniform probability distribution of the alphabets/symbols

2. Geometric Shaping (GS): Equally or unequally spaced (due to correlated and/or unequal power distribution between quadrature components of the signals) symbol constellation in a distinct geometric envelop

3. Orthogonal/Non-Orthogonal Sharing: Assigning unequal orthogonal/non-orthogonal resources to users in a MU environment

4. Hybrid Signaling (HS)

5. Deep Learning

Asymmetry provides new tools for the constellation design and can be considered as choices much as symbol separation, the number of bits transmitted per symbol, or power. Interestingly, the introduction of asymmetry into the signal set offers another design freedom which neither affects the bandwidth nor the power requirements of the system [139]. Surprisingly, the digital modulation schemes yield cyclostationary signals with periodic mean, auto covariance and auto complementary-covariance functions. Thus, improved performance can be attained by utilizing this cyclostationary property besides impropriety [178, 334, 335]. In the following, we highlight the famous shaping techniques and procedures to design appropriate asymmetric signaling for a given application.

8.1.1 Probabilistic Shaping

Given a fixed number of symbols and the symbol locations, an asymmetric constellation can be obtained by adjusting the symbol probabilities [100]. Therefore, PS maps equally distributed input bits into constellation symbols with non-uniform distribution [101]. Intuitively, manipulating symbol probabilities and deviation from uniform distribution will result not only in some entropy loss but also added complexity in the encoding/decoding process. Despite this implementation penalty, the attained performance gains are totally worth it. So, what should be the ideal non-uniform probability distribution and how can we attain it? Intriguingly, the Gaussian probability distribution is the ultimate goal to approach the channel capacity bounds but this comes with a number of practical problems. Therefore, multiple transformations are presented to tackle this issue including prefix codes [196, 197], many-to-one mappings combined with a turbo code [198], distribution matching [199, 202] and cut-and-paste method [200]. Coded modulation scheme with PS aims to remove the shaping gap and coarse mode granularity problems [202]. Interested reader can read [132] for the design guidelines of asymmetric signaling in the coherent Gaussian channel with equal signal energies and unequal a priori probabilities. Probabilistic amplitude shaping is

another concept that can only be used for square QAM, which greatly limits its application [203].

8.1.2 Geometric Shaping

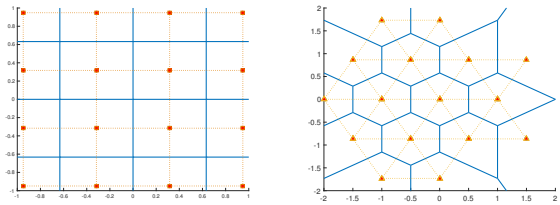
Geometric shaping can be characterized in two distinct ways. First, uniformly spaced symbols within distinct geometric envelop. Second, non-uniformly spaced symbols pertaining to either non i.i.d quadrature components or intentional asymmetric placement in rectilinear constellations. GS requires unconventional partitioning without any loss in entropy. Isaka *et al.* emphasize unconventional signal set partitioning for asymmetric constellations to achieve unequal error protection capability with multilevel coding and multistage decoding [141]. Nonetheless, moving some symbols close to each other may result in more erroneous symbol decisions depending on the underlying application [100]. GS can be realized using the following well-known methodologies.

Widely Linear Transformation

WLT is the most popular way of transforming proper signaling to improper one in order to exploit the additional freedom offered by complementary covariance matrix [336]. The extension of WLP from Gaussian code books to discrete constellations is paving the way for its practical utilization in different applications. The simplest design of asymmetric constellations with complex symbols $v = v_{\mathrm{I}} + iv_{\mathrm{Q}}$ from standard symmetric discrete constellation with symbols $x = x_{\mathrm{I}} + ix_{\mathrm{Q}}$ for a given circularity coefficient κ and circularity angle ϕ can be attained by the following WLT

$$v = \sqrt{\frac{1}{2}\left(1+\alpha\right)}x + \sqrt{\frac{1}{2}\left(1-\alpha\right)}\exp^{i\phi} x^{*}, \tag{8.1}$$

where, $\alpha = \sqrt{(1-\kappa^2)}$ and $\phi \in [0, \pi/2]$. The optimal κ and ϕ can transform a given M-ary symmetric constellation to M-ary asymmetric constellation [96]. This transformation



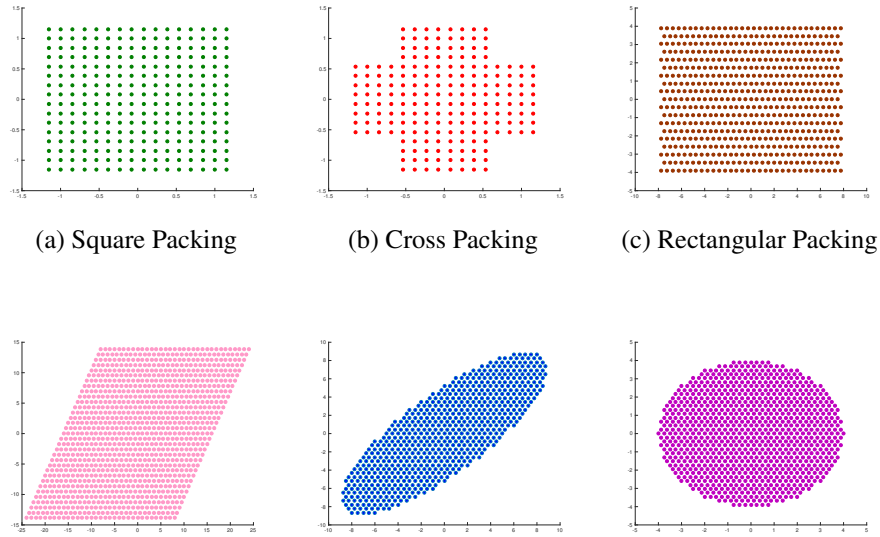
(a) Square Lattice (b) Hexagonal Lattice

Figure 8.2: Celebrated 16-QAM Lattice Structures

has been proven helpful in various applications e.g., WL digital beamforming is employed in massive antenna arrays suffering from IQI in RF chain. Therefore, WL extension of the well-established minimum variance distortionless response beamformer for M-QAM modulation is a promising candidate for complex Gaussian interference, unwanted mirror beam and RF imperfections suppression [69]. Furthermore, adaptive algorithms by combining WL processing and set-membership filtering techniques are proven to improve sensor array processing when the signals under study are second-order NC such as BPSK modulated signals [337]. Likewise, WLP in spatially multiplexed MIMO-FBMC system with OQAM provides lower BER as compared to LP at high SINR [338]. MU two-way MIMO-AF relay system exploits the NC transmitted signals like BPSK using WLP to achieve improved system performance with minimal relay power [181]. Furthermore, single user and MU MIMO communications systems employing asymmetric modulation like M-ary ASK and OQPSK depict superior performance with improved ZF and MMSE precoders without loss of spectrum efficiency [339].

Lattice Packing

Numerous studies have supported the concept of efficient modulation technique by packing a particular lattice structure in some geometric shape [197, 340]. Lattice packing is a two-fold procedure. Firstly, lattice structure is chosen from the square, rhombic, triangular or hexagonal lattices [341, 342]. A square and rhombic lattice is the periodic arrangement of discrete constellation points/symbols at the corners of the square (as depicted in Fig. 8.2a)



(a) Square Packing (b) Cross Packing (c) Rectangular Packing

(d) Parallelogram Packing (e) Elliptical Packing (f) Circular Packing

Figure 8.3: Geometrical Packing/Mask Structures

and rhombus shape [341]. Unlike these quadrilaterals, triangular lattice has constellation points at the vertexes of contiguous equilateral triangles. This lattice is sometimes referred as hexagonal lattice owing to the hexagonal Voronoi decision region around internal lattice points as shown in Fig. 8.2b [342]. Square QAM (SQAM) is preferred for the simple ML detection mechanism whereas triangular QAM (TQAM) is preferred for power efficiency [342, 343]. Park *et al.* report the asymptotic power gain of 0.5799dB with TQAM over SQAM, identical peak-to-average-power ratio for significantly large constellation size M and a significant reduction in EP with tolerable detection complexity in an AWGN channel [343]. Therefore, hexagonal QAM (HQAM) is preferred for various applications, including advanced channel coding [344], multi-antenna systems [345], multicarrier systems [346], physical-layer network coding [347], and optical communications [348].

Next step is the distribution of lattice around origin, which can be packed in square, cross, elliptical, circular, rectangular, or parallelogram envelops [197, 340, 349]. For instance, Fig. 8.3a and 8.3b depict square and cross packing of the square lattice, respectively. However, Fig. 8.3c-8.3f pack hexagonal lattice in elliptical, circular, rectangular,

and parallelogram envelops, respectively. Asymmetric constellations formulated by the elliptical packing are based on the geometric interpretation of the circularity quotient [57]. Accordingly, the modulus and phase of its principal square-root depict the eccentricity and angle of orientation of the ellipse defined by the covariance matrix. Thus, vanishing eccentricity implies zero circularity quotient (symmetric case) whereas maximum eccentricity implies that the circularity quotient exists on the unit complex circle [57].

Recently, Fernández *et al.* suggest elliptical packing of the hexagonal lattice to realize optimal asymmetric constellations for a certain circularity coefficient [340]. The optimal constellation is the overlap between the translated hexagonal lattice and the rotated ellipse (as per circularity quotient). The idea is to capture exactly M-constellation points in the intersection area and then apply WLT to transform it to quasi-hexagonal constellation rendering the required circularity coefficient. They are optimal as they yield lowest EP under average power constraint in an AWGN channel especially at high SNR and large constellation size. Authors claim significant SNR gains resulting from this design scheme as compared to WL transformation for two reasons: *shaping loss* and *packing loss* [340]. Shaping loss of WL transformed constellations is around 1.53dB with respect to the IGS which is equivalent to proper square M-QAM constellation. This limitation is addressed by the hexagonal lattice which offers 0.6dB gain over a rectangular one with the same boundary being the densest lattice in 2D [197]. On the other hand, packing loss is pertinent to the parallelogram envelope of WLT asymmetric constellation in place of the optimal elliptical envelope.

Beamforming/Precoding

Numerous other transformations are based on the fine tuning of the basic constellation parameters to achieve minimum EP with a given power budget. For instance, Zhang *et al.* propose asymmetric 4-PSK constellation design with TCM by calculating the optimum angle $\alpha \in [0, \pi]$ (with $\alpha = \pi/2$ yielding the symmetric 4 PSK) that minimizes the EP

bound [350]. Similarly, Subramaniam *et al.* introduce asymmetry in 8-PSK to increase the minimum product distance reducing error events in TCM [351]. For colored noise compensation, Taubock prefers optimally rotated rectangular symbol constellations over common quadratic QAM in order to minimize capacity loss and SEP [74].

Unlike the contributions which propose modulation specific designs, [35] and [305] propose asymmetric constellation design based on the minimum Euclidean distance aiming at minimizing the maximum PEP or SER irrespective of the underlying modulation scheme. Nguyen *et al.* optimize the precoding matrices of all users in order to minimize the maximum PEP or SER in MU SISO-IC. Alternately, precoding matrices can be designed in order to minimize the sum MSE, maximum MSE, minimize interference leakage or maximize SINR. However, the schemes based on minmax-MSE perform inferior to minmax PEP/SER especially for higher modulation transmission with higher interference [305]. Moreover, Javed *et al.* design asymmetric modulation for HWI transceiver systems by separately optimizing the rotation and translation matrices which jointly formulate the transmit precoding matrix [35]. Likewise, pairwise optimization algorithm transforms conventional constellations like 8PSK star-8QAM, set-partitioning-8QAM and circular-8QAM to asymmetric optimal constellation in order to minimize the BER [352].

In general, researchers mainly rely on maximizing the minimum Euclidean distance of the constellation ($d_{\min}$) as it is the building block of various performance metrics like mutual information, MMSE and SEP. They all have asymptotic behavior which is proportional to the Gaussian Q-function $Q\left(\sqrt{\mathrm{SNR}}d_{\min}/2\right)$ [96].

8.1.3 Orthogonal/Non-Orthogonal Sharing

Another form of asymmetry, which can be induced to attain added benefits, is through the non-uniform allocation of orthogonal resources like time slots and frequency bands. For instance, asymmetric time sharing is assumed to be a potential candidate in future generation of wireless communications for enhanced performance in some interference limited

scenarios. On the other hand, some NOMA schemes like sparse carrier multiple access that is going to be an essential part of 5G communications for many excellent properties i.e., shaping and diversity gain by sparse codebooks, resilient to inter-user interference, and robust to codebook collision can also be considered as a form of asymmetric signaling [353].

8.1.4 Hybrid Signaling

We characterize the joint employment of any two or more types of asymmetric signaling schemes as hybrid signaling. There is an ongoing debate on the superiority of one type of asymmetric scheme over the other depending on the underlying application and employed modulation schemes. Both PS and GS were successfully employed in optical communications [198]. From one perspective, PS outperforms circular based GS in terms of shaping gain for the same number of constellation points [101]. However, another point of view is the superior performance of GS 16-QAM constellations over its PS counterpart [354]. A striking way out of this debate is to employ hybrid probabilistic and geometric shaping (PS/GS) concept to bridge the gap to the Shannon limit [355]. Albeit hybrid PS/GS is popular in optical communications but this concept is yet to find its standings in the wireless communication systems. In contrast to the PS, where redundancy improves power efficiency, Hybrid PS/GS capitalize on redundancy by introducing a transmitted signal structure that improves Euclidean distance and reduces SER [101]. Hybrid PS/GS of any two-dimensional signal constellation outperform the probabilistically shaped as well as regular constellations with universal distribution matchers for asymmetric M-QAM [356] and multi-dimensional coded modulation format using amplitude-phase shift keying for single-stage [101] and multi-stage [357] nonlinearity compensation. Other optimal strategies involving the combined PS/GS to shape circular 64-QAM constellation attain 1dB sensitivity gain and 28% gain in transmission reach over compared to conventional 64-QAM in both linear and nonlinear regime of wavelength division multiplexed systems [358].

To conclude, dense hexagonal packing with optimal circular boundaries yields around

1dB improvement over PAM in a band-limited channel [197]. Moreover, source coding with non-uniform probabilities provide more than 1dB gain which is alternately achievable using higher dimensional modulation in an uncoded system. Furthermore, channel coding with simple block/trellis coding can render coding gains of the order of 3dB [197].

8.1.5 Deep Learning

Although communications field has matured over time, however there are scenarios when the accurate mathematical modeling and rigorous analysis seems unattainable or intractable. Such scenario may arise while modeling and equalizing various types of channels and hardware imperfections. Equivalently, the optimal signaling design and detection schemes is another major challenge. To address this concern, machine learning especially its branch deep learning has demonstrated some promising results. For instance, a communication system can be trained as an auto-encoder treating TX, channel and RX as one deep neural network. The auto-encoder learns appropriate symbol representations of information messages to tackle channel imperfections i.e., noise, distortion, and fading, etc. in order to attain small error PoD. Interestingly, the performance enhancement is achieved when the auto-encoder learns asymmetrical constellation in 2U IC setting [359].

8.2 Asymmetric Signal Recovery

Communication systems with asymmetric transmission and/or improper noise require appropriate treatment in terms of equalization, estimation, filtering and detection to account for the induced improperness. Therefore, WL processing models are incorporated to design several estimation [31], filtering [37], and detection [360] algorithms. WL processing is utilized in communication systems that apply asymmetric constellations e.g., [65, 339, 361–363] , and/or encounter improper noise, e.g., [75]. Additionally, it is exploited in linear-dispersion STBC, e.g., [360, 364, 365], and, recently, in IC and BC [164, 172].

8.2.1 Equalization

In band-limited, high data rate digital communication systems, equalizers flatten the channel frequency response in order to minimize channel distortion on the transmitted information symbols. Asymmetric transmission, improper noise/interference, and/or frequency-selective channel necessitate the employment of WL-equalizers in place of SL-equalizers. For example, real-valued data transmission over complex-valued frequency-selective channels producing ISI necessitates MMSE based WL equalizers for RXs with or without decision feedback [65]. Besides this, STBC transmission with conventional equalizers and decoders require equal receiver and transmitter antennas for detection. However, this kills the main purpose of STBC to achieve pure transmit diversity. Employment of MMSE based WL equalizers with/without decision feedback can overcome this restriction to further enhance the achievable data rates [365]. Multicarrier systems with finite impulse response (FIR) linear precoders and asymmetric constellations offer an intrinsic source of redundancy, which aids in efficient design of WL-ZF universal equalizers for immaculate symbol recovery in FIR-channel with narrowband interference [339].

8.2.2 Estimation

Estimation involves approximating/estimating the value of an entity from a sequence of observations or measurements. Conventional linear estimation aims at designing $\mathbf{u}$ to approximate y from a set of observations stacked in $\mathbf{x}$ using the linearly combination $\hat{y} = \mathbf{u}^{\mathrm{H}}\mathbf{x}$. Still it fails to exploit the information hidden in the correlation between the observations and their complex conjugate. Thus, WLE designs $\mathbf{v}$ and $\mathbf{w}$ to better approximate y as $\hat{y} = \mathbf{v}^{\mathrm{H}}\mathbf{x} + \mathbf{w}^{\mathrm{H}}\mathbf{x}^{*}$. WLE is advantageous in various applications in signal processing [36, 366], communications [59], power systems [53], biomedical engineering [48] and renewable energy [37]. WLE can be carried out using various variants like WL-MMSE estimation [328], WL minimum variance distortionless response (WLMVDR) estimation [26], [27], [29], [77], and the WL-LMS algorithm. The practical application of these estimators

in modern digital communication systems with asymmetric constellations and/or improper noise/interference are overwhelming. The important class of space-time block-codes for MIMO channels i.e., linear dispersion codes is constructed from linear combination of input symbols and their complex conjugates. Linear dispersion codes like orthogonal, quasi-orthogonal and V-BLAST codes utilize WL-MMSE estimates of transmitted symbols as the sufficient statistics for ML detection of these symbols [360]. IQI aware WL-MMSE RX with channel estimation and data detection capability outperforms its linear counterpart in uplink multicell massive MIMO systems. It jointly suppresses MU interference, pilot contamination, and IQI while performing closely to the MMSE RX in a perfect system without IQI [82]. A linear RX cannot reap the maximum benefit from WL precoding in a MU MIMO-BC. Therefore, WL estimator is required to maximize the weighted sum-rate with limited power budget of the participating base station [163].

8.2.3 Filtering

Filtering separates any entity or group of entities from a mixture/amalgam based on their distinct characteristics. Analogous to other signal recovery schemes, optimal WL filtering is superior to SL filtering in the NC context. It is widely used in applications such as detection [367], prediction [95], modeling [55], interference cancellations like co-channel interference (CCI) [368] and narrow-band interference (NBI) and equalization for SISO [59,65], MIMO [365], and DS-CDMA systems [369, 370]. Asymmetric complex nature of BPSK signals has motivated researchers to apply WL filtering at the RX for improved detection [55]. Moreover, co-antenna interference in the generalized MIMO systems (transmitting complex conjugates along with actual data) can be effectively attenuated using an iterative RX with WL filters. In such linear space time mappings, WL filter also accounts for the non-circularity arising inherently within an iterative RX [371]. Additionally, multiple CCI cancellation in PAM/QAM modulated SIMO is achieved using WL filtering for demodulation. Surprisingly, interference cancellation ability of WL-ML RXs is only dependent on

interferer modulation type and RX antennas but irrespective of modulation scheme of the desired signal [372]. Similarly, fully WL-MMSE filter opens new perspectives for intra-network and external interferences management for 4th generation radio communication cellular networks using the Alamouti scheme [373]. WL array RXs are also considered optimal for the demodulation of BPSK, MSK, and GMSK Signals under improper interferences. Effective single antenna interference cancellation can be attained for these modulation schemes in GSM cellular networks using WL filtering [59]. Furthermore, Hellings *et al.* suggest block-Hankel-skew-circulant structured matrices for appropriate processing of WL filters and asymmetric signaling in MIMO IC [374].

8.2.4 Detection

Unlike the IGS and impropriety detection, asymmetric detection involves the recovery of the transmitted signals inheriting asymmetric structure from asymmetric transmission and/or asymmetric noise/interference. The gains of such WL detection are two folds: 1) It accounts for the asymmetric characteristics for improved detection. 2) It efficiently suppresses the effects of asymmetric noise/interference. These perks cannot be attained by the conventional detectors neglecting SOS. For example, WL structures for blind MUD in synchronous DS/CDMA systems can effectively suppress both narrowband and wideband multiple-access interference [140]. Similarly, WL RX for OQPSK modulated DS/CDMA system employing least-mean square algorithm efficiently mitigates both symmetric/asymmetric interference [60]. Furthermore, asynchronous DS/CDMA systems employing BPSK requires a new family of MMSE detectors to jointly suppress multi-access and external interference by fully exploiting the SOS through poly periodical processing [56]. Direct application of WL processing may not be suitable in some scenarios e.g., MIMO FBMC/OQAM. Therefore, a two step RX; first with linear and later WL processing may help in removing the intrinsic interference which keeps us from taking the full advantage of WL-MMSE RX [375].

Discussion

WL extension of various RXs may or may not be helpful in attaining optimum performance depending on the type of application, employed modulation/coding schemes and degrading noise/interference. For instance, WL matched filter solution does not take account noise and is still linear solution [36]. Alternately, optimum ML detection and WL-MVDR estimation only requires adjustments for improper noise/interference and is irrespective of the symmetric/asymmetric transmission. Whereas, other detectors like WL-MMSE, WL-ZF, and WL-LMS filtering may require the propriety characterization of the transmitted signals as well as received noise. Importantly, the performance gains of WL model diminish when the underlying system is linear [36].

8.3 Error Probability Analysis

EP is a tangible measure used to fairly judge the performance of communication systems. EP captures prime system details (e.g., modulation scheme, RX type, symbol constellation, etc.) and is considered the most revealing metric about the communication system performance [288]. It is the probability of receiving the erroneous information and can be studied through pairwise EP (PEP), bit EP (BEP), symbol EP (SEP), and frame/block EP (FEP) etc. In the context of wireless networks, EP has mainly been studied and conducted for symmetric transmission and additive white proper Gaussian noise. This section summarizes the contributions that tackle EP analysis for asymmetric constellations with/without improper noise/interference in various system configurations.

8.3.1 Multiuser Direct Sequence Multiple Access Systems

Various trellis-coded systems with asymmetric PSK transmission outperform their counterparts with symmetric PSK signaling. For example, the performance of the DS spread-spectrum multiple access (SSMA) system employing trellis coded modulation (TCM) with

asymmetric signal constellations outperforms its competitor symmetric signal constellations by 58.02% for 3U DS/SSMA system and offers 1dB SNR reduction at 1e-10 BER. Similarly, WL MUD offers perks over SL MUD in a DS/CDMA system with QPSK modulation suffering from BPSK NBI. BER reduction of 100% and 96% is observed by WL-MUD RX w.r.t conventional matched filter and L-MUD, respectively, for the first user out of 6 CDMA users. Alternately it offers EbNo gains of 2dB at 1e-5 BER w.r.t L-MUD RX [140].

8.3.2 Coded Communication Systems

Interestingly, deep learning for the physical layer suggests the joint optimization of TX and RX components by modeling a communications system as an auto-encoder and training it using stochastic gradient descent. This approach to minimize block error rate (BLER) outperforms all the existing well-known schemes by yielding the asymmetric modulation as the optimum one. Surprisingly, for a communications system employing BPSK modulation and a Hamming code, the trained auto-encoder outperforms the uncoded and coded scheme with hard decision (HD) by 98.64% and 96.25%, respectively. It further offers EbNo gain of 2dB at 8e-4BLER and 1.5dB at 3e-4BLER over uncoded and coded scheme with HD, respectively. The gains further increase with increasing the batch size while reducing the learning rate during training. Unanticipated asymmetric learned constellations of auto-encoder performs equally good as the coded scheme with ML detector without any prior knowledge [359]. Similarly, for a 2U IC, the auto-encoder and time-sharing have identical BLERs for (1, 1) and (2, 2), the former yields significant gains of around 0.7dB for (4, 4) and 1dB for (4, 8) at a BLER of 1e-3. O'Shea *et al.* further claim that auto-encoder with RTN can outperform differential BPSK with ML estimation and Hamming (7,4) code in a multipath fading environment.

Asymmetric TCM (ATCM) provides better SNR gains over the traditional symmetric TCM (STCM) for very low data rate systems especially with small number of users in high

SNR and less drastic fading circumstance [376]. Similarly, asymmetric 8-PSK signal sets in 4-state and 8-state rate 2/3 TCM outperforms symmetric 8-PSK TCM owing to the increase in the minimum product distance. ATCM reduces BER up to 20%, 76.67% and 50% relative to STCM for Rayleigh, Rician and light shadowed Rician (S-Rician) channels, respectively. Moreover, an improvement of about 0.3-0.4dB around BER 1e-5 is observed over the Rayleigh and shadowed Rician channels. Whereas, it reduces to 0.2-0.3dB for the Rician channels. Subramaniam *et al.* claim that the improvement of 0.2-0.4dB signifies 5-10% savings in power, owing to the enlarged minimum product distance and free Euclidean distance by the ATCM [351]. Another approach to design ATCM is to minimize BER (or BER bound) instead of maximizing the free Euclidean distance. Zhang *et al.* present 75% and 55% reduction in BER with 2-state and 4-state ATCM relative to STCM. Furthermore, the optimum ATCM achieves 0.5dB SNR gain at 1e-5 BER and 0.1dB SNR gain at 1e-7 BER with 2-state code and 4-state codes, respectively [350]. Moreover, asymmetric signals with non-uniform spacing outperform symmetric signals with uniform spacing in the trellis-coded systems. For instance, two-state trellis-coded optimum asymmetric 4PSK offers 99.83% less BER and 0.5dB EbNo gain at 1e-5 BER as compared to its symmetric counterpart. Additionally, two-state trellis-coded asymmetric 4-AM 98.33% less BER and 1dB EbNo gain at 1e-5 BER as compared to its symmetric counterpart [139]. Divsalar *et al.* emphasize that the asymmetric signaling does not affect power or bandwidth requirements of the system. Thus, rendering BER performance gains at little or no cost.

8.3.3 Multiple Antenna Systems

Conventional iterative RXs which are optimal for the symmetric modulations, such as M-QAM and M-PSK, are suboptimal for the asymmetric modulations, such as M-ary ASK, OQPSK (for which $\mathrm{E}\left[\mathbf{s}\mathbf{s}^{\mathrm{T}}\right] \neq \mathbf{0}$) in uncoded MIMO systems. Therefore, Xiao *et al.* proposed a novel iterative RX with various decoding strategies like ZF, MMSE and SIC etc. Accommodating the asymmetric behavior of the 4-ASK and OQPSK signal constellations

offer BER percentage decrease up to 99.25% and 99.58% with ZF RX and 97.78% and 97.50% with MMSE RX, respectively. In other terms, the proposed scheme SNR gains of 8dB at 2e-3 and 12dB at 1.5e-3 with ZF RX and 6dB at 2e-4 and 8.5dB at 3.5e-5 with MMSE RX, respectively, with 4-ASK and OQPSK modulations in 4x4 uncoded MIMO system. The percentage decrease in BER with 4-ASK modulation employing MMSE based novel iterative RX in 4x3, 4x4 and 4x5 uncoded MIMO system is 98.9%, 98% and 94%, respectively. Interestingly, accounting for asymmetry not only achieves superior performance and faster convergence than the conventional systems but renders equivalent effect of increasing receive diversity order [377].

Co-antenna interference suppression in generalized MIMO systems with linear dispersion codes like V-BLAST achieve up to 30% less frame error rate with WL detector (WLD) as compared to SL detector with iterative RX. Furthermore, it renders 0.3dB SNR gain at 2e-2 FER [371].

WL filtering is beneficial for effective demodulation of PAM/QAM modulated SIMO systems suffering from multiple data-like CCIs. Kuchi *et al.* argue the tradeoff between diversity advantage and interference cancellation in a flat Rayleigh fading channel. Interestingly, WLF offers SER reduction up to 83.53%, 98.15%, 80% and 42% for desired-CCI combinations of 1x1 QPSK-BPSK, 1x1 BPSK-BPSK, 1x1 BPSK-PAM+QAM and 1x2 BPSK-2BPSK, respectively. In other words, the respective SNR gain of these four examples are given by 7dB at 8e-4 SER, 12dB at 1.3e-1 SER, 4dB at 1.4e-3 SER and 3dB at 1e-3 SER, respectively. Compelling WL RX with N-antennas is capable of rejecting any combination of M_1 PAM and M_2 QAM interferers satisfying $M_1 + 2M_2 < 2N$, whereas SL RX can only reject up to $M_1 + M_2$ interferers with $M_1 + M_2 < N$ [372]. In a nutshell, WL RX can handle more CCIs relative to SL RX while maintaining a certain SER.

8.3.4 Multi-carrier and Single-carrier Systems

WL processing with asymmetric transmission is advantageous in both multi-carrier and single-carrier systems. For example, WL-MMSE is beneficial over LMMSE for unique word (UW)-OFDM systems with ASK modulation which introduces asymmetry in the system. Interestingly, WL-MMSE yields up to 97.4% and 94.5% less BER relative to LMMSE for 8-ASK modulation over AWGN and IEEE indoor channels, respectively. Additionally, it offers SNR reduction of 1dB at 2e-5 BER for 16-ASK modulation and 3dB at 1e-6 BER for 4-ASK modulation over AWGN and IEEE indoor channels, respectively [233]. Filter bank multicarrier (FBMC) scheme to combat frequency selective fading in MIMO systems employing OQAM modulation require WL processing to account for NC transmitted signals. Caus *et al.* present two conditions i.e., number of streams $(S) \leq N_{\mathrm{R}} \leq 3, N_{\mathrm{T}} \geq N_{\mathrm{R}}$ and Coherence BW $>>$ Subcarrier Spacing, when linear processing is superior to WL processing. Otherwise, linear processing gives an error floor and thus, WL processing is the preferred choice in terms of BER, especially in low noise. Intuitively, for a $N_{\mathrm{R}} = 2$, $N_{\mathrm{T}} = 3$ and $S = 2$ MIMO system, WL processing to minimize the sum MSE outperforms it's counterpart by 91.25% and 99.9% when the underlying channel follows international telecommunication union (ITU): Pedestrian A guidelines (PAG) and Vehicular A guidelines (VAG), respectively. Moreover, the respective SNR gains are obtained as 2.5dB at 8e-6 BER and 5dB at 5e-4 BER, respectively, at high SNR regime [338]. Alternately, Cheng *et al.* suggest two step RX based on linear and WL processing to reap the benefits of both domains. First step cancels intrinsic interference that prevents to reap maximum benefit from WL processing and second step employs WL-MMSE RX to cater for RSI. This technique renders 97.6% and 99.6% BER reduction with partial and complete interference cancellation in first step, respectively, in 4x4 MIMO FBMC/OQAM system. Equivalently, the respective SNR gains are 9dB and 10.5dB at 2.5e-3 BER [375]. Similarly, MIMO single carrier (SC) block transmission suffering from inter-symbol interference (ISI) and data-like CCI requires WL filters with feed forward (FF) and noise prediction feedback (NP-FB) at-

tributes to outperform its linear counterparts. Both suboptimal and optimal SC frequency domain equalizers with WL filters are proposed for OQPSK modulated data and CCI. Surprisingly, WL filters outperform SL filters by 98.50% and 99.67% with optimal equalizer and by 97.10% and 99.75% with low-complexity suboptimal equalizer in the presence and absence of CCI, respectively. Moreover, both equalizers offer EbNo gains of around 20dB at 6e-2 and 8dB at 1e-5 with and without CCI, respectively, for a 2x2 MIMO SC block transmission with a cyclic prefix [151].

8.3.5 Systems with Frequency Selective Channels

equalization schemes based on WL processing outperform their linear counterparts in frequency selective channels (FSC) with underlying ASK, OQAM, BMSK type modulation schemes. Gerstacker *et al.* propose FIR filters for WL-MMSE equalization without and with decision feedback , termed as MMSE-WLE and MMSE-WDFE, respectively. The novel MMSE-WLE and MMSE-WDFE RXs provide BER reduction up to 96.07% and 56% at 20dB EbNo relative to their linear counterparts MMSE-LE and MMSE-DFE, respectively. Alternately, they offer respective EbNo gains of 5dB and 0.7dB at 1e-4 BER for real-valued transmission over complex ISI channel [65]. On the other hand, if the zeros of the channel are far from unit circle then WLP can only offer small-to-moderate gains. Gerstacker *et al.* further extended their findings of WLE to STBC over frequency selective channels. Numerical examples show BER reduction of 93.5%, 72.5% and 74.29% with 2x1 STBC WLE employing 8PSK, 2x1 STBC WDFE employing BPSK and 2x2 STBC WDFE employing 8PSK relative to SISO-LE, SISO-DFE and 1x2 SIMO-DFE over rural area (RA), hilly terrain (HT) and typical urban area (TU), respectively [365].

8.3.6 Systems with Improper Noise

Similarly for general improper Gaussian noise (IGN), Alsmadi *et al.* present an optimal detector for MIMO system with space shift keying (SSK) RX. The accommodation of im-

proper nature of the additive noise in optimal detection results in the average EP percentage reduction up to 0%, 20%, 32% and 44% for RXs affected by proper Gaussian noise, non-identical uncorrelated IGN , identical correlated IGN and non-identical correlated IGN, respectively, relative to the traditional ML detector considering PGN. Moreover, respective SNR gains are given by 0dB, 1dB, 2dB and 3dB at 5e-3, respectively [19].

8.3.7 Hardware Impaired Systems

Evidently, HWI systems under IQI and additive distortions alter the symmetry of the signals under study. Thus, accounting for the induced asymmetry in the detection process presents benefits in terms of average BER reduction up to 8.33%, 38.64% and 52.78% for low, moderate and high transmit distortion levels, respectively. Alternately, it renders SNR gain up to 10dB at 1.2e-1, 9dB at 4.4e-2 and 4dB at 1.7e-3, respectively, for the three aforementioned transmit distortion cases [378]. Asymmetric signaling is favorable in suppressing improper accumulated noise and self-interference resulting from the hardware imperfections like IQI and non-linear distortions. Thus, Javed *et al.* suggest to employ asymmetric QAM to minimize maximum PEP or SEP. The proposed asymmetric transmission renders up to 84% and 97.8% reduced average SER with respect to symmetric transmission with ML detector at 10dB and 20dB SNR, respectively. Additionally, asymmetric transmission attains the same SER performance (8e-3) with 10dB SNR as that of symmetric modulation with 20dB SNR [35]. Moreover, Canbilen *et al.* advocate the effectiveness of optimal MLD which incorporates the asymmetric characteristics over suboptimal MLDs in a dual-hop AF relay system suffering from IQI. They claim 37.50%, 95.48% and 55.56% reduction in average SEP at 30dB SNR when the system is subject to RX only, TX only and both TX-RX IQI. This reduction exhibits 2dB, 3.5dB and 3dB SNR gain at 1e-2 average SEP for RX only, TX only and both TX-RX IQI systems [379].

8.3.8 Multiuser Interference Channel

Asymmetric signaling is also beneficial in MU IC when interference is treated as noise. Thus, Nguyen *et al.* optimize the precoding matrices in MU SISO ICs to minimize maximum pairwise EP (PEP) and symbol error rate (SER) [305]. They present the advantageous asymmetric signaling under both AWGN and cellular networks with or without channel coding. Asymmetric signaling offers 100% and 99% decrease in SER with minmax-PEP/minmax-SER objectives relative to symmetric signaling with power control for 3U SISO IC in AWGN and cellular setup, respectively. Furthermore, it renders 5dB SNR gain at 8e-2 BER in a cellular setup with 3-edge users employing low-density-parity-check (LDPC) coding with QPSK modulation [305].

In a nutshell, the generalized approach to account for system asymmetry instead of an unrealistic symmetry assumption equip system designers with appropriate tools and additional design freedom to achieve lower EP, depending on the considered application, as emphasized in Table 8.1.

Table 8.1: Error Probability Reduction with Appropriate Modeling (Signaling, Filtering, Estimation or Detection)

System	Transmitter	Detector	Competitor	Percentage Reduction	SNR Gain	Ref
Generalized MIMO	16-QAM with V-Blast	WLD iterative RX	WL filter versus SL filter	30%	0.3dB at 0.02 FER	[371]
Multicarrier Transmission (UW-OFDM)	M-ASK modulation	WL-MMSE	WL-MMSE versus LMMSE	AWGN: 97.4% IEEE indoor channel: 94.5%	1dB at 2e-5 BER 3dB at 1e-6 BER	[233]
4- and 8-state TCM schemes	Asymmetric 8-PSK	Viterbi decoder	ATCM versus STCM	Rayleigh: 20% Rician: 76.67% S-Rician: 50%	0.3dB at 1e-5 BER 0.4dB at 1e-5 BER 0.2dB at 1e-5 BER	[351]
2- and 4-state TCM schemes	Asymmetric 4-PSK			2 states: 75% 4 states: 55%	0.5dB at 1e-5 BER 0.1dB at 1e-7 BER	[350]
2-16 states trellis coding	Asymmetric M-PSK, M-AM		Asymmetric versus Symmetric Signaling	4PSK:99.83% 4AM:98.33%	0.5dB at 1e-5 BER 1dB at 1e-5 BER	[139]
DS/SSMA trellis coding	Asymmetric M-PSK			3U: 58.02%	1dB at 1e-10 BER	[376]
MU IC	Asymmetric M-PSK or M-QAM	Symbol -by- Symbol		AWGN: 100% Cellular: 99%	- 5dB at 8e-2 max(BER)	[305]
Phase estimation without pilot	Asymmetric 8-PSK	Coherent detection		97.50%	3dB at 1e-2 SEP	[100]
Uncoded MIMO systems	M-ASK, QPSK/ OQPSK	ZF RX		ASK: 99.25% OQPSK: 99.58%	8dB at 2e-3 BER 12dB at 1.5e-3 BER	[377]
		MMSE RX		ASK: 97.78% OQPSK: 97.50%	6dB at 2e-4 BER 8.5dB at 3.5e-5 BER	[377]
Deep learning for commun. system	Learned coded modulation by auto-encoder	Neural networks	Autoencoder versus Conventional	Uncoded: 98.64% Coded:96.25%	0.7dB at 1e-3 BLER 1dB at 1e-3 BLER	[359]
DS/CDMA systems	QPSK modulation	MMSE MUD	L-MUD versus WL-MUD	MF: 100% L-MUD: 96%	- 2dB at 1e-5BER	[140]
Frequency selective channels	ASK, OQAM or BMSK	WL-MMSE	LE versus WLE	MMSE-LE:96.07% MMSE-DFE:56%	5dB at 1e-4BER 0.7dB at 1e-4BER	[65]
	STBC with M-PSK modulation			RA SISO-LE: 93.5% HT SISO-DFE:72.5% TU SIMO-DFE:74.29%	7dB at 2e-3 BER 2dB at 1e-3 BER 2.3dB at 7e-4 BER	[365]
SIMO with multiple CCIs	PSK data with PAM/ QAM CCI	WL ML RX	WL RX versus Usual RX	Desired CCI QPSK-BPSK: 83.53% BPSK-BPSK: 98.15% BPSK-PAM+QAM: 80% BPSK-2BPSK: 42%	7dB at 8e-4 SER 12dB at 1.3e-1 SER 4dB at 1.4e-3 SER 3dB at 1e-3 SER	[372]

HWI AF Relaying	QAM with Optimal power allocation	MED and MLD	Optimal versus suboptimal detectors	RX only IQI: 37.50% TX only IQI: 95.48% TX& RX IQI: 55.56%	2dB at 1e-2ASEP 3.5dB at 1e-2ASEP 3dB at 1e-2ASEP	[379]
HWI System	Asymmetric GPSK, QAM		Asymmetric versus Symmetric	10dB SNR: 84% 20dB SNR: 97.8%	10dB at 8e-2 SER	[35]
	Grey coded M-QAM modulation		Optimal MLD with IGN versus regular MLD with PGN	σ_{t}^2 = 0.001: 8.33% σ_{t}^2 = 0.010: 38.64% σ_{t}^2 = 0.100: 52.78%	10dB at 1.2e-1 BER 9dB at 4.4e-2 BER 4dB at 1.7e-3 BER	[378]
MIMO System under IGN	SSK modulation	ML RX		PGN: 0% n.id IGN: 20% correlated IGN:32% correlated and n.id IGN: 44%	0dB at 5e-3 EP 1dB at 5e-3 EP 2dB at 5e-3 EP 3dB at 5e-3 EP	[19]
MIMO with data like CCI and ISI in FSC	SC block transmission of OQPSK	WL-MMSE with FF, NP-FB	Linear versus WL FF and NP-FB filters	w/ CCI: 98.50% w/o CCI: 99.67%	20dB at 6e-2 BER 8dB at 1e-5 BER	[151]
				w/ CCI: 97.10% w/o CCI: 99.75%	20dB at 6.2e-2 BER 8dB at 1e-5 BER	[151]
MIMO FBMC systems	FBMC modulation based on offset QAM	MSE RX	SL versus WL processing	ITU:PAG: 91.25% ITU:VAG: 91.9%	2.5dB at 8e-6 BER 5dB at 5e-4 BER	[338]
		LP and WLP	MMSE versus WL-MMSE	Partial IC: 97.6% Full IC: 99.6%	9dB at 2.5e-3 BER 10.5dB at 2.5e-3 BER	[375]

Chapter 9

Applications

Previously, we discussed various scenarios and sources which arise impropriety in the system. This impropriety can be exploited to achieve performance gain in numerous diverse fields including but not limited to medicine, communication, geology and computer vision. The existing impropriety exploitation and/or introduced impropriety utilization have wide applications in data analysis, signal processing and communications [31,36,137,172]. Intuitively, asymmetric signaling and WL processing are the concepts that go hand in hand [58] and are beneficial in the various settings as demonstrated in Figure 9.1.

9.1 Data Analysis

Impropriety incorporation and appropriate processing provide tremendous advantages in data analysis and characterization. Key data analysis tools such as ICA and principal component analysis (PCA) demonstrate enhanced performance owing to NC characterization [135, 380, 381]. Other miscellaneous data analysis techniques with impropriety incorporation successfully categorize seismic, oceanographic and weather data sets [36, 382, 383].

9.1.1 Independent Component Analysis

ICA is a relatively new statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals. It particularly aims at the blind recovery of the source signal from the observations [135]. Many complex ICA algorithms either assume that the underlying sources are circular or rely on a magnitude-only

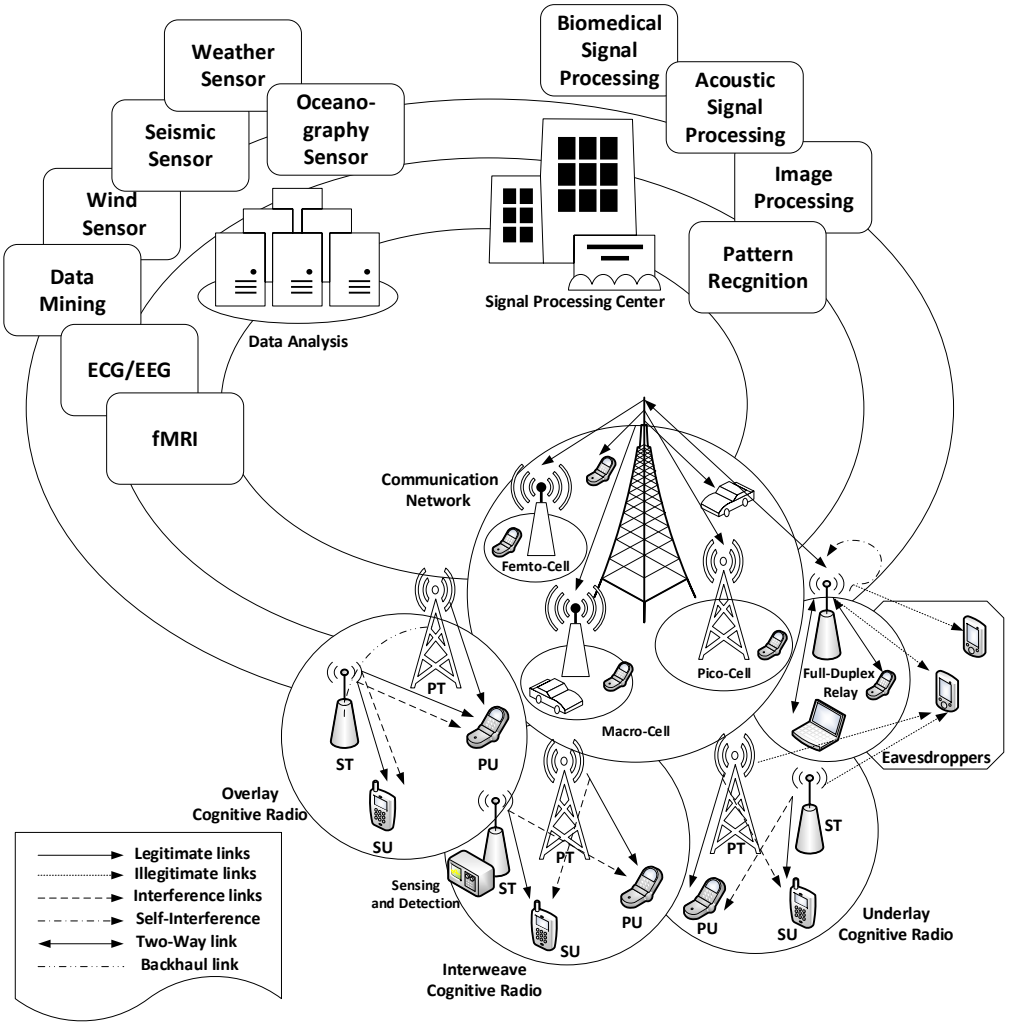


Figure 9.1: Applications of Impropriety in Data Analysis, Signal Processing and Communication Networks

model. However, this greatly limits the performance of ICA [32]. Thus, NC characterization enabled ICA has found real-world applications in various diverse fields such as medicine, economics, and data mining [40, 380, 384, 385].

Biomedical Data Analysis

Human brain functioning can be thoroughly examined using complex-valued fMRI data which measures the electrical and magnetic activity of the human brain [380]. However, fMRI poses tremendous challenges for data analysis techniques, including the design of ro-

bust yet flexible framework to capture the richness of human brain activities [384]. Adaptive noncircular ICA algorithms can effectively address these challenges rendering vast applications [40, 135]. For instance, feature extraction in electrocardiograms and fMRI data analysis lead to the improved neural activity estimation [31–33]. Moreover, real-time brain-computer interfacing relies on the extraction of eye muscle activity: electrooculogram (EOG) from electroencephalogram (EEG) recordings [41]. EEG records electrical potentials at various locations on the scalp and can render immense WL predictability by using blind source extraction algorithms like ICA. Furthermore, a robust ICA technique is also proposed to extract atrial activity in atrial fibrillation electrocardiograms (ECGs) [386].

Econometric and Data Mining

One major concern in econometric is the identification of underlying independent causes of a phenomenon, e.g., economic indicators, interest rates, and psychological factors of the exchange rates. These causes are quite insightful and can be identified by the decomposition of the financial time series analysis using appropriate ICA algorithms [380]. Another diverse application of such data analysis tool is in data mining, such as latent variable decomposition, multivariate time series analysis and prediction, extracting hidden signals in satellite images, text document data analysis, and weather data mining [135, 385].

9.1.2 Principal Component Analysis

PCA transforms correlated observed variables into a subset of uncorrelated variables, that account for total variance. PCA identifies patterns in data based on the correlation between features and thus it is less stringent than ICA [380]. PCA is a classical technique in statistical data analysis for pattern recognition [387–390], feature extraction [391, 392], data compression [393, 394], data reduction [395–397], data visualization [398], noise reduction [399, 400], factor analysis [401, 402], model selection [381], rank reduction [55], dimensionality reduction [403], etc. Classical PCA for real-valued systems relies on SOS

i.e., variance maximization. A staright forward extension to the proper complex-valued systems i.e., circular PCA (cPCA) relies on Hermitian covariance matrix [17]. However, a more general extension to the complex systems which can be circular/NC is based on:

- Both covariance and pseudo-covariance estimates to maximize likelihood function in complex representation rendering NC PCA (ncPCA) [404].
- The augmented covariance matrices yielding WL PCA (wlPCA) for direct extension from the cPCA [405].
- The real-composite covariance matrices resulting composite real PCA (crPCA) [403].

Aforementioned forms of general complex PCA exhibit their own merits and demerits based on the underlying applications and their linear/non-linear models.

Rank, Dimensionality and Data Reduction

Rank reduction finds balance between model bias and model variance to reconstruct signals from noisy observations. It can be achieved using PCA with eigen analysis of complex vectors. WL transformations take full advantage of SOS when compared to SL transformation. The wlPCA offers concentrated signal variance in first few principal components relative to ncPCA for rank reduction [55]. Interestingly, crPCA demonstrates its superiority in dimensionality reduction owing to the finer granularity and lower computational complexity than that of wlPCA [403]. PCA is widely used as a preliminary step for data reduction in various biomedical applications e.g., to extract and differentiate biomechanical features of gait waveform data related to knee osteoarthritis [395], in order selection for complex NC fMRI data [384], in data analysis for complementary DNA microarray experiments [396] and for genome data analysis in bioinformatics [397]. PCA offers numerous other applications including data visualization for biplot graphic display of matrices [398] and data compression of meteorological parameters obtained from high-resolution infrared spectra [394].

Feature Extraction and Pattern Recognition

Feature selection selects a subset of the original features, whereas feature extraction constructs a new feature subspace based on the feature set information. Feature extraction of image sequences by PCA can be later used for classification and recognition process in quality control applications [391] and fast iris recognition [392], respectively. Moreover, PCA is inevitably useful in pattern recognition like handwritten digits recognition [387], face recognition [389] and automated diagnosis of cardiac health using principal components of segmented ECG beats [390].

Model Selection, Digital Filtering and Data Denoising

Some complex mixing models i.e., DoA estimation, BSS and NC signal detection cannot employ wlPCA owing to their linear models. Thus, ncPCA is particularly advantageous for applications like model selection which aim to determine the subspace order and the number of noncircular signals. Furthermore, it is capable of detecting circular and noncircular signals and estimating signal subspace [381]. ncPCA is preferred for scenarios with high SNR, large number of samples and high degree of noncircularity. On the other hand, nonlinear component analysis requires WL PCA with complex kernel (wlPCA-ck) [406] for the design of digital filters and regression frameworks [407]. Furthermore, statistical PCA ia also employed for data denoising especially for image denosing on multi-exponential MRI relaxometry [399] and signal denoising in stock market trading [400].

Conclusively, non-circularity exploitation in complex PCA is particularly significant when the underlying entities are improper.

9.1.3 Complex Least Mean Square Analysis

LMS extenstion to the complex domain i.e., complex-LMS (CLMS) with non-circularity incorporation leads to variants like augmented CLMS (ACLMS) and complementary CLMS. These analysis techniques can be employed for adaptive estimation with numerous appli-

cations in system identification [366], real-time impropriety detection [46], communications [408], signal processing [409], renewable energy [410], power systems [29,411], and medicine [412].

CLMS algorithm for SL estimation with general second-order noncircular (improper) Gaussian input is found useful in identifying system coefficients which formed a strictly linear FIR channel [366]. Moreover, collaborative adaptive filters trained by the CLMS can detect and track improperness in real-time unlike competing static detectors [46]. Additionally, multiple access interference in DS-CDMA systems can be efficiently supressed using WL LMS algorithms [408]. Interestingly, ACLMS algorithm renders lower steady-state mean-squared error than conventional CLMS in adaptive beamforming for multi-port antenna arrays [409, 413]. Furthermore, ACLMS usefulness in wind modelling and forecasting is unprecedented in renewable energy domain [410, 414]. Similarly, the approximate uncorrelating transform improved adaptive frequency estimations using ACLMS in three-phase power grid systems [411]. Last but not the least, a hybrid filter with standard CLMS and ACLMS algorithms can discriminate between discrete states of brain consciousness i.e., coma and quasi-brain-death using nonlinear features in EEG [412].

9.1.4 Miscellaneous

Numerous other data analysis techniques which incorporate complete SOS of the complex improper observation data lead to the improved estimations of seismic traces [382], wind measurements [25, 36], and oceanographic velocity measurements [415].

WL complex auto-regressive processing of the seismic signals helps to capture essential data characteristics like elliptical oscillations [45]. Other climate and seismology applications with improper complex-valued stochastic models can be efficiently simulated using circulant embedding [416]. Compressive sensing of weather sensor network application can effectively exploit the asymmetrical features for energy estimation to cool down a given structure [383].

9.2 Signal Processing

Augmented and WL processing have demonstrated remarkable performance gains in different signal processing domains, e.g., signal estimation [31, 417], filtering [37], and detection [360]. Therefore, it is a leading competitor rendering vast applications in neuroscience, image processing, pattern recognition, and computer vision [336, 380].

9.2.1 Array Processing

The merits of impropriety adaptation in array processing algorithms are widely studied in [418–422]. For example, coherent processing (incorporating complementary covariance) for detection and estimation enjoys a 3dB gain over non-coherent processing [63]. Similarly, estimation accuracy is substantially enhanced by employing NC signal constellations in 1-D and 2-D DoA estimation [64]. Furthermore, such enhanced DoA estimation and identification methods for mixed circular and NC sources also improve the resolution capacity [423].

9.2.2 System Identification and Feature Extraction

Superior system identification can be achieved using WL adaptive estimation of general IGS using augmented complementary least mean square analysis [366, 417]. Similarly, impropriety incorporation is crucial in blind source identification or accurate estimation and then separation or equalization in NC mixing arrangements like acoustic sources and fault diagnosis [39, 51, 135, 424]. Additionally, noncircular ICA has demonstrated huge feature extracting potential in neuro-science, image processing, and vision research where we aim to find features that are as independent as possible [380].

9.2.3 Pattern Recognition and Image Processing

Major signal processing tasks such as compression, denoising, classification, feature extraction, image processing, and pattern recognition require sophisticated generalized models [380]. These problems can be efficiently solved using a statistical generative model based on NC ICA or require statistical measures like Bhattacharyya coefficient/distance and Kullback-Leibler divergence for appropriate modeling. These measures are well defined for the real signals or proper complex signals. However, their extension to more generalized scenarios of improper complex-valued Gaussian densities has enabled superior and reliable performance in the aforementioned applications [336]. Similarly, target detection in multi-band spectral images suffering from improper Gaussian noise is only possible with the advancement in impropriety literature [248].

9.3 Communication Systems

Communication systems can reap tremendous benefits from the asymmetric signaling in various interference limited scenarios. In this way, enhanced system performance in terms of improved achievable data rates and more reliable communication with lower EP can be achieved without exhausting the already saturated resources.

9.3.1 Cellular Networks

Improper/asymmetric signaling techniques are implemented in GSM [425, 426] and 3GPP networks [427, 428]. Similarly, in mobile MU communications, non-circularity characterization can render better tradeoff between power consumption and spectral efficiency [36]

9.3.2 Cognitive Radios

To address the sparse temporal and spatial utilization of spectrum bands, cognitive radio settings permit a secondary network to opportunistically utilize the spectrum resources of

a licensed primary network [61]. Secondary network senses the network availability and transmits under primary network QoS constraints. Cognitive radio settings are broadly categorized as underlay, overlay and interweave. Underlay cognitive system limits the transmission power of secondary network to maintain licensed users' QoS [191]. Overlay cognitive enjoys part of spectrum resources for its transmission while assisting primary network transmission. This scheme maintains primary network QoS through assistance with minimal interference and its own QoS by effectively canceling primary interference at its RX [192]. Lastly, interweave setting utilizes unused spectrum holes for its transmission as long as it is available [294]. Interestingly, efficient interference management by IGS permits secondary network to effectively utilize spectrum resources and enhance their system efficiency while maintaining primary network QoS with PGS. It is evident from the fact that improper interference from secondary network to primary network is way less deteriorating than its counterpart proper interference. Surprisingly, the gains reaped by secondary network with IGS transmission over PGS transmission are conditional. For instance, SU rate improvement with IGS is only feasible if the fraction of the squared modulus between the SU-PU interference link and the SU direct link surpass a threshold in underlay [190] and overlay cognitive radio paradigm [192]. Moreover, for underlay MAC, IGS is optimal if the accumulative IC gains exceed a certain threshold [188]. On the other hand, with adequate detection capabilities SU can employ IGS with maximum power to unconditionally improve its rate while satisfying PU QoS [294]. Apart from the traditional role of IGS to enhance data rate or reduce EP in communication systems, it can also be employed to enhance the secrecy performance. Consequently, unlicensed user achieves lower SOP with IGS employment in cognitive radio setup [189].

9.3.3 Full Duplex Systems

The performance of full duplex systems with simultaneous transmission and reception is invariably limited by the inherent self-interference (SI). This can be efficiently mitigated with

the optimal asymmetric signaling transmission for in-band full-duplex capable transceivers with [187] or without [154] spectrum sharing, SISO [179] and MIMO [182, 429] full duplex relaying. as well as heterogeneous multi-tier network involving cellular and D2D full duplex communications [186]. Improper transmission is even more beneficial for the joint compensation of SI and hardware imperfections in full duplex HWI systems. For instance, asymmetric transmission is capable of suppressing SI along with transmitter power amplifier nonlinear distortion and transceiver IQI [78] and asymmetric hardware distortions [83]. Besides full-duplex relaying, WL processing is also rewarding for other relay networks including two-way AF-MIMO relayed MU systems [181] and multi-layered relay systems [180]. Surprisingly, IGS is also favorable for the alternate relaying systems which mimic as full-duplex systems [159].

9.3.4 Hardware Impairments Mitigation

Improper/asymmetric signaling is a promising candidate for the compensation of various hardware imperfections including asymmetric hardware distortions in receive diversity systems [108] and IQI in space-time coded transmit diversity systems [430]. Severe performance losses caused by the IQI (which leads to improper received signals) can be efficiently compensated by WL RX for uplink multi-cell massive MIMO [82], WL precoding for Large scale MIMO [77], WL beamforming of linear antenna arrays [431], and massive antenna arrays [69], and circularity based compensators for wideband direct-conversion RXs [68] and OFDM based WLAN transmitters [70]. Improper transmission is also recognized to jointly compensate multiple HWIs such as additive hardware distortions and transceiver IQI in single antenna [35] and multi-antenna systems [20]. Interestingly, impairments in I/Q modulators are also accurately modelled using widely non-linear model using compressed sensing [80]. Moreover, the expectation maximization based ML channel estimation in multicarrier scenarios under phase distortion namely, phase noise and carrier frequency offset holds true for both proper and improper signaling [432].

9.3.5 Phase Estimation

In the absence of pilot training sequence, conventional symmetric signaling employ differential coding scheme for phase estimation. On contrary, asymmetric signaling can do the needful without differential coding hence saving around 3dB loss in SNR at 1e-2 SER. However, the performance gain up to 97.50% reduced SEP along with the absolute phase estimation comes at some cost. It is very small reduction in entropy and/or minimum distance owing to the unequally probable symbols and/or unequal symbol spacing, respectively, introducing asymmetry in the constellation [100].

9.3.6 Interference Channels

An interesting question is the suitability of improper signaling and WL filters when the system under study is proper (i.e., information and noise signals are proper). Among many others, Cadambe *et al.* in their pioneering work demonstrated the superiority of IGS in interference alignment scenarios [168, 433] counter intuitive to the PGS and linear filtering optimality in P2P communications [434]. Performance gains are also reported for a variety of interference-limited settings spanning SISO Z-IC [176, 178], MIMO Z-IC [177], MIMO P2P [169, 175], MU SISO X-IC [170–172, 185, 293, 305], MU MISO IC [173, 295], MU MIMO IC [81, 174], MIMO-BC [163, 164, 295, 306], MIMO-IBC [161, 162], cognitive MAC with primary P2P [188], cognitive P2P with primary MAC [166], MU diversity systems [296], multicarrier systems [363, 375, 435], multi-antenna systems [184], and single-/multi- antenna NC interference cancellation [59, 373]. Other forms of interference i.e., co-channel interference [151, 372], intra- and intersystem multiple access interference in radio navigation satellite services [436], and wideband multiple access and narrow band interference in CDMA systems [56, 140] can also be suppressed using improper characteristics. The application of the analysis to HCN multi-tier deployment with one macro eNode BS and multiple small eNode BS render large gains with guaranteed rate improvement for all SeNBs [177]. Additionally, IGS with symbol extension can outperform PGS for inter-

ference alignment within the context of linear precoding schemes where all interference is treated as noise [143, 168].

9.3.7 Noisy Channels

For MU Gaussian MIMO P2P, MAC, BC and IC with proper Gaussian per-user input signals, proper noise is the worst case for the rates under any constraint on the noise covariance matrices [437]. However, there are instances when noisy channels render improper/asymmetric signatures. The treatment of improper complex noise is carried out in various systems including CDMA [75, 438], discrete multitoned systems [74] and spectral image target detection [248]. Asymmetric noise characterization is necessary for appropriate estimation [329], detection [133, 156, 439], filtering [333], processing [137, 334], and compensation [19].

9.3.8 Trelli's Coding

A common perception is the optimality of the symmetric discrete signal constellations for both coded and uncoded communication systems. Although this holds true for the uncoded transmission, but may stand false for the coded systems [139]. Various contributions have supported this statement by designing asymmetric signal constellations to obtain a performance gain with trellis coding. Trellis-based detection is essential to close the gap between suboptimum DFE and optimum but computational complex MLSE. Trellis-based detection with WL preprocessing enables better suppression of noise and ISI for improper transmission in frequency selective channels [65]. Moreover, Divsalar *et al.* proposed joint design of $n/(n+1)$ trellis codes and 2^{n+1}-point asymmetric signaling, with same bandwidth requirement as an uncoded 2^n-point symmetric signal constellation. The joint treatment depicts significant improvement in minimum free Euclidean distance of the TCM which corresponds to maximum reduction in EbNo for a given BER [139]. Similar studies aim to design TCM parameters not only to increase the effective length but also the mini-

mum product distance of the code with Rayleigh or Rician fading channels. Interestingly, the later asymmetric TCM scheme offers gains without additional bandwidth or power requirements [351]. Other contributions optimize asymmetric constellations to target minimal BER of TCM systems over Gaussian channels unlike the conventional criterion of maximizing the free Euclidean distance [350]. Equivalently, asymmetric signal constellation with Trellis coding outperforms conventional symmetric signaling in terms of BER in DS/SSMA systems [376].

Chapter 10

Conclusion

This chapter summarizes the main contributions and signifies the lessons learned during the challenges faced in this work. It also conveys some concluding remarks and way forward for the future research work.

10.1 Summary

This report addresses the three-fold lacking and limitations of the existing literature i.e., 1) Inaccurate complex analysis 2) Inappropriate propriety characterization and 3) Ideal hardware assumption in wireless communication system modeling. It identifies numerous inherent and induced sources of impropriety in various disciplines of science to draw a motivation for the accurate modeling and analysis. The main focus of the report is on the accurate aggregate hardware impairment modeling, it's meticulous impropriety characterization, explicit analysis, appropriate signaling design, and efficient adaptive practical implementation.

We presented the comprehensive and exhaustive technical framework of the complex and quaternion stochastic data models followed by their complete impropriety characterization, testing, and appropriate treatment in terms of the transformations and operations. Then, we developed an accurate statistical model for communication links with non-ideal transceiver blocks, which captures the asymmetric property of the aggregate hardware impairments. Later on, we presented a case study of multihop decode-and-forward full-duplex relay systems suffering from HWDs and RSI. Compelled by improper nature of the un-

desired interference and inspired by the beneficial IGS, we adopted IGS to improve the end-to-end achievable rate using both joint and distributed optimization approaches. However, given the inevitable problems arising in the practical implementation of IGS such as unbounded peak-to-average power ratio and complicated detection process, we adopt asymmetric signaling as the finite discrete counterpart of IGS. This is achieved using the structural or/and stochastic shaping of the M-ary QAM constellation which is capable of improving the error performance and throughput of the underlying system even in the presence of improper HWIs. The performance of the proposed geometric, probabilistic, and hybrid shaping schemes to realize asymmetric signaling are analyzed, compared and validated using the Monte-Carlo simulations.

Afterwards, we surveyed the literature on impropriety characterization and elaborated the existing work from theoretical analysis and performance limits to practical realization and implementation. Theoretical analysis covers the achievable rate, outage probability, power efficiency, and DoF analysis followed by the broad IGS design, detection, and estimation guidelines. On the other hand, practical realization encompasses the detailed asymmetric signal design and recovery techniques. It also quantifies the improved error performance of multi-user direct sequence multiple access systems, coded communication systems, multiple antenna and multi-carrier systems, frequency selective and multi-user interference channels, and systems suffering from improper noise or HWIs.

This study is particularly significant owing to its vast applications in the diverse domains of data analysis, signal processing and communications. We have presented few applications in the fields of medicine, acoustics, geology, oceanography, economics, bioinformatics, forensics, image processing, computer vision, communication systems and power grids. This applications endorse the importance of propriety characterization for appropriate modeling, accurate analysis and targeted treatments.

10.2 Inferences

We can infer the following conclusions from the research carried out for this dissertation.

- Theoretically speaking, IGS is the preferred choice over PGS in an interference-limited scenario. Its effectiveness is proven in the multi-hop decode-and-forward full-duplex relaying (MH-DF-FDR) systems under residual self-interference (RSI) and hardware distortions (HWD). We have presented, analyzed and illustrated two realization schemes namely joint and distributed optimization schemes. Distributed-IGS is further categorized as per the cluster size as well as its relative position in the system network. These distinct forms of IGS-scheme can be adopted as per suitability in the underlying system configurations. For a small system configuration with fewer hops joint-IGS is the preferred choice. However, for a larger-hops system, the joint-IGS renders sub-optimal results pertaining to the inevitable processing and round-trip delays back and forth from the central-node at the cost of increased system complexity and communication overhead. Therefore, distributed-IGS is the preferred approach for large system configurations. Furthermore, distributed-IGS with bigger odd-cluster along with the optimal cluster placement is the preferred choice as per the acquired simulation results. In a nutshell, all forms of IGS are proven to be promising candidates for next generation networks that can significantly improve the overall achievable rate under various HWD and RSI levels, which have asymmetric signatures on the useful signal.

- Beyond HWDs, we demonstrated the significance of incorporating multiple HWIs in accurate system modeling and analysis. We mainly focused on the detailed modeling of I/Q imbalance and several distortion noises at the transceiver. We also proposed an optimal and a sub-optimal linear receiver which incorporates improper interference characteristics. Both transmitter and receiver I/Q imbalance render SI information bearing signal whereas only receiver I/Q imbalance is responsible for transforming

AWGN to improper Gaussian noise. The transmitter distortion is subject to channel fading while the receiver added impairment is not. We carried out the appropriate error probability analysis which can accommodate the correlated and power variant noise components. The derived Chernoff bounds and numerical integration approximations of the average error probability are in close agreement with the simulated average BER trends, particularly for the lower impairments levels. To this end, system reliability can be enhanced using the accurate model, analysis and optimal detection procedures. Further performance improvement can be achieved using the proposed asymmetric modulation schemes, which outperform the existing symmetric signaling with or without the optimal receiver. Moreover, the maximal asymmetric scheme can be a fairly good candidate to achieve better performance without rendering any optimization expenses for highly impaired systems.

- Probabilistic and hybrid shaping are proposed as yet other ways to realize asymmetric signaling in digital wireless communication systems suffering from improper HWD. Instinctively, all forms of asymmetric shaping are capable of decreasing the BER, and this performance gain improves with increasing SNR and/or increasing HWD levels with respect to NS. However, PS outperforms GS and performs equally well as HS. We can achieve more than 50% BER reduction with PS/HS over traditional GS. The perks of PS come at the cost of increased complexity in the design and decoding process. The HS scheme is capable of improving the system performance in terms of the BER as well as throughput. However, for less HWD levels and low EbNo, the benefits of HS over PS are limited while requiring additional complications in optimization, modulation, and detection procedures. Therefore, PS emerges as the best choice for asymmetric signaling in the trade-off between enhanced performance and added complexity.

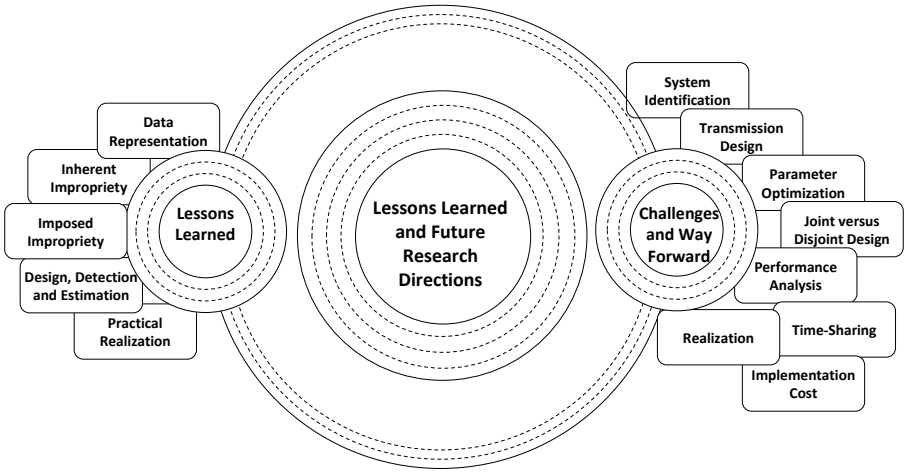


Figure 10.1: Lessons Learned and Way Forward

10.3 Lessons Learned

This section focuses on the challenges and limitations in getting the maximum benefit from impropriety and asymmetry concepts. These limitations open new research directions in numerous fields, especially communication theory as categorized in Fig. 10.1.

The importance of complete SOS cannot be advocated more, but this is just the beginning. It completes the analysis for Gaussian distribution; however, others may require higher-order statistics and rotational invariance characteristics for their complete description. It may bring added benefits at the expense of increased computational cost in terms of resources such as time and energy.

10.3.1 Data Representations

To summarize, three different yet interchangeable data representations with their own merits are presented. Studies advocate complex representation for easy and compact analysis, real composite representation for straightforward geometrical interpretations, and augmented representation for complete characterization, transformation, and analysis. Moreover, real-composite can identify improperness but fails to provide the DoI, unlike the rest of the two representations. DoI is used to identify the extent of impropriety and quantify

the entropy loss. Notably, circularity implies propriety whereas impropriety implies non-circularity but not vice versa. Propriety can vary from generalized to strict whereas circularity of a RV may range from marginal, weak, strong to total circularity. It is important to identify the extent of properness using propriety tests in order to apply the simplified form of processing as properness is only preserved under SL or affine transformation. Impropriety incorporation requires extended definitions of differential entropy and joint distribution for a RV e.g., Gaussian RV (2.38)-(2.48).

10.3.2 Inherent Impropriety

Exploiting the inherent impropriety in various configurations such as non-circular modulation, linear time-space coding, iterative receiver, improper interference and hardware impairments in communication and improper empirical data in fluid dynamics etc. can reap numerous benefits. For instance, it can pay significant rewards in the diverse fields like system fault identification in power systems, feature extraction and enhance estimation of EOG, EEG, ECG, and fMRI in biomedical engineering, quantum OCT in optics, speech recognition in acoustics, seismic and wind fields estimation in geophysics, and ocean-current spectra in oceanography etc. Impropriety incorporation is essential for appropriate modeling, analysis and optimum performance. Considering wireless communication systems, augmented representation is crucial for accurate SNR, achievable rate, and outage analysis.

10.3.3 Imposed Impropriety

The substantial performance improvement can be achieved by imposing impropriety in various interference-limited communication scenarios. To summarize the comparative study carried out in Table IV, underlay cognitive setup reaps the maximum benefit of IGS when compared to overlay and interweave cognitive setups. Similarly, IGS is proven to be more beneficial, with a substantial increase in achievable rate, in IC relative to BC in a MIMO

setup. SISO and SIMO follow the same trend for achievable rate whereas MISO reverses this trend by offering more rate-region improvement for BC and IBC relative to the IC. Furthermore, IGS has an added advantage of suppressing RSI in FD relaying and thus rendering higher achievable rates when compared to the HD relay mode. As per the IGS merits in outage analysis, multiple antenna systems depict lower outage than single-antenna systems. Moreover, percentage improvement relative to maximal PGS is far more than the optimal PGS in dual-hop DF-FDR. Also, the FD underlay cognitive mode renders lower rate outage than its HD counterpart whereas significant secrecy outage improvement is observed in the presence of an eavesdropper.

10.3.4 Design, Detection, and Estimation

System performance gains can only be attained by the appropriate design of IGS as per the underlying application. The enumerated design guidelines and tools signify the importance of problem identification. For instance, simple convex problems can either be solved in closed form or using algorithms like IPM. Alternately, for non-convex problems, semi-definite relaxation or sequential convex programming can help to convexify and find the approximate solution. Other times, it is difficult to solve a joint optimization problem. It is then recommended to employ alternate optimization like ADMM if the problem is convex for the subgroups of variables. If this condition fails, then separate optimization is a better alternate with suboptimal solution. For minimal dimensions, line search or even exhaustive search can give promising results whereas other algorithms are needed for NP-hard problems. Eventually, resorting to maximal IGS in place of conventional PGS can also be beneficial for other intractable optimization problems. As for the detection, the presence or DoI of an improper signal can be identified in the presence of proper, improper, or colored noise. Estimation is required to approximate the value of such improper signal at any instance. Furthermore, the separation, filtering and feature extraction of such sources can also be of great interest which can be efficiently carried out using BSS, ICA and PCA

etc.

10.3.5 Practical Realization

Out of the theoretical discussion, the practical realization is of utmost interest. The crux of the matter is that evaluation of the superiority of one form of asymmetric signaling over the other is critical yet tricky. Thus, hybrid signaling can reap tremendous payoffs in terms of improved achievable rates, energy efficiency, DoF and reduced outage and error probabilities at the expense of increased computation complexity in the modulation and detection phase. Error probability analysis depicts significant performance enhancement by asymmetric signaling in trellis coded systems especially with DS/SSMA scheme. Apart from the coded systems, uncoded MIMO with iterative receiver or with CCI and ISI in frequency selective channels reap the maximum benefit with asymmetric signaling. WL extension of various asymmetric signal recovery methods like equalization, estimation, filtering, and detection outperform their SL counterparts when dealing with asymmetric systems. Interestingly, asymmetric signaling can open new dimensions for user separation based on their asymmetric characteristics.

In a nutshell, the impropriety characterization in data analysis and signal processing renders numerous applications in medicine, economy, geology, oceanography, data mining, data denoising, data compression, dimensionality reduction, array processing, feature extraction, pattern recognition, and image processing etc. Focusing on the communication systems, the improper signaling has played a vital role in improving the system performance in cellular networks, cognitive radio setups, full-duplex communication, multi-antenna, multi-user and multi-cell setups. Additionally, it can effectively compensate the drastic effects of hardware impairments, interference and noisy channels. Although the journey from improper signaling to asymmetric signaling is quite appealing, it comes with few challenges and limitations as discussed next.

10.4 Challenges and Way Forward

Throughout this article, we have discussed the advantages of exploiting or incorporating non-circularity and impropriety. However, this performance comes with few challenges in terms of applicability, suitability, and practicality, etc which need further investigations. In this subsection, we discuss these limitations and suggest some way forward.

10.4.1 Impairment Modeling and Incorporation

A possible extension to this paper would be to consider the impact of aggregated HWIs on the massive MIMO systems, which is a key concept to attain higher area throughput in future wireless networks. Interestingly, the studies have shown that the huge degrees of freedom offered by the massive densification provide robustness to only some of the impairments. For example, [289] proved that the concentrated antennas deployment offers immunity to the hardware distortions but not the phase drifts through closed-form achievable rate performance analysis. Similarly, [290] demonstrated that the effects of impairments and noise at the massive-antenna fusion center vanish while the sensor impairment dominates the achievable distributed detection performance, in the limit of an infinite number of antennas and infinite sensors reporting power budget.

10.4.2 System Identification

Improper/asymmetric signaling techniques are already implemented in GSM and 3GPP networks [172]. Now we need to systematically scrutinize all the applications/systems/scenarios in numerous diverse fields where impropriety characterization can beat the traditional processing. It is noteworthy that not all the systems benefit from improper/asymmetric signaling and even if they do, the advantages may be conditional. In wireless communications, the thorough examination to determine the superiority conditions of improper signaling or WL processing over proper signaling or SL processing is as crucial as sys-

tem scrutinizing itself. As an example, MIMO-BC with proper Gaussian noise can achieve sum rate capacity under a sum power constraint with dirty paper coding and PGS in place of IGS [58, 440]. Similarly, IGS in combination with WL transceivers is beneficial in MIMO-BC under specific scenarios which are not yet understood [441]. Moreover, the perks of IGS over PGS for SU transmission are conditional in underlay and overlay cognitive setups [190, 192]. On the other hand, IGS is the all time favorite in other interference limited, hardware impaired or improper/non-circular noise based systems [81, 82, 108, 155] Therefore, a major challenge is to assess the usefulness of improper transmission in the underlying system which is not straightforward. Nonetheless, we would like to highlight a broader guideline as proper/symmetric signaling is the preferred choice in the noise limited regime whereas improper/asymmetric signaling is favorable in the interference limited regime [172]. Similarly, circular models are favored with small number of samples, low signal-to-noise ratio, or minimal degree of non-circularity [39].

10.4.3 Transmission Design

Considering the favorable scenarios, when impropriety can reap benefits, the majority of the studies advertise the employment of IGS transmission which is practically not feasible. Thus, we need to resort to discrete asymmetric transmission, which poses new challenges. The main challenge is choosing the optimal asymmetric signaling scheme based on the underlying system. For example, probabilistic shaping is widely applied in optical communications, whereas geometrical shaping is recently introduced for wireless communication systems. Performance superiority of one over the other in a particular application is yet to be investigated. Therefore, we advocate the employment of hybrid signaling in order to meet the upcoming demands of the communication systems for the internet-of-things era. Conclusively, hybrid geometrical and probabilistic asymmetric constellation designs can return significant performance merits while closely approaching Shannon limits.

10.4.4 Parameter Optimization

The next challenge is the optimization of the opted asymmetric signaling to fine tune transmission parameters, e.g., prior probabilities for probabilistic shaping, optimal rotation/translation, lattice and envelope for geometric shaping, non-uniform allocation of orthogonal/non-orthogonal resources, or some/all of these for hybrid signaling. The intricate search for an optimal solution is especially complicated by the significant number of feasible transmission strategies, mainly for a large number of participating users and/or antennas [176]. For instance, joint optimization of transmission parameters is doable for SISO-IC but we have to resort to suboptimal solutions when it comes to MISO- and MIMO-IC [81, 295]. Similarly, maximal IGS is adopted owing to the intractable optimization in MIMO hardware impaired systems [20]. In fact, either there is a lack of optimization tools for non-convex or NP-hard structural problems, or the existing algorithms render suboptimal solutions with excessive computational surcharge. Thus, low complexity algorithms with the near-optimal performance are required to fill the gap opened by the lack of optimal solutions in the complex systems [176]. Consequently, the search for least-complex near-to-optimal optimization strategy is an open research area.

10.4.5 Joint and Disjoint Design

Another conflict is the choice of cooperative or non-cooperative signaling in cases like multi-antenna, multi-user and multi-cell configurations. Non-cooperative signaling may render suboptimal performance whereas cooperative scheme requires the global knowledge of system parameters to yield optimal performance. This may lead to excessive communication overhead besides increased computational cost. For example, the performance comparison of multi-hop DF FDR communication under HWIs with distributed optimization framework reveals enhanced performance gains with increasing cluster size. The maximum gain is achieved with joint optimization of all nodes, however, this performance comes at the expense of increased complexity, communication overhead and processing delays. As

a general guideline, distributed optimization approach is the favorable choice for large systems whereas joint optimization is preferred for relatively small systems. Another concern is the unavailability or inaccurate estimation of few system parameters on IGS performance. For example, can IGS gains surpass PGS in large scale MU and/or multi-antenna systems in the absence of instantaneous CSIT? [304]. Future research may address the limitation arising from the imperfection or lack of instantaneous/average system parameters while tuning improper transmission parameters.

10.4.6 Performance Analysis

Accurate analysis is the key to design appropriate system parameters which can attain the expected system performance. Most of the studies employ complex representation relying on the covariance matrices for SNR analysis. Adopting such representation in the analysis ignores the correlation between the entities and their respective conjugates. Therefore, we suggest the employment of complex augmented covariance matrices to evaluate SNR and subsequent SNR outage performance. For instance, the accurate rate analysis and rate outage with augmented representations is advocated for multi-antenna ICs [81]. Similarly, error probability analysis should exploit the improper noise characteristics which is particularly emphasized in a hardware imapired system configuration [35]. Nevertheless, another limitation is the lack of numerical tools for exact performance analysis. The design of improper/asymmetric signaling parameters highly depends on the objective function which can be maximizing achievable rates like average achievable rate, achievable sum rate, minimum achievable rate or achievable rate region, minimizing outage probabilities like rate outage, SNR outage, or secrecy outage, or minimizing error probabilities. Thus, the accurate analysis of these performance metrics will dictate the optimality of improper/asymmetric transmission. For instance, the asymmetric system design is mostly based on the derived bounds instead of exact EP analysis, which yields loosely fitted model parameters [350]. Thus, tools are inevitably required to derive exact EP based on accurately

estimated parameters.

10.4.7 Time-Sharing

Most of the studies focused on the perks of improper/asymmetric signaling in the absence of time sharing (TS). Comparison study between PGS and IGS in the MIMO-BC when TIN at high SNR reveals different trends with or without TS. In the absence of TS, gains due to IGS occurs both in systems with enough antennas at the base station and in overloaded systems. Whereas, if TS is allowed then IGS cannot bring any gains in a system with enough antennas at the BS as opposed to overloaded system where it is still advantageous. Similarly, IGS with TS is yet to find it's standings in a MU multicell MIMO IBC [301]. Likewise, the superiority of IGS over PGS to enlarge rate region in SISO Z-IC with TS is subject to the underlying assumptions. IGS is only beneficial under short-term average power constraints, whereas it cannot bring any gains under the long-term average power constraints [442]. Although improper rate TS can outperform proper rate TS for SISO Z-IC, the investigation of this trend in a general SISO-IC is an open research problem. Nevertheless, situation is altogether different with three or more users where IGS is bound to bring the benefits even with TS owing to the added DoF [443]. Therefore, the superiority examination of IGS in a TS context under certain assumptions is an open research area in various interference-limited systems [301, 441].

10.4.8 Realization

The realization of the optimal asymmetric complex signaling along with the appropriate detection mechanism (to cater for the induced asymmetry) is also one the challenges the obstruct the journey from IGS to asymmetric signals. Undoubtedly, the asymmetric modulation does not work in isolation in modern communication setups. Interestingly, the survey of asymmetric discrete constellation with Trellis coding clarifies that the coding and modulation schemes are sometimes interrelated and hence cannot be treated independently [139].

Moreover, the optimal detection should exploit the non-i.i.d. noise components e.g., maximum a posterior detection in a HWI system outperforms regular ML or minimum Euclidean distance detection [35]. Similarly, asymmetric model of the aggregate HWIs in a wireless communication system will help in accurate system analysis and design [20]. Such appropriate modeling will also dictate the requisite resources to tackle and deal with the interference challenges. For example, asymmetric HWIs generate more errors requiring special buffer management approaches in wireless networks to maintain QoS [444, 445]. Efficient buffering and queue management schemes are inevitably required to tackle the latency issues in delay-sensitive applications [446]. Consequently, the appropriate realization of the system containing asymmetric signatures is immensely important to achieve the target performance.

10.4.9 Implementation Cost

The existing infrastructure employs SL transceivers, and thus the up-gradation to WL transceivers is a sequential process [174]. Last but not least is the evaluation of the tradeoff between performance gains and computational/implementation complexity. Performance gains in terms of increased capacity, reduced outage, and minimized EP are attained at the cost of added communication overhead to transfer system parameters, computational complexity to find a near-to-optimal solution, and implementation complexity to practically realize/detect the asymmetric transmission. Another intriguing concern is the power saving affair as emphasized in [139], i.e., introducing asymmetry does not affect the power or bandwidth needs of the systems. Nevertheless, a fair comparison is required to inspect whether the power saved by exploiting additional design freedom offered by asymmetric constellation is greater than the computational power spent on its fine tuning or not?

These are few challenges and limitations which need dedicated efforts for comprehensive treatment and effective realization in order to attain the maximum benefit from the rising asymmetry concept.

10.4.10 Recommendations

- We highlight the significance of incorporating hardware impairments to accurately model and efficiently design the future system configurations. It is crucial for the precise system analysis and exact system performance investigation. This thesis can form a basis for any study that aims to study the statistical hardware impairments system model and the developed analytical expressions can serve as performance metric to investigate practical system performance.

- Our work strongly emphasize the need to revise the existing literature and various performance metrics such as achievable rate, system outage and error probability analysis especially for the future generation of interference limited networks.

- Various numerous optimization problems have been studied for the joint optimization of various parameters to attain improved system performance. However, they improve one or other aspect of the system performance rendering a trade-off between different performance metrics. Therefore, we suggest Pareto analysis to achieve optimal system performance and efficiently utilize the given resources.

- In addition, the proposed adaptive scheme lays emphasis on the development of similar alternate sub-optimal solutions which can improve system performance while reducing computational overhead and system complexity.

- Eventually, the degrading effect of various hardware impairments becomes significant at higher data-rates and SNR values. Therefore, more effort needs to be placed in the design of highly sensitive and accurate RF transceiver blocks including HPA/LNA, ADC/DAC , BPF and LOs. Highly synchronized oscillators are inevitable for the mitigation of self interfering signals resulting from IQ imbalance.

10.5 Concluding Remarks

The journey from proper signaling to improper signaling and then from improper Gaussian to asymmetric discrete constellation is captured and summarized in this work. Various complex data representations, their complete SOS characterization and appropriate processing models are presented for comprehensive illustration. Furthermore, some intrinsic sources of impropriety as well as the vast applications of asymmetric signaling in various diverse fields i.e., medicine, communication, geology and computer vision are elaborated. This review article takes readers from the theoretical achievable bounds to practical realization of impropriety concepts. One of the notable contribution of this work is the performance comparison of improper signaling versus traditional proper signaling in terms of achievable rate, system outage and EP in numerous system configurations. The comparison captures all the necessary details including maximum achievable percentage improvement, transceiver types, design metric, employed strategies and optimization procedure etc. We believe that this survey along with the presented challenges and future research directions will not only compel readers to incorporate propriety concepts but also increase the activity in this critical realm.

REFERENCES

[1] L. Chettri and R. Bera, "A comprehensive survey on internet of things (IoT) toward 5G wireless systems," *IEEE Internet of Things J.*, vol. 7, no. 1, pp. 16–32, Jan. 2020.

[2] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, "A journey from improper Gaussian signaling to asymmetric signaling," vol. 22, no. 3, pp. 1539–1591, Apr. 2020.

[3] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," vol. 18, no. 3, pp. 11–21, May 2015.

[4] S. A. R. Naqvi, S. A. Hassan, H. Pervaiz, and Q. Ni, "Drone-aided communication as a key enabler for 5G and resilient public safety networks," vol. 56, no. 1, pp. 36–42, Jan. 2018.

[5] B. Razavi, "Design considerations for direct-conversion receivers," vol. 44, no. 6, pp. 428–435, June 1997.

[6] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," vol. 30, no. 12, pp. 1399–1410, Dec. 1995.

[7] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," vol. 34, no. 4, pp. 697–709, Apr. 2016.

[8] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.

[9] R. Krishnan, *On the Impact of Phase Noise in Communication Systems—Performance Analysis and Algorithms*. Chalmers University of Technology, Apr. 2015.

[10] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-rf impairments," in *Int. ITG Workshop on Smart Antennas (WSA),*. IEEE, Feb. 2010, pp. 189–196.

[11] M. Matthaiou, A. Papadogiannis, E. Bjornson, and M. Debbah, "Two-way relaying under the presence of relay transceiver hardware impairments," vol. 17, no. 6, pp. 1136–1139, June 2013.

[12] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," vol. 63, no. 24, pp. 6565–6580, Dec. 2015.

[13] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," vol. 60, no. 11, pp. 7112–7139, Sep. 2014.

[14] T. T. Duy, T. Q. Duong, D. B. da Costa, V. N. Q. Bao, and M. Elkashlan, "Proactive relay selection with joint impact of hardware impairment and co-channel interference," vol. 63, no. 5, pp. 1594–1606, May 2015.

[15] H. Mehrpouyan, M. Matthaiou, R. Wang, G. K. Karagiannidis, and Y. Hua, "Hybrid millimeter-wave systems: A novel paradigm for HetNets," vol. 53, no. 1, pp. 216–221, Jan. 2015.

[16] B. Picinbono, "Second-order complex random vectors and normal distributions," vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[17] ——, "On circularity," vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[18] Y. Xia, C. C. Took, and D. P. Mandic, "An augmented affine projection algorithm for the filtering of noncircular complex signals," *Signal Processing*, vol. 90, no. 6, pp. 1788–1799, Jun. 2010.

[19] M. M. Alsmadi, N. A. Ali, and S. S. Ikki, "SSK in the presence of improper Gaussian noise: Optimal receiver design and error analysis," in *IEEE Wireless Commun. Netw. Conf. (WCNC)*. Barcelona, Spain: IEEE, Apr. 2018, pp. 1–6.

[20] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Multiple antenna systems with hardware impairments: New performance limits," vol. 68, no. 2, pp. 1593–1606, Feb. 2019.

[21] B. I. Erkmen and J. H. Shapiro, "Optical coherence theory for phase-sensitive light," in *Quantum Communications and Quantum Imaging IV*, vol. 6305. International Society for Optics and Photonics, Aug. 2006, p. 63050G.

[22] B. Rivet, L. Girin, and C. Jutten, "Log-rayleigh distribution: A simple and efficient statistical representation of log-spectral coefficients," *IEEE Trans. on Audio, Speech, Language Process.*, vol. 15, no. 3, pp. 796–802, Mar. 2007.

[23] M. T. Taner, F. Koehler, and R. Sheriff, "Complex seismic trace analysis," *Geophysics*, vol. 44, no. 6, pp. 1041–1063, Jun. 1979.

[24] J. Calman, "On the interpretation of ocean current spectra. Part 1: The kinematics of three-dimensional vector time series," *Journal of Physical Oceanography*, vol. 8, no. 4, pp. 627–643, Jul. 1978.

[25] A. Khalili, A. Rastegarnia, and W. Bazzi, "A collaborative adaptive algorithm for the filtering of noncircular complex signals," in *7th Int. Symp. Telecommun. (IST)*. Tehran, Iran: IEEE, Sep. 2014, pp. 96–99.

[26] M. I. Knight and M. A. Nunes, "Long memory estimation for complex-valued time series," *Statistics and Computing*, pp. 1–20, Jul. 2018.

[27] W. V. Burt, T. Cummings, and C. A. Paulson, "The mesoscale wind field over the ocean," *Journal of Geophysical Research*, vol. 79, no. 36, pp. 5625–5632, 1974.

[28] A. M. Sykulski, S. C. Olhede, J. M. Lilly, and J. J. Early, "Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals," vol. 65, no. 12, pp. 3136–3151, Jun. 2017.

[29] Y. Xia and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems," vol. 61, no. 1, pp. 74–83, Jan. 2012.

[30] D. H. Dini and D. P. Mandic, "Widely linear modeling for frequency estimation in unbalanced three-phase power systems," vol. 62, no. 2, pp. 353–363, Feb. 2013.

[31] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: The theory of improper and noncircular signals*. Cambridge University Press, 2010.

[32] D. B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *Neuroimage*, vol. 25, no. 4, pp. 1310–1324, May 2005.

[33] P. A. Rodriguez, V. D. Calhoun, and T. Adalı, "De-noising, phase ambiguity correction and visualization techniques for complex-valued ICA of group fMRI data," *Pattern Recognition*, vol. 45, no. 6, pp. 2050–2063, Jun. 2012.

[34] X. Zhang, M. Matthaiou, E. Björnson, M. Coldrey, and M. Debbah, "On the MIMO capacity with residual transceiver hardware impairments," in *IEEE Int. Conf. Commun. (ICC)*. IEEE, Jun. 2014, pp. 5299–5305.

[35] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric modulation for hardware impaired systems-Error probability analysis and receiver design," vol. 18, no. 3, pp. 1723–1738, Mar. 2019.

[36] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[37] D. P. Mandic and V. S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*. John Wiley & Sons, 2009, vol. 59.

[38] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the Kurtosis contrast with algebraic optimal step size," vol. 21, no. 2, pp. 248–261, Feb. 2010.

[39] T. Adali and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," vol. 31, no. 5, pp. 112–128, Sep. 2014.

[40] H. Li, N. M. Correa, P. A. Rodriguez, V. D. Calhoun, and T. Adali, "Application of independent component analysis with adaptive density model to complex-valued fMRI data," vol. 58, no. 10, pp. 2794–2803, Oct. 2011.

[41] S. Javidi, D. P. Mandic, and A. Cichocki, "Complex blind source extraction from noisy mixtures using second-order statistics," vol. 57, no. 7, pp. 1404–1416, Jul. 2010.

[42] P. Clark, I. P. Kirsteins, and L. Atlas, "Existence and estimation of impropriety in real rhythmic signals," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. Kyoto, Japan: IEEE, Mar. 2012, pp. 3713–3716.

[43] J. Gonella, "A rotary-component method for analysing meteorological and oceanographic vector time series," in *Deep Sea Research and Oceanographic Abstracts*, vol. 19, no. 12. Elsevier, Dec. 1972, pp. 833–846.

[44] C. N. Mooers, "A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants," in *Deep Sea Research and Oceanographic Abstracts*, vol. 20, no. 12. Elsevier, Dec. 1973, pp. 1129–1141.

[45] A. M. Sykulski, S. C. Olhede, and J. M. Lilly, "A widely linear complex autoregressive process of order one," vol. 64, no. 23, pp. 6200–6210, Dec. 2016.

[46] B. Jelfs, D. P. Mandic, and S. C. Douglas, "An adaptive approach for the identification of improper complex signals," *Signal Processing*, vol. 92, no. 2, pp. 335–344, Feb. 2012.

[47] T. Adalı and H. Li, "Complex-valued adaptive signal processing," 2010.

[48] Y. Xia, B. Jelfs, M. M. Van Hulle, J. C. Príncipe, and D. P. Mandic, "An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals," vol. 22, no. 1, pp. 74–83, Jan. 2011.

[49] H. P. Yuen, "Two-photon coherent states of the radiation field," *Physical Review A*, vol. 13, no. 6, p. 2226, 1976.

[50] B. I. Erkmen and J. H. Shapiro, "Phase-conjugate optical coherence tomography," *Physical Review A*, vol. 74, no. 4, p. 041601, 2006.

[51] D. Looney and D. P. Mandic, "Augmented complex matrix factorisation," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Prague, Czech Republic: IEEE, May 2011, pp. 4072–4075.

[52] M. I. Layton and M. J. Gales, "Augmented statistical models for speech recognition," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Toulouse, France: IEEE, May 2006.

[53] Y. Xia, S. C. Douglas, and D. P. Mandic, "Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators," vol. 29, no. 5, pp. 44–54, Sep. 2012.

[54] R. Arablouei, S. Werner, and K. Doğançay, "Adaptive frequency estimation of three-phase power systems with noisy measurements," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. Vancouver, BC, Canada: IEEE, May 2013, pp. 2848–2852.

[55] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," vol. 51, no. 3, pp. 714–725, Mar. 2003.

[56] S. Buzzi, M. Lops, and A. M. Tulino, "A new family of MMSE multiuser receivers for interference suppression in DS/CDMA systems employing BPSK modulation," vol. 49, no. 1, pp. 154–167, Jan. 2001.

[57] E. Ollila, "On the circularity of a complex random variable," vol. 15, pp. 841–844, Nov. 2008.

[58] C. Hellings and W. Utschick, "Block-skew-circulant matrices in complex-valued signal processing," vol. 63, no. 8, pp. 2093–2107, Apr. 2015.

[59] P. Chevalier and F. Pipon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences-application to SAIC," vol. 54, no. 3, pp. 870–883, Mar. 2006.

[60] A. Mirbagheri, K. N. Plataniotis, and S. Pasupathy, "An enhanced widely linear CDMA receiver with OQPSK modulation," vol. 54, no. 2, pp. 261–272, Feb. 2006.

[61] J. H. Yeo and J. H. Cho, "Optimal presence detection of improper-complex second-order cyclostationary random signal for spectrum sensing in cognitive radio," in

GLOBECOM Workshops (GC Wkshps). Houston, TX, USA: IEEE, Dec. 2011, pp. 959–963.

[62] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," vol. 16, no. 8, pp. 1451 – 1458, Oct. 1998.

[63] P. J. Schreier, L. L. Scharf, and C. T. Mullis, "Detection and estimation of improper complex random signals," vol. 51, no. 1, pp. 306–312, Jan. 2005.

[64] F. Roemer and M. Haardt, "Efficient 1-D and 2-D DOA estimation for non-circular sources with hexagonal shaped ESPAR arrays," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*, vol. 4. Toulouse: IEEE, May 2006, pp. IV–IV.

[65] H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," vol. 51, no. 9, pp. 1512–1523, Sep. 2003.

[66] A. Mohammadi and K. N. Plataniotis, "Distributed widely linear multiple-model adaptive estimation," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 1, no. 3, pp. 164–179, Sep. 2015.

[67] Y. Xia and D. Mandic, "Augmented performance bounds on strictly linear and widely linear estimators with complex data," vol. 66, no. 2, pp. 507–514, Jan. 2018.

[68] L. Anttila, M. Valkama, and M. Renfors, "Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers," vol. 57, no. 4, pp. 2099–2113, Jul. 2008.

[69] A. Hakkarainen, J. Werner, K. R. Dandekar, and M. Valkama, "Widely-linear beamforming and RF impairment suppression in massive antenna arrays," *J. Commun. and Netw.*, vol. 15, no. 4, pp. 383–397, Aug. 2013.

[70] Z. Li, Y. Xia, W. Pei, K. Wang, Y. Huang, and D. P. Mandic, "Noncircular measurement and mitigation of I/Q imbalance for OFDM-Based WLAN transmitters," vol. 66, no. 3, pp. 383–393, Mar. 2017.

[71] M. Mokhtar, A. Gomaa, and N. Al-Dhahir, "OFDM AF relaying under I/Q imbalance: Performance analysis and baseband compensation," vol. 61, no. 4, pp. 1304–1313, Apr. 2013.

[72] J. Li, M. Matthaiou, and T. Svensson, "I/Q imbalance in AF dual-hop relaying: Performance analysis in nakagami-m fading," vol. 62, no. 3, pp. 836–847, Mar. 2014.

[73] P. J. Schreier and L. L. Scharf, "Stochastic time-frequency analysis using the analytic signal: Why the complementary distribution matters," vol. 51, no. 12, pp. 3071–3079, Dec. 2003.

[74] G. Taubock, "Complex noise analysis of DMT," vol. 55, no. 12, pp. 5739–5754, Dec. 2007.

[75] Y. C. Yoon and H. Leib, "Maximizing SNR in improper complex noise and applications to CDMA," vol. 1, no. 1, pp. 5–8, Jan. 1997.

[76] E. Bjornson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying:Performance limits with hardware impairments," vol. 61, no. 11, pp. 4512–4525, Nov. 2013.

[77] W. Zhang, R. C. de Lamare, C. Pan, M. Chen, J. Dai, B. Wu, and X. Bao, "Widely linear precoding for large-scale MIMO with IQI: Algorithms and performance analysis," vol. 16, no. 5, pp. 3298–3312, May 2017.

[78] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[79] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Impact of improper Gaussian signaling on hardware impaired systems," in *IEEE Int. Conf. on Commun. (ICC)*. Paris, France: IEEE, May 2017, pp. 1–6.

[80] M. Madero-Ayora, J. Reina-Tosina, C. Crespo-Cadenas, and J. Becerra-González, "On the robustness of a widely nonlinear approach to model impairments in I/Q modulators," in *Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*. Taormina, Italy: IEEE, Oct. 2015.

[81] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[82] S. Zarei, W. H. Gerstacker, J. Aulin, and R. Schober, "I/Q imbalance aware widely-linear receiver for uplink multi-cell massive MIMO systems: Design and sum rate analysis," vol. 15, no. 5, pp. 3393–3408, May 2016.

[83] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, "Improper Gaussian signaling for hardware impaired multihop full-duplex relaying systems," vol. 67, no. 3, pp. 1858–1871, Mar. 2019.

[84] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Improper signaling for SISO two-user interference channels with additive asymmetric hardware distortion," Sep. 2019.

[85] M. M. Alsmadi, A. E. Canbilena, S. S. Ikki, E. Basar, S. S. Gultekin, and I. Develi, "Imperfect CSI and improper Gaussian noise effects on SSK: Optimal detection and

error analysis," in *2018 IEEE Global Communi. Conf. (GLOBECOM)*. Abu Dhabi, United Arab Emirates: IEEE, Dec. 2018, pp. 1–6.

[86] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" vol. 32, no. 6, pp. 1065–1082, June 2014.

[87] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5g systems," vol. 52, no. 5, pp. 94–101, May 2014.

[88] L. Kong, L. Ye, F. Wu, M. Tao, G. Chen, and A. V. Vasilakos, "Autonomous relay for millimeter-wave wireless communications," vol. 35, no. 9, pp. 2127–2136, June 2017.

[89] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16 J relay-based wireless access networks: an Overview," vol. 15, no. 5, Oct. 2008.

[90] O. Amin, S. Bavarian, and L. Lampe, "Cooperative techniques for energy-efficient wireless communications," in *Green Radio Communication Networks*. Cambridge University Press, 2012, pp. 125–149.

[91] G. Farhadi and N. C. Beaulieu, "On the ergodic capacity of multi-hop wireless relaying systems," vol. 8, no. 5, Apr. 2009.

[92] H. Ju, S. Lim, D. Kim, H. V. Poor, and D. Hong, "Full duplexity in beamforming-based multi-hop relay networks," vol. 30, no. 8, pp. 1554–1565, Sep. 2012.

[93] A. Masmoudi and T. Le-Ngoc, "Residual self-interference after cancellation in full-duplex systems," in *Proc. 2014 IEEE Intern. Conf. Commun. (ICC 2014)*. Sydney, Australia: IEEE, June 2014, pp. 4680–4685.

[94] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[95] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," vol. 45, no. 2, pp. 411–420, Feb. 1997.

[96] I. Santamaria, P. M. Crespo, C. Lameiro, and P. J. Schreier, "Information-theoretic analysis of a family of improper discrete constellations," *Entropy*, vol. 20, no. 1, p. 45, Jan. 2018.

[97] J. A. López-Fernández, R. G. Ayestarán, I. Santamaria, and C. Lameiro, "Design of asymptotically optimal improper constellations with hexagonal packing," vol. 67, no. 8, pp. 5445–5457, Aug. 2019.

[98] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric modulation for hardware impaired systems - Error probability analysis and receiver design," vol. 18, no. 3, pp. 1723–1738, Feb. 2019.

[99] G. Foschini, R. Gitlin, and S. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," vol. 22, no. 1, pp. 28–38, Jan. 1974.

[100] T. Thaiupathump, C. D. Murphy, and S. A. Kassam, "Asymmetric signaling constellations for phase estimation," in *10th IEEE Workshop on Statistical Signal and Array Processing*. Pocono Manor, PA, USA: IEEE, Aug. 2000, pp. 161–165.

[101] H. Batshon, M. Mazurczyk, J.-X. Cai, O. Sinkin, M. Paskov, C. Davidson, D. Wang, M. Bolshtyansky, and D. Foursa, "Coded modulation based on 56APSK with hybrid shaping for high spectral efficiency transmission," in *2017 European Conf. on Optical Commun. (ECOC)*. Gothenburg, Sweden: IEEE, Sep. 2017.

[102] P. Schulte and G. Böcherer, "Constant composition distribution matching," vol. 62, no. 1, pp. 430–434, Jan. 2016.

[103] P. Schulte, "Zero error fixed length distribution matching," Ph.D. dissertation, M.S. thesis, Inst. Commun. Eng., Technische Universität München, Munich, Germany, May 2015.

[104] A. Cirino, "Design and analysis of a Trellis-based syndrome distribution matching algorithm," Master's thesis, University of Bologna, 2018.

[105] A. Elzanaty, A. Giorgetti, and M. Chiani, "Lossy compression of noisy sparse sources based on syndrome encoding," vol. 67, no. 10, pp. 7073–7087, Oct. 2019.

[106] M. Dia, V. Aref, and L. Schmalen, "A compressed sensing approach for distribution matching," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*. Vail, CO, USA: IEEE, June 2018, pp. 1266–1270.

[107] A. Elzanaty, A. Giorgetti, and M. Chiani, "Limits on sparse data acquisition: RIC analysis of finite Gaussian matrices," vol. 65, no. 3, pp. 1578–1588, Mar. 2019.

[108] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric hardware distortions in receive diversity systems: Outage performance analysis," *IEEE Access*, vol. 5, pp. 4492–4504, Apr. 2017.

[109] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," vol. 17, no. 1, pp. 91–94, Jan. 2013.

[110] S. Javed, O. Amin, S. S. Ikki, and M. S. Alouini, "On the achievable rate of hardware-impaired transceiver systems," in *in Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[111] A.-A. Boulogeorgos, V. Kapinas, R. Schober, and G. Karagiannidis, "I/Q-imbalance self-interference coordination," vol. 15, no. 6, pp. 4157–4170, June 2016.

[112] S. Li and R. D. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," vol. 13, no. 9, pp. 4794–4806, Sep. 2014.

[113] A.-A. A. Boulogeorgos, P. C. Sofotasios, B. Selim, S. Muhaidat, G. K. Karagiannidis, and M. Valkama, "Effects of RF impairments in communications over cascaded fading channels," vol. 65, no. 11, pp. 8878–8894, Jan. 2016.

[114] M. Mokhtar, A.-A. A. Boulogeorgos, G. K. Karagiannidis, and N. Al-Dhahir, "OFDM opportunistic relaying under joint transmit/receive I/Q imbalance," vol. 62, no. 5, pp. 1458–1468, May 2014.

[115] M. Mokhtar, N. Al-Dhahir, and R. Hamila, "OFDM full-duplex DF relaying under I/Q imbalance and loopback self-interference," vol. 65, no. 8, pp. 6737–6741, Aug. 2016.

[116] N. Kolomvakis, M. Matthaiou, and M. Coldrey, "IQ imbalance in multiuser systems: Channel estimation and compensation," vol. 64, no. 7, pp. 3039–3051, July 2016.

[117] A.-A. A. Boulogeorgos, N. D. Chatzidiamantis, and G. K. Karagiannidis, "Energy detection spectrum sensing under RF imperfections," vol. 64, no. 7, pp. 2754–2766, July 2016.

[118] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," vol. 17, no. 4, pp. 2017–2046, Feb. 2015.

[119] I. Harjula, R. Wichman, K. Pajukoski, E. Lahetkangas, E. Tiirola, and O. Tirkkonen, "Full duplex relaying for local area," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*. London, UK: IEEE, Sep. 2013, pp. 2684–2688.

[120] D. Ramirez and B. Aazhang, "Optimal routing and power allocation for wireless networks with imperfect full-duplex nodes," vol. 12, no. 9, pp. 4692–4704, Nov. 2013.

[121] B. Mahboobi and M. Ardebilipour, "Joint power allocation and routing in full-duplex relay network: An outage probability approach," vol. 17, no. 8, pp. 1497–1500, July 2013.

[122] T. K. Baranwal, D. S. Michalopoulos, and R. Schober, "Outage analysis of multihop full duplex relaying," vol. 17, no. 1, pp. 63–66, Nov. 2013.

[123] G. J. Gonzalez, F. H. Gregorio, J. Cousseau, T. Riihonen, and R. Wichman, "Performance analysis of full-duplex AF relaying with transceiver hardware impairments," in *Proc. 22th European Wireless Conf.* Oulu, Finland: VDE, May 2016, pp. 1–5.

[124] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward MIMO relaying," *arXiv preprint arXiv:1703.10209*, 2017.

[125] M. Windisch and G. Fettweis, "On the impact of I/Q imbalance in multi-carrier systems for different channel scenarios," in *IEEE International Symposium on Circuits and Systems (ISCAS 2007)*. New Orleans: IEEE, May 2007, pp. 33–36.

[126] J. Qi and S. Aissa, "Analysis and compensation of I/Q imbalance in MIMO transmit-receive diversity systems," vol. 58, no. 5, May 2010.

[127] R. Krishnan, M. R. Khanzadi, T. Eriksson, and T. Svensson, "Soft metrics and their performance analysis for optimal data detection in presence of strong oscillator phase noise," vol. 61, no. 6, pp. 2385–2395, June 2013.

[128] A. Bouhlel, S. G. Domouchtsidis, S. Ikki, and A. Sakly, "Performance of OFDM-IM under joint hardware impairments and channel estimation errors over correlated fading channels," *IEEE Access*, Nov. 2017.

[129] A. T. Walden and P. Rubin-Delanchy, "On testing for impropriety of complex-valued Gaussian vectors," vol. 57, no. 3, pp. 825–834, Mar. 2009.

[130] A. Mohammadi and K. N. Plataniotis, "Improper complex-valued multiple-model adaptive estimation." vol. 63, no. 6, pp. 1528–1542, Mar. 2015.

[131] E. Ollila and V. Koivunen, "Generalized complex elliptical distributions," in *Proc. Sensor Array and Multichannel Signal Process. Workshop*. Barcelona, Spain: IEEE, Jul. 2004, pp. 460–464.

[132] B. Dunbridge, "Asymmetric signal design for the coherent Gaussian channel," vol. 13, no. 3, pp. 422–431, Jul. 1967.

[133] S. Kassam, G. Moustakides, and J. Shin, "Robust detection of known signals in asymmetric noise," vol. 28, no. 1, pp. 84–91, Jan. 1982.

[134] A. van den Bos, "The multivariate complex normal distribution-A generalization," vol. 41, no. 2, pp. 537–539, Mar. 1995.

[135] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[136] J. Navarro-Moreno, M. D. Estudillo-Martínez, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Estimation of improper complex-valued random signals in colored noise by using the Hilbert space theory," vol. 55, no. 6, pp. 2859–2867, Jun. 2009.

[137] G. Taubock, "Complex-valued random vectors and channels: entropy, divergence, and capacity," vol. 58, no. 5, pp. 2729–2744, May 2012.

[138] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," vol. 56, no. 7, pp. 3502–3515, Jul. 2010.

[139] D. Divsalar, M. Simon, and J. Yuen, "Trellis coding with asymmetric modulations," vol. 35, no. 2, pp. 130–141, Feb. 1987.

[140] G. Gelli, L. Paura, and A. R. Ragozini, "Blind widely linear multiuser detection," vol. 4, no. 6, pp. 187–189, Jun. 2000.

[141] M. Isaka, M. P. Fossorier, R. H. Morelos-Zaragoza, S. Lin, and H. Imai, "Multilevel coded modulation for unequal error protection and multistage decoding.II. Asymmetric constellations," vol. 48, no. 5, pp. 774–786, May 2000.

[142] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," vol. 60, no. 11, pp. 5597–5625, Nov. 2012.

[143] S. A. Jafar *et al.*, "Interference alignment–A new look at signal dimensions in a communication network," *Foundations and Trends® in Commun. and Inf. Theory*, vol. 7, no. 1, pp. 1–134, 2011.

[144] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," vol. 17, no. 4, pp. 1941–1988, Sep. 2015.

[145] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," vol. 29, no. 4, pp. 61–80, Jul. 2012.

[146] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust statistics for signal processing*. Cambridge University Press, 2018.

[147] J. P. Delmas and H. Abeida, "Survey and some new results on performance analysis of complex-valued parameter estimators," *Signal Processing*, vol. 111, pp. 210–221, Jun. 2015.

[148] E. Moreau and T. Adali, *Blind identification and separation of complex-valued signals*. Wiley Online Library, 2013.

[149] S. Lagen, A. Agustin, and J. Vidal, "Decentralized widely linear precoding design for the MIMO interference channel," in *IEEE Globecom Workshops*. Austin, TX, USA: IEEE, Dec. 2014, pp. 771–776.

[150] I. A. Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-rate maximization of NOMA systems under imperfect successive interference cancellation," vol. 23, no. 3, pp. 474–477, Mar. 2019.

[151] M. Kim, J. H. Cho, and J. S. Lehnert, "Asymptotically optimal low-complexity SC-FDE with noise prediction in data-like improper-complex interference," vol. 15, no. 3, pp. 2090–2103, Mar. 2016.

[152] C. Lameiro, I. Santamaría, W. Utschiclk, and P. J. Schreier, "Maximally improper interference in underlay cognitive radio networks," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Shanghai, China: IEEE, Mar. 2016, pp. 3666–3670.

[153] S. Javed, O. Amin, and M.-S. Alouini, "Full-duplex relaying under I/Q imbalance using improper Gaussian signaling," in *Proc. 42th Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. New Orleans, LA, USA: IEEE, Mar. 2017.

[154] R. Sornalatha, N. Janakiraman, and T. Balasharmitha, "Modeling of widely linear digitally controlled self-interference canceller for full-duplex transceiver," in *Int. Conf. on Wireless Commun., Signal Process. and Netw. (WiSPNET)*. Chennai, India: IEEE, Mar. 2017, pp. 229–235.

[155] C. Hellings, "Reduced-entropy signals in MIMO communication systems," Ph.D. dissertation, Technische Universität München, 2017.

[156] J. K. Tugnait, "Multisensor detection of improper signals in improper noise," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. New Orleans, LA, USA: IEEE, Mar. 2017.

[157] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Impact of improper Gaussian signaling on hardware impaired systems," in *Proc. 2017 IEEE Intern. Conf. Commun. (ICC 2017)*, Paris, France, May 2017, pp. 1–6.

[158] C. Kim, E.-R. Jeong, Y. Sung, and Y. H. Lee, "Asymmetric complex signaling for full-duplex decode-and-forward relay channels," in *Int. Conf. on ICT Convergence (ICTC)*. Jeju Island, South Korea: Citeseer, Oct. 2012, pp. 28–29.

[159] M. Gaafar, O. Amin, A. Ikhlef, A. Chaaban, and M.-S. Alouini, "On alternate relaying with improper Gaussian signaling," vol. 20, no. 8, pp. 1683–1686, Aug. 2016.

[160] M. Gaafar, M. G. Khafagy, O. Amin, R. F. Schaefer, and M. S. Alouini, "Full-duplex relaying with improper Gaussian signaling over Nakagami-m fading channels," vol. PP, no. 99, pp. 1–1, Oct. 2017.

[161] H.-Y. Shin, S.-H. Park, H. Park, and I. Lee, "Achievable degrees of freedom for interference broadcast channels with asymmetric complex signaling," in *73rd Veh. Technol. Conf. (VTC Spring)*. Yokohama, Japan: IEEE, May 2011, pp. 1–5.

[162] J.-Y. Lin, R. Y. Chang, C.-H. Lee, H.-W. Tsao, and H.-J. Su, "Multi-agent distributed beamforming with improper Gaussian signaling for MIMO interference broadcast channels," vol. 18, no. 1, pp. 136–151, Jan. 2019.

[163] H. Bai, F. Gu, S. Wang, and J. Xiong, "Optimal design of widely linear pre-coding and estimation in multi-user MIMO broadcast channel," in *10th Int. Conf.on Wireless Commun. and Signal Process. (WCSP)*. Hangzhou, China: IEEE, Oct. 2018, pp. 1–7.

[164] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[165] C. Hellings and W. Utschick, "Iterative algorithms for transceiver design in MIMO broadcast channels with improper signaling," in *10th Int. ITG Conf. on Systems, Commun. and Coding (SCC)*. Hamburg, Germany: VDE, Feb. 2015, pp. 1–6.

[166] A. Kariminezhad, A. Chaaban, and A. Sezgin, "Improper signaling and symbol extensions: How far can we go with Gaussian P2P codebooks in the interfering MAC with TIN?" *arXiv preprint arXiv:1607.01995*, 2016.

[167] ——, "Interference MAC: Impact of improper Gaussian signaling on the rate region pareto boundary," in *25th European Signal Process. Conf. (EUSIPCO)*. Kos, Greece: IEEE, Sep. 2017, pp. 2536–2540.

[168] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signalingSettling the Høst-Madsen–Nosratinia conjecture," vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

[169] L. Yang and W. Zhang, "Interference alignment with asymmetric complex signaling on MIMO X channels," vol. 62, no. 10, pp. 3560–3570, Oct. 2014.

[170] C. Lameiro and I. Santamaría, "Degrees-of-freedom for the 4-user SISO interference channel with improper signaling." in *IEEE Int. Conf. on Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 3053–3057.

[171] J. Kim, J. Yeo, and J. H. Cho, "Potential of improper-complex signaling in communications over two-user interference channel," in *6th Int. Workshop on Signal Design and Its Appl. in Commun.* Tokyo, Japan: IEEE, Nov. 2013, pp. 4–7.

[172] Z. K. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," vol. 11, no. 9, pp. 3194–3203, Sep. 2012.

[173] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "MISO interference channel with improper Gaussian signaling," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Vancouver, BC, Canada: IEEE, May 2013, pp. 4374–4378.

[174] S. Lagen, A. Agustin, and J. Vidal, "Coexisting linear and widely linear transceivers in the MIMO interference channel," vol. 64, no. 3, pp. 652–664, Feb. 2016.

[175] ——, "Improper Gaussian signaling for the Z-interference channel," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. Florence, Italy: IEEE, May 2014, pp. 1140–1144.

[176] E. Kurniawan and S. Sun, "Improper Gaussian signaling scheme for the Z-interference channel," vol. 14, no. 7, pp. 3912–3923, Jul. 2015.

[177] S. Lagén Morancho, A. Agustín de Dios, and J. Vidal Manzano, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," vol. 64, no. 8, pp. 3350–3368, Aug. 2016.

[178] C. Lameiro, I. Santamaría, and P. J. Schreier, "Rate region boundary of the SISO Z-interference channel with improper signaling," vol. 65, no. 3, pp. 1022–1034, Mar. 2017.

[179] M. Gaafar, M. G. Khafagy, O. Amin, R. F. Schaefer, and M.-S. Alouini, "Full-duplex relaying with improper Gaussian signaling over Nakagami-m fading channels," vol. 66, no. 1, pp. 64–78, Jan. 2018.

[180] Z. Ho, A. Zappone, and E. Jorswieck, "Optimal non-regenerative relay processing with improper signals," in *10th Int. Symp. Wireless Commun. Sys. (ISWCS),*. Ilmenau, Germany: VDE, Aug. 2013, pp. 1–5.

[181] J. Zhang and M. Haardt, "Widely linear signal processing for two-way relaying with MIMO amplify and forward relays," in *10th Int. Symp. Wireless Commun. Sys.(ISWCS)*. Ilmenau, Germany: VDE, Aug. 2013, pp. 1–5.

[182] A. Kariminezhad, A. Sezgin, and M. Pesavento, "Power efficiency of improper signaling in MIMO full-duplex relaying for K-user interference networks," in *IEEE Int. Conf. on Commun. (ICC)*. Paris, France: IEEE, May 2017, pp. 1–6.

[183] J. K. Tugnait, "Multiantenna spectrum sensing for improper signals over frequency selective channels," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Shanghai, China: IEEE, Mar. 2016, pp. 3781–3785.

[184] C. Lameiro, I. Santamaría, and P. J. Schreier, "Performance analysis of maximally improper signaling for multiple-antenna systems," in *IEEE Wireless Commun. and Netw. Conf. (WCNC)*. Barcelona, Spain: IEEE, Apr. 2018, pp. 1–6.

[185] H. Park, S.-H. Park, J.-S. Kim, and I. Lee, "SINR balancing techniques in coordinated multi-cell downlink systems," vol. 12, no. 2, pp. 626–635, Feb. 2013.

[186] A. Kariminezhad and A. Sezgin, "Heterogeneous multi-tier networks: Improper signaling for joint rate-energy optimization," vol. 18, no. 1, pp. 680–694, Jan. 2019.

[187] M. Gaafar, O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay spectrum sharing techniques with in-band full-duplex systems using improper Gaussian signaling," vol. 16, no. 1, pp. 235–249, Jan. 2017.

[188] C. Lameiro, I. Santamaría, and P. J. Schreier, "Improper Gaussian signaling for multiple-access channels in underlay cognitive radio," vol. 67, no. 3, pp. 1817–1830, Mar. 2019.

[189] G. Oliveira, E. Fernandez, S. Mafra, and S. Montejo-Sánchez, "Physical layer security in cognitive radio networks using improper Gaussian signaling," vol. 22, no. 9, pp. 1886–1889, Sep. 2018.

[190] C. Lameiro, I. Santamaría, and P. J. Schreier, "Benefits of improper signaling for underlay cognitive radio," vol. 4, no. 1, pp. 22–25, Feb. 2015.

[191] O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay cognitive radio systems with improper Gaussian signaling: Outage performance analysis," vol. 15, no. 7, pp. 4875–4887, Jul. 2016.

[192] ——, "Overlay spectrum sharing using improper Gaussian signaling," vol. 35, no. 1, pp. 50–62, Jan. 2017.

[193] W. Hedhly, O. Amin, and M.-S. Alouini, "Benefits of improper Gaussian signaling in interweave cognitive radio with full and partial CSI," *IEEE Trans. Cognitive Commun. and Netw.*, 2020.

[194] N. S. Loghin, J. Zöllner, B. Mouhouche, D. Ansorregui, J. Kim, and S.-I. Park, "Non-uniform constellations for ATSC 3.0," vol. 62, no. 1, pp. 197–203, Feb. 2016.

[195] J. Zöllner and N. Loghin, "Optimization of high-order non-uniform QAM constellations," in *Proc. IEEE Int. Symp. on Broadband Multimedia Syst. and Broadcast. (BMSB)*. London, UK: IEEE, June 2013, pp. 1–6.

[196] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," vol. 39, no. 3, pp. 913–929, May 1993.

[197] G. Forney, R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," vol. 2, no. 5, pp. 632–647, Sep. 1984.

[198] M. P. Yankov, D. Zibar, K. J. Larsen, L. P. Christensen, and S. Forchhammer, "Constellation shaping for fiber-optic channels with QAM and high spectral efficiency," *IEEE Photon. Technol. Lett.*, vol. 26, no. 23, pp. 2407–2410, Dec. 2014.

[199] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 2016.

[200] J. Cho, S. Chandrasekhar, R. Dar, and P. J. Winzer, "Low-complexity shaping for enhanced nonlinearity tolerance," in *ECOC 2016; 42nd European Conf on Optical Commun.* Dusseldorf, Germany: VDE, Sep. 2016.

[201] J.-X. Cai, H. G. Batshon, M. V. Mazurczyk, O. V. Sinkin, D. Wang, M. Paskov, W. Patterson, C. R. Davidson, P. Corbett, G. Wolter, T. Hammon, M. Bolshtyansky, D. Foursa, and A. Pilipetskii, "70.4 Tb/s capacity over 7,600 km in C+L band using coded modulation with hybrid constellation shaping and nonlinearity compensation," in *Proc. Opt. Fiber Comm. Conf. and Exhibit. (OFC)*. Los Angeles, CA, USA: Optical Society of America, Mar. 2017, pp. Th5B–2.

[202] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[203] C. Pan and F. R. Kschischang, "Probabilistic 16-QAM shaping in WDM systems," *J. Lightw. Technol.*, vol. 34, no. 18, pp. 4285–4292, Sep. 2016.

[204] A. Elzanaty and M. Alouini, "Adaptive coded modulation for IM/DD free-space optical backhauling: A probabilistic shaping approach," pp. 1–1, July 2020.

[205] Y. C. Gültekin, W. J. van Houtum, S. Şerbetli, and F. M. Willems, "Constellation shaping for IEEE 802.11," in *Proc. IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. Montreal, QC, Canada: IEEE, Oct. 2017, pp. 1–7.

[206] H. Iimori, R.-A. Stoica, and G. T. F. de Abreu, "Constellation shaping for rate maximization in AWGN channels with non-linear distortion," in *Proc. IEEE 7th*

Int. Workshop on Comput. Advances in Multi-Sensor Adaptive Process. (CAMSAP). Curacao, Netherlands Antilles: IEEE, Dec. 2017, pp. 1–5.

[207] S. Javed, O. Amin, S. S. Ikki, and M. S. Alouini, "On the optimal detection and error performance analysis of the hardware impaired systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–7.

[208] W. H. Hayt, J. E. Kemmerly, and S. M. Durbin, *Engineering circuit analysis*. McGraw-Hill New York, 1986.

[209] J. G. Proakis and M. Salehi, *Digital communications*. McGraw-Hill, 2008.

[210] P. Rubin-Delanchy and A. T. Walden, "Kinematics of complex-valued time series," vol. 56, no. 9, pp. 4189–4198, Sep. 2008.

[211] D. P. Mandic, S. Javidi, G. Souretis, and V. S. Goh, "Why a complex valued solution for a real domain problem," in *IEEE workshop on machine learning for signal processing*. Thessaloniki, Greece: IEEE, Aug. 2007, pp. 384–389.

[212] S. C. Olhede and A. T. Walden, "Noise reduction in directional signals using multiple morse wavelets illustrated on quadrature Doppler ultrasound," vol. 50, no. 1, pp. 51–57, Jan. 2003.

[213] L. Fortuna, G. Muscato, and M. G. Xibilia, "A comparison between HMLP and HRBF for attitude control," vol. 12, no. 2, pp. 318–328, Mar. 2001.

[214] A. J. Hanson, *Visualizing Quaternions*. San Francisco: The Morgan Kaufmann Series in Interactive 3D Technology, 2006.

[215] S. Miron, N. Le Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," vol. 54, no. 4, pp. 1218–1229, Apr. 2006.

[216] T. Bulow and G. Sommer, "Hypercomplex signals-a novel extension of the analytic signal to the multidimensional case," vol. 49, no. 11, pp. 2844–2852, Nov. 2001.

[217] J. Seberry, K. Finlayson, S. S. Adams, T. A. Wysocki, T. Xia, and B. J. Wysocki, "The theory of quaternion orthogonal designs," vol. 56, no. 1, pp. 256–265, Jan. 2008.

[218] S. Buchholz and N. Le Bihan, "Polarized signal classification by complex and quaternionic multi-layer perceptrons," *Int. J. Neural Sys.*, vol. 18, no. 02, pp. 75–85, May 2008.

[219] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," vol. 57, no. 4, pp. 1316–1327, Apr. 2009.

[220] J. Navarro-Moreno, R. Marı, J. C. Ruiz-Molina *et al.*, "A quaternion widely linear model for nonlinear Gaussian estimation," vol. 62, no. 24, pp. 6414–6424, Dec. 2014.

[221] M. Xiang, S. Kanna, and D. P. Mandic, "Performance analysis of Quaternion-valued adaptive filters in nonstationary environments," vol. 66, no. 6, pp. 1566–1579, Mar. 2018.

[222] R. Krupiński, "Generating augmented Quaternion random variable with generalized Gaussian distribution," *IEEE Access*, vol. 6, pp. 34 608–34 615, Jul. 2018.

[223] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing," *Signal Processing*, vol. 84, no. 7, pp. 1177–1199, Jul. 2004.

[224] K. Daniilidis, "Hand-eye calibration using dual quaternions," *Int. J. of Robotics Research*, vol. 18, no. 3, pp. 286–298, Mar. 1999.

[225] P.-O. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," in *Int.Conf. Mathematics (IMA) in Signal Process.*, Cirencester, 2004, pp. 23–26.

[226] E. Ollila, J. Eriksson, and V. Koivunen, "Complex elliptically symmetric random variablesgeneration, characterization, and circularity tests," vol. 59, no. 1, pp. 58–69, Jan. 2011.

[227] W. R. Hamilton, "Theory of quaternions," *Proc. Royal Irish Academy (1836-1869)*, vol. 3, pp. 1–16, Nov. 1844.

[228] F. Zhang, "Quaternions and matrices of quaternions," *Linear algebra and its appl.*, vol. 251, pp. 21–57, Jan. 1997.

[229] H. Coxeter, "Quaternions and reflections," *The American mathematical monthly*, vol. 53, no. 3, pp. 136–146, 1946.

[230] P. Ginzberg and A. T. Walden, "Testing for quaternion propriety," vol. 59, no. 7, pp. 3025–3034, Jul. 2011.

[231] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214–224, Feb. 2011.

[232] J. Via, D. P. Palomar, and L. Vielva, "Generalized likelihood ratios for testing the properness of quaternion Gaussian vectors," vol. 59, no. 4, pp. 1356–1370, Apr. 2011.

[233] S. Trampitsch, "Complex-valued data estimation," Ph.D. dissertation, Master Thesis. Alpen-Adria University of Klagenfurt, Apr. 2013.

[234] P. J. Schreier, "Bounds on the degree of impropriety of complex random vectors," vol. 15, pp. 190–193, Jan. 2008.

[235] K. Fang, S. Kotz, and K. Ng, *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall, 1990.

[236] H. Abeida and J.-P. Delmas, "MUSIC-like estimation of direction of arrival for non-circular sources," vol. 54, no. 7, pp. 2678–2690, Jul. 2006.

[237] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," vol. 13, no. 7, pp. 433–436, Jul. 2006.

[238] E. Ollila, V. Koivunen, and H. V. Poor, "Complex-valued signal processingEssential models, tools and statistics," in *Inf. Theory and Appl. (ITA) Workshop*. La Jolla, CA, USA: IEEE, Feb. 2011, pp. 1–10.

[239] N. N. Vakhania, "Random vectors with values in quaternion hilbert spaces," *Theory of Probability & Its Applications*, vol. 43, no. 1, pp. 99–115, 1999.

[240] T. M. Cover and J. A. Thomas, *Elements of information theory*. Hoboken, NJ, USA : John Wiley & Sons, 2006.

[241] L. Zetterberg and H. Brandstrom, "Codes for combined phase and amplitude modulated signals in a four-dimensional space," vol. 25, no. 9, pp. 943–950, Sep. 1977.

[242] W. McGill, "Multivariate information transmission," vol. 4, no. 4, pp. 93–111, Sep. 1954.

[243] R. Fisher, *Precoding and Signal Shaping for Digital Communications*. John Wiley & Sons, New York, 2002.

[244] J.-P. Delmas, A. Oukaci, and P. Chevalier, "On the asymptotic distribution of GLR for impropriety of complex signals," *Signal Processing*, vol. 91, no. 10, pp. 2259–2267, Oct. 2011.

[245] D. Ramírez, P. J. Schreier, J. Vía, and I. Santamaría, "Testing blind separability of complex Gaussian mixtures," *Signal Processing*, vol. 95, pp. 49–57, Feb. 2014.

[246] S. Chandna and A. T. Walden, "A frequency domain test for propriety of complex-valued vector time series." vol. 65, no. 6, pp. 1425–1436, Mar. 2017.

[247] J. K. Tugnait and S. A. Bhaskar, "On testing for impropriety of multivariate complex-valued random sequences," vol. 65, no. 11, pp. 2988–3003, Jun. 2017.

[248] G. M. Matalkah, J. M. Sabatier, and M. M. Matalgah, "A generalized likelihood ratio test for detecting targets in multiple-band spectral images with improper complex Gaussian noise," in *15th IEEE International Conference on Image Processing (ICIP)*. San Diego, CA, USA: IEEE, Oct. 2008, pp. 1856–1859.

[249] E. Ollila, V. Koivunen, and H. V. Poor, "A robust estimator and detector of circularity of complex signals," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Prague, Czech Republic: IEEE, May 2011, pp. 3620–3623.

[250] J. P. Delmas and H. Abeida, "Asymptotic distribution of circularity coefficients estimate of complex random variables," *Signal Processing*, vol. 89, no. 12, pp. 2670–2675, Dec. 2009.

[251] L. Pralon, M. Pralon, B. Pompeo, and G. Vasile, "On a probabilistic approach to detect Noise Radar random transmit waveforms based on a simple circularity test," in *Radar Conf.* Philadelphia, PA, USA: IEEE, May 2016, pp. 1–4.

[252] C. Hellings, M. Koller, and W. Utschick, "An impropriety test based on block-skew-circulant matrices," in *Proc. 19th Int. ITG Workshop on Smart Antennas (WSA)*. Ilmenau, Germany,: VDE, Mar. 2015, pp. 1–6.

[253] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis*. New York: Academic, 1979.

[254] S. M. Kay and J. R. Gabriel, "An invariance property of the generalized likelihood ratio test," vol. 10, no. 12, pp. 352–355, Dec. 2003.

[255] M. Novey, E. Ollila, and T. Adali, "On testing the extent of noncircularity," vol. 59, no. 11, pp. 5632–5637, Nov. 2011.

[256] A.-A. A. Boulogeorgos, V. M. Kapinas, R. Schober, and G. K. Karagiannidis, "I/Q-imbalance self-interference coordination," vol. 15, no. 6, pp. 4157–4170, June 2016.

[257] M. Wenk, *MIMO-OFDM-testbed: challenges, implementations, and measurement results*. ETH Zurich, 2010.

[258] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Hardware impairments in large-scale MISO systems: Energy efficiency, estimation, and capacity limits," in *Proc. 18th Intern. Conf. Digital Signal Process.*, Santorini, Oct. 2013, pp. 1–6.

[259] P. Zetterberg, "Experimental investigation of TDD reciprocity-based zero-forcing transmit precoding," *EURASIP J. Adv.S ignal Process.*, vol. 2011, p. 5, Dec. 2011.

[260] K. Papadopoulos, *Channel Estimation for Large MIMO Systems Under Hardware Impairments*. McGill University Libraries, 2016.

[261] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," vol. 26, no. 6, Aug. 2008.

[262] A. Behzad, K. A. Carter, H.-M. Chien, S. Wu, M.-A. Pan, C. P. Lee, Q. Li, J. C. Leete, S. Au, M. S. Kappes *et al.*, "A fully integrated MIMO multiband direct conversion CMOS transceiver for wlan applications (802.11 n)," vol. 42, no. 12, pp. 2795–2808, Nov. 2007.

[263] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," vol. 49, no. 10, pp. 2335–2344, Oct. 2001.

[264] L. Anttila, M. Valkama, and M. Renfors, "Frequency-selective I/Q mismatch calibration of wideband direct-conversion transmitters," vol. 55, no. 4, pp. 359–363, Apr. 2008.

[265] A. Kiayani, L. Anttila, O. Mylläri, and M. Valkama, "Prototype implementation and rf performance measurements of DSP based transmitter I/Q imbalance calibration," in *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*. IEEE, Jul. 2010, pp. 484–489.

[266] B. Narasimhan, D. Wang, S. Narayanan, H. Minn, and N. Al-Dhahir, "Digital compensation of frequency-dependent joint Tx/Rx I/Q imbalance in OFDM systems under high mobility," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 3, pp. 405–417, Jun. 2009.

[267] A. Tarighat and A. H. Sayed, "MIMO OFDM receivers for systems with IQ imbalances," vol. 53, no. 9, pp. 3583–3596, Sep. 2005.

[268] M. Gaafar, M. G. Khafagy, O. Amin, and M.-S. Alouini, "Improper gaussian signaling in full-duplex relay channels with residual self-interference," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpure, May 2016.

[269] M. Duarte, *Full-duplex wireless: Design, implementation and characterization*. Rice University, 2012.

[270] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*. IEEE, 2010, pp. 1558–1562.

[271] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," vol. 11, no. 12, pp. 4296–4307, Nov. 2012.

[272] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," vol. 32, no. 9, pp. 1637–1652, Sept. 2014.

[273] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[274] X. Li, C. Tepedelenlioğlu, and H. Şenol, "Channel estimation for residual self-interference in full-duplex amplify-and-forward two-way relays," vol. 16, no. 8, pp. 4970–4983, Aug. 2017.

[275] T. M. Kim and A. Paulraj, "Outage probability of amplify-and-forward cooperation with full duplex relay," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. Paris, France: IEEE, Apr. 2012, pp. 75–79.

[276] X. Li, C. Tepedelenlioğlu, and H. Şenol, "Optimal training for residual self-interference for full duplex one-way relays," *arXiv preprint arXiv:1709.06140*, 2017.

[277] N. H. Tran, L. J. Rodríguez, and T. Le-Ngoc, "Optimal power control and error performance for full-duplex dual-hop AF relaying under residual self-interference," vol. 19, no. 2, pp. 291–294, Feb. 2015.

[278] G. Zheng, "Joint beamforming optimization and power control for full-duplex MIMO two-way relay channel," vol. 63, no. 3, pp. 555–566, Feb. 2015.

[279] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[280] T. Kwon, S. Lim, S. Choi, and D. Hong, "Optimal duplex mode for DF relay in terms of the outage probability," vol. 59, no. 7, pp. 3628–3634, Sep. 2010.

[281] S. Boyd, "Sequential convex programming," *Lecture Notes, Stanford University*, July 2008.

[282] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[283] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.

[284] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[285] S. Javed, O. Amin, S. S. Ikki, and M. S. Alouini, "Error probability analysis of hardware impaired systems with asymmetric transmission," *Submitted to IEEE Trans. Wireless Commun.*, 2018.

[286] W. Zhang, R. C. de Lamare, C. Pan, and M. Chen, "Joint TX/RX IQ imbalance parameter estimation using a generalized system model," in *IEEE International Conference on Communications (ICC-2015)*. London: IEEE, 2015, pp. 4704–4709.

[287] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Corporation, 1964, vol. 55.

[288] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, 2005, vol. 95.

[289] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[290] G. Ding, X. Gao, Z. Xue, Y. Wu, and Q. Shi, "Massive MIMO for distributed detection with transceiver impairments," vol. 67, no. 1, pp. 604–617, Jan. 2018.

[291] J. J. Moré and D. C. Sorensen, "Computing a trust region step," *J. Sci. Stat. Comput.*, vol. 4, no. 3, pp. 553–572, Sept. 1983.

[292] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "Approximate solution of the trust region problem by minimization over two-dimensional subspaces," *Math. Program.*, vol. 40, no. 1-3, pp. 247–263, Jan. 1988.

[293] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Improper gaussian signaling for the K-user SISO interference channel," in *IEEE Int. Conf. Commun. (ICC)*. Budapest, Hungary: IEEE, Jun. 2013, pp. 5219–5223.

[294] W. Hedhly, O. Amin, and M.-S. Alouini, "Interweave cognitive radio with improper Gaussian signaling," in *IEEE Global Commun. Conf. (GLOBECOM)*. Singapore: IEEE, Dec. 2017, pp. 1–6.

[295] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper Gaussian signaling in the k-user MISO interference channel," vol. 12, no. 12, pp. 6303–6313, Dec. 2013.

[296] H.-Y. Shin, S.-H. Park, H. Park, and I. Lee, "A new approach of interference alignment through asymmetric complex signaling and multiuser diversity," vol. 11, no. 3, pp. 880–884, Mar. 2012.

[297] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," vol. 11, no. 1, pp. 116–130, Mar. 2009.

[298] M. Gaafar, O. Amin, W. Abediseid, and M.-S. Alouini, "Spectrum sharing opportunities of full-duplex systems using improper Gaussian signaling," in *26th Annual Int.Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*. Hong Kong, China: IEEE, Sep. 2015, pp. 244–249.

[299] O. Amin, W. Abediseid, and M.-S. Alouini, "Impact of improper Gaussian signaling on the achievable rate of overlay cognitive radio," in *Wireless Commun. and Netw. Conf. (WCNC)*. San Francisco, CA, USA: IEEE, Mar. 2017.

[300] M. Soleymani, P. J. Schreier, and I. Santamaria, "Performance analysis of MIMO K-user interference channels with hardware impairments," *arXiv preprint arXiv:2001.10403*, 2020.

[301] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Improper Gaussian signaling for broadcast interference networks," vol. 26, no. 6, pp. 808–812, Jun. 2019.

[302] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper Gaussian signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 496–507, Jun. 2019.

[303] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Robust improper signaling for two-user SISO interference channels," vol. 67, no. 7, pp. 4709–4723, Jul. 2019.

[304] M. Soleymani, I. Santamaria, C. Lameiro, and P. J. Schreier, "Ergodic rate for fading interference channels with proper and improper Gaussian signaling," *Entropy*, vol. 21, no. 10, p. 922, Sep. 2019.

[305] H. D. Nguyen, R. Zhang, and S. Sun, "Improper signaling for symbol error rate minimization in k-user interference channel," vol. 63, no. 3, pp. 857–869, Mar. 2015.

[306] C. Hellings and W. Utschick, "Performance gains due to improper signals in MIMO broadcast channels with widely linear transceivers," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Vancouver, BC, Canada: IEEE, May 2013, pp. 4379–4383.

[307] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Commun. and Inf. Theory*, vol. 11, no. 3-4, pp. 185–396, 2014.

[308] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," vol. 59, no. 3, pp. 1142–1157, Mar. 2011.

[309] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," vol. 2, no. 1, pp. 57–73, 2008.

[310] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," vol. 57, no. 5, pp. 1917–1935, May 2009.

[311] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," vol. 24, no. 5, pp. 1074–1084, May 2006.

[312] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[313] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[314] W. Utschick and J. Brehmer, "Monotonic optimization framework for coordinated beamforming in multicell networks," vol. 60, no. 4, pp. 1899–1909, Apr. 2012.

[315] L. P. Qian, Y. J. Zhang, and J. Huang, "MAPEL: Achieving global optimality for a non-convex wireless power control problem," vol. 8, no. 3, pp. 1553–1563, Mar. 2009.

[316] J. L. Gross and J. Yellen, *Handbook of graph theory*. CRC press, 2004.

[317] P. C. Pendharkar, "Exhaustive and heuristic search approaches for learning a software defect prediction model," *Eng. Appl. of Artificial Intell.*, vol. 23, no. 1, pp. 34–40, 2010.

[318] R. Dijkman, M. Dumas, and L. García-Bañuelos, "Graph matching algorithms for business process model similarity search," in *Int. Conf. Business Process Manag.* Springer, 2009, pp. 48–63.

[319] S. Lee, J. Kim, and S. Cho, "Resource allocation for NOMA based D2D system using genetic algorithm with continuous pool," in *Int. Conf. Inf. and Commun. Technol. Convergence (ICTC)*. IEEE, Oct 2019, pp. 705–707.

[320] C. Lameiro, I. Santamaría, P. J. Schreier, and W. Utschick, "Maximally improper signaling in underlay MIMO cognitive radio networks," vol. 67, no. 24, pp. 6241–6255, 2019.

[321] T. Ando, "Majorizations and inequalities in matrix theory," *Linear algebra and its Appl.*, vol. 199, no. 1, pp. 17–67, Mar. 1994.

[322] M. Novey, T. Adali, and A. Roy, "A complex generalized Gaussian distributioncharacterization, generation, and estimation," vol. 58, no. 3, pp. 1427–1433, Mar. 2010.

[323] ——, "Circularity and Gaussianity detection using the complex generalized Gaussian distribution," vol. 16, no. 11, pp. 993–996, Nov. 2009.

[324] E. Ollila and V. Koivunen, "Adjusting the generalized likelihood ratio test of circularity robust to non-normality," in *10th Workshop on Signal Process.Advances in Wireless Commun.* Perugia, Italy: IEEE, Jun. 2009, pp. 558–562.

[325] D. Dini, C. Jahanchahi, and D. Mandic, "Kalman filtering for widely linear complex and quaternion valued bearings only tracking," *IET Signal Processing*, vol. 6, no. 5, pp. 435–445, Jul. 2012.

[326] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Trans. on Aeros. Electron. Syst.*, vol. 34, no. 1, pp. 103–123, Jan. 1998.

[327] Y. Zhang and X. R. Li, "Detection and diagnosis of sensor and actuator failures using IMM estimator," *IEEE Trans. on Aeros. Electron. Syst.*, vol. 34, no. 4, pp. 1293–1313, Oct. 1998.

[328] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[329] O. Lang and M. Huemer, "Classical widely linear estimation of real valued parameter vectors in complex valued environments," *arXiv preprint arXiv:1704.08825*, Apr. 2017.

[330] J. Navarro-Moreno, J. Moreno-Kayser, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Widely linear estimation algorithms for second-order stationary signals," vol. 57, no. 12, pp. 4930–4935, Dec. 2009.

[331] A. Napolitano and M. Tanda, "Doppler-channel blind identification for noncircular transmissions in multiple-access systems," vol. 52, no. 12, pp. 2073–2078, Dec. 2004.

[332] E. C. Mengüç and N. Acır, "An augmented complex-valued least-mean kurtosis algorithm for the filtering of noncircular signals," vol. 66, no. 2, pp. 438–448, Jan. 2018.

[333] A. Mohammadi and K. N. Plataniotis, "Complex-valued Gaussian sum filter for non-linear filtering of non-Gaussian/non-circular noise," vol. 22, no. 4, pp. 440–444, Apr. 2015.

[334] B. W. Han and J. H. Cho, "Capacity of second-order cyclostationary complex Gaussian noise channels," vol. 60, no. 1, pp. 89–100, Jan. 2012.

[335] J. Yeo, J. H. Cho, and J. S. Lehnert, "Joint transmitter and receiver optimization for improper-complex second-order stationary data sequence," *J. Commun. Netw.*, vol. 17, no. 1, pp. 1–11, Feb. 2015.

[336] A. Mohammadi and K. N. Plataniotis, "Improper complex-valued Bhattacharyya distance." *IEEE Trans. Neural Netw. Learning Syst.*, vol. 27, no. 5, pp. 1049–1064, May 2016.

[337] R. Zelenovsky and R. C. de Lamare, "Set-membership constrained widely linear beamforming algorithms," *Sensor Signal Processing for Defence (SSPD)*, Sep. 2012.

[338] M. Caus and A. I. Perez-Neira, "Comparison of linear and widely linear processing in MIMO-FBMC systems," in *10th Int. Symp.on Wireless Commun. Sys. (ISWCS)*. Ilmenau, Germany: VDE, Aug. 2013, pp. 1–5.

[339] P. Xiao and M. Sellathurai, "Improved linear transmit processing for single-user and multi-user MIMO communications systems," vol. 58, no. 3, pp. 1768–1779, Mar. 2010.

[340] J. A. Lpez-Fernndez, R. G. Ayestarn, I. Santamaria, and C. Lameiro, "Design of asymptotically optimal improper constellations with hexagonal packing," vol. 67, no. 8, pp. 5445–5457, Aug. 2019.

[341] D. Wubben, D. Seethaler, J. Jalden, and G. Matz, "Lattice reduction," vol. 28, no. 3, pp. 70–91, May 2011.

[342] L. Rugini, "Symbol error probability of hexagonal QAM," vol. 20, no. 8, pp. 1523–1526, Aug. 2016.

[343] S.-J. Park, "Performance analysis of triangular quadrature amplitude modulation in AWGN channel," vol. 16, no. 6, pp. 765–768, Jun. 2012.

[344] M. Tanahashi and H. Ochiai, "A multilevel coded modulation approach for hexagonal signal constellation," vol. 8, no. 10, pp. 4993–4997, Oct. 2009.

[345] K. P. Srinath and B. S. Rajan, "Fast-decodable MIDO codes with large coding gain," vol. 60, no. 2, pp. 992–1007, Feb. 2014.

[346] S. H. Han, J. M. Cioffi, and J. H. Lee, "On the use of hexagonal constellation for peak-to-average power ratio reduction of an ODFM signal," vol. 7, no. 3, pp. 781–786, Mar. 2008.

[347] M. Hekrdla and J. Sykora, "Hexagonal constellations for adaptive physical-layer network coding 2-way relaying," vol. 18, no. 2, pp. 217–220, Feb. 2014.

[348] C. R. Doerr, L. Zhang, P. Winzer, and A. H. Gnauck, "28-Gbaud InP square or hexagonal 16-QAM modulator," in *Opt. Fiber Commun. Conf.* Los Angeles, CA, USA: Optical Society of America, Mar. 2011, p. OMU2.

[349] M. Abdelaziz and T. A. Gulliver, "Triangular constellations for adaptive modulation," vol. 66, no. 2, pp. 756–766, Feb. 2018.

[350] X. Zhang, Y. Zhao, and L. Zou, "Optimum asymmetric constellation design for trellis-coded modulation over Gaussian channels," vol. 13, no. 7, pp. 528–530, Jul. 2009.

[351] L. V. Subramaniam, B. S. Rajan, and R. Bahl, "Performance of 4-and 8-state TCM schemes with asymmetric 8-PSK in fading channels," vol. 49, no. 1, pp. 211–219, Jan. 2000.

[352] S. Zhang, F. Yaman, E. Mateo, T. Inoue, K. Nakamura, and Y. Inada, "A generalized pairwise optimization for designing multi-dimensional modulation formats," in *Optical Fiber Commun. Conf. and Exhibition (OFC).* Los Angeles, CA, USA: IEEE, Mar. 2017.

[353] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," vol. 53, no. 9, pp. 74–81, Sep. 2015.

[354] Z. Qu and I. B. Djordjevic, "Geometrically shaped 16QAM outperforming probabilistically shaped 16QAM," in *European Conf.on Optical Commun.(ECOC).* Gothenburg, Sweden: IEEE, Sep. 2017.

[355] H. G. Batshon, I. B. Djordjevic, L. Xu, and T. Wang, "Iterative polar quantization-based modulation to achieve channel capacity in ultrahigh-speed optical communication systems," *IEEE Photonics J.*, vol. 2, no. 4, pp. 593–599, Aug. 2010.

[356] Z. Qu, S. Zhang, and I. B. Djordjevic, "Universal hybrid probabilistic-geometric shaping based on two-dimensional distribution matchers," in *Optical Fiber Commun. Conf. and Exposition (OFC).* San Diego, CA, USA: IEEE, Mar. 2018.

[357] J.-X. Cai, H. Batshon, M. Mazurczyk, O. Sinkin, D. Wang, M. Paskov, W. Patterson, C. Davidson, P. Corbett, G. Wolter *et al.*, "70.4 Tb/s capacity over 7,600 km in C+L

band using coded modulation with hybrid constellation shaping and nonlinearity compensation," in *Optical Fiber Commun. Conf. (OFC)*. Los Angeles, CA, USA: IEEE, Mar. 2017.

[358] F. Jardel, T. A. Eriksson, F. Buchali, W. Idler, A. Ghazisaeidi, C. Measson, and J. J. Boutros, "Experimental comparison of 64-QAM and combined geometric-probabilistic shaped 64-QAM," in *European Conf.on Optical Commun.(ECOC)*. Gothenburg, Sweden: IEEE, Sep. 2017.

[359] T. OShea and J. Hoydis, "An introduction to deep learning for the physical layer," vol. 3, no. 4, pp. 563–575, Dec. 2017.

[360] A. S. Aghaei, K. N. Plataniotis, and S. Pasupathy, "Widely linear MMSE receivers for linear dispersion space-time block-codes," vol. 9, no. 1, pp. 8–13, Jan. 2010.

[361] J. Steinwandt, R. C. de Lamare, L. Wang, N. Song, and M. Haardt, "Widely linear adaptive beamforming algorithm based on the conjugate gradient method," in *2011 Int. ITG Workshop on Smart Antennas*. Aachen, Germany: IEEE, Feb. 2011, pp. 1–4.

[362] D. Mattera, L. Paura, and F. Sterle, "Widely linear MMSE transceiver for real-valued sequences over MIMO channel," in *14th European Signal Process. Conf.* Florence, Italy: IEEE, Sep. 2006.

[363] D. Darsena, G. Gelli, and F. Verde, "Universal linear precoding for NBI-proof widely linear equalization in MC systems," *Eurasip J. Wireless Commun.and Netw.*, vol. 2008, no. 1, p. 321450, Sep. 2007.

[364] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[365] W. H. Gerstacker, F. Obernosterer, R. Schober, A. T. Lehmann, A. Lampe, and P. Gunreben, "Equalization concepts for Alamouti's space-time block code," vol. 52, no. 7, pp. 1178–1190, Jul. 2004.

[366] Y. Xia and D. P. Mandic, "A full mean square analysis of CLMS for second-order noncircular inputs," vol. 65, no. 21, pp. 5578–5590, Nov. 2017.

[367] B. Picinbono and P. Chevalier, "Extensions of the minimum variance method," *Signal processing*, vol. 49, no. 1, pp. 1–9, Feb. 1996.

[368] G. Latouche, D. Pirez, and P. Vila, "MMSE cyclic equalization," in *Proc. IEEE Military Commun. Conf. (MILCOM 98)-(Cat. No. 98CH36201)*, vol. 1. Boston, MA, USA: IEEE, Oct. 1998, pp. 150–154.

[369] G. Gelli, L. Paura, and A. M. Tulino, "Cyclostationarity-based filtering for narrowband interference suppression in direct-sequence spread-spectrum systems," vol. 16, no. 9, pp. 1747–1755, Dec. 1998.

[370] S. Buzzi, M. Lops, and A. M. Tulino, "A generalized minimum-mean-output-energy strategy for CDMA systems with improper MAI," vol. 48, no. 3, pp. 761–767, Mar. 2002.

[371] M. Witzke, "Linear and widely linear filtering applied to iterative detection of generalized MIMO signals," in *Annales des télécommunications*, vol. 60, no. 1-2. Springer, Feb. 2005, pp. 147–168.

[372] K. Kuchi and V. K. Prabhu, "Performance evaluation for widely linear demodulation of PAM/QAM signals in the presence of rayleigh fading and co-channel interference," vol. 57, no. 1, pp. 183–193, Jan. 2009.

[373] P. Chevalier and F. Dupuy, "Widely linear Alamouti receiver for the reception of real-valued constellations corrupted by interferencesThe Alamouti-SAIC/MAIC concept," vol. 59, no. 7, pp. 3339–3354, Jul. 2011.

[374] C. Hellings and W. Utschick, "Two different real representations of complex matrices for describing widely linear systems," in *International Symposium on Wireless Communication Systems (ISWCS)*. Brussels, Belgium: IEEE, Aug. 2015, pp. 641–645.

[375] Y. Cheng and M. Haardt, "Widely linear processing in MIMO FBMC/OQAM systems." in *10th Int. Symp.on Wireless Commun. Sys. (ISWCS)*. Ilmenau, Germany: VDE, Aug. 2013, pp. 1–5.

[376] S. I. Park, K. S. Kim, H.-M. Kim, and I. Song, "The DS/SSMA systems using trellis coding with asymmetric PSK signal constellation," *Signal processing*, vol. 67, no. 2, pp. 173–187, Jun. 1998.

[377] P. Xiao and M. Sellathurai, "Iterative receiver design for MIMO systems with improper signal constellations," in *IEEE Int. Conf. Commun. (ICC'09)*. Dresden, Germany: IEEE, Jun. 2009, pp. 1–5.

[378] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "On the optimal detection and error performance analysis of the hardware impaired systems," in *IEEE Global Commun. Conf. (GLOBECOM)*. Singapore: IEEE, Dec. 2017, pp. 1–7.

[379] A. E. Canbilen, S. S. Ikki, E. Basar, S. S. Gultekin, and I. Develi, "Impact of I/Q imbalance on amplify-and-forward relaying: Optimal detector design and error performance," vol. 67, no. 5, pp. 3154–3166, May 2019.

[380] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent component analysis. john willey & sons," *Inc., New York*, pp. 165–202, 2001.

[381] X.-L. Li, T. Adali, and M. Anderson, "Noncircular principal component analysis and its application to model selection," vol. 59, no. 10, pp. 4516–4528, Oct. 2011.

[382] J. Samson, "Pure states, polarized waves, and principal components in the spectra of multiple, geophysical time-series," *Geophysical Journal International*, vol. 72, no. 3, pp. 647–664, Mar. 1983.

[383] S. Wang and N. Rahnavard, "A framework for compressive sensing of asymmetric signals using Normal and Skew-Normal mixture prior." vol. 63, no. 12, pp. 5062–5072, Dec. 2015.

[384] T. Adali and V. D. Calhoun, "Complex ICA of brain imaging data," vol. 24, no. 5, pp. 136–139, Sep. 2007.

[385] F. Wang, H. Li, and R. Li, "Data mining with independent component analysis," in *6th World Congress on Intelligent Control and Automation*, vol. 2. Dalian, China: IEEE, Jun. 2006, pp. 6043–6047.

[386] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," vol. 21, no. 2, pp. 248–261, 2009.

[387] B. Zhang, M. Fu, and H. Yan, "A nonlinear neural network model of mixture of local principal component analysis: application to handwritten digits recognition," *Pattern Recognition*, vol. 34, no. 2, pp. 203–214, 2001.

[388] N. Kwak, "Principal component analysis based on L1-norm maximization," vol. 30, no. 9, pp. 1672–1680, 2008.

[389] H. Zhao, P. C. Yuen, and J. T. Kwok, "A novel incremental principal component analysis and its application for face recognition," vol. 36, no. 4, pp. 873–886, 2006.

[390] R. J. Martis, U. R. Acharya, K. Mandana, A. K. Ray, and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Expert Syst. with Appl.*, vol. 39, no. 14, pp. 11 792–11 800, 2012.

[391] B. Xiao, "Principal component analysis for feature extraction of image sequence," in *Int. Conf. on Comput. and Commun. Tech. in Agriculture Eng.*, vol. 1. IEEE, 2010, pp. 250–253.

[392] M. Suganthy and P. Ramamoorthy, "Principal component analysis based feature extraction, morphological edge detection and localization for fast iris recognition," *J. Comput. Sci.*, vol. 8, no. 9, p. 1428, 2012.

[393] I. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer, 2002.

[394] H.-L. Huang and P. Antonelli, "Application of principal component analysis to high-resolution infrared measurement compression and retrieval," *Journal of Applied Meteorology*, vol. 40, no. 3, pp. 365–388, 2001.

[395] K. Deluzio and J. Astephen, "Biomechanical features of gait waveform data associated with knee osteoarthritis: an application of principal component analysis," *Gait & Posture*, vol. 25, no. 1, pp. 86–93, 2007.

[396] M. Crescenzi and A. Giuliani, "The main biological determinants of tumor line taxonomy elucidated by a principal component analysis of microarray data," *Febs Lett.*, vol. 507, no. 1, pp. 114–118, 2001.

[397] N. Duforet-Frebourg, K. Luu, G. Laval, E. Bazin, and M. G. Blum, "Detecting genomic signatures of natural selection with principal component analysis: application to the 1000 genomes data," *Molecular biology and evolution*, vol. 33, no. 4, pp. 1082–1093, 2016.

[398] K. R. Gabriel, "The biplot graphic display of matrices with application to principal component analysis," *Biometrika*, vol. 58, no. 3, pp. 453–467, 1971.

[399] M. D. Does, J. L. Olesen, K. D. Harkins, T. Serradas-Duarte, D. F. Gochberg, S. N. Jespersen, and N. Shemesh, "Evaluation of principal component analysis image denoising on multi-exponential MRI relaxometry," *Magnetic resonance in medicine*, vol. 81, no. 6, pp. 3503–3514, 2019.

[400] H. El-Amir and M. Hamdy, *Deep Learning Pipeline*. Springer, 2020.

[401] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: Computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[402] F. Guimet, J. Ferré, R. Boqué, and F. X. Rius, "Application of unfold principal component analysis and parallel factor analysis to the exploratory analysis of olive oils by means of excitation–emission matrix fluorescence spectroscopy," *Analytica Chimica Acta*, vol. 515, no. 1, pp. 75–85, 2004.

[403] C. Hellings, P. Gogler, and W. Utschick, "Composite real principal component analysis of complex signals," in *23rd European Signal Process. Conf. (EUSIPCO)*. Nice, France: IEEE, Sep. 2015, pp. 2216–2220.

[404] X.-L. Li, M. Anderson, and T. Adali, "Principal component analysis for noncircular signals in the presence of circular white gaussian noise," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*. IEEE, 2010, pp. 1796–1801.

[405] P. J. Schreier and L. L. Scharf, "The Karhunen-Loeve expansion of improper complex random signals with applications in detection," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*, vol. 6. Hong Kong, China: IEEE, Apr. 2003, pp. VI–717.

[406] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *Int. Conf. Artificial Neural Netw.* Springer, 1997, pp. 583–588.

[407] A. Papaioannou and S. Zafeiriou, "Principal component analysis with complex kernel: The widely linear model," *IEEE Transactions on Neural networks and Learning systems*, vol. 25, no. 9, pp. 1719–1726, 2013.

[408] R. Schober, W. H. Gerstacker, and L.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," vol. 52, no. 3, pp. 746–756, 2004.

[409] S. C. Douglas and D. P. Mandic, "Performance analysis of the conventional complex LMS and augmented complex LMS algorithms," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. IEEE, 2010, pp. 3794–3797.

[410] D. Mandic, S. Javidi, S. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 34, no. 1, pp. 196–201, 2009.

[411] D. P. Mandic, S. Kanna, and S. C. Douglas, "Mean square analysis of the CLMS and ACLMS for non-circular signals: The approximate uncorrelating transform approach," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. IEEE, 2015, pp. 3531–3535.

[412] L. Li, Y. Xia, B. Jelfs, J. Cao, and D. P. Mandic, "Modelling of brain consciousness based on collaborative adaptive filters," *Neurocomputing*, vol. 76, no. 1, pp. 36–43, 2012.

[413] Y. R. Krishna, P. K. Prasad, P. Subbaiah, and B. P. Rao, "A hybrid model of CLMS and ACLMS algorithms for smart antennas," pp. 93–104, 2013.

[414] S. R. Aali, M. R. Besmi, and M. H. Kazemi, "Adaptive filtering with robust controller for improvement of inertial response of wind turbine," *Int. Trans. Elect. Energy Syst.*, vol. 29, no. 10, p. e12089, 2019.

[415] J. Lilly and J.-C. Gascard, "Wavelet ridge diagnosis of time-varying elliptical signals with application to an oceanic eddy," *Nonlinear Processes in Geophysics*, vol. 13, no. 5, pp. 467–483, Sep. 2006.

[416] A. M. Sykulski and D. B. Percival, "Exact simulation of noncircular or improper complex-valued stationary Gaussian processes using circulant embedding," in *26th Int.Workshop on Machine Learning for Signal Processing (MLSP)*. Salerno, Italy: IEEE, Sep. 2016, pp. 1–6.

[417] Y. Xia and D. P. Mandic, "Complementary mean square analysis of augmented CLMS for second order noncircular gaussian signals," vol. 24, no. 9, pp. 1413–1417, Sep. 2017.

[418] T. McWhorter and P. Schreier, "Widely-linear beamforming," in *37th Asilomar Conf. on Signals, Systems & Computers*, vol. 1. CA, USA: IEEE, Nov. 2003, pp. 753–759.

[419] F. Roemer and M. Haardt, "Multidimensional unitary tensor-ESPRIT for non-circular sources," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*. Taipei, Taiwan: IEEE, Apr. 2009, pp. 3577–3580.

[420] J.-P. Delmas, "Asymptotically minimum variance second-order estimation for non-circular signals with application to DOA estimation," vol. 52, no. 5, pp. 1235–1241, May 2004.

[421] P. Chevalier and A. Blin, "Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences," vol. 55, no. 11, pp. 5323–5336, Nov. 2007.

[422] P. Chargé, Y. Wang, and J. Saillard, "A non-circular sources direction finding method using polynomial rooting," *Signal Processing*, vol. 81, no. 8, pp. 1765–1770, Aug. 2001.

[423] C. Xian-Ning, W. Qun, D. Heng, S. X. Feng, and Y. Wan-Lin, "Direction of arrival estimation and identification of mixed circular and non-circularity sources," in *Int. Conf. on Commun., Circuits and Sys.* Fujian, China: IEEE, May 2008, pp. 918–921.

[424] L. De Lathauwer and B. De Moor, "On the blind separation of non-circular sources," in *11th European Signal Process. Conf.* Toulouse, France: IEEE, Sep. 2002.

[425] B. Otterseten, M. Kristensson, and D. Astely, "Receiver and method for rejecting cochannel interference," *Google Patents*, 2005.

[426] A. Mostafa, R. Kobylinski, I. Kostanic, and M. Austin, "Single antenna interference cancellation (SAIC) for GSM networks," in *58th Veh. Techn. Conf. (VTC)-Fall (IEEE Cat. No. 03CH37484)*, vol. 2. Orlando, FL, USA: IEEE, Oct. 2003, pp. 1089–1093.

[427] "MUROS downlink receiver performance for interference and sensitivity," *GP-090115 3GPP TSG Geran Tdoc, GERAN no. 41*, 2009.

[428] "MUROS uplink receiver performance," *GP-090114 3GPP TSG Geran Tdoc, GERAN no. 41*, 2009.

[429] K. Liu, M. Tao, and X. Yuan, "Optimal DoF region for the asymmetric two-pair MIMO two-way relay channel," in *IEEE Global Commun. Conf. (GLOBECOM)*. Washington, DC, USA: IEEE, Dec. 2016.

[430] Y. Zou, M. Valkama, and M. Renfors, "Digital compensation of I/Q imbalance effects in space-time coded transmit diversity systems," vol. 56, no. 6, pp. 2496–2508, Jun. 2008.

[431] A. Hakkarainen, J. Werner, and M. Valkama, "RF imperfections in antenna arrays: Response analysis and widely-linear digital beamforming," in *IEEE Radio and Wireless Symposium (RWS)*. Austin, TX, USA: IEEE, Jan. 2013, pp. 187–189.

[432] R. Carvajal, J. C. Aguero, B. I. Godoy, and G. C. Goodwin, "EM-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," vol. 62, no. 1, pp. 152–160, Jan. 2013.

[433] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[434] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European J. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[435] T. Fusco and M. Tanda, "ML frequency offset and carrier phase estimation in OFDM systems with noncircular transmissions," in *12th European Signal Process. Conf.* Vienna, Austria: IEEE, Sep. 2004, pp. 897–900.

[436] C. Enneking, F. Antreich, L. Krieger, and A. L. de Almeida, "Gaussian approximations for intra-and intersystem interference in RNSS," vol. 23, no. 7, pp. 1198–1201, Jul. 2019.

[437] C. Hellings and W. Utschick, "On the worst-case noise in Gaussian MIMO systems with proper and with improper signaling," in *21st Int. ITG Workshop on Smart Antennas (WSA)*. Berlin, Germany: VDE, Mar. 2017, pp. 1–7.

[438] Y. C. Yoon and H. Leib, "Matched filtering in improper complex noise and applications to DS-CDMA," in *6th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'95)*, vol. 2. Toronto, Ontario, Canada: IEEE, Sep. 1995, pp. 701–705.

[439] A. S. Aghaei, K. N. Plataniotis, and S. Pasupathy, "Maximum likelihood binary detection in improper complex Gaussian noise," in *IEEE Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*. Las Vegas, NV, USA: IEEE, Apr. 2008, pp. 3209–3212.

[440] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[441] C. Hellings, F. Weisser, and W. Utschick, "High-SNR analysis of improper signaling in the MIMO broadcast channel with TIN," in *12th Int. ITG Conf. on Sys. Commun. and Coding (SCC)*. Rostock, Germany: VDE, Feb. 2019, pp. 1–6.

[442] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the SISO Z-interference channel," vol. 21, no. 11, pp. 2432–2435, Nov. 2017.

[443] ——, "Proper time-sharing as a baseline for studying improper signaling in interference channels," in *22nd Int. ITG Workshop on Smart Antennas (WSA)*. Bochum, Germany: VDE, Mar. 2018, pp. 1–6.

[444] A. Showail, K. Jamshaid, and B. Shihada, "Buffer sizing in wireless networks: challenges, solutions, and opportunities," vol. 54, no. 4, pp. 130–137, Apr. 2016.

[445] B. Shihada and K. Jamshaid, "Buffer sizing for multi-hop networks," Jan. 2014, uS Patent 8,638,686.

[446] N. Bouacida and B. Shihada, "Practical and dynamic buffer sizing using LearnQueue," vol. 18, no. 8, pp. 1885–1897, Aug. 2019.

[447] J. P. M. Salehi, *Digital Communications 5e*. McGraw-Hill Inc., 2008.

[448] C. N. Moore, "On the uniform convergence of the developments in Bessel functions," *Transactions of the American Mathematical Society*, vol. 12, no. 2, pp. 181–206, Apr. 1911.

[449] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, June 1963.

APPENDICES

A Computational Complexity of the Joint Algorithm

For a MH-DF-FDR system accommodating a source and k relays, the total number of optimization variables in joint optimization are $N_1 = 2k + 2$, considering two optimization variables (real and imaginary component of the transmit pseudo-variance) per transmitting node. Moreover, the number of inequalities representing constraints of the joint optimization problem is defined as $M_1 = 2(k + 1)$, involving two constraints (power and rate constraints) per transmitting node, as given in **4-P6**. In addition, $F_1 = N_1(1 + M_1) + N_1^2 M_1$ is the cost of evaluating the first and second derivatives of the objective and constraint functions in the joint algorithm. The computational complexity of the joint algorithm can be simplified as follows:

$$O\left(I_1^{\mathrm{SCP}}\left(N_1^2 k + \alpha \max\left(N_1^3, N_1^2 M_1, F_1\right)\right)\right) \tag{A.1}$$

Where, α is assumed to be between 10 and 100 for the interior point method [282]. Considering $N_1^2 k \leq \alpha \max\left(N_1^3, N_1^2 M_1, F_1\right)$, we can simplify the complexity expression as

$$O\left(\alpha I_1^{\mathrm{SCP}} \max\left(N_1^3, N_1^2 M_1, F_1\right)\right) \tag{A.2}$$

Using $N_1^3 = N_1^2 M_1 = 8k^3 + 24k^2 + 24k + 8$ and $F_1 = 8k^3 + 28k^2 + 34k + 14$, depicts $F_1 = \max\left(N_1^3, N_1^2 M_1, F_1\right)$. Thus, the complexity reduces to $O\left(\alpha I_1^{\mathrm{SCP}} F_1\right)$. Considering

the dominant term in F_1 renders the desired complexity of order $O\left(\alpha I_1^{\mathrm{SCP}} k^3\right)$.

B Computational Complexity of the Distributed Algorithm

Equivalently, the computational complexity of the distributed algorithm with cluster size Γ, the total number of optimization variables $N_2 = 2\Gamma - 2$, the number of inequalities representing constraints $M_2 = 2(\Gamma - 1)$ and the costs of evaluating the first and second derivatives of the objective and constraint functions $F_2 = N_2(1 + M_2) + N_2^2 M_2$ can be simplified to $O\left(\alpha I_2^{\mathrm{SCP}} \max\left(N_2^3, N_2^2 M_2, F_2\right)\right)$. Following the similar steps as in Appendix A with $F_2 = \max\left(N_2^3, N_2^2 M_2, F_2\right)$ and using $F_2 = 8\Gamma^3 - 20\Gamma^2 + 18\Gamma - 6$, yields the complexity of order $O\left(\alpha I_2^{\mathrm{SCP}} \Gamma^3\right)$.

C Derivation of $F_\vartheta\left(\vartheta\right)$

Consider the Rayleigh fading channel where real and imaginary components of the channel coefficient are distributed as $h_j \sim \mathcal{CN}(0, \lambda/2, 0); j \in \{r, i\}$. This implies h_j^2 to be distributed as unnormalized chi-square variable with first degree of freedom i.e. $f_{h_j^2}\left(x\right) \sim \frac{1}{\sqrt{\pi\lambda x}} e^{\frac{-x}{\lambda}}; x \geq 0;$. Using (5.26), the CDF of ϑ is defined as $F_\vartheta\left(\vartheta\right) = \Pr\left\{\frac{\alpha h_{\mathrm{r}}^2 + \beta h_{\mathrm{i}}^2}{1+\gamma\left(h_{\mathrm{r}}^2 + h_{\mathrm{i}}^2\right)} \leq \vartheta\right\}$ which can also be represented as the probability of the event

$$F_\vartheta\left(\vartheta\right) = \Pr\left\{\left(\alpha - \gamma\vartheta\right) h_{\mathrm{r}}^2 + \left(\beta - \gamma\vartheta\right) h_{\mathrm{i}}^2 \leq \vartheta\right\} = \Pr\left\{S \leq \vartheta\right\} \tag{C.1}$$

PDF of $\left(\alpha - \gamma\vartheta\right) h_{\mathrm{r}}^2$ is evaluated using careful transformations.

$$\left(\alpha - \gamma\vartheta\right) h_{\mathrm{r}}^2 \sim \frac{1}{\sqrt{\pi\lambda\left(\alpha - \gamma\vartheta\right) x}} e^{\frac{-x}{\left(\alpha - \gamma\vartheta\right)\lambda}}; \quad x \geq 0 \tag{C.2}$$

Similarly, the distribution of $(\beta - \gamma\vartheta)\, h_{\mathrm{i}}^2$ is identical to (C.2) and can be obtained by replacing α with β in (C.2) pertaining to iid channel coefficients. Next, we convolve the derived PDFs under the assumption of uncorrelated real and imaginary channel coefficients. Then employing Mathematica integrator to compute the convolution, we obtain

$$f_S(s) = \frac{1}{k_1(\vartheta)} e^{-sk_2(\vartheta)} \mathcal{I}_0\left\{sk_3(\vartheta)\right\}; s \geq 0 \tag{C.3}$$

where, $k_1(\vartheta) = \lambda\sqrt{(\alpha - \gamma\vartheta)(\beta - \gamma\vartheta)}$, $k_2(\vartheta) = \frac{\lambda(\alpha+\beta-2\gamma\vartheta)}{2[k_1(\vartheta)]^2}$, $k_3(\vartheta) = \frac{\lambda(\beta-\alpha)}{2[k_1(\vartheta)]^2}$ and $\mathcal{I}_0\{x\}$ is the zero-order modified Bessel function, which has the following power series representation [447]

$$\mathcal{I}_0\{x\} = \sum_{m=0}^{\infty} \frac{x^{2m}}{[m!]^2 2^{2m}}. \tag{C.4}$$

From (C.1), the distribution function $F_\vartheta(\vartheta)$ is given as

$$F_\vartheta(\vartheta) = \frac{1}{k_1(\vartheta)} \int_0^{\vartheta} e^{-sk_2(\vartheta)} \mathcal{I}_0\left\{sk_3(\vartheta)\right\} ds, \tag{C.5}$$

Eventually, substituting $t = sk_3(\vartheta)$ and the power series representation of $\mathcal{I}_0\{t\}$ yields

$$F_\vartheta(\vartheta) = \frac{1}{k_1(\vartheta)\, k_3(\vartheta)} \int_0^{\vartheta k_3(\vartheta)} e^{-t\frac{k_2(\vartheta)}{k_3(\vartheta)}} \sum_{m=0}^{\infty} \frac{t^{2m}}{[m!]^2 2^{2m}}\, dt \tag{C.6}$$

where $\delta = \max(\alpha, \beta)/\gamma$ and $0 \leq \vartheta \leq \delta$. Based on the uniform convergence [448], swapping the order of integral and summation yields the closed form distribution function $F_\vartheta(\vartheta)$ as

$$F_\vartheta(\vartheta) = \frac{1}{k_1(\vartheta)} \sum_{m=0}^{\infty} \frac{2^{-2m} k_3(\vartheta)^{2m}}{[m!]^2 k_2(\vartheta)^{2m+1}} \left[2m! - \Gamma\left(2m+1, \vartheta k_2(\vartheta)\right)\right]. \tag{C.7}$$

where $\Gamma(a, x)$ is the upper incomplete gamma function [287, (6.5.12)].

D Chernoff Bound on Average PEP

Substituting $f_\varrho(\varrho)$ from (5.39) to (5.40) yields

$$\bar{P}_{\mathrm{s}} \leq \frac{\psi}{2\lambda\sqrt{\alpha\beta}} \int_0^\infty e^{-\frac{\varrho}{2\lambda}\left(\lambda+\frac{1}{\alpha}+\frac{1}{\beta}\right)} \mathcal{I}_0 \left\{ \frac{\varrho}{2\lambda} \left(\frac{1}{\alpha} - \frac{1}{\beta} \right) \right\} d\varrho \tag{D.1}$$

Substituting $t = \varrho(\beta - \alpha) / (2\lambda\alpha\beta)$ yields

$$\bar{P}_{\mathrm{s}} \leq \frac{\sqrt{\alpha\beta}\psi}{|\beta - \alpha|} \int_0^\infty e^{-\frac{t(\alpha+\beta+\alpha\beta\lambda)}{|\beta-\alpha|}} \mathcal{I}_0\{t\} dt \tag{D.2}$$

Using the power series representation of $\mathcal{I}_0\{t\}$ from (C.4) in (D.2) and swapping integration and summation based on the uniform convergence yields

$$\bar{P}_{\mathrm{s}} \leq \frac{\sqrt{\alpha\beta}\psi}{|\beta - \alpha|} \sum_{m=0}^{\infty} \frac{1}{(m!)^2 2^{2m}} \int_0^\infty e^{-\frac{t(\alpha+\beta+\alpha\beta\lambda)}{|\beta-\alpha|}} t^{2m} dt \tag{D.3}$$

Finally, we utilize the well-known integral substitution as given in (D.4) to comprehend (D.3).

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}} \tag{D.4}$$

Thus, (D.3) reduces to (D.5) and we can obtain the desired closed form expression as given in (5.41) after some trivial simplifications.

$$\bar{P}_{\mathrm{s}} \leq \frac{\sqrt{\alpha\beta}\psi}{|\beta - \alpha|} \sum_{m=0}^{\infty} \frac{(2m)!}{(m!)^2 2^{2m}} \left(\frac{\alpha\beta(\alpha + \beta + \alpha\beta\lambda)}{|\beta - \alpha|} \right)^{-(2m+1)} \tag{D.5}$$

E Moment Generating Function of ϱ

Let $\varrho = \Theta_r + \Theta_i$ where $\Theta_p = k_p \hat{h}_p^2$ with $p \in \{r, i\}$, $k_r = \frac{\alpha}{2}$ and $k_i = \frac{\beta}{2}$. Thus, $\hat{h}_p^2$ is a chi-square RV $\chi^2(1)$ and the MGF of Θ_p is given as $\mathrm{M}_{\Theta}(s) = \frac{1}{\sqrt{1-k_p s}}$. Furthermore, the independent Θ_r and Θ_i dictate $M_{\varrho}(s) = \mathrm{M}_{\Theta_\mathrm{r}}(s)\,\mathrm{xM}_{\Theta_i}(\mathrm{s})$ as given in (5.47).

F Derivation of $f_{\Omega}(\Omega)$

Using (5.49), the CDF of Ω is $F_{\Omega}(\Omega) = \Pr\left\{\frac{\alpha|h|^2}{1+\gamma|h|^2} \leq \Omega\right\} = \Pr\left\{(\alpha - \gamma\Omega)|h|^2 \leq \Omega\right\}$. Rayleigh fading channel assumption yields $f_{|h|^2}(x) = \lambda e^{-\lambda x}$. This implies $(\alpha - \gamma\Omega)|h|^2 \sim \mathrm{Exp}\left[(\alpha - \gamma\Omega)/\lambda\right]$ Thus, $F_{\Omega}(\Omega) = 1 - e^{\frac{-\lambda\Omega}{\alpha-\gamma\Omega}}$, that can be differentiated giving $f_{\Omega}(\Omega)$ as given in (5.50).

G Optimality of Problem 5-P2

Considering $\mathbf{D} \in \mathbb{C}^{2x2}$ and $\mathbf{D} = \widetilde{\mathbf{H}}\mathbf{Y}\widetilde{\mathbf{H}}^T$. Clearly $\mathbf{Y}$ is a symmetric positive definite matrix owing to the positive trace $\mathrm{Tr}(\mathbf{Y}) = 1/\sigma_{\mathrm{I}}^2 + 1/\sigma_{\mathrm{Q}}^2$ and the positive determinant $|\mathbf{Y}| = (1-\rho^2)/\sigma_{\mathrm{I}}^2\sigma_{\mathrm{Q}}^2$. This implies that $\mathbf{D}$ is also a symmetric positive definite matrix. Assuming $\mathbf{q}_{mn} = \mathbf{R}\mathbf{s}_{mn}$, the constraint becomes

$$\begin{aligned}&\mathbf{s}_{mn}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\tilde{\mathbf{H}}\mathbf{Y}\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{A}\mathbf{R}\mathbf{s}_{mn} = \mathbf{q}_{mn}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{D}\mathbf{A}\mathbf{q}_{mn} = (1+\eta)\times \\ &(\mathbf{q}_{mn}^1)^2\mathbf{D}_{11} + (1-\eta)(\mathbf{q}_{mn}^2)^2\mathbf{D}_{22} + 2\sqrt{1-\eta^2}\mathbf{q}_{mn}^1\mathbf{q}_{mn}^2\mathbf{D}_{12}\end{aligned} \tag{G.1}$$

where $\mathbf{q}_{mn}^1$ and $\mathbf{q}_{mn}^2$ are the first and second element of vector $\mathbf{q}_{mn}$ respectively. They represent the distance between the real and imaginary components of symbols m and n in a rotated signal constellation respectively. Conclusively, (G.1) demonstrates the concave nature of the **5-P2** in η.

H Optimailty of Problem 5-P3

Considering the constraint in **5-P3**,

$$f(\eta) = Q\left(\sqrt{h(\eta)}\right) = Q\left(\sqrt{\frac{p\mathbf{s}_{mn}^T\mathbf{R}^T\mathbf{A}^T\widetilde{\mathbf{H}}\mathbf{Y}\widetilde{\mathbf{H}}^T\mathbf{A}\mathbf{R}\mathbf{s}_{mn}}{4(1-\rho^2)}}\right) \tag{H.1}$$

where $h(\eta)$ is proven to be concave in η in Appendix G. Furthermore, $\sqrt{h(\eta)}$ is also concave owing to the positive nature of $h(\eta)$. The composite function in (H.1) is proven to be convex as

$$f''(\eta) = Q''\left(\sqrt{h(\eta)}\right)\left(\left(\sqrt{h(\eta)}\right)'\right)^2 + Q'\left(\sqrt{h(\eta)}\right)\left(\sqrt{h(\eta)}\right)'' \tag{H.2}$$

where, $Q''\left(\sqrt{h(\eta)}\right) \geq 0$ and $Q'\left(\sqrt{h(\eta)}\right) \leq 0$ owing to convex decreasing nature of $Q(x)$ in $\mathbb{R}_+$. Moreover, $\left(\sqrt{h(\eta)}\right)'' \leq 0$ yields $f''(\eta) \geq 0$ proving convex constraints and **5-P3** optimality.

I Statistical Characterization of Aggregate Noise

The superposed Gaussian distributions render the accumulative noise $z \sim \mathcal{CN}(0, v, \tilde{v})$, where $v = \alpha|g|^2\kappa + \sigma_w^2$ and $\tilde{v} = \alpha g^2\tilde{\kappa}$. Exploiting the relation between the v, $\tilde{v}$ and the variances of $\sigma_I^2 = E\{z_I^2\}$, $\sigma_Q^2 = E\{z_Q^2\}$ and their mutual correlation $r_{z_I z_Q} = E\{z_I z_Q\}$,

we get

$$v = E\left\{|z|^2\right\} = \sigma_I^2 + \sigma_Q^2. \tag{I.1}$$

$$\tilde{v} = E\left\{z^2\right\} = \sigma_I^2 - \sigma_Q^2 + i2r_{z_I z_Q}. \tag{I.2}$$

Their inter relation enables us to evaluate σ_I^2, σ_Q^2, and $r_{z_I z_Q}$ from v and $\tilde{v}$ as

$$\sigma_I^2 = \frac{v + \tilde{v}_I}{2} = \frac{\alpha|g|^2\kappa + \sigma_w^2 + \alpha\Re\left(g^2\tilde{\kappa}\right)}{2}, \tag{I.3}$$

$$\sigma_Q^2 = \frac{v - \tilde{v}_I}{2} = \frac{\alpha|g|^2\kappa + \sigma_w^2 - \alpha\Re\left(g^2\tilde{\kappa}\right)}{2}, \tag{I.4}$$

$$r_{z_I z_Q} = \frac{\tilde{v}_Q}{2} = \frac{\alpha\Im\left(g^2\tilde{\kappa}\right)}{2}. \tag{I.5}$$

Finally, (I.3)-(I.5) allow us to find the correlation coefficient between z_I and z_Q as

$$\rho_z = \frac{r_{z_I z_Q}}{\sigma_I \sigma_Q} = \frac{\alpha\Im\left(g^2\tilde{\kappa}\right)}{\sqrt{\left(\alpha|g|^2\kappa + \sigma_w^2\right)^2 - \left(\alpha\Re\left(g^2\tilde{\kappa}\right)\right)^2}}. \tag{I.6}$$

J Translation within power budget

In this appendix we present the proof of Remark 1. It is straightforward to prove that the translation $\mathbf{v} = \mathbf{A}\mathbf{w}$ does not change the variance/power but only introduce asymmetry/improperness. Considering the transformation caused by the translation $v = \sqrt{1+\zeta}w_I + i\sqrt{1-\zeta}w_Q$, the power/variance is given by

$$\sigma_v^2 = (1+\zeta)\sigma_{w_I}^2 + (1-\zeta)\sigma_{w_Q}^2. \tag{J.1}$$

Using the symmetric nature of r.v. w i.e., $\sigma^2_{w_I} = \sigma^2_{w_Q}$, it is clear that $\sigma^2_v = \sigma^2_w$. On the other hand, the pseudo-variance can be calculated as

$$\tilde{\sigma}^2_v = (1+\zeta)\sigma^2_{w_I} - (1-\zeta)\sigma^2_{w_Q} + i2\sqrt{1-\zeta^2}E\left\{w_I w_Q\right\}. \tag{J.2}$$

Again, the symmetry implies $E\left\{w_I w_Q\right\} = 0$. Thus, the circularity coefficient can be derived from (J.2) i.e., $|\tilde{\sigma}^2_v|/\sigma^2_v = \zeta$.

The same concept can be extended to the symmetric discrete constellations with uniform prior probabilities. Considering the transformation caused by the translation $v_m = \sqrt{1+\zeta}x_{mI} + i\sqrt{1-\zeta}x_{mQ}$, the power of the transformed constellation is given by

$$P = \frac{1}{M}\left((1+\zeta)\sum_{m=1}^{M} x^2_{mI} + (1-\zeta)\sum_{m=1}^{M} x^2_{mQ}\right). \tag{J.3}$$

Using the symmetric property of the original discrete constellation $\sum\limits_{m=1}^{M} x^2_{mI} = \sum\limits_{m=1}^{M} x^2_{mQ}$, it is clear that the power is preserved as $P = \frac{2}{M}\sum\limits_{m=1}^{M} x^2_{mI}$. Moreover, the non-zero pseudo-variance is given by

$$\tilde{P} = \zeta P + \frac{2i}{M}\sqrt{1-\zeta^2}\sum_{m=1}^{M} x_{mI}\, x_{mQ}. \tag{J.4}$$

Again, the symmetry implies $\sum\limits_{m=1}^{M} x_{mI}x_{mQ} = 0$. Thus, the circularity coefficient can be

Table K.1: First Order Necessary KKT Conditions

Index	KKT Conditions	Satisfied with	Reason
$1:M_{\text{nu}}$	$\nabla_{\mathbf{p}}\mathcal{L}\left(\mathbf{p}^*,\lambda^*\right)$=0, $\forall m$	$\nabla_{\mathbf{p}}\mathcal{L}\left(\mathbf{p}^*,\lambda_1^*,\lambda_2^*,\lambda_3^*\right)$=0	Saddle point of dual problem
M_{nu}+1	$\lambda_1^*\left(\sum_{m=1}^{M_{\text{nu}}}\left\|x_m\right\|^2 p_m^*-1\right)$=0	$\sum_{m=1}^{M_{\text{nu}}}\left\|x_m\right\|^2 p_m^*=1$, $\lambda_1^*\geq 0$	Maximum power transmission
M_{nu}+2	$\lambda_2^*\left(\sum_{m=1}^{M_{\text{nu}}} p_m^*-1\right)$=0	$\sum_{m=1}^{M_{\text{nu}}} p_m^*=1$, $\lambda_2^*\geq 0$	Equality constraint
M_{nu}+3	$\lambda_3^*\left(\log_2\left(M_{\text{u}}\right)-\text{H}(\mathbf{p}^*)\right)$=0	$\text{H}(\mathbf{p}^*)=\log_2\left(M_{\text{u}}\right)$, $\lambda_3^*\geq 0$	BER-Rate tradeoff

derived from (J.4), i.e., $|\tilde{P}|/P=\zeta$.

K KKT Conditions

The convex non-linear constraint problem **6-P1a** can be efficiently solved using the first order necessary KKT conditions. We begin by writing the Lagrangian function $\mathcal{L}$ as

$$\mathcal{L}\left(\mathbf{p},\lambda_1,\lambda_2,\lambda_3\right)=\tilde{\text{P}}_{\text{b}}^{\text{UB}}\left(\mathbf{p},\mathbf{p}^{(k)}\right)+\lambda_1\left(\sum_{m=1}^{M}\left|x_m\right|^2 p_m-1\right)+\lambda_2\left(\sum_{m=1}^{M} p_m-1\right)+\lambda_3\left(\log_2\left(M_{\text{u}}\right)-\text{H}(\mathbf{p})\right), \tag{K.1}$$

where the Lagrange multipliers are $\lambda_1,\lambda_2,\lambda_3\geq 0$. Next. we evaluate the gradient of the (K.1) with respect to the optimization variables in $\mathbf{p}$

$$\nabla_{\mathbf{p}}\mathcal{L}=\left[\frac{\partial\mathcal{L}}{\partial p_1}\quad\frac{\partial\mathcal{L}}{\partial p_2}\quad\cdots\quad\frac{\partial\mathcal{L}}{\partial p_{M_{\text{nu}}}}\right], \tag{K.2}$$

where the partial derivative of $\mathcal{L}$ with respect to p_m is given by

$$\frac{\partial\mathcal{L}}{\partial p_m}=\frac{\partial\text{P}_{\text{b}}^{\text{UB}}\left(\mathbf{p}^{(k)}\right)}{\partial p_m}+\lambda_1\left|x_m\right|^2+\lambda_2+\lambda_3\left(\frac{1}{\ln(2)}+\log_2\left(p_m\right)\right),\quad\forall\,1\leq m\leq M_{\text{nu}} \tag{K.3}$$

Suppose that there is a local solution $\mathbf{p}^*$ of **6-P1a** and the objective function $\tilde{\text{P}}_{\text{b}}^{\text{UB}}\left(\mathbf{p},\mathbf{p}^{(k)}\right)$ along with the constraints (6.22b) and (6.22c) are continuously differentiable. Then, there

exists a Lagrange multiplier vector λ^*, with components λ_i, where $i \in (1,2,3)$, such that the necessary first order KKT conditions (as presented in Table K.1) are satisfied at $(\mathbf{p}^*, \lambda^*)$. Interestingly, the KKT conditions are satisfied with

$$\nabla_{\mathbf{p}} \mathcal{L}\left(\mathbf{p}^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\right) = 0, \tag{K.4}$$

$$\sum_{m=1}^{M} |x_m|^2 p_m^* = 1, \tag{K.5}$$

$$\sum_{m=1}^{M} p_m^* = 1, \tag{K.6}$$

$$\mathrm{H}(\mathbf{p}^*) = \log_2\left(M_{\mathrm{u}}\right). \tag{K.7}$$

owing to the maximum transmission power preference, equality constraint and BER -Rate trade-off , respectively. Thus, the $M_{\mathrm{nu}}+3$ solution parameters $\left(p_1^*, p_2^*, \ldots, p_{M_{\mathrm{nu}}}^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\right)$ can be efficiently obtained by solving equations (K.4)-(K.7) using Levenberg-Marquardt algorithm [449].

L Gradient for Optimization

The gradient of the upper bound on BER w.r.t GS parameters is given as

$$\nabla_{\mathcal{G}} \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}} = \begin{bmatrix} \dfrac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial \zeta} & \dfrac{\partial \mathrm{P}_{\mathrm{b}}^{\mathrm{UB}}}{\partial \theta} \end{bmatrix} = \frac{1}{\log_2(M)} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} \Delta_{mn} \begin{bmatrix} \dfrac{\partial \gamma_{mn}}{\partial \zeta} & \dfrac{\partial \gamma_{mn}}{\partial \theta} \end{bmatrix}, \tag{L.1}$$

where Δ_{mn} is the common part in both partial derivatives.

$$\Delta_{mn} = \frac{p_m \gamma_{mn}^{-3/2}}{2\sqrt{2\pi}} \sqrt{\frac{1-\rho_z^2}{\alpha}} e^{-\frac{\Omega_{mn}^2}{2}} \left(\ln\left(\frac{p_m}{p_n}\right) - \frac{1}{2\beta_{mn}^2} \right). \tag{L.2}$$

Moreover, the partial derivative of γ_{mn} with respect to the translation parameter ζ is given as

$$\frac{\partial \gamma_{mn}}{\partial \zeta} = \frac{\bar{\xi}^2_{mnI}}{\sigma_{\mathrm{I}}^2} + \frac{2\rho_z \bar{\xi}_{mnI}\bar{\xi}_{mnQ}}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}} \frac{\zeta}{\sqrt{1-\zeta^2}} - \frac{\bar{\xi}^2_{mnQ}}{\sigma_{\mathrm{Q}}^2}, \tag{L.3}$$

where, $\bar{\xi}_{mnI} = \xi_{mnI}\cos(\theta) - \xi_{mnQ}\sin(\theta)$ and $\bar{\xi}_{mnQ} = \xi_{mnI}\sin(\theta) + \xi_{mnQ}\cos(\theta)$. Furthermore, the partial derivative of γ_{mn} with respect to the rotation parameter is evaluated as

$$\begin{aligned}\frac{\partial \gamma_{mn}}{\partial \theta} &= 2\frac{1+\zeta}{\sigma_{\mathrm{I}}^2}\left(\xi_{mnI}\cos(\theta) - \xi_{mnQ}\sin(\theta)\right)\left(-\xi_{mnI}\sin(\theta) - \xi_{mnQ}\cos(\theta)\right) + \\ &+2\frac{1-\zeta}{\sigma_{\mathrm{Q}}^2}\left(\xi_{mnI}\sin(\theta) + \xi_{mnQ}\cos(\theta)\right)\left(\xi_{mnI}\cos(\theta) - \xi_{mnQ}\sin(\theta)\right) + \\ &-\frac{2\rho_z}{\sigma_{\mathrm{I}}\sigma_{\mathrm{Q}}}\sqrt{1-\zeta^2}\left({\xi_{mnI}}^2\cos(2\theta) - {\xi_{mnQ}}^2\cos(2\theta) - 2\xi_{mnI}\xi_{mnQ}\sin(2\theta)\right).\end{aligned} \tag{L.4}$$

M Papers Published and Submitted

Journal Papers

Published

• Sidrah Javed, Osama Amin, Basem Shihada and Mohamed-Slim Alouini, "A Journey from Improper Gaussian Signaling to Asymmetric Signaling", *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1539-1591, Apr. 2020.

• Sidrah Javed, Osama Amin, Basem Shihada and Mohamed-Slim Alouini, "Improper Gaussian signaling for hardware impaired multihop full-duplex relaying systems", *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1858-1871, Mar. 2019.

• Sidrah Javed, Osama Amin, Salama S. Ikki, and Mohamed-Slim Alouini, "Asymmetric modulation for hardware impaired systems - Error probability analysis and receiver design", *IEEE Trans. Wireless Commun.*, , vol. 18, no. 3, pp. 1723-1738, Feb. 2019.

Submitted

• Sidrah Javed, Ahmed Elzanaty, Osama Amin, Basem Shihada and Mohamed-Slim Alouini, "When Probabilistic Shaping Realizes Improper Signaling for Hardware Distortion Mitigation", submitted to *IEEE Trans. Commun.*, Sept. 2020.

Conference Papers

- Sidrah Javed, Osama Amin, Salama S. Ikki, and Mohamed-Slim Alouini, "On the achievable rate of hardware-impaired transceiver systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. Singapore: IEEE, Dec. 2017, pp. 1-6.

- Sidrah Javed, Osama Amin, Salama S. Ikki, and Mohamed-Slim Alouini, "On the optimal detection and error performance analysis of the hardware impaired systems," in *IEEE Global Commun. Conf. (GLOBECOM).* Singapore: IEEE, Dec. 2017, pp. 1-7.