A Journey from Improper Gaussian Signaling to Asymmetric Signaling

Sidrah Javed [1], Osama Amin [2], Basem Shihada [2], and Mohamed-Slim Alouini [2]

[1]KAUST
[2]Affiliation not available

October 30, 2023

Abstract

p.p1 {margin: 0.0px 0.0px 0.0px 0.0px; font: 9.0px Helvetica}
The deviation of continuous and discrete complex
random variables from the traditional proper and symmetric
assumption to a generalized improper and asymmetric characterization
(accounting correlation between a random entity
and its complex conjugate), respectively, introduces new design
freedom and various potential merits. As such, the theory of
impropriety has vast applications in medicine, geology, acoustics,
optics, image and pattern recognition, computer vision, and
other numerous research fields with our main focus on the
communication systems. The journey begins from the design of
improper Gaussian signaling in the interference-limited communications
and leads to a more elaborate and practically
feasible asymmetric discrete modulation design. Such asymmetric
shaping bridges the gap between theoretically and practically
achievable limits with sophisticated transceiver and detection
schemes in both coded/uncoded wireless/optical communication
systems. Interestingly, introducing asymmetry and adjusting the
transmission parameters according to some design criterion
render optimal performance without affecting the bandwidth or
power requirements of the systems. This dual-flavored article
initially presents the tutorial base content covering the interplay
of reality/complexity, propriety/impropriety and circularity/noncircularity
and then surveys majority of the contributions in this
enormous journey.

A Journey from Improper Gaussian Signaling to Asymmetric Signaling

Sidrah Javed, Osama Amin, Basem Shihada and Mohamed-Slim Alouini

Abstract

The deviation of continuous and discrete complex random variables from the traditional proper and symmetric assumption to a generalized improper and asymmetric characterization (accounting correlation between a random entity and its complex conjugate), respectively, introduces new design freedom and various potential merits. As such, the theory of impropriety has vast applications in medicine, geology, acoustics, optics, image and pattern recognition, computer vision, and other numerous research fields with our main focus on the communication systems. The journey begins from the design of improper Gaussian signaling in the interference-limited communications and leads to a more elaborate and practically feasible asymmetric discrete modulation design. Such asymmetric shaping bridges the gap between theoretically and practically achievable limits with sophisticated transceiver and detection schemes in both coded/uncoded wireless/optical communication systems. Interestingly, introducing asymmetry and adjusting the transmission parameters according to some design criterion render optimal performance without affecting the bandwidth or power requirements of the systems. This dual-flavored article initially presents the tutorial base content covering the interplay of reality/complexity, propriety/impropriety and circularity/non-circularity and then surveys majority of the contributions in this enormous journey.

Index Terms

Achievable rate, asymmetric signaling, circularity, data analysis, detection, DoF, equalization, error probability, estimation, filtering, IGS, Impropriety, outage probability, power efficiency, separation, signal processing, widely linear processing, wireless communications

I. Introduction

The ever growing demands of energy and bandwidth efficient systems in diverse fields of science and technology have compelled researchers to think beyond the traditional approaches

S. Javed, O. Amin, B. Shihada and M.-S. Alouini are with CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia. E-mail: {sidrah.javed, osama.amin, basem.shihada, slim.alouini}@kaust.edu.sa

and techniques. 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.). A proper complex r.v. is thus 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 [1]–[7]. The absence of this property renders non-circular (NC) complex r.v. and the evaluation study to determine the circular and NC nature of complex entities is termed as circularity and non-circularity, respectively.

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. For instance, improper nature of complex modelled entities has been proven for the accumulative additive thermal noise model in communication systems [8], [9], complex envelop of the scalar optical fields [10], empirical speech model [11], complex traces of the seismic signals [12], ocean-current spectra [13], wind fields [14]–[16] and fluid dynamics [17], complex valued model of unbalanced three-phase voltage in power systems [18], [19] and neural activity in brain/spinal cord as measured by functional magnetic resonance imaging (fMRI) [6], [20], [21].

Besides covering the broader aspects of various disciplines, this survey is focused on the communication systems with intentional/unintentional improper signatures. Various discrete modulation schemes that realize the asymmetric complex characteristics include Binary Phase Shift Keying (BPSK) [4], [22], Pulse Amplitude Modulation (PAM) [5], Gaussian Minimum Shift Keying (GMSK) [23], [24], Offset Quadrature Phase Shift Keying (OQPSK) or staggered QPSK [25], single sideband (SSB), vestigial sideband (VSB) [26] and baseband (but not passband) Orthogonal Frequency Division Multiplexing (OFDM) or discrete multitone (DMT) modulation [27]. On the other hand, coding schemes like space-time block coding (STBC) also results in asymmetric signaling [28]. Apart from these naturally occurring improper signals, some phenomenon like transceiver in-phase and quadrature-phase (I/Q) imbalance (IQI) transforms

proper signal to improper signal as well as proper thermal and distortion noise to improper accumulative noise in the wireless communication systems [29].

A. Background

Complete second-order statistical (SOS) characterization of the random signal vector s involves the covariance matrix $\mathrm{E}\left[\mathbf{ss}^{\mathrm{H}}\right]$ as well as pseudo-covariance matrix $\mathrm{E}\left[\mathbf{ss}^{\mathrm{T}}\right]$ analysis besides statistical mean $\mathrm{E}\left[\mathbf{s}\right]$. The traditional complex-valued signal processing assumes a vanishing pseudo-covariance, which is inadequate for many real problems where the signals involved are frequently improper [30]. This superficial assumption not only leads to misleading analysis and incorrect results but also deprives us from seeking the maximum benefit of the additional design freedom offered by improper/asymmetric signaling. Thus, much can be gained if the information contained in the complementary/pseudo covariance is also exploited [7], [31].

Surprisingly, the propriety concept is not unheard-of and vast majority of contributions were put forward in one domain or the other. For instance, improper processes and models have seen growing interest in the statistics community [32]–[34]. Similarly, the pioneering works in signal processing community which emphasized the significance of detailed propriety characterization include [2], [4], [5], [7], [35], [36]. Moreover, the significant contributions from the information theory group comprises of [1], [37]–[43]. Furthermore, few early contributions from the communication circle include [22], [44]–[48].

The naturally existing improper phenomenon and advanced statistical tools supported various studies to exploit the complete SOS characterization of the involved signals. Paradoxically, researchers resort to the more tedious two-dimensional real modeling of the complex variables for their statistical investigation [35] and the complex theory lost most of its beauty, elegance, and interest [2].

B. Motivation

The convenience of modeling multivariate data as a complex entity owing to its compact form, simple notation and fast computation has endorsed it's utility in various applications. Nonetheless, the complete statistical characterization is inevitable to not only exploit the information contained in complementary covariance matrix but also utilize the additional design freedom. Unlike the conventional assumption, complementary covariance function has generally no reason to be equal to zero [3]. This is especially significant when dealing with the concepts of distribution

function and entropy quantification. Circularity is an assumption that was originally introduced for the definition of the probability distribution function [2]. For instance, the probability density function (PDF) of a complex Gaussian random vector (RV) assumes the anticipated and familiar natural form only for proper RVs [1], [4]. Moreover, the definitions of independence and/or correlation are inadequate without complete SOS characterization. For example, it is shown that contrary to the real case, uncorrelated complex normal r.vs. are not generally independent [35]. 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.

The conceptual analysis of properness in various signal processing and information theoretic studies can render significant benefits when extended to communication applications. Initially, people argued the superiority of the proper complex signaling as it maximizes the entropy of a given system [1]. However, the reduced-entropy improper Gaussian signaling (IGS) can be beneficial in unlicensed spectrum-sharing techniques with minimal interference to the legitimate users in underlay [49]–[53], overlay [54] and interweave [55] cognitive radio setups. Followed by the opinion of dominant improper/asymmetric signaling when the system is contaminated by the improper complex interference [56], [57], self-interference [58]–[60], improper noise [8], [45], [61], [62], asymmetric noise/distortions [38], [63], [64] and NC hardware imperfections [9], [65]–[69].

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) [70], [71], broadcast channel (BC) [72]–[74], multiple access channel (MAC) [75], [76], cross-interference channel (X-IC or IC) [77]–[84], one-sided interference channel (Z-IC) [85]–[88], relaying systems [89]–[93] multi-antenna systems [94], [95], multi-cell systems [96], and multi-tier networks [97], etc.

Besides communication systems, general estimation, detection and filtering techniques of circular cases differ from the NC cases. Thus, significant performance gains can be reaped by exploiting circularity/non-circularity in, e.g., wireless transceivers with beamformers or DoA algorithms [98]–[100], source identification [26], [101]–[103], blind source separation [4], [40], [104]–[106], estimation [32], [41], [107]–[111], detection [112]–[114], linear and nonlinear filtering [115]–[117], and Cramer-Rao performance bounds [118], etc.

TABLE I: Summary of Relevant Contributions Surveying Impropriety Concepts

Focal Area	Description	Year	Contributions	Ref
Filtering	Nonlinear adaptive filtering	2009	Focused on the complex nonlinear adaptive filters incorporating non-circularity, widely linear processing and neural models.	[31]
Theoretical Fundamentals	Theory of improper and NC signals	2010	Basics of improper and NC complex-valued signals with comprehensive illustration of the main applications, covering detection, estimation, and signal analysis.	[6]
Interference mitigation	Interference alignment in communication networks	2011	Provides a tutorial while surveying the state-of-the-art interference alignment schemes with a special emphasis on asymmetric complex signal alignment	[119]
Gaussianity Deviations	Complex elliptically symmetric distributions	2012	Comprehensive survey and overview of applications involving complex elliptically symmetric distributions	[120]
Estimation	Statistical Robust Estimation	2012-2018	A tutorial-style treatment of fundamental concepts for robust statistical estimation in signal processing integrating deviations from the distributional assumptions	[121], [122]
	Source identification and separation	2013	Blind identification and separation of general complex-valued signals	[123]
	Optimization and Estimation	2014	Thorough review of the necessary tools in filtering and blind source separation emphasizing the use of the full statistical information in complex-valued optimization and estimation.	[124]
	Parameter Estimation	2015	Presents a meticulous statistical study of complex-valued parameter estimators with a special emphasis on properness	[125]
Detection	MIMO Detection	2015	Exhaustive survey on the fifty years of MIMO detection entailing emerging massive/large-scale MIMO detection schemes and briefly highlighting the importance of appropriate treatment of improper complex-valued systems	[126]

Conclusively, all of these and many other relevant contributions motivate the incorporation of propriety characterization for accurate system analysis and improved system design. The significant performance gains marked by various applications in numerous disciplines helped as a precursor to put forward this extensive survey.

C. Contributions

Various studies discussed the significance of a complete characterization and appropriate treatment of the systems involving improper and NC signals, as summarized in Table I. These studies deal with impropriety concepts that are focused on Gaussianity deviants [120], interference mitigation [119], filtering [31], detection [126], estimation [121], [122], [125], source identification [123] and separation [124]. Furthermore, a comprehensive treatment of theoretical fundamentals of improper and NC signals along with their diverse applications is carried out in [6]. Despite of all these contributions, an exhaustive yet comprehensive survey is required which not only encompasses all these domains but also furnishes complete evolution from the theoretical aspects to the realization ones. Furthermore, this article bridges the interdisciplinary gap between the fields of information theory and signal processing for the wireless communication applications pertaining to signal characterization. The main contributions of this paper in a chronological order include:

- Emphasizing the significance of propriety characterization with three different data representations and their inter-relations.
- Highlighting main differences between the intermingled terms of impropriety and circularity along with the corresponding measures to determine their extent.
- Presenting the generalized entropy and probability distribution functions of the complex Gaussian distributions followed by various tests for impropriety.
- Taxonomy of literature to feature naturally occurring sources of impropriety and their consequences.
- 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.
- Various design guidelines covering suitable optimization tools for the IGS design in addition to the relevant impropriety detection, estimation, filtering and separation procedures.
- Encompassing the journey from theoretical IGS to practical asymmetric discrete signaling and corresponding asymmetric signal recovery methodologies namely equalization, estimation, filtering and detection.
- Error probability (EP) analysis to figure out 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.

- Summarizing 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.

D. Organization and Road map

The paper covers the journey of improper/asymmetric signaling from theory to practical design by first encompassing the preliminaries and overview of the propriety concepts in Section II. Section III encompasses the detailed technical framework for the theoretical analysis. Followed by the brief review of naturally occurring sources of impropriety in various diverse fields in Section IV. Next, Section V covers the theoretical analysis and performance bounds achieved by the IGS in terms of achievable rates, outage probability, power efficiency and degrees of freedom (DoF) in numerous systems. Furthermore, it elaborates the general guidelines for appropriate IGS design algorithms and recovery procedures. The motivation drawn from the theoretical analysis is used as a precursor for practical realization and design of asymmetric discrete constellation in Section VI. This further includes the more revealing metric i.e., error rate analysis to emphasize the superiority of the adopted/advertised signaling or widely linear (WL) processing in several contributions. Later, Section VII classifies numerous applications utilizing the impropriety concepts to further enhance their performance. Subsequently, Section VIII and IX furnish the lessons learned, challenges/limitations and critical future research directions followed by the concluding remarks. The comprehensive outline of the paper organization is detailed in Fig. 1.

E. Notation

Throughout the paper, lower-case letter x denotes variables, while boldface lower-case letter $\mathbf{x}$ and boldface upper-case letter $\mathbf{X}$ denote vectors and matrices, respectively. $\mathbf{0}_N$ and $\mathbf{I}_N$ depict the NxN square null and identity matrix, respectively. The complex-conjugate, inverse conjugate, transpose and conjugate-transpose operators are denoted as $\{.\}^*$, $\{.\}^{-*}$, $\{.\}^{\mathrm{T}}$ and $\{.\}^{\mathrm{H}}$, respectively. The complex augmented random vector (RV) $\underline{\mathbf{x}}$ is defined as $\underline{\mathbf{x}} = [\mathbf{x}^{\mathrm{T}}\ \mathbf{x}^{\mathrm{H}}]^{\mathrm{T}}$. The
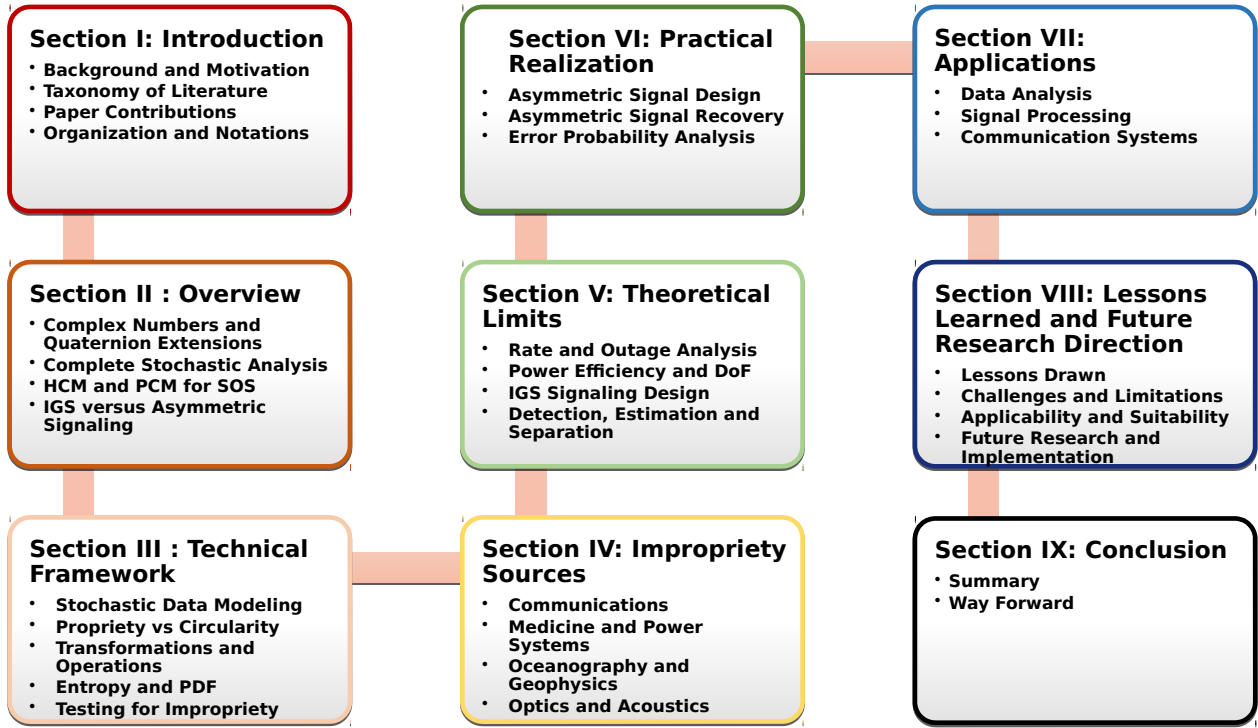
Fig. 1: Paper Organization and Contents

L2-norm and absolute value of a scalar x are represented as $\|\mathbf{x}\|_2$ and $|x|$, respectively. The notation $\mathbf{X} \in \mathbb{F}^{M \times N}$ describes that the $\mathbf{X}$ is a $M \times N$ matrix with entries in $\mathbb{F}$, where $\mathbb{F}$ can be the field of real numbers $\mathbb{R}$, the field of complex numbers $\mathbb{C}$, or the skew-field of quaternion numbers $\mathbb{H}$. Moreover, the trace, determinant, Frobenius norm, and pseudo-inverse of $\mathbf{X}$ are given as $\mathrm{Tr}\,(\mathbf{X})$, $|\mathbf{X}|$, $\|\mathbf{X}\|$ and $\mathbf{X}^\dagger$, respectively. Similarly, for a square invertible matrix $\mathbf{X}$, the inverse and square-root matrix are represented by $\mathbf{X}^{-1}$ and $\mathbf{X}^{1/2}$, respectively. Additionally, $\mathbf{X} \succeq 0$ and $\mathbf{X} \succ 0$ denote that $\mathbf{X}$ is a positive semi-definite (PSD) matrix and positive definite (PD) matrix, respectively. The square block diagonal matrix $\mathbf{J} = \mathbf{Block\text{-}Diag}\,(\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_3)$ comprises of square matrices $\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_3$ of any size as the diagonal elements and the off-diagonal elements are all zero. The subscripts $\{.\}_+$ and $\{.\}_{++}$ are used to define non-negative and strictly positive real numbers of a set, respectively. The expected value operator is expressed as $\mathrm{E}\,[.]$ and the probability of occurrence of an event $\mathbf{\Omega}$ is given by $\mathrm{Pr}\,\{\mathbf{\Omega}\}$. The operation $\odot$ represents a point-wise product. The pure imaginary unit i and the Euler's number e are defined by $i^2 = 1$ and the limit of $(1 + 1/n)^n$ as n approaches infinity, respectively. The term 'iff' is used for the condition 'if and only if' and the notations $=, \triangleq$ and $\equiv$ describe the general

mathematical equality, equality in distribution and equivalence, respectively.

II. Overview

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 [127]. 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 PAM signals in its inphase and quadrature phase components [128]. 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 [129]). 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 [33], [130], [131].

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 [132], computer graphics [133], signal array processing [134], Fourier transforms of images [135], design of orthogonal polarized STBC [136], wave separation [137], wind forecasting [138], nonlinear estimation [139], adaptive filtering [140], [141], and vector sensors [142]. One compelling application is the unified treatment of the relative position and orientation in hand-eye calibration of a robot [143].

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. [144]. Therefore, complex and quaternion r.v. require appropriate and complete characterization to be fully understood and applied.

At the outset of complex theory, the complex stochastic analysis was straightforwardly derived from the real stochastic analysis. First-order statistics were given by the statistical average/mean, whereas SOS were provided by the Hermitian covariance matrix (HCM) in complex analysis as opposed to the symmetric covariance matrix in real analysis. Interestingly, this completes SOS for real analysis but soon realized to be highly insufficient for the complex analysis. HCM accounts for the correlation between different entries of a RV but not the correlation between the two quadrature components within a complex r.v. Moreover, the diagonal elements of HCM render the cumulative variance of the real and imaginary components of the corresponding r.vs but fail to provide the difference of energy between them. Furthermore, the non-diagonal elements of HCM render

- The aggregate covariance between reals and imaginaries of two different r.vs, respectively but does not show the difference between them
- The difference in the cross-covariances which are evaluated between the real of one r.v. and the imaginary of the other r.v. and vice versa. However, it is unsuccessful in declaring the accumulative cross-covariances values.

The lackings mentioned above necessitated adopting another matrix, to complement the missing information in HCM, later named as *pseudo-covariance matrix* (PCM) [1], [5], *conjugate covariance matrix* [6], *complementary covariance matrix* [4], *relation matrix* [35] or *pseudo-scatter matrix* [36].

The incorporation of PCM for the complete characterization is essential to account for the cases when real and imaginary compositions of a r.v. exhibit different variances/powers and/or correlated with each other. Additionally, a null PCM not only refers to uncorrelated components of r.vs but also dictates the correlation between the reals and the imaginaries of two different r.vs in a RV. Moreover, the cross-PCM inclusion is motivated by the fact that the regular definition of two uncorrelated Gaussian RVs requires cross-HCM to be zero. However, two uncorrelated Gaussian RVs may not be independent pertaining to non-zero cross-PCM, which is counterintuitive [6]. Therefore, non-zero PCM must be considered for consonance. Remarkably, a Gaussian RV can be completely described using mean vector, HCM, and PCM, where a null PCM designates proper Gaussian RV and a non-zero PCM dictates an improper Gaussian R.V. A stronger version of propriety is termed as circularity and it accounts for the rotationally invariant PDF and all existing moments. Thus, circularity implies zero-mean and propriety but not vice

versa [7]. However, they are exactly same for a zero-mean Gaussian RV.

This work highlights the importance of SOS characterization for complex analysis of Gaussian RVs and higher order statistics for other RVs. Many real-world problems render improper random signals in diverse fields requiring appropriate analysis and treatment e.g., in medicine [6], [20], [21], [102], [124], [145]–[147], oceanography [13], [148], [149], geology [12], [14]–[16], [103], [150]–[152], optics [10], [153], [154], acoustics [11], [155], [156], power systems [18], [19], [157], [158]. and communication systems [4]–[7], [9], [22]–[26], [28], [32], [48], [67]–[69], [107], [108], [159]–[161]. Focusing on communication systems, IGS with appropriate design and treatment is proven to improve the theoretical bounds of various system configurations. The journey begins with the theoretical IGS analysis and leads to the practical asymmetric signaling scheme for the modern communication systems. Asymmetric signaling is the discrete counterpart of IGS as symmetric signaling is of PGS. This expedition encompasses various forms of asymmetric signaling along with their vast applications in numerous fields.

III. TECHNICAL FRAMEWORK

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.

A. 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 [7], [33], [130]–[134].

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 RVs [6], [7]. Based on these representations, we will introduce the first-order and SOS characteristics of the RVs and their intuitive meanings.

a) 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

$$\mu_{\mathbf{z}} = \mathrm{E}\{\mathbf{z}\} = \mathrm{E}\{\mathbf{x}\} + i\mathrm{E}\{\mathbf{y}\} = \mu_{\mathbf{x}} + i\mu_{\mathbf{y}}, \tag{1}$$

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

$$\mathbf{R_{zz}} = \mathrm{E}\{(\mathbf{z} - \mu_{\mathbf{z}})(\mathbf{z} - \mu_{\mathbf{z}})^{\mathrm{H}}\}$$
$$= \mathbf{R_{xx}} + \mathbf{R_{yy}} + i(\mathbf{R_{xy}^T} - \mathbf{R_{xy}}), \tag{2}$$

where $\mathbf{R_{xx}}$ and $\mathbf{R_{yy}}$ are the auto covariance matrices of the real and imaginary components, respectively and $\mathbf{R_{xy}}$ is the cross covariance matrix between them. Furthermore, $\mathbf{R_{xx}}, \mathbf{R_{yy}}, \mathbf{R_{xy}} \in \mathbb{R}^{N \mathrm{x} N}$ and $\mathbf{R_{zz}} \in \mathbb{C}^{N \mathrm{x} N}$. The covariance matrices in (2) are found from

$$\mathbf{R_{uv}} = \mathrm{E}\{(\mathbf{u} - \mu_{\mathbf{u}})(\mathbf{v} - \mu_{\mathbf{v}})^{\mathrm{T}}\}, \tag{3}$$

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* [1], [5].

$$\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathrm{E}\{(\mathbf{z} - \mu_{\mathbf{z}})(\mathbf{z} - \mu_{\mathbf{z}})^{\mathrm{T}}\}$$
$$= \mathbf{R_{xx}} - \mathbf{R_{yy}} + i(\mathbf{R_{xy}^T} + \mathbf{R_{xy}}). \tag{4}$$

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 $\mathbf{R_{zz}}$, the following three conditions are necessary and sufficient for $\mathbf{R_{zz}}$ and $\tilde{\mathbf{R}}_{\mathbf{zz}}$ to be the covariance and pseudo-covariance matrices of $\mathbf{z}$ [35]

- $\mathbf{R_{zz}}$ is Hermitian and PSD.
- $\tilde{\mathbf{R}}_{\mathbf{zz}}$ is symmetric
- the Schur complement $\mathbf{R_{zz}} - \tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R_{zz}^{-*}}\tilde{\mathbf{R}}_{\mathbf{zz}}^*$ is PSD.

SOS are sufficient for the widely known Gaussian RVs as they can completely characterize their distribution, characteristic function, and higher order moments [35].

Definition 1a. *A multivariate complex Gaussian RV* $\mathbf{z}$ *can be fully characterized using complete SOS properties as* $\mathcal{CN}\left(\mu_{\mathbf{z}}, \mathbf{R_{zz}}, \tilde{\mathbf{R}}_{\mathbf{zz}}\right)$ *[9].*

Consequently, a scalar Gaussian r.v. $z = x + iy$ can be fully described as $\mathcal{CN}(\mu_z, \sigma_z^2, \tilde{\sigma}_z^2)$ [5], [162], 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 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* $\tilde{\mathbf{R}}_{\mathbf{z}_1\mathbf{z}_2} = \mathrm{E}\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_2 - \mu_{\mathbf{z}_2})^{\mathrm{T}}\}$ *[1].*

b) 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 [4], [7], [35].

$$\mu_{\mathbf{u}} = \mathrm{E}\{\mathbf{u}\} = \begin{bmatrix} \mu_{\mathbf{x}} \\ \mu_{\mathbf{y}} \end{bmatrix}, \ \mathbf{R}_{\mathbf{uu}} = \begin{bmatrix} \mathbf{R}_{\mathbf{xx}} & \mathbf{R}_{\mathbf{xy}} \\ \mathbf{R}_{\mathbf{xy}}^{\mathrm{T}} & \mathbf{R}_{\mathbf{yy}}. \end{bmatrix} \tag{5}$$

Definition 1b. *A complex Gaussian RV* $\mathbf{z}$ *with alternate representation* $\mathbf{u}$ *can be fully described as* $\mathcal{N}(\mu_{\mathbf{u}}, \mathbf{R}_{\mathbf{uu}})$ *[75].*

Definition 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 [1].*

c) 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 [4], [7], [35]

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_N & i\mathbf{I}_N \\ \mathbf{I}_N & -i\mathbf{I}_N \end{bmatrix}. \tag{6}$$

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 [35]

$$\mu_{\underline{\mathbf{z}}} = \mathrm{E}\{\underline{\mathbf{z}}\} = \begin{bmatrix} \mu_{\mathbf{x}} + i\mu_{\mathbf{y}} \\ \mu_{\mathbf{x}} - i\mu_{\mathbf{y}} \end{bmatrix}, \ \mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{zz}} & \tilde{\mathbf{R}}_{\mathbf{zz}} \\ \tilde{\mathbf{R}}_{\mathbf{zz}}^* & \mathbf{R}_{\mathbf{zz}}^* \end{bmatrix}. \tag{7}$$

The matrix $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$, which was first used in [39], has the following features [6]

- 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 [4]. $\mathbf{R_{\underline{z}\underline{z}}}$ can be connected with $\mathbf{R_{uu}}$ as

$$\mathbf{R_{\underline{z}\underline{z}}} = \mathrm{E}\{\underline{\mathbf{z}}\underline{\mathbf{z}}^{\mathrm{H}}\} = 2\mathrm{E}\{\mathbf{Tuu}^{\mathrm{H}}\mathbf{T}^{\mathrm{H}}\} = 2\mathbf{TR_{uu}T}^{\mathrm{H}}. \tag{8}$$

Definition 1c. *A complex Gaussian RV* $\mathbf{z}$ *can alternately be fully characterized as* $\mathcal{CN}\left(\mu_{\underline{\mathbf{z}}}, \mathbf{R_{\underline{z}\underline{z}}}\right)$.

Definition 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* $\mathbf{R_{\underline{z}_1\underline{z}_2}} = \mathrm{E}\{(\underline{\mathbf{z}}_1 - \mu_{\underline{\mathbf{z}}_1})(\underline{\mathbf{z}}_2 - \mu_{\underline{\mathbf{z}}_2})^{\mathrm{H}}\}$ *vanishes.*

2) Quaternion Random Vectors: The algebra of quaternions was invented by Sir W. R. Hamilton in 1844 [163] 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 [164]. Analogous to the case of complex numbers, we present three different representations of the quaternion RV.

a) 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 $\mathbf{q}^* = \Re\left(\mathbf{q}\right) - \Im\left(\mathbf{q}\right) = \mathbf{r}_1 - i\mathbf{r}_i - j\mathbf{r}_j - k\mathbf{r}_k$ and the involution of $\mathbf{q}$ over a pure unit quaternion $\alpha \in \{i,j,k\}$ is defined as $\mathbf{q}^{(\alpha)} = -\alpha\mathbf{q}\alpha$ [43], [165]. 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 $\mathbf{R_{qq}} = \mathrm{E}\{\mathbf{qq}^{\mathrm{H}}\}$ and three pseudo-covariance matrices $\mathbf{R_{qq^{(i)}}}, \mathbf{R_{qq^{(j)}}}$ and $\mathbf{R_{qq^{(k)}}}$. These pseudo-covariance matrices quantify the correlation between $\mathbf{q}$ and its involutions $\mathbf{q}^{(\alpha)}$ over three pure unit quaternions [166]

$$\mathbf{R_{qq^{(\alpha)}}} = \mathrm{E}\{\mathbf{qq}^{(\alpha)\mathrm{H}}\} = -\mathrm{E}\{\mathbf{q}\alpha\mathbf{q}^{\mathrm{H}}\alpha\}. \tag{9}$$

Definition 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* $\mathcal{QN}\left(\mathbf{R_{qq}}, \mathbf{R_{qq^{(i)}}}, \mathbf{R_{qq^{(j)}}}, \mathbf{R_{qq^{(k)}}}\right)$.

b) 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 [166]

$$
\mathbf{R_{vv}} = \begin{bmatrix}
\mathbf{R_{r_1 r_1}} & \mathbf{R_{r_1 r_i}} & \mathbf{R_{r_1 r_j}} & \mathbf{R_{r_1 r_k}} \\
\mathbf{R_{r_i r_1}} & \mathbf{R_{r_i r_i}} & \mathbf{R_{r_i r_j}} & \mathbf{R_{r_i r_k}} \\
\mathbf{R_{r_j r_1}} & \mathbf{R_{r_j r_i}} & \mathbf{R_{r_j r_j}} & \mathbf{R_{r_j r_k}} \\
\mathbf{R_{r_k r_1}} & \mathbf{R_{r_k r_i}} & \mathbf{R_{r_k r_j}} & \mathbf{R_{r_k r_k}}
\end{bmatrix}, \tag{10}
$$

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

Definition 3b. *A zero-mean quaternion Gaussian RV $\mathbf{q}$ with quadrivariate real composite representation $\mathbf{v}$ is completely characterized using the symmetric correlation matrix $\mathbf{R_{vv}}$.*

c) Augmented Representation: Let $\underline{\mathbf{q}}$ denotes the augmented representation of $\mathbf{q}$ and its involutions as

$$
\underline{\mathbf{q}} = \begin{bmatrix}
\mathbf{q} \\
\mathbf{q}^{(i)} \\
\mathbf{q}^{(j)} \\
\mathbf{q}^{(k)}
\end{bmatrix} = \begin{bmatrix}
\mathbf{r}_1 + i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k \\
\mathbf{r}_1 + i\mathbf{r}_i - j\mathbf{r}_j - k\mathbf{r}_k \\
\mathbf{r}_1 - i\mathbf{r}_i + j\mathbf{r}_j - k\mathbf{r}_k \\
\mathbf{r}_1 - i\mathbf{r}_i - j\mathbf{r}_j + k\mathbf{r}_k
\end{bmatrix}. \tag{11}
$$

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 [167]

$$
\mathbf{A}_N = \frac{1}{2} \begin{bmatrix}
\mathbf{I}_N & i\mathbf{I}_N & j\mathbf{I}_N & k\mathbf{I}_N \\
\mathbf{I}_N & i\mathbf{I}_N & -j\mathbf{I}_N & -k\mathbf{I}_N \\
\mathbf{I}_N & -i\mathbf{I}_N & j\mathbf{I}_N & -k\mathbf{I}_N \\
\mathbf{I}_N & -i\mathbf{I}_N & -j\mathbf{I}_N & k\mathbf{I}_N
\end{bmatrix}. \tag{12}
$$

The SOS characterization of $\underline{\mathbf{q}}$ is given by the following augmented covariance matrix [168]

$$
\mathbf{R_{\underline{q}\underline{q}}} = \begin{bmatrix}
\mathbf{R_{qq}} & \mathbf{R_{qq^{(i)}}} & \mathbf{R_{qq^{(j)}}} & \mathbf{R_{qq^{(k)}}} \\
\mathbf{R_{qq^{(i)}}^{(i)}} & \mathbf{R_{qq}^{(i)}} & \mathbf{R_{qq^{(k)}}^{(i)}} & \mathbf{R_{qq^{(j)}}^{(i)}} \\
\mathbf{R_{qq^{(j)}}^{(j)}} & \mathbf{R_{qq^{(k)}}^{(j)}} & \mathbf{R_{qq}^{(j)}} & \mathbf{R_{qq^{(i)}}^{(j)}} \\
\mathbf{R_{qq^{(k)}}^{(k)}} & \mathbf{R_{qq^{(j)}}^{(k)}} & \mathbf{R_{qq^{(i)}}^{(k)}} & \mathbf{R_{qq}^{(k)}}
\end{bmatrix}, \tag{13}
$$

where $\mathbf{R_{qq^{(\alpha)}}^{(\beta)}} = -\beta \mathbf{R_{qq^{(\alpha)}}} \beta$ with $\beta \in \{i, j, k\}$. The correlation matrix $\mathbf{R_{qq^{(\alpha)}}^{(\alpha)}} = \mathbf{R_{qq^{(\alpha)}}^{(\alpha)\mathrm{H}}}$ 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 [167]. $\mathbf{R_{\underline{q}\underline{q}}}$ is linked with $\mathbf{R_{vv}}$ as

$$
\mathbf{R_{\underline{q}\underline{q}}} = \mathrm{E}\{\underline{\mathbf{q}}\underline{\mathbf{q}}^{\mathrm{H}}\} = 4\mathrm{E}\{\mathbf{A}_N \mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{A}_N^{\mathrm{H}}\} = 4\mathbf{A}_N \mathbf{R_{vv}}\mathbf{A}_N^{\mathrm{H}}. \tag{14}
$$

Definition 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 $\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2}^{(\alpha)} = -\alpha\mathbf{R}_{\mathbf{q}_1\mathbf{q}_2}\alpha$ along with all possible combinations of cross pseudo-covariance matrices $\mathrm{E}\{\mathbf{q}_1^{(\gamma)}\mathbf{q}_2^{(\varsigma)\mathrm{H}}\}$ 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.

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 [35]. 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 $\mathbf{R}_{\underline{\mathbf{z}}_1\underline{\mathbf{z}}_2} = \mathbf{0}$ [6]. 2) Pertaining to the block structure of the $\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}$, interestingly it can be invertible even when $\mathbf{R}_{\mathbf{z}\mathbf{z}}$ and $\tilde{\mathbf{R}}_{\mathbf{z}\mathbf{z}}$ are not [4]. 3) It is a powerful tool for WL transformations (refer to Section III-C) [7]. 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 [144]. 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) [167].

B. 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 [2], [4], [5] and information theory [1]. It has also been extended to quaternion valued r.v. [144] and vectors [43], [167] 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.

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.

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

Definition 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* $\tilde{\mathbf{R}}_{\mathbf{zz}}$ *[1]. Properness may also be seen as the uncorrelation of* z *with its complex conjugate* z^* *[3], [109]. These equivalent propriety definitions are sometimes referred as Strict Propriety [4].*

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 $k \neq l$ [1]. 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 $\tilde{\mathbf{R}}_{\underline{\mathbf{zz}}}$. However, proper $\underline{\mathbf{z}}$ does not imply vanishing $\tilde{\mathbf{R}}_{\underline{\mathbf{zz}}}$ rather it implies a block diagonal structure of $\mathbf{R}_{\underline{\mathbf{zz}}} = \mathbf{Block\text{-}Diag}\,(\mathbf{R_{zz}}, \mathbf{R_{zz}^*})$. 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 $\mathbf{R_{zz}}$) and have even multiplicity (paired because of diagonal blocks $\mathbf{R_{zz}}$ and $\mathbf{R_{zz}^*}$). 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 5 (Generalized Propriety). *A complex RV is termed as generalized proper if all eigenvalues of* $\mathbf{R}_{\underline{\mathbf{zz}}}$ *have even multiplicity [4].*

The EVD of $\mathbf{R}_{\underline{\mathbf{zz}}}$ takes on the form $\mathbf{R}_{\underline{\mathbf{zz}}} = \mathbf{V}\left(\mathbf{T\Lambda T}^{\mathrm{H}}\right)\mathbf{V}^{\mathrm{H}}$ with diagonal $\mathbf{T\Lambda T}^{\mathrm{H}}$ iff $\mathbf{z}$ is generalized proper, otherwise block matrix $\mathbf{T\Lambda T}^{\mathrm{H}}$ with diagonal blocks [4]. Moreover, the complex RVs $\mathbf{z}_1$ and $\mathbf{z}_2$ are *cross proper* iff the cross pseudo-covariance matrix $\tilde{\mathbf{R}}_{\mathbf{z}_1\mathbf{z}_2}$ vanishes. Notably, they are *jointly proper* if the composite RV having $\mathbf{z}_1$ and $\mathbf{z}_2$ as subvectors is proper

[1] or if they are proper and cross proper [43].

Definition 6 (Improper RV). *A complex RV $\mathbf{z}=\mathbf{x}+i\mathbf{y}$ is called improper if any of the following statements holds [169]*

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

Absence of properness of a RV is termed as *improperness* and the extent of improperness is argued by the eigenvalue spread of $\mathbf{R_{\underline{zz}}}$ [169]. For instance, a RV $\mathbf{z}_1$ is *less improper* than another RV $\mathbf{z}_2$ if the eigenvalues of $\mathbf{R_{\underline{z}_1\underline{z}_1}}$ are majorized by (less spread out) those of $\mathbf{R_{\underline{z}_2\underline{z}_2}}$ [32].

Definition 7 (Maximally Improper RV). *For a given $\mathbf{R_{zz}}$, the vector whose $\mathbf{R_{\underline{zz}}}$ has least eigenvalue spread must be proper $\tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$ whereas the vector whose $\mathbf{R_{\underline{zz}}}$ has maximum possible eigenvalue spread must be maximally improper [6].*

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

Definition 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 [3], [7].*

For a scalar z, is equivalent to spherical symmetry of the corresponding real composite vector $u = [x\ y]^{\mathrm{T}}$ [170]. 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 [2], [151]:

- *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 $\mathbf{a}$

Fig. 2: Extents of Circularity and their Implications

- *Total Circular:* If its components are independent and circular. The circularity characteristics are related to each other as shown in Fig. 2.

    c) 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 $\rightleftharpoons$ Propriety, Impropriety $\rightleftharpoons$ 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:

$$\tilde{\mathbf{R}}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} = \mathrm{E}\left[\hat{\mathbf{z}}\hat{\mathbf{z}}^{\mathrm{T}}\right] = e^{j2\alpha}\tilde{\mathbf{R}}_{\mathbf{zz}} \quad \forall\alpha. \tag{15}$$

    Thus, (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 PDF and thus all existing moments to be rotationally invariant. Thus, circularity implies zero mean and propriety, but not vice versa [7].
- In the light of the aforementioned arguments, an improper RV with rotationally variant second moment is essentially 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.
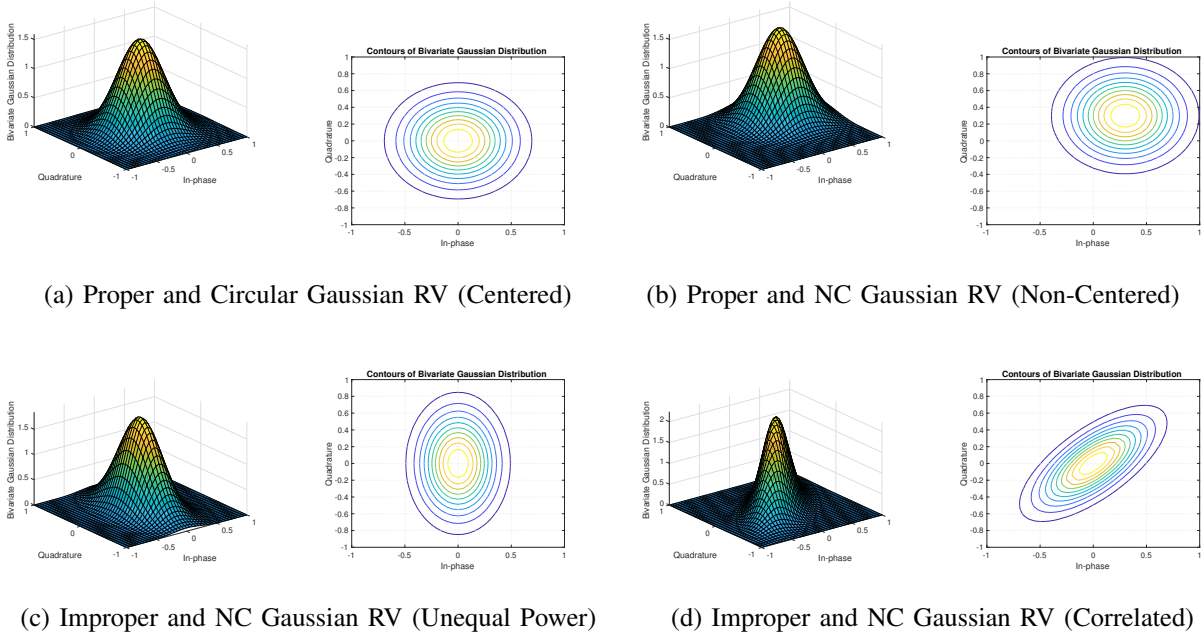
(a) Proper and Circular Gaussian RV (Centered)

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

(c) Improper and NC Gaussian RV (Unequal Power)

(d) Improper and NC Gaussian RV (Correlated)

Fig. 3: 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 Fig. 3a, whereas the same distribution with non-zero mean is proper but not circular in Fig. 3b. Furthermore, the deviation from equal power distribution and uncorrelation result in improper as well as NC complex r.v. as shown in Fig. 3c and Fig. 3d, respectively. The quadrature component has more power/variance than the in-phase component in Fig. 3c and non-perpendicular distribution contours to x- or y-axis depict the correlation between I/Q components in Fig. 3d.

For other zero-mean RVs propriety and circularity are related as shown in Fig. 4. 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 $\mathbf{R_{\underline{zz}}} = \mathbf{Block\text{-}Diag}\left(\mathbf{R_{zz}}, \mathbf{R_{zz}^*}\right)$ 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 6 leads to improper RV. An improper RV with maximum eigenvalue spread of $\mathbf{R_{\underline{zz}}}$ is termed as maximal improper RV. Moreover, improperness always implies non-circularly as it nullifies the condition of rotational invariant second-moment.
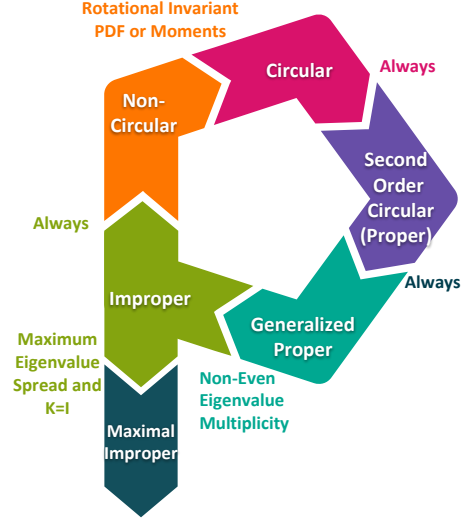
Fig. 4: Extents of Propriety and their Implications

d) 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 9. *Circularity quotient ρ_z is defined as the fraction between the pseudo-variance and the variance*

$$\rho_z = \frac{\tilde{\sigma}_z^2}{\sigma_z^2} = ke^{i\phi}, \tag{16}$$

where, DoI is measured by the circularity coefficient $k = |\tilde{\sigma}_z^2|/\sigma_z^2$ and the circularity angle is given by ϕ [5].

The term *circularity coefficient* for k is originated from [40] while the terms *non-circularity rate* and *non-circularity phase* are also used for k and ϕ, respectively [171]. In reality, the circularity coefficient is the *canonical correlation* between z and z^* [172]. 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 [5]. 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) [173]. This quantifies DoI for a r.v., next we present DoI of a RV as:

Definition 10. *For a complex RV* $\mathbf{z}$*, the coherence matrix* $\mathbf{C} = \mathbf{R}_{\mathbf{zz}}^{-1/2}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-T/2}$ *with Takagi factorization* $\mathbf{C} = \mathbf{FKF}^{T}$ *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 [169]*

$$\rho_1 = 1 - \prod_{i=1}^{r} \left(1 - k_i^2\right) \overset{(r=N)}{=} 1 - |\mathbf{R}_{\underline{\mathbf{zz}}}| \, |\mathbf{R}_{\mathbf{zz}}|^{-2}, \tag{17}$$

$$\rho_2 = \prod_{i=1}^{r} k_i^2 \overset{(r=N)}{=} \left|\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-*}\tilde{\mathbf{R}}_{\mathbf{zz}}^{*}\right| |\mathbf{R}_{\mathbf{zz}}|^{-1}, \tag{18}$$

$$\rho_3 = \frac{1}{N}\sum_{i=1}^{r} k_i^2 \overset{(r=N)}{=} \frac{1}{N}\text{Tr}\left(\mathbf{R}_{\mathbf{zz}}^{-1}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-*}\tilde{\mathbf{R}}_{\mathbf{zz}}^{*}\right), \tag{19}$$

where $r = \text{rank}(\mathbf{R}_{\underline{\mathbf{zz}}})$. The set of canonical correlations $\{k_i\}_{i=1}^{N}$ is also referred to as *circularity spectrum* [40]. DoI must satisfy $0 \leq \rho_i \leq 1$, ranging from proper signal ($\rho_i = 0$) to maximally improper signal ($\rho_i = 1$) [5]. 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 [7] (refer to Section III-D) and it is also used for impropriety likelihood-ratio testing [172] (refer to Section III-E) 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.

e) 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) Quaternion Random Vectors: Impropriety characterization is not limited to the complex domain. For instance, Vakhania studied the concept of properness in quaternion domain [174]; 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 [144]. The evolved and refined definitions of quaternions propriety rely on the vanishing properties of the pseudo-covariance matrices as:

Definition 11 ($\mathcal{R}^\alpha$-properness)**.** *A quaternion RV* $\mathbf{q}$ *is* $\mathcal{R}^\alpha$-*proper iff the pseudo-covariance matrix* $\mathbf{R}_{\mathbf{qq}^{(\alpha)}}$ *vanishes [43].*

Definition 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.,* $\mathbf{q}^{(\bar{\alpha})}$ *[167].*

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 [43]. Equivalently, $\mathcal{C}^j$-proper quaternion exhibits the following $\mathbf{R}_{\mathbf{vv}}$ structure [166]

$$\mathbf{R}_{\mathbf{vv}} = \begin{bmatrix} \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_i} \\ \mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_i} \end{bmatrix}, \tag{20}$$

Proof. Evidently, (20) is obtained from (10) by exploiting the definitions of vanishing $\mathbf{R}_{\mathbf{q},\mathbf{q}^{(i)}}$

TABLE II: Consequences of Vanishing Pseudo-Variances

$\mathbf{R_{q,q^{(i)}}} = 0$	$\mathbf{R_{q,q^{(j)}}} = 0$	$\mathbf{R_{q,q^{(k)}}} = 0$
$\mathbf{R_{r_1 r_1}} + \mathbf{R_{r_i r_i}} = \mathbf{R_{r_j r_j}} + \mathbf{R_{r_k r_k}}$	$\mathbf{R_{r_1 r_1}} + \mathbf{R_{r_j r_j}} = \mathbf{R_{r_i r_i}} + \mathbf{R_{r_k r_k}}$	$\mathbf{R_{r_1 r_1}} + \mathbf{R_{r_k r_k}} = \mathbf{R_{r_i r_i}} + \mathbf{R_{r_j r_j}}$
$\mathbf{R_{r_i r_1}} - \mathbf{R_{r_1 r_i}} = \mathbf{R_{r_k r_j}} - \mathbf{R_{r_j r_k}}$	$\mathbf{R_{r_1 r_i}} + \mathbf{R_{r_i r_1}} = -\mathbf{R_{r_k r_j}} - \mathbf{R_{r_j r_k}}$	$\mathbf{R_{r_i r_1}} + \mathbf{R_{r_1 r_i}} = \mathbf{R_{r_k r_j}} + \mathbf{R_{r_j r_k}}$
$\mathbf{R_{r_1 r_j}} + \mathbf{R_{r_j r_1}} = \mathbf{R_{r_i r_k}} + \mathbf{R_{r_k r_i}}$	$\mathbf{R_{r_j r_1}} - \mathbf{R_{r_1 r_j}} = \mathbf{R_{r_i r_k}} - \mathbf{R_{r_k r_i}}$	$\mathbf{R_{r_1 r_j}} + \mathbf{R_{r_j r_1}} = -\mathbf{R_{r_i r_k}} - \mathbf{R_{r_k r_i}}$
$\mathbf{R_{r_1 r_k}} + \mathbf{R_{r_k r_1}} = -\mathbf{R_{r_i r_j}} - \mathbf{R_{r_j r_i}}$	$\mathbf{R_{r_1 r_k}} + \mathbf{R_{r_k r_1}} = \mathbf{R_{r_i r_j}} + \mathbf{R_{r_j r_i}}$	$\mathbf{R_{r_1 r_k}} - \mathbf{R_{r_k r_1}} = \mathbf{R_{r_i r_j}} - \mathbf{R_{r_j r_i}}$

and $\mathbf{R_{q,q^{(k)}}}$, which are expressed as

$$\mathbf{R_{q,q^{(i)}}} = -\mathrm{E}\{(\mathbf{r}_1 + i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k)(\mathbf{r}_1^\mathrm{T} - i\mathbf{r}_i^\mathrm{T} + j\mathbf{r}_j^\mathrm{T} + k\mathbf{r}_k^\mathrm{T})\}.$$

$$\mathbf{R_{q,q^{(k)}}} = -\mathrm{E}\{(\mathbf{r}_1 + i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k)(\mathbf{r}_1^\mathrm{T} + i\mathbf{r}_i^\mathrm{T} + j\mathbf{r}_j^\mathrm{T} - k\mathbf{r}_k^\mathrm{T})\}. \tag{21}$$

Equating these pseudo-variances to zero render the set of four real-valued equations as shown in Table II. In case of $\mathcal{C}^j$-proper quaternion, the four set of equations each from $\mathbf{R_{q,q^{(i)}}}$ and $\mathbf{R_{q,q^{(k)}}}$ are simultaneously solved. The solution implies $\mathbf{R_{r_j r_j}} = \mathbf{R_{r_1 r_1}}$, $\mathbf{R_{r_k r_k}} = \mathbf{R_{r_i r_i}}$, $\mathbf{R_{r_k r_j}} = \mathbf{R_{r_i r_1}}$, $\mathbf{R_{r_j r_k}} = \mathbf{R_{r_1 r_i}}$, $\mathbf{R_{r_1 r_j}} = -\mathbf{R_{r_j r_1}}$, $\mathbf{R_{r_i r_k}} = -\mathbf{R_{r_k r_i}}$, $\mathbf{R_{r_1 r_k}} = -\mathbf{R_{r_j r_i}}$, and $\mathbf{R_{r_i r_j}} = -\mathbf{R_{r_k r_1}}$. Consequently, substituting these relations in (10) yields $\mathbf{R_{vv}}$ in (20). $\qquad\square$

Definition 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 [144].*

Definition 13 ($\mathcal{Q}$-properness)**.** *A quaternion RV $\mathbf{q}$ is $\mathcal{Q}$-proper iff all three pseudo-covariance matrices $\mathbf{R_{qq^{(i)}}}$, $\mathbf{R_{qq^{(j)}}}$ and $\mathbf{R_{qq^{(k)}}}$ vanish [43]. It also implies that $\mathbf{q}$ is uncorrelated with its three vector involutions. Moreover, the corresponding $\mathbf{R_{\underline{qq}}}$ is real-valued, positive definite, and symmetric [167].*

$\mathcal{Q}$-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 [144]. Analogous to the $\mathcal{C}^j$-proper case, the equivalent $\mathbf{R_{vv}}$ for $\mathcal{Q}$-proper quaternion is obtained by simultaneously solving the 12 set of equations (as given in Table II) obtained by setting $\mathbf{R_{qq^{(i)}}}$, $\mathbf{R_{qq^{(j)}}}$ and

$\mathbf{R}_{\mathbf{qq}^{(k)}}$ to zero [166].

$$\mathbf{R}_{\mathbf{vv}} = \begin{bmatrix} \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} \\ \mathbf{R}_{\mathbf{r}_i\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} \\ \mathbf{R}_{\mathbf{r}_j\mathbf{r}_1} & \mathbf{R}_{\mathbf{r}_j\mathbf{r}_i} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} \\ \mathbf{R}_{\mathbf{r}_k\mathbf{r}_1} & -\mathbf{R}_{\mathbf{r}_i\mathbf{r}_k} & \mathbf{R}_{\mathbf{r}_k\mathbf{r}_j} & \mathbf{R}_{\mathbf{r}_1\mathbf{r}_1} \end{bmatrix}. \tag{22}$$

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) [166].

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 [43].

a) 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 $\mathbf{R}_{\underline{\mathbf{qq}}}$ is evaluated as

$$\mathcal{P}_\Pi = \min_{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} \in \mathfrak{R}_\Pi} D\left(\mathbf{R}_{\underline{\mathbf{qq}}} \parallel \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}\right), \tag{23}$$

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

$$\mathfrak{R}_\mathcal{Q} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(i)}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(j)}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(k)}} = \mathbf{0}\}, \tag{24}$$

$$\mathfrak{R}_{\mathcal{C}^\alpha} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha')}} = \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha'')}} = \mathbf{0}\}, \tag{25}$$

$$\mathfrak{R}_{\mathcal{R}^\alpha} = \{\hat{\mathbf{R}}_{\underline{\mathbf{qq}}} | \hat{\mathbf{R}}_{\underline{\mathbf{qq}}^{(\alpha)}} = \mathbf{0}\}. \tag{26}$$

Moreover, $D\left(\mathbf{R}_{\underline{\mathbf{qq}}} \parallel \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}\right)$ is the Kullback-Leibler divergence [175] between two zero-mean quaternion Gaussian distributions with $\mathbf{R}_{\underline{\mathbf{qq}}}$ and $\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}$ [43, 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 V [43].

b) 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 [176].

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 $k \neq l$ [1] 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 propriety in both complex and quaternion RVs. Additionally, 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 [168]. Moreover, the complex domain DoI is also useful in blind source separation [6] 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.

C. Transformations and Operations

Impropriety classification helps to identify the simplified form of processing and transformation in terms of computational complexity [138]. This subsection highlights suitable processing models for complex as well as quaternion RVs based on their propriety characterization.

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' impropriety and extract the information embedded in them.

a) Transformations: Three forms of transformations, i.e., *real linear transformation* (RLT) or WL *transformation* (WLT), *complex linear transformation* (CLT) and *widely unitary trans-*

formation (WUT) are reviewed as:

- RLT $\mathbf{L} \in \mathbb{R}^{2M \times 2N}$ 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

$$\mathbf{u}_{\text{RLT}} = \begin{bmatrix} \mathbf{x}_{\text{RLT}} \\ \mathbf{y}_{\text{RLT}} \end{bmatrix} = \mathbf{L}\mathbf{u} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \tag{27}$$

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

$$\underline{\mathbf{w}} = \begin{bmatrix} \mathbf{w} \\ \mathbf{w}^* \end{bmatrix} = \sqrt{2}\mathbf{T}\mathbf{u}_{\text{RLT}} = \sqrt{2}\mathbf{T}\mathbf{L}\mathbf{u}$$

$$= \{\mathbf{T}\mathbf{L}\mathbf{T}^{\mathrm{H}}\}\{\sqrt{2}\mathbf{T}\mathbf{u}\} = \underline{\mathbf{N}}\mathbf{z}, \tag{28}$$

where,

$$\underline{\mathbf{N}} = \begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_2 \\ \mathbf{N}_2^* & \mathbf{N}_1^* \end{bmatrix}. \tag{29}$$

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

$$\mathbf{N}_1 = \frac{1}{2} \left[\mathbf{L}_{11} + \mathbf{L}_{22} + i \left(\mathbf{L}_{21} - \mathbf{L}_{12} \right) \right], \tag{30}$$

$$\mathbf{N}_2 = \frac{1}{2} \left[\mathbf{L}_{11} - \mathbf{L}_{22} + i \left(\mathbf{L}_{21} + \mathbf{L}_{12} \right) \right]. \tag{31}$$

- 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

$$\mathbf{z}_{\text{CLT}} = \begin{bmatrix} \mathbf{x}_{\text{CLT}} \\ \mathbf{y}_{\text{CLT}} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ -\mathbf{L}_{12} & \mathbf{L}_{11} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \tag{32}$$

To summarize, RLT on $\mathbb{R}^{2N}$ are linear on $\mathbb{C}^N$ if they have the matrix structure as in (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 [6].

- Interestingly, the EVD of $\mathbf{R}_{\mathbf{zz}}$ is different from EVD of ordinary matrices and takes the unfamiliar form $\mathbf{R}_{\mathbf{zz}} = \underline{\mathbf{U}}\Lambda\underline{\mathbf{U}}^{\mathrm{H}}$ 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 [4], [169]

$$\underline{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_2^* & \mathbf{U}_1^* \end{bmatrix}. \tag{33}$$

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

$$\underline{\Lambda} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{\mathrm{H}} = \frac{1}{2} \begin{bmatrix} \Lambda^{(1)} + \Lambda^{(2)} & \Lambda^{(1)} - \Lambda^{(2)} \\ \Lambda^{(1)} - \Lambda^{(2)} & \Lambda^{(1)} + \Lambda^{(2)} \end{bmatrix}, \tag{34}$$

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

Definition 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* $\hat{\mathbf{A}} \in \mathbb{C}^{M\times N}$ *and* $\hat{\mathbf{b}} \in \mathbb{C}^M$ *are constant, is also proper. Thus, properness is preserved under affine transformations [1]. On contrary, the generalized propriety is also preserved by unitary WLT and arbitrary SLT [4]*

b) 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

$$\mathbf{u}_1^{\mathrm{T}}\mathbf{u}_2 = \frac{1}{2}\underline{\mathbf{z}}_1^{\mathrm{H}}\mathbf{T}\mathbf{T}^{\mathrm{H}}\underline{\mathbf{z}}_2 = \frac{1}{2}\underline{\mathbf{z}}_1^{\mathrm{H}}\underline{\mathbf{z}}_2 = \Re\{\mathbf{z}_1^{\mathrm{H}}\mathbf{z}_2\}. \tag{35}$$

- Moreover, the real-valued quadratic operation can be described using different representations as [7], [35].

$$\mathbf{u}^{\mathrm{T}}\mathbf{L}\mathbf{u} = \frac{1}{2}\underline{\mathbf{z}}^{\mathrm{H}}\underline{\mathbf{N}}\underline{\mathbf{z}} = \Re\{\mathbf{z}^{\mathrm{H}}\mathbf{N}_1\mathbf{z} + \mathbf{z}^{\mathrm{H}}\mathbf{N}_2\mathbf{z}^*\}. \tag{36}$$

c) Discussion: Remarkably, the maximal invariants of $\mathbf{R}_{\underline{\mathbf{zz}}}$ depend on the underlying transformations. Meaning thereby, any function of $\mathbf{R}_{\underline{\mathbf{zz}}}$ that is invariant under particular transformation must be a function of these maximal invariants only. For example, the eigenvalues of $\mathbf{R}_{\underline{\mathbf{zz}}}$ constitute a maximal invariant for $\mathbf{R}_{\underline{\mathbf{zz}}}$ under WUT whereas the circularity coefficients constitute a maximal invariant for $\mathbf{R}_{\underline{\mathbf{zz}}}$ under non-singular SLT [6].

2) Quaternion Random Vectors: The type of quaternion properness defines several types of linear processing [168]

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

$$\mathbf{q}_{\mathrm{f}} = \mathbf{F}_1^{\mathrm{H}}\mathbf{q} + \mathbf{F}_i^{\mathrm{H}}\mathbf{q}^{(i)} + \mathbf{F}_j^{\mathrm{H}}\mathbf{q}^{(j)} + \mathbf{F}_k^{\mathrm{H}}\mathbf{q}^{(k)}, \tag{37}$$

where, $\mathbf{q}_f \in \mathbb{H}^{P\text{x}1}$ and $\mathbf{F}_1^H, \mathbf{F}_i^H, \mathbf{F}_j^H, \mathbf{F}_k^H \in \mathbb{H}^{P\text{x}N}$. Full-WLT for augmented representation takes the form $\underline{\mathbf{q}}_f = \mathbf{F}_{\underline{\mathbf{q}}}^H \underline{\mathbf{q}}$ with $\mathbf{F}_{\underline{\mathbf{q}}} = \begin{bmatrix} \mathbf{F}_1^T & \mathbf{F}_i^T & \mathbf{F}_j^T & \mathbf{F}_k^T \end{bmatrix}^T$, whereas for the real composite representation it becomes $\mathbf{v}_f = \mathbf{F}_{\mathbf{v}}^T \mathbf{v}$ with $\mathbf{F}_{\mathbf{v}} = \mathbf{A}_N^H \mathbf{F}_{\underline{\mathbf{q}}} \mathbf{A}_P$.

- *Semi-widely* linear processing is for the $\mathcal{C}^\alpha$-proper vectors and takes the form

$$\mathbf{q}_s = \mathbf{F}_1^H \mathbf{q} + \mathbf{F}_\alpha^H \mathbf{q}^{(\alpha)}. \tag{38}$$

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

Notably, $\mathcal{R}^\alpha$-proper quaternion RVs do not result in a simplified linear processing model [168]. A $\mathcal{Q}$-proper RV is invariant under affine transformation [167] whereas the $\mathcal{C}^\alpha$-properness is invariant under semi-WLT [43].

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 [138], [140].

D. 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.

1) Complex Random Vectors: This part contains the definitions of differential entropy and probability distributions for a general complex Gaussian RV.

a) 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 [4]

$$H(\mathbf{u}) = \frac{1}{2} \log \left[(2\pi e)^{2N} \left| \mathbf{R_{uu}} \right| \right]. \tag{39}$$

Equivalently, the differential entropy of $\mathbf{z}$ in terms of the augmented covariance matrix $\mathbf{R_{\underline{z}\underline{z}}}$ is [7]

$$H(\mathbf{z}) = \frac{1}{2} \log \left[(\pi e)^{2N} \left| \mathbf{R_{\underline{z}\underline{z}}} \right| \right]. \tag{40}$$

Definition 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 [1], where it is given by*

$$H_P(\mathbf{z}) = \log \left[(\pi e)^N \left| \mathbf{R_{zz}} \right| \right]. \tag{41}$$

Remark: For a scalar, the differential entropy is appropriately defined as the joint differential entropy of its real and imaginary parts [1]. 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 [7]

$$H_I(\mathbf{z}) = H_P(\mathbf{z}) - \mathcal{I}(\mathbf{z}; \mathbf{z}^*), \tag{42}$$

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 (17).

b) 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 [39]. 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

$$p(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{2N} \left| \mathbf{R_{uu}} \right|}} \exp \left\{ -\frac{1}{2} (\mathbf{u} - \mu_{\mathbf{u}})^{\mathrm{T}} R_{\mathbf{uu}}^{-1} (\mathbf{u} - \mu_{\mathbf{u}}) \right\}. \tag{43}$$

Equivalently, the PDF of a complex RV $\mathbf{z}$ with augmented covariance matrix $\mathbf{R_{\underline{z}\underline{z}}}$ is written as [39]

$$p(\mathbf{z}) = \frac{1}{\sqrt{\pi^{2N} \left| \mathbf{R_{\underline{z}\underline{z}}} \right|}} \exp \left\{ -\frac{1}{2} (\underline{\mathbf{z}} - \mu_{\underline{\mathbf{z}}})^{\mathrm{H}} \mathbf{R}_{\underline{z}\underline{z}}^{-1} (\underline{\mathbf{z}} - \mu_{\underline{\mathbf{z}}}) \right\}. \tag{44}$$

2) Quaternion Random Vector: This subsection contains the definitions of differential entropy and probability distributions for a general multivariate quaternion RV.

a) Differential entropy: The differential entropy of a quaternion RV $\mathbf{q}$ with real represen-
tation $\mathbf{v}$ and the respective covariance matrix $\mathbf{R_{vv}}$ is given by [43]

$$H_{\mathbf{q}}\left(\mathbf{R_{vv}}\right) = 2N \ln\left(2\pi e\right) + \frac{1}{2} \ln |\mathbf{R_{vv}}|. \tag{45}$$

Equivalently, the augmented representation renders [167]

$$H_{\mathbf{q}}\left(\mathbf{R_{\underline{q}\underline{q}}}\right) = 2N \ln\left(\frac{\pi e}{2}\right) + \frac{1}{2} \ln |\mathbf{R_{\underline{q}\underline{q}}}|. \tag{46}$$

Definition 16. *The differential entropy of any quaternion RV is upper bounded by the differential
entropy of a centered $\mathcal{Q}$-proper Gaussian RV with $\mathbf{R_{vv}} = \sigma^2 \mathbf{I}_{4N}$ and is given by [167]*

$$H_{\mathbf{q}} \leqslant H_{\mathcal{Q}-\text{proper}} = 2N \ln\left(2\pi e \sigma^2\right). \tag{47}$$

Beyond the mutual information between two complex RVs, its generalization to higher dimen-
sions is termed as *interaction information* [177]. 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 [167, 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 $\mathcal{P}_{\mathcal{Q}} = H_{\mathbf{q}}\left(\mathbf{R_{qq}}\right) - H_{\mathbf{q}}\left(\mathbf{R_{\underline{q}\underline{q}}}\right)$.
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 [43]. 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.

b) Quaternion-valued Gaussian Distribution: The PDF of a zero-mean Gaussian RV $\mathbf{q}$
with real representation $\mathbf{R_{vv}}$ is given by [43]

$$p\left(\mathbf{v}\right) = \frac{\exp\left(-\frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{R_{vv}^{-1}}\mathbf{v}\right)}{\left(2\pi\right)^{2N} |\mathbf{R_{vv}}|^{1/2}}. \tag{48}$$

Equivalently, it can be expressed using the augmented representation as [167]

$$p\left(\underline{\mathbf{q}}\right) = \frac{\exp\left(-\frac{1}{2}\underline{\mathbf{q}}^{\mathrm{H}}\mathbf{R_{\underline{q}\underline{q}}^{-1}}\underline{\mathbf{q}}\right)}{\left(\pi/2\right)^{2N} |\mathbf{R_{\underline{q}\underline{q}}}|^{1/2}}. \tag{49}$$

In the Gaussian case, the distribution of a $\mathcal{H}$-proper variable is contained in a 4D hypersphere
[144].

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 [178]. 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.

E. 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 III. They include hypothesis testing based on likelihood ratio tests (LRT) [33], [36], [62], [113], [166], [168], [172], [173], [179]–[182], [185], maximum likelihood (ML) estimation of circularity coefficients [184], Wald's type detector [162], [183] and invariant testing [33] etc. For instance, Schreier *et al.* [172] and Olilla *et al.* [36] independently proposed a binary hypothesis test for impropriety of the complex Gaussian data based on generalized LRT (GLRT) and simulated test thresholds. Whereas, [33] proposed analytical test threshold based on the theoretical analysis of the null asymptotic distribution of the test statistic. In [186], 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 [179]. A robust Walds type ML (WTML)-detector for propriety is also presented, robust to deviations from the Gaussianity

TABLE III: 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 [36] [172]
			independent but non-identical samples [179]
			i.i.d complex elliptically symmetric distributions [173]
			i.i.d samples of Gaussian RV and analytical test threshold [180]
	Frequency	Complex vector time-series	stationary Gaussian vector sequence with spectral identification [181]
		Complex vector sequences	stationary non-Gaussian vector sequence [182]
			Stationary non-Gaussian signal with improper/colored noise [62]
	Spectral image	2D-spectrum	Stationary non-Gaussian image with improper Gaussian noise [113]
	Conventional	Quaternion RV	i.i.d sample of quaternion Gaussian RV [168]
LRT with real representation	Conventional	Quaternion RV	i.i.d sample of quaternion Gaussian RV [166]
GLRT with real representation		Complex RV	i.i.d samples of Gaussian RV and analytical test threshold [33]
Wald's type detector			i.i.d CES distributions [183] [162]
Circularity coefficients estimation			asymptotic analysis for arbitrary distribution [184]
Invariant testing			i.i.d samples of Gaussian RV and analytical test threshold [33]

assumption, based on multiple i.i.d samples in the broad class of complex elliptically symmetric (CES) distributions [162], [183].

Previous parametric impropriety tests are limited to a sequence of independent RV. To address this concern, [181] and [182] extended the results to stationary Gaussian and non-Gaussian sequences, respectively, using power spectral representations. Interestingly, [181] specifies the frequencies causing improperness when propriety is invalid, whereas the approach in [182] 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 [166], [168] and spectral images [113]. 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 [185].

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 [33], [36], [187]. Moreover, GLR is well-known for its invariance characteristics as the hypothesis test (like propriety) must be invariant to LT but not WLT [172], [188].

Let $\mathbf{z} \in \mathbb{C}^N$ be a complex Gaussian RV with the PDF given by (50). 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 $\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}$. 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

$$p\left(\mathbf{Z}\right) = \pi^{-MN} \left|\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}\right|^{-M/2} \mathrm{x}$$

$$\exp\left\{-\frac{1}{2}\sum_{m=1}^{M}\left(\underline{\mathbf{z}}_m - \mu_{\underline{\mathbf{z}}}\right)^{\mathrm{H}} \mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}^{-1}\left(\underline{\mathbf{z}}_m - \mu_{\underline{\mathbf{z}}}\right)\right\}, \tag{50}$$

$$= \pi^{-MN} \left|\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}\right|^{-M/2} \exp\left\{-\frac{M}{2}\mathrm{Tr}(\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}^{-1}\hat{\mathbf{R}}_{\underline{\mathbf{z}\mathbf{z}}})\right\}, \tag{51}$$

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

$$\hat{\mathbf{R}}_{\underline{\mathbf{z}\mathbf{z}}} = \frac{1}{M}\sum_{m=1}^{M}(\underline{\mathbf{z}}_m - \hat{\mu}_{\underline{\mathbf{z}}})(\underline{\mathbf{z}}_m - \hat{\mu}_{\underline{\mathbf{z}}})^{\mathrm{H}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{z}\mathbf{z}} & \hat{\tilde{\mathbf{R}}}_{\mathbf{z}\mathbf{z}} \\ \hat{\tilde{\mathbf{R}}}_{\mathbf{z}\mathbf{z}}^* & \hat{\mathbf{R}}_{\mathbf{z}\mathbf{z}}^*, \end{bmatrix} \tag{52}$$

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 IV. GLRT is the ratio of likelihood with $\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}$ constrained to have zero off-diagonal blocks i.e., $\tilde{\mathbf{R}}_{\mathbf{z}\mathbf{z}} = \mathbf{0}$, to likelihood with unconstrained $\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}$. Thus, GLRT is testing the block-diagonal structure of the augmented covariance matrix as [36], [172]

$$\lambda = \frac{\max\limits_{\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}, \tilde{\mathbf{R}}_{\mathbf{z}\mathbf{z}}=\mathbf{0}} p\left(\mathbf{Z}\right)}{\min\limits_{\mathbf{R}_{\underline{\mathbf{z}\mathbf{z}}}} p\left(\mathbf{Z}\right)}. \tag{53}$$

In a GLR, the unknown parameters are replaced by ML estimates [187]. It is well known that

TABLE IV: Test Thresholds for $\mathbb{C}-$GLR Hypothesis Tests

Null Hypothesis (Proper)	Alternate Hypothesis (Improper)	Test Threshold	Ref
$H_0 : \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$	$H_1 : \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \tilde{\mathbf{R}}_{\mathbf{zz}}^{\dagger}$	Absent	[36]
$H_0 : \ \tilde{\mathbf{R}}_{\mathbf{zz}} = \mathbf{0}$	$H_1 : \ \tilde{\mathbf{R}}_{\mathbf{zz}} \neq \mathbf{0}$	Simulated	[172]
$H_0 : \ \text{all } \delta_n = 0$	$H_1 : \ \text{all } \delta_n \neq 0$	Approximated	[33]
$H_0 : \ \rho = 0$	$H_1 : \ \rho \neq 0$	Asymptotic	[162]

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 $\hat{\mathbf{R}}_0$, respectively, where $\hat{\mathbf{R}}_0$ is defined as

$$\hat{\mathbf{R}}_0 = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{zz}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{R}}_{\mathbf{zz}}^* \end{bmatrix}. \tag{54}$$

Therefore, GLR can be expressed as

$$l = \lambda^{\frac{2}{M}} = \left| \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{R}}_{\underline{\mathbf{zz}}} \right| \left(\exp \left\{ -\frac{M}{2} \text{Tr}(\hat{\mathbf{R}}_0^{-1} \hat{\mathbf{R}}_{\underline{\mathbf{zz}}} - \mathbf{I}) \right\} \right)^{\frac{2}{M}}$$

$$= \frac{\left| \hat{\mathbf{R}}_{\underline{\mathbf{zz}}} \right|}{\left| \hat{\mathbf{R}}_{\mathbf{zz}} \right|^2} = \frac{\left| \hat{\mathbf{R}}_{\mathbf{zz}} - \hat{\tilde{\mathbf{R}}}_{\mathbf{zz}} \hat{\mathbf{R}}_{\mathbf{zz}}^{-*} \hat{\tilde{\mathbf{R}}}_{\mathbf{zz}}^* \right|}{\left| \hat{\mathbf{R}}_{\mathbf{zz}} \right|}. \tag{55}$$

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 [36] and the hypothesis testing involves the test for null $\tilde{\mathbf{R}}_{\mathbf{zz}}$ signifying proper RV $\mathbf{z}$ or known $\tilde{\mathbf{R}}_{\mathbf{zz}}^{\dagger}$ signifying improper RV $\mathbf{z}$. On the other hand, [172] presents slightly different hypothesis test as given in Table IV 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 10. We then have

$$l = \left| \mathbf{I} - \hat{\mathbf{K}} \hat{\mathbf{K}}^{\text{H}} \right| = \prod_{n=1}^{N} \left(1 - \hat{k}_n^2 \right). \tag{56}$$

This explains that the GLR is a function of the squared canonical correlations which make up a complete/maximal set of invariants under LT. Contributions like [184] 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 [33] 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 $\mathbf{R}_{\mathbf{uu}} = \mathbf{G}\Delta_1\mathbf{G}^{\mathrm{T}}$ with $\Delta_1 = \mathbf{Diag}\left((\mathbf{I} + \Delta_2), (\mathbf{I} - \Delta_2)\right)$ and $\Delta_2 = \mathbf{Diag}\left(\delta_1, \delta_2, \ldots, \delta_N\right)$. They accept H_0 iff

$$T_1 \equiv \prod_{n=1}^{N} \left(1 - \hat{\delta}_n^2\right) \geq c_1\left(\eta_1, M, N\right), \tag{57}$$

$$T_2 \equiv \sum_{n=1}^{N} \left(\hat{\delta}_n^2\right) \leq c_2\left(\eta_1, M, N\right), \tag{58}$$

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 [33]. 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 $N \geq 2$ [33]. 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 [33], [173], [179]. 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 [162] 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

$$M|\hat{\rho}|^2/\hat{\varsigma}_0 \geq \chi_{2;1-\eta_2}^2 \tag{59}$$

where $\chi_{2;1-\eta_2}^2$ 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 [162, (20)]. An interesting extension is to look for the extent of non-circularity when the null-hypothesis is rejected e.g., [173] 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 [186]. One of the proposed efficient implementation involves the following test statistic

$$r = \frac{2^{2N}\left|\hat{\mathbf{R}}_{\mathbf{uu}}\right|}{\left|\hat{\mathbf{R}}_{\mathbf{zz}}\right|^2}, \tag{60}$$

TABLE V: Quaternion GLR Hypothesis Tests and Decision Criterion

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

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 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 (55), owing to its inherent redundancy.

2) GLRT for Quaternion Random Vectors: Like complex RV, propriety testing has also been expanded for the quaternion RV based on its real composite [166] as well as augmented representation [168]. 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 V. 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 $\mathfrak{R}_{\mathcal{Q}}$ (24) and $\mathfrak{R}_{\mathcal{C}^\alpha}$ (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 [168]

$$\hat{\mathbf{D}}_{\mathcal{Q}} = \textbf{Block-Diag}\left(\hat{\mathbf{R}}_{\underline{\mathbf{qq}}}, \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}^{(i)}, \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}^{(j)}, \hat{\mathbf{R}}_{\underline{\mathbf{qq}}}^{(k)}\right), \tag{61}$$

$$\hat{\mathbf{D}}_{\mathcal{C}^\alpha} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{qq}} & \hat{\mathbf{R}}_{\mathbf{qq}^{(i)}} & \mathbf{0}_N & \mathbf{0}_N \\ \hat{\mathbf{R}}_{\mathbf{qq}^{(i)}}^{(i)} & \hat{\mathbf{R}}_{\mathbf{qq}}^{(i)} & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \hat{\mathbf{R}}_{\mathbf{qq}}^{(j)} & \hat{\mathbf{R}}_{\mathbf{qq}^{(i)}}^{(j)} \\ \mathbf{0}_N & \mathbf{0}_N & \hat{\mathbf{R}}_{\mathbf{qq}^{(i)}}^{(k)} & \hat{\mathbf{R}}_{\mathbf{qq}}^{(k)} \end{bmatrix}. \tag{62}$$

GLRT comparisons for $\mathcal{Q}$-properness, $\mathcal{C}^\alpha$-properness and improperness are established with predefined fixed thresholds $\gamma_{\mathcal{Q}}, \gamma_{\mathcal{C}^\alpha}$ and $\gamma_{\mathcal{R}^\alpha}$, respectively.

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 [33]. Most of the contributions describe the GLRT of circularity assuming complex normal data [36], [172], 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 [189]. Similarly, GLRT for quaternions is focused on the identification of $\mathcal{Q}$-proper, $\mathcal{C}^{\alpha}$-proper or possibly improper RV.

IV. Impropriety Sources and Consequences

The most common speculation in the complex analysis assumes propriety and/or circularity of the r.v. under investigation. However, this phenomenon assumes uncorrelated real and imaginary components of the complex entities with/without identical variance which is generally not true. 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. 5.

A. Communications

A necessary, yet insufficient condition for impropriety is that the real-valued random process be at least nonstationary [190]. 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) [32]. Thus, the potentially 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 [4], [22], PAM [5], GMSK [23], [24], OQPSK or staggered QPSK [25], SSB, VSB [26] and baseband (but not passband) OFDM [27]. On the other hand, coding schemes like STBC also result in improper signals [28]. 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 [48] and general simplex signals. In many
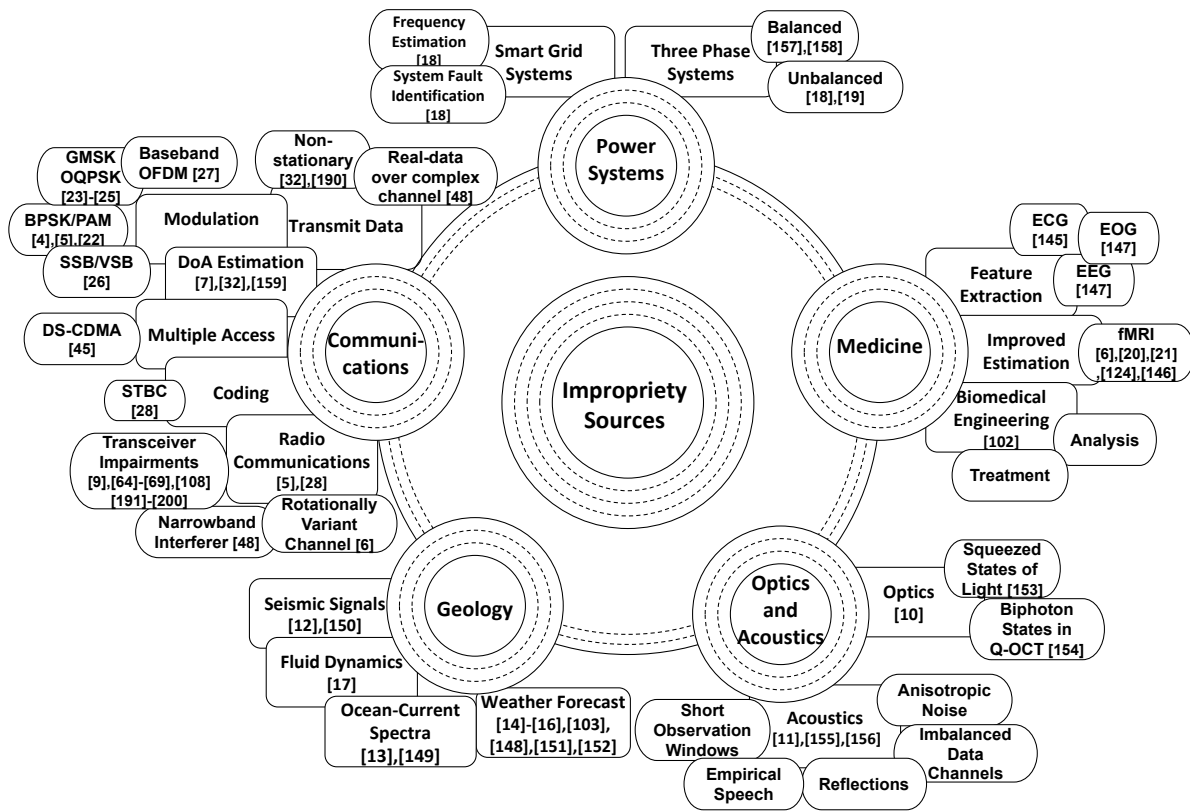
Fig. 5: Interdisciplinary Sources of Improperness

sensor applications, DoA estimation with electronically steerable antenna arrays is improved by employing NC signal constellations [7], [32], [159]. In addition, multiple access scheme like direct-sequence code-division multi-access (DS-CDMA) [45] with narrowband interferer results in NC interference [48]. 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 [107].

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 [108]. This can be caused by transceiver imperfections [9], [67]–[69], communication channels that are not rotationally invariant [6] or NC interference from other sources [5]. Therefore, WL detection of the multiple-input multiple-output (MIMO) signals is emphasized by arguing the non-circularity of the data signal arising inherently within an iterative receiver [28].

B. 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 [18], [19]. 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 estimation of instantaneous frequency in a three-phase system under both balanced and unbalanced conditions [157], [158]. The complete second-order information offers next-generation solutions for accurate, adaptive and robust frequency estimation as well as system fault identification [18].

C. Medicine

In medical data analysis, non-circularity is exploited for features extraction and improved estimations in electrocardiograms (ECG) [145] and functional magnetic resonance imaging (fMRI) [124], [146], respectively. The fMRI measured neural activities in the brain or spinal cord are recently recognized as improper signals [6], [20], [21]. Thus, the NC probability distributions of the sources enable real-time extraction of eye muscle activity: electrooculogram (EOG) from the electroencephalogram (EEG) recordings using some blind source extraction algorithms [147]. Rhythms of brain waves are non-stationary signals [102] 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.

D. 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 [10]. 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, LED and lasers have trivial phase-sensitive correlation whereas advanced non-linear optical sources like squeezed states of light possess phase sensitive correlation [153]. 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 [154]. 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 [155], [156]. It is based on the fact that speech can be empirically improper in the frequency domain [11].

E. Oceanography and Geophysics

Complex trace analysis of seismic signals [12], interpretation of ocean-current spectra [13] and the analysis of wind fields [14]–[16] exploit the improperness of their respective complex representations. The impropriety concept is well-known and appreciated by meteorologists and oceanographers since the early 1970s [148], [149]. Recently, the improper nature of the real-time wind data along with the tracking of degree of circularity is identified for renewable energy applications [103], [151]. Similarly, adaptive filtering of the real-world wind signals is based on its NC and nonstationary observations [152]. Others employ impropriety concepts to further physical applications with asymmetric/anisotropic data structures e.g., in fluid dynamics [17] and seismic data signals [150] to overcome modeling challenges. Most of the above mentioned contributions are backed by [31], which claims the improper nature of signals involved in various real problems.

F. Hardware Imperfections

Hardware impairments (HWIs) like non-linearity, IQI and phase noise rise in various radio frequency (RF) and baseband blocks such as analog-to-digital convertors, high power and low noise amplifiers, low-pass and band-pass filters etc. [191], [192]. These impairments accumulate at both the transmitter (TX) and receiver (RX) RF chains and yield undesired effects. For instance, non-linearities result in additive Gaussian distortion noise [193]–[195], whereas IQI not only induces phase and amplitude errors but also mixes the desired and image signals [59], [66], [68],

(a) Symmetric 16-PSK   (b) Asymmetric 16-PSK

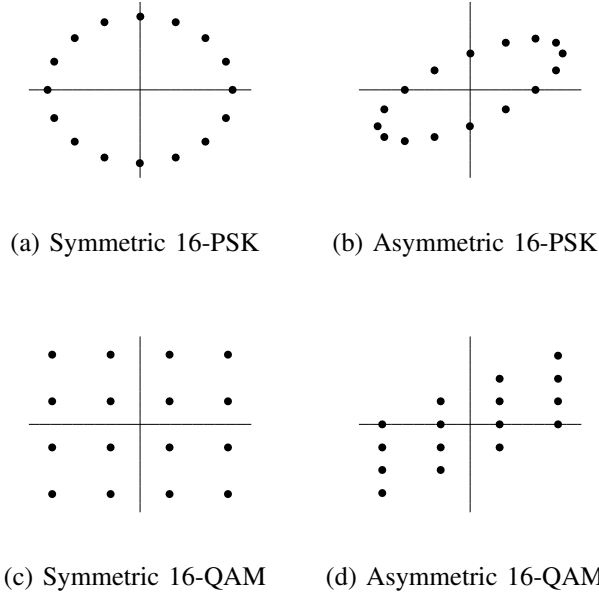(c) Symmetric 16-QAM   (d) Asymmetric 16-QAM

Fig. 6: Widely Linear Transformed Signal Constellations

[192]. Considering the information bearing transmitted signal $x_0 \sim \mathcal{CN}(0, \sigma_x^2, 0)$, the noiseless down-converted IQI signal is modeled as WLT [196]

$$x_{\text{IQI}} = \alpha_1 x_0 + \alpha_2 x_0^*, \tag{63}$$

where, α_1 and α_2 are complex scalars which capture the amplitude and phase errors. The WL modeling of IQI is thoroughly investigated and analyzed in the literature [197]. However, only few works have acknowledged the asymmetric characteristics induced by this WL structure of IQI [9], [65], [67], [197]. The WLT of IQI transforms PGS to IGS, which can be verified by finding non-zero pseudo-variance of the resulting signal $\tilde{\sigma}_{x_{\text{IQI}}}^2 = 2\alpha_1 \alpha_2 \sigma_x^2$ [198]. Similarly, the RX-IQI is mainly responsible for the improper Gaussian additive noise [160], [161]. Therefore, transceiver impairments specifically IQI play a vital role in transforming proper transmitted signal to improper received signal as well as transforming proper Gaussian noise to improper Gaussian noise [9], [64], [69], [199], [200]. An example of baseband equivalence of WLT of discrete 16-PSK and 16-QAM owing to TX-IQI is depicted in Fig. 6. Besides wireless communications, other noise models such as underwater propeller noise from maritime ship also demonstrate an improper nature [102].
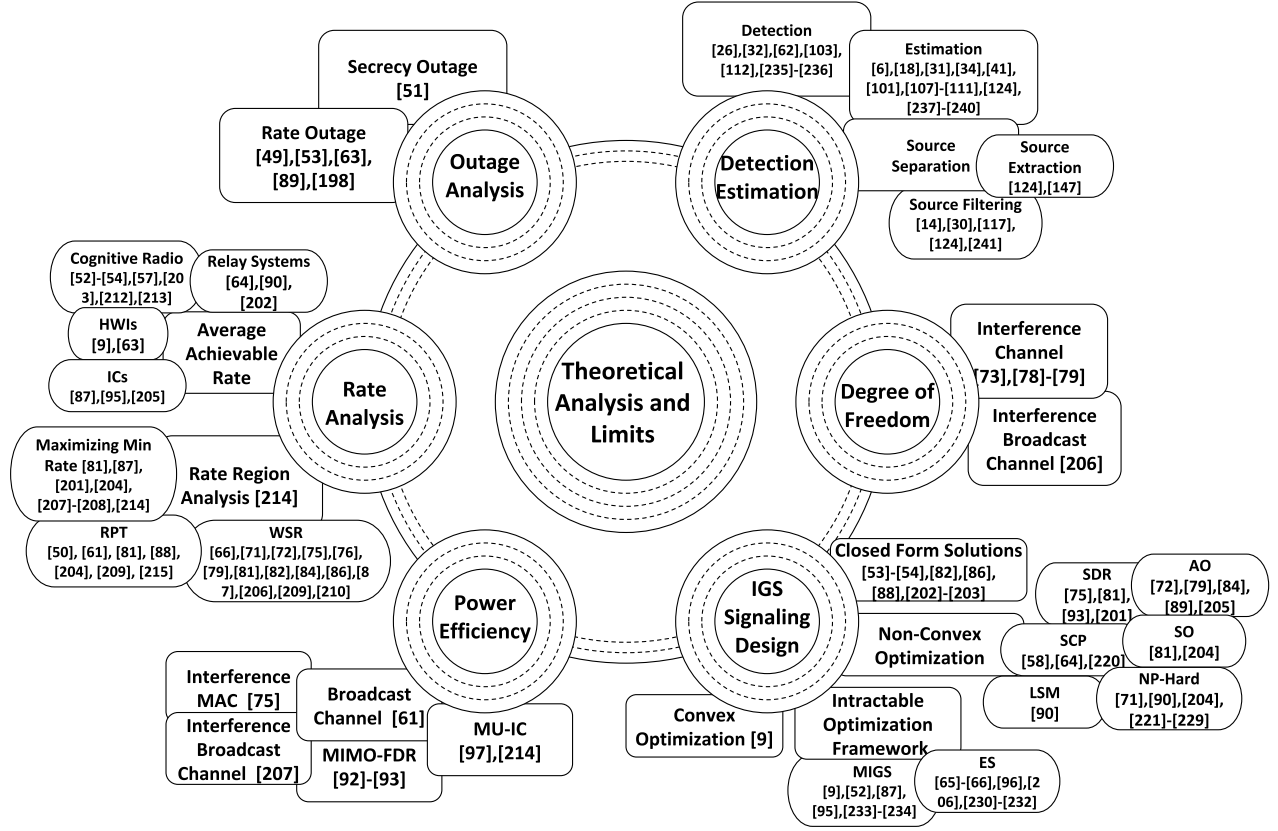
Fig. 7: Overview and Configuration of Theoretical Analysis

V. 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 [175]. 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 [42]. 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 Fig. 7.

A. 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 IGS [9], [50], [52], [54], [55], [64], [82], [86], [88], [201]–[204], WL precoding of PGS [66], [71], [72], [79], [81], [84], [87], [95], [205], or complex beamforming resulting in dependent real and imaginary parts [90], [206].

As an illustrative example, consider the most general case of K-user MIMO-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

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{x}_k + \sum_{j \neq k, j=1}^{K} \mathbf{H}_{kj}\mathbf{x}_j + \mathbf{n}_k, \tag{64}$$

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 $\mathbf{H}_{kj} \in \mathbb{C}^{N \times M}$ 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 [6].

$$H(\mathbf{z}) = \log\left((\pi e)^N |\mathbf{R}_{\mathbf{zz}}|\right) + \frac{1}{2}\log\left(\left|\mathbf{I} - \mathbf{R}_{\mathbf{zz}}^{-1}\tilde{\mathbf{R}}_{\mathbf{zz}}\mathbf{R}_{\mathbf{zz}}^{-T}\tilde{\mathbf{R}}_{\mathbf{zz}}^{H}\right|\right).$$

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]

$$R_k = \mathcal{I}\left(\mathbf{y}_k; \mathbf{x}_k\right) = H(\mathbf{y}_k) - H(\mathbf{y}_k|\mathbf{x}_k) = H(\mathbf{y}_k) - H(\mathbf{w}_k)$$

$$= \frac{1}{2}\log \frac{\mathbf{R}_{\underline{\mathbf{y}}_k \underline{\mathbf{y}}_k}}{\mathbf{R}_{\underline{\mathbf{w}}_k \underline{\mathbf{w}}_k}} = \underbrace{\log \frac{\left|\sigma^2 \mathbf{I} + \sum\limits_{j=1}^{K} \mathbf{H}_{kj} \mathbf{R}_{\mathbf{x}_j \mathbf{x}_j} \mathbf{H}_{kj}^{\mathrm{H}}\right|}{\left|\sigma^2 \mathbf{I} + \sum\limits_{j \neq k, j=1}^{K} \mathbf{H}_{kj} \mathbf{R}_{\mathbf{x}_j \mathbf{x}_j} \mathbf{H}_{kj}^{\mathrm{H}}\right|}}_{\triangleq R_{k,\mathrm{proper}}\left(\mathbf{R}_{\mathbf{x}_j \mathbf{x}_j}\right)}$$

$$+ \frac{1}{2}\log \frac{\left|\mathbf{I} - \mathbf{R}_{\mathbf{y}_k \mathbf{y}_k}^{-1} \tilde{\mathbf{R}}_{\mathbf{y}_k \mathbf{y}_k} \mathbf{R}_{\mathbf{y}_k \mathbf{y}_k}^{-\mathrm{T}} \tilde{\mathbf{R}}_{\mathbf{y}_k \mathbf{y}_k}^{\mathrm{H}}\right|}{\left|\mathbf{I} - \mathbf{R}_{\mathbf{w}_k \mathbf{w}_k}^{-1} \tilde{\mathbf{R}}_{\mathbf{w}_k \mathbf{w}_k} \mathbf{R}_{\mathbf{w}_k \mathbf{w}_k}^{-\mathrm{T}} \tilde{\mathbf{R}}_{\mathbf{w}_k \mathbf{w}_k}^{\mathrm{H}}\right|}. \tag{65}$$

The achievable rate R_k at RX-k has two components, the first term $R_{k,\mathrm{proper}}\left(\mathbf{R}_{\mathbf{x}_j \mathbf{x}_j}\right)$ 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 (65) 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 [201], SISO Z-IC [88], single-input multiple-output (SIMO) systems with hardware distortions (HWDs) [63], MU multiple-input single-output (MISO)-IC [204], MU MIMO-IC [84], MU MIMO-BC [72], MIMO-IBC [71], cognitive radio settings [50], [52], [54], [203], point-to-point (P2P) systems with asymmetric HWIs [9], relaying systems [202], full-duplex systems [64] 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].

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 [52], [54], [203], multi- antenna or MU systems [9], [87], [95], [205], and relay systems [64], [90], [202] 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 [26], [207]. 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 [52], [53]. 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 [52], [57], [208].

- 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 [54], [209].

- 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 [203].

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 [64], [90], [202]. 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 [90]. 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 [64]. 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 [202].

IGS superiority is also proven in multi-antenna setup to mitigate various **HWIs** [9], [63]. 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 [9].

Other contributions highlight the perks of IGS in single user [87] and **MU MIMO-ICs** [95], [205]. 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 [87]. 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 [95]. 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 [205].

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 [88].*

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 [210].

a) 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 [75], [76], Z-IC [86], and X-IC [82]. Apart from the complete boundary characterization, some studies focused on the weighted sum-rate (WSR) maximization achieved by IGS in various system settings like cognitive radio [50], massive MIMO with IQI [66], X-IC [79], [81], [84], [213], Z-IC [86], [87], BC [72], [214], and IBC [71], [206].

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 [75], [76]. Secondly, underlay 2U **secondary-MAC** (S-MAC) exploited IGS to improve average sum rates up to 200% [50]. 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 [66]. They argued the achievability of same multiplexing gains and WSR with WL-ZF and WL-

TABLE VI: Achievable Rate Improvement by IGS over PGS in Different Settings

System	Tech	Desc	IGS-TX	Metric	Procedure	Improvement	Ref
Cognitive Radio	Underlay	P2P	SU	Max SU-Rate with PU-Rate Constraint	Maximal IGS	up to 256%	[52]
		S-MAC	S-MAC Users	Rate Region Boundary Maximize Sum Rate	Algorithm: Closed-Form Expressions	R1:133%, R2:117% up to 200%	[50]
	Overlay	P2P	SU	Max SU-Rate with PU-QoS Constraint	Piecewise Closed-Form Solution	up to 191.7%	[54]
	Interweave	P2P	SU			up to 8.3%	[203]
MIMO	Massive	MU DL	BS	Maximize Sum Rate	Numerical Evaluation	ZF/MMSE:75% BD:97%	[66]
	X-IC	2U	SU	Maximize SU-Rate	Maximal IGS	R1:195.5% increase R2:8.7% decrease	[95]
		MU	All	Minimize MSE	AO	cent-IM:10.6% decent-IM:13.9% no-IM:22.9%	[205]
		LP/WLP Users	Only WL Users	Maximize WSR ($\equiv$ Minimize WMSE)	BCD with AO	WLP:133%, HetTX(1):84.2% HetTX(K/2):28.9%	[84]
	Z-IC	OPA	Interferer	Average Sum Rate Min-Rate Performance	Maximal IGS	Sum rate:55.2% Min rate:640%	[87]
	BC	Multicell N/w	BS	Maximize WSR	BCD with AO	up to 11.8%	[72]
	IBC	HCRAN	Femto BSs in all cells		WMMSE and ADMM for AO	WMMSE-IGS:12.5% WMMSE-PGS:21.7%	[71]
		MU	BS	Maximize min rate	Path following Algo	30dBm: 37.93%	[211]
	P2P-I	UPA/OPA	MIMO TX	Maximize Rate Minimize MSE	Majorization Theory Tools	Rate Gap:1.55 MSE GAP:0.2 Rate Gain:115%	[87]
MISO	IC	MU P2P	Users	Pareto Rate Region Max-Min Rate	SDR with SOCP and GR	R1:70%, R2:25% up to 42.8%	[204]
	BC	2U	BS	Pareto Rate Region	Separate Optimization with Bisection Search	R1:150%, R2:100%	[204]
		3U	BS	Pareto Rate Region	Gradient based Rate Balancing	R1:100%, R2/R3:42.86%	[61]
	NOMA	Multicell N/w	BS	Max-min Fairness	LMI based Path following Algorithm	NOMA: 87.50% OMA:121.05%	[212]
SIMO	P2P	HWI	TX	Maximize Achievable Rate	Convex QCQP using IPM	up to 10%	[9]
SISO	X-IC		1U	Rate Region	Closed-form	R1:133.7%,R2:86.6%	[213]
		2U		Max-Min Fairness Proportional Fairness	Closed-Form Solution	up to 83% up to 633%	[82]
				Max-Min Rate	SDR	up to 228.57%	[204]
			All users	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		Maximize Sum Rate	Alternating Minimization	up to 35.2%	[79]
		MU		Max-Min Weighted Rate	SDR and GR	2U:221.4%, 3U:304.3%	[201]

SISO	Z-IC	P2P		Maximize Sum Rate	Closed-Form Solution	up to 30.8%	[86]
				Pareto Rate Region		R1:83%, R2:150%	[88]
						R1:30.9%,R2:36.2%	[213]
	BC-NOMA	2U	BS	Max Sum Rate	KKT conditions	$\Delta P = 0.4P_t$: 23.60%	[214]
	IBC	MU in 3-Cells	Each BS in all Cells	Maximize Sum Rate	Exhaustive Search	K=5: 13.7% K=10:17.6% K=20:20.6%	[206]
DF Relays	FD	Multi-hop with HWIs	Source and Relays	Maximize E2E Rate	SCP for QCLP	Centralized: up to 355.6% Distributed: up to 80%	[64]
		Dual-hop		Throughput	Line Search Method	up to 107%	[90]
	HD	Alternate Relaying	Relays only	Maximize Total Achievable Rate	Piecewise Closed-Form Solution	SUD:161.5% SUD-SD:66.7%	[202]

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 [66].

Rate region improvement of **Z-IC** has also been studied extensively [86]. This contribution employs real-composite representation for easy optimization and characterizes only one point of rate region [86]. 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 [86].

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 [82]. 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 [82]. 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 [213]. 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 [79]. 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) [84]. 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 [84].

Apart from IC, WSR maximization problem is also extensively studied for **BC** and **IBC** [71], [72], [206], [214]. 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 (P_t), respectively [214]. 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 [72]. In contrast to the IC and BC, few studies focused on analyzing the potential benefits of IGS in the combined

IBC [71], [206]. 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 [206]. WSR maximization problem in MIMO-IBC (e.g., heterogeneous cloud radio access network (H-CRAN)) is a non-trivial extension of MIMO-IC [84] and MIMO-BC [74]. 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 [71].

b) 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 [50], Z-IC [88], [213], X-IC [81], [204], [213], [215], and BC [61], [204].

Lameiro *et al.* extended their work in underlay P2P cognitive system [52] 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 [50].

Rate region improvement of **Z-IC** has also been studied extensively [88]. The contribution [88] 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 [88]. 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 [213].

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], [204], [213], [215]. 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 [213]. 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 [215]. 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 [204].

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 improvement [204]. Similarly, for a 3U MISO-BC, IGS offers R2/R3:42.86% and R1:100% increase relative to PGS [61].

c) 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], [201], MU MISO-IBC [212], MU MISO-IC [204], MU MIMO Z-IC [87], MU MIMO-IC [210], and MU MIMO-IBC [211].

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 [204]. 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 [201].

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 [212].

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** [87]. Moreover, a hardware impaired MU MIMO-IC depicts more than 80% improvement in fairness rate with IGS over PGS in a 10U setup [210]. 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 [211].

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 VI.

B. 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.

1) Rate Outage Probability (ROP): ROP is generally defined as the probability of the event $\Pr\{\text{Rate} < \text{Threshold}\}$. Various contributions have demonstrated substantial decrease in the ROP when IGS transmission is adopted as opposed to PGS. These studies analyzed various systems including but not limited to cognitive radio [49], [53], multi-antenna [63] and relay systems [89] 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 [53]. 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 [49]. 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 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 [89].

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 [63] and SISO [198] system, respectively. The ROP gain in

TABLE VII: Outage Probability Reduction with IGS Relative to PGS

System	Technology	Description	Improper TX	Metric	Procedure	Improvement	Ref
Multiple Antennas	SIMO	HWDs	Single TX	Minimize Rate Outage	Closed-Form Solution	SISO:85% SIMO:100%	[63]
DF Relays	FDR	Dual-hop	Relays	Minimize E2E Outage	Coordinate Descent	MPGS: 90% OPGS: 68.18%	[89]
Cognitive Radio Setting	Underlay	HD-PU	SU only	Minimize Rate Outage	Closed-Form Solution	up to 77.5%	[53]
		FD-PU			Algorithm Design	up to 91.43%	[49]
		Eavesdropper		Minimize Secrecy Outage	Numerically	up to 95%	[51]

imperfect hardware system is especially significant with increased distortion levels and more receiver streams N_{R} for any given threshold rate.

2) Secrecy Outage Probability (SOP): 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., $\mathrm{Pr}\{\mathcal{I}_{\mathrm{AB}} - \mathcal{I}_{\mathrm{AE}} < \mathrm{Threshold}\}$. 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 SOP. It is evident from the fact that IGS demonstrate lower differential entropy 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 [51].

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 VII and Fig. 8.
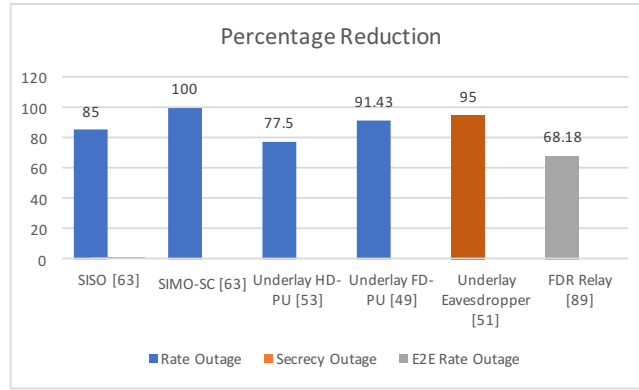
Fig. 8: Percentage Reduction in Outage Probability with IGS

C. 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 [75]. 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 [211].

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 [97]. 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 [61].

IGS is beneficial to improve the power efficiency of one-way DF-FD MIMO relay [93] and two-way AF-FD MIMO relay [92] 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 [93]. 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 [92]. 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 [210].

D. 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 [78], [79], BC [73]and IBC [206]. For instance, Lameiro *et al.* illustrated improved DoF $\approx 4/3$ (requiring minimal symbol extensions) with IGS and linear IA for 4U **SISO-IC** [79]. 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 [78]. Moreover, in **MIMO-BC** system with WL transceivers IGS per-user transmit signals outperform PGS counterpart under QoS constraints without time-sharing [73]. IGS guarantees faster convergence of effective DoF to its upper bound relative to PGS for a given rate requirement [73]. 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 [206].

E. IGS Signaling Design

Theoretical analysis of the IGS revealed tremendous payoffs in various MU scenarios such as the X-IC [81], [82], [84], [216], Z-IC [85]–[88], BC [83], [217], cognitive radio networks [52]–[54], [203], HWI systems [29], [63], [194], [199] and relay channels [64], [89], [202]. The vast majority of them assumed Gaussian codebooks for more efficient IM [218]. 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.

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) [82], [88] and maximizing sum-rate in SISO Z-IC [86] yield closed form solutions. Alternately, other optimization problems produce piece-wise closed form solutions. Consider an **overlay cognitive radio** [54] or **interweave cognitive radio** [203] 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 $R_p(p_s, \mathcal{C}_x) \geq R_{\min}$ (R_p is the achievable rate of the PU). The optimization problem is similar to **P1** and tunes IGS transmission parameters of SU i.e., transmit power p_s and circularity coefficient $\mathcal{C}_x$.

$$\textbf{P1}: \max_{p_s, \mathcal{C}_x} \quad R_s(p_s, \mathcal{C}_x)$$
$$\text{s. t.} \quad R_p(p_s, \mathcal{C}_x) \geq R_{\min},$$
$$0 \leq p_s \leq p_{s,\max},$$
$$0 \leq \mathcal{C}_x \leq 1.$$

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 $p_{s,\max}$ 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 [202] is equivalent to **P1** with maximum power transmission eliminating any primary rate constraint and can also be solved as piece-wise closed form solution.

Another form of **P1** is the ROP minimization problem **P2** in **cognitive radio** setup. The SU optimizes IGS transmission parameters p_s and $\mathcal{C}_x$ to minimize its ROP i.e., $P_{\text{out},s}(p_s, \mathcal{C}_x) =$

$\Pr\{R_s(p_s, \mathcal{C}_x) \leq R_{\min}\}$ while maintaining PU QoS $P_{\text{out},p}(p_s, \mathcal{C}_x) \leq P_{\text{out,th}}$ [53].

$$\textbf{P2}: \min_{p_s, \mathcal{C}_x} \quad P_{\text{out},s}(p_s, \mathcal{C}_x)$$

$$\text{s. t.} \quad P_{\text{out},p}(p_s, \mathcal{C}_x) \leq P_{\text{out,th}},$$

$$0 \leq p_s \leq p_{s,\max},$$

$$0 \leq \mathcal{C}_x \leq 1.$$

Interestingly, problems similar to **P2** can also be solved in closed form by investigating the monotonic trend of the objective function with respect to the optimization variables.

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 **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 [9].

$$\textbf{P3}: \max_{\mathbf{s}} \quad \text{R}_{\text{SIMO-IGS}}(\mathbf{s})$$

$$\text{s. t.} \quad \mathbf{A}_1 \mathbf{s} \leqslant \mathbf{b},$$

$$\mathbf{s}^{\text{T}} \mathbf{A}_2 \mathbf{s} \leqslant \mathbf{0},$$

where $\mathbf{A}_1 = \text{diag}\,[0 \ 0 \ 1]^{\text{T}}$ and $\mathbf{b} = [0 \ 0 \ \text{P}_{\text{T}}]^{\text{T}}$ signify the transmission power constraint whereas $\mathbf{A}_2 = \text{diag}\,[1 \ 1 \ -1]^{\text{T}}$ 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.

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.

a) 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 **P4** in case of **2U Gaussian SISO-IC** [81].

$$\mathbf{P4}: \max_{\mathbf{R}_{x_k x_k}, \tilde{\mathbf{R}}_{x_k x_k}, R} \quad R$$

$$\text{s. t.} \quad R_k \geq \alpha_k R, \; \forall k$$

$$0 \leq \text{Tr}\left(\mathbf{R}_{x_k x_k}\right) \leq P_k, \; \forall k$$

$$\mathbf{R}_{\underline{x}_k \underline{x}_k} \succeq \mathbf{0}, \text{ or } 0 \leq |\tilde{\mathcal{C}}_{x_k}|^2 \leq \mathcal{C}_{x_k}^2 \; \forall k$$

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 **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 [201].

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 **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 [75].

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 [219] is suggested [93].

b) 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 [64]. Considering the achievable link rate between nodes m and n as $R_{nm}\left(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2\right)$, **P5** aims at maximizing the minimum link rate, i.e.,

$$\mathbf{P5}: \quad \text{maximize} \quad \min_n \left\{ R_{nm}\left(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2\right) \right\}$$

$$\text{subject to} \quad 0 \le \left|\tilde{\sigma}_m^2\right| \le \sigma_m^2, \ \forall m.$$

The joint optimization **P5** turned out to be a max-min fractional programming problem which can be solved in two different ways.

1) Generalized Dinkelbach algorithm (GDA) [220] can transform non-linear fractional programming to non-linear parametric programming and SCP can transform the resultant DC problem to QCQP convex optimization problem [58]

2) Alternately, the logarithmic properties can transform fractional programming problem to DC problem and eliminate GDA step. However, SCP is inevitable to solve further [64].

SCP for DC programming approximates second concave to its first-order Taylor series expansion and solves the resultant convex problem iteratively.

c) 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 [84] and decentralized processing [205]. 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 equivalently minimizing W-MSE with weights μ_k, i.e.,

$$\mathbf{P6}: \quad \min_{\substack{\mathbf{T}_{1,k}, \mathbf{T}_{2,k} \\ \mathbf{R}_{1,k}, \mathbf{R}_{2,k}}} \sum_{k \in \mathcal{K}} \frac{\mu_k}{2} \log_2 \left|\mathbf{E}_k \mathbf{F}_k^*\right|$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\mathbf{T}_{1,k}\mathbf{T}_{1,k}^{\mathrm{H}} + \mathbf{T}_{2,k}\mathbf{T}_{2,k}^{\mathrm{H}}\right) \le P_k \ \forall k,$$

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 **P6** with BCD and AO [72]. 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 [89]. 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** [79].

d) 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 $\mathbf{R}_{x_k x_k}$ and pseudo-covariances $\tilde{\mathbf{R}}_{x_k x_k}$ for $k = 1, 2$ can be carried out in two ways [81].

1) Exclusive optimization

2) Optimizing $\mathbf{R}_{x_k x_k}$ assuming zero $\tilde{\mathbf{R}}_{x_k x_k}$ and then obtaining $\tilde{\mathbf{R}}_{x_k x_k}$ with given $\mathbf{R}_{x_k x_k}$.

As an illustration, **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 $\mathbf{R}_{x_k x_k}$ 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 **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) [204].

e) 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 [90].

f) 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 [221]. Surprisingly, some subclasses of the general NP-hard problem can still be solved in polynomial time [222]. Whereas, others are solved using suboptimal/approximation algorithms e.g., game-theory based algorithm [223], interference pricing based algorithm [224], gradient descent algorithm [225] with line search methods for unconstrained optimization [90], iterative weighted MMSE based algorithm [226],

monotonic optimization frameworks [227], [228], graph theory for combinatorial optimization [229] and SDR for solving non-convex QCQPs with GR by restricting to rank-1 solutions [204]. 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 [71]. 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.

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.

a) Exhaustive Search (ES): IGS parameters can be fine tuned using ES in the feasible domain. For example, the parameters of WLP in downlink [66] and WL RX for uplink [65] in **multi-cell massive MIMO** systems can be chosen using ES. Similarly, exhaustive user scheduling algorithms combined with IGS in **SISO IBC** can improve its sum-rate performance [206]. 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 [96]. Although, ES leads to a near-optimal solution but it comes at the expense of factorial time complexity [230]–[232]. Therefore, a rule of thumb is to employ ES when nothing else works.

b) 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 [9]. 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 [52]. 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 [233]. 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) [87] and multi-antenna systems (specifically 2U MIMO-IC) [95] exploit MTT to demonstrate the superiority of MIGS [234]. 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 [7].

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.

F. IGS Detection and Estimation

In the engineering sciences, the three main branches of statistical signal processing are estimation, detection, and signal analysis [6]. Therefore, various contributions have addressed these issues related to IS as enumerated in Fig. 9.

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

$$\mathsf{H}_0: \ r(t) = n(t), \qquad \mathsf{H}_1: \ r(t) = s(t) + n(t)$$

a) 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 [32]. The performance metric in terms of deflection yields double performance gain when pseudo-covariance is taken into account,
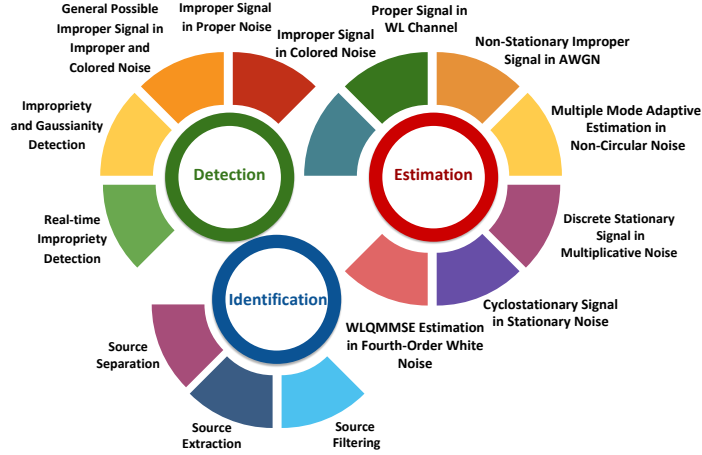
Fig. 9: Conceptual Classification of IGS Detection, Estimation and Identification

generalizing the 3-dB gain of coherent processing over non-coherent processing [32]. 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 [26].

b) 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 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 [62].

c) 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) [235] i.e.,

$$p\left(\underline{\mathbf{z}}, c, \mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}\right) = \frac{\beta(c)}{\sqrt{|\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}|}} \exp\left\{-\alpha\left(c\right)\left(\mathbf{z}^{\mathrm{H}}\mathbf{R}_{\underline{\mathbf{z}}\underline{\mathbf{z}}}^{-1}\mathbf{z}\right)\right\}^{c},$$ (66)

where, $\beta(c) = \Gamma(2/c)/\pi\Gamma(1/c)^2$ and $\alpha(c) = c\Gamma(2/c)/2\Gamma(1/c)$ with Gamma function $\Gamma(.)$ 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 [112]

<div align="center">

Non-circularity   Non-Gaussianity

$\mathsf{H}_0 : \mathbf{R_{\underline{zz}}} = \sigma_z^2\mathbf{I},$ $\mathsf{H}_0 : c = 1, \mathbf{R_{\underline{zz}}} = \sigma_z^2\mathbf{I},$

$\mathsf{H}_1 : \mathbf{R_{\underline{zz}}} \neq \sigma_z^2\mathbf{I}.$ $\mathsf{H}_1 : c \neq 1, \mathbf{R_{\underline{zz}}} \neq \sigma_z^2\mathbf{I}.$

</div>

The adjusted-GLRT detector [236] performs fairly good as CGGD [112] 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 [112].

d) 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 [103].

e) 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.

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 [31], passive radar tracking [237], target tracking [238], power systems frequency estimation [18] and fault diagnosis [239] 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 [6].

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 [124]. 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 [6]

- For joint circularity between observations however correlated observations and estimandum, WL-MSE still offers better estimates than linear MSE [110]

- 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 [110]

- 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 [6]

- The WL-MMSE estimate of a real signal from a complex signal is always real whereas the LMMSE estimate is generally complex [111]

This section further characterizes estimation problems of different signals with different types of noise using SL and WL estimators based on MMSE criterion.

a) 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 [240].

b) 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 [101].

c) 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 [6].

d) 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 [34]. Mohammadi *et al.* also extended their work to distributed estimation using diffusion strategies, when a system is observed distributively using an agent/sensor network [107].

e) 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 [41].

f) 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 [109].

g) 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 [6].

h) Discussion: The performance of various estimators can be distinguished using complementary MSE analysis which quantifies the DoI of the SL and WL estimation errors [108]. 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 [124].

3) Source Separation: Separating one source from a mixture of noisy sources can be carried out using source extraction or source filtering.

a) 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 [147]. 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 [124]:

- 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.

b) 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 [30]

- WL least mean squares to minimize MSE [124]

- WL least stochastic entropy algorithm [124]

- Complex augmented least-mean kurtosis algorithm [241]

- Incremental augmented complex LMS [14]

- Complex-valued Gaussian sum filter [117]

   c) 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.

VI. 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 [218]. 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. 10.

A. 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 between IGS capacity and the envelop of capacity
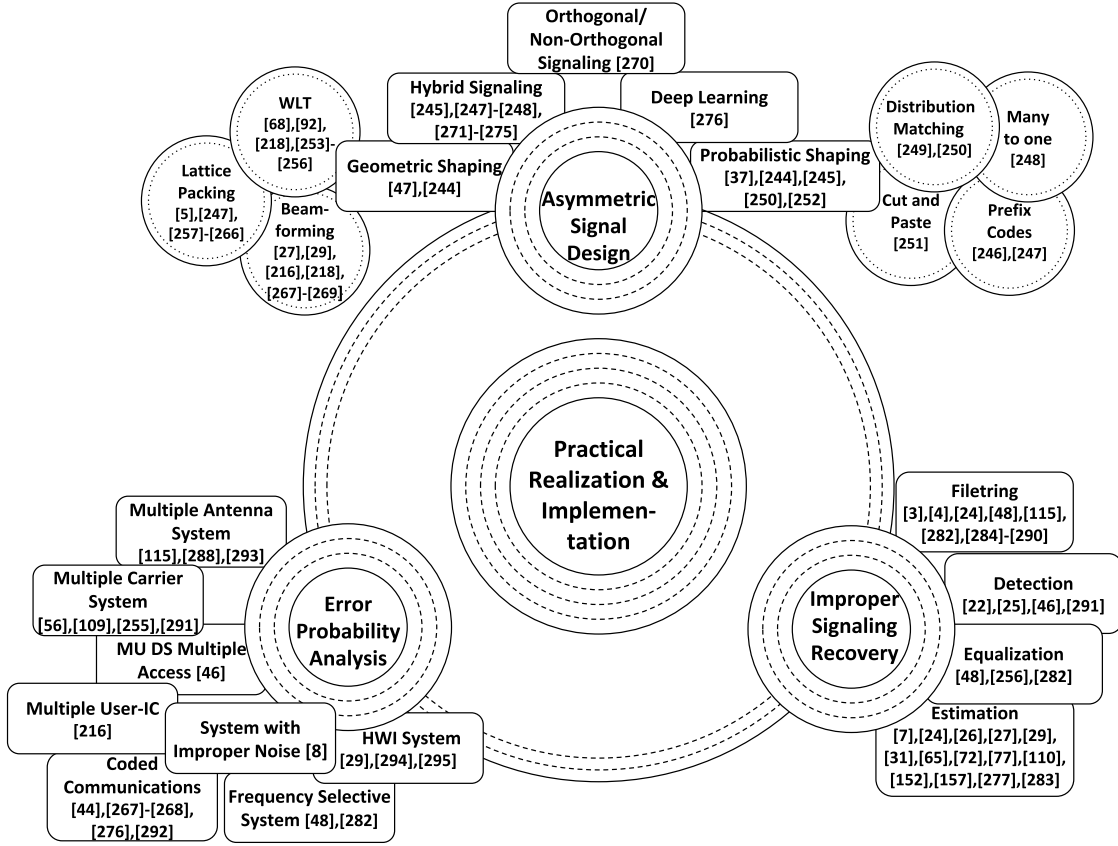
Fig. 10: Practical Implementation of IS: From Realization to Recovery and Analysis

curves with sufficiently large M-ary constellations [218]. 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 [44]. 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 [88], [242], [243]. In the following, we highlight the famous shaping techniques and procedures to design appropriate asymmetric signaling for a given application.

1) Probabilistic Shaping: Given a fixed number of symbols and the symbol locations, an asymmetric constellation can be obtained by adjusting the symbol probabilities [244]. Therefore, PS maps equally distributed input bits into constellation symbols with non-uniform distribution [245]. 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 [246], [247], many-to-one mappings combined with a turbo code [248], distribution matching [249], [250] and cut-and-paste method [251]. Coded modulation scheme with PS aims to remove the shaping gap and coarse mode granularity problems [250]. Interested reader can read [37] 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 [252].

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 [47]. Nonetheless, moving some symbols close to each other may result in more erroneous symbol decisions depending on the underlying application [244]. GS can be realized using the following well-known methodologies.

a) 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 [253]. 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

$$v = \sqrt{\frac{1}{2}(1+\alpha)}x + \sqrt{\frac{1}{2}(1-\alpha)}\exp^{i\phi}x^*, \tag{67}$$

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 [218]. This transformation 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 [68]. 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 [254]. Likewise, WLP in spatially multiplexed MIMO-FBMC system with OQAM provides lower BER as compared to LP at high SINR [255]. 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 [92]. 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 [256].

b) Lattice Packing: Numerous studies have supported the concept of efficient modulation technique by packing a particular lattice structure in some geometric shape [247], [257]. Lattice packing is a two-fold procedure. Firstly, lattice structure is chosen from the square, rhombic, triangular or hexagonal lattices [258], [259]. A square and rhombic lattice is the periodic arrangement of discrete constellation points/symbols at the corners of the square (as depicted in Fig. 11a) and rhombus shape [258]. 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

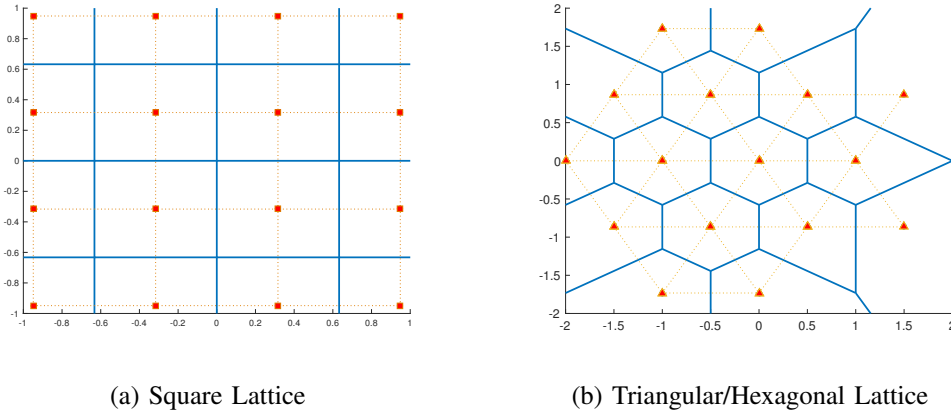(a) Square Lattice   (b) Triangular/Hexagonal Lattice

Fig. 11: Celebrated 16-QAM Lattice Structures

as shown in Fig. 11b [259]. Square QAM (SQAM) is preferred for the simple ML detection mechanism whereas triangular QAM (TQAM) is preferred for power efficiency [259], [260]. 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 [260]. Therefore, hexagonal QAM (HQAM) is preferred for various applications, including advanced channel coding [261], multi-antenna systems [262], multicarrier systems [263], physical-layer network coding [264], and optical communications [265].

Next step is the distribution of lattice around origin, which can be packed in square, cross, elliptical, circular, rectangular, or parallelogram envelops [247], [257], [266]. For instance, Fig. 12a and 12b depict square and cross packing of the square lattice, respectively. However, Fig. 12c-12f 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 [5]. 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 [5].

Recently, Fernández *et al.* suggest elliptical packing of the hexagonal lattice to realize optimal asymmetric constellations for a certain circularity coefficient [257]. The optimal constellation is

(a) Square Packing   (b) Cross Packing   (c) Elliptical Packing

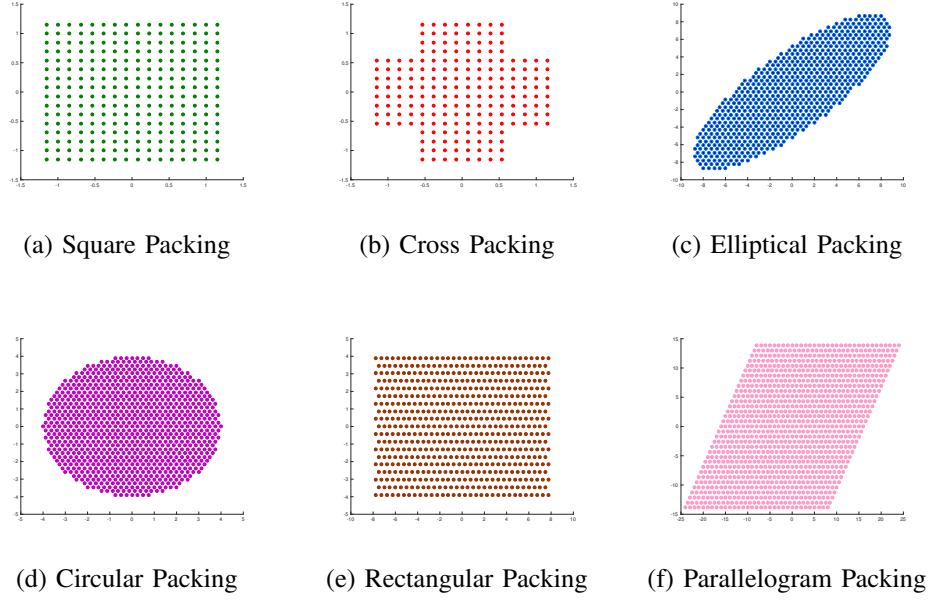(d) Circular Packing   (e) Rectangular Packing   (f) Parallelogram Packing

Fig. 12: Geometrical Packing/Mask Structures

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* [257]. 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 [247]. On the other hand, packing loss is pertinent to the parallelogram envelope of WLT asymmetric constellation in place of the optimal elliptical envelope.

    c) 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 [267]. Similarly, Subramaniam *et al.* introduce asymmetry in 8-PSK to increase the

minimum product distance reducing error events in TCM [268]. For colored noise compensation, Taubock prefers optimally rotated rectangular symbol constellations over common quadratic QAM in order to minimize capacity loss and SEP [27].

Unlike the contributions which propose modulation specific designs, [29] and [216] 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 [216]. 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 [29]. 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 [269].

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 $Q\left(\sqrt{\mathrm{SNR}}d_{\min}/2\right)$ [218].

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 [270].

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 [248]. From one perspective, PS outperforms circular based GS in terms of shaping gain for

the same number of constellation points [245]. However, another point of view is the superior performance of GS 16-QAM constellations over its PS counterpart [271]. 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 [272]. 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 [245]. 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 [273] and multi-dimensional coded modulation format using amplitude-phase shift keying for single-stage [245] and multi-stage [274] 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 [275].

To conclude, dense hexagonal packing with optimal circular boundaries yields around 1dB improvement over PAM in a band-limited channel [247]. 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 [247].

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 [276].

B. 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 [6], filtering [31], and detection [277] algorithms. WL processing is utilized in communication systems that apply asymmetric constellations e.g., [48], [256], [278]–[280] , and/or encounter improper noise, e.g., [45]. Additionally, it is exploited in linear-dispersion STBC, e.g., [277], [281], [282], and, recently, in IC and BC [73], [82].

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 [48]. 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 [282]. 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 [256].

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 [7], [283], communications [24], power systems [157], biomedical engineering [152] and renewable energy [31]. WLE can be carried out using various variants like WL-MMSE estimation [110], 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 [277]. 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 [65]. 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 [72].

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 [284], prediction [3], modeling [4], interference cancellations like co-channel interference (CCI) [285] and narrow-band interference (NBI) and equalization for SISO [24], [48], MIMO [282], and DS-CDMA systems [286], [287]. Asymmetric complex nature of BPSK signals has motivated researchers to apply WL filtering at the RX for improved detection [4]. 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 [115]. 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 [288]. 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 [289]. 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 [24]. Furthermore, Hellings *et al.* suggest block-Hankel-skew-circulant structured matrices for appropriate processing of WL

filters and asymmetric signaling in MIMO IC [290].

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 [46]. Similarly, WL RX for OQPSK modulated DS/CDMA system employing least-mean square algorithm efficiently mitigates both symmetric/asymmetric interference [25]. 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 [22]. 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 [291].

a) 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 [7]. 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 [7].

C. 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 [292]. 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.

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 [46].

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 [276]. 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 [293]. 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 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 [268]. 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 [267]. 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 [44]. 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.

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 $\mathrm{E}\left[\mathbf{ss}^{\mathrm{T}}\right] \neq \mathbf{0}$) 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 [294].

Co-antenna interference suppression in generalized MIMO systems with linear dispersion

TABLE VIII: 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	[115]
Multicarrier Transmission (UW-OFDM)	M-ASK modulation with UW-OFDM	WL-MMSE	WL-MMSE versus LMMSE	AWGN: 97.4% IEEE indoor channel: 94.5%	1dB at 2e-5 BER 3dB at 1e-6 BER	[109]
2-16 states trellis coding	Asymmetric M-PSK,M-AM modulation	Viterbi decoder	Asymmetric versus symmetric signaling	4PSK:99.83% 4AM:98.33%	0.5dB at 1e-5 BER 1dB at 1e-5 BER	[44]
DS/SSMA using trellis coding	Asymmetric M-PSK	Viterbi decoder	Asymmetric signaling versus Symmetric signaling	3U: 58.02%	1dB at 1e-10 BER	[293]
4- and 8-state TCM schemes	Asymmetric 8-PSK	Viterbi decoder	ATCM versus STCM to minimize Euclidean distance	Rayleigh: 20% Rician: 76.67% Shadowed Rician: 50%	0.3dB at 1e-5 BER 0.4dB at 1e-5 BER 0.2dB at 1e-5 BER	[268]
2- and 4-state TCM schemes	Asymmetric 4-PSK	Viterbi decoder	ATCM versus STCM for BER bound minimization	2 states: 75% 4 states: 55%	0.5dB at 1e-5 BER 0.1dB at 1e-7 BER	[267]
MU IC	Asymmetric M-PSK or M-QAM	Symbol-by-Symbol detection	Symmetric versus asymmetric for max PEP/SER Minimization	AWGN: 100% Cellular: 99%	- 5dB at 8e-2 max(BER)	[216]
Deep learning for commun. system	Learned coded modulation by auto-encoder	Neural networks	Auto-encoder versus conventional modulation/ coding to Min BLER	Uncoded: 98.64% Coded:96.25%	0.7dB at 1e-3 BLER 1dB at 1e-3 BLER	[276]
Phase estimation without pilot	Asymmetric 8-PSK	Coherent detection	Asymmetric 8-PSK versus symmetric 8-DPSK	97.50%	3dB at 1e-2 SEP	[244]
Uncoded MIMO systems	M-ASK, QPSK/OQPSK	ZF RX	Iterative RX with asymmetric versus symmetric modulation	ASK: 99.25% OQPSK: 99.58%	8dB at 2e-3 BER 12dB at 1.5e-3 BER	[294]
		MMSE RX		ASK: 97.78% OQPSK: 97.50%	6dB at 2e-4 BER 8.5dB at 3.5e-5 BER	[294]
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	[295]
HWI system	Asymmetric GPSK, QAM		Asymmetric versus symmetric modulation to minimize max PEP/SEP	10dB SNR: 84% 20dB SNR: 97.8%	10dB at 8e-2 SER	[29]
	Grey coded M-QAM modulation		Optimal MLD with IGN versus traditional MLD with PGN	$\sigma_t^2 = 0.001$: 8.33% $\sigma_t^2 = 0.010$: 38.64% $\sigma_t^2 = 0.100$: 52.78%	10dB at 1.2e-1 BER 9dB at 4.4e-2 BER 4dB at 1.7e-3 BER	[296]
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	[8]
MIMO with data like CCI and ISI in FS channel	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	[56]
				w/ CCI: 97.10% w/o CCI: 99.75%	20dB at 6.2e-2 BER 8dB at 1e-5 BER	[56]

				Desired CCI QPSK-BPSK: 83.53%		
SIMO with multiple CCIs	PSK data modulation and PAM/QAM CCI	WL ML RX	WL RX versus conventional RX	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	[288]
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	[255]
		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	[291]
DS/CDMA systems	QPSK modulation	MMSE MUD	L-MUD versus WL-MUD	MF: 100% L-MUD: 96%	- 2dB at 1e-5BER	[46]
Frequency selective channels	ASK, OQAM or BMSK	WL-MMSE	LE versus WLE with or without decision feedback	MMSE-LE: 96.07% MMSE-DFE: 56%	5dB at 1e-4BER 0.7dB at 1e-4BER	[48]
	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	[282]

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 [115].

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 $M_1 + M_2 < N$ [288]. In a nutshell, WL RX can handle more CCIs relative to SL RX while maintaining a certain SER.

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 [109]. 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 [255]. 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 [291]. 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) attributes 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 [56].

5) Systems with Frequency Selective Channels: equalization schemes based on WL processing outperform their linear counterparts in frequency selective channels 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 [48]. 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 [282].

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 improper 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 [8].

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 [296]. 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 [29]. 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

[295].

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) [216]. 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 [216].

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 VIII.

VII. 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 [6], [7], [42], [82]. Intuitively, asymmetric signaling and WL processing are the concepts that go hand in hand [23] and are beneficial in the various settings as demonstrated in Fig. 13.

A. 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 [40], [297], [298]. Other miscellaneous data analysis techniques with impropriety incorporation successfully categorize seismic, oceanographic and weather data sets [7], [299], [300].
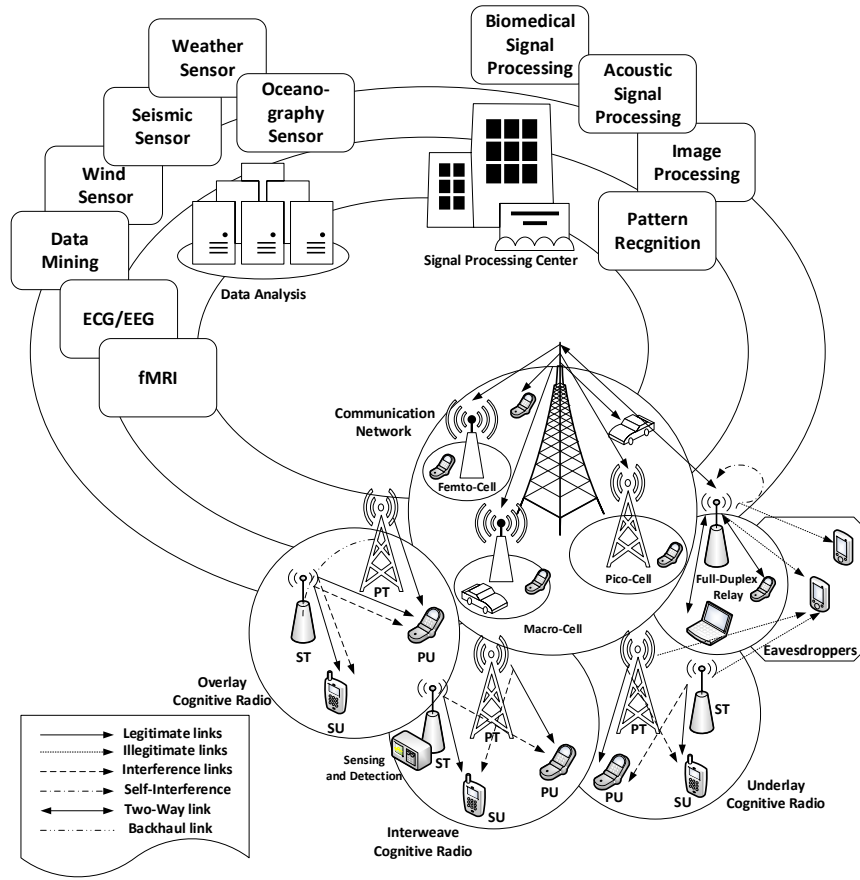
Fig. 13: Applications of Impropriety in Data Analysis, Signal Processing and Communication Networks

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 [40]. Many complex ICA algorithms either assume that the underlying sources are circular or rely on a magnitude-only model. However, this greatly limits the performance of ICA [20]. Thus, NC characterization enabled ICA has found real-world applications in various diverse fields such as medicine, economics, and data mining [146], [297], [301], [302].

a) 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 [297]. However, fMRI poses tremendous challenges for data analysis techniques, including the design of robust yet flexible framework to capture the richness of human brain activities

[301]. Adaptive noncircular ICA algorithms can effectively address these challenges rendering vast applications [40], [146]. For instance, feature extraction in electrocardiograms and fMRI data analysis lead to the improved neural activity estimation [6], [20], [21]. Moreover, real-time brain-computer interfacing relies on the extraction of eye muscle activity: electrooculogram (EOG) from electroencephalogram (EEG) recordings [147]. 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) [303].

b) 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 [297]. 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 [40], [302].

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 [297]. PCA is a classical technique in statistical data analysis for pattern recognition [304]–[307], feature extraction [308], [309], data compression [310], [311], data reduction [312]–[314], data visualization [315], noise reduction [316], [317], factor analysis [318], [319], model selection [298], rank reduction [4], dimensionality reduction [320], 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 [2]. 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) [321].
- The augmented covariance matrices yielding WL PCA (wlPCA) for direct extension from the cPCA [322].
- The real-composite covariance matrices resulting composite real PCA (crPCA) [320].

Aforementioned forms of general complex PCA exhibit their own merits and demerits based on the underlying applications and their linear/non-linear models.

a) 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 [4]. Interestingly, crPCA demonstrates its superiority in dimensionality reduction owing to the finer granularity and lower computational complexity than that of wlPCA [320]. 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 [312], in order selection for complex NC fMRI data [301], in data analysis for complementary DNA microarray experiments [313] and for genome data analysis in bioinformatics [314]. PCA offers numerous other applications including data visualization for biplot graphic display of matrices [315] and data compression of meteorological parameters obtained from high-resolution infrared spectra [311].

b) 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 [308] and fast iris recognition [309], respectively. Moreover, PCA is inevitably useful in pattern recognition like handwritten digits recognition [304], face recognition [306] and automated diagnosis of cardiac health using principal components of segmented ECG beats [307].

c) 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 [298]. ncPCA is preferred for scenarios with high SNR, large number of samples and high degree of noncircularity. On the other hand, non-linear component analysis requires WL PCA with complex kernel (wlPCA-ck) [323] for the design of digital filters and regression frameworks [324]. Furthermore, statistical PCA ia also employed for data denoising especially for image denosing on multi-exponential MRI relaxometry [316] and signal denoising in stock market trading [317].

Conclusively, non-circularity exploitation in complex PCA is particularly significant when the

underlying entities are improper.

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 applications in system identification [283], real-time impropriety detection [103], communications [325], signal processing [326], renewable energy [327], power systems [18], [328], and medicine [329].

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 [283]. Moreover, collaborative adaptive filters trained by the CLMS can detect and track improperness in real-time unlike competing static detectors [103]. Additionally, multiple access interference in DS-CDMA systems can be efficiently supressed using WL LMS algorithms [325]. Interestingly, ACLMS algorithm renders lower steady-state mean-squared error than conventional CLMS in adaptive beamforming for multi-port antenna arrays [326], [330]. Furthermore, ACLMS usefulness in wind modelling and forecasting is unprecedented in renewable energy domain [327], [331]. Similarly, the approximate uncorrelating transform improved adaptive frequency estimations using ACLMS in three-phase power grid systems [328]. 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 [329].

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 [299], wind measurements [7], [14], and oceanographic velocity measurements [332].

WL complex auto-regressive processing of the seismic signals helps to capture essential data characteristics like elliptical oscillations [150]. Other climate and seismology applications with improper complex-valued stochastic models can be efficiently simulated using circulant embedding [333]. Compressive sensing of weather sensor network application can effectively exploit the asymmetrical features for energy estimation to cool down a given structure [300].

B. Signal Processing

Augmented and WL processing have demonstrated remarkable performance gains in different signal processing domains, e.g., signal estimation [6], [334], filtering [31], and detection

[277]. Therefore, it is a leading competitor rendering vast applications in neuro-science, image processing, pattern recognition, and computer vision [253], [297].

1) Array Processing: The merits of impropriety adaptation in array processing algorithms are widely studied in [99], [335]–[338]. For example, coherent processing (incorporating complementary covariance) for detection and estimation enjoys a 3dB gain over non-coherent processing [32]. Similarly, estimation accuracy is substantially enhanced by employing NC signal constellations in 1-D and 2-D DoA estimation [159]. Furthermore, such enhanced DoA estimation and identification methods for mixed circular and NC sources also improve the resolution capacity [339].

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 [283], [334]. 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 [40], [124], [155], [340]. 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 [297].

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 [297]. 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 [253]. Similarly, target detection in multi-band spectral images suffering from improper Gaussian noise is only possible with the advancement in impropriety literature [113].

C. 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.

1) Cellular Networks: Improper/asymmetric signaling techniques are implemented in GSM [341], [342] and 3GPP networks [343], [344]. Similarly, in mobile MU communications, non-circularity characterization can render better tradeoff between power consumption and spectral efficiency [7]

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 [26]. 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 [53]. 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 [54]. Lastly, interweave setting utilizes unused spectrum holes for its transmission as long as it is available [203]. 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 [52] and overlay cognitive radio paradigm [54]. Moreover, for underlay MAC, IGS is optimal if the accumulative IC gains exceed a certain threshold [50]. On the other hand, with adequate detection capabilities SU can employ IGS with maximum power to unconditionally improve its rate while satisfying PU QoS [203]. 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 [51].

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 [49] or without [60] spectrum sharing, SISO [89] and MIMO [93],

[345] full duplex relaying. as well as heterogeneous multi-tier network involving cellular and D2D full duplex communications [97]. 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 [59] and asymmetric hardware distortions [64]. Besides full-duplex relaying, WL processing is also rewarding for other relay networks including two-way AF-MIMO relayed MU systems [92] and multi-layered relay systems [91]. Surprisingly, IGS is also favorable for the alternate relaying systems which mimic as full-duplex systems [202].

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 [63] and IQI in space-time coded transmit diversity systems [346]. 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 [65], WL precoding for Large scale MIMO [66], WL beamforming of linear antenna arrays [347], and massive antenna arrays [68], and circularity based compensators for wideband direct-conversion RXs [67] and OFDM based WLAN transmitters [69]. Improper transmission is also recognized to jointly compensate multiple HWIs such as additive hardware distortions and transceiver IQI in single antenna [29] and multi-antenna systems [9]. Interestingly, impairments in I/Q modulators are also accurately modelled using widely non-linear model using compressed sensing [197]. 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 [348].

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 [244].

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 [77], [349] counter intuitive to the PGS and linear filtering optimality in P2P communications [350]. Performance gains are also reported for a variety of interference-limited settings spanning SISO Z-IC [86], [88], MIMO Z-IC [87], MIMO P2P [78], [85], MU SISO X-IC [79], [80], [82], [96], [201], [216], MU MISO IC [83], [204], MU MIMO IC [81], [84], MIMO-BC [72], [73], [204], [217], MIMO-IBC [70], [71], cognitive MAC with primary P2P [50], cognitive P2P with primary MAC [75], MU diversity systems [206], multicarrier systems [280], [291], [351], multi-antenna systems [95], and single-/multi-antenna NC interference cancellation [24], [289]. Other forms of interference i.e., co-channel interference [56], [288], intra- and intersystem multiple access interference in radio navigation satellite services [352], and wideband multiple access and narrow band interference in CDMA systems [22], [46] 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 [87]. Additionally, IGS with symbol extension can outperform PGS for interference alignment within the context of linear precoding schemes where all interference is treated as noise [77], [119].

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 [353]. 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 [45], [116], discrete multitoned systems [27] and spectral image target detection [113]. Asymmetric noise characterization is necessary for appropriate estimation [111], detection [38], [62], [114], filtering [117], processing [42], [242], and compensation [8].

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 [44]. 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 [48]. 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
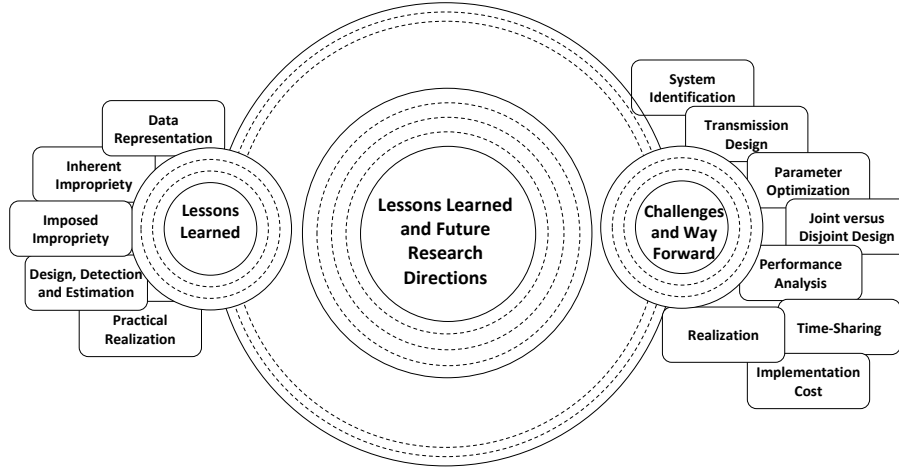
Fig. 14: Lessons Learned and Way Forward

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 [44]. Similar studies aim to design TCM parameters not only to increase the effective length but also the minimum 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 [268]. 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 [267]. Equivalently, asymmetric signal constellation with Trellis coding outperforms conventional symmetric signaling in terms of BER in DS/SSMA systems [293].

VIII. Lessons Learned and Future Research Directions

This section summarizes the valuable lessons learned throughout this survey. It also highlights 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. 14.

A. Lessons Learned

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.

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 (39)-(49).

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.

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.

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.

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 asym-

metric 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.

B. 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.

1) System Identification: Improper/asymmetric signaling techniques are already implemented in GSM and 3GPP networks [82]. 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 system 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 [23], [354]. Similarly, IGS in combination with WL transceivers is beneficial in MIMO-BC under specific scenarios which are not yet understood [355]. Moreover, the perks of IGS over PGS for SU transmission are conditional in underlay and overlay cognitive setups [52], [54]. On the other hand, IGS is the all time favorite in other interference limited, hardware impaired or improper/non-circular noise based systems [61], [63], [65], [81] 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 [82]. Similarly, circular models are favored with small number of samples, low signal-to-noise ratio, or minimal degree of non-circularity [124].

2) 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.

3) 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 [86]. 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], [204]. Similarly, maximal IGS is adopted owing to the intractable optimization in MIMO hardware impaired systems [9]. 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 [86]. Consequently, the search for least-complex near-to-optimal optimization strategy is an open research area.

4) 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? [215]. Future research may address the limitation arising from the imperfection or lack of instantaneous/average system parameters while tuning improper transmission parameters.

5) 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 [29]. 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 [267]. Thus, tools are inevitably required to derive exact EP based on accurately estimated parameters.

6) Time-Sharing (TS): Most of the studies focused on the perks of improper/asymmetric signaling in the absence of 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 [211]. 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 [356]. 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 [357]. Therefore, the superiority examination of IGS in a TS context under certain assumptions is an open research area in various interference-limited systems [211], [355].

7) 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 [44]. 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 [29]. Similarly, asymmetric model of the aggregate HWIs in a wireless communication system will help in accurate system analysis and design [9]. 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 [358], [359]. Efficient buffering and queue management schemes are inevitably required to tackle the latency issues in delay-sensitive applications [360]. Consequently, the appropriate realization of the system containing asymmetric signatures is immensely important to achieve the target performance.

8) Implementation Cost: The existing infrastructure employs SL transceivers, and thus the up-

gradation to WL transceivers is a sequential process [84]. 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 [44], 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.

IX. CONCLUSION

The journey from proper signaling to improper signaling and then from improper Gaussian to asymmetric discrete constellation is captured and summarized in this article. 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] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[2] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[3] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[4] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[5] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, Nov. 2008.

[6] 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.

[7] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[8] 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.

[9] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Multiple antenna systems with hardware impairments: New performance limits," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1593–1606, Feb. 2019.

[10] 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.

[11] 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.

[12] M. T. Taner, F. Koehler, and R. Sheriff, "Complex seismic trace analysis," *Geophysics*, vol. 44, no. 6, pp. 1041–1063, Jun. 1979.

[13] 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.

[14] 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.

[15] M. I. Knight and M. A. Nunes, "Long memory estimation for complex-valued time series," *Statistics and Computing*, pp. 1–20, Jul. 2018.

[16] 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.

[17] A. M. Sykulski, S. C. Olhede, J. M. Lilly, and J. J. Early, "Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3136–3151, Jun. 2017.

[18] Y. Xia and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 74–83, Jan. 2012.

[19] D. H. Dini and D. P. Mandic, "Widely linear modeling for frequency estimation in unbalanced three-phase power systems," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 2, pp. 353–363, Feb. 2013.

[20] D. B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *Neuroimage*, vol. 25, no. 4, pp. 1310–1324, May 2005.

[21] 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.

[22] 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," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 154–167, Jan. 2001.

[23] C. Hellings and W. Utschick, "Block-skew-circulant matrices in complex-valued signal processing," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2093–2107, Apr. 2015.

[24] 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," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 870–883, Mar. 2006.

[25] A. Mirbagheri, K. N. Plataniotis, and S. Pasupathy, "An enhanced widely linear CDMA receiver with OQPSK modulation," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 261–272, Feb. 2006.

[26] 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.

[27] G. Taubock, "Complex noise analysis of DMT," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5739–5754, Dec. 2007.

[28] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451 – 1458, Oct. 1998.

[29] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric modulation for hardware impaired systems-Error probability analysis and receiver design," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1723–1738, Mar. 2019.

[30] 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.

[31] 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.

[32] P. J. Schreier, L. L. Scharf, and C. T. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.

[33] A. T. Walden and P. Rubin-Delanchy, "On testing for impropriety of complex-valued Gaussian vectors," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 825–834, Mar. 2009.

[34] A. Mohammadi and K. N. Plataniotis, "Improper complex-valued multiple-model adaptive estimation." *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1528–1542, Mar. 2015.

[35] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[36] 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.

[37] B. Dunbridge, "Asymmetric signal design for the coherent Gaussian channel," *IEEE Trans. Inf. Theory*, vol. 13, no. 3, pp. 422–431, Jul. 1967.

[38] S. Kassam, G. Moustakides, and J. Shin, "Robust detection of known signals in asymmetric noise," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 84–91, Jan. 1982.

[39] A. van den Bos, "The multivariate complex normal distribution-A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

[40] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[41] 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," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2859–2867, Jun. 2009.

[42] G. Taubock, "Complex-valued random vectors and channels: entropy, divergence, and capacity," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2729–2744, May 2012.

[43] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.

[44] D. Divsalar, M. Simon, and J. Yuen, "Trellis coding with asymmetric modulations," *IEEE Trans. Commun.*, vol. 35, no. 2, pp. 130–141, Feb. 1987.

[45] Y. C. Yoon and H. Leib, "Maximizing SNR in improper complex noise and applications to CDMA," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 5–8, Jan. 1997.

[46] G. Gelli, L. Paura, and A. R. Ragozini, "Blind widely linear multiuser detection," *IEEE Commun. Lett.*, vol. 4, no. 6, pp. 187–189, Jun. 2000.

[47] 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," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 774–786, May 2000.

[48] H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1512–1523, Sep. 2003.

[49] M. Gaafar, O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay spectrum sharing techniques with in-band full-duplex systems using improper Gaussian signaling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 235–249, Jan. 2017.

[50] C. Lameiro, I. Santamaría, and P. J. Schreier, "Improper Gaussian signaling for multiple-access channels in underlay cognitive radio," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1817–1830, Mar. 2019.

[51] G. Oliveira, E. Fernandez, S. Mafra, and S. Montejo-Sánchez, "Physical layer security in cognitive radio networks using improper Gaussian signaling," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1886–1889, Sep. 2018.

[52] C. Lameiro, I. Santamaría, and P. J. Schreier, "Benefits of improper signaling for underlay cognitive radio," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 22–25, Feb. 2015.

[53] O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay cognitive radio systems with improper Gaussian signaling: Outage performance analysis," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4875–4887, Jul. 2016.

[54] ——, "Overlay spectrum sharing using improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 50–62, Jan. 2017.

[55] 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.

[56] M. Kim, J. H. Cho, and J. S. Lehnert, "Asymptotically optimal low-complexity SC-FDE with noise prediction in data-like improper-complex interference," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2090–2103, Mar. 2016.

[57] 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.

[58] 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.

[59] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[60] 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.

[61] C. Hellings, "Reduced-entropy signals in MIMO communication systems," Ph.D. dissertation, Technische Universität München, 2017.

[62] 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.

[63] 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.

[64] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, "Improper Gaussian signaling for hardware impaired multihop full-duplex relaying systems," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1858–1871, Mar. 2019.

[65] 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," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3393–3408, May 2016.

[66] 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," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3298–3312, May 2017.

[67] L. Anttila, M. Valkama, and M. Renfors, "Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2099–2113, Jul. 2008.

[68] 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.

[69] 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," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 383–393, Mar. 2017.

[70] 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.

[71] 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," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 136–151, Jan. 2019.

[72] 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.

[73] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[74] 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.

[75] 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.

[76] ——, "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.

[77] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signalingSettling the Høst-Madsen–Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

[78] L. Yang and W. Zhang, "Interference alignment with asymmetric complex signaling on MIMO X channels," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3560–3570, Oct. 2014.

[79] 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.

[80] 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.

[81] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[82] Z. K. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3194–3203, Sep. 2012.

[83] 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.

[84] S. Lagen, A. Agustin, and J. Vidal, "Coexisting linear and widely linear transceivers in the MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 652–664, Feb. 2016.

[85] ——, "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.

[86] E. Kurniawan and S. Sun, "Improper Gaussian signaling scheme for the Z-interference channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3912–3923, Jul. 2015.

[87] S. Lagén Morancho, A. Agustín de Dios, and J. Vidal Manzano, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3350–3368, Aug. 2016.

[88] C. Lameiro, I. Santamaría, and P. J. Schreier, "Rate region boundary of the SISO Z-interference channel with improper signaling," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1022–1034, Mar. 2017.

[89] 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," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 64–78, Jan. 2018.

[90] 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.

[91] 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.

[92] 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.

[93] 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.

[94] 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.

[95] 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.

[96] H. Park, S.-H. Park, J.-S. Kim, and I. Lee, "SINR balancing techniques in coordinated multi-cell downlink systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 626–635, Feb. 2013.

[97] A. Kariminezhad and A. Sezgin, "Heterogeneous multi-tier networks: Improper signaling for joint rate-energy optimization," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 680–694, Jan. 2019.

[98] M. Haardt and F. Romer, "Enhancements of unitary ESPRIT for non-circular sources," in *Intern. Conf. Acoustics, Speech, and Signal Process.(ICASSP)*, vol. 2. Montreal, Que., Canada: IEEE, May 2004.

[99] 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.

[100] H. Abeida and J.-P. Delmas, "Statistical performance of MUSIC-like algorithms in resolving noncircular sources," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4317–4329, Sep. 2008.

[101] A. Napolitano and M. Tanda, "Doppler-channel blind identification for noncircular transmissions in multiple-access systems," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2073–2078, Dec. 2004.

[102] 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.

[103] 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.

[104] H. Li and T. Adali, "A class of complex ICA algorithms based on the kurtosis cost function," *IEEE Trans. Neural Netw.*, vol. 19, no. 3, pp. 408–420, Mar. 2008.

[105] S. C. Douglas, "Fixed-point algorithms for the blind separation of arbitrary complex-valued non-Gaussian signal mixtures," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 83–83, Dec. 2007.

[106] M. Novey and T. Adali, "On extending the complex FastICA algorithm to noncircular sources," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 2148–2154, May 2008.

[107] 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.

[108] Y. Xia and D. Mandic, "Augmented performance bounds on strictly linear and widely linear estimators with complex data," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 507–514, Jan. 2018.

[109] S. Trampitsch, "Complex-valued data estimation," Ph.D. dissertation, Master Thesis. Alpen-Adria University of Klagenfurt, Apr. 2013.

[110] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[111] 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.

[112] M. Novey, T. Adali, and A. Roy, "Circularity and Gaussianity detection using the complex generalized Gaussian distribution," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 993–996, Nov. 2009.

[113] 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.

[114] 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.

[115] 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.

[116] 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.

[117] A. Mohammadi and K. N. Plataniotis, "Complex-valued Gaussian sum filter for nonlinear filtering of non-Gaussian/non-circular noise," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 440–444, Apr. 2015.

[118] A. K. Jagannatham and B. D. Rao, "Cramer-rao lower bound for constrained complex parameters," *IEEE Signal Process. Lett.*, vol. 11, no. 11, pp. 875–878, Nov. 2004.

[119] 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.

[120] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5597–5625, Nov. 2012.

[121] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Process. Mag.*, vol. 29, no. 4, pp. 61–80, Jul. 2012.

[122] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust statistics for signal processing*.   Cambridge University Press, 2018.

[123] E. Moreau and T. Adali, *Blind identification and separation of complex-valued signals*.   Wiley Online Library, 2013.

[124] T. Adali and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 112–128, Sep. 2014.

[125] 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.

[126] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, Sep. 2015.

[127] W. H. Hayt, J. E. Kemmerly, and S. M. Durbin, *Engineering circuit analysis*.   McGraw-Hill New York, 1986.

[128] J. G. Proakis and M. Salehi, *Digital communications*.   McGraw-Hill, 2008.

[129] P. Rubin-Delanchy and A. T. Walden, "Kinematics of complex-valued time series," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4189–4198, Sep. 2008.

[130] 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.

[131] S. C. Olhede and A. T. Walden, "Noise reduction in directional signals using multiple morse wavelets illustrated on quadrature Doppler ultrasound," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 1, pp. 51–57, Jan. 2003.

[132] L. Fortuna, G. Muscato, and M. G. Xibilia, "A comparison between HMLP and HRBF for attitude control," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 318–328, Mar. 2001.

[133] A. J. Hanson, *Visualizing Quaternions*.   San Francisco: The Morgan Kaufmann Series in Interactive 3D Technology, 2006.

[134] S. Miron, N. Le Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.

[135] T. Bulow and G. Sommer, "Hypercomplex signals-a novel extension of the analytic signal to the multidimensional case," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2844–2852, Nov. 2001.

[136] J. Seberry, K. Finlayson, S. S. Adams, T. A. Wysocki, T. Xia, and B. J. Wysocki, "The theory of quaternion orthogonal designs," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 256–265, Jan. 2008.

[137] 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.

[138] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.

[139] J. Navarro-Moreno, R. Marı, J. C. Ruiz-Molina *et al.*, "A quaternion widely linear model for nonlinear Gaussian estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6414–6424, Dec. 2014.

[140] M. Xiang, S. Kanna, and D. P. Mandic, "Performance analysis of Quaternion-valued adaptive filters in nonstationary environments," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1566–1579, Mar. 2018.

[141] R. Krupiński, "Generating augmented Quaternion random variable with generalized Gaussian distribution," *IEEE Access*, vol. 6, pp. 34 608–34 615, Jul. 2018.

[142] 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.

[143] K. Daniilidis, "Hand-eye calibration using dual quaternions," *Int. J. of Robotics Research*, vol. 18, no. 3, pp. 286–298, Mar. 1999.

[144] 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.

[145] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the Kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, no. 2, pp. 248–261, Feb. 2010.

[146] 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," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2794–2803, Oct. 2011.

[147] S. Javidi, D. P. Mandic, and A. Cichocki, "Complex blind source extraction from noisy mixtures using second-order statistics," *IEEE Trans. Circuits Syst.*, vol. 57, no. 7, pp. 1404–1416, Jul. 2010.

[148] 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.

[149] 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.

[150] A. M. Sykulski, S. C. Olhede, and J. M. Lilly, "A widely linear complex autoregressive process of order one," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6200–6210, Dec. 2016.

[151] T. Adalı and H. Li, "Complex-valued adaptive signal processing," 2010.

[152] 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," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 74–83, Jan. 2011.

[153] H. P. Yuen, "Two-photon coherent states of the radiation field," *Physical Review A*, vol. 13, no. 6, p. 2226, 1976.

[154] B. I. Erkmen and J. H. Shapiro, "Phase-conjugate optical coherence tomography," *Physical Review A*, vol. 74, no. 4, p. 041601, 2006.

[155] 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.

[156] 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.

[157] Y. Xia, S. C. Douglas, and D. P. Mandic, "Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 44–54, Sep. 2012.

[158] 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.

[159] 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.

[160] M. Mokhtar, A. Gomaa, and N. Al-Dhahir, "OFDM AF relaying under I/Q imbalance: Performance analysis and baseband compensation," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1304–1313, Apr. 2013.

[161] J. Li, M. Matthaiou, and T. Svensson, "I/Q imbalance in AF dual-hop relaying: Performance analysis in nakagami-m fading," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 836–847, Mar. 2014.

[162] E. Ollila, J. Eriksson, and V. Koivunen, "Complex elliptically symmetric random variablesgeneration, characterization, and circularity tests," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 58–69, Jan. 2011.

[163] W. R. Hamilton, "Theory of quaternions," *Proc. Royal Irish Academy (1836-1869)*, vol. 3, pp. 1–16, Nov. 1844.

[164] F. Zhang, "Quaternions and matrices of quaternions," *Linear algebra and its appl.*, vol. 251, pp. 21–57, Jan. 1997.

[165] H. Coxeter, "Quaternions and reflections," *The American mathematical monthly*, vol. 53, no. 3, pp. 136–146, 1946.

[166] P. Ginzberg and A. T. Walden, "Testing for quaternion propriety," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3025–3034, Jul. 2011.

[167] 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.

[168] J. Via, D. P. Palomar, and L. Vielva, "Generalized likelihood ratios for testing the properness of quaternion Gaussian vectors," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1356–1370, Apr. 2011.

[169] P. J. Schreier, "Bounds on the degree of impropriety of complex random vectors," *IEEE Signal Process. Lett.*, vol. 15, pp. 190–193, Jan. 2008.

[170] K. Fang, S. Kotz, and K. Ng, *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall, 1990.

[171] H. Abeida and J.-P. Delmas, "MUSIC-like estimation of direction of arrival for noncircular sources," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2678–2690, Jul. 2006.

[172] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 433–436, Jul. 2006.

[173] 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.

[174] N. N. Vakhania, "Random vectors with values in quaternion hilbert spaces," *Theory of Probability & Its Applications*, vol. 43, no. 1, pp. 99–115, 1999.

[175] T. M. Cover and J. A. Thomas, *Elements of information theory*. Hoboken, NJ, USA : John Wiley & Sons, 2006.

[176] L. Zetterberg and H. Brandstrom, "Codes for combined phase and amplitude modulated signals in a four-dimensional space," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 943–950, Sep. 1977.

[177] W. McGill, "Multivariate information transmission," *IEEE Trans. Inf. Theory*, vol. 4, no. 4, pp. 93–111, Sep. 1954.

[178] R. Fisher, *Precoding and Signal Shaping for Digital Communications*. John Wiley & Sons, New York, 2002.

[179] 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.

[180] 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.

[181] S. Chandna and A. T. Walden, "A frequency domain test for propriety of complex-valued vector time series." *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1425–1436, Mar. 2017.

[182] J. K. Tugnait and S. A. Bhaskar, "On testing for impropriety of multivariate complex-valued random sequences," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2988–3003, Jun. 2017.

[183] 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.

[184] 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.

[185] 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.

[186] 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.

[187] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis.* New York: Academic, 1979.

[188] S. M. Kay and J. R. Gabriel, "An invariance property of the generalized likelihood ratio test," *IEEE Signal Process. Lett.*, vol. 10, no. 12, pp. 352–355, Dec. 2003.

[189] M. Novey, E. Ollila, and T. Adali, "On testing the extent of noncircularity," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5632–5637, Nov. 2011.

[190] P. J. Schreier and L. L. Scharf, "Stochastic time-frequency analysis using the analytic signal: Why the complementary distribution matters," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3071–3079, Dec. 2003.

[191] R. Krishnan, *On the Impact of Phase Noise in Communication Systems—Performance Analysis and Algorithms.* Chalmers University of Technology, Apr. 2015.

[192] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation.* Springer Science & Business Media, 2008.

[193] E. Bjornson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying:Performance limits with hardware impairments," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4512–4525, Nov. 2013.

[194] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6565–6580, Dec. 2015.

[195] 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," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1594–1606, May 2015.

[196] A.-A. A. Boulogeorgos, V. M. Kapinas, R. Schober, and G. K. Karagiannidis, "I/Q-imbalance self-interference coordination," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4157–4170, June 2016.

[197] 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.

[198] 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.

[199] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Improper signaling for SISO two-user interference channels with additive asymmetric hardware distortion," *IEEE Trans. Commun.*, Sep. 2019.

[200] 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.

[201] 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.

[202] M. Gaafar, O. Amin, A. Ikhlef, A. Chaaban, and M.-S. Alouini, "On alternate relaying with improper Gaussian signaling," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1683–1686, Aug. 2016.

[203] 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.

[204] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper Gaussian signaling in the k-user MISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6303–6313, Dec. 2013.

[205] 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.

[206] H.-Y. Shin, S.-H. Park, H. Park, and I. Lee, "A new approach of interference alignment through asymmetric complex signaling and multiuser diversity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 880–884, Mar. 2012.

[207] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, Mar. 2009.

[208] 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.

[209] 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.

[210] 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.

[211] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Improper Gaussian signaling for broadcast interference networks," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 808–812, Jun. 2019.

[212] 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.

[213] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Robust improper signaling for two-user SISO interference channels," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4709–4723, Jul. 2019.

[214] I. A. Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-rate maximization of NOMA systems under imperfect successive interference cancellation," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 474–477, Mar. 2019.

[215] 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.

[216] H. D. Nguyen, R. Zhang, and S. Sun, "Improper signaling for symbol error rate minimization in k-user interference channel," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 857–869, Mar. 2015.

[217] 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.

[218] 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.

[219] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[220] 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.

[221] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1142–1157, Mar. 2011.

[222] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," vol. 2, no. 1, pp. 57–73, 2008.

[223] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[224] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[225] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[226] 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," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[227] W. Utschick and J. Brehmer, "Monotonic optimization framework for coordinated beamforming in multicell networks," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1899–1909, Apr. 2012.

[228] L. P. Qian, Y. J. Zhang, and J. Huang, "MAPEL: Achieving global optimality for a non-convex wireless power control problem," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.

[229] J. L. Gross and J. Yellen, *Handbook of graph theory*. CRC press, 2004.

[230] 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.

[231] 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.

[232] 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.

[233] C. Lameiro, I. Santamaría, P. J. Schreier, and W. Utschick, "Maximally improper signaling in underlay MIMO cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6241–6255, 2019.

[234] T. Ando, "Majorizations and inequalities in matrix theory," *Linear algebra and its Appl.*, vol. 199, no. 1, pp. 17–67, Mar. 1994.

[235] M. Novey, T. Adali, and A. Roy, "A complex generalized Gaussian distributioncharacterization, generation, and estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1427–1433, Mar. 2010.

[236] 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.

[237] 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.

[238] 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.

[239] 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.

[240] J. Navarro-Moreno, J. Moreno-Kayser, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Widely linear estimation algorithms for second-order stationary signals," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4930–4935, Dec. 2009.

[241] E. C. Mengüç and N. Acır, "An augmented complex-valued least-mean kurtosis algorithm for the filtering of noncircular signals," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 438–448, Jan. 2018.

[242] B. W. Han and J. H. Cho, "Capacity of second-order cyclostationary complex Gaussian noise channels," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 89–100, Jan. 2012.

[243] 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.

[244] 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.

[245] 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.

116

[246] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993.

[247] G. Forney, R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 5, pp. 632–647, Sep. 1984.

[248] 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.

[249] 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.

[250] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[251] 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.

[252] 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.

[253] 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.

[254] R. Zelenovsky and R. C. de Lamare, "Set-membership constrained widely linear beamforming algorithms," *Sensor Signal Processing for Defence (SSPD)*, Sep. 2012.

[255] 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.

[256] P. Xiao and M. Sellathurai, "Improved linear transmit processing for single-user and multi-user MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1768–1779, Mar. 2010.

[257] J. A. Lpez-Fernndez, R. G. Ayestarn, I. Santamaria, and C. Lameiro, "Design of asymptotically optimal improper constellations with hexagonal packing," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5445–5457, Aug. 2019.

[258] D. Wubben, D. Seethaler, J. Jalden, and G. Matz, "Lattice reduction," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 70–91, May 2011.

[259] L. Rugini, "Symbol error probability of hexagonal QAM," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1523–1526, Aug. 2016.

[260] S.-J. Park, "Performance analysis of triangular quadrature amplitude modulation in AWGN channel," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 765–768, Jun. 2012.

[261] M. Tanahashi and H. Ochiai, "A multilevel coded modulation approach for hexagonal signal constellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4993–4997, Oct. 2009.

[262] K. P. Srinath and B. S. Rajan, "Fast-decodable MIDO codes with large coding gain," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 992–1007, Feb. 2014.

[263] 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," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 781–786, Mar. 2008.

[264] M. Hekrdla and J. Sykora, "Hexagonal constellations for adaptive physical-layer network coding 2-way relaying," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 217–220, Feb. 2014.

[265] 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.

[266] M. Abdelaziz and T. A. Gulliver, "Triangular constellations for adaptive modulation," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 756–766, Feb. 2018.

[267] X. Zhang, Y. Zhao, and L. Zou, "Optimum asymmetric constellation design for trellis-coded modulation over Gaussian channels," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 528–530, Jul. 2009.

[268] L. V. Subramaniam, B. S. Rajan, and R. Bahl, "Performance of 4-and 8-state TCM schemes with asymmetric 8-PSK in fading channels," *IEEE Trans. Veh. Technol.*, vol. 49, no. 1, pp. 211–219, Jan. 2000.

[269] 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.

[270] 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," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[271] 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.

[272] 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.

[273] 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.

[274] 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.

[275] 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.

[276] T. OShea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[277] A. S. Aghaei, K. N. Plataniotis, and S. Pasupathy, "Widely linear MMSE receivers for linear dispersion space-time block-codes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 8–13, Jan. 2010.

[278] 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.

[279] 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.

[280] 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.

[281] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[282] W. H. Gerstacker, F. Obernosterer, R. Schober, A. T. Lehmann, A. Lampe, and P. Gunreben, "Equalization concepts for Alamouti's space-time block code," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1178–1190, Jul. 2004.

[283] Y. Xia and D. P. Mandic, "A full mean square analysis of CLMS for second-order noncircular inputs," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5578–5590, Nov. 2017.

[284] B. Picinbono and P. Chevalier, "Extensions of the minimum variance method," *Signal processing*, vol. 49, no. 1, pp. 1–9, Feb. 1996.

[285] 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.

[286] G. Gelli, L. Paura, and A. M. Tulino, "Cyclostationarity-based filtering for narrowband interference suppression in direct-sequence spread-spectrum systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1747–1755, Dec. 1998.

[287] S. Buzzi, M. Lops, and A. M. Tulino, "A generalized minimum-mean-output-energy strategy for CDMA systems with improper MAI," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 761–767, Mar. 2002.

[288] 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," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 183–193, Jan. 2009.

[289] P. Chevalier and F. Dupuy, "Widely linear Alamouti receiver for the reception of real-valued constellations corrupted by interferencesThe Alamouti-SAIC/MAIC concept," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3339–3354, Jul. 2011.

[290] 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.

[291] 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.

[292] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, 2005, vol. 95.

[293] 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.

[294] 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.

[295] 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," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3154–3166, May 2019.

[296] 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.

[297] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent component analysis. john willey & sons," *Inc., New York*, pp. 165–202, 2001.

[298] X.-L. Li, T. Adali, and M. Anderson, "Noncircular principal component analysis and its application to model selection," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4516–4528, Oct. 2011.

[299] 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.

[300] S. Wang and N. Rahnavard, "A framework for compressive sensing of asymmetric signals using Normal and Skew-Normal mixture prior." *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5062–5072, Dec. 2015.

[301] T. Adali and V. D. Calhoun, "Complex ICA of brain imaging data," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 136–139, Sep. 2007.

[302] 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.

[303] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, no. 2, pp. 248–261, 2009.

[304] 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.

[305] N. Kwak, "Principal component analysis based on L1-norm maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 9, pp. 1672–1680, 2008.

[306] H. Zhao, P. C. Yuen, and J. T. Kwok, "A novel incremental principal component analysis and its application for face recognition," *IEEE Trans. Syst., Man, Cybern.*, vol. 36, no. 4, pp. 873–886, 2006.

[307] 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.

[308] 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.

[309] 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.

[310] I. Jolliffe, *Principal Component Analysis*, 2nd ed.   New York: Springer, 2002.

[311] 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.

[312] 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.

[313] 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.

[314] 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.

[315] K. R. Gabriel, "The biplot graphic display of matrices with application to principal component analysis," *Biometrika*, vol. 58, no. 3, pp. 453–467, 1971.

[316] 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.

[317] H. El-Amir and M. Hamdy, *Deep Learning Pipeline*.   Springer, 2020.

[318] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: Computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[319] 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.

[320] 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.

[321] 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.

[322] 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.

[323] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *Int. Conf. Artificial Neural Netw.*   Springer, 1997, pp. 583–588.

[324] 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.

[325] 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," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, 2004.

[326] 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.

[327] 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.

[328] 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.

[329] 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.

[330] 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.

[331] 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.

[332] 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.

[333] 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.

[334] Y. Xia and D. P. Mandic, "Complementary mean square analysis of augmented CLMS for second order noncircular gaussian signals," *IEEE Signal Process. Lett.*, vol. 24, no. 9, pp. 1413–1417, Sep. 2017.

[335] 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.

[336] 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.

[337] J.-P. Delmas, "Asymptotically minimum variance second-order estimation for noncircular signals with application to DOA estimation," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1235–1241, May 2004.

[338] P. Chevalier and A. Blin, "Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5323–5336, Nov. 2007.

[339] 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.

[340] 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.

[341] B. Otterseten, M. Kristensson, and D. Astely, "Receiver and method for rejecting cochannel interference," *Google Patents*, 2005.

[342] 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.

[343] "MUROS downlink receiver performance for interference and sensitivity," *GP-090115 3GPP TSG Geran Tdoc, GERAN no. 41*, 2009.

[344] "MUROS uplink receiver performance," *GP-090114 3GPP TSG Geran Tdoc, GERAN no. 41*, 2009.

[345] 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.

[346] Y. Zou, M. Valkama, and M. Renfors, "Digital compensation of I/Q imbalance effects in space-time coded transmit diversity systems," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2496–2508, Jun. 2008.

[347] 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.

[348] R. Carvajal, J. C. Aguero, B. I. Godoy, and G. C. Goodwin, "EM-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 152–160, Jan. 2013.

[349] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[350] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European J. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[351] 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.

[352] C. Enneking, F. Antreich, L. Krieger, and A. L. de Almeida, "Gaussian approximations for intra-and intersystem interference in RNSS," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1198–1201, Jul. 2019.

[353] 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.

[354] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[355] 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.

[356] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the SISO Z-interference channel," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2432–2435, Nov. 2017.

[357] ——, "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.

[358] A. Showail, K. Jamshaid, and B. Shihada, "Buffer sizing in wireless networks: challenges, solutions, and opportunities," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 130–137, Apr. 2016.

[359] B. Shihada and K. Jamshaid, "Buffer sizing for multi-hop networks," Jan. 2014, uS Patent 8,638,686.

[360] N. Bouacida and B. Shihada, "Practical and dynamic buffer sizing using LearnQueue," *IEEE Trans. Mobile Comput.*, vol. 18, no. 8, pp. 1885–1897, Aug. 2019.
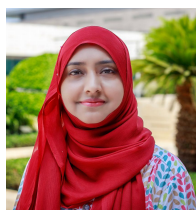
Sidrah Javed (S'16) received the B.E. degree in Electrical (Telecommunication) Engineering from National University of Science and Technology (NUST), Islamabad, Pakistan, in 2012. From 2012 to 2015, she has worked as a Research Engineer in National Radio and Telecommunication Corporation, Pakistan. She is currently pursuing M.S./Ph.D. degree at the Computer, Electrical and Mathematical Sciences & Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST). Her research interests include hardware impairment and interference management using asymmetric signaling, constellation shaping, spatial modulation, and cooperative communications in wireless communications, vehicular tracking/monitoring and solar energy harvesting.

Osama Amin (S'07, M'11, SM'15) received the B.Sc. degree in electrical and electronics engineering from Aswan University, Aswan, Egypt, in 2000, the M.Sc. degree in electrical engineering from Assiut University, Assiut, Egypt, in 2004, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2010. In 2012, he joined the Electrical and Electronics Engineering Department, Assiut University, as an Assistant Professor. He is currently a Research Scientist with the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. His general research interests lie in communications systems and signal processing for communications with special emphasis on wireless applications, optical wireless communication, molecular communications, terahertz communications, green communications, and cognitive radio.

Dr. Amin has served as a technical program committee (TPC) member for ICC, GLOBECOM, IEEE VTC, CROWNCOM, PIMRC, and ISSPIT conferences. He has served also as a Co-Organizer and a Co-Chair of the Next Generation Green ICT and 5G Networking 2015 at the IEEE International Conference on Ubiquitous Wireless Broadband in Montreal, Canada. In 2018, he served as the chair of the IoT, M2M, Sensor Networks, and Ad-Hoc Networking track in the VTC2018-fall held in Chicago, USA. He serves as an Associate Editor for the IEEE COMMUNICATIONS LETTERS.

Basem Shihada (SM'12) is an associate & founding professor in the Computer, Electrical and Mathematical Sciences & Engineering (CEMSE) Division at King Abdullah University of Science and Technology (KAUST). He obtained his PhD in Computer Science from University of Waterloo.

In 2009, he was appointed as visiting faculty in the Department of Computer Science, Stanford University. In 2012, he was elevated to the rank of Senior Member of IEEE. His current research covers a range of topics in energy and resource allocation in wired and wireless networks, software defined networking, internet of things, data networks, network security, and cloud/fog computing.

Mohamed-Slim Alouini (S'94, M'98, SM'03, F'09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.