DIGITAL COMMUNICATIONS IN ADDITIVE WHITE SYMMETRIC ALPHA-STABLE NOISE

AHMED MAHMOOD
(B.E.), (M.S.), National University of Sciences & Technology (NUST), Pakistan

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Ahmed Mahmood

20$^{\text{th}}$ January 2014

ACKNOWLEDGEMENTS

Living these last few years has been as intense as a roller-coaster ride. I have had my share of ups and downs; late-night studying marathons, the stresses involved in meeting multiple deadlines, the euphoria that comes with a sudden research breakthrough. Some say that graduating with a PhD degree is a self-accomplishment - a direct consequence of the sheer number of arduous hours one has clocked. In my case, I was very fortunate to be associated with amazing people throughout my life. I am a firm believer, that their combined effort and support has propelled me and my career to this point in time. To them, I am forever grateful.

Life has a weird way of unraveling itself. The decisions one takes may inadvertently present surprising opportunities in time. My arrival in NUS was one such decision. I am indebted to my supervisors Mandar Chitre and Marc Andre Armand. They have my utmost respect and gratitude. Without their support, encouragement and expertise, this work would never have achieved completion. It was a true privilege to be under their tutelage. Over the course of these few years, I am glad to have developed a strong friendship with either of them. In retrospect, if I had the chance of starting my PhD afresh, I would have definitely chosen the same path all over again.

Where would one be without family? First and foremost, I am grateful to my mother whose unconditional love, prayers and wishes have been with me since day one. She always says that she knew well-beforehand that I would pursue higher education. I want to thank her from the bottom of my heart for all the sacrifices she has made for me. To my late father, who without any doubt would have been so proud to witness this moment. His support and advice were a paramount factor in my decision to enroll into the PhD program at NUS. Your influence in my life will always be remembered and cherished. I miss you so very much. Special thanks to my loving wife Saima who has been - and still is - such an inspiration to me. She has supported me in every way I can think possible. My sisters, Meerah and Zahra, who have been there through thick and thin and have done way more than I ever asked of. My gratitude goes to my in-laws for regularly bucking me up and believing in me. I wish them nothing but the best.

Arriving at a new university in an alien country seemed simultaneously exciting and daunting. Many thanks and appreciation goes out to all my friends at NUS who have truly made my time here remarkable. Special shout-outs go to Shankar (Shanx), Mansoor, Mohammad, Katayoun (Katy), Amna, Rohit, Bhavya, Shahzor, Rizwan, Qasim and Wu Tong for all of the enlightening and joyful discussions we have had. Finally, I want to thank all my teachers, friends and colleagues I have had over the years. Thanks for being there!

Contents

Abstract

The conventional additive white Gaussian noise (AWGN) model adequately simulates many noisy environments. The performance of digital communication schemes in the presence of AWGN has been widely studied and optimized. However if the noise is impulsive, this model fails to mirror the physical attributes of the channel effectively. Impulsive noise is non-Gaussian in nature and is modeled well by random processes based on heavy-tailed symmetric α-stable (SαS) distributions. If the noise samples are independent and identically distributed (IID), the additive white SαS noise (AWSαSN) model may be used to simulate the channel.

System performance is conventionally analyzed at the baseband level. Therefore we investigate characteristics of complex noise derived from passband AWSαSN using conventional (linear) passband-to-baseband conversion schemes. We use a characteristic function (CF) based approach to analyze the noise statistics as the probability density functions (PDF) of SαS random variables cannot be (generally) expressed in closed form. When converted to its complex baseband form, the resulting noise is radically different from its Gaussian counterpart. By varying certain physical parameters, such as the passband sampling rate and the carrier frequency, we may attain different anisotropic (yet symmetric) distributions. Furthermore, the real and imaginary components of the converted noise may be dependent. The bivariate distribution of each

complex noise sample takes on a star-like geometrical configuration. Given that the in-phase and quadrature (I & Q) components are decoded separately, we prove that the uncoded error performance for baseband noise with independent components is the best amongst all possible statistical configurations. We highlight a sampling criterion that guarantees independent noise components.

Using the anisotropy offered by the baseband distribution, efficient placement of signal points on constellation maps for phase-shift keying (PSK) and quadrature amplitude modulation (QAM) are proposed. It is shown that good constellations significantly improve the uncoded error performance of the system under Maximum-Likelihood (ML) detection. Also, as ML detection may be difficult to implement due to the lack of closed-form SαS PDFs, we introduce analytic baseband detectors that achieve near-ML performance.

Though error performance may be enhanced using a discretized linear passband-to-baseband conversion block, further analysis reveals that this is a lossy (sub-optimal) process in non-Gaussian AWSαSN. Therefore, the next logical step is to modify the passband-to-baseband conversion block at the receiver. We discuss and investigate the performance of non-linear schemes based on the myriad filter, L_p-norm and the asymptotic PDF of SαS variables. In conjunction with efficient constellations and suitable baseband detectors, the error performance is significantly better than conventional (linear) receivers. It is shown that if the receiver bandwidth is large enough relative to the symbol rate, impulsive noise may be effectively countered using 'good' decoding methodologies.

We extend our research to multi-carrier communications. In orthogonal frequency-division multiplexing (OFDM) a single impulse will corrupt several symbols in the same block. In conjunction with a modified linear passband-to-baseband conversion block, we show how ML baseband detection performance improves as the number of sub-carriers increases in non-Gaussian AWSαSN. Results are presented for Rayleigh block fading and pure noise scenarios with emphasis on binary and quadrature phase-shift keying (BPSK/QPSK) constellations. As the number of carriers increases, the ML detector error performance actually tends towards the Gaussian noise error curve irrespective of the noise impulsiveness. On the downside, the detection complexity increases exponentially with the number of carriers and is therefore unrealisable. Using results for the single-carrier case, we develop a theory for practically realizable near-optimal receiver schemes for OFDM signals in AWSαSN.

List of Tables

List of Figures

List of Acronyms

AWGN	Additive White Gaussian Noise
AWSαSN	Additive White Symmetric α-Stable Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CF	Characteristic Function
CLT	Central Limit Theorem
CS	Compressed Sensing
DFT	Discrete Fourier Transform
DTFT	Discrete-Time Fourier Transform
FIR	Finite Impulse Response
GBG	Gaussian-Bernoulli-Gaussian
GCLT	Generalized Central Limit Theorem
GGD	Generalized Gaussian Distribution
GSNR	Generalized Signal-to-Noise Ratio
I	In-phase
IDFT	Inverse Discrete Fourier Transform
IID	Independent and Identically Distributed
ISI	Inter-Symbol Interference
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMyF	Matched Myriad Filter
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak-to-Average Power Ratio
PDF	Probability Density Function
PSD	Power Spectral Density
Q	Quadrature
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
SαS	Symmetric α-Stable
SER	Symbol Error Rate
SEP	Symbol Error Probability

SNR	Signal-to-Noise Ratio

List of Notation

$\stackrel{d}{=}$	equal in distribution
$*$	linear convolution operator
$\approx$	approximately equal to
$\sim$	distributed by
$\gg$	much greater than
$\ll$	much lesser than
$a \equiv b \pmod N$	a and b are congruent modulo N
$\langle \cdot, \cdot \rangle$	scalar dot-product
$\Gamma(\cdot)$	the Gamma function
$\Lambda(\cdot)$	the spectral measure of a stable random vector
$E[\cdot]$	the expectation operator
$\Re\{x\}$	real part of x
$\Im\{x\}$	imaginary part of x
$\|x\|$	absolute value of x, i.e., $\|x\| = \sqrt{\Re\{x\}^2 + \Im\{x\}^2}$
$\angle x$	the phase of x, i.e., $\angle x = \tan^{-1}(\Im\{x\}/\Re\{x\})$
$\mathtt{diag}[x_1, \ldots, x_N]$	the $N \times N$ diagonal matrix with i^{th} entry equal to x_i
$\mathbf{I}_N$	the $N \times N$ identity matrix
$\mathbf{0}_N$	the $N \times N$ null matrix
$\mathbf{0}_{K\times N}$	the $K \times N$ null matrix
$\mathbf{x}^*$	element-wise conjugate of the vector $\mathbf{x}$
$\mathbf{x}^{\mathsf{H}}$	hermitian transpose of the vector $\mathbf{x}$
$\mathbf{x}^{\mathsf{T}}$	transpose of the vector $\mathbf{x}$
$\mathbf{X}^*$	element-wise conjugate of the matrix $\mathbf{X}$
$\mathbf{X}^{\mathsf{H}}$	hermitian transpose of the matrix $\mathbf{X}$
$\mathbf{X}^{\mathsf{T}}$	transpose of the matrix $\mathbf{X}$
$\mathcal{S}(\alpha, \beta, \delta, \mu)$	a univariate stable distribution with characteristic exponent α, skew β, scale δ and location μ
$\mathcal{S}(\alpha, \delta)$	univariate SαS distribution with characteristic exponent α and scale δ
$\mathcal{N}(\mu, \sigma^2)$	univariate normal distribution with mean μ and variance σ^2
$\mathcal{CN}(0, \sigma^2)$	circular-symmetric complex normal distribution with variance σ^2
$\mathcal{CN}(\mathbf{0}_{N\times 1}, \mathbf{C})$	N-dimensional circular-symmetric complex normal distribution with covariance matrix $\mathbf{C}$
$\Phi_x(\theta)$	CF of the random variable x
$\Phi_{\mathbf{x}}(\boldsymbol{\theta})$	CF of the random vector $\mathbf{x}$
N_0	the one-sided PSD of an AWGN process
$\|\mathbf{x}\|$	L_2 norm of the vector $\mathbf{x}$

$\|\mathbf{x}\|_p$	L_p norm of the vector $\mathbf{x}$
$\lfloor x \rfloor$	the floor operation
$\{\ldots\}$	elements in a set
$\lvert\mathbb{S}\rvert$	cardinality of the alphabet $\mathbb{S}$
$\mathbb{R}$	set of real numbers
$\mathbb{R}^N$	set of N-tuples such that each element lies in $\mathbb{R}$
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^N$	set of N-tuples such that each element lies in $\mathbb{C}$
$\mathbb{Z}$	set of integers
$\mathbb{Z}^+$	set of positive integers
$\mathtt{arg\,max}\,\rho(\mathbf{x})$	the instance of $\mathbf{x}$ for which $\rho(\mathbf{x})$ is at its maximum
$\mathtt{arg\,min}\,\rho(\mathbf{x})$	the instance of $\mathbf{x}$ for which $\rho(\mathbf{x})$ is at its minimum
`s.t`	such that
$\log(x)$	natural logarithm of x
$\log_{10}(x)$	logarithm to the base 10 of x
$\exp(x)$	exponential of x
$\mathrm{rect}(x)$	the rectangular function
$\mathrm{sign}(x)$	the sign function
$\mathrm{sinc}(x)$	the normalized sinc function: $\frac{\sin(\pi x)}{\pi x}$
$H(X)$	entropy of the distribution X
$H(Y\lvert X)$	conditional entropy of Y given the distribution X
$I(X;Y)$	mutual information between the distributions Y and X

Chapter 1

Introduction

1.1 Motivation

The justification of using the well-known additive white Gaussian noise (AWGN) model stems from the central limit theorem (CLT), which states that for a fixed power constraint, the sum of N independent and identically distributed (IID) random variables tends to a Gaussian distribution as $N \to \infty$ [1]. The AWGN model is a good approximation for the cumulative effect of random noise producing phenomenon encountered in practical communication scenarios [1],[2]. If however the noise is impulsive in nature, i.e., there are sudden high deviations (spikes) in the amplitude of subsequent noise samples, then the AWGN model does not work as well [3], [4]. Therefore, techniques optimized for AWGN cannot be blindly extended to impulsive noise. In certain practical scenarios, impulsive noise dominates the available spectrum. To name a few: the shallow underwater channel [3],[4], communication over power lines [5], digital subscriber line transmission [6] etc. Therefore, a solid understanding of its impact on digital receivers is required. This will further propel the development of new error mitigation techniques in impulsive noise. Though these topics individually cover a vast range of specialized research areas, we try our best to provide adequate discussion to both in this thesis.

Gaussian distributions are part of the larger general family of *stable* distributions [8]. If the power constraint is removed from the CLT, the sum of N IID random variables tends to a *stable* random variable as $N \to \infty$. This is called the *generalized central limit theorem* (GCLT) [7]–[9]. Non-Gaussian stable distributions are heavy-tailed and therefore model impulses much more effectively [8], [9]. In this thesis, we use the additive white symmetric α-stable noise (AWSαSN) channel to model impulsive noise in all our analysis. This approximation is suitable when the impulsive noise samples are IID [10]. Further still, the stability property allows exact tractability of the noise statistics within linear systems.

In wireless communications, signals are transmitted in the passband [1], [2]. However, in the literature, a digital communications system is typically designed and analyzed for a given baseband model [1], [2]. This is done mainly due to the fact that transmitted information is embedded in the baseband signal and therefore most operations are performed in the baseband [1]. In such a scenario, the received signal is implicitly assumed to have gone through a passband-to-baseband conversion process. In the presence of passband AWGN, the optimal passband-to-baseband conversion block is a linear system that only retains the in-band noise information [1]. The corresponding baseband noise samples are circular symmetric complex Gaussian random variables [1], [2]. This noise model is well-known and has been employed vastly in the literature [1], [2]. However, if AWSαSN is passed through this block, the resulting noise statistics are still not truly understood. This thesis offers much needed clarity in this regard. With new insight into the baseband noise model, a communication system can be designed to be more robust in AWSαSN. In due course, we show that the improvement in error performance over conventional linear receivers is significant.

It is well-known that linear systems are far from optimal in AWSαSN [3], [9]. Though characterizing and analyzing baseband noise in linear receivers offers much insight into developing robust systems in impulsive noise, the receiver is still suboptimal. A major part of this thesis explores non-linear design methodologies for a communications receiver. The resulting schemes far outperform any linear receiver in AWSαSN.

1.2 Thesis Goals

The aims of this thesis can be succinctly summarized as follows:

1. To provide a solid understanding of the effects of impulsive noise (modeled by AWSαSN) in a single and multi-carrier digital communications receiver.

2. To propose new mechanisms that mitigate the effect of impulsive noise for both linear and non-linear receivers.

3. To harness modern advances in signal processing that allow robust yet low complexity reception of digital signals in impulsive noise.

1.3 Research Contribution

A digital communications receiver is made up of a number of crucial parts. Our work focuses primarily on the front-end of the physical-layer. The contributions of this thesis are summarized as follows:

1. The baseband statistics of impulsive noise (modeled by AWSαSN)

is shown to take on a plethora of anisotropic symmetric star-like statistical configurations on introducing uniform passband sampling to the conventional (linear) receiver.

2. The probability density functions (PDFs) of stable distributions do not exist in closed-form. The statistics of the baseband noise are therefore derived using a characteristic function (CF) approach that results in analytic expressions.

3. If the real and imaginary components of the received signal are processed separately, we prove that the case with IID real and imaginary noise components offers the best error performance. This results in a bivariate four-tailed symmetric distribution per complex noise sample. Further, a sampling rule is introduced that guarantees this scenario.

4. New constellation design rules are proposed that harness the true potential of the anisotropic baseband noise. The resulting gains for single-carrier systems are large.

5. As closed-form expressions of stable density functions are unavailable, maximum-likelihood (ML) detection may be cumbersome. We therefore analyze the error performance of various analytic cost functions that offer near-ML performance.

6. The linearity of the system may be sacrificed to harness even larger performance gains. This is shown by invoking various non-linear estimation schemes on the passband signal that output soft-values of the transmitted

symbol.

7. If joint-detection is performed directly on the passband samples, this offers the best performance. Further still, this is also exempted from the proposed sampling rule.

8. An orthogonal frequency division multiplexing (OFDM) system with a large number of carriers is shown to naturally mitigate the adverse influence of impulsive noise. This is true even if all carriers are transmitting data. We highlight this by presenting the ML error performance of an OFDM system with increasing number of carriers. The dependency on the optimal constellation is also shown to reduce with the number of carriers for both pure-noise and block fading channels.

9. Due to its computational complexity, ML detection for OFDM in impulsive noise is unfeasible for large number of carriers. Robust approaches based on compressive sensing (CS) and convex programming are employed to generate near-ML estimates of the transmitted OFDM symbol albeit under some additional constraints.

10. Like its single-carrier counterpart, we show that the sampled passband OFDM signal may be processed directly through an estimator to output soft-values of the transmitted symbol block. This approach is devoid of any sampling constraint besides Nyquist's criterion.

11. Various signal-to-noise (SNR) measures have been introduced in the literature. A complete SNR analysis of the discussed schemes are

presented. It is analytically shown how baseband conversion via linear receivers actually reduces the operational SNR of the system.

1.4 Organization

A study of related works is important to provide an understanding of the impact of our research. This is presented in Chapter 2. In Chapter 3 we briefly introduce preliminary concepts and notations that are used in our analysis. Chapter 4 presents an in-depth analysis of the structure of baseband noise in passband AWSαSN in linear receivers. Using the results in Chapter 4, we analyze various single-carrier linear and non-linear schemes that mitigate impulsive noise in Chapter 5. We also discuss new constellation design rules that, along with the new receivers, are necessary to enhance overall error performance. Chapter 6 extends inferences and results in the single-carrier case to OFDM. Using modern tools like CS and convex programming, we show that near-ML performance may be achieved at relatively much lower computational cost than the optimal detector by introducing a few constraints. Finally, we wrap up this work by presenting our conclusions and highlighting future problems in Chapter 7.

Chapter 2

Literature Review

2.1 Impulsive Noise Modeling

Due to the CLT, the cumulative effect of thermal noise (and even external interference from sources with finite power) is modeled by Gaussian distributions [11]. If the samples are IID, then the AWGN channel is used to model the noise process [1]. Though this is a limiting argument, it has been adequately backed up via experimental data. Further still, decades worth of in-depth performance analysis for a large variety of digital communication schemes operating in Gaussian-inspired noise models have been amassed in the literature [1]. This is still an ongoing research area and is the reason we see communication systems as they presently are.

In certain communication scenarios, impulsive noise dominates the available transmission spectrum. Examples include the shallow underwater channel littered with snapping shrimp [4], [10], [12]–[15], communication over power lines [5], [16], [17], digital subscriber line transmission [6], [18]–[21] and atmospheric noise [22]–[25]. Though not as widely prevalent as thermal noise, it can cause severe degradation if not specifically accounted for [10]. The nature of the noise may vary for different scenarios. For example, in power lines the observed noise has two components; an asynchronous process and one that is cyclic and

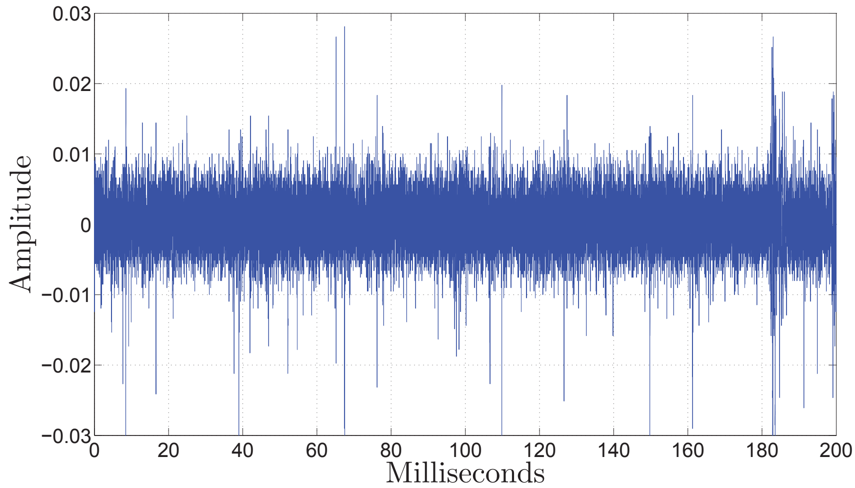


Figure 2.1: A realization of ambient noise in shallow coastal waters in Singapore.

bursty [5],[17]. The findings in one, however, may be directly or indirectly applied to another. In Fig. 2.1, we present the ambient noise heard by an underwater acoustic receiver operating in shallow waters of the coast of Singapore. The data was collected during sea-trials by members of the Acoustic Research Laboratory at the National University of Singapore [26]. The sampling frequency was 500kHz. One can clearly see the impulses (or outliers) in the noise process.

In the literature, several noise models have been used to simulate impulsive noise environments. Analytic representations coupled with 'sufficiently good' empirical fits to practical data have propelled the use of the Middleton class A,B and C models [27]–[29]. These have been used extensively in the literature for several scenarios (for e.g. [17], [30], [31]). Mixtures such as the Cauchy-Gaussian [32]–[34], the Gaussian mixture [14], [17], [33] and the Gaussian-Bernoulli-Gaussian (GBG) [35]–[38] models have also been widely employed to model impulsive noise. Mixture models may have good attributes such as closed-form PDFs or finite first and second-order moments (in the case of the Gaussian mixture/GBG model), however they may not truly depict the noise characteristics at the tails (and hence the impulses) [14]. They also lack the stability property due to which limiting arguments may be needed to characterize the resulting distribution after processing. The same can be stated about the Middleton noise models. Another commonly employed model is based on the heavy-tailed generalized Gaussian distribution [39]–[41]. Though it also has an analytical PDF and finite first and second-order moments, it is devoid of the stability property.

As discussed in Chapter 1, the motivation for using non-Gaussian stable distributions to model impulsive noise in this thesis stems not only from the fact that they are heavy-tailed, but from the GCLT as well [8], [9]. Further still, they model practical impulsive noise models well [3], [9], [14] and have been employed vastly in the literature. They do have a few drawbacks such as the general unavailability of closed-form PDFs and the lack of finite second-order moments [7]–[9]. However, efficient numerical approximations of the density function do exist [42]–[44] and may be used accordingly. Likewise, closed-form approximations of the PDFs have been employed in many instances in the literature. In this thesis we design and analyze systems based on the AWSαSN model. An updated list of the literature employing stable distributions is kept on J. P. Nolan's website [45].

2.2 Communication in Impulsive Noise

Single-carrier systems are extensively used in modern day wireless standards. Research on their performance and spectral efficiencies have been studied thoroughly for various fading and noise channels [1], [2]. Due to a number of favorable properties, multi-carrier schemes (namely OFDM) are being increasingly endorsed in new and emerging wireless standards [1], [2], [46]. Coupled with multiple-input multiple-output (MIMO) techniques, they offer robustness to fading and higher spectral efficiency to a communications system.

In the literature, capacity analysis has been performed for some impulsive noise models. However, there still is much to do. Very recently, [47] and [48] proposed new capacity bounds and results for the GBG channel model. These works focused on single-input single-output systems at the complex baseband level. For the α-stable scenario, [49] offered new results on the capacity of SαS and skewed-stable random variables. The authors considered a real additive noise channel model. There are many open problems for capacity analysis in impulsive noise. For example, a direct extension of the aforementioned research would target passband transmission and MIMO systems.

Various signal processing techniques for stable distributions have been presented in the literature. Common approaches are based on ML [50], fractional lower order moments [9], [51], [52] and CF [9], [53]–[55]. As closed-form PDFs do not exist and the objective is to achieve near-ML performance, analytic approximations to stable densities are sometimes employed. In this regard, the mixture distributions previously discussed have been used to develop near-optimal techniques for AWSαSN [34], [56]. In additive noise scenarios, the location vector of the joint-PDF corresponding to the received observations is a function of the transmitted symbol. Therefore robust estimates of location form an integral part of good receivers. The L_p-norm for $0 < p < 2$ [10], [57], the myriad filter [58]–[62], the generalized Cauchy estimator [59], [63] and the meridian filter [64] are a few examples of measures that mitigate the effect of the impulsive component in received observations. The motivation of using these techniques stems from robust statistics and/or generalized ML estimation (or M-estimation) theory [65].

In digital communications, performance analysis is typically performed at the baseband level [1]. A linear passband-to-baseband conversion process is assumed. However, we show that linear conversion actually reduces the operational SNR at the receiver in the presence of passband non-Gaussian AWSαSN. Due to the lack of finite second-order moments, the conventional definition of SNR does not extend to the general AWSαSN model [9]. Therefore, a suitable SNR measure needs to be introduced to analyze system performance in such scenarios [9], [10], [58], [66].

Another inherent issue with current research trends in mitigating impulsive noise is the implicit assumption that the baseband noise vector is isotropic or has IID samples [9], [62], [67]. Yet the underlying theory that substantiates this assumption has never been formulated. We highlight and address these issues in this thesis. In fact, we show that within the framework of linear passband-to-baseband conversion, one can achieve either statistical configuration by setting the system parameters appropriately. We also show which noise configuration can be exploited to achieve better results.

The effect of impulsive noise in single and multi-carrier systems has been studied [9], [36], [37], [68], [69]. Schemes based on blanking (threshold-based erasure), clipping or clipping-blanking are easy to implement and mitigate the effect of impulsive noise [70]–[75]. Though they outperform conventional techniques tailor-made for AWGN, they are suboptimal. Also, capacity analysis and the ultimate error performance of these schemes in impulsive noise is still an open problem. More recently, error control coding has been used for impulsive noise mitigation in OFDM by taking advantage of the nulls and pilots that are typically located within the transmitted symbol block [37], [38], [76]. However, this method is computationally extensive [36]. The inherently sparse nature of impulsive noise allows using a CS approach to estimate the noise process, which is then used to cancel out the impulses [36], [77]. Due to powerful convex optimization algorithms [78], the computational cost of this approach is not as high [36]. In this thesis, we discuss the CS technique in the light of M-estimation of the passband noise samples. A thorough error analysis of the CS-based receiver is conducted and the results are compared with ML detection.

Chapter 3

Summary of Concepts: Stable Distributions

3.1 Univariate Stable Distributions

3.1.1 Stable Random Variables

A random variable X is classified as stable or (α-stable) if and only if

$$a_1 X^{(1)} + a_2 X^{(2)} \stackrel{d}{=} cX + d \tag{3.1}$$

where $X^{(1)}$ and $X^{(2)}$ are IID copies of X and a_1, a_2, c and d are real numbers [7]–[9]. The symbol $\stackrel{d}{=}$ implies equality in distribution. By induction, we can extend (3.1) to a sum of K random variables. Formally, if $X^{(i)}\ \forall\ i \in \{1, 2, \ldots, K\}$ are IID copies of X and

$$\sum_{i=1}^{K} a_i X^{(i)} \stackrel{d}{=} cX + d \tag{3.2}$$

where $K \in \mathbb{Z}^+$ and $a_i, c, d \in \mathbb{R}$, then X is a stable random variable [8]. If $d = 0$, then X is termed as *strictly stable*. With the exception of the Gaussian and Cauchy cases, a closed-form expression for the PDF of a stable random variable does not exist [8], [9]. On the other hand the CF of such a variable has a closed form [7], [8]. The CF $\Phi_X(\theta)$ of a random variable X is the Fourier transform of its PDF $f_X(x)$ and is defined as

$$\Phi_X(\theta) = E[e^{j\theta X}] = \int_{-\infty}^{+\infty} f_X(x) e^{j\theta x} \, \mathrm{d}x \tag{3.3}$$

where $E[\cdot]$ is the expectation operator and $\theta \in \mathbb{R}$ is the frequency domain variable [11]. Because of their relationship, the CF is a suitable replacement for the PDF to statistically characterize any random variable. For stable random variables, there are different parameterizations of $\Phi_X(\theta)$ which are summarized in [7], each of which have their own desirable properties. We stick to a commonly used convention [7], [8]:

$$\Phi_X(\theta) = \begin{cases} \exp\left(-\delta^\alpha |\theta|^\alpha (1 - j\beta(\text{sign } \theta) \tan \frac{\pi\alpha}{2}) + j\mu\theta\right) & \text{for } \alpha \neq 1 \\ \exp\left(-\delta|\theta|(1 + j\beta\frac{2}{\pi}(\text{sign } \theta) \log|\theta|) + j\mu\theta\right) & \text{for } \alpha = 1 \end{cases} \tag{3.4}$$

The parameters α, β, δ and μ are real and completely define the distribution of X which in turn is denoted by $\mathcal{S}(\alpha, \beta, \delta, \mu)$. α is the *characteristic exponent* and determines the heaviness of the tails for the distribution. The *skew parameter* β alters the symmetry. δ controls the spread and is consequently termed the *scale parameter*. Finally, the value of μ determines the position and is the *location parameter* of the distribution. A summary of these parameters is listed in Table 3.1.

When $\alpha = 2$, (3.4) is the CF of a Gaussian random variable with distribution $\mathcal{N}(\mu, 2\delta^2)$, where μ and $2\delta^2$ are the mean and variance of the distribution respectively [7]–[9]. Notice that when $\alpha = 2$, the skew parameter β is nullified and has no effect on the distribution. For $\alpha = 1$ the CF in (3.4) is that of a

TABLE 3.1: PARAMETER DESCRIPTIONS FOR STABLE DISTRIBUTIONS.

Parameter	Name	Range
α	characteristic exponent	$(0, 2]$
β	skew parameter	$[-1, +1]$
δ	scale parameter	$(0, +\infty)$
μ	location parameter	$(-\infty, +\infty)$

Cauchy random variable.

3.1.2 Symmetric α-Stable Random Variables

A random variable is symmetric α-stable (SαS) if β and μ are equal to zero [8], [9]. The distribution of such a variable reduces to $\mathcal{S}(\alpha, 0, \delta, 0)$. The term 'symmetric' stems from the fact that $f_X(x) = f_X(-x)$ where $f_X(x)$ is the distribution function of X. Further still, as $f_X(x) \in \mathbb{R}$, then from the properties of the Fourier transform we have $\Phi_X(\theta)$ real and symmetric about θ, i.e., $\Phi_X(\theta) = \Phi_X(-\theta) = \Phi_X^*(\theta)$. This relationship between a PDF and its CF is unique to symmetric distributions and is an appropriate test to validate if a stable distribution is indeed SαS or not [8]. We can see this by substituting $\beta = 0$ and $\mu = 0$ in (3.4) to get the CF of X [8]:

$$\Phi_X(\theta) = \exp\left(-\delta^\alpha |\theta|^\alpha\right). \tag{3.5}$$

Any SαS random variable *is also* strictly stable, the converse *does not* hold when $\alpha = 1$ but *holds* otherwise [8]. Therefore, if X is SαS, then from (3.2)

$$\sum_{i=1}^{K} a_i X^{(i)} \stackrel{d}{=} cX. \tag{3.6}$$

The relationship between the coefficients in (3.2) is [7]–[9]

$$c^\alpha = \sum_{i=1}^{K} |a_i|^\alpha. \tag{3.7}$$

From (3.5), the PDF of an SαS random variable is completely parameterized by α and δ. We therefore denote it using the abridged notation $\mathcal{S}(\alpha, \delta)$. For the Gaussian case, $\mathcal{S}(2, \delta)$ is equivalent to $\mathcal{N}(0, 2\delta^2)$, i.e., the zero-mean Gaussian distribution with variance $2\delta^2$:

$$f_X(x) = \frac{1}{\sqrt{4\pi\delta^2}} \exp\left(-\frac{x^2}{\delta^2}\right). \tag{3.8}$$

For the Cauchy case, the PDF corresponding to the CF in (3.5) is

$$f_X(x) = \frac{\delta}{\pi(x^2 + \delta^2)}. \tag{3.9}$$

An SαS PDF is termed as *standard* if $\delta = 1$. Do note that this is different from the conventionally applied definition for a standard Gaussian PDF, i.e., $\mathcal{N}(0, 1) \stackrel{d}{=} \mathcal{S}(0, \frac{1}{\sqrt{2}})$.

Besides stability, a defining characteristic of *non-Gaussian* SαS distributions is that they have algebraic (heavy) tails [8]. The heaviness of these tails is characterized by α. As $\alpha \to 0$, the tails become *increasingly* heavier. This effect can be seen from the asymptotic convergence of an SαS PDF for $\alpha \neq 2$ as $|x| \to +\infty$ [7]:

$$f_X(x) \approx \left(\frac{\alpha\delta^\alpha \sin(\pi\alpha/2)\Gamma(\alpha)}{\pi}\right) |x|^{-\alpha-1}. \tag{3.10}$$

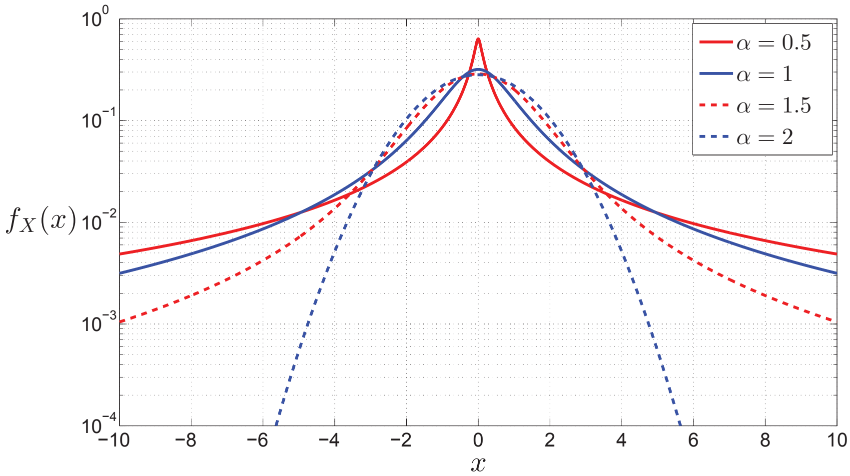


Figure 3.1: Comparison of standard SαS PDFs for different α.

Here, $\Gamma(\cdot)$ denotes the gamma function. As $\int_\epsilon^\infty x^q \, \mathrm{d}x$ is divergent for $q \geq -1$ for all finite $\epsilon > 0$, then from (3.10), we clearly observe that the p^{th}-order moments for $p > \alpha$ are infinite. Thus, second-order moments are infinite for all non-Gaussian SαS random variables. Further still, for $\alpha \leq 1$ even the first-order moment (mean) is infinite.

We highlight various standard SαS PDFs in Fig. 3.1. One can clearly see that as $\alpha \to 2$, the tails of the PDFs get increasingly lighter. Do note that as $\alpha \to 0$, the peak gets more prominent at $x = 0$.

3.2 Multivariate Stable Distributions

3.2.1 Stable Random Vectors

The expression in (3.2) can be extended to define a stable random *vector* $\mathbf{x} = [x_1, x_2, \ldots, x_N]^\mathsf{T}$ where $\mathbf{x} \in \mathbb{R}^N$ and $N \in \mathbb{Z}^+$ such that

$$\sum_{i=1}^{K} a_i \mathbf{x}^{(i)} \stackrel{d}{=} c\mathbf{x} + \mathbf{d} \tag{3.11}$$

and $\mathbf{x}^{(i)}$ $\forall$ $i \in \{1, 2, \ldots, K\}$ are IID copies of $\mathbf{x}$. Here, $a_i, c \in \mathbb{R}$ and $\mathbf{d} \in \mathbb{R}^{\mathbb{N}}$. If $\mathbf{d} = \mathbf{0}_{N\times 1}$, then $\mathbf{x}$ is *strictly stable* [8], [9]. The joint-CF of $\mathbf{x}$ is evaluated by

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = E\left[\exp\left(j\boldsymbol{\theta}^{\mathsf{T}}\mathbf{x}\right)\right] \tag{3.12}$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_N]^{\mathsf{T}}$ and $\theta_k \in \mathbb{R}$ is the frequency domain variable corresponding to the k^{th} element in $\mathbf{x}$. To get the marginal CF corresponding to x_k from (3.12), one needs to substitute $\theta_l = 0$ $\forall$ $l \neq k$. Unlike the univariate case, $\Phi_{\mathbf{x}}(\boldsymbol{\theta})$ generally does not have a closed form and is given by

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\int_{\mathbb{S}_N} |\langle\boldsymbol{\theta}, \mathbf{s}\rangle|^{\alpha}\left(1 - j\,\mathrm{sign}(\langle\boldsymbol{\theta}, \mathbf{s}\rangle)\tan\frac{\pi\alpha}{2}\right)\Lambda(\mathbf{s})\;\mathrm{d}\mathbf{s} + j\langle\boldsymbol{\theta}, \boldsymbol{\mu}\rangle\right) \tag{3.13}$$

for $\alpha \neq 1$ and

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\int_{\mathbb{S}_N} |\langle\boldsymbol{\theta}, \mathbf{s}\rangle|^{\alpha}\left(1 + j\frac{2}{\pi}\mathrm{sign}(\langle\boldsymbol{\theta}, \mathbf{s}\rangle)\log|\langle\boldsymbol{\theta}, \mathbf{s}\rangle|\right)\Lambda(\mathbf{s})\;\mathrm{d}\mathbf{s} + j\langle\boldsymbol{\theta}, \boldsymbol{\mu}\rangle\right) \tag{3.14}$$

for $\alpha = 1$, where $\mathbb{S}_N$ represents all points on the $(N-1)$-dimensional unit circle lying in N-dimensional space and $\langle\cdot,\cdot\rangle$ denotes the inner product of two vectors [8], [9]. Mathematically,

$$\mathbb{S}_N = \{\mathbf{s}|\mathbf{s} \in \mathbb{R}^N, \|\mathbf{s}\| = 1\}. \tag{3.15}$$

Here, $\boldsymbol{\mu} \in \mathbb{R}^N$ is the *location vector* and $\Lambda(\mathbf{s}) \in \mathbb{R}$ is a finite *spectral measure.* $\Lambda(\mathbf{s})$ contains information related to the scale and skewness of the distribution and is non-zero only over $\mathbf{s} \in \mathbb{S}_N$. The integration in (3.13) and (3.14) is performed over all points $\mathbf{s} \in \mathbb{S}_N$.

3.2.2 SαS Random Vectors

The PDF $f_{\mathbf{x}}(\mathbf{x})$ of a *stable* random vector $\mathbf{x}$ is SαS if $f_{\mathbf{x}}(\mathbf{x}) = f_{\mathbf{x}}(-\mathbf{x})$. This implies that the CF is real and symmetric about $\boldsymbol{\theta} = \mathbf{0}$, i.e,

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \Phi_{\mathbf{x}}(-\boldsymbol{\theta}) = \Phi^*_{\mathbf{x}}(\boldsymbol{\theta}). \tag{3.16}$$

Like the univariate case, this is an appropriate test to validate if a stable distribution is SαS or not. If $\mathbf{x}$ is SαS, it is also strictly stable. The converse, however, is not true [8]. On applying the condition in (3.16) to (3.13) and (3.14), we see that the CF of a SαS random vector $\mathbf{x}$ is given by

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\int_{\mathbb{S}_N} |\langle\boldsymbol{\theta}, \mathbf{s}\rangle|^{\alpha} \Lambda(\mathbf{s})\ \mathrm{d}\mathbf{s}\right). \tag{3.17}$$

We note that $\Lambda(\mathbf{s})$ is equal for *any two* antipodal vectors $\mathbf{s}$, i.e., $\Lambda(\mathbf{s}) = \Lambda(-\mathbf{s})$, and assigns weights to $|\langle\boldsymbol{\theta}, \mathbf{s}\rangle|^{\alpha}$ [8]. To better understand the relationship between $\Lambda(\mathbf{s})$ and the configuration of an SαS PDF, we briefly discuss two special cases:

The Isotropic Case

If $\mathbf{x}$ is an *isotropic* SαS vector with each component $\mathcal{S}(\alpha, \delta)$, the CF in (3.17) reduces to

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\delta^{\alpha}\|\boldsymbol{\theta}\|^{\alpha}\right) \tag{3.18}$$

where $\|\cdot\|$ represents the Euclidean norm , i.e., $\|\boldsymbol{\theta}\| = (\sum_{i=1}^{N} \theta_i^2)^{1/2}$. We note that the CF is solely a function of the magnitude of the frequency domain vector $\boldsymbol{\theta}$. On comparison with (3.17) we see that $\Lambda(\mathbf{s})$ is constant over all $\mathbf{s} \in \mathbb{S}_N$ [8].

IID Components

If the components of $\mathbf{x}$ are IID copies of $X \sim \mathcal{S}(\alpha, \delta)$, the joint-CF is given by the multiplication of N individual copies of the expression in (3.5), i.e.,

$$\begin{aligned} \Phi_{\mathbf{x}}(\boldsymbol{\theta}) = E\left[\exp\left(j\sum_{i=1}^{N}\theta_i x_i\right)\right] &= \prod_{i=1}^{N} E\left[\exp\left(j\theta_i x_i\right)\right] \\ &= \prod_{i=1}^{N} \Phi_X(\theta_i) \\ &= \exp\left(-\delta^{\alpha}\left(\sum_{i=1}^{N}|\theta_i|^{\alpha}\right)\right). \end{aligned} \tag{3.19}$$

As (3.19) satisfies the condition in (3.18), it is SαS. Here, $\Lambda(\mathbf{s})$ is non-zero only for a finite number of $\mathbf{s} \in \mathbb{S}_N$. Precisely, it is a sum of N-dimensional equal-weighted Dirac delta functions located at the Cartesian axis intercepts with the $(N-1)$-dimensional unit circle. An example is the univariate SαS case, which in essence is a 1-dimensional random vector with a single $\mathcal{S}(\alpha, \delta)$ distributed component. In this case $\mathbb{S}_1 = \{-1, 1\}$ and $\Lambda(s) = \delta^{\alpha}/2(D(s-1) + D(s+1))$ where $D(s)$ is the Dirac delta function.

Contrary to the univariate case, closed-form CFs generally do not exist for multivariate SαS distributions. However, there are certain subclasses that are exceptions to this rule, with one of them being the sub-Gaussian α-stable vector family [8].

3.2.3 Sub-Gaussian α-Stable Random Vectors

If a non-Gaussian SαS random vector can be factored into

$$\mathbf{x} = A^{1/2}\mathbf{g} \tag{3.20}$$

where $\mathbf{g}$ is a zero-mean Gaussian vector of dimension equal to that of $\mathbf{x}$ and A is a totally right-skewed stable *random variable* independent of $\mathbf{g}$, then $\mathbf{x}$ is sub-Gaussian α-stable or α-sub-Gaussian [8], [9]. Further still, the distribution of A will be $\mathcal{S}(\alpha/2, 1, (\cos\frac{\pi\alpha}{4})^{2/\alpha}, 0)$.

The density of $\mathbf{x}$ in (3.20) shares structural similarities with that of the underlying Gaussian vector. For e.g., an N-dimensional non-degenerate α-sub-Gaussian vector implies an underlying non-degenerate N-dimensional Gaussian vector and will have its equiprobable density surfaces shaped as N-dimensional ellipsoids. These surfaces become spherical if the elements of the underlying Gaussian vector are IID [9]. This concept may be extended to the degenerate case. Also note that due to A in (3.20), the elements of $\mathbf{x}$ will always be dependent, irrespective of the elements of $\mathbf{g}$ being independent or not [8]. The CF of $\mathbf{x}$ is further given as

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\left|\frac{1}{2}\boldsymbol{\theta}^{\mathsf{T}}\mathbf{R}\boldsymbol{\theta}\right|^{\alpha/2}\right) \tag{3.21}$$

where $\mathbf{R}$ is the covariance matrix of $\mathbf{g}$ and α is the characteristic exponent of $\mathbf{x}$ [8], [9]. For $\alpha = 1$ and $\alpha = 2$ the joint CF in (3.21) reduces to that of an α-sub-Gaussian Cauchy and a zero-mean Gaussian vector, respectively. As $\mathbf{R}$ is a covariance matrix, it is positive semi-definite [1], [11]. Therefore, we can omit $|\cdot|$ and write (3.21) as

$$\Phi_{\mathbf{x}}(\boldsymbol{\theta}) = \exp\left(-\left(\frac{1}{2}\boldsymbol{\theta}^{\mathsf{T}}\mathbf{R}\boldsymbol{\theta}\right)^{\alpha/2}\right) \tag{3.22}$$

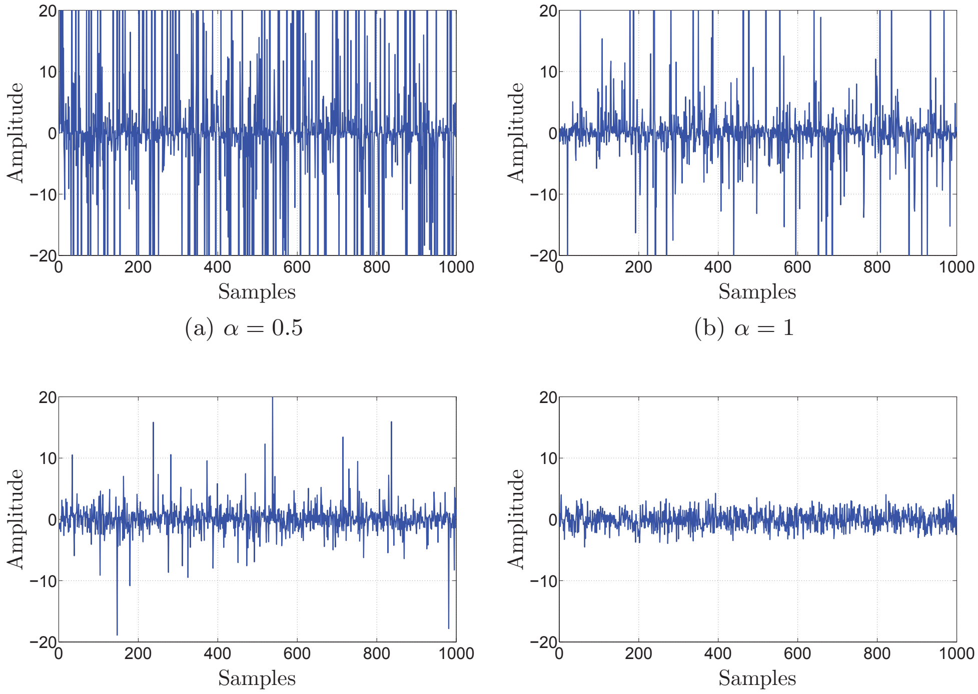
For the univariate case, (3.20) can be written as

$$X = A^{1/2}G \tag{3.23}$$

where $G \sim \mathcal{N}(0, 2\delta^2)$ and $A \sim \mathcal{S}(\alpha/2, 1, (\cos\frac{\pi\alpha}{4})^{2/\alpha}, 0)$ and are mutually independent. We see that a SαS random variable is α-sub-Gaussian. The converse is true as well. This is observed by comparing (3.22) for scalar $\mathbf{x}$ to (3.5) and noting that they are equal. Multivariate SαS distributions, however, may not be α-sub-Gaussian.

3.2.4 The Additive White Symmetric α-Stable Noise Model

The AWSαSN channel has been used in the literature to model practical impulsive noise channels [3], [4], [9], [10]. By definition, samples of AWSαSN are real and IID copies of $X \sim \mathcal{S}(\alpha, \delta)$. This implies that the AWSαSN process is *stationary*. If the N-tuple $\mathbf{x} = [x_1, x_2, \ldots, x_N]^{\mathsf{T}}$ consists of dissimilar samples of an AWSαSN process, the joint-CF of $\mathbf{x}$ is given by (3.19) and therefore $\mathbf{x}$ is SαS. However, on comparison with (3.22) we see the joint-CF in (3.19) is not sub-Gaussian. If we denote the PDF of X by $f_X(\cdot)$, the joint-PDF of $\mathbf{x}$ is given



(a) $\alpha = 0.5$ (b) $\alpha = 1$

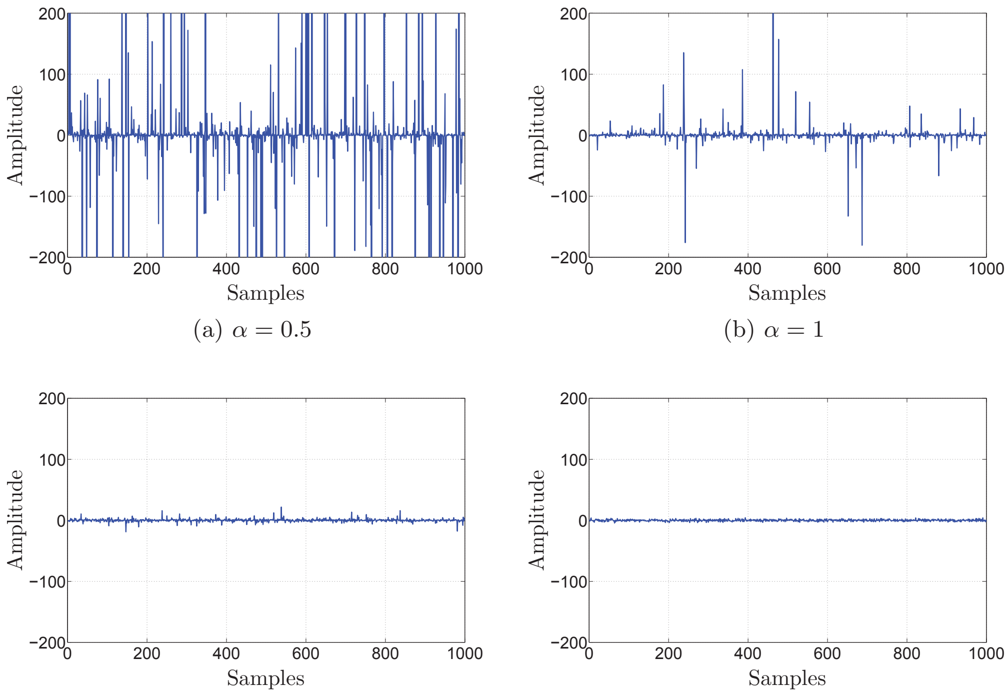
(c) $\alpha = 1.5$ (d) $\alpha = 2$, AWGN

Figure 3.2: Realizations of AWSαSN for different values of α and $\delta = 1$.

by

$$f(\mathbf{x}) = \prod_{i=1}^{N} f_X(x_i). \tag{3.24}$$

When $\alpha = 2$, the more general AWSαSN channel reduces to the AWGN model. The term 'white' implies a flat power spectral density (PSD) spanning over all frequencies for the Gaussian case. It should be noted that this definition does not hold when associated with non-Gaussian AWSαSN. This is due to the fact that second-order moments of stable non-Gaussian distributions are infinite [8], [9]. The term is maintained because it asserts independence of noise samples in AWGN which is what is implied in the case of non-Gaussian AWSαSN.



(a) $\alpha = 0.5$

(b) $\alpha = 1$

(c) $\alpha = 1.5$

(d) $\alpha = 2$, AWGN

Figure 3.3: Realizations of AWSαSN for different values of α and $\delta = 1$ on a larger scale.

In Fig. 3.2, realizations of AWSαSN for different values of α have been plotted. Each sample is distributed by $\mathcal{S}(\alpha, 1)$. One can clearly observe the varying impulsive behavior of the noise processes. The same realizations are reproduced on a larger scale in Fig. 3.3 to highlight the difference in amplitudes of the impulses.

3.2.5 Complex SαS Random Vectors

If $\mathbf{x}_c = [x_{c_1}, x_{c_2}, \ldots, x_{c_N}]^\mathsf{T}$ is an N-dimensional *complex* random vector, its statistics can be completely characterized by a $2N$-dimensional PDF [11]. Denoting $\mathbf{x} = [\Re\{\mathbf{x}_c\}^\mathsf{T}\ \Im\{\mathbf{x}_c\}^\mathsf{T}]^\mathsf{T}$, i.e., $\mathbf{x} \in \mathbb{R}^{2N}$, we state that $\mathbf{x}_c$ is a complex SαS random vector if $\mathbf{x}$ is symmetric and satisfies (3.11).

Chapter 4

Characterization of Complex Baseband SαS Noise

In this chapter we derive a general bivariate-CF for complex noise under conventional (linear) passband-to-baseband conversion for passband AWSαSN. We prove that the baseband noise is SαS. Using the derived expressions, we extract useful insight into the characteristics of the resultant noise. It is well known that the baseband noise derived from AWGN is isotropic and its components are IID Gaussian [11]. For the non-Gaussian AWSαSN case, the components may or may not be independent. Further still, the noise might not even be isotropic. In fact, by varying system parameters one may achieve a variety of baseband noise distributions. Due to these differences, techniques that are optimized for Gaussian noise scenarios might not be effective in the presence of impulsive noise. The work presented in this chapter has been published in [79]–[81].

4.1 Linear Passband-to-Baseband Conversion

The relationship of a continuous-time passband signal $s(t)$, indexed by $t \in \mathbb{R}$, to its baseband form $\tilde{s}(t)$ is represented by the well known expression [1]

$$s(t) = \Re\left\{\tilde{s}(t)\exp\left(j2\pi f_c t\right)\right\}, \tag{4.1}$$

where f_c is the carrier frequency. In a typical digital communications scheme, the information is usually embedded in the baseband signal. However, the actual transmission is performed in the passband [1]. This is done primarily to allow effective propagation through the communication medium [1], [2]. Examples of such mediums are the wireless electromagnetic channel and the underwater acoustic channel [1], [2]. Another inherent advantage of passband transmission is that desired frequency bands can be allocated by modifying f_c. This allows efficient use of the available spectrum. If (4.1) is sampled, we get

$$s[n] = \Re\left\{\tilde{s}[n] \exp\left(j2\pi\frac{f_c}{f_s}n\right)\right\}, \tag{4.2}$$

where f_s and $\tilde{s}[n]$ are the *passband sampling frequency* and the *upsampled baseband signal*, respectively. Also, $n \in \mathbb{Z}$ is the discrete-time index. The square bracket notation is used to denote discrete-time signals, i.e., $s[n] = s(t/f_s)$ and so forth. We note that (4.1) is the limiting case of (4.2) as $f_s \to +\infty$.

At the receiver, the signal is converted back to its baseband form. To do this, one essentially has to shift $s[n]$ by f_c in the spectral domain and pass the result through a low pass filter, the impulse response of which we denote by $v[n]$. The filter is of L-taps and of bandwidth equal to the normalized message signal bandwidth B/f_s where B is the baseband sampling frequency. We assume the order of the filter, and hence L, to be sufficiently large so that there is an adequate low-pass filtering effect. In the presence of a passband signal the Nyquist criterion should be satisfied, i.e., $f_s > 2f_c + B$ [1]. Mathematically, the passband-to-baseband conversion is given by

$$\tilde{s}[n] = 2v[n] * s[n]e^{-j2\pi\frac{f_c}{f_s}n}, \tag{4.3}$$

where $*$ denotes the linear convolution operation. The scale factor of 2 is appended so that the relationship in (4.4) is maintained [1].

If $s[n]$ and hence $\tilde{s}[n]$ are random processes, then in addition to the relationship in (4.2), a joint-PDF is associated with either of these signals. Adhering to convention in standard texts [1], [11], we use capitalized letters to represent random processes. We define $W[n]\ \forall\ n \in \mathbb{Z}$ to be samples of a *real* passband AWSαSN process and further state $\tilde{W}[n]$ to be its upsampled baseband counterpart. Therefore, from (4.2) and (4.3), we have

$$W[n] = \Re\left\{\tilde{W}[n]\exp\left(j2\pi\frac{f_c}{f_s}n\right)\right\} \tag{4.4}$$

and

$$\tilde{W}[n] = 2v[n] * W[n]e^{-j2\pi\frac{f_c}{f_s}n}, \tag{4.5}$$

respectively. As (4.5) consists only of linear operations, $\tilde{W}[n]$ may be analyzed irrespective of the transmitted signal [1]. This is done next.

We note that though $W[n]$ is a real process, $\tilde{W}[n]$ is complex. The shifting operation in (4.5) is given by

$$\begin{aligned} W^{+}[n] &= W[n]e^{-j2\pi\frac{f_c}{f_s}n} \\ &= \left(\cos(2\pi\frac{f_c}{f_s}n) - j\sin(2\pi\frac{f_c}{f_s}n)\right) W[n]. \end{aligned} \tag{4.6}$$

We can write (4.6) in vector form as

$$\mathbf{w}^{+}[n] = \begin{bmatrix} W_R^{+}[n] \\ \\ W_I^{+}[n] \end{bmatrix} = \begin{bmatrix} \cos(2\pi\frac{f_c}{f_s}n) \\ \\ -\sin(2\pi\frac{f_c}{f_s}n) \end{bmatrix} W[n], \tag{4.7}$$

where $W_R^{+}[n]$ and $W_I^{+}[n]$ are the real and imaginary components of $W^{+}[n] \in \mathbb{C}$, respectively. We use the $^{+}$ symbol in our notation to highlight that the positive band of the passband signal is shifted to zero. The subsequent filtering operation in (4.5) is expressed as

$$\tilde{W}[n] = 2\sum_{k=0}^{L-1} v[n]W^{+}[n-k]. \tag{4.8}$$

As $\tilde{W}[n] \in \mathbb{C}$ is complex, it can be written in vector form as well:

$$\begin{aligned} \tilde{\mathbf{w}}[n] &= \begin{bmatrix} \tilde{W}_R[n] \\ \\ \tilde{W}_I[n] \end{bmatrix} \\ &= 2\sum_{k=0}^{L-1} v[k] \begin{bmatrix} W_R^{+}[n-k] \\ \\ W_I^{+}[n-k] \end{bmatrix} \\ &= 2\sum_{k=0}^{L-1} v[k]\mathbf{w}^{+}[n-k], \end{aligned} \tag{4.9}$$

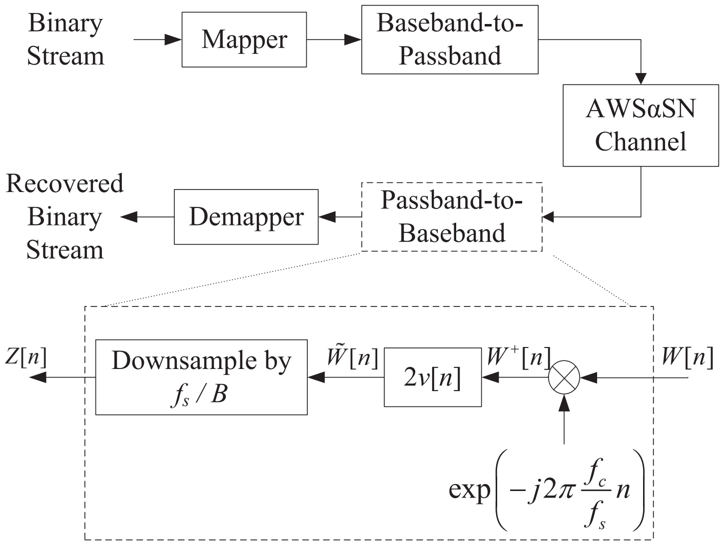


Figure 4.1: A schematic of an uncoded digital communication system with AWSαSN along with a descriptive block diagram of the passband-to-baseband conversion block.

where $\tilde{W}_R[n]$ and $\tilde{W}_I[n]$ are the real and imaginary components of $\tilde{W}[n]$, respectively. Therefore, $\tilde{\mathbf{w}}[n] \in \mathbb{R}^2$. To get the actual baseband signal $Z[n] \in \mathbb{C}$, $\tilde{W}[n]$ is downsampled by a factor of f_s/B, i.e.,

$$Z[n] = \tilde{W}[f_s n/B]. \tag{4.10}$$

Similarly, in vector form,

$$\mathbf{z}[n] = \begin{bmatrix} Z_R[n] \\ Z_I[n] \end{bmatrix} = \tilde{\mathbf{w}}[f_s n/B]. \tag{4.11}$$

where $Z_R[n]$ and $Z_I[n]$ are the real and imaginary components of $Z[n]$, respectively.

A schematic for an uncoded digital communication system is shown in Fig. 4.1 along with an elaborate diagram of the passband-to-baseband conversion block. The mapper converts a sequence of M information bits to a symbol that is represented as a signal point on a constellation diagram. The total number of symbols is consequently assumed to be 2^M. The operation of the demapper is the inverse of the mapper.

4.2 Complex Baseband SαS Noise

In this section we derive the bivariate CF of complex baseband SαS noise with the assumption that the passband noise is AWSαSN. We will first characterize $\mathbf{w}^+[n]$. On the basis of that we will derive the CF of $\tilde{\mathbf{w}}[n]$ and $\mathbf{z}[n]$. We assume the passband samples $W[n]$ are each distributed by $\mathcal{S}(\alpha, \delta_w)$.

From the discussion in Section 3.2.3, the samples $W[n]$ are *individually* sub-Gaussian as they are each univariate SαS. Using (3.23) we can express $W[n]$ as

$$W[n] = A^{\frac{1}{2}}[n]G[n], \tag{4.12}$$

where $A[n] \sim \mathcal{S}(\frac{\alpha}{2}, 1, (\cos\frac{\pi\alpha}{4})^{2/\alpha}, 0)$ and $G[n] \sim \mathcal{N}(0, 2\delta_w^2)$ are independent of each other. As the samples $W[n]$ are IID, so will be the samples $A[n]$ and $G[n]$ for all $n \in \mathbb{Z}$.

<u>*Proposition 1:*</u> For any n and $\alpha \in (0, 2]$, $\mathbf{w}^+[n]$ is α-sub-Gaussian with the

covariance matrix of the underlying Gaussian vector $\mathbf{g}[n]$ being:

$$\mathbf{R}[n] = 2\delta_w^2 \begin{bmatrix} \cos^2(2\pi\frac{f_c}{f_s}n) & -\frac{1}{2}\sin(4\pi\frac{f_c}{f_s}n) \\ -\frac{1}{2}\sin(4\pi\frac{f_c}{f_s}n) & \sin^2(2\pi\frac{f_c}{f_s}n) \end{bmatrix}. \tag{4.13}$$

<u>*Proof:*</u> By substituting (4.12) in (4.7) we get

$$\begin{aligned} \mathbf{w}^+[n] &= \begin{bmatrix} \cos\left(2\pi\frac{f}{f_s}n\right) \\ -\sin\left(2\pi\frac{f}{f_s}n\right) \end{bmatrix} W[n] \\ &= A^{\frac{1}{2}}[n] \begin{bmatrix} \cos\left(2\pi\frac{f}{f_s}n\right) \\ -\sin\left(2\pi\frac{f}{f_s}n\right) \end{bmatrix} G[n] \\ &= A^{\frac{1}{2}}[n]\mathbf{g}[n], \end{aligned} \tag{4.14}$$

where

$$\mathbf{g}[n] = \begin{bmatrix} \cos\left(2\pi\frac{f}{f_s}n\right) \\ -\sin\left(2\pi\frac{f}{f_s}n\right) \end{bmatrix} G[n].$$

We see that $\mathbf{g}[n]$ is zero-mean bivariate Gaussian and therefore due to the form in (4.14), $\mathbf{w}^+[n]$ is α-sub-Gaussian. The covariance matrix of $\mathbf{g}[n]$ is calculated by evaluating $\mathbf{R}[n] = E[\mathbf{g}[n]\mathbf{g}[n]^\mathsf{T}]$, which results in (4.13). $\square$

As the rank of the covariance matrix in (4.13) is 1 $\forall\ n \in \mathbb{Z}$, $\mathbf{g}[n]$ and hence $\mathbf{w}^+[n]$ are degenerate.

Corollary 1: The characteristic function of $\mathbf{w}^{+}[n]$ is

$$\Phi_{\mathbf{w}^{+}[n]}(\boldsymbol{\theta}) = \exp\left(-\left(\frac{1}{2}\boldsymbol{\theta}^{\mathsf{T}}\mathbf{R}[n]\boldsymbol{\theta}\right)^{\alpha/2}\right). \tag{4.15}$$

Eq. (4.15) results from substituting (4.13) for $\mathbf{R}$ in (3.22).

Proposition 2: For all $n \in \mathbb{Z}$ and $\alpha \in (0, 2]$, the random vector $\tilde{\mathbf{w}}[n]$ is SαS and has the following joint-CF:

$$\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) = \exp\left(-\sum_{k=0}^{L-1}\left|2v^2[k]\boldsymbol{\theta}^{\mathsf{T}}\mathbf{R}[n-k]\boldsymbol{\theta}\right|^{\alpha/2}\right). \tag{4.16}$$

Proof: We use (4.9) and the whiteness of passband noise samples to get:

$$\begin{aligned}
\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) &= E\left[\exp\left(j\boldsymbol{\theta}^{\mathsf{T}}\tilde{\mathbf{w}}[n]\right)\right] \\
&= E\left[\exp\left(j2\sum_{k=0}^{L-1}v[n]\boldsymbol{\theta}^{\mathsf{T}}\mathbf{w}^{+}[n-k]\right)\right] \\
&= \prod_{k=0}^{L-1}E\left[\exp\left(j2v[n]\boldsymbol{\theta}^{\mathsf{T}}\mathbf{w}^{+}[n-k]\right)\right] \\
&= \prod_{k=0}^{L-1}E\left[\exp\left(j\left(2v[n]\boldsymbol{\theta}\right)^{\mathsf{T}}\mathbf{w}^{+}[n-k]\right)\right].
\end{aligned} \tag{4.17}$$

From (3.12), we have

$$\Phi_{\mathbf{w}^{+}[n-k]}\left(2v[n]\boldsymbol{\theta}\right) = E\left[\exp\left(j\left(2v[n]\boldsymbol{\theta}\right)^{\mathsf{T}}\mathbf{w}^{+}[n-k]\right)\right]. \tag{4.18}$$

Therefore, we can write (4.17) as

$$\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) = \prod_{k=0}^{L-1} \Phi_{\mathbf{w}^+[n-k]}\left(2v[n]\boldsymbol{\theta}\right) \tag{4.19}$$

On substituting (4.15) in (4.19), we get

$$\begin{aligned} \Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) &= \prod_{k=0}^{L-1} \exp\left(-\left(2v^2[k]\boldsymbol{\theta}^\mathsf{T}\mathbf{R}[n-k]\boldsymbol{\theta}\right)^{\alpha/2}\right) \\ &= \exp\left(-\sum_{k=0}^{L-1}\left(2v^2[k]\boldsymbol{\theta}^\mathsf{T}\mathbf{R}[n-k]\boldsymbol{\theta}\right)^{\alpha/2}\right). \end{aligned} \tag{4.20}$$

To see if the distribution of $\tilde{\mathbf{w}}[n]$ is bivariate SαS, we merely note that (4.20) is real and symmetric about $\boldsymbol{\theta}$, i.e., it satisfies (3.16). □

4.2.1 Marginal Distributions

We can get the marginal CFs $\Phi_{\tilde{W}_R[n]}(\theta)$ and $\Phi_{\tilde{W}_I[n]}(\theta)$ by substituting $\boldsymbol{\theta} = [\theta, 0]^\mathsf{T}$ and $\boldsymbol{\theta} = [0, \theta]^\mathsf{T}$ into (4.20), respectively.

Corollary 2: The CFs of the marginal distributions of $\tilde{\mathbf{w}}[n]$ are

$$\Phi_{\tilde{W}_R[n]}(\theta) = \exp\left(-\sum_{k=0}^{L-1}\left(4v^2[k]\theta^2\cos^2\left(2\pi\frac{f_c}{f_s}(n-k)\right)\delta_w^2\right)^{\alpha/2}\right) \tag{4.21}$$

$$\Phi_{\tilde{W}_I[n]}(\theta) = \exp\left(-\sum_{k=0}^{L-1}\left(4v^2[k]\theta^2\sin^2\left(2\pi\frac{f_c}{f_s}(n-k)\right)\delta_w^2\right)^{\alpha/2}\right) \tag{4.22}$$

From (3.5), (4.21) and (4.22), we note that both $\tilde{W}_R[n]$ and $\tilde{W}_I[n]$ are SαS random variables. Their scale parameters are

$$\delta_{\tilde{W}_R[n]} = \left(\sum_{k=0}^{L-1} \left| 2v[k] \cos\left(2\pi \frac{f_c}{f_s}(n-k) \right) \delta_w \right|^{\alpha} \right)^{1/\alpha} \quad \text{and} \tag{4.23}$$

$$\delta_{\tilde{W}_I[n]} = \left(\sum_{k=0}^{L-1} \left| 2v[k] \sin\left(2\pi \frac{f_c}{f_s}(n-k) \right) \delta_w \right|^{\alpha} \right)^{1/\alpha}, \tag{4.24}$$

respectively. It is observed that the relationship between the CFs of $Z[n]$ and $\tilde{W}[n]$, i.e., $\Phi_{\mathbf{z}[n]}(\boldsymbol{\theta}) = \Phi_{\tilde{\mathbf{w}}[f_s n/B]}(\boldsymbol{\theta})$, extends to the marginal distributions of $Z[n]$ corresponding to (4.21) and (4.22).

4.2.2 An Example: The AWGN Case

The validity of the joint-CF in (4.16) and its marginals in (4.21) and (4.22) may be verified by applying the results to the case where $W[n]$ is an AWGN process. The following facts of the resulting baseband noise are already known [1]:

1. For a given sample $\tilde{W}[n]$, the real and imaginary components are IID. Therefore, the bivariate distribution of any complex baseband sample is isotropic.

2. All complex baseband samples $Z[n]$ are IID. Hence, the distribution does not vary with time.

We first calculate the CF of the marginal distribution of $\tilde{W}_R[n]$ for this case. From (4.21) we have

$$\begin{aligned}\Phi_{\tilde{W}_R[n]}(\theta) &= \Phi_{\tilde{\mathbf{w}}[n]}(\theta_1 = \theta,\ \theta_2 = 0)\\ &= \exp\left(-\sum_{k=0}^{L-1} 4v^2[k]\theta^2\cos^2(2\pi\frac{f_c}{f_s}(n-k))\delta_w^2\right)\\ &= \exp\left(-4\theta^2\delta_w^2 \underbrace{\sum_{k=0}^{L-1} v^2[k]\cos^2(2\pi\frac{f_c}{f_s}(n-k))}_{\text{convolution}}\right).\end{aligned} \tag{4.25}$$

We know that $v[n]$ is a fixed low-pass filter which allows frequencies within $[-\frac{B}{2f_s}, \frac{B}{2f_s}]$ to pass through. Therefore, $v^2[n]$ is in essence also a low-pass filter as the magnitude of the frequency response of $v^2[n]$ is a *triangular* function scaled by B/f_s and lies within $[-B/f_s, B/f_s]$. Looking at the convolution term in (4.25), we see that $v^2[n]$ succeeds to terminate the high frequency component in $\cos^2(2\pi\frac{f_c}{f_s}n) = \cos(4\pi\frac{f_c}{f_s}n)/2 + 1/2$ and retains the latter term after scaling it by B/f_s. Therefore (4.25) is independent of n (time-invariant) and reduces to:

$$\Phi_{\tilde{W}_R[n]}(\theta) = \Phi_{\tilde{W}_R}(\theta) = \exp\left(-\frac{2B\delta_w^2\theta^2}{f_s}\right). \tag{4.26}$$

Using the same arguments we also evaluate

$$\begin{aligned}\Phi_{\tilde{W}_I[n]}(\theta) &= \Phi_{\tilde{\mathbf{w}}[n]}(\theta_1 = 0,\ \theta_2 = \theta)\\ &= \Phi_{\tilde{W}_I}(\theta) = \exp\left(-\frac{2B\delta_w^2\theta^2}{f_s}\right).\end{aligned} \tag{4.27}$$

From the discussion in Section 3.1.2, we see that the individual distributions of the real and imaginary components of $\tilde{W}[n]$ coincide with $\mathcal{N}(0, 4B\delta_w^2/f_s)$. Also, the marginal CFs are independent of the sample index n, highlighting the fact that the distributions of the components of $\tilde{\mathbf{w}}[n]$ do not vary with time. We can substitute δ_w^2 by $N_0 f_s/4$ in (4.26) and (4.27) where $N_0/2$ is the two-sided PSD of the passband AWGN process to get the marginal CFs in terms of N_0:

$$\Phi_{\tilde{W}_R}(\theta) = \Phi_{\tilde{W}_I}(\theta) = \exp\left(-\frac{BN_0\theta^2}{2}\right). \tag{4.28}$$

Now to see if the real and imaginary parts of $\tilde{W}[n]$ are mutually independent at any n, we apply the same principle used in simplifying the convolution term in (4.25) to the joint-CF in (4.16):

$$\begin{aligned}\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) &= \exp\left(-\sum_{k=0}^{L-1}\left(2v^2[k]\boldsymbol{\theta}^\mathsf{T}\mathbf{R}[n-k]\boldsymbol{\theta}\right)\right)\\
&= \exp\Bigg(-4\theta_1^2\delta_w^2\underbrace{\sum_{k=0}^{L-1}v^2[k]\cos^2(2\pi\frac{f_c}{f_s}(n-k))}_{=B/2f_s}\\
&\qquad -4\theta_2^2\delta_w^2\underbrace{\sum_{k=0}^{L-1}v^2[k]\sin^2(2\pi\frac{f_c}{f_s}(n-k))}_{=B/2f_s}\\
&\qquad\qquad +4\theta_1\theta_2\delta_w^2\underbrace{\sum_{k=0}^{L-1}v^2[k]\sin(4\pi\frac{f_c}{f_s}(n-k))}_{=0}\Bigg).\end{aligned} \tag{4.29}$$

Therefore,

$$\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) = \exp\left(-\frac{2B\delta_w^2\theta_1^2}{f_s} - \frac{2B\delta_w^2\theta_2^2}{f_s}\right).$$

On comparison with (4.26) and (4.27), we have

$$\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta}) = \Phi_{\tilde{W}_R}(\theta_1)\Phi_{\tilde{W}_I}(\theta_2). \tag{4.30}$$

So the real and imaginary components of $\tilde{W}[n]$ for any n are also independent. We further see that $\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta})$ is independent of n, thus showing that the bivariate distribution of all baseband samples are identical.

Finally, we note that the impulse response of the FIR filter will have a similar form to

$$v[n] = \begin{cases} \frac{B}{f_s} \operatorname{sinc}\left(\frac{B}{f_s}\left(n - \frac{L}{2}\right)\right) & \text{for } 0 \leq n \leq L-1 \\ 0 & \text{otherwise,} \end{cases} \tag{4.31}$$

where $\operatorname{sinc} = \frac{\sin(\pi x)}{\pi x}$ $\forall\, x \in \mathbb{R}$ is the *normalized sinc* function. From (4.31), when $n - \frac{L}{2}$ is a multiple of f_s/B, then $v[n] = 0$, except at $n = 0$. The baseband samples are mutually independent because of the whiteness of passband samples, the placement of nulls in the impulse response $v[n]$ and the fact that we downsample by f_s/B after filtering to generate the baseband signal. Precisely, the convolution operation in conjunction with the downsampling block allows each $Z[n]$ to be expressed as a projection of $W^+[k]$ over

$$\begin{aligned} \ell_n[k] &= 2v[f_s n/B - k] \\ &= \begin{cases} \frac{2B}{f_s} \operatorname{sinc}\left(n - \frac{B}{f_s}\left(k + \frac{L}{2}\right)\right) & \text{for } f_s n/B - L + 1 \leq k \leq f_s n/B \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{4.32}$$

From (4.8) we have

$$Z[n] = \tilde{W}[f_s n/B] = 2\sum_{k=0}^{L-1} v[k]W^{+}[f_s n/B - k]. \tag{4.33}$$

Without loss of generality, we may express (4.33) as

$$\begin{aligned} Z[n] &= 2\sum_{k=-\infty}^{+\infty} v[k]W^{+}[f_s n/B - k] \\ &= 2\sum_{k=-\infty}^{+\infty} v[f_s n/B - k]W^{+}[k] \\ &= \sum_{k=-\infty}^{+\infty} \ell_n[k]W^{+}[k]. \end{aligned} \tag{4.34}$$

For $v[n]$ in (4.31), the set of $\ell_n[k]$ $\forall$ $n \in \mathbb{Z}$ are *orthogonal* functions over $k \in \mathbb{Z}$ for sufficiently large L. This coupled with the fact that $W^{+}[n]$ $\forall$ $n \in \mathbb{Z}$ are independent random variables ensures that $Z[n]$ $\forall\, n \in \mathbb{Z}$ are independent random variables for the Gaussian case [1].

4.2.3 Analysis of non-Gaussian SαS Noise Samples

The AWGN example was a verification exercise for the joint-CF. We will now analyze the statistics of $Z[n]$ for the non-Gaussian scenario.

Independence of Samples

Though the orthogonal argument associated with (4.34) is sufficient to guarantee independence of $Z[n]$ $\forall$ $n \in \mathbb{Z}$ for the Gaussian case, it *does not* extend to the remaining SαS family. By setting

$$L \leq f_s/B, \tag{4.35}$$

one ensures no overlap (of non-zero values) of the orthogonal function set $\ell_n[k]$ over $k \in \mathbb{Z}$. In other words, this condition ensures that any passband noise sample $W[n]$ is involved in creating only one baseband sample $Z[n]$. As $W[n]\ \forall\, n \in \mathbb{Z}$ are IID random variables, this guarantees the mutual independence of all $Z[n]\ \forall\ n \in \mathbb{Z}$.

Identical Samples

To see if the baseband samples are *identical*, the expression in (4.16) may be rewritten as

$$\begin{aligned}\log(\Phi_{\tilde{\mathbf{w}}[n]}(\boldsymbol{\theta})) &= -\sum_{k=0}^{L-1}\left(2v^2[k]\boldsymbol{\theta}^\mathsf{T}\mathbf{R}[n-k]\boldsymbol{\theta}\right)^{\alpha/2} \\ &= -p[n] * q[n], \end{aligned} \tag{4.36}$$

where

$$p[n] = (2v^2[n])^{\alpha/2} \tag{4.37}$$

$$q[n] = \left(\boldsymbol{\theta}^\mathsf{T}\mathbf{R}[n]\boldsymbol{\theta}\right)^{\alpha/2} \tag{4.38}$$

and $*$ represents the convolution operator. Using (4.31), $p[n]$ becomes

$$p[n] = \begin{cases} \left|\frac{\sqrt{2}B}{f_s}\operatorname{sinc}\left(\frac{B}{f_s}\left(n - \frac{L}{2}\right)\right)\right|^\alpha & \text{for } 0 \le n \le L-1 \\ 0 & \text{otherwise.} \end{cases} \tag{4.39}$$

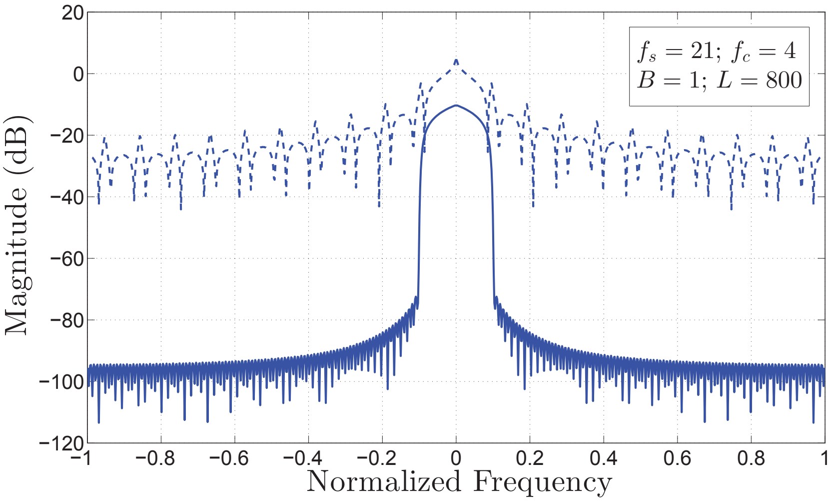
It is observed that $p[n]$ is in essence a low-pass filter. To depict this, the magnitude response of (4.39) for $\alpha = 1$ and 2 are presented in Fig. 4.2a with

$f_s = 21$ Hz, $f_c = 4$ Hz, $B = 1$ Hz and $L = 800$. In practical scenarios, typical estimates of α lie within $[1.5, 2)$ [3]. The low-pass characteristics of $p[n]$ may be extended to other values of α within this range. We also note that $q[n]$ is a periodic signal as the term $\boldsymbol{\theta}^{\mathsf{T}}\mathbf{R}[n]\boldsymbol{\theta}$ in (4.38) may be expanded as

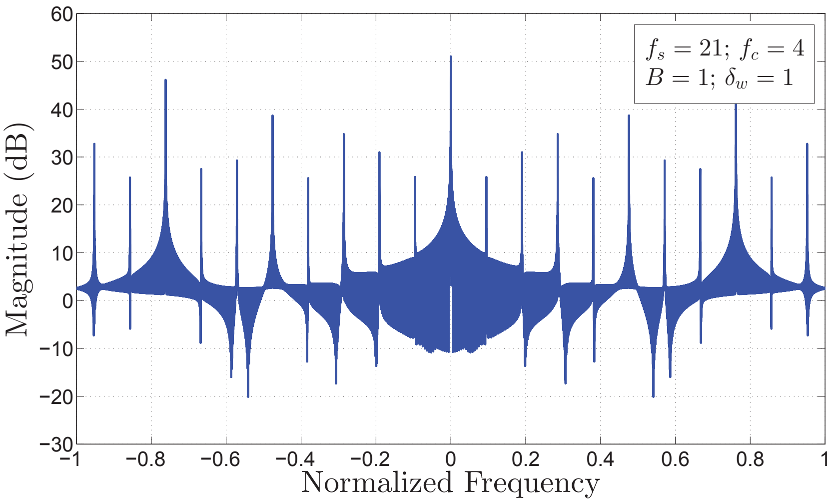
$$q[n] = \left(2\delta_w^2\left(\theta_1^2\cos^2\left(2\pi\frac{f_c}{f_s}n\right) + \theta_2^2\sin^2\left(2\pi\frac{f_c}{f_s}n\right) - \theta_1\theta_2\sin\left(4\pi\frac{f_c}{f_s}n\right)\right)\right)^{\alpha/2}. \tag{4.40}$$

Any function of a periodic signal is periodic as well and in turn may be represented as a Fourier series. It should be noted that the number of harmonics of $q[n]$ is equivalent to $f_s/\gcd(4f_c, f_s)$ (where gcd is the *greatest common divisor*) and does not depend on α, $\boldsymbol{\theta}$ and/or δ_z. For $\theta_1 = \theta_2 = 1$ and $\delta_z = 1$, we have plotted the magnitude response of $q[n]$ for the Cauchy case in Fig. 4.2b. The result in (4.36) may be visualized as the multiplication of the respective frequency responses of $p[n]$ and $q[n]$. For *any* combination of $\boldsymbol{\theta}$, it has been evaluated that the convolution in (4.36) (after subsequent downsampling) is independent of n, i.e., all harmonics of $q[n]$ are effectively suppressed. This can be seen from the instances of the magnitude frequency response of $p[n]$ and $q[n]$ presented in Fig. 4.2. Thus the distribution of all samples $Z[n]$ are identical. We should highlight that these arguments are valid only if L is large enough to induce an effective low-pass filtering effect.

On the other hand we know that $L \leq f_s/B$ for $Z[n]$ to be independent in time. If L is constrained to f_s/B, we still get a low-pass frequency response for $p[n]$ in (4.39). We depict this for the Cauchy and Gaussian cases in Fig. 4.3



(a)



(b)

Figure 4.2: In part (a), the magnitude frequency response of $p[n]$ for the Gaussian case (solid line) and the Cauchy case (dashed line) are shown. Part (b) presents the magnitude response of $q[n]$ for the Cauchy case with $\theta_1 = \theta_2 = 1$.

for a certain instance of system parameters. Therefore, it is possible to get IID $Z[n]$ $\forall$ $n \in \mathbb{Z}$.

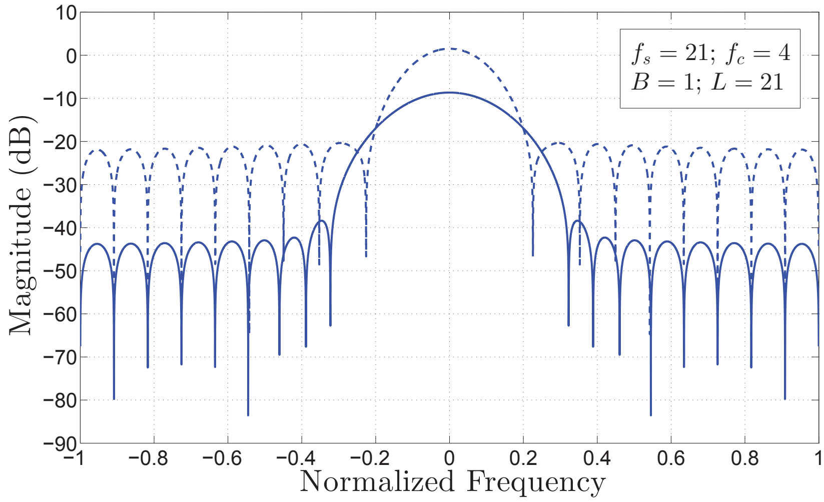


Figure 4.3: the magnitude frequency response of $p[n]$ for the Gaussian case (solid line) and the Cauchy case (dashed line) are shown for $L = f_s/B$.

Independence of Components

In the Gaussian case, it was determined that the real and imaginary components of $Z[n]$ are always independent. This is generally not true for non-Gaussian baseband SαS noise. For the components of $Z[n]$ to be independent, (4.16) has to break up into a product of its two marginal CFs in (4.21) and (4.22).

Corollary 3: For any given sample $Z[n]$, the real and imaginary components are independent if and only if $f_s = 4f_c$.

Corollory 3 follows from the fact that only for $f_s = 4f_c$ does the matrix $\mathbf{R}[n]$ in (4.16) become diagonal for any n. Further still, only one of the diagonal elements will be non-zero for all $n \in \mathbb{Z}$. The joint-CF then reduces to the product of its marginal CFs which proves independence of components.

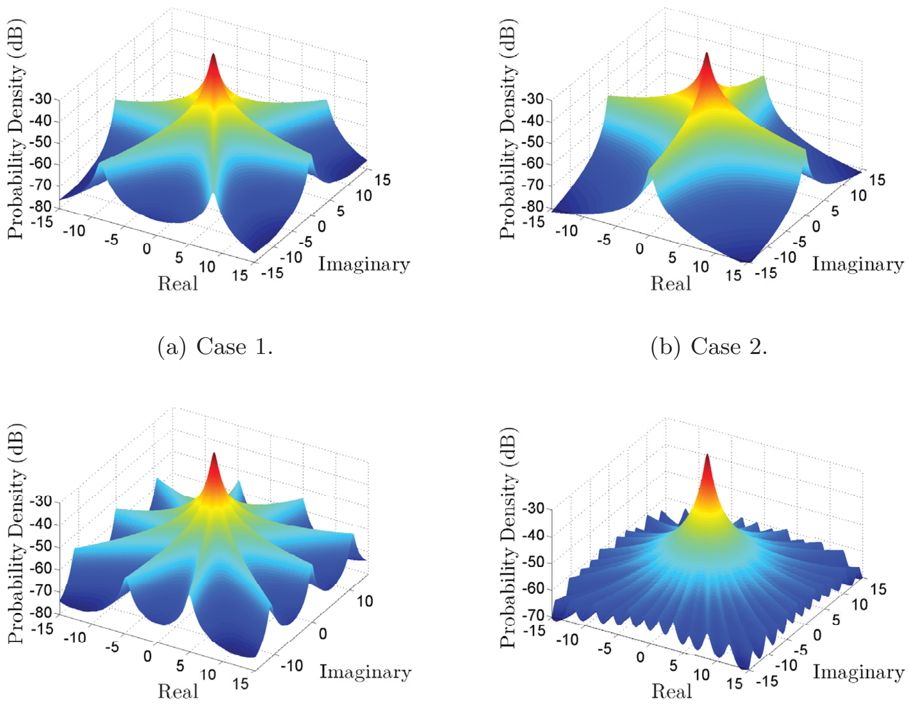
Bivariate Tail Statistics

A question pertaining to the structure of the bivariate PDF of non-Gaussian $\mathbf{z}[n]$ arises, which unlike the Gaussian case (isotropic), varies for different ratios of f_c/f_s. An intuitive look into the expression in (4.7) reveals that the joint-PDF of $\mathbf{w}^+[n]$ is degenerate and lies along an angle of $2\pi\frac{f_c}{f_s}n$ from the positive real axis. The filtering operation in (4.9) essentially scales and sums the independent vectors $\mathbf{w}^+[n]$, which results in a 2-dimensional convolution of these rotated degenerate PDFs. Due to the heavy tail phenomenon accompanying stable random variables, one would expect the resultant bivariate PDF of $\tilde{\mathbf{w}}[n]$ to have tails along angles that are multiples of $2\pi f_c/f_s$ from the positive real axis. If the ratio f_c/f_s to be rational, the number of tails will be *finite* and will be *uniformly distributed* around the origin, hence resulting in non-isotropic distributions. The angle between the tails is given as

$$\psi_z = \begin{cases} \frac{2\pi \gcd(f_c, f_s)}{f_s} & \text{if } f_s \text{ is an even multiple of } f_c \\ \frac{\pi \gcd(f_c, f_s)}{f_s} & \text{otherwise.} \end{cases} \tag{4.41}$$

The total number of tails is noted to be equal to $2\pi/\psi_z$.

Fig. 4.4 presents the bivariate density functions for the Cauchy case ($\alpha = 1$). The different system parameters used to obtain these plots are summarized in Table 4.1. For all cases, the order of the FIR filter was 800. The PDFs were evaluated by taking the inverse Fourier transform of (4.16). For Fig. 4.4b the real and imaginary components are independent following Corollary 3. For all other two cases, the real and imaginary parts are dependent. In Fig. 4.4d, the



(a) Case 1.

(b) Case 2.

(c) Case 3.

(d) Case 4.

Figure 4.4: Bivariate PDFs of complex baseband SαS noise are presented for the Cauchy case ($\alpha = 1$) under the assumption that the passband noise is AWSαSN. The parameters that generate each of these plots are summarized in Table 4.1.

baseband noise is near-isotropic. From the trends in Fig. 4.4, as $2\pi/\psi_z \to \infty$, it would be reasonable to expect the PDF of $\mathbf{z}[n]$ to converge to an isotropic distribution. In wireless communications, passband-to-baseband conversion is performed in the continuous time domain, i.e., with $f_s \to \infty$ [1]. In Chapter 5 we prove that for $f_s \to \infty$ and finite f_c the resultant PDF of $\mathbf{z}[n]$ is indeed isotropic.

Identical Components

On a final note, the marginal distributions of $\mathbf{z}[n]$, although are time-invariant, are also *not* exactly identical. From Fig. 4.4, we observe that they are only

TABLE 4.1: PARAMETER SETTINGS FOR GENERATING THE DENSITY FUNCTIONS IN FIG. 4.4.

Case	f_c	f_s	$\gcd(f_c, f_s)/f_s$	B
1.	4	12	1/3	1
2.	4	16	1/4	1
3.	4	20	1/5	1
4.	4	21	1/21	1

identical if there exists a tail along the imaginary axis as in Fig. 4.4b or if the number of tails is large, i.e., ψ_z is small. One way to get around this is by modifying the definition of a baseband signal in (4.2) to

$$s[n] = \Re\left\{\tilde{s}[n] \exp\left(j(2\pi\frac{f_c}{f_s}n - \frac{\pi}{4})\right)\right\}. \tag{4.42}$$

This ensures that the tails of the bivariate distribution are uniformly distributed about both the real and imaginary axis.

4.3 Bounds on the Baseband Scale Parameter

An important relationship is that of the baseband scale parameter with the noise impulsiveness and system parameters. As the marginal distributions are not exactly identical we restrict our analysis to (4.23). We adopt a limiting approach that is also applicable to (4.24). As per the discussion in the previous section, $\tilde{W}[n]$ (and therefore $Z[n]$) are IID complex samples $\forall\ n \in \mathbb{Z}$. From (4.23), we have

$$\delta_{Z_R}^{\alpha} = (2\delta_w)^{\alpha}\left(\sum_{k=0}^{L-1}\left|v[k]\cos\left(2\pi\frac{f_c}{f_s}(n-k)\right)\right|^{\alpha}\right), \tag{4.43}$$

where δ_{Z_R} is the scale parameter of $Z_R[n]$. We note that δ_{Z_R} changes linearly with δ_w. We comment on (4.43) for two special cases:

1. *The Gaussian Case:* We have discussed the baseband Gaussian CF in detail in Section 4.2.2. From the results in (4.26) and (4.28), (4.43) reduces to

$$\delta_{Z_R}^2 = (2\delta_w)^2 \frac{B}{f_s} = \frac{BN_0}{2} \tag{4.44}$$

   for $\alpha = 2$. Thus, $\delta_{\tilde{W}_R} \propto \sqrt{B/f_s}$ or $N_0/2$.

2. *Extremely Impulsive Noise:* As $\alpha \to 0$, the passband noise becomes increasingly impulsive. In the limit, (4.43) converges to

$$\delta_{Z_R}^\alpha \to L \tag{4.45}$$

   for $\alpha = 0$. From (4.45), we note that $\delta_{Z_R} \propto L^{1/\alpha}$. The order of the FIR filter thus plays an important role in evaluating δ_{Z_R} as $\alpha \to 0$.

We will now analyze $\delta_{Z_R}^\alpha$ for the general SαS case. Eq. (4.43) may be written as

$$\delta_{Z_R}^\alpha = (2\delta_w)^\alpha \left(a[n] * b[n]\right), \tag{4.46}$$

where

$$a[n] = |v[n]|^{\alpha} \quad \text{and} \tag{4.47}$$

$$b[n] = \left| \cos\left(2\pi \frac{f_c}{f_s} n\right) \right|^{\alpha} \tag{4.48}$$

As (4.43) is time invariant, (4.46) simplifies to:

$$\delta_{Z_R}^{\alpha} = (2\delta_w)^{\alpha} \left(A_D B_D\right), \tag{4.49}$$

where A_D and B_D are the dc terms of $a[n]$ and $b[n]$, respectively. Noting that $b[n]$ is periodic, we have from the Fourier transform

$$B_D = \frac{1}{N} \sum_{n=0}^{N-1} b[n], \tag{4.50}$$

where $N = f_s/\mathrm{gcd}(2f_c, f_s)$ is the period of $b[n]$ and signifies the number of tails in distribution of $Z[n]$. This can be seen by comparing N to ψ_z in (4.41). As $N \to \infty$, B_D converges to

$$\begin{aligned} B_D &= \frac{f_s}{N} \sum_{n=0}^{N-1} b[n]/f_s = \frac{f_s}{N} \sum_{n=0}^{N-1} b\left(n/f_s\right)/f_s \\ &\to \frac{f_s}{N} \int_0^{\frac{N}{f_s}} b(t)\, \mathrm{d}t. \end{aligned} \tag{4.51}$$

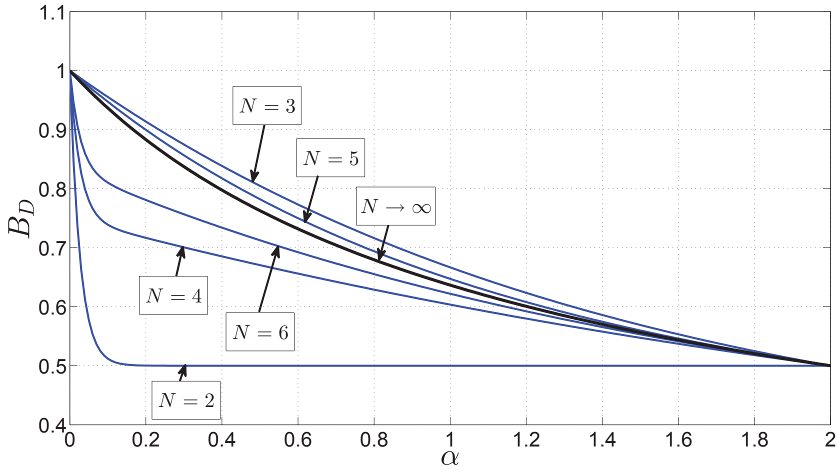


Figure 4.5: B_D against α for different values of N.

Using the inherent structure of $b(t)$, (4.51) simplifies to

$$\begin{aligned} B_D &= 4f_c \int_0^{\frac{1}{4f_c}} b(t)\,\mathrm{d}t \\ &= \frac{\Gamma\left(\frac{1+\alpha}{2}\right)}{\sqrt{\pi}\Gamma\left(1+\frac{\alpha}{2}\right)}. \end{aligned} \tag{4.52}$$

Although (4.52) is evaluated for the limit $N \to \infty$, it offers a good approximation for a large range of N. In Fig. 4.5, we highlight this by plotting (4.52) and (4.50) against α for increasing values of N. In the limit, B_D depends only on the impulsiveness, which is quantified by α, and not on any of the system parameters B, f_c, f_s and $L-1$.

We note that (4.52) also extends to its counterpart in (4.24), i.e, if

$$b[n] = \left|\sin\left(2\pi\frac{f_c}{f_s}n\right)\right|^{\alpha}, \tag{4.53}$$

then (4.52) is true as $N \to \infty$. This implies that $\delta_{Z_R} = \delta_{Z_I}$ for large N and therefore the marginal CFs of $Z[n]$ are identical.

We will now focus on evaluating an analytical expression for A_D. From Fourier transform properties:

$$A_D = \sum_{k=0}^{L-1} a[k]. \tag{4.54}$$

Using the same sequence of steps in (4.51) and noting that $a(t)$ is symmetric about $t = \frac{L-1}{2f_s}$,

$$\begin{aligned} A_D \to f_s \int_0^{\frac{L}{f_s}} a(t)\, dt &= 2f_s \int_{\frac{L-1}{2f_s}}^{\frac{2(L-1)+1}{2f_s}} a(t)\, \mathrm{d}t \\ &= 2f_s \int_0^{\frac{L}{2f_s}} a\left(t + \frac{L-1}{2f_s}\right) \mathrm{d}t. \end{aligned} \tag{4.55}$$

As $f_s >> B$, (4.55) is a good approximation for A_D. Evaluating (4.55) is still not trivial. We accomplish this by introducing a tight upper-bound $\tilde{a}(t)$ for $a\left(t + \frac{L-1}{2f_s}\right)$,

$$\tilde{a}(t) = \begin{cases} \left|\frac{B}{f_s}\right|^\alpha & 0 \le t < \frac{1}{2B} \\ \left|\frac{1}{\pi f_s t}\right|^\alpha & \frac{1}{2B} \le t < \frac{L}{2f_s} \\ 0 & \text{otherwise.} \end{cases} \tag{4.56}$$

In Fig. 4.6, we compare both $a(t + (L-1)/(2f_s))$ and $\tilde{a}(t)$ for $\alpha = 1$. We observe that $\tilde{a}(t)$ correctly highlights the decay in $a(t+(L-1)/(2f_s))$ as $t \to L/f_s$. The bound becomes tighter as $\alpha \to 0$. On substituting $a(t + (L-1)/(2f_s))$ by

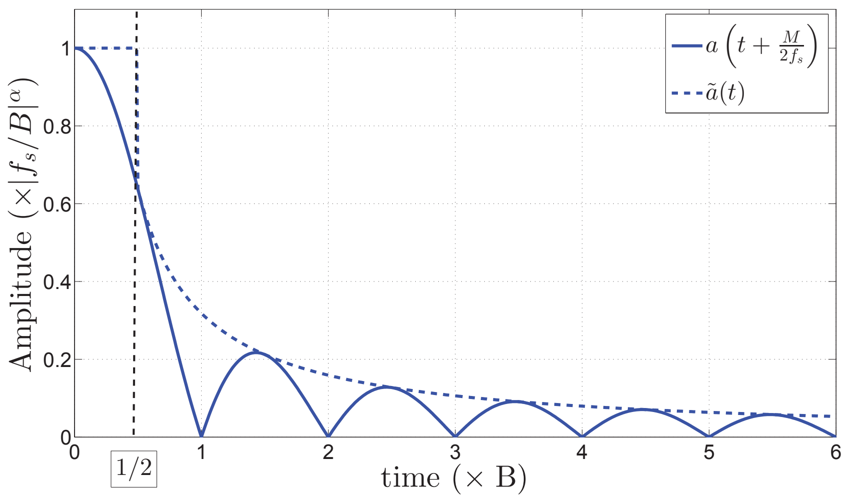


Figure 4.6: $a\left(t+\frac{L-1}{2f_s}\right)$ and $\tilde{a}(t)$ against time for $\alpha = 1$.

$\tilde{a}(t)$ in (4.55) and solving, we get the following upper bound on A_D:

$$\tilde{A}_D = \left(\frac{B}{f_s}\right)^{\alpha-1}\left(1+\frac{2^\alpha}{\pi^\alpha(\alpha-1)}\right) - \left(\frac{1}{L}\right)^{\alpha-1}\left(\frac{2^\alpha}{(\alpha-1)\pi^\alpha}\right). \tag{4.57}$$

To ensure adequate low-pass filtering in the passband-to-baseband conversion process, $L-1$ has to be large enough. In practice, there is a limit to how large $L-1$ can be as it adds to the complexity of the system. We may factor L into

$$L = \frac{2f_s}{B}\tilde{L}, \tag{4.58}$$

where $\tilde{L}$ is a measure of the number of lobes of the sinc function in $v[n]$. For $\tilde{L} = 1$, $v[n]$ consists of only the main lobe. For $\tilde{L} = 2$, the main lobe and its two adjacent side lobes (one on either side) constitute $v[n]$. Usually $\tilde{L} > 1$ to ensure good low-pass filtering. In Fig. 4.7, we plot the error between $\tilde{A}_D$ and A_D against α for various values of $\tilde{L}$ with $B/f_s = 0.05$. The error curve is

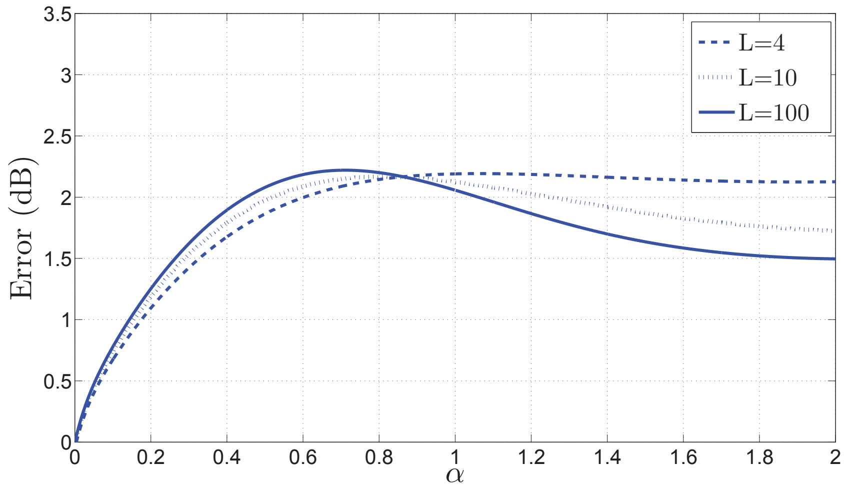


Figure 4.7: Error (dB) between A_D and $\tilde{A}_D$ against α.

almost the same for any $f_s >> B$. Also, increasing $\tilde{L}$ further hardly results in any difference. This shows that (4.57) tracks the transitions in A_D consistently for all possible combinations of the system parameters.

The bound in (4.57) is a fruitful result. For any given α, $\tilde{A}_D$ is a function of $1/L^{\alpha-1}$ and $(B/f_s)^{\alpha-1}$. As (4.57) is a tight bound, the trends observed in $\tilde{A}_D$ against the system parameters can be extended A_D. We analyze these trends for three different cases:

1. *Gaussian-Like*: If $L-1$ is large enough to guarantee adequate low-pass filtering, increasing $L-1$ any further will not affect A_D significantly. For this case, B/f_s plays a larger role in determining δ_{Z_R}. From (4.44), it is known that $L-1$ plays no part in the evaluation of the baseband noise spread for the Gaussian case. We may substitute (4.58) in (4.57) to get

$$A_D < c_{\alpha,\tilde{L}} \left(\frac{B}{f_s}\right)^{\alpha-1}, \tag{4.59}$$

where

$$c_{\alpha,\tilde{L}} = 1 + \frac{2^{\alpha}}{(\alpha-1)\pi^{\alpha}}\left(1 - \frac{1}{(2\tilde{L})^{\alpha-1}}\right).$$

For α within the vicinity of 2, $c_{\alpha,\tilde{L}}$ is hardly affected by increasing $\tilde{L}$ or α and is almost constant.

2. *Cauchy-Like*: For the special case of $\alpha = 1$, (4.57) is

$$A_D < 1 + \frac{2\log(\frac{B}{f_s}) + 2\log L}{\pi}. \tag{4.60}$$

It is observed that $\tilde{A}_D$ increases logarithmically with B/f_s and $L-1$. Using (4.58), we may simplify this further to

$$A_D < 1 + \frac{2\log(2) + 2\log(\tilde{L})}{\pi}. \tag{4.61}$$

Thus $\tilde{A}_D$ may be expressed solely as a function of $\tilde{L}$.

3. *Very Impulsive Noise*: As $\alpha \to 0$, the order $L-1$ plays a more significant role than B/f_s in evaluating δ_{Z_R}. This can be seen from (4.45). Further still, as $\alpha \to 0$, (4.57) converges to (4.54). We may rewrite (4.57) as

$$A_D < d_{\alpha,\tilde{L}} L^{1-\alpha}, \tag{4.62}$$

where

$$d_{\alpha,\tilde{L}} = \frac{1}{(2\tilde{L})^{1-\alpha}} + \frac{2^{\alpha}}{(1-\alpha)\pi^{\alpha}}\left(1 - \frac{1}{(2\tilde{L})^{1-\alpha}}\right). \tag{4.63}$$

As $\alpha \to 0$, $d_{\alpha,\tilde{L}}$ depends less on $\tilde{L}$ and α and is almost constant.

On plugging (4.57) and (4.52) in (4.49) we see a direct relationship between δ_{Z_R}, δ_w, B/f_s and $L-1$. As $\delta_{Z_R}^{\alpha} \propto A_D$, the results in (4.59), (4.61) and (4.62) for A_D are easily extended to $\delta_{Z_R}^{\alpha}$.

Practical impulsive noise is usually approximated well by AWSαSN for α in the range of 1.5 and 1.9 [82]. In this range, B/f_s plays a pivotal role in determining δ_{Z_R}, as depicted by (4.59). In the literature, various SNR measures have been introduced to analyze error performance of digital communication systems [58]. These measures are inversely proportional to $\delta_{Z_R}^2$. Thus for practical impulsive noise,

$$\mathrm{SNR}_{\mathrm{measure}} \propto \left(\frac{f_s}{B}\right)^{\frac{2(\alpha-1)}{\alpha}} = \left(\frac{f_s}{B}\right)\left(\frac{f_s}{B}\right)^{\frac{\alpha-2}{\alpha}}. \tag{4.64}$$

In comparison to the traditional Gaussian SNR, which varies proportionally with $\frac{f_s}{B}$, we see that there is an additional term. As $f_s > B$, this term incurs an SNR *loss* on the order of

$$\mathrm{SNR}_{\mathrm{loss}} = 10\left(\frac{2}{\alpha} - 1\right)\log_{10}\left(\frac{f_s}{B}\right) \tag{4.65}$$

on the decibel scale. For a given $\frac{f_s}{B}$, the loss in SNR varies linearly with $\left(\frac{2}{\alpha} - 1\right)$. Therefore, the ratio $\frac{f_s}{B}$ should be taken under account in receiver design. This is covered in more detail in the next chapter. Similarly, if the noise is more impulsive, we may use (4.60) or (4.62) to see how the SNR varies in impulsive noise scenarios.

4.4 Summary

In this chapter we have analyzed complex baseband noise derived from passband AWSαSN. All baseband noise samples are proven to be IID and any given noise sample is shown to be SαS. The characteristics of the resulting noise are dissimilar to those obtained in the Gaussian case. It has been shown that the real and imaginary components of each sample are generally dependent and the distribution of each sample may be non-isotropic. The baseband noise samples are identically SαS and have star-like distributions. The distribution is completely determined by the system parameters. Varying these parameters allows constructing multi-tailed bivariate PDF structures with the tails always being uniformly distributed around the origin. The scale parameter of the baseband noise distribution has been analyzed and bounds have been proposed that show it to be a function of the system parameters and noise impulsiveness.

Chapter 5

Receiver Design for Single-Carrier Systems

In a single-carrier scheme, a symbol is selected from an M-point constellation lying in the complex plane. The in-phase and quadrature (I & Q) components are the real and imaginary parts of the symbol, respectively. These are then upsampled, pulse-shaped, converted to passband and transmitted. At the receiver, the exact opposite is performed to estimate the transmitted symbol which now also has the added element of complex noise. Conventionally, the receiver is optimized for AWGN. This results in the linear system discussed in Section 4.2 [1]. The passband-to-baseband conversion process is also performed in the continuous-time domain. If pure AWSαSN is passed through this receiver, the resulting baseband noise samples are complex SαS *circular symmetric* (or *isotropic*) random variables. In Section 4.2, we presented a CF approach to characterize the joint-distribution of a baseband noise sample. The isotropic configuration was intuitively explained as a limiting argument of the number of tails in the resulting bivariate PDF. We discuss this with more mathematical detail in Section 5.2.1.

In the previous chapter we investigated the statistics of complex baseband noise derived from passband non-Gaussian AWSαSN and found the resulting distribution to be radically different from its Gaussian counterpart due to the introduction of *uniform passband sampling*. By varying certain physical parameters we may attain different non-isotropic distributions. A question of choosing the best statistical configuration that minimizes the error performance arises. Amongst these PDFs, the one with *independent* components is of special interest as it can be exploited to attain the best possible ML error performance if the I and Q channels are processed separately. We show this in Section 5.2.2. By introducing efficient constellations, suitable baseband detectors and passband sampling, the uncoded error performance of the conventional (linear) receiver can be enhanced given the I and Q components of the transmitted signal are decoded separately.

We further show that there is non-negligible SNR degradation if the passband-to-baseband process is linear. The performance may be improved by sacrificing the linearity of the system. Various non-linear estimation and joint-detection schemes are discussed and their error performance analyzed. It is shown that if the receiver bandwidth is large enough, impulsive noise may be effectively countered.

It is pertinent to mention, that in some practical scenarios, implementing passband sampling may not be feasible. For example, in wireless communications the signal is transmitted via RF waves operating at high-frequencies (hundreds/thousands of MHz) [1]. Due to the limitation of current technologies, sampling at the Nyquist rate will result in costly hardware implementation. However, in underwater communications, signal transmission is performed via acoustic waves [13]. Typical ranges for carrier frequencies run into tens of kHz, a significant difference from those adopted in wireless communications [82].

Sampling the passband signal is therefore practically feasible and is employed in some underwater modems [82].

The work presented in this chapter has been published in [79], [80], [83]. We now introduce notation and concepts specific to single-carrier communication that will allow us to construct good receivers in AWSαSN.

5.1 Transmission & Reception via Orthonormal Signaling

Assuming memoryless modulation, the passband transmit-receive equation is given by

$$r(t) = s_i(t) + w(t) \tag{5.1}$$

where $s_i(t)\ \forall\ i \in \{0, 1, \ldots, M-1\}$ is the transmitted passband signal corresponding to the i^{th} symbol in the constellation of size M, $w(t)$ is a continuous-time AWSαSN process and $r(t)$ is the corresponding received signal. We denote the PDF of any passband noise sample by $f_W(\cdot)$ where $W \sim \mathcal{S}(\alpha, \delta_w)$. In an M-QAM signaling scheme, $s_i(t)$ is written as

$$\begin{aligned} s_i(t) &= \Re\left\{\sqrt{\frac{2\mathcal{E}_{x_i}}{\mathcal{E}_g}} g(t) e^{j\phi_i} e^{j2\pi f_c t}\right\} \\ &= \sqrt{\frac{2\mathcal{E}_{x_i}}{\mathcal{E}_g}} g(t)\cos(2\pi f_c t + \phi_i) \end{aligned} \tag{5.2}$$

where $0 \leq t < T$, $g(t)$ is a real baseband pulse-shaping signal of duration T and f_c is the carrier frequency [1]. The symbol rate is $1/T$ and $f_c = \xi/T$ for some $\xi \in \mathbb{Z}^+$, i.e., the carrier frequency is a multiple of the baseband symbol rate. In the spectral domain, $g(t)$ is band limited to $[-\frac{\beta}{2T}, \frac{\beta}{2T}]$, where $\beta \geq 1$ is a measure of the *excess* bandwidth relative to $1/T$. Though no signal can be time-limited and band-limited simultaneously, practically only the *significant* part of the spectrum is considered. Therefore, β is assumed a finite value. To avoid distortion in the passband signal, $f_c > \frac{\beta}{2T}$ and therefore $\xi > \beta/2$. Typically, f_c is set to be *at least* a few multiples greater than $\frac{\beta}{2T}$. The energies of $s_i(t)$ and $g(t)$ over $t \in [0, T)$ are denoted by $\mathcal{E}_{x_i}$ and $\mathcal{E}_g$, respectively. The baseband symbol $\sqrt{\mathcal{E}_{x_i}} e^{j\phi_i}$ is represented by the constellation point $(x_{I_i}, x_{Q_i}) = (\sqrt{\mathcal{E}_{x_i}}\cos(\phi_i), \sqrt{\mathcal{E}_{x_i}}\sin(\phi_i))$ in the complex plane.

Conventionally, orthonormal signaling is used to represent the passband modulated signal [1]. This is highlighted below:

$$\begin{aligned} s_i(t) &= \sqrt{\mathcal{E}_{x_i}}\cos(\phi_i)\ell_I(t) + \sqrt{\mathcal{E}_{x_i}}\sin(\phi_i)\ell_Q(t) \\ &= x_{I_i}\ell_I(t) + x_{Q_i}\ell_Q(t). \end{aligned} \tag{5.3}$$

On comparison with (5.2),

$$\ell_I(t) = \sqrt{\frac{2}{\mathcal{E}_g}}g(t)\cos(2\pi f_c t) \quad \text{and} \tag{5.4}$$

$$\ell_Q(t) = -\sqrt{\frac{2}{\mathcal{E}_g}}g(t)\sin(2\pi f_c t). \tag{5.5}$$

We observe that $\ell_I(t)$ and $\ell_Q(t)$ is an orthonormal basis over $t \in [0, T)$, i.e.,

$$\int_0^T \ell_I^2(t) = \int_0^T \ell_Q^2(t) = 1, \quad \int_0^T \ell_I(t)\ell_Q(t) = 0. \tag{5.6}$$

As $\ell_I(t)$ and $\ell_Q(t)$ are periodic over $t \in [kT, (k+1)T)\ \forall\ k \in \mathbb{Z}$ and $w(t)$ is stationary, the transmit-receive equation in (5.1) can be mapped onto the interval $[0, T)$ for any $t \in \mathbb{R}$. Thus we restrict our analysis to this interval.

The elegance of the representation in (5.3) is that one determines the I and Q components of the transmitted symbol by mere inspection. At the receiver, one may retrieve x_{I_i} and x_{Q_i} by multiplying (5.3) with $\ell_I(t)$ and $\ell_Q(t)$, respectively, and integrating over $t \in [0, T)$ [1]. If the same process is applied to the corrupted signal in (5.1), the resultant output can be expressed in the following form:

$$\mathbf{y} = \mathbf{x}_i + \mathbf{z} \tag{5.7}$$

where

$$\mathbf{y} = \begin{bmatrix} y_I \\ y_Q \end{bmatrix}, \; \mathbf{x}_i = \begin{bmatrix} x_{I_i} \\ x_{Q_i} \end{bmatrix} \text{ and } \mathbf{z} = \begin{bmatrix} z_I \\ z_Q \end{bmatrix}.$$

Thus the continuous signal form in (5.1) is converted to the vector form in (5.7) which is termed as the baseband transmit-receive equation. This is then followed by baseband detection to estimate the transmitted symbol. Given equiprobable symbols and (5.7), ML detection is optimal in reducing the error probability at the receiver. Mathematically, this is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\mathtt{arg\,max}} f_{\mathbf{z}}(\mathbf{y} - \mathbf{x}_l) \tag{5.8}$$

where $f_{\mathbf{z}}(\cdot)$ is the bivariate PDF of $\mathbf{z}$ and $\mathbb{M}$ is the set of all symbols in the constellation.

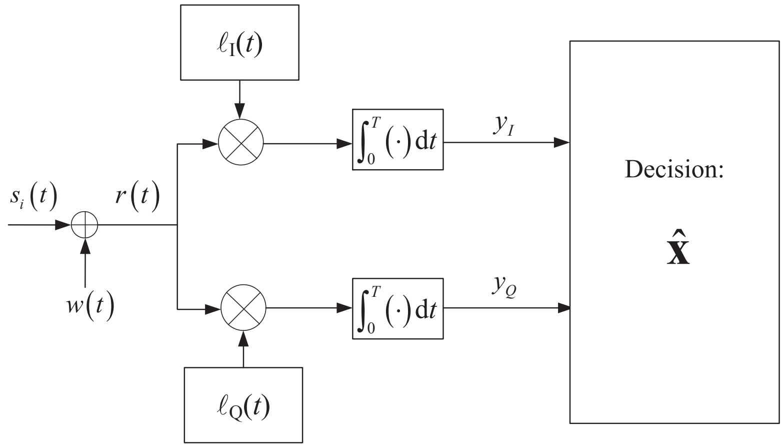


Figure 5.1: Conventional continuous-time correlator-based receiver implementation.

In Fig. 5.1 we present the receiver block structure based on the orthonormal signaling concepts discussed above. We term this as the *conventional receiver* and note it to be a linear system. The scheme is optimal in the ML-sense if $w(t)$ is an AWGN process [1]. In this case, $\mathbf{z}$ is an isotropic Gaussian random vector. The components of $\mathbf{z}$ are IID $\mathcal{N}(0, N_0/2)$ where $N_0/2$ is the two-sided PSD of the AWGN channel. In non-Gaussian AWSαSN, $\mathbf{z}$ is also an isotropic SαS vector. This is shown in Section 5.2.1.

An isotropic SαS PDF $f_{\mathbf{z}}(\mathbf{x})$ has favorable geometric properties. Mathematically, $f_{\mathbf{z}}(\mathbf{x})$ is a function of $\|\mathbf{x}\|$, the Euclidean norm of $\mathbf{x}$ [9]. This implies that its equiprobable density contours are in the form of concentric circles around the origin. Further still, the marginal PDFs are *identical* [9]. In the Gaussian case, an isotropic distribution is only possible if $\mathbf{z}$ has independent components [1]. However, for $\alpha \neq 2$, $\mathbf{z}$ has *dependent* components [8], [9]. The conventional receiver performs poorly in non-Gaussian AWSαSN as it is a linear system and does not exploit the dependency between the I and Q components [9], [79]. It is therefore imperative that the receiver be designed more robust to impulsive noise.

In this chapter, we study robust single-carrier receivers that fall into two broad categories:

1. Soft-Estimates & Baseband Detection: Soft-estimates of the transmitted symbol are generated via a linear or non-linear operation on the passband samples. This is the passband-to-baseband conversion process. The result is then given to a detector which maps the estimate onto a constellation point. The conventional receiver falls under this category.

2. Joint-Detection: Instead of initially converting to baseband, one can directly map the passband samples onto a constellation point.

We designate a section to either category.

5.2 Soft-Estimates & Baseband Detection

5.2.1 Conventional Passband-to-Baseband Conversion

If the schematic in Fig. 5.1 is employed in passband AWSαSN, then $\mathbf{z}$ is a bivariate *isotropic* SαS vector. These properties are proven below:

Proposition 4: If $w(t)$ is a continuous-time real AWSαSN process and the conventional receiver is employed, then $\mathbf{z}$ is *isotropic* in the limit $f_c \to \infty$.

Proof: We can express $\mathbf{z}$ as

$$z_I + jz_Q = \int_0^T w(t)(\ell_I(t) + j\ell_Q(t))\mathrm{d}t \tag{5.9}$$

$$= \sqrt{\frac{2}{\mathcal{E}_g}} \int_0^T w(t)g(t)e^{j2\pi f_c t}\mathrm{d}t. \tag{5.10}$$

The integral in (5.10) may be divided into a sum of ξ integrals over adjacent intervals:

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int\limits_{\lambda T/\xi}^{(\lambda+1)T/\xi} w(t)g(t)e^{j2\pi f_c t}\mathrm{d}t \tag{5.11}$$

We note that $\exp(j2\pi f_c t)$ is periodic in t with period T/ξ, thus we may express (5.11) as

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int\limits_{\lambda T/\xi}^{(\lambda+1)T/\xi} w(t)g(t)e^{j2\pi f_c (t-\lambda T/\xi)}\mathrm{d}t \tag{5.12}$$

Applying a change of variable from $t - \lambda T/\xi$ to t results in

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int_0^{T/\xi} w\left(t + \frac{\lambda T}{\xi}\right) g\left(t + \frac{\lambda T}{\xi}\right) e^{j2\pi f_c t}\mathrm{d}t \tag{5.13}$$

The integration and summation operations may be interchanged to get

$$\sqrt{\frac{2}{\mathcal{E}_g}} \int_0^{T/\xi} \sum_{\lambda=0}^{\xi-1} w\left(t + \frac{\lambda T}{\xi}\right) g\left(t + \frac{\lambda T}{\xi}\right) e^{j2\pi f_c t}\mathrm{d}t \tag{5.14}$$

From (3.2) and (3.7), we note that

$$\sum_{\lambda=0}^{\xi-1} w\left(t+\frac{\lambda T}{\xi}\right) g\left(t+\frac{\lambda T}{\xi}\right) \stackrel{d}{=} w(t)c(t;\alpha,\xi,g(t)) \tag{5.15}$$

where

$$c(t;\alpha,\xi,g(t)) = \left(\sum_{\lambda=0}^{\xi-1}\left|g\left(t+\frac{\lambda T}{\xi}\right)\right|^{\alpha}\right)^{1/\alpha} \tag{5.16}$$

We observe that (5.16) corresponds to *sampling* $|g(t)|^{\alpha}$ at a rate of $f_c = \xi/T$. Therefore, the summation term in (5.16) is the sum of ξ samples of $g(t)$ which are uniformly spread over the interval $t \in [0, T)$. Consequently, as $\xi \to \infty$, we may express (5.16) as

$$c(t;\alpha,\xi,g(t)) = c(\alpha,\xi,g(t)) = \left(\frac{\xi}{T}\int_0^T |g(t)|^{\alpha}\,\mathrm{d}t\right)^{1/\alpha}. \tag{5.17}$$

Finally, from (5.15) and (5.17), we may express (5.14) as

$$z_I + jz_Q \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}}c(\alpha,\xi,g(t))\int_0^{T/\xi} w(t)e^{j2\pi f_c t}\mathrm{d}t. \tag{5.18}$$

Now for $\mathbf{z}$ to be isotropic,

$$z_I + jz_Q \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}}c(\alpha,\xi,g(t))\int_0^{T/\xi} w(t)e^{j(2\pi f_c t+\phi)}\mathrm{d}t \tag{5.19}$$

for all $\phi \in \mathbb{R}$. We note that the phasors $\exp(j2\pi f_c t)$ in (5.18) and $\exp(j2\pi f_c t + j\phi)$ in (5.19) complete a full rotation in the complex plane over $t \in [0, T/\xi)$ for any ϕ. This, coupled with stationary $w(t)$ proves that (5.18) and (5.19) are statistically equivalent, thus $\mathbf{z}$ is isotropic. □

Note that the formulation in (5.10)-(5.19) may be applied to *any* IID noise process (albeit with some modifications), irrespective of its samples being symmetric or asymmetric. Due to the transition in (5.16)-(5.17), proposition 4 is valid for $\xi \to \infty$ (the narrowband case as $f_c \gg \frac{1}{T}$). However, this condition can be somewhat relaxed for any band-limited $g(t)$. For example, if $g(t)$ is a rectangular pulse then (5.16) and (5.17) are both *equal* to $\xi^{1/\alpha}$ *for all* ξ. For general $g(t)$, f_c needs to be greater than the Nyquist rate of $|g(t)|^\alpha$ for (5.17) to be equivalent to (5.16). Two cases are highlighted below:

- For $\alpha = 2$, The Fourier transform of $g^2(t)$ can be expressed as the convolution of the spectra of $g(t)$ with itself. Consequently, $g^2(t)$ lies within $[-\beta/T, \beta/T]$ as its bandwidth is roughly twice that of $g(t)$. When sampled at f_c, the Nyquist criterion is satisfied for $\xi > 2\beta$.

- As $\alpha \to 0$, $|g(t)|^\alpha$ tends to the unit-amplitude rectangular pulse and therefore (5.17) is exact for any $\xi \in \mathbb{Z}^+$.

For other values of $\alpha \in (0, 2]$, ξ need only be a few multiples greater than β for $\mathbf{z}$ to be *sufficiently* isotropic. Increasing ξ would cause negligible change in the distribution of $\mathbf{z}$.

Proposition 5: If $w(t)$ is a continuous-time real AWSαSN process, then $\mathbf{z}$ is a SαS vector in the conventional receiver.

Proof: We may write (5.9) in vector form:

$$\mathbf{z} = \begin{bmatrix} \int_0^T w(t)\ell_I(t)\mathrm{d}t \\ \int_0^T w(t)\ell_Q(t)\mathrm{d}t \end{bmatrix}. \tag{5.20}$$

Let $\mathbf{z}^{(i)}$ $\forall\ i \in \{1, 2, \ldots, K\}$ be IID copies of $\mathbf{z}$, then from (3.7) we have

$$\begin{aligned} \sum_{i=1}^{K} a_i \mathbf{z}^{(i)} &= \sum_{i=1}^{K} a_i \begin{bmatrix} \int_0^T w^{(i)}(t)\ell_I(t)\mathrm{d}t \\ \int_0^T w^{(i)}(t)\ell_Q(t)\mathrm{d}t \end{bmatrix} \\ &= \begin{bmatrix} \int_0^T \left(\sum_{i=1}^{K} a_i w^{(i)}(t)\right) \ell_I(t)\mathrm{d}t \\ \int_0^T \left(\sum_{i=1}^{K} a_i w^{(i)}(t)\right) \ell_Q(t)\mathrm{d}t \end{bmatrix} \end{aligned} \tag{5.21}$$

where $w^{(i)}(t)$ $\forall\ i \in \{1, 2, \ldots, K\}$ are IID copies of $w(t)$. On invoking (3.6), we get

$$\sum_{i=1}^{K} a_i \mathbf{z}^{(i)} \stackrel{d}{=} \begin{bmatrix} \int_0^T cw(t)\ell_I(t)\mathrm{d}t \\ \int_0^T cw(t)\ell_Q(t)\mathrm{d}t \end{bmatrix} = c\mathbf{z} \tag{5.22}$$

where c may be evaluated from (3.7):

$$c = \left(\sum_{i=1}^{K} |a_i|^{\alpha}\right)^{1/\alpha} \tag{5.23}$$

Therefore, we conclude that $\mathbf{z}$ is SαS with characteristic exponent α. □

Due to the linearity of the conventional receiver, the characteristic exponent of $\mathbf{z}$ is equivalent to that of the passband noise samples. From the discussion in Section 3.2.2, the components of isotropic $\mathbf{z}$ for $\alpha \neq 2$ are identically distributed but dependent. In Appendix A.1 we evaluate the scale parameters of z_I and z_Q as a function of the bandwidth. The conventional receiver does not exploit the dependency between the components of $\mathbf{z}$. As shown in the next Section, this severely inhibits error performance of the system. This situation can be avoided entirely by introducing passband sampling (under some constraints) at the receiver.

5.2.2 Where Conventional Conversion Fails

We have discussed the baseband noise statistics of the *discretized* receiver in Chapter 4. The noise statistics vary with respect to the system parameters. From an information theoretic perspective, we now explain why the conventional receiver does not perform well and which baseband noise configuration can be exploited to give us the best error performance:

We assume that the transmitted information is fully preserved in the conversion from $r(t)$ to $\mathbf{y}$. By Nyquist's theorem, $s(t)$ may be sampled at any $f_s > 2f_c + \beta/T \Rightarrow Tf_s > 2\xi + \beta$ to avoid aliasing and hence loss of information. If this axiom is satisfied, then irrespective of whatever Tf_s may be, $\mathbf{x}_i$ can be fully recovered from $s_i(n/f_s)\ \forall\ n \in \{0, 1, \ldots, \lfloor Tf_s \rfloor - 1\}$. If the accompanying noise samples $w(n/f_s)\ \forall\ n \in \{0, 1, \ldots, \lfloor Tf_s \rfloor - 1\}$ are passed through the estimator, the information within $\mathbf{z}$ remains the same for *any given* Tf_s. In other words, only the *noise component* that affects the transmitted symbol is retained. However, this information may vary for *different* non-lossy estimation schemes. Mathematically, the retained noise information is quantified by the joint-entropy $H(z_I, z_Q)$ of the components of $\mathbf{z}$. This may be expressed as

$$H(z_I, z_Q) = H(z_I) + H(z_Q) - I(z_I; z_Q) \tag{5.24}$$

where $H(z_I)$ and $H(z_Q)$ are the self-entropies of z_I and z_Q, respectively, and $I(z_I; z_Q)$ is the information shared between them [84]. For *any given* unbiased estimator and Tf_s, $H(z_I, z_Q)$ will be *constant*. In the presence of non-zero $I(z_I; z_Q)$, the self-entropies increase correspondingly to maintain equality in (5.24). Therefore, receivers that process the I and Q channels separately (like the one in Fig. 5.1) perform better when $I(z_I; z_Q) = 0$ as $H(z_I)$ and $H(z_Q)$ will be at their respective minimums. For $f_s = 4f_c$, the Nyquist criteria is fulfilled if $\xi > \beta/2$. Further still, it is only in this case that z_I and z_Q are independent for *any* IID noise process $w(t)$ (see corollary 3, Section 4.2.3). This guarantees $H(z_I)$ and $H(z_Q)$ to be at their respective minimums.

From (3.24), if $\mathbf{z}$ is an SαS vector with IID components, its PDF may be split as

$$f_{\mathbf{z}}(\mathbf{x}) = f_Z(x_1) f_Z(x_2). \tag{5.25}$$

where $z_I \overset{d}{=} z_Q \overset{d}{=} Z$. In Section 4.2, an intuitive discussion pertaining to the identicalness of z_I and z_Q for general f_c and f_s was presented. We show that z_I and z_Q are in fact identical for $f_s = 4f_c$ in Section 5.2.3. The PDF corresponding to (5.25) will have four 'tails' in the non-Gaussian case. These tails are positioned along both the positive and negative directions of each axis in the complex plane.

An instance of this is shown for the standard Cauchy case in Fig. 4.4b.

5.2.3 Linear Baseband Conversion with Passband Sampling

As discussed, if the I and Q channels are processed separately, the case of $\mathbf{z}$ with independent components offers the best error performance over all possible statistical structures. We therefore focus on this particular scenario. We modify our initial definition slightly to use the square bracket notation to denote discrete-time signals sampled at $f_s = 4f_c$. In the remainder of this chapter, we assume $f_s = 4f_c$ unless explicitly stated otherwise. For $f_s = 4f_c$, (5.3) reduces to

$$s_i[n] = x_{I_i}\ell_I[n] + x_{Q_i}\ell_Q[n] \tag{5.26}$$

where

$$\ell_I[n] = \sqrt{\frac{2}{\mathcal{E}_g}}g[n]\cos(\pi n/2) \quad \text{and} \tag{5.27}$$

$$\ell_Q[n] = -\sqrt{\frac{2}{\mathcal{E}_g}}g[n]\sin(\pi n/2). \tag{5.28}$$

We note that only one of the functions in (5.27) and (5.28) is non-zero at any given $n \in \mathbb{Z}$. This implies that any *sample* of $s_i[n]$ consists of either the I or Q component, never a combination of both. We can split (5.26) into a superimposition of two distinct sequences:

$$s_i[n] = \begin{cases} x_{I_i}\ell_I[n] & \forall\, n \in \{0, 2, \ldots, 4\xi - 2\} \\ x_{Q_i}\ell_Q[n] & \forall\, n \in \{1, 3, \ldots, 4\xi - 1\} \end{cases} \tag{5.29}$$

By substituting variables in (5.29), we can separate the I and Q components completely:

$$
\begin{aligned}
s_i[2n] &= x_{I_i}\ell_I[2n] \quad \text{and} \\
s_i[2n+1] &= x_{Q_i}\ell_Q[2n+1]
\end{aligned}
\tag{5.30}
$$

$\forall\ n \in \{0, 1, \ldots, 2\xi - 1\}$. From (5.1), the sampled received signal is

$$
r[n] = s_i[n] + w[n] \tag{5.31}
$$

$\forall\ n \in \{0, 1, \ldots, 4\xi - 1\}$. Following (5.29) and (5.30), the transmit-receive equation can be expressed as two parallel channels:

$$
r_I[n] = r[2n] = s_i[2n] + z_I[n] \tag{5.32}
$$

$$
r_Q[n] = r[2n+1] = s_i[2n+1] + z_Q[n] \tag{5.33}
$$

where

$$
\begin{aligned}
z_I[n] &= w[2n] \quad \text{and} \\
z_Q[n] &= w[2n+1]
\end{aligned}
\tag{5.34}
$$

$\forall\ n \in \{0, 1, \ldots, 2\xi - 1\}$. The expression in (5.34) can be thought of two independent (yet similar) AWSαSN processes. As the noise samples contaminating $s_i[2n]$ and $s_i[2n+1]$ are mutually independent, sampling at $f_s = 4f_c$ and separately processing (5.32) and (5.33) is *sufficient* to ensure that $\mathbf{z}$ will have independent components. This corroborates with corollary 3 in

Section 4.2.3. Further still, the arguments in this section may be extended to any noise process $w(t)$ that has IID samples.

Mathematically, the integrals in (5.6) reduce to the following sums:

$$\sum_{n=0}^{4\xi-1} \ell_I^2[n] = \sum_{n=0}^{4\xi-1} \ell_Q^2[n] = f_s \tag{5.35}$$

$$\sum_{n=0}^{4\xi-1} \ell_I[n]\ell_Q[n] = 0 \tag{5.36}$$

Eq. (5.36) is straightforward, we however prove (5.35) below.

$$\begin{aligned}\sum_{n=0}^{4\xi-1} \ell_I^2[n] &= \frac{2}{\mathcal{E}_g}\sum_{n=0}^{4\xi-1} g^2[n]\cos^2(\pi n/2) \\ &= \frac{2}{\mathcal{E}_g}\sum_{n=0}^{4\xi-1} \frac{g^2[n]}{2}(1+\cos(\pi n)) \\ &= \frac{2}{\mathcal{E}_g}\sum_{n=0}^{4\xi-1} \frac{g^2[n]}{2}. \end{aligned} \tag{5.37}$$

As discussed in Section 5.2.1, the Nyquist criteria for $g^2(t)$ is met if it is sampled at a rate greater than $2\beta/T$. Therefore, if $f_s > 2\beta/T \Rightarrow \xi > \beta/2$, we may express $\mathcal{E}_g$ as

$$\mathcal{E}_g = \frac{1}{f_s}\sum_{n=0}^{4\xi-1} g^2[n]. \tag{5.38}$$

Substituting this back into (5.37) gives us

$$\sum_{n=0}^{4\xi-1} \ell_I^2[n] = \frac{2}{\mathcal{E}_g} \times \frac{\mathcal{E}_g f_s}{2} = f_s. \tag{5.39}$$

Like its continuous counterpart, the goal of the linear discrete receiver is

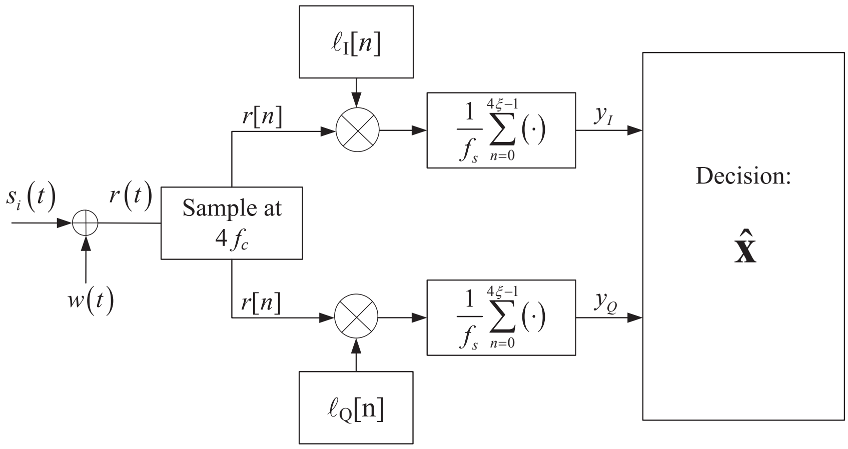


Figure 5.2: Linear receiver schematic with $f_s = 4f_c$.

to re-acquire the I and Q components of the transmitted symbol. This is accomplished by initially estimating $\mathbf{y}$ from $r[n]$. The properties of $\ell_I[n]$ and $\ell_Q[n]$ in (5.35) and (5.36) may be exploited to achieve this. We present the receiver structure in Fig. 5.2 and term this as the *discretized linear receiver.*

Following (5.34), $\mathbf{z}$ will have independent components if the passband noise is AWSαSN. These are expressed as

$$z_I = \frac{1}{f_s}\sum_{n=0}^{4\xi-1} w[n]\ell_I[n], \quad z_Q = \frac{1}{f_s}\sum_{n=0}^{4\xi-1} w[n]\ell_Q[n]. \tag{5.40}$$

From (3.2), z_I and z_Q are each SαS random variables with characteristic exponent α. This in turn implies that $\mathbf{z}$ is an SαS vector. Denoting the scale parameters of z_I and z_Q by δ_{z_I} and δ_{z_Q}, respectively, from (3.2) and (3.7) we can express z_I as

$$z_I \overset{d}{=} \frac{1}{f_s}\left(\sum_{n=0}^{4\xi-1} |\ell_I[n]|^{\alpha}\right)^{1/\alpha} w[n].$$

Therefore, we can express δ_{z_I} as

$$\delta_{z_I} = \frac{\delta_w}{f_s}\left(\sum_{n=0}^{4\xi-1} |\ell_I[n]|^\alpha\right)^{1/\alpha}. \tag{5.41}$$

On substituting $\ell_I[n]$ with (5.27), we get

$$\begin{aligned}\delta_{z_I} &= \frac{\delta_w}{f_s}\sqrt{\frac{2}{\mathcal{E}_g}}\left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha\right)^{1/\alpha} \\ &\approx \frac{\delta_w}{f_s}\sqrt{\frac{2}{\frac{2}{f_s}\sum_{n=0}^{2\xi-1} g^2[2n]}}\left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha\right)^{1/\alpha} &(5.42)\\ &= \frac{\delta_w}{\sqrt{f_s}}\frac{\left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha\right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g^2[2n]\right)^{1/2}}. &(5.43)\end{aligned}$$

We observe that $g[2n]$ results from sampling $g(t)$ at a rate of $f_s/2$. The approximation for $\mathcal{E}_g$ in (5.42) is valid as long as $f_s/2 > 2\beta/T \Rightarrow \xi > \beta$. The form in (5.43) is intuitive as it depicts δ_{z_I} varying proportionally with $\delta_w/\sqrt{f_s}$ for all $\xi > \beta$. Do note that f_s is the available receiver bandwidth. Similarly, δ_{z_Q} may also be evaluated from (5.28) and (5.40):

$$\delta_{z_Q} = \frac{\delta_w}{\sqrt{f_s}}\frac{\left(\sum_{n=0}^{2\xi-1} |g[2n+1]|^\alpha\right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g^2[2n+1]\right)^{1/2}}. \tag{5.44}$$

Given that the Nyquist criterion is also satisfied for $|g(t)|^\alpha$ with sampling rate $f_s/2$, we note that (5.43) and (5.44) are equivalent. Therefore, $\mathbf{z}$ has IID components and a PDF of the form in (5.25) with $Z \sim \mathcal{S}(\alpha, \delta_z)$.

Specifically for the Gaussian case, $\delta_z = \delta_w/\sqrt{f_s}$. Therefore, z_I and z_Q are each $\mathcal{N}(0, 2\delta_w^2/f_s)$. The variance or *power* of a band-limited AWGN channel can be written as a product of its PSD ($N_0/2$) and bandwidth (f_s). In our case $2\delta_w^2 = N_0 f_s/2$, which implies that $z_I \overset{d}{=} z_Q \sim \mathcal{N}(0, N_0/2)$. Thus, $\mathbf{z}$ is statistically equivalent to its counterpart in the conventional receiver in Fig. 5.1. Intuitively, this is of no surprise as the transmitted information is kept intact in the sampling process and the operations in both receivers are identical. In fact, if $w[n]$ are IID Gaussian, the ML soft estimates of x_{I_i} and x_{Q_i} in (5.31) are determined by the linear correlator [1]:

$$y_I = \frac{\sum_{n=0}^{4\xi-1} r[n]\ell_I[n]}{\sum_{n=0}^{4\xi-1} \ell_I^2[n]} = \frac{\sum_{n=0}^{4\xi-1} r[n]\ell_I[n]}{f_s} \tag{5.45}$$

$$y_Q = \frac{\sum_{n=0}^{4\xi-1} r[n]\ell_Q[n]}{\sum_{n=0}^{4\xi-1} \ell_Q^2[n]} = \frac{\sum_{n=0}^{4\xi-1} r[n]\ell_Q[n]}{f_s} \tag{5.46}$$

On comparing Fig. 5.2 and the expressions above, we note that the implementation is indeed ML-based.

5.2.4 Non-Linear Baseband Conversion

As $w(n/f_s)\ \forall\ n \in \{0, 1, \ldots, \lfloor Tf_s \rfloor - 1\}$ are IID, we note that $\mathbf{y}$ can be evaluated from the minimization

$$\mathbf{y} = \underset{\boldsymbol{\mu}}{\texttt{arg min}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} -\log \rho\left(r\left(\frac{n}{f_s}\right) - \mu_I \ell_I\left(\frac{n}{f_s}\right) - \mu_Q \ell_Q\left(\frac{n}{f_s}\right)\right) \tag{5.47}$$

where $\boldsymbol{\mu} = [\mu_I, \mu_Q]^T \in \mathbb{R}^2$, $\rho(x) \in \mathbb{R}^+ \ \forall\ x \in \mathbb{R}$ and $f_s > 2f_c + \beta/T$. The expression in (5.47) stems from robust generalized ML estimation (or M-estimator) theory [65]. If $\rho(\cdot) = f_W(\cdot)$, then $\mathbf{y}$ is the ML estimate of $\boldsymbol{\mu}$. In the context of digital communications, $\mathbf{y}$ is the soft-ML estimate of $\mathbf{x}_i$. Substituting $\rho(\cdot)$ with the Gaussian PDF corresponding to $\mathcal{S}(2, \delta_w)$ in (5.47), results in two single-variable minimizations:

$$y_I = \underset{\mu_I}{\texttt{arg min}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} \frac{\mu_I^2}{f_s} - 2\mu_I \ell_I \left(\frac{n}{f_s}\right) r \left(\frac{n}{f_s}\right)$$

$$y_Q = \underset{\mu_Q}{\texttt{arg min}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} \frac{\mu_Q^2}{f_s} - 2\mu_Q \ell_Q \left(\frac{n}{f_s}\right) r \left(\frac{n}{f_s}\right)$$

Therefore, the ML estimator of $\boldsymbol{\mu}$ *separately processes* the I and Q channels *for all* f_s in the Gaussian case and its form is similar to (5.45) and (5.46). However, (5.47) cannot be split into separate minimizations of μ_I and μ_Q if $w(t)$ is non-Gaussian AWSαSN.

Proposition 6: For $f_s = 4f_c$, the bivariate minimization in (5.47) is equivalent to individually evaluating

$$y_I = \underset{\mu_I}{\texttt{arg min}} \sum_{n=0}^{2\xi-1} -\log \rho(r_I[n] - \mu_I \ell_I[2n]) \quad \text{and} \tag{5.48}$$

$$y_Q = \underset{\mu_Q}{\texttt{arg min}} \sum_{n=0}^{2\xi-1} -\log \rho(r_Q[n] - \mu_Q \ell_Q[2n+1]) \tag{5.49}$$

for *all* $\rho(x)$.

Proof: For $f_s = 4f_c$, (5.47) becomes

$$\mathbf{y} = \underset{\boldsymbol{\mu}}{\texttt{arg min}} \sum_{n=0}^{4\xi-1} -\log \rho\left(r[n] - \mu_I \ell_I[n] - \mu_Q \ell_Q[n]\right) \tag{5.50}$$

From the discussion in Section 5.2.3, we know that $\ell_I[n]$ is non-zero for $n \in \{0, 2, \ldots, 4\xi - 2\}$ and $\ell_Q[n]$ for $n \in \{1, 3, \ldots, 4\xi - 1\}$. Therefore, (5.50) may be

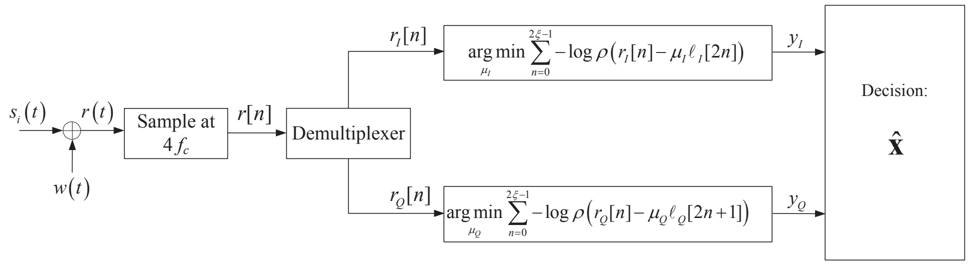


Figure 5.3: General receiver schematic with $f_s = 4f_c$.

rewritten as

$$\mathbf{y} = \underset{\boldsymbol{\mu}}{\texttt{arg min}} \left(\sum_{n=0}^{2\xi-1} -\log\rho\left(r[2n] - \mu_I \ell_I[2n]\right) \right. \\ \left. + \sum_{n=0}^{2\xi-1} -\log\rho\left(r[2n+1] - \mu_Q \ell_Q[2n+1]\right) \right). \quad (5.51)$$

We observe that evaluating (5.51) is equivalent to *individually* minimizing over μ_I and μ_Q to get y_I and y_Q, respectively. Since no assumption was made about $\rho(x)$, (5.48) and (5.49) hold for all $\rho(x)$. □

Corollary 4: The separation in (5.51) is possible for all $\rho(x)$ *if and only if* $f_s = 4f_c$. This is a direct consequence of the fact that $r(n/f_s)$ splits into (5.32) and (5.33) *if and only if* $f_s = 4f_c$.

Therefore, $f_s = 4f_c$ is a *sufficient* condition for any scheme to achieve the ML estimate of $\boldsymbol{\mu}$ in (5.47) if the estimation is done individually for the I and Q components. In Fig. 5.3, we present a general uncoded receiver schematic that optimizes error performance if the I and Q channels are processed separately.

From an implementation perspective, the ML estimator of $\boldsymbol{\mu}$ in AWSαSN may not be desirable due to the lack of closed form SαS PDFs. As highlighted in Chapter 2, numerical methods to exist to evaluate the PDF, but implementing them in real-time may still be cumbersome. Therefore, the next step is to find good functions for $\rho(\cdot)$ that approach the ML estimator performance. Closed-form expressions also offer an intuitive feel to the design of a system. We use this in Section 5.4.2 where we discuss good methodologies for symbol placement in constellations.

Our focus for the remainder of this section will be on analyzing (5.48) and (5.49). As the expressions are similar, we drop the subscripts and deal with the general expression

$$y = \underset{\mu}{\texttt{arg}\,\texttt{min}} \sum_{n=0}^{2\xi-1} - \log \rho(x[n] - \mu\ell[n]). \tag{5.52}$$

The PDF of a Cauchy random variable $X \sim \mathcal{S}(1, \delta_w)$ is given by (3.9). On substituting this for $\rho(\cdot)$ in (5.52) and simplifying, we get the ML Cauchy estimator for μ:

$$y = \underset{\mu}{\texttt{arg}\,\texttt{min}} \sum_{n=0}^{2\xi-1} \log \left(\delta_w^2 + (x[n] - \mu\ell[n])^2\right). \tag{5.53}$$

The cost function in (5.53) consists of multiple local minimas/maximas in μ. To observe this, we can rewrite (5.53) equivalently as

$$y = \underset{\mu}{\texttt{arg}\,\texttt{min}} \prod_{n=0}^{2\xi-1} \left(\delta_w^2 + (x[n] - \mu\ell[n])^2\right). \tag{5.54}$$

Clearly, (5.54) is a 4ξ order polynomial in μ. Under certain regularity conditions, y tends towards a Gaussian distribution as $\xi \to +\infty$ [85], [86]. The ML

Cauchy estimator and its variants have been employed vastly in the literature to combat impulsive noise [59]. This approach is intuitively gratifying as Cauchy distributions share the heavy-tailed property associated with impulsive noise distributions. However, these estimates are still sub-optimal when $\alpha \neq 1$ and are not supported by any underlying theory.

The MMyF Estimate

The ML Cauchy estimate for μ in (5.53) may be seen in the light of a class of robust M-estimators; namely the *matched myriad filter* (MMyF) [58], [59], [61]. The MMyF estimate $y(K)$ of μ with *linearity parameter* $K \in \mathbb{R}^+$ is given by

$$y(K) = \underset{\mu_I}{\texttt{arg}\,\texttt{min}} \sum_{n=0}^{2\xi-1} \log\left(K^2 + \ell^2[n]\left(\frac{x[n]}{\ell[n]} - \mu\right)^2\right). \tag{5.55}$$

From observation, (5.55) is equivalent to the cost function in (5.53) when $K = \delta_w$, i.e., $y(\delta_w)$ is the Cauchy ML estimate of μ. By appropriately tuning K, the MMyF offers robustness in impulsive noise for all α. We highlight the following aspects of the MMyF [58]:

1. As $K \to +\infty$, the MMyF converges to the linear correlator, which is the optimum ML estimate in Gaussian noise, i.e., $\alpha = 2$.

2. As $K \to 0$, the MMyF becomes a mode-selector, i.e., the estimate is equal to the element in $\{x[n]/\ell[n]\}\ \forall\ n \in \{0, 1, \ldots, 2\xi - 1\}$ that has the largest frequency of repetition. If there is no repetition of elements, any one element is selected as the estimate. This is usually chosen from within a cluster of closely spaced values. The mode-selector is the optimal (ML)

estimator in extremely impulsive noise, i.e, $\alpha \to 0$.

Thus by varying K, one may achieve ML optimality for three scenarios within the SαS framework. We observe that by decreasing K the estimate of μ is made more robust to impulsive noise. Similarly, for mildly impulsive scenarios, we can consider higher values of K to achieve better results. Therefore, we may express the linearity parameter $K = K(\alpha, \delta_w)$ as a monotonically increasing function of α and δ_w that attains the three points of optimality: $K(2, \delta_w) = +\infty$, $K(1, \delta_w) = \delta_w$ and $K(0, \delta_w) = 0$. Further still, if $K(\alpha, \delta_w)$ offers the optimal estimate of μ for all $\alpha \in (0, 2]$ in (5.55), the scale parameter is separable, i.e. $K(\alpha, \delta_w) = K(\alpha)\delta_w$ [[59], Eq. 31]. The MMyF estimate in (5.55) can now be written as

$$y(K) = \underset{\mu}{\mathtt{arg\,min}} \sum_{n=0}^{2\xi-1} \log \left(K^2(\alpha)\delta_w^2 + \ell^2[n] \left(\frac{x[n]}{\ell[n]} - \mu \right)^2 \right). \tag{5.56}$$

In the literature, a heuristic function has been proposed for $K(\alpha)$ that works well for all $\alpha \in (0, 2]$ [59]:

$$K(\alpha) = \sqrt{\frac{\alpha}{2-\alpha}}. \tag{5.57}$$

The MMyF offers good near-optimal estimates of $\boldsymbol{\mu}$ in the general AWSαSN case. Like (5.53), the cost function in (5.56) has at most 4ξ minimas/maximas in μ. As the number of samples in (5.56) increases, $y(K)$ converges to a normal distribution for all $\alpha \in (0, 2]$ [53], [58]. Keeping this in mind, it is correct to assume that $\mathbf{y}$ has an isotropic Gaussian distribution for large values of ξ.

Thus, as $\xi \to +\infty$, Euclidean detection will be optimal given MMyF estimation. However, there still is a *residual* impulsive noise component when ξ is small. In such a case, the statistics of $\mathbf{z}$ needs to be determined before invoking ML (or near-ML) detection.

The L_p-Norm Estimator

Besides the Cauchy estimator and the MMyF, other functions known to perform well in impulsive noise may also be used as $\rho(x)$ in (5.52). As the objective is to approximate ML estimation as close as possible, it is logical to find *analytic* functions $f_{\bar{W}}(x)$ that closely resemble $f_W(x)$ and substitute them for $\rho(x)$. An example is the PDF

$$f_{\bar{W}}(x) = \frac{d_1}{\delta_{\bar{w}}} \exp(-d_2 \left| \frac{x}{\delta_{\bar{w}}} \right|^p) \tag{5.58}$$

where $0 < p < \alpha$. Here, d_1 and d_2 are positive (normalizing) constants and $\delta_{\bar{w}}$ is the scale parameter of the distribution. On substituting (5.58) for $\rho(x)$ in (5.52) and simplifying, we get

$$\begin{aligned} y &= \underset{\mu}{\texttt{arg}\,\texttt{min}} \sum_{n=0}^{2\xi-1} |x[n] - \mu\ell[n]|^p \\ &= \underset{\mu}{\texttt{arg}\,\texttt{min}} \|x[n] - \mu\ell[n]\|_p^p \end{aligned} \tag{5.59}$$

where $\|\cdot\|_p$ is the L_p-norm. Thus the L_p-norm is based on the approximation of $f_W(x)$ that is provided by (5.58). The L_p-norm for $0 < p < \alpha$ is convergent in the ergodic sense and is known to perform very well in impulsive noise [9], [10], [57]. By changing p one can tweak the 'tails' of the general PDF in (5.58). For $p = 2$, (5.58) is a Gaussian PDF. On the other hand, as $p \to 0$, $\bar{f}_W(x)$ becomes a constant (zero), i.e., $f_{\bar{W}}(x) \to (d_1/\delta_{\bar{w}})\exp(-d_2)$. This implies that the tails of (5.58) become increasingly heavier as $p \to 0$. In the medium-to-high SNR regime, errors are predominantly determined by the tail probabilities of the impulsive noise distribution. In this region, the value of p for which the estimate of μ is optimized will depend on α and ξ.

The Log-Norm Estimator

Similarly, the asymptotic PDF expression in (3.10) may be employed as $\rho(x)$ to get

$$y = \underset{\mu}{\texttt{arg}\,\texttt{min}} \sum_{n=0}^{2\xi-1} \log|x[n] - \mu\ell[n]|. \tag{5.60}$$

We term this as the *log-norm estimator*. On comparison with (5.56), we note that (5.60) is the MMyF estimate with $K(\alpha) = 0$ which corresponds to a mode-type estimator. This is not surprising as (3.10) assumes each $w[n]$ to be an impulse.

The cost functions in (5.56), (5.59) and (5.60) are in analytic forms. However, the estimators themselves cannot be represented in closed form. Therefore, the minimizations have to be numerically evaluated. One issue that arises is that the global minima cannot be generally found for a small number of samples as the cost functions will have multiple local minimums (traps). An exception to this is the L_p-norm for $p \geq 1$, as it is convex and may be readily solved by convex programming [78] irrespective of the number of samples. For larger samples, the MMyF cost function 'smooths' out and may be solved via unconstrained descent. The number of samples for which *sufficient* smoothing is attained depends on

α. For example, in the case of AWGN, the MMyF does not have any local traps as it is equivalent to the L_2-norm.

As $\mathbf{y}$ is of the form in (5.7), the statistics of $\mathbf{z}$ need to be known before the detection stage. The components of $\mathbf{z}$ are independent as $f_s = 4f_c$. From (5.48) and (5.49), it is not hard to convince ourselves that $z_I \stackrel{d}{=} z_Q$, therefore $\mathbf{z}$ has IID components. The list of near-optimal robust estimators certainly does not exhaust here [9], [63], [64]. We have discussed popular schemes and due to the lack of space, we cease further discussion on non-linear estimators.

5.2.5 Baseband Detection

Till now, we have discussed how $r(t)$ can be processed to get $\mathbf{y}$. The next step is to detect the transmitted symbol within $\mathbb{M}$ from $\mathbf{y}$. We comment on a few detectors and their performance in conjunction with linear and non-linear estimation of $\boldsymbol{\mu}$ in non-Gaussian AWSαSN with $f_s = 4f_c$.

Maximum-Likelihood Detection

In non-Gaussian AWSαSN, $\mathbf{z}$ is radically different for the conventional and discretized receivers. It is isotropic with dependent components in the first case and possesses a four-tailed symmetric PDF (similar to Fig. 4.4b) with IID components in the latter. The sampling process keeps the transmitted information intact, yet it statistically changes $\mathbf{z}$. We may rewrite the ML detector in (5.8) as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\mathtt{arg\,max}} f_Z(y_I - x_{I_l}) f_Z(y_Q - x_{Q_l}). \tag{5.61}$$

where $f_Z(\cdot)$ is the marginal PDF corresponding to $z_I \overset{d}{=} z_Q \overset{d}{=} Z \sim \mathcal{S}(\alpha, \delta_z)$. As $f_Z(\cdot)$ does not generally exist in closed form, numerical evaluations such as those in [42], [43] are employed to evaluate (5.61). The statistics of $\mathbf{z}$ need to be fully known to evaluate (5.61). In the non-linear case, the statistics of $\mathbf{z}$ depend on α, ξ and the estimator. As the estimators are based on good approximations of $f_W(\cdot)$, then from the discussion in Section-5.2.4 $\mathbf{z}$ should approximate a Gaussian vector with increasing ξ. Therefore $\mathbf{z}$ should be near-isotropic.

The Euclidean Detector

The detection rule for this is

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg}\,\texttt{min}} \|(\mathbf{y} - \mathbf{x}_l)\|^2 = \underset{\mathbf{x}_l}{\texttt{arg}\,\texttt{min}} \left(\mathcal{E}_{x_l} - 2\mathbf{y}^{\mathsf{T}}\mathbf{x}_l\right) \tag{5.62}$$

and is optimal in the ML sense for unimodal isotropic $\mathbf{z}$. This is true for the conventional receiver and the MMyF estimator as $\xi \to \infty$.

The Myriad Detector

We may invoke the myriad detector at the output of the linear estimator. The detection rule is

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg}\,\texttt{min}} \left(\log|K(\alpha)\delta_z^2 + (y_I - x_{I_l})^2| + \log|K(\alpha)\delta_z^2 + (y_Q - x_{Q_l})^2|\right) \tag{5.63}$$

We note that α and δ_z need to be estimated to invoke the myriad detector.

The L_p-Norm Detector

For the discretized linear receiver, the L_p-norm detector for $p < \alpha$ is defined as

$$\begin{aligned}\hat{\mathbf{x}} &= \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg min}} \|(\mathbf{y} - \mathbf{x}_l)\|_p^p \\ &= \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg min}} \left(|y_I - x_{I_l}|^p + |y_Q - x_{Q_l}|^p\right). \end{aligned} \tag{5.64}$$

As shown later, the optimal value of p depends on the SNR. At low SNR, the errors are determined by the background noise in the system (not by the impulses) for any α. This phenomenon is a characteristic of Gaussian noise and therefore p close to 2 performs well in this regime. In the medium-to-high SNR regime, the impulses predominantly determine the errors and thus the optimal p is close to zero.

The Log-Norm Detector

Like its estimator counterpart, the asymptotic detector is based on (3.10). On substituting (3.10) in place of $f_Z(\cdot)$ in (5.61), we get

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg min}} \left(|y_I - x_{I_l}||y_Q - x_{Q_l}|\right) \tag{5.65}$$

or equivalently

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg min}} \left(\log |y_I - x_{I_l}| + \log |y_Q - x_{Q_l}|\right). \tag{5.66}$$

We note that (5.66) is merely the logarithm of the cost function in (5.65). Either one may be used. The log-norm detector may be employed in the linear case as $\mathbf{z}$

is SαS. Like the L_p and myriad detectors, it also offers near-optimal performance. However, it has the added advantage of not requiring any knowledge about α and δ_z.

5.3 Joint-Detection

Till now we have focused on a mechanism consisting of passband-to-baseband conversion followed by detection in the complex plane. If the soft-values are not required, we may perform joint-detection of $\mathbf{x}_i$ directly from $r(n/f_s)$. Analogous to (5.47), the joint-detector is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}_l \in \mathbb{M}}{\texttt{arg min}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} - \log \rho \left(r\left(\frac{n}{f_s}\right) - x_{I_l}\ell_I\left(\frac{n}{f_s}\right) - x_{Q_l}\ell_Q\left(\frac{n}{f_s}\right)\right). \tag{5.67}$$

From the discussion in Section 5.2.4, if the passband noise is impulsive, one can use $\rho(x) = \log(K^2(\alpha) + x^2)$ (the myriad detector) for robust detection. Similarly, the L_p-norm for $0 < p < \alpha$ and the log-norm detectors can be used by substituting $-\log \rho(\cdot)$ by (5.58) and (3.10), respectively. Though cumbersome, one may also substitute $\rho(\cdot)$ with $f_W(\cdot)$ to implement ML joint-detection. In this case, the pdf will have to be numerically evaluated for each of its arguments in (5.67). Do note that (5.67) corresponds to a $\lfloor Tf_s \rfloor$-dimensional detection problem.

From an implementation perspective, the joint-detector is preferred as evaluating μ in (5.52) for non-Gaussian AWSαSN requires a numerical technique even if $\rho(\cdot)$ is in analytic form. The computational cost on solving (5.52) depends on the minimization algorithm. Also, the $f_s = 4f_c$ constraint, which is required to reduce (5.47) into two single variable minimizations, does not significantly reduce complexity in joint-detection and may therefore be discarded. More importantly, converting $r(n/f_s)$ to the vector form in (5.7), even if $\mathbf{y}$ is the ML estimate of $\boldsymbol{\mu}$, may result in loss of information (due to the local traps) and thereby is sub-optimal. Intuitively, the conversion corresponds to simplifying a $\lfloor Tf_s \rfloor$-dimensional problem to a 2-dimensional one and therefore is optimal only in specific scenarios such as minimizing the L_p-norm for $p \geq 1$ or the MMyF with large ξ. Even if their are no traps, there will be some loss at the detection stage as the statistics of $\mathbf{z}$ are not truly known and are assumed to be isotropic.

Though joint-detection is advantageous in both performance and implementation, it lacks the flexibility of integrating it with other schemes such as equalizers and soft-decoders as it does not output soft-values. One way to ensure compatibility with soft-decoders is to use the costs in (5.67) to generate approximates to the log-likelihood ratios (LLRs) involved. For example, in the case of binary modulation, the LLR measure would be given by

$$\log \frac{\prod_{n=0}^{\lfloor Tf_s \rfloor - 1} \rho\left(r\left(\frac{n}{f_s}\right) - x_{I_0}\ell_I\left(\frac{n}{f_s}\right) - x_{Q_0}\ell_Q\left(\frac{n}{f_s}\right)\right)}{\prod_{n=0}^{\lfloor Tf_s \rfloor - 1} \rho\left(r\left(\frac{n}{f_s}\right) - x_{I_1}\ell_I\left(\frac{n}{f_s}\right) - x_{Q_1}\ell_Q\left(\frac{n}{f_s}\right)\right)}.$$

The LLR is exact for $\rho(\cdot) = f_W(\cdot)$. Do note that though the LLRs can be approximated, we do not have soft-estimates of the transmitted symbol which may be necessary for processing in baseband. To evaluate the soft-estimates, the PDFs of z_I and z_Q are also required. Before we present a performance comparison of the discussed receiver mechanisms, we discuss the importance of constellation design in non-Gaussian AWSαSN.

5.4 Efficient Constellation Design

Signal constellations are conventionally designed for isotropic $\mathbf{z}$. This is reasonable as the passband noise process is typically modeled by AWGN and the receiver in Fig. 5.1 is employed. As per the discussion in Section 5.2.1, such constellations would be effective if the conventional receiver is employed in AWSαSN due to isotropic $\mathbf{z}$. This approach is also validated when the MMyF is employed in Fig. 5.3 with large ξ in AWSαSN. However, as highlighted previously, $\mathbf{z}$ is anisotropic with IID components if the discretized linear receiver is used in non-Gaussian AWSαSN. Similarly, if non-linear passband-to-baseband conversion is employed for small ξ, it is reasonable to assume that $\mathbf{z}$ still retains some impulsiveness. In such a case $f_{\mathbf{z}}(\cdot)$ will be anisotropic and of the form in Fig. 4.4b.

Statistically, the symmetry of the four-tailed PDF is given by

$$z_I + jz_Q \stackrel{d}{=} z_I - jz_Q \stackrel{d}{=} (z_I + jz_Q)e^{jk\pi/2} \tag{5.68}$$

$\forall\ k \in \mathbb{Z}$. If the constellation $x_i = x_{I_i} + jx_{Q_i}\ \forall\ i \in \{0, 1, \ldots, M-1\}$ has a certain error performance, then from (5.7) and (5.68), the symbol sets $x_i e^{jk\pi/2}$ and $x_i^* e^{jk\pi/2}\ \forall\ i \in \{0, 1, \ldots, M-1\}$ offer similar performance for any $k \in \mathbb{Z}$. Finding the constellation that offers globally optimal/near-optimal performance is a problem of interest, especially if the gains are large. Given (5.7), the symbol error probability (SEP) is evaluated by

$$\text{SEP} = \frac{1}{M} \sum_{i=0}^{M-1} \int_{\mathbf{y} \notin \mathbb{M}_i} f_{\mathbf{z}}(\mathbf{y} - \mathbf{x}_i) \mathrm{d}\mathbf{y} \tag{5.69}$$

where $\mathbb{M}_i \in \mathbb{R}^2$ is the set of points (determined by the detection rule) that lie in the decision region of $\mathbf{x}_i$. Optimizing the constellation corresponds to minimizing (5.69) with respect to $\mathbf{x}_i \ \forall \ i \in \{0, 1, \ldots, M-1\}$. As shown in Section 5.4.1, the ML decision regions for non-Gaussian SαS $f_{\mathbf{z}}(\cdot)$ with IID components are complex and cannot be expressed in closed form. Intuitively, given $\mathbf{x}_i$ is transmitted, one would want the tails of $f_{\mathbf{z}}(\mathbf{y} - \mathbf{x}_i)$ *directed away* from any other constellation point as there is a significant amount of probability in the tails. Due to the symmetry in (5.68), this ensures that the tails do not point towards each other, hence avoiding complete tail overlap and allowing an impulse/s to lie within the right decision region. We propose minimizing a simpler cost function:

$$\sum_{i=0}^{M-1} \sum_{\substack{l=0,\\ l \neq i}}^{M-1} f_{\mathbf{z}}(\mathbf{x}_l - \mathbf{x}_i) \tag{5.70}$$

or equivalently

$$\sum_{i=0}^{M-1} \sum_{\substack{l=0,\\ l \neq i}}^{M-1} \log f_{\mathbf{z}}(\mathbf{x}_l - \mathbf{x}_i). \tag{5.71}$$

Eq. (5.70) is merely (5.69) with $\mathbb{M}_i$ restricted to only $\mathbf{x}_i$. This ensures that the tails are diverted away from the constellation points. We validated this approach by Monte Carlo simulations and found the resulting constellations to work very well.

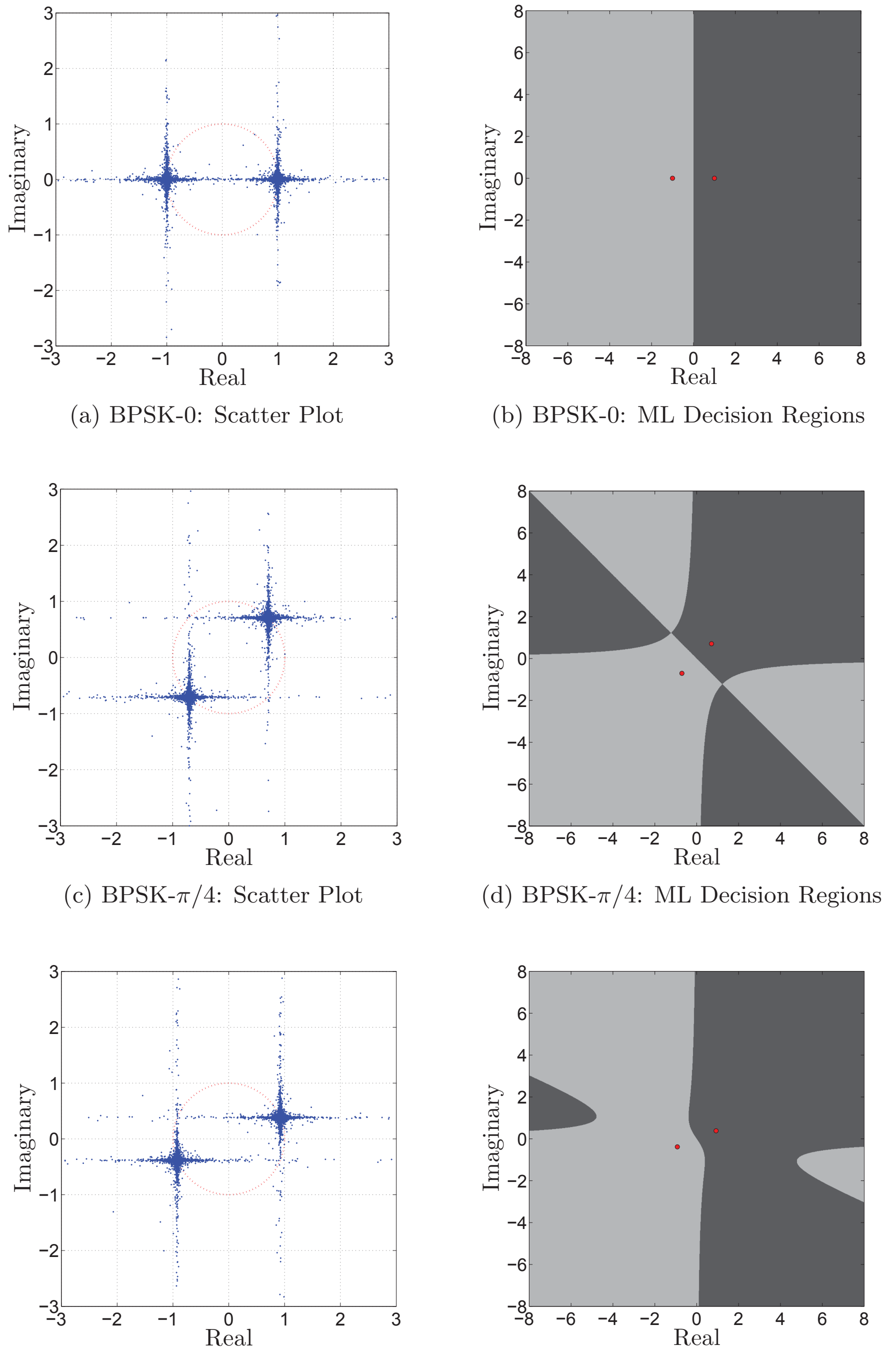
5.4.1 Rotated PSK Maps & Decision Regions

The optimal decision boundaries for isotropic baseband noise are evaluated from the Euclidean distance between signal points. This implies that for a given constellation map and any one of its *rotated version*, there is no advantage in terms of error rate between them as the optimum decision boundaries rotate accordingly. By a 'rotated version' we imply that all signal points in the constellation have been rotated by a similar angle. The same deductions do not hold for the non-Gaussian case in the discretized receiver as the bivariate distribution of $\mathbf{z}$ is anisotropic (four-tailed). If phase shift keying (PSK) is adopted as the modulation scheme, then from (5.71), the optimal angle of rotation is determined by

$$\phi_{\min} = \underset{\phi \in [0,\pi/4)}{\mathtt{arg\,min}} \underbrace{\sum_{i=0}^{M-1} \sum_{\substack{l=0, \\ l \neq i}}^{M-1} \log f_{\mathbf{z}}((\mathbf{x}_l - \mathbf{x}_i)e^{j\phi})}_{J(\phi)} . \tag{5.72}$$

where ϕ is the *rotation angle* or the angle (in radians) of the signal point in the first quadrant from the positive real axis. Do note that ϕ will lie in $[0, \pi/4)$. This is a direct consequence of the symmetry rule in (5.68).

In Figs. 5.4 & 5.5 we show scatter plots and the corresponding ML decision regions (via the rule in (5.61)) for various rotated BPSK and QPSK schemes, respectively, for $\alpha = 1$. We denote these schemes by BPSK-ϕ and QPSK-ϕ. Constellation points are signified by the red dots plotted on top of the decision



(a) BPSK-0: Scatter Plot

(b) BPSK-0: ML Decision Regions

(c) BPSK-$\pi/4$: Scatter Plot

(d) BPSK-$\pi/4$: ML Decision Regions

(e) BPSK-$\pi/8$: Scatter Plot

(f) BPSK-$\pi/8$: ML Decision Regions

Figure 5.4: Scatter Plots and Optimum decision regions for the Cauchy case ($\alpha = 1$) with independent baseband noise components for various rotated BPSK schemes.

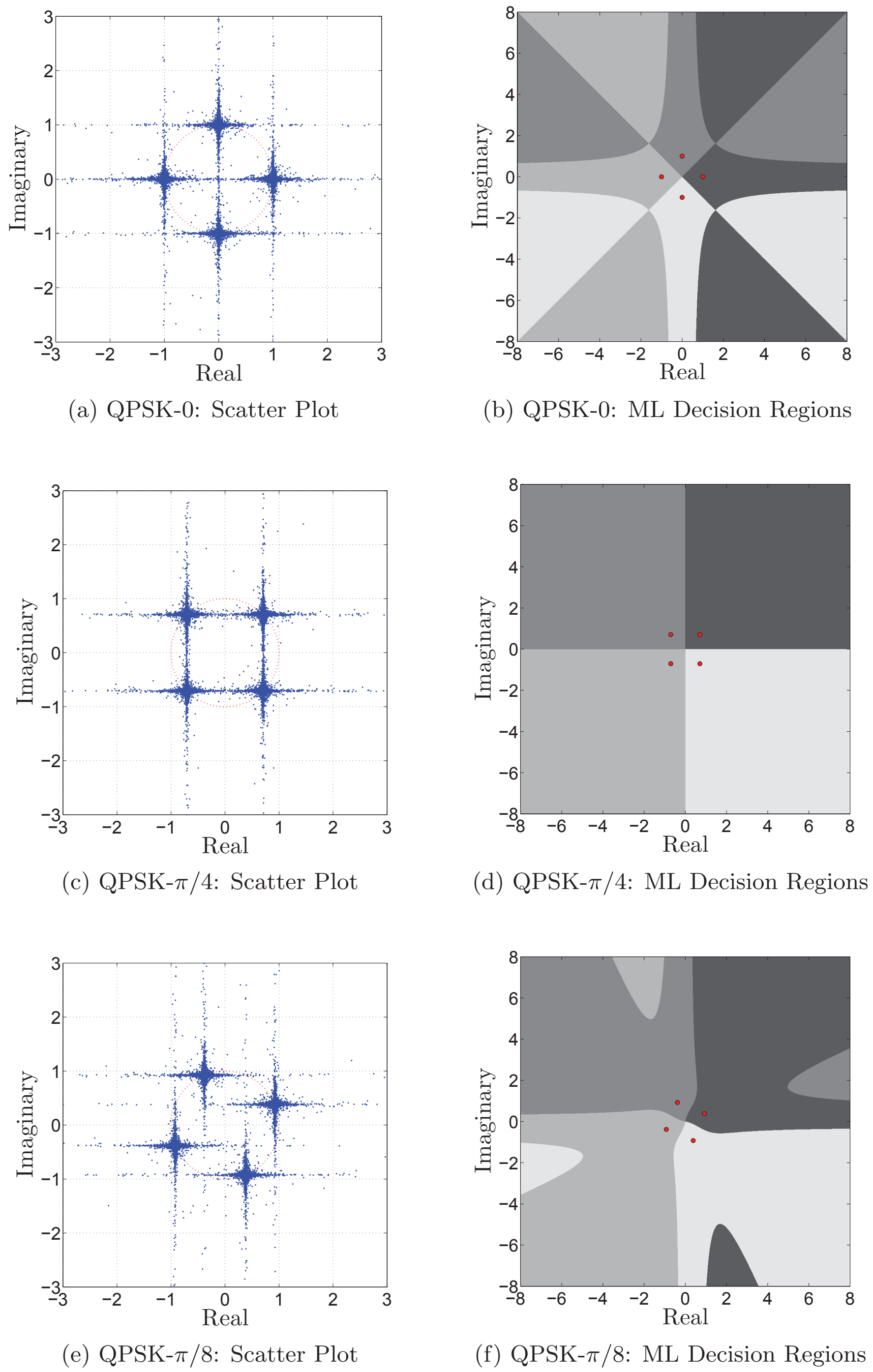


Figure 5.5: Scatter Plots and Optimum decision regions for the Cauchy case ($\alpha = 1$) with independent baseband noise components for various rotated QPSK schemes.

regions. The discretized linear receiver is employed which is why the received observations depict a four-tailed geometry. For the plots, we assume the signal points lie on the unit circle and $\delta_z = 1$. We note that the decision regions in Figs. 5.4b & 5.5d are the same as the isotropic case, but it will be seen later that they are not efficient in terms of error probability. The reason for this can be seen in the corresponding scatter plots in Figs. 5.4a & 5.5c, which clearly show tails directed towards other constellation points. Similarly, Fig. 5.5a also depicts tails being directed towards adjacent symbols. This makes BPSK-0, QPSK-0 and QPSK-$\pi/4$ undesirable. Do note that results from the Cauchy case may be intuitively extended to other non-Gaussian SαS scenarios due to the fact that $f_{\mathbf{z}}(\cdot)$ is a four-tailed PDF.

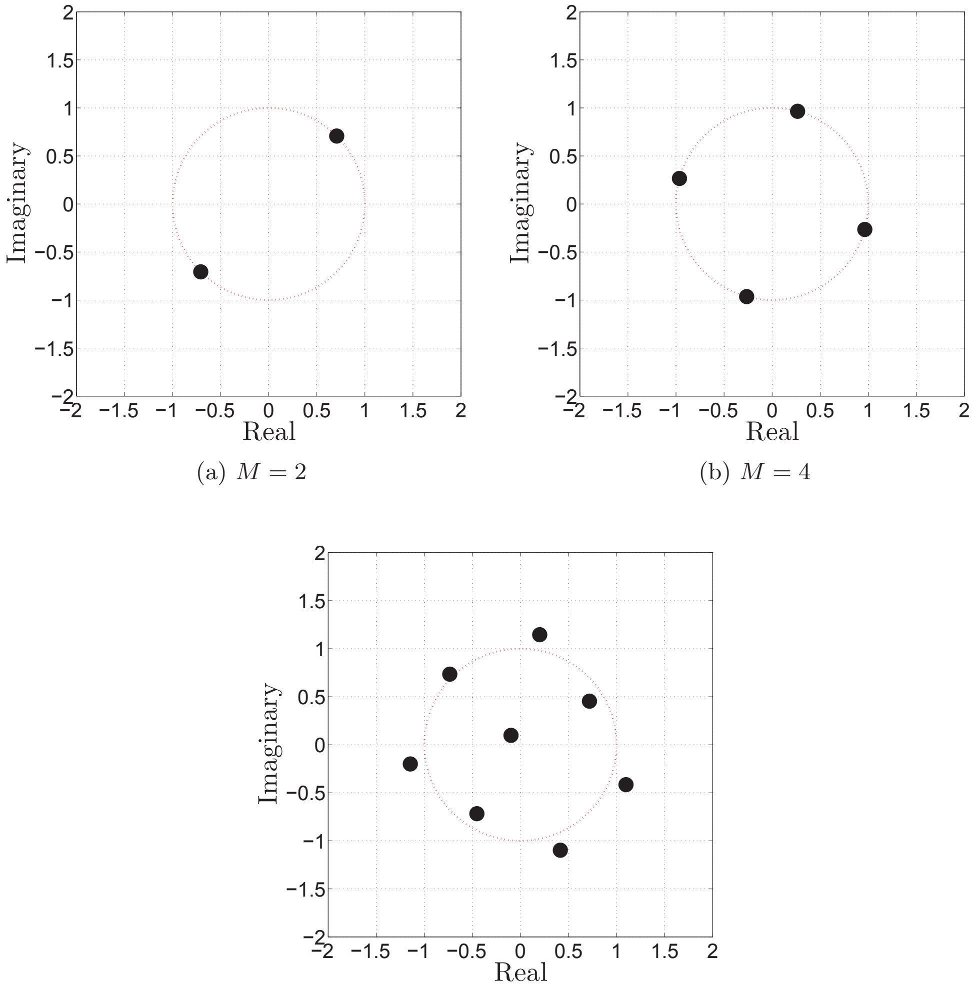
5.4.2 Globally Optimal QAM Constellations

For $\alpha = 2$, optimizing (5.69) corresponds to maximizing the *minimum* Euclidean distance between all points. Extending this concept to non-Gaussian $\mathbf{z}$, it is reasonable to maximize a *measure* within the points of the constellation. This can be observed by splitting $f_{\mathbf{z}}(\cdot)$ in (5.71) into a product of its IID marginals and substituting (3.10) in place of $f_Z(\cdot)$:

$$\sum_{i=0}^{M-1} \sum_{\substack{l=0,\\ l\neq i}}^{M-1} -\log|x_{I_l} - x_{I_i}| - \log|x_{Q_l} - x_{Q_i}|. \tag{5.73}$$

Similarly we may use (5.58) (with $\delta_{\bar{w}} = \delta_z$) for $f_Z(\cdot)$ to get

$$\sum_{i=0}^{M-1} \sum_{\substack{l=0,\\ l\neq i}}^{M-1} -\|\mathbf{x}_l - \mathbf{x}_i\|_p^p. \tag{5.74}$$



(a) $M = 2$

(b) $M = 4$

(c) $M = 8$

Figure 5.6: Optimal Constellations for various M for medium-to-high SNR.

Therefore, minimizing (5.71) can be interpreted as maximizing the combined log-norm or L_p-norm between the constellation points. As (5.73) is independent of α, the resultant constellation will be efficient for all non-Gaussian $\mathbf{z}$. For (5.74), we need to set a suitable p before the minimization takes place. Eq. (5.73) and (5.74) (for small p) get more accurate with increasing SNR. The minimization has to be performed over $2M$ variables; $\{x_{I_i}, x_{Q_i}\}$ $\forall$ $i \in \{1, 2, \ldots, M\}$. This may be accomplished via general search methods like

TABLE 5.1: OPTIMAL SYMBOL PLACEMENT.

$M=2$		$M=4$		$M=8$	
$\sqrt{\mathcal{E}_{x_i}}$	ϕ_i	$\sqrt{\mathcal{E}_{x_i}}$	ϕ_i	$\sqrt{\mathcal{E}_{x_i}}$	ϕ_i
1	45	1	-15.3679	1.1736	-20.7183
1	225	1	74.6321	1.1736	-69.2818
-	-	1	164.632	1.1627	80.0703
-	-	1	-105.368	1.1627	-170.0700
-	-	-	-	1.0392	135.0000
-	-	-	-	0.8491	-122.2640
-	-	-	-	0.8491	32.2645
-	-	-	-	0.1386	135.0000

Differential Evolution [87] or Simulated Annealing [88],[89]. Though the optimal constellation generally varies with SNR for $\alpha \neq 2$, it is almost constant in the medium-to-high regime where the errors are predominantly determined by the tails of $f_{\mathbf{z}}(\cdot)$.

In Fig. 5.6 we present constellations for various M that offer the best error performance for Cauchy $\mathbf{z}$ with IID components for $\delta_z = 0.001$ and $E[\mathcal{E}_{x_i}] \leq 1$ (this corresponds to an SNR of 30 dB). The unit circle is also plotted for comparison. In complex form, the i^{th} point is $\sqrt{\mathcal{E}_{x_i}} e^{j\phi_i}$. In Table 5.1, we have listed down $\sqrt{\mathcal{E}_{x_i}}$ and ϕ_i (in degrees) $\forall\ i \in \{0, 1, \ldots, M-1\}$ for each of the constellations in Fig. 5.6. There are no noticeable changes in the optimal constellation as the SNR decreases to as low as 10 dB. Further still, the results can be extended to any $\alpha \neq 2$ as the tail directions are similar [79]. We show this by presenting numerical results later. On a final note, it is worth mentioning that the globally optimal constellation for the $M = 2$ and $M = 4$ case is that of a rotated BPSK and QPSK, respectively.

5.5 Error Performance: Linear Receivers

Before we present the error performance of various receiver schemes, it is important to identify a suitable SNR measure. This is done next.

5.5.1 SNR Measures

For digital communication systems in Gaussian noise scenarios, the bit error rate (BER) or symbol error rate (SER) curves are conventionally plotted against the SNR per information bit ($\mathcal{E}_b/N_0$), where $\mathcal{E}_b$ is the energy of an information bit and $N_0/2$ is the two-sided PSD of passband AWGN [1]. Recall, the concept of PSD does not extend to other stable random variables as their respective second moments are infinite. Consequently, one needs a suitable equivalent measure for the non-Gaussian stable case. To our knowledge, there are two different measures that have been proposed in the literature. In essence, there is no difference between the both of them except for an α-dependent scaling parameter. We highlight these measures for the discretized linear receiver. It is worth recalling from the discussion in Section 5.2.3 that $z_I \stackrel{d}{=} z_Q \stackrel{d}{=} Z \sim \mathcal{S}(\alpha, \delta_z)$.

For the first approach we represent $\mathcal{E}_b/N_0$ for the Gaussian case in terms of the scale parameter δ_z instead of N_0 as the concept of second-order does not extend to stable distributions. We may then directly use this form for other stable random variables as δ_z exists for each of them. This measure has been used in [3], [10]. The conversion is done as follows:

$$\frac{\mathcal{E}_b}{N_0} = \frac{\mathcal{E}_b}{4\delta_z^2} \tag{5.75}$$

As $\mathcal{E}_b = E[\mathcal{E}_{x_i}]/\log_2 M$ where $\log_2 M$ is the number of information bits per message symbol, we get

$$\frac{\mathcal{E}_b}{N_0} = \frac{E[\mathcal{E}_{x_i}]}{4\delta_z^2 \log_2 M} \tag{5.76}$$

The second is based on the *geometric signal-to-noise ratio* (GSNR) approach. This was first proposed in [58] and has been used in [66], [90]–[93]. To explain what the GSNR is, we have to define the *geometric power* of an SαS random variable. The geometric power of an SαS random variable X is defined as:

$$S_0 = e^{E[\log|X|]} = \delta C_g^{\frac{1}{\alpha}-1} \tag{5.77}$$

where δ is the scale parameter of X and C_g is the exponential of Euler's constant and is approximately 1.7811. It has been proven that $E[\log|X|]$ exists for stable X [9]. Given (5.7), the GSNR is defined as

$$\text{GSNR} = \frac{1}{2C_g}\left(\frac{E[\mathcal{E}_{x_i}]}{S_0}\right)^2 = \frac{1}{2C_g}\left(\frac{E[\mathcal{E}_{x_i}]}{\delta_z C_g^{\frac{1}{\alpha}-1}}\right)^2 = \frac{E[\mathcal{E}_{x_i}]^2}{2\delta_z^2 C_g^{\frac{2}{\alpha}-1}} \tag{5.78}$$

The GSNR is designed such that for $\alpha = 2$ (the Gaussian case),

$$\text{GSNR} = \frac{E[\mathcal{E}_{x_i}]^2}{2\delta_z^2} = \frac{E[\mathcal{E}_{x_i}]^2}{N_0}. \tag{5.79}$$

From (5.75), this implies that

$$\frac{\mathcal{E}_b}{N_0} = \frac{\text{GSNR}}{\log_2 M} \tag{5.80}$$

for the Gaussian case. Extending the definition in (5.80) to $\alpha \neq 2$, we have

$$\frac{\mathcal{E}_b}{N_0} = \frac{E[\mathcal{E}_{x_i}]^2}{\delta_z^2 C_g^{\frac{2}{\alpha}-1} \log_2 M}. \tag{5.81}$$

On comparing (5.81) and (5.76), we see that they only differ by the scale factor $1/C_g^{\frac{2}{\alpha}-1}$. For a given α the scale factor is constant. No expression holds any advantage over the other; in fact, the two different derivations arrive at a somewhat similar result, showing the consistency of both measures. Some authors prefer using $\gamma_z^{1/\alpha}$ instead of δ_z as the way it was initially proposed in [9], where γ_z is the *dispersion parameter* of z. We choose the measure in (5.76) for its simpler mathematical form.

5.5.2 Simulations

We start off by presenting results for rotated PSK schemes and graphically depict the accuracy of (5.72) as a rule for constellation design. Fig. 5.7 presents the BER/SER curves for the Cauchy AWSαSN under the discretized linear receiver for rotated versions of QPSK using the ML decision regions highlighted in Fig. 5.5 and Gray coding. The results were generated for a minimum of 4000 errors for high BER/SER ($> 10^{-3}$) and 1000 errors for low BER/SER. It is observed that when the distribution tails are directed away from the signal points, the BER/SER falls drastically. The QPSK-$\pi/8$ scheme has therefore better BER/SER performance than its QPSK-0 and QPSK-$\pi/4$ counterparts. An interesting observation is that of the SER and BER for QPSK-0. We see that the SER and BER are almost equal as the tails are pointed exactly towards the opposing neighbors for each signal point. The gain between the worst and

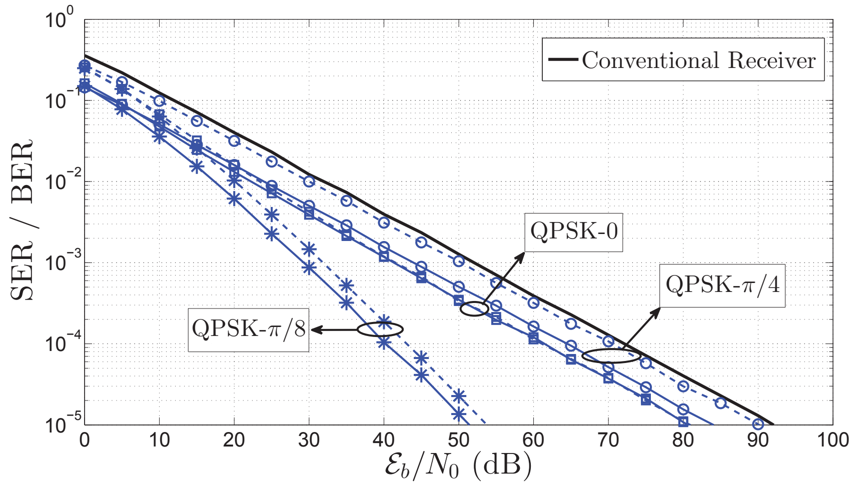


Figure 5.7: BER/SER curves for various QPSK schemes are presented for the Cauchy case in the discretized linear receiver under ML detection. The dashed lines represent the SER. ML decision boundaries were used for decoding.

optimum cases for is over 35 dB at an error rate of 10^{-5}. For comparison, we also plot the SER performance of QPSK in the conventional receiver in Fig. 5.7. An analytical expression for $\mathcal{E}_b/N_0$ of the conventional receiver is given in Appendix A.2. One can clearly see the advantage in performance gain of the discretized receiver over the conventional receiver.

Estimates for α within the AWSαSN framework for practical underwater ambient noise have been recorded to be as low as $\alpha = 1.5$ [4]. We also present SER plots in Fig. 5.8 for independent components with $\alpha = 1.5$ in the discretized and conventional receivers. It is observed that the trends encountered in the Cauchy case extend to this case as well due to their common heavy-tailed property. Similarly, we may extend results to other values of α as well.

Fig. 5.9 depicts the variation of the uncoded SER for the Cauchy case with independent components against the rotation angle assuming QPSK for

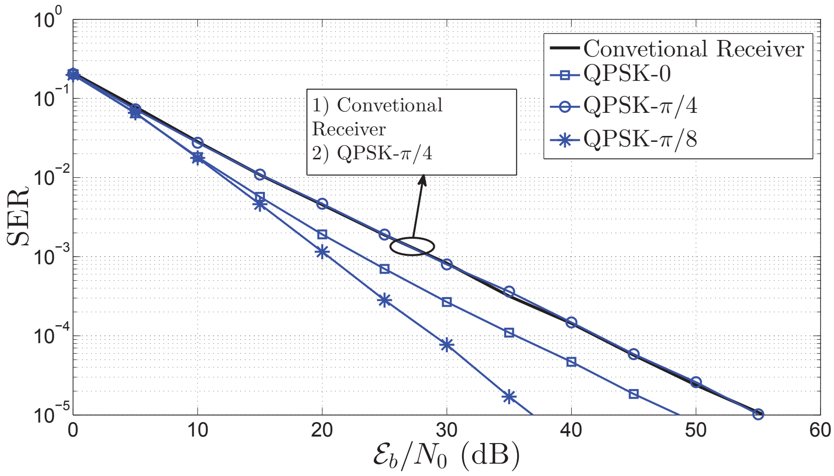


Figure 5.8: SER curves for various QPSK schemes are presented for $\alpha = 1.5$ in the discretized linear receiver. ML decision boundaries were used for decoding.

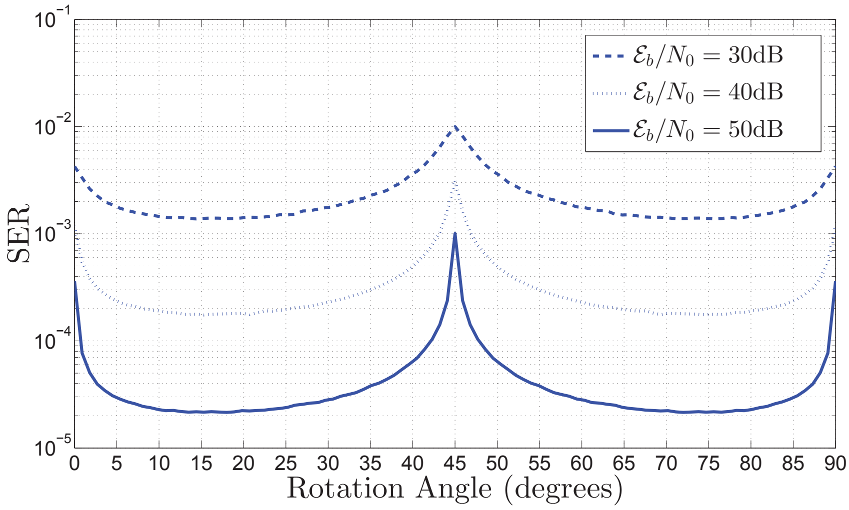


Figure 5.9: The SER for the Cauchy case with independent components plotted against the rotation angle for three different values of $\mathcal{E}_b/N_0$ (dB).

three different values of $\mathcal{E}_b/N_0$. Each curve was evaluated using Monte Carlo simulations for a minimum of 3000 errors for selected rotation angle. One can observe that there is an optimum angle (albeit not unique) for uncoded QPSK transmission that ensures minimum error probability. The plot clearly

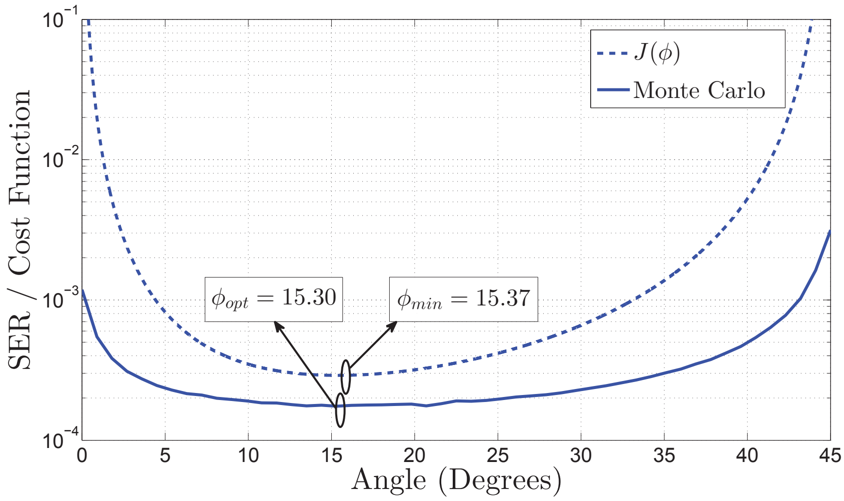


Figure 5.10: SER / $J(\phi)$ vs. ϕ for the Cauchy case with independent noise components for $\mathcal{E}_b/N_0 = 40$dB.

shows that the SER is a periodic function of ϕ with period $\pi/2$ radians and is symmetric about the vertical line $\phi = \pi/4$. This is a direct consequence of the symmetry of $f_{\mathbf{z}}(\cdot)$ in (5.68). From the discussion in Section 5.4, one would want to direct the tails towards the gaps between the constellation points. Intuitively, for high $\mathcal{E}_b/N_0$, one would expect the tails to align a little towards the opposite constellation point as it is further away in comparison to the adjacent points, i.e., the optimal ϕ (or ϕ_{opt}) is less than $\pi/8$. On the other hand δ_z is high at low $\mathcal{E}_b/N_0$. This results in thickening of the tails. The small relative distance between the adjacent and opposite points becomes inconsequential allowing the tails to bisect the spaces between the points equally, i.e., the QPSK-$\pi/8$ (22.5 degrees of rotation) case would be optimal. As the SNR increases, it was numerically determined that ϕ_{opt} converges to 15.3 degrees.

To validate the proximity of the expression in (5.72), we show how the cost function in (5.72) varies with ϕ for QPSK in Fig. 5.10. Also highlighted is

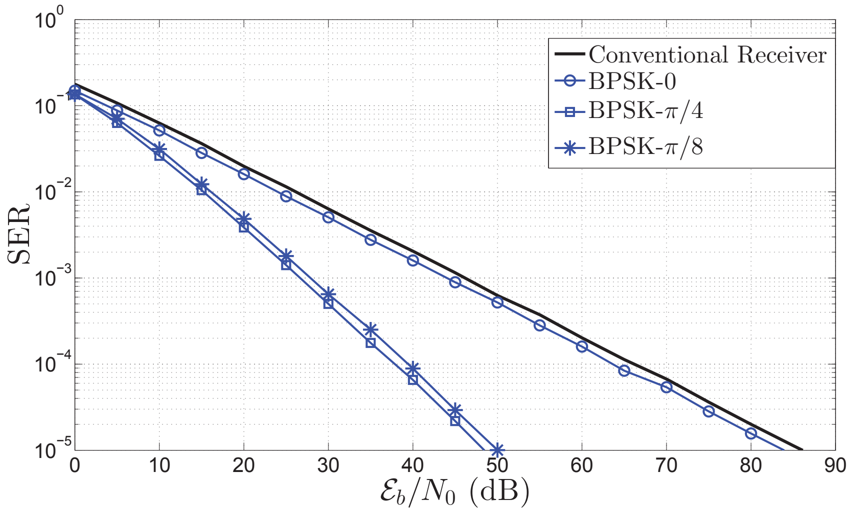


Figure 5.11: SER vs. $\mathcal{E}_b/N_0$ (dB) for the Cauchy case for various rotated BPSK schemes.

TABLE 5.2: PERFORMANCE GAIN (dB) OVER THE CONVENTIONAL RECEIVER FOR VARIOUS ROTATED PSK MAPS IN THE LINEAR DISCRETIZED RECEIVER. THE RESULTS ARE COMPILED FOR ML DETECTION AT SER= 10^{-5}.

α	QPSK-0	QPSK-$\pi/4$	QPSK-$\pi/8$	BPSK-0	BPSK-$\pi/4$	BPSK-$\pi/8$
1	2.0	11.0	38.0	2.0	36.0	37.5
1.5	0.5	7.0	18.0	-	-	-

$\phi_{\min}$. The solid line depicts the variation of the SER against ϕ. Both curves were generated for $\mathcal{E}_b/N_0 = 40$ dB. This value is of practical interest as the corresponding SER is approximately 10^{-4} for ϕ equal or close to ϕ_{opt} (see Fig. 5.7). We see that ϕ_{opt} is closely approached by the approximation $\phi_{\min}$ in (5.72). It is observed that the SER is approximately the same for a certain range of ϕ, i.e., between 10 to 30 degrees approximately. Any ϕ chosen from this interval gives good results.

To wrap up the discussion on rotated-PSK maps, we plot in Fig. 5.11 the BER performance of various rotated BPSK schemes in the conventional and

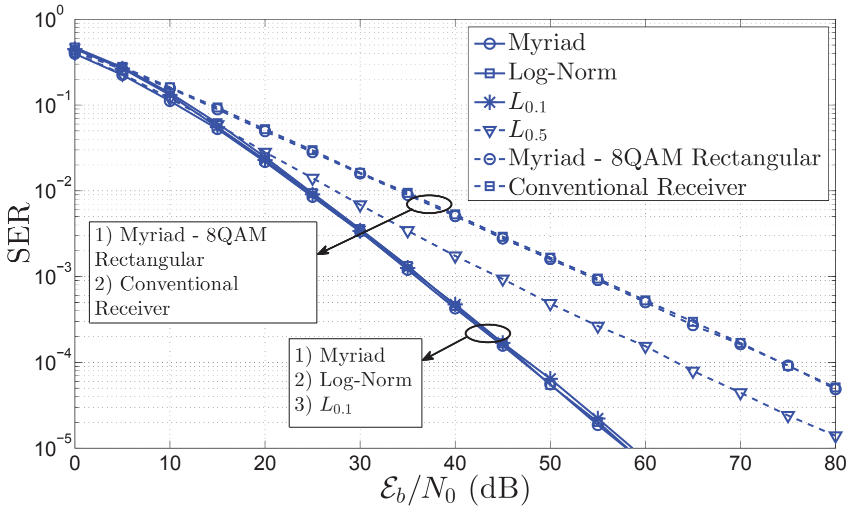


Figure 5.12: SER for various detectors in the discretized linear receiver in Cauchy AWSαSN for $M = 8$.

discretized linear receivers for the Cauchy case with ML detection. The trends encountered in QPSK can be clearly seen here as well. We summarize the performance gain at SER=10^{-5} of all schemes in Figs. 5.7, 5.8 and 5.11 over the corresponding conventional receiver in Table 5.2. The results are rounded to the nearest 0.5 dB.

We plot the SER for various *detectors* in the discretized linear receiver for $\alpha = 1$ and $\alpha = 1.5$ in Fig. 5.12 and Fig. 5.13, respectively. The 8-QAM constellation in Fig. 5.6 is employed. For comparison we have also plotted the SER for the conventional receiver with the same constellation. Similarly, for $M = 4$, we have presented results for the globally optimal constellation for the Cauchy and $\alpha = 1.5$ case in Fig. 5.14 and Fig. 5.15, respectively. The increase in performance due to sampling at $f_s = 4f_c$ and invoking good constellations and detectors over the conventional receiver is clear. The myriad, log-norm and L_p-norm (as $p \to 0$) detectors perform very well over a large range of $\mathcal{E}_b/N_0$. To emphasize

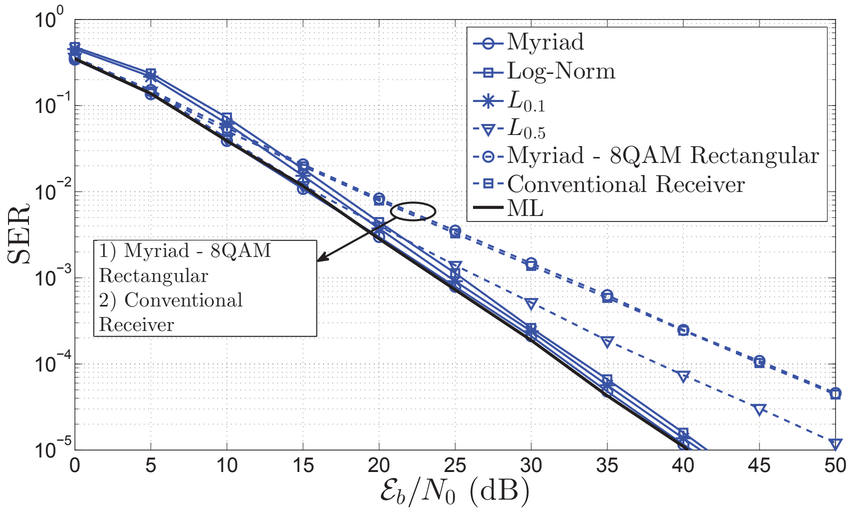


Figure 5.13: SER for various detectors in the discretized linear receiver in AWSαSN with $\alpha = 1.5$ and $M = 8$.

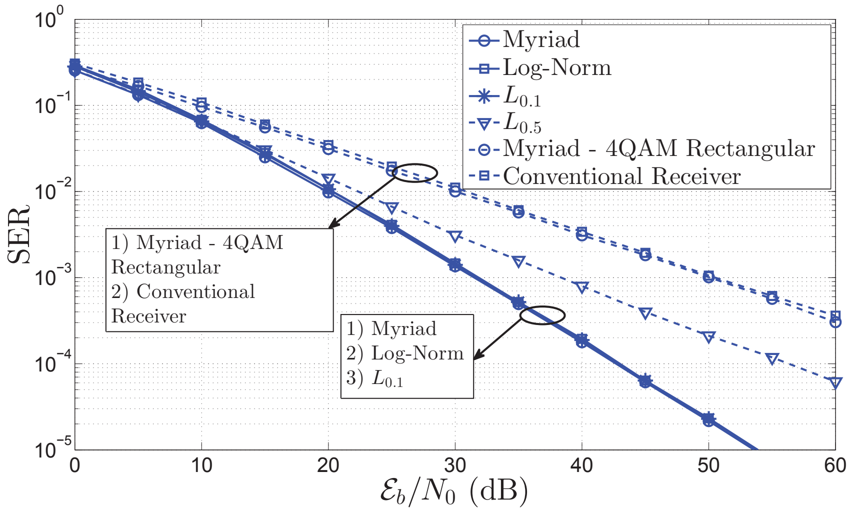


Figure 5.14: SER for various detectors in the discretized linear receiver in Cauchy AWSαSN for $M = 4$.

the importance of constellation design, we have also plotted results of the myriad detector for the well-known 8-QAM and 4-QAM rectangular maps for both the Cauchy and $\alpha = 1.5$ cases. For Cauchy noise, myriad detection corresponds to ML detection. We have additionally presented the ML detector performance

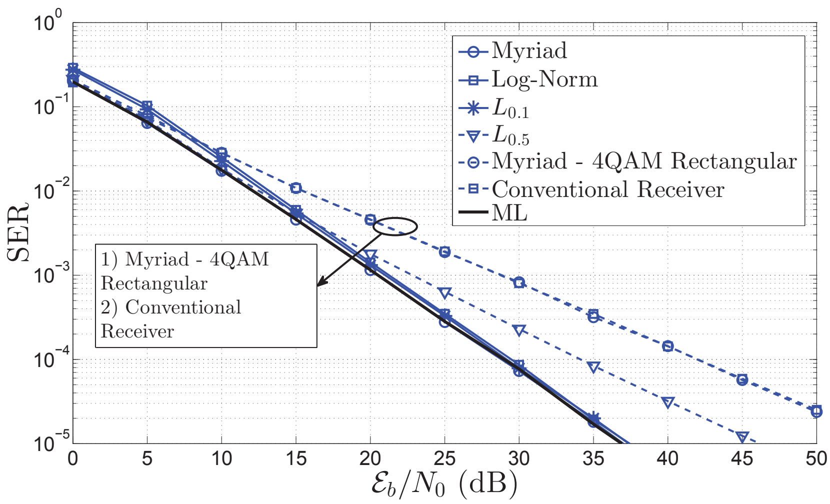


Figure 5.15: SER for various receiver schemes in AWSαSN with $\alpha = 1.5$ and $M = 4$.

TABLE 5.3: PERFORMANCE GAIN (dB) OVER THE CONVENTIONAL RECEIVER AT SER $= 10^{-5}$ FOR VARIOUS DETECTORS IN THE LINEAR DISCRETIZED RECEIVER WITH OPTIMAL CONSTELLATIONS.

α	M	Myriad	$L_{0.1}$	$L_{0.5}$	Log-Norm
1	4	37.5	37.0	16.0	37.0
	8	36.5	36.0	11.5	36.0
1.5	4	18.5	18.0	9.5	18.0
	8	18.0	17.5	8.5	17.0

for the $\alpha = 1.5$ case in Figs. 5.13 & 5.15. As $\alpha \to 2$, the SER performance between the discretized linear receiver with ML detection and its conventional counterpart converge.

As there are many results, we summarize the performance gain at SER$=10^{-5}$ of various detectors in the discretized linear receiver over the conventional receiver in Table 5.3. Optimal constellations are employed and ML detection is used for the conventional receiver. Results are rounded off to the nearest 0.5 dB.

5.6 Error Performance: Joint-Detection & Non-Linear Receivers

5.6.1 SNR Measures

For fair comparison, the performance of each receiver needs to be analyzed for the passband AWSαSN process that amounts to (5.76) in the discretized linear receiver. Therefore, we need to evaluate $\mathcal{E}_b/N_0$ as a function of δ_w (the scale parameter of the passband noise samples). From (5.43),

$$\delta_z = \frac{d(\alpha, \xi, g[n])}{\sqrt{f_s}} \delta_w \tag{5.82}$$

where

$$d(\alpha, \xi, g[n]) = \frac{\left(\sum_{n=0}^{2\xi-1} |g[2n]|^{\alpha}\right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g[2n]^2\right)^{1/2}} \tag{5.83}$$

We note that (5.83) is a ratio of the L_α and L_2 norms of $g[2n] \; \forall \, n \in \{0, 1, \ldots, 2\xi - 1\}$ and is solely a function of α, ξ and the *sampled* baseband shaping pulse. Finally, on substituting (5.82) in (5.76) we get

$$\frac{\mathcal{E}_b}{N_0} = \underbrace{\frac{E[\mathcal{E}_{x_i}] f_s}{4\delta_w^2 \log_2 M}}_{\text{SNR}} \times \frac{1}{d(\alpha, \xi, g[n])} \tag{5.84}$$

or in dB

$$10 \log_{10} \frac{\mathcal{E}_b}{N_0} = \text{SNR (dB)} - 20 \log_{10} d(\alpha, \xi, g[n]). \tag{5.85}$$

Do note that the SNR term in (5.84) and (5.85) is *not* the actual SNR

$\left(\frac{\text{signal power}}{\text{noise power}}\right)$ at the receiver. It is in fact a measure of $\mathcal{E}_b/N_0$. This is due to the fact that $d(\alpha, \xi, g[n]) = 1$ for all $g(t)$ and $\xi \in \mathbb{Z}^+$ when $\alpha = 2$. Therefore, $\mathcal{E}_b/N_0 =$ SNR in the Gaussian case. The term in (5.84) is denoted as SNR to differentiate between the definition of $\mathcal{E}_b/N_0$ in (5.76).

As $\|\mathbf{x}\|_p \geq \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^{2\xi}$ and $0 < p \leq 2$, we see that $d(\alpha, \xi, g[n]) \geq 1$. In the Gaussian case, $d(2, \xi, g[n]) = 1$ for all $g(t)$ and $\xi \in \mathbb{Z}^+$. Analogous to the Gaussian case, $4\delta_w^2/f_s$ may be interpreted as the *pseudo*-PSD of the AWSαSN process. For a given pseudo-PSD, we see that $\mathcal{E}_b/N_0$ varies with $d(\alpha, \xi, g[n])$ for $\alpha \neq 2$ in the discretized linear receiver. In fact, increasing $d(\alpha, \xi, g[n])$ for $\alpha \neq 2$ *decreases* the operational SNR. In the special case of $g(t)$ being a rectangular pulse, (5.85) reduces to

$$10 \log_{10} \frac{\mathcal{E}_b}{N_0} = \text{SNR (dB)} - 10 \left(\frac{2}{\alpha} - 1\right) \log_{10} \xi \tag{5.86}$$

We note that (5.86) decreases linearly with $10 \log_{10} \xi$ at a rate proportional to $\frac{2}{\alpha} - 1$. In essence, one can arbitrarily *reduce* the SNR by *increasing* ξ. For α close to 2, increasing ξ causes no significant effect in SNR. However, as α decreases, the reduction in SNR becomes apparent. As the non-linear performance can be made arbitrarily better than that of the linear receiver by increasing ξ (and hence decreasing (5.86)), we plot the SER against

$$\text{SNR (dB)} = 10 \log_{10} \frac{E[\mathcal{E}_{x_i}] f_s}{4\delta_w^2 \log_2 M}. \tag{5.87}$$

To compare these results with those of the conventional/discretized linear receivers, one needs only to shift the latter's results by $20\log_{10} d(\alpha, \xi, g[n])$ to the right.

Remark: The previous discussion raises the question of finding the optimum $g[n]$ in the discretized linear receiver that reduces the SNR. From (5.85), we note that the latter term is equivalent to zero when $g[n]$ is a single impulse. In the spectral domain, this implies that the baseband signal spans the bandwidth f_s. Therefore, the SNR of the system is maximized for a given pseudo-PSD when the receiver bandwidth is *equivalent* to the signal bandwidth.

5.6.2 Simulations

We present the *joint-detector* performance for different non-linear receivers in Cauchy AWSαSN for $Tf_s = 40$ in Fig. 5.16 and $Tf_s = 400$ in Fig. 5.17. All plots are generated for $g(t)$ a rectangular pulse and the 8-QAM constellation in Fig. 5.6. For comparison, we have also plotted the Gaussian error curve for the same constellation in both figures.

We observe that for $Tf_s = 400$, the myriad joint-detector performance (which corresponds to ML detection in the Cauchy case) actually *converges* to the Gaussian error curve. We tested this empirically for even larger values of Tf_s. Increasing Tf_s expands the *available bandwidth* f_s relative to the symbol rate. To show the effect of constellation design, we plot the joint-myriad detection performance for the 8-QAM rectangular map for $Tf_s = 40$ in Fig. 5.16. We can clearly see the degradation in error performance. Similar effects have also been empirically observed for other joint-detectors for $Tf_s = 40$. In the $Tf_s = 400$ case, there is no significant gain for the optimized 8-QAM constellation over its

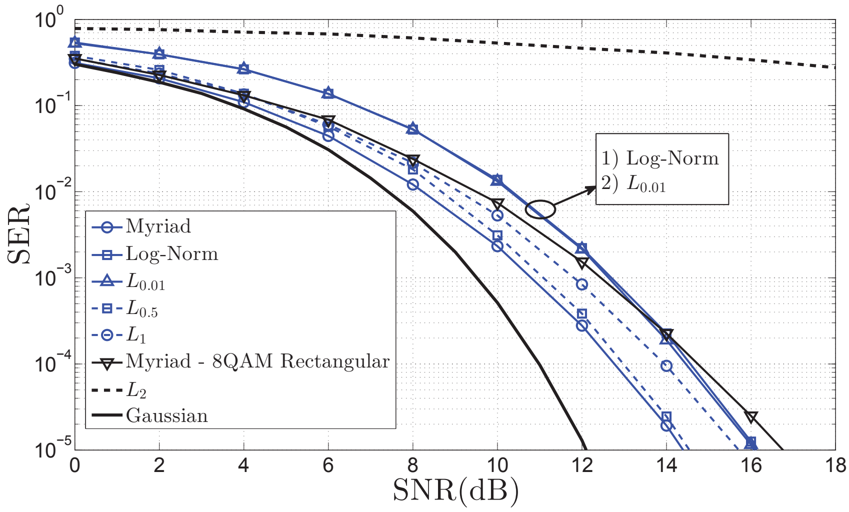


Figure 5.16: SER for various joint-detection schemes in Cauchy AWSαSN with $Tf_s = 40$ and $M = 8$.

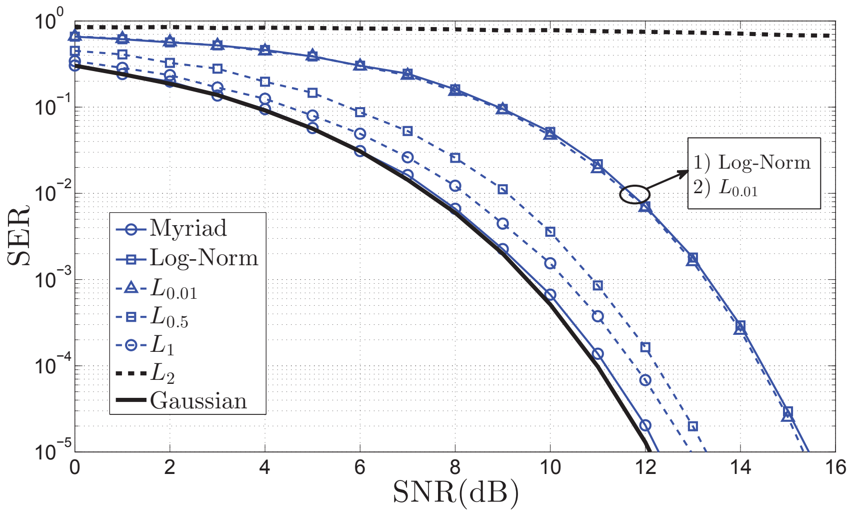


Figure 5.17: SER for various joint-detection schemes in Cauchy AWSαSN with $Tf_s = 400$ and $M = 8$.

rectangular counterpart.

In Fig. 5.18 we present SER results for different non-linear passband-to-baseband conversion schemes with isotropic baseband detection. We show results for $Tf_s = 40$ ($\xi = 10$) and $Tf_s = 400$ ($\xi = 100$) with the added

Table 5.4: Performance Loss (dB) with respect to the Gaussian error curve at SER $= 10^{-5}$ for the Joint-Detection schemes in Figs. 5.16 and 5.17.

Tf_s	Myriad	$L_{0.01}$	$L_{0.5}$	L_1	Log-Norm
40	2.3	4.0	2.4	3.6	4.0
400	0.2	3.2	1.2	0.9	3.3

Table 5.5: Performance Loss (dB) with respect to the Gaussian error curve at SER $= 10^{-5}$ for the Non-Linear schemes in Fig. 5.18.

Tf_s	Myriad	L_1
40	4.8	4.6
400	0.2	0.8

constraint of $f_s = 4f_c$. These results can be compared to their joint-detector counterparts in Fig. 5.16 and Fig. 5.17. The isotropic assumption of $\mathbf{z}$ causes slight performance degradation for the L_1-norm based passband-to-baseband conversion for both $Tf_s = 40$ and $Tf_s = 400$. However, for the MMyF based conversion, there is no error for the $Tf_s = 400$ case. We see a deviation for the MMyF curve for $Tf_s = 40$ after 13 dB. This is because the MMyF cost function is not smooth enough and the local traps hinder the soft-value estimation process.

In Table 5.4, we present the performance loss (in decibels) of all joint-detection schemes in Figs. 5.16 and 5.17 with respect to the Gaussian error curve at SER=10^{-5}. Results are rounded of to the nearest 0.1 dB. Similarly, we evaluated the performance loss for the non-linear schemes in Fig. 5.18 and presented them in Table 5.5.

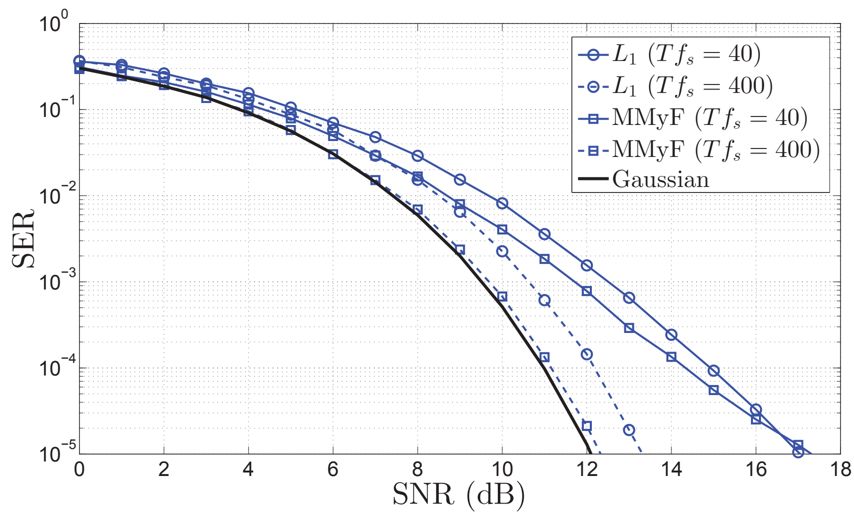


Figure 5.18: SER for various receiver schemes in Cauchy AWSαSN with $M = 8$.

5.7 A Practical Implementation: Rotated PSK Schemes

From a practical perspective, the rotation of a constellation map can be accomplished at the receiver *without* actually transmitting the rotated constellation symbols themselves. This of course is only of interest if the baseband noise components are anisotropic. We propose a simple mechanism that not only incorporates constellation rotation at the receiver, but also generates baseband noise with independent components assuming passband AWSαSN.

Let us assume that a single-carrier scheme is to be implemented over an impulsive noise channel and the transmitted symbols are chosen from the QPSK-ϕ_1 configuration. Also, let the optimal constellation map for this particular realization of the channel be QPSK-ϕ_{opt}. Each symbol in QPSK-ϕ_1 can be mapped on to a unique point in QPSK-ϕ_{opt} by multiplying it with $\exp(j\Delta\phi)$ where $\Delta\phi = \phi_{\text{opt}} - \phi_1$. This mapping corresponds to a rotation of the constellation points in QPSK-ϕ_1 to attain QPSK-ϕ_{opt}.

The relationship between the transmitted passband signal and its baseband counterpart is given in (5.2). Let $\tilde{s}_i(t) = \sqrt{\frac{2\mathcal{E}_{x_i}}{\mathcal{E}_g}} g(t)e^{j\phi_i}$, then from (5.2) we have

$$s(t) = \Re\{\tilde{s}_i(t)\exp(j2\pi f_c t)\} \tag{5.88}$$

Note that (5.88) is the continuous-time version of (4.2). As $\tilde{s}_i(t)$ is the baseband signal corresponding to (a sequence of) symbols in QPSK-ϕ_1, $\tilde{s}_i(t)\exp(j\Delta\phi)$ will be the baseband signal if the symbols are chosen from QPSK-ϕ_{opt}. We can rewrite (5.88) as

$$s(t) = \Re\big\{\tilde{s}_i(t)\exp(j\Delta\phi)\underbrace{\exp\big(j(2\pi f_c t - \Delta\phi)\big)}_{\text{carrier}}\big\} \tag{5.89}$$

On comparing (5.88) with (5.89), we observe that given $s(t)$, $\tilde{s}_i(t)\exp(j\Delta\phi)$ can be acquired if the carrier (or clock) at the receiver lags that of the transmitter by $\Delta\phi$. For example, if QPSK-$\pi/4$ is the transmitted constellation and $\phi_{\text{opt}} = \pi/8$, the QPSK-$\pi/8$ constellation map can be generated by letting the receiver clock lag the transmitter clock by $\Delta\phi = -\pi/8$.

Though we have concocted a mechanism that rotates the constellation at the receiver, independence of baseband noise components is only ensured if the received signal $s(t)$ is sampled at

$$t = \frac{n}{f_s} + \frac{\Delta\phi}{2\pi f_c} \tag{5.90}$$

where $f_s = 4f_c$ and n is the discrete-time index. The sampling rule in (5.90)

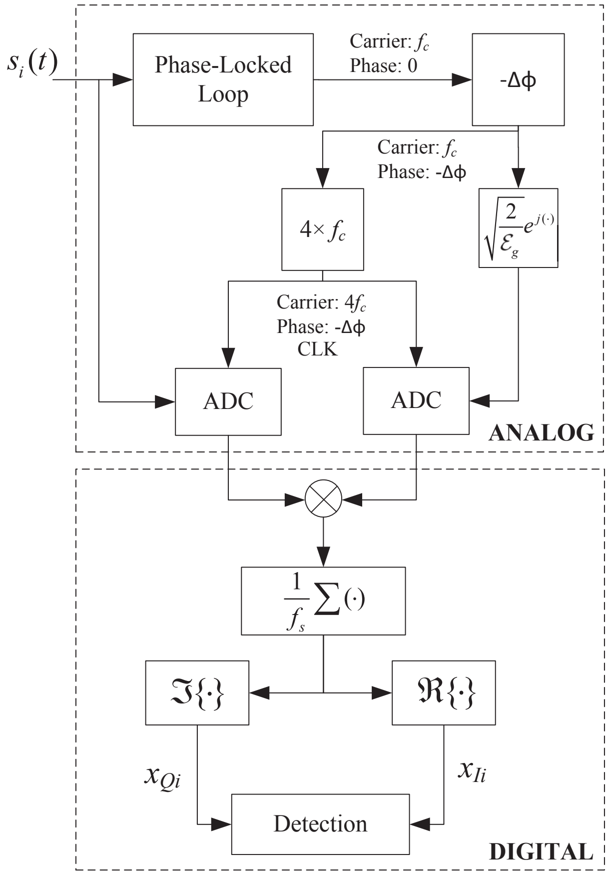


Figure 5.19: Practical implementation of a single-carrier receiver employing rotated-constellations.

does not effect the constellation rotation at the receiver. On substituting (5.90) in (5.89) we get

$$\acute{s}[n] = \Re\left\{\tilde{\tilde{s}}_i[n]\exp\left(j2\pi\frac{f_c}{f_s}n\right)\right\} \tag{5.91}$$

where

$$\begin{aligned}\acute{s}[n] &= s\left(\frac{n}{f_s}+\frac{\Delta\phi}{2\pi f_c}\right) \quad \text{and}\\ \tilde{\tilde{s}}_i[n] &= \tilde{s}_i[n]\left(\frac{n}{f_s}+\frac{\Delta\phi}{2\pi f_c}\right)\exp\left(j\Delta\phi\right).\end{aligned}$$

In reality, $\acute{s}(n)$ is a sampled version of $s(t)$, which in turn consists of the transmitted signal corrupted with impulsive noise. As (5.91) is the same as (4.2), the additive noise in $\tilde{\tilde{s}}_i(f_s n/B)$ has independent I and Q components under the AWSαSN framework.

In Fig. 5.19 we present a schematic that depicts a practical implementation of passband-to-baseband conversion with constellation rotation at the receiver. By setting $f_s = 4f_c$, independent baseband noise components are guaranteed. This scheme is applicable for any constellation map that requires rotation while ensuring independence of noise components. The analog and digital blocks of the receiver are also highlighted.

5.8 On Fading Channels and AWSαSN

Till now we have discussed and analyzed digital communication schemes in a pure AWSαSN channel. If the analysis is extended to incorporate fading as well, then (5.2) modifies to

$$
\begin{aligned}
s_i(t) &= \Re\left\{\sqrt{\frac{2\mathcal{E}_{x_i}}{\mathcal{E}_g}}g(t)he^{j\phi_i}e^{j2\pi f_c t}\right\} && (5.92)\\
&= \Re\{\sqrt{\mathcal{E}_{x_i}}e^{j\phi_i}h\}\ell_I(t) + \Im\{\sqrt{\mathcal{E}_{x_i}}e^{j\phi_i}h\}\ell_Q(t). && (5.93)
\end{aligned}
$$

where $h \in \mathbb{C}$ is a circular symmetric complex Gaussian random variable. This channel is termed as a Rayleigh flat-fading channel. If h is independent over subsequent symbol transmissions, then (5.93) depicts block fading as well. We will analyze this channel in detail for multi-carrier transmission in Chapter 6, for which a single-carrier scheme is a special case. However, it is pertinent to comment briefly on the changes that need to be made to the mechanisms introduced in this chapter in the presence of fading.

From (5.93), do note that though $\mathbb{M}$ is the constellation employed at the transmitter, the received symbols are rotated by $\angle h$ and scaled by $|h|$. The alphabet at the receiver is $\bar{\mathbb{M}}$ such that $\sqrt{\mathcal{E}_{x_i}}e^{j\phi_i}h \in \bar{\mathbb{M}}\ \forall\ i \in \{1, 2, \ldots, M\}$. As the constellation structure is critical for the $f_s = 4f_c$ case, the random rotation caused by h will degrade error performance. From the discussion in Section 5.7, once the receiver is synchronized with the transmitter, the constellation may be rotated at will by introducing a sampling-offset at the receiver. Though the discussion revolves around QPSK, it can be extended to any constellation map. Therefore, if the channel is known at the receiver one needs only to rotate the constellations by $-\angle h$. If not catered for, then on the average $\angle h$ will produce a near-optimal rotation as highlighted for QPSK in Fig. 5.10. If the channel estimate is error-prone, then from Fig. 5.10 it is also noted that the scheme is robust to rotational error in the constellation. Divergence from the optimal rotation will result in a very small degradation in performance for a large range of rotational error as seen in Fig. 5.10.

For large M, the rotation principle at the receiver via sampling-offset still holds. Though the range of 'good' rotations is limited due to the packing of more symbols, one can still find a suitable range of sampling offsets to work with. The higher the number of points in the constellation, the more stringent the requirement of getting the exact rotation and therefore estimating $\angle h$. Thus, the limit to the size of a good constellation will be determined by the channel estimation scheme employed at the receiver.

With the constellations now rectified, the next step is to update the arguments of the cost functions used for detection in Sections 5.2.5 & 5.3. Instead of searching over $\mathbb{M}$ one needs to search over $\bar{\bar{\mathbb{M}}}$, where $\sqrt{\mathcal{E}_{x_i}}e^{j\phi_i}|h| \in \bar{\bar{\mathbb{M}}}\ \forall\ i \in \{1, 2, \ldots, M\}$.

5.9 Summary

In this chapter we have discussed and analyzed features of a good communications receiver for single-carrier modulation in impulsive noise. The analysis covers several schemes under the following design methodologies,

1. Soft-Estimates & Baseband Detection

2. Joint-Detection

The conventional (continuous-time) receiver was shown to perform poorly in non-Gaussian AWSαSN. By introducing a passband sampling criteria, suitable baseband detectors and efficient constellations, the error performance of the receiver is enhanced significantly while maintaining linearity of the system. From a practical perspective, linear receivers are easy to implement due to the availability of closed-form estimates. Further still, existing baseband signal processing techniques that assume system linearity, such as equalizers [1], may be used (without modification) in the receiver prior to the detection stage. A drawback of linear systems in AWSαSN is that they are far from optimal. We have categorized this mathematically as SNR degradation at the receiver.

If the linearity of the system is sacrificed, the error performance is enhanced further by employing suitable passband-to-baseband conversion schemes that generate more *robust* soft-estimates of the transmitted symbol. In terms of implementability, a cost-function needs to be minimized with respect to a complex variable every time a new estimate is generated. For $f_s = 4f_c$, we have shown that the bivariate minimization problem is reduced to two single-variable minimizations. Even then, this will be computationally taxing when large data rates are required. Further still, if the channel changes due to multipath and fading, the cost-functions need to be updated accordingly.

If soft-estimates are not required in the baseband, joint-detection can be used directly on the passband signal. Compared to the non-linear receiver schemes mentioned above, the error performance is better as the prior assumes isotropic baseband noise for detection purposes. Joint-detection is computationally efficient as minimizing the cost requires evaluating it for a small finite set of points and choosing the minimum within. However, due to the lack of soft-estimates, this approach has a drawback in terms of integration and flexibility with baseband processing techniques such as soft-decoders and equalizers. This can be circumvented by evaluating approximate LLRs. Like non-linear soft-estimation, any change in the channel requires modifying the cost-function for subsequent transmission.

We also propose an innovative but simple implementation of a single-carrier system that takes advantage of the baseband noise anisotropy by rotating constellations solely at the receiver side while also ensuring independence of noise components.

Chapter 6

OFDM in Impulsive Noise

Previously, we have highlighted the effect of AWSαSN in a single-carrier communication system. Statistical analysis of the noise process was thoroughly conducted after conversion to its complex baseband form. Conventional matched-filter conversion warrants the resulting noise to be heavy-tailed and isotropic. Under some rules and modification to the passband-to-baseband conversion mechanism, the in-phase and quadrature components of the resulting noise can have IID components. This provides significantly better maximum-likelihood (ML) detection performance over the conventional conversion case. These results open up a number of worthy problems including its application and analysis to multicarrier systems, which is the purpose of this chapter.

Orthogonal frequency-division multiplexing (OFDM) has garnered significant attention from the research community these last few decades. Modern day multicarrier systems are increasingly being incorporated with OFDM as it provides a number of advantages in terms of implementation and performance over other digital modulation schemes [1], [46]. High spectral efficiency, low inter-symbol interference (ISI) due to a guard interval, single-tap equalization and efficient implementation via the fast Fourier transform (FFT) algorithm are some of its attractive properties. The effects of impulsive noise on an OFDM system are established and well-known [37]. At the receiver, an N-point FFT is invoked on the received vector to generate an OFDM symbol block. The same operation would cause an impulse in the noise vector to be mapped onto an N-point complex sinusoid that affects all symbols in the OFDM block. Therefore, the *transformed* noise vector will have dependent components. As the FFT is a linear operation, multiple impulses in the noise vector will result in a summation of sinusoids with varying frequencies and amplitudes in the transformed noise vector. The frequency and amplitude of each of these sinusoids depend on the corresponding impulse location and weight, respectively, in the original noise vector.

A number of well-written articles discuss methods to mitigate this effect. The most applied concept is noise cancellation. Though sub-optimal, this is a valuable technique. In a practical OFDM system, a few symbols are reserved as nulls and pilots due to various constraints. Taking advantage of this, conceptual similarities between the OFDM transmit-receive equation and the syndromes in a block code are highlighted in [37], [38]. Reed-Solomon codes are then used to exploit the information in these symbols to estimate the noise with good effect. Similarly in [36], the authors have employed the relatively new concept of compressed sensing (CS) along with efficient and robust convex programmes developed in [77] to estimate the noise. The quoted articles focus on the Gaussian-Bernoulli-Gaussian (GBG) noise model and inherently assume the application of a linear passband-to-baseband conversion mechanism (like the one discussed in Chapter 4) that preserves the impulsiveness of the noise. In all impulsive noise cancellation techniques the objective is to remove (cancel) the impulses from the received vector so that Gaussian detection/decoding may be performed.

In this chapter we analyze the performance of ML detection in baseband OFDM for a channel contaminated with AWSαSN. Simulation results surprisingly depict that the achieved error rates *approach* the Gaussian error curve as the number of carriers increase. The results can be used as benchmarks for various schemes developed to mitigate impulsive noise. We provide insight to the relationship between ML detection and the CS approach in [78]. The pros and cons of the latter are discussed.

One of the main assumptions in the current literature is that the noise vector at the baseband level is sparse. Though a passband noise process may be impulsive (and thus sparse), this does not guarantee sparsity in the baseband. Building on the concepts acquired in Chapters 4 and 5, we highlight the design constraints within a linear passband-to-baseband conversion block that are sufficient to induce sparsity in the baseband noise vector for OFDM. It is shown that linear passband-to-baseband conversion is suboptimal and reduces the SNR at the receiver. We show how this can be entirely avoided by estimating the transmitted symbol block directly from the passband samples. The work presented in this chapter has been published in [94], [95].

6.1 The Baseband OFDM System Model

In digital communications, analysis is typically performed at the baseband level [1], [2]. Let $\mathbf{z} = [z_1, z_2, \ldots, z_N]^\mathsf{T}$ be the complex noise vector, i.e., $\mathbf{z} \in \mathbb{C}^N$.

Also define $\mathbf{x} = [x_1, x_2, \ldots, x_N]^\mathsf{T}$ as the $N \times 1$ OFDM symbol vector and $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N]$ the N-point *unitary* discrete Fourier transform (DFT) matrix with columns $\mathbf{a}_k$. Each $x_k \ \forall \ k \in \{1, 2, \ldots, N\}$ is selected from an M-symbol constellation. The baseband transmit-receive OFDM equation is then

$$\mathbf{y} = \mathbf{H}_c \mathbf{A}^\mathsf{H} \mathbf{x} + \mathbf{z} \tag{6.1}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_N]^\mathsf{T}$ is the received vector and $\mathbf{H}_c$ is the $N \times N$ complex circulant channel matrix. We consider zero-Doppler and Rayleigh block fading, therefore, $\mathbf{H}_c$ is time-invariant. The use of a cyclic-prefix and a sufficient-length guard interval is assumed. From the properties of $\mathbf{A}$ and $\mathbf{H}_c$, the latter can be diagonalized by $\mathbf{H} = \mathbf{A}\mathbf{H}_c\mathbf{A}^\mathsf{H}$, where $\mathbf{H} = \mathtt{diag}[h_1, h_2, \ldots, h_N]$ [46],[96]. We can thus rewrite (6.1) as

$$\mathbf{y} = \mathbf{A}^\mathsf{H} \mathbf{H} \mathbf{x} + \mathbf{z} \tag{6.2}$$

where $h_k \sim \mathcal{CN}(0, \sigma_h^2) \ \forall \ k \in \{1, 2, \ldots, N\}$ and all h_k are IID. The notation $\mathcal{CN}(0, \sigma_h^2)$ implies a circular-symmetric complex normal distribution with variance σ_h^2. Typically, an OFDM symbol block consists of data, pilots and nulls. The locations of these within $\mathbf{x}$ are known. Also, the receiver is assumed to have *complete knowledge* of the channel.

Practical systems assign nulls and pilots within an OFDM block due to various design/channel constraints [46]. For our analysis, we use $0 < K \leq N$ data-carriers and $N - K$ nulls. As the pilots are known, they can easily be accommodated within our problem formulation. We discuss this briefly in the last paragraph of Section 6.4.1. Other techniques such as forward error correction, time/frequency interleaving and reduction of the peak-to-average power ratio (PAPR) are typically employed to enhance OFDM performance. As our goal is to analyze the ML performance in impulsive noise at the baseband level, these schemes are considered as independent problems and are not discussed in this chapter.

Eq. (6.2) can be expressed in terms of only the actual transmitted data. Defining $\mathcal{L}_{\mathbf{x}} = \{\ell_1, \ell_2, \ldots, \ell_K\}$ as the set whose elements are the locations (indices) of the data symbols in $\mathbf{x}$ and $\mathbf{x}_{(1)} = [x_{\ell_1}, x_{\ell_2}, \ldots, x_{\ell_K}]^{\mathsf{T}}$ as the K-tuple data vector, we have from (6.2)

$$\mathbf{y} = \bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\mathbf{x}_{(1)} + \mathbf{z} \tag{6.3}$$

where $\bar{\mathbf{A}}^{\mathsf{H}} = [\mathbf{a}^*_{\ell_1}, \mathbf{a}^*_{\ell_2}, \ldots, \mathbf{a}^*_{\ell_K}]$ is of size $N \times K$ and $\bar{\mathbf{H}} = \mathtt{diag}[h_{\ell_1}, h_{\ell_2}, \ldots, h_{\ell_K}]$. The notation $\mathbf{a}^*_{\ell_1}$ denotes the complex conjugate of all elements in the vector $\mathbf{a}_{\ell_1}$. Similarly, we can combine the columns of $\mathbf{A}^*$ whose indices are not in $\mathcal{L}_{\mathbf{x}}$ to form the $N \times (N-K)$ matrix $\bar{\bar{\mathbf{A}}}^{\mathsf{H}}$. These columns correspond to the position of the nulls in $\mathbf{x}$. Like its vector counterpart, the notation $\mathbf{A}^*$ denotes the complex conjugate of all elements in $\mathbf{A}$. As the columns of $\mathbf{A}$ are orthonormal, we get

$$\begin{aligned} &\bar{\mathbf{A}}\bar{\mathbf{A}}^{\mathsf{H}} = \mathbf{I}_K, \\ &\bar{\bar{\mathbf{A}}}\bar{\bar{\mathbf{A}}}^{\mathsf{H}} = \mathbf{I}_{N-K} \text{ and} \\ &\bar{\mathbf{A}}\bar{\bar{\mathbf{A}}}^{\mathsf{H}} = \mathbf{0}_{K\times(N-K)}. \end{aligned} \tag{6.4}$$

where $\mathbf{I}_K$ and $\mathbf{0}_{K\times(N-K)}$ represent the $K \times K$ identity matrix and the $K \times (N-K)$ all-zero matrix, respectively.

In Section 5.4, the performance of single-carrier phase-shift keying with non-Gaussian SαS $\mathbf{z}$ with IID components was analyzed. Large error performance gains were achieved by exploiting the baseband PDF via rotated constellations. The optimal rotation angle and performance gain depends on the employed constellation. An extension of this scheme to the multi-carrier case, could comprise of rotating the symbols on every carrier by a certain angle. For the purpose of comparison, we will also show a few results for the case of no fading and near-optimal per-component rotation with $K = N$. The channel model for this is

$$\mathbf{y} = \mathbf{A}^{\mathsf{H}}\mathbf{H}_{\phi}\mathbf{x} + \mathbf{z} \tag{6.5}$$

where $\mathbf{H}_{\phi} = \texttt{diag}(e^{j\phi_1}, e^{j\phi_2}, \ldots, e^{j\phi_N})$ and ϕ_k is the rotation angle for the k^{th} carrier $\forall\ k \in \{1, 2, \ldots, N\}$. From (6.5), we observe that $\mathbf{H}_{\phi}\mathbf{x}$ is the *transmitted* OFDM symbol and therefore the receiver will have full knowledge of $\mathbf{H}_{\phi}$.

In (6.5), the optimal rotation angles are functions of the SNR, the per-carrier constellation pattern, the number of carriers and the noise impulsiveness. Evaluating a suitable $\mathbf{H}_{\phi}$ may not be feasible as practical channels introduce a random phase to each carrier via fading. This is mathematically characterized by $\mathbf{H}$ in (6.2). Also, transferring channel information to the transmitter is usually not an option. This problem is augmented by the fact that the channel may change by the time $\mathbf{H}$ is estimated and delivered. Further still, calculation of the optimal angle is computationally very complex. On the bright side, the range of angles for which the system performs at near-optimum levels is large and most *random* instances of $\mathbf{H}_\phi$ will offer good performance with high probability. As we will see later, the performance gain as N increases does not warrant the cost of calculating the optimal $\mathbf{H}_\phi$.

A more appropriate approach would allow analyzing the *average* error rates over all possible combinations of $\mathbf{H}$ in (6.2) and $\mathbf{H}_\phi$ in (6.5). From a practical point-of-view, this would provide a benchmark for the error performance of *any instance* of $\mathbf{H}$ the channel introduces or a random $\mathbf{H}_\phi$. The statistical characteristics of $\mathbf{z}$ will be briefly discussed next.

6.2 Statistical Characterization of the Complex Noise Vector

A typical passband-to-baseband conversion block is a linear system which retains in-band information [1]. As discussed in Chapter 4, this is optimal (in the ML sense) for AWGN and may be implemented in either continuous-time or on a non-lossy sampled version of the passband signal. Regardless of the implementation, the statistics of $\mathbf{z}$ do not change for passband AWGN. Precisely, if the double-sided noise PSD is $N_0/2$, then $z_n \sim \mathcal{CN}(0, N_0)\ \forall\ n \in \{1, 2, \ldots, N\}$, i.e., z_n has IID real and imaginary components such that $\Re\{z_n\} \overset{d}{=} \Im\{z_n\} \overset{d}{=} Z \sim \mathcal{N}(0, N_0/2)\ \forall\ n \in \{1, 2, \ldots, N\}$. Further still, the components of $\mathbf{z}$ are independent. As $\mathbf{z}$ is a Gaussian vector, this in turn implies it is *isotropic* [1]. For this case, without losing any information of the noise component contaminating

$\mathbf{x}_{(1)}$, one may multiply (6.3) with $\bar{\mathbf{A}}$ to get

$$\acute{\mathbf{y}} = \bar{\mathbf{H}}\mathbf{x}_{(1)} + \acute{\mathbf{z}}. \tag{6.6}$$

Here $\acute{\mathbf{y}} = \bar{\mathbf{A}}\mathbf{y}$ and $\acute{\mathbf{z}} = \bar{\mathbf{A}}\mathbf{z}$. The elements of $\acute{\mathbf{z}}$ are also $\mathcal{CN}(0, N_0)$ and independent [46].

For non-Gaussian AWSαSN, the distribution of $\mathbf{z}$ varies significantly with the passband sampling frequency. Extending the argument in Section 5.2.2 to the multi-carrier case, we see that the case of $\mathbf{z}$ with IID real and imaginary components offers the best error performance. In such a case, $\Re\{z_n\} \overset{d}{=} \Im\{z_n\} \overset{d}{=} Z \sim \mathcal{S}(\alpha, \delta_z)\ \forall\ n \in \{1, 2, \ldots, N\}$ and the components of $\mathbf{z}$ are independent. This implies isotropy if $\mathbf{z}$ is Gaussian. However, the joint-PDF for $\alpha \neq 2$ has tails directed along the positive and negative directions of each axis. Though still heavy-tailed, this should offer good system performance as $\mathbf{z}$ retains the sparsity (to some extent) of the passband AWSαSN process. This is the case primarily considered in the remainder of the chapter.

From the discussion in Section 4.2.3, if a continuous-time implementation is adopted, z_n is a complex isotropic SαS random variable in non-Gaussian AWSαSN. Therefore, $\Re\{z_n\} \overset{d}{=} \Im\{z_n\} \overset{d}{=} Z \sim \mathcal{S}(\alpha, \delta_z)\ \forall\ n \in \{1, 2, \ldots, N\}$ [9]. Unlike the Gaussian case, $\Re\{z_n\}$ and $\Im\{z_n\}$ are dependent [8],[9]. Note that this is *not* equivalent to defining $\mathbf{z}$ as an isotropic random vector with CF in (3.18). The dependency *within the components* of $\mathbf{z}$ depends on the length of the low-pass filter in the passband-to-baseband conversion block. Though it is possible to ensure independence within the components of $\mathbf{z}$, the isotropy between the I and

Q components of z_n essentially reduces the sparsity of $\mathbf{z}$ by a factor of two in comparison to the IID case. This is due to the fact that information in the tails within each sub-carrier are lost. As pointed out in Section 5.5, the effect of this on a single-carrier system is severe.

In the literature, baseband analysis in impulsive noise has been conducted both for the per-carrier isotropic case with independent components and for $\mathbf{z}$ with IID real and imaginary components [37], [58], [67]. For easy reference, we term these configurations as $\mathbf{z}_{\text{ISO}}$ and $\mathbf{z}_{\text{IID}}$, respectively. Though our primary focus is on the latter, we discuss ML detection results for both in the next section to see their similarities and differences.

6.3 Performance Analysis of Baseband OFDM

6.3.1 ML Detection

The ML detection rule of the OFDM symbol block in (6.3) is given by

$$\hat{\mathbf{x}}_{(1)} = \underset{\boldsymbol{\zeta} \in \mathbb{M}}{\texttt{arg}\,\texttt{max}}\; f_{\mathbf{z}}(\mathbf{y} - \bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\zeta}) \tag{6.7}$$

where $f_{\mathbf{z}}(\cdot)$ is the $2N$-dimensional joint-PDF of $\mathbf{z}$ and $\mathbb{M}$ is the set of all possible OFDM symbols such that $\mathbf{x}_{(1)} \in \mathbb{M}$. Given that $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$ and denoting $\bar{\mathbf{A}} = [\bar{\mathbf{a}}_1, \bar{\mathbf{a}}_2, \ldots, \bar{\mathbf{a}}_N]$, we have

$$\begin{aligned} \hat{\mathbf{x}}_{(1)} &= \underset{\boldsymbol{\zeta} \in \mathbb{M}}{\texttt{arg}\,\texttt{min}} \left(\prod_{n=1}^{N} f_z(y_n - \bar{\mathbf{a}}_n^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\zeta}) \right)^{-1} &(6.8)\\ &= \underset{\boldsymbol{\zeta} \in \mathbb{M}}{\texttt{arg}\,\texttt{min}} \sum_{n=1}^{N} -\log f_z(y_n - \bar{\mathbf{a}}_n^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\zeta}) &(6.9) \end{aligned}$$

where $f_z(\cdot) = f_Z(\Re\{\cdot\})f_Z(\Im\{\cdot\})$ is the bivariate PDF of $z_n \ \forall \ n \in \{1, 2, \ldots, N\}$. The expressions in (6.8) and (6.9) are equivalent as the cost function in (6.8) is strictly positive and $\log(\cdot)$ is a monotonically increasing function in this region.

For $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{ISO}}$, we note that $f_z(\cdot)$ has algebraic tails when z is non-Gaussian SαS and therefore is an *algebraic function* of $\|\cdot\|$ due to the isotropy. The rule in (6.9) then simplifies to

$$\hat{\mathbf{x}}_{(1)} = \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}} \sum_{n=1}^{N} \log |y_n - \bar{\mathbf{a}}_n^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\zeta}|^2. \tag{6.10}$$

For the Gaussian case we note that $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{IID}} \overset{d}{=} \mathbf{z}_{\mathrm{ISO}}$. From (6.6), we can simplify (6.9) to

$$\begin{aligned}
\hat{\mathbf{x}}_{(1)} &= \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{max}}\ f_{\mathbf{z}}(\acute{\mathbf{y}} - \bar{\mathbf{H}}\boldsymbol{\zeta}) \\
&= \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}} \sum_{k=1}^{K} -\log f_z(\acute{y}_k - h_{\ell_k}\zeta_k) \\
&= \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}} \sum_{k=1}^{K} |\acute{y}_k - h_{\ell_k}\zeta_k|^2 = \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}}\ \|\acute{\mathbf{y}} - \bar{\mathbf{H}}\boldsymbol{\zeta}\|^2
\end{aligned} \tag{6.11}$$

where $\acute{\mathbf{y}} = [\acute{y}_1, \acute{y}_2, \ldots, \acute{y}_K]^{\mathsf{T}}$, $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \ldots, \zeta_K]^{\mathsf{T}}$ and $\|\cdot\|$ is the Euclidean norm. In fact, (6.11) is the ML-detection rule for *any* unimodal isotropic $\acute{\mathbf{z}}$ as its PDF is a monotonically decreasing function of $\|\acute{\mathbf{z}}\|$. Though (6.11) is a combinatorial problem, the computational cost increases linearly with the number of carriers [46]. This is because the cost function is a sum of individual terms for each k and therefore each term can be independently minimized. This makes it easy to perform even for moderately large N. Do note that the cost function in (6.9) is a sum of N elements, while that in (6.11) is a sum of K. This is due to the fact that the information in the null carriers is irrelevant for the Gaussian case, but not in general.

6.3.2 Optimizing Constellations

One key attribute of $\mathbf{z} = \mathbf{z}_{\text{IID}}$ is that its joint-PDF will have tails directed along each Cartesian axis in both the positive and negative directions. This has already been presented for the single-carrier ($N = 1$) Cauchy case in Fig. 4.4d. The real and imaginary axis correspond to $\Re\{z_1\}$ and $\Im\{z_1\}$, respectively, and are IID Cauchy random variables. For values of α near 2, the tails are still visible but less pronounced than those in Fig. 4.4d.

A strong result from the discussion in Section 3.2.2 is that $\Lambda(\mathbf{s})$ corresponding to $\mathbf{z}$ is non-zero only when $\mathbf{s}$ is directed along each positive and negative Cartesian axis. In other words, as $\mathbf{z} \in \mathbb{C}^N$, there are $4N$ unique vectors $\mathbf{s} \in \mathbb{S}_{2N}$ (where $\mathbb{S}_{2N} = \{\mathbf{s}|\mathbf{s} \in \mathbb{R}^{2N}, \|\mathbf{s}\| = 1\}$) and their respective weights $\Lambda(\mathbf{s})$ that completely characterize the statistics of $\mathbf{z}$. We denote the set of these vectors by $\mathbb{T}$. The relationship between the PDF's tails and $\mathbf{s} \in \mathbb{T}$ is very clear: The tails are directed along the vectors in $\mathbb{T}$. If $\mathbf{X}$ has a CF of the form in (3.19), then any linear mapping of $\mathbf{X}$ results in a similar transformation in the tails of its PDF, and hence the vectors in $\mathbb{T}$. The example below depicts the tail transformation of an SαS random vector under linear mapping:

Example

Let $\mathbf{Y} = \mathbf{BX}$, where $\mathbf{Y}$ and $\mathbf{X}$ are real N-dimensional SαS vectors with IID components and $\mathbf{B}$ is a real $N \times N$ matrix with its i^{th} column denoted by $\mathbf{b}_i$.

Then

$$\begin{aligned}\Phi_{\mathbf{Y}}(\theta) &= \Phi_{\mathbf{X}}(\mathbf{B}^{\mathsf{T}}\theta) \\ &= \Phi_{\mathbf{X}}(\theta) = \exp\left(-\delta^{\alpha}\left(\sum_{i=1}^{N}|\mathbf{b}_i^{\mathsf{T}}\theta|^{\alpha}\right)\right).\end{aligned} \tag{6.12}$$

On comparison with (3.17) and (3.19), we observe that $\Lambda(\mathbf{s})$ consists of a finite sum of Dirac delta functions located at $\mathbf{s}_i = \pm\mathbf{b}_i/\|\mathbf{b}_i\|$ with weight $\delta^{\alpha}\|\mathbf{b}_i\|/2$ $\forall\ i \in \{1, 2, \ldots, N\}$. The weight of each transformed tail is proportional to $\|\mathbf{b}_i\|$. If $\mathbf{B}$ is invertible then the number of tails in the PDF of $\mathbf{Y}$ is the same as that of $\mathbf{X}$. □

We know that an impulse in $\mathbf{z}$ affects all symbols in $\mathbf{x}$ after the FFT operation. From a statistical perspective, this phenomenon is represented by the tails in the PDF of $\mathbf{H}^{-1}\mathbf{A}\mathbf{z}$ and $\mathbf{H}_{\phi}^{\mathsf{H}}\mathbf{A}\mathbf{z}$ in (6.2) and (6.5), respectively. With suitable values of ϕ_i in (6.5), the tail vectors in $\mathbb{T}$ of the baseband noise PDF (shifted to $\mathbf{A}^{\mathsf{H}}\mathbf{H}_{\phi}\mathbf{x}$) do not point directly towards any other OFDM constellation point. One may use a geometric approach using $\mathbb{T}$ and the OFDM constellation to accomplish this. However, we derive $\mathbf{H}_{\phi}$ by minimizing a cost function based on the SEP.

For $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$, the ML detection rule for (6.5) given $\mathbf{x}$ is transmitted is

$$\hat{\mathbf{x}} = \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\mathtt{arg\,min}} \sum_{n=1}^{N} -\log f_z(y_n - \mathbf{a}_n^{\mathsf{H}}\mathbf{H}_{\phi}\boldsymbol{\zeta}) \tag{6.13}$$

where $f_z(\cdot) = f_Z(\Re\{\cdot\})f_Z(\Im\{\cdot\})$ is the bivariate PDF of z_n $\forall\ n \in \{1, 2, \ldots, N\}$. Analogous to (5.69), if the OFDM symbols are equiprobable, the symbol error probability for the detection scheme in (6.13) is given by

$$\mathrm{SEP}(\mathbf{H}_\phi) = \frac{1}{M^N} \sum_{i=1}^{M^N} \int_{\mathbf{y}:\hat{\mathbf{x}}\neq\mathbf{x}_i} \prod_{l=1}^{N} f_z(y_l - \mathbf{a}_l^{\mathsf{H}} \mathbf{H}_\phi \mathbf{x}_i) \mathrm{d}\mathbf{y} \tag{6.14}$$

where $\mathbf{x}_i \ \forall \ i \in \{1, 2, \ldots, M^N\}$ is the i^{th} N-tuple in the set $\mathbb{M}$. The integration is performed over all $\mathbf{y}$ such that $\hat{\mathbf{x}} \neq \mathbf{x}_i$ where $\mathbf{x}_i$ is the transmitted OFDM symbol.

The expression in (6.14) is not solvable as the integration is performed over complex areas. This is augmented by the fact that $f_z(\cdot)$ is not available in closed form with the exception of the Cauchy case and averaging needs to be performed over *all* transmitted symbols $\mathbf{x}_i \in \mathbb{M}$. On the other hand, by extrapolating the results in Section 5.4, the ML error performance for (6.13) is almost constant (near-optimal) for a large range of ϕ_i at high SNR. Thus evaluating the optimal $\mathbf{H}_\phi$ at every SNR instance does not provide adequate gain. Therefore $\mathbf{H}_\phi$ may be evaluated just once for a given SNR where error performance meets requirements.

Calculating $\mathbf{H}_\phi$ from (6.14) is not trivial and requires further simplifications. Using a similar line of reasoning in (5.70), we propose minimizing the following cost function:

$$J(\mathbf{H}_\phi) = \sum_{i=1}^{M^N} \prod_{l=1}^{N} \sum_{\substack{m=0,\\ m\neq i}}^{M^N} f_z\left(\mathbf{a}_l^{\mathsf{H}} \mathbf{H}_\phi \left(\mathbf{x}_m - \mathbf{x}_i\right)\right). \tag{6.15}$$

Observe that $J(\cdot)$ is a normalized version of $\mathrm{SEP}(\cdot)$ with the detection regions limited to the points $\mathbf{y} = \mathbf{a}_i^H \mathbf{H}_\phi \mathbf{x}_m$ where $m \neq i$. The expression in (6.15) is convex in $\mathbf{H}_\phi$ and can be minimized via conventional techniques such as gradient descent. However, this is still very complex to solve for large N. Though many sub-optimal schemes may be designed to evaluate $\mathbf{H}_\phi$, it is a different problem which will not be pursued in this work.

6.3.3 Simulations

Following the discussion in Section 5.5.1 and the expression in (5.76), we use the following definitions for $\mathcal{E}_b/N_0$:

$$\frac{\mathcal{E}_b}{N_0} = \frac{\mathcal{E}_x \sigma_h^2}{4\delta_z^2 \log_2 M} \tag{6.16}$$

for (6.2) and

$$\frac{\mathcal{E}_b}{N_0} = \frac{\mathcal{E}_x}{4\delta_z^2 \log_2 M} \tag{6.17}$$

for (6.5). where $\mathcal{E}_x = E[\|\mathbf{x}\|^2]/K$ is the average energy per-carrier.

All simulations and analysis in this section are conducted for Cauchy $\mathbf{z}$ with $K = N$. The Cauchy distribution shares the heavy-tailed property common to all non-Gaussian SαS distributions. Thus, results for this case can be intuitively extended to all other heavy-tailed SαS cases. The simulations are conducted via the Monte Carlo method. At least 4000 errors are accumulated for BER/SER$<10^{-3}$ and 1000 errors otherwise.

In Fig. 6.1 we present the BER curves for the model in (6.2). The per-carrier constellation is BPSK and $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$ is a Cauchy random vector. The BER is averaged over all possible instances of $\mathbf{H}$ and the receiver is assumed to have full knowledge of the channel. It is observed that the error rates *tend towards* the Gaussian error curve as N increases. We see a similar trend for the SER in Fig. 6.2 when the per-carrier constellation is QPSK. Though not presented here, similar convergence is expected of other constellations.

Intuitively, we know that the information within an impulse is scattered over a large bandwidth. This will be *larger* than the available bandwidth per-carrier. Increasing N essentially allows the scheme to access a larger bandwidth. In fact, the increase in bandwidth is directly proportional to N. This allows the detection process to harness more noise information and thus enhance performance in non-Gaussian AWSαSN. Increasing N in AWGN does not improve error performance as the noise information that affects any sub-carrier is constrained to the latter's bandwidth. According to this reasoning, one can expect the detection performance in impulsive noise to *converge* to a certain level when *sufficient* information is harnessed. This, in turn, is accomplished by increasing the OFDM symbol bandwidth (and therefore N).

For the model in (6.13), we present results in Fig. 6.3 for the Cauchy case with per-carrier constellation BPSK. We evaluate $\mathbf{H}_\phi$ at $\mathcal{E}_b/N_0$ (dB)= 20dB for each scheme by minimizing (6.15) via gradient descent. The error performance is then calculated over all $\mathcal{E}_b/N_0$ values with the same $\mathbf{H}_\phi$. As with Fig. 6.1, the error performance becomes better with increasing N. This trend however is more pronounced in Fig. 6.3.

Though $\mathbf{H}_\phi$ significantly influences the error performance in the single-carrier case, the range of values for which it performs well enhances with increasing N. We plot the *average* BER over all possible values of $\mathbf{H}_\phi$ for $N = 2$, 4 and 8 in Fig. 6.4. For comparison, the near-optimal error curves in Fig. 6.3

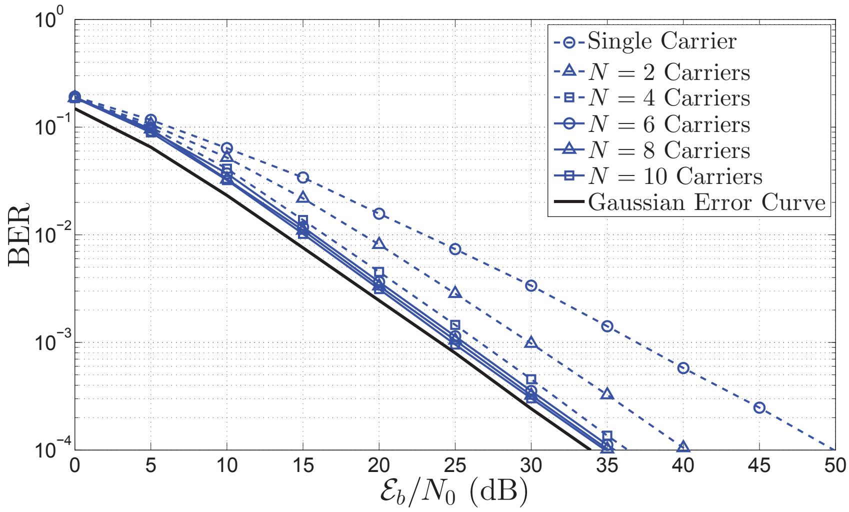


Figure 6.1: ML detection BER performance averaged over $\mathbf{H}$ with Cauchy $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{IID}}$ in (6.2). The curves are generated for various $K = N$ with per-carrier constellation BPSK.

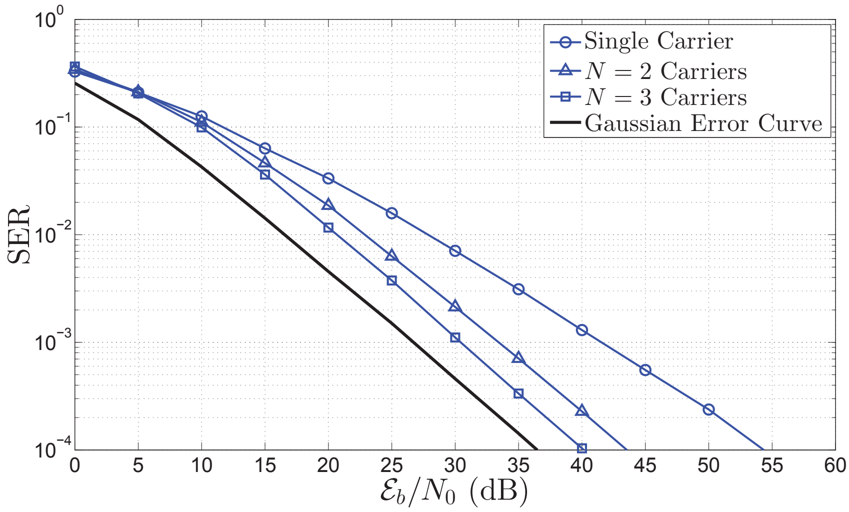


Figure 6.2: ML detection SER performance averaged over $\mathbf{H}$ with Cauchy $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{IID}}$ in (6.2). The curves are generated for various $K = N$ with per-carrier constellation QPSK.

corresponding to these values of N are redrawn in Fig. 6.4. We note that as the number of carriers increase, the difference between the near-optimal and average error performance decreases substantially. Thus the probability of a *random* $\mathbf{H}_\phi$

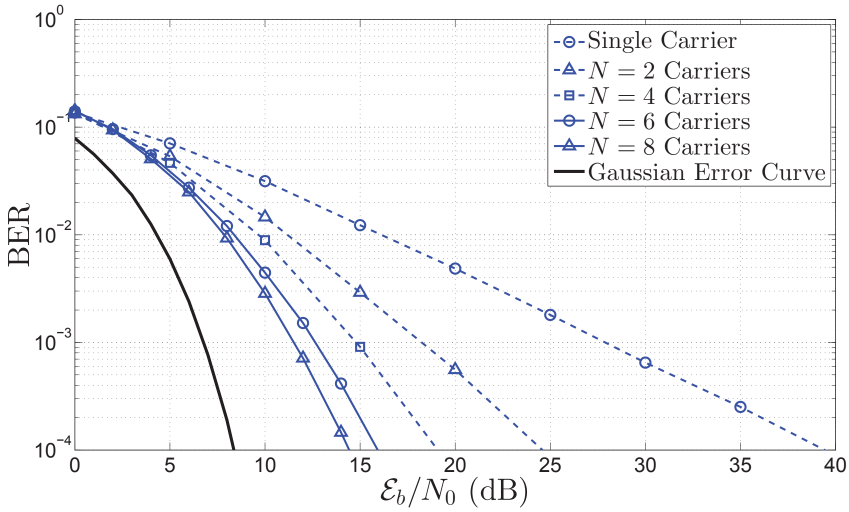


Figure 6.3: ML detection BER performance for various $K = N$ with Cauchy $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{IID}}$ in (6.5). The per-carrier constellation is BPSK and $\mathbf{H}_\phi$ is optimized for $\mathcal{E}_b/N_0$ (dB)= 20dB.

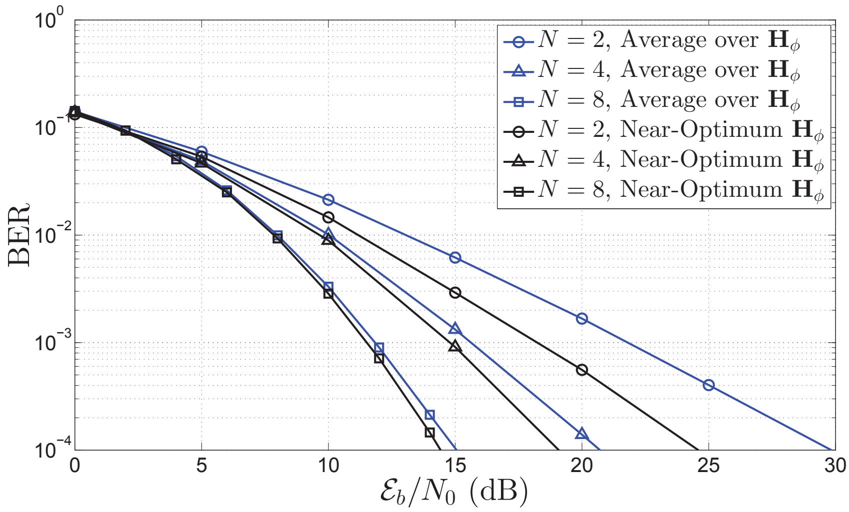


Figure 6.4: ML detection BER performance averaged over $\mathbf{H}_\phi$ with Cauchy $\mathbf{z} \overset{d}{=} \mathbf{z}_{\mathrm{IID}}$ in (6.5). The curves are generated for various $K = N$ with per-carrier constellation BPSK.

producing near-optimal error performance in PSK increases with N. In essence, the dependence on the constellation structure actually *reduces* with increasing N in OFDM. These trends can be extended to the fading model in (6.2) where

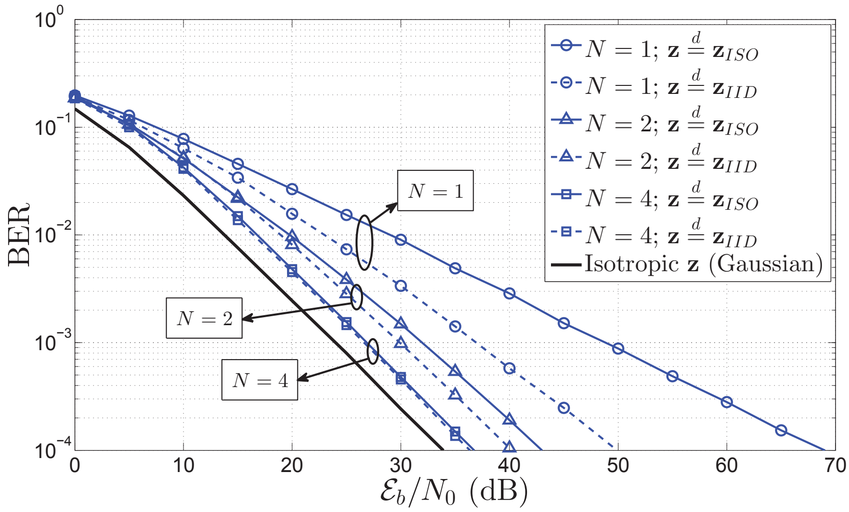


Figure 6.5: ML detection BER performance averaged over $\mathbf{H}$ for Cauchy $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{ISO}}$ (solid lines) in (6.2). The curves are generated for various $K = N$ with per-carrier constellation BPSK and compared with those in Fig. 6.1 (dashed lines).

$\mathbf{H}_\phi$ is essentially replaced by the random channel matrix $\mathbf{H}$.

In Fig. 6.5, we compare the BER for $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{ISO}}$ to their counterparts in Fig. 6.1 for the model in (6.2). One can clearly see the performance difference between the two statistical configurations of $\mathbf{z}$ especially when $N = 1$ and $N = 2$. Further still, for $N = 4$ the error performance is almost identical, implying that $\mathbf{z}_{\text{IID}}$ and $\mathbf{z}_{\text{ISO}}$ offer almost equal information about the impulses under ML detection. In either case, however, there is a remarkable improvement in error performance in comparison to their single-carrier ($N = 1$) counterpart. This is clearly observed even for small $N > 1$. The increase in performance is attributed to the fact that the DFT operation spreads the transmitted information amongst the carriers. If a received sample is affected by an impulse, joint-detection takes advantage of the spread in information to output more robust estimates of the transmitted symbol block.

The results in this section clearly highlight the advantages of OFDM in AWSαSN. As N increases the error performance of the system improves and approaches the Gaussian error curve, even for $K = N$, for both $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$ and $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{ISO}}$ for the channel model in (6.2). Further still, the optimal constellation structure, which is a significant design characteristic in the single-carrier case, offers decreasing performance advantage with increasing N.

A problem associated with the ML-detection rules in (6.9) and (6.10) is that the computational cost increases exponentially with K. This is not an issue when K is small. In (6.9), the issue is further compounded due to the unavailability of closed-form SαS PDFs. Therefore, estimating $\mathbf{x}_{(1)}$ for large K becomes computationally inefficient and eventually, intractable. Further still, one needs to estimate α and δ associated with z_n before ML detection can be truly applied. In the next section, we lay out an approach that is not only unhampered by these problems but (under some constraints) results in near-ML performance when K is large.

6.4 Baseband OFDM Receiver Design

Our analysis will be primarily based on $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$. We use the results to comment on the $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{ISO}}$ case as well.

6.4.1 Problem Formulation

Instead of performing detection directly in (6.9), one can first try to evaluate *soft-estimates* of $\mathbf{x}_{(1)}$. The detection stage may then be employed subsequently.

We can modify (6.9) to get the ML estimate of $\mathbf{x}_{(1)}$:

$$\hat{\mathbf{x}}_{(1)} = \underset{\boldsymbol{\mu}\in\mathbb{C}^K}{\texttt{arg}\,\texttt{min}} \sum_{n=1}^{N} -\log f_z(y_n - \bar{\mathbf{a}}_n^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\mu}). \tag{6.18}$$

Do note how $\boldsymbol{\mu}$ spans the entire $\mathbb{C}^K$ space. By using a change of variables $\gamma_n = y_n - \bar{\mathbf{a}}_n^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\mu}$, we can convert the unconstrained problem in (6.18) into a constrained one with linear equalities:

$$\begin{aligned} \hat{\mathbf{x}}_{(1)}, \hat{\mathbf{z}} = \quad & \underset{\boldsymbol{\mu},\boldsymbol{\gamma}}{\texttt{arg}\,\texttt{min}} \quad \textstyle\sum_{n=1}^{N} -\log f_z(\gamma_n) \\ & \texttt{s.t} \qquad \mathbf{y} = \bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\mu} + \boldsymbol{\gamma}. \end{aligned} \tag{6.19}$$

Here γ_n is the n^{th} element of $\boldsymbol{\gamma} \in \mathbb{C}^N$. The vector $\boldsymbol{\gamma} = \hat{\mathbf{z}}$ is an estimate of $\mathbf{z}$, and along with $\boldsymbol{\mu} = \hat{\mathbf{x}}_{(1)}$, minimizes the cost function in (6.19). As any one-to-one mapping of the constraints (or cost function) does not influence the minimization process [78], we may express (6.19) as

$$\begin{aligned} \hat{\mathbf{x}}_{(1)}, \hat{\mathbf{z}} = \quad & \underset{\boldsymbol{\mu},\boldsymbol{\gamma}}{\texttt{arg}\,\texttt{min}} \quad \textstyle\sum_{n=1}^{N} -\log f_z(\gamma_n) \\ & \texttt{s.t} \qquad \mathbf{A}\mathbf{y} = \mathbf{A}\bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\boldsymbol{\mu} + \mathbf{A}\boldsymbol{\gamma}. \end{aligned} \tag{6.20}$$

From the equalities in (6.4), we can further simplify (6.20) to

$$\begin{aligned} \hat{\mathbf{x}}_{(1)}, \hat{\mathbf{z}} = \quad & \underset{\boldsymbol{\mu},\boldsymbol{\gamma}}{\texttt{arg}\,\texttt{min}} \quad \textstyle\sum_{n=1}^{N} -\log f_z(\gamma_n) \\ & \texttt{s.t} \qquad \bar{\mathbf{A}}\mathbf{y} = \bar{\mathbf{H}}\boldsymbol{\mu} + \bar{\mathbf{A}}\boldsymbol{\gamma} \\ & \qquad\qquad \bar{\bar{\mathbf{A}}}\mathbf{y} = \bar{\bar{\mathbf{A}}}\boldsymbol{\gamma}. \end{aligned} \tag{6.21}$$

Do note that there are two sets of equalities in (6.21); the first consists of the data vector and the latter just the nulls. We can express $\hat{\mathbf{x}}_{(1)}$ explicitly in terms of $\mathbf{x}_{(1)}$ and the estimation error $\mathbf{e}$. From (6.3), $\mathbf{y} = \bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\hat{\mathbf{x}}_{(1)} + \hat{\mathbf{z}}$, so

$$\begin{aligned}\hat{\mathbf{x}}_{(1)} &= \bar{\mathbf{H}}^{-1}\bar{\mathbf{A}}\,(\mathbf{y} - \hat{\mathbf{z}}) \\ &= \mathbf{x}_{(1)} + \underbrace{\bar{\mathbf{H}}^{-1}\bar{\mathbf{A}}\,(\mathbf{z} - \hat{\mathbf{z}})}_{\text{estimation error}} \\ &= \mathbf{x}_{(1)} + \mathbf{e}. \end{aligned} \tag{6.22}$$

ML estimation theory in reference to stable distributions and their parameterizations have been covered well in [50], [97], [98]. Under certain regularity conditions, the limiting properties generally associated with ML estimates extend to stable parameters: they are efficient, consistent and asymptotically normal [99]. In the limit $N \to \infty$, $\mathbf{e}$ is a Gaussian vector with *independent* real and imaginary components, and is given by

$$\mathbf{e} \sim \mathcal{CN}(\mathbf{0}_{N\times 1}, \frac{2\delta_z^2}{\mathcal{I}^{(0)}}(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})^{-1}), \tag{6.23}$$

where $\mathcal{I}^{(0)}$ is the Fisher information of the location parameter provided by *one* real noise sample with distribution $\mathcal{S}(\alpha, 1)$ [100]. A proof is provided in Appendix-A.3. Given (6.22) and (6.23), the optimal detection rule is

$$\begin{aligned}\acute{\mathbf{x}}_{(1)} &= \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}}\ \|\bar{\mathbf{H}}(\hat{\mathbf{x}}_{(1)} - \boldsymbol{\zeta})\| \\ &= \underset{\boldsymbol{\zeta}\in\mathbb{M}}{\texttt{arg}\,\texttt{min}}\ \sum_{k=1}^{K} |\hat{x}_{\ell_k} - \zeta_k|^2. \end{aligned} \tag{6.24}$$

where $\hat{\mathbf{x}}_{(1)} = [\hat{x}_{\ell_1}, \hat{x}_{\ell_2}, \ldots, \hat{x}_{\ell_K}]^{\mathsf{T}}$. Analogous to (6.11), the minimization in (6.24) is equivalent to minimizing per-carrier and is therefore computationally easy to perform. As N is finite in practical OFDM systems, (6.23) may not truly represent the distribution of $\mathbf{e}$. Moreover, as α decreases, the convergence to (6.23) is *increasingly slower* [50]. However, the reason for generating soft-values in the first place is to allow for low-complexity at the detection stage. Also, the approximation in (6.24) offers good error performance for practical values of α and moderately large N. This is justified by the BER results in our simulations - see Section 6.4.3.

In the Gaussian case, the ML estimate of $\mathbf{x}_{(1)}$ is evaluated from (6.11) by substituting $\boldsymbol{\zeta}$ with $\boldsymbol{\mu} \in \mathbb{C}^K$ and is in analytical form:

$$\begin{aligned} \hat{\mathbf{x}}_{(1)} &= \bar{\mathbf{H}}^{-1}\acute{\mathbf{y}} \\ &= \mathbf{x}_{(1)} + \underbrace{\bar{\mathbf{H}}^{-1}\acute{\mathbf{z}}}_{\mathbf{e}}. \end{aligned} \tag{6.25}$$

This is also the linear least square solution of (6.2). From the discussion on (6.6), $\acute{\mathbf{z}}$ is a Gaussian vector. Therefore,

$$\begin{aligned} \mathbf{e} &\sim \mathcal{CN}(\mathbf{0}_{N\times 1}, E[\mathbf{e}\mathbf{e}^{\mathsf{H}}]) \\ &= \mathcal{CN}(\mathbf{0}_{N\times 1}, 4\delta_z^2(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})^{-1}) \end{aligned}$$

for *all* N. Given (6.25), one can then employ isotropic (per-carrier) detection via the rule in (6.24). This overall process is equivalent to the joint-detection rule in (6.11).

As discussed in Chapter 3, $f_z(\cdot)$ is generally not in closed form. Further still, the cost function in (6.21) is *not* convex as $f_Z(\cdot) \approx D_{\alpha,\delta_z}|\cdot|^{-\alpha-1}$ at the tails where D_{α,δ_z} is a positive constant dependent on α and δ_z [8], [9]. Therefore, solving (6.21) (even for small N) may not be practically feasible. From generalized ML estimation (or M-estimation) theory [65], $f_z(\cdot)$ in (6.21) is replaced by a more general function $\rho(\cdot) \in \mathbb{R}^+$, i.e.,

$$\begin{aligned} \hat{\mathbf{x}}_{(1)}, \hat{\mathbf{z}} = \quad & \underset{\boldsymbol{\mu},\boldsymbol{\gamma}}{\texttt{arg min}} \quad \textstyle\sum_{n=1}^{N} -\log \rho(\gamma_n) \\ & \texttt{s.t} \quad \bar{\mathbf{A}}\mathbf{y} = \bar{\mathbf{H}}\boldsymbol{\mu} + \bar{\mathbf{A}}\boldsymbol{\gamma} \\ & \qquad\;\; \bar{\bar{\mathbf{A}}}\mathbf{y} = \bar{\bar{\mathbf{A}}}\boldsymbol{\gamma}. \end{aligned} \tag{6.26}$$

To achieve near-ML performance, $\rho(\cdot)$ should approximate $f_z(\cdot)$ very well. Though the efficiency of the estimator reduces, choosing a suitable $\rho(\cdot)$ may significantly lessen the computational cost of evaluating $\hat{\mathbf{x}}_{(1)}$. This is discussed next.

On a final note, to accommodate the pilot symbols in the formulation, one needs to add the additional equalities $\mathbf{A}_{\mathsf{P}}\mathbf{y} = \mathbf{H}_{\mathsf{P}}\mathbf{x}_{\mathsf{P}} + \mathbf{A}_{\mathsf{P}}\boldsymbol{\gamma}$ to (6.26). Analogous to the constructions of $\bar{\mathbf{A}}$ and $\mathbf{x}_{(1)}$, $\mathbf{A}_{\mathsf{P}}^{\mathsf{H}}$ consists of the columns of $\mathbf{A}^{\mathsf{H}}$ corresponding to the locations of the pilot symbols $\mathbf{x}_{\mathsf{P}}$ in $\mathbf{x}$. Similarly, $\mathbf{H}_{\mathsf{P}}$ is the diagonal submatrix of $\mathbf{H}$ with entries corresponding to the locations of the elements of $\mathbf{x}_{\mathsf{P}}$ in $\mathbf{x}$.

6.4.2 The L_p-norm as a Cost Function

In the literature, the L_p-norm for $p < 2$ has been used effectively to counter impulsive noise with IID samples [9], [10], [57]. Substituting $-\log \rho(\cdot) = |\Re\{\cdot\}|^p +$

$|\Im\{\cdot\}|^p$ in (6.26), we get

$$\begin{aligned}\hat{\mathbf{x}}_{(1)}, \hat{\mathbf{z}} = \quad & \underset{\boldsymbol{\mu},\boldsymbol{\gamma}}{\texttt{argmin}} \quad \|\boldsymbol{\gamma}\|_p \\ & \texttt{s.t} \quad \bar{\mathbf{A}}\mathbf{y} = \mathbf{H}\boldsymbol{\mu} + \bar{\mathbf{A}}\boldsymbol{\gamma} \\ & \qquad\;\; \bar{\bar{\mathbf{A}}}\mathbf{y} = \bar{\bar{\mathbf{A}}}\boldsymbol{\gamma}\end{aligned} \tag{6.27}$$

where $\|\cdot\|_p$ denotes the L_p-norm. The L_p-norm for $1 \leq p \leq 2$ is a convex function and may be readily solved via low-complexity numerical techniques [78]. From another perspective, (6.27) arises from approximating $f_z(\cdot)$ by

$$\begin{aligned}f_z(\cdot) &\approx f_G(\Re\{\cdot\}) f_G(\Im\{\cdot\}) \\ &= C^2_{p,\delta_z} \exp\left(-\frac{|\Re\{\cdot\}|^p + |\Im\{\cdot\}|^p}{\delta_z^p}\right),\end{aligned} \tag{6.28}$$

where

$$f_G(x) = C_{p,\delta_z} \exp\left(-\frac{|x|^p}{\delta_z^p}\right) \tag{6.29}$$

is the zero-mean univariate PDF of a generalized Gaussian distribution (GGD) with scale δ_z, shape parameter $p \in \mathbb{R}^+$ and C_{p,δ_z} is a positive constant dependent on p and δ_z. The GGD is heavy-tailed for $p < 2$.

Unlike (6.21), it is observed that the cost function in (6.27) is not dependent on δ_z. As the q^{th} order moment of an SαS distribution is finite if and only if $q < \alpha$ [9], (6.28) converges (in the ergodic-sense) to a finite value for $p < \alpha$. For the problem to be simultaneously convex and convergent, $1 \leq p < \alpha$. It is desirable for p to lie within this range. This is justifiable too as $\alpha \geq 1.5$ is typically a good fit for practical impulsive noise scenarios [3], [4].

From the discussion in Section 6.2, one aspect of $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$ is that the noise realizations will be sparse. Drawing insights from compressed sensing (CS) theory [77], [101], the L_1-norm recovery of $\mathbf{z}$ given the $N-K$ complex samples $\bar{\bar{\mathbf{A}}}\mathbf{y}$ is

$$\begin{aligned} \hat{\mathbf{z}} = \; & \underset{\boldsymbol{\gamma}}{\arg\min} \quad \|\boldsymbol{\gamma}\|_1 \\ & \texttt{s.t} \qquad \bar{\bar{\mathbf{A}}}\mathbf{y} = \bar{\bar{\mathbf{A}}}\boldsymbol{\gamma}. \end{aligned} \tag{6.30}$$

From (6.3) and (6.4), one can subsequently evaluate the soft-estimate of the OFDM symbol:

$$\hat{\mathbf{x}}_{(1)} = \bar{\mathbf{H}}^{-1}\bar{\mathbf{A}}\left(\mathbf{y} - \hat{\mathbf{z}}\right). \tag{6.31}$$

If the CS approach is compared with (6.27) for $p = 1$, we see that though both may be readily solved via linear programming, the latter has more computational cost due to the added equality constraints. Do note that the CS or L_p-norm estimation schemes do not require any information about α and δ_z and are therefore non-parametric. However, p may be optimized as a function of α in the latter. Although (6.27) contains added information and $\mathbf{z}$ is not *truly* sparse as the probability of any $z_n \; \forall \; n \in \{1, 2, \ldots, N\}$ to be equal to zero is infinitely small, our simulations showed that both techniques perform *at par* for any K and N. For the extreme case $K = N$, $\hat{\mathbf{z}} = \mathbf{0}_{N\times 1}$ in (6.30) and therefore (6.31) is equivalent to (6.25). The application of CS in OFDM to combat impulsive noise is not a new concept [36]. However, its performance and relationship with the

ML detection problem have not been discussed before.

Though we have highlighted computationally efficient ways of evaluating $\hat{\mathbf{x}}_{(1)}$ via (6.27) and (6.30), there is still the problem of detecting the transmitted OFDM symbol from (6.22). In the ML estimation case, we know that $\mathbf{e}$ is asymptotically normal with distribution (6.23). Therefore, $\bar{\mathbf{H}}\mathbf{e}$ should approximate a Gaussian vector with IID real and imaginary components for large N. If the L_1-norm minimization is employed, we see that $\bar{\mathbf{H}}\mathbf{e}$ is a near-Gaussian vector if $(N-K)/N$ is large enough. We therefore employ the Euclidean detector in (6.24) to compute BER in the next section.

6.4.3 Performance Analysis

Though ML joint-detection offers much better performance in OFDM over a single-carrier system, it is important to know how the CS or L_1-norm minimization problem compare. Do note that one can directly apply the joint-detection rule in (6.9) for small K as computational complexity is low. Therefore, we only test the proposed estimation approaches when K is sufficiently large.

In Fig. 6.6, we present the BER performance for $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$ with real and imaginary components for $\alpha = 1.5$ and $N = 32$. The results are plotted for varying null carriers. The percentage of nulls is given by $\frac{N-K}{N} \times 100$. The L_p-norm estimation scheme in (6.27) with $p = 1$ was employed with Euclidean detection. Our simulations further revealed that estimation with the L_1-norm and the CS approach offers *almost similar* performance over a variety of K and N combinations. For clarity, we plot only one BER curve instead of two for each

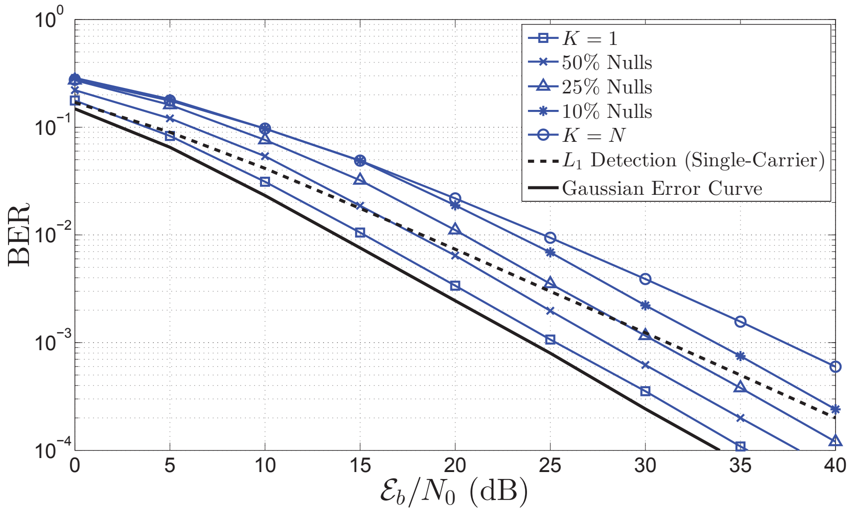


Figure 6.6: L_1-norm BER performance for BPSK-OFDM averaged over $\bar{\mathbf{H}}$ for $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$ and $\alpha = 1.5$. The curves are generated for $N = 32$.

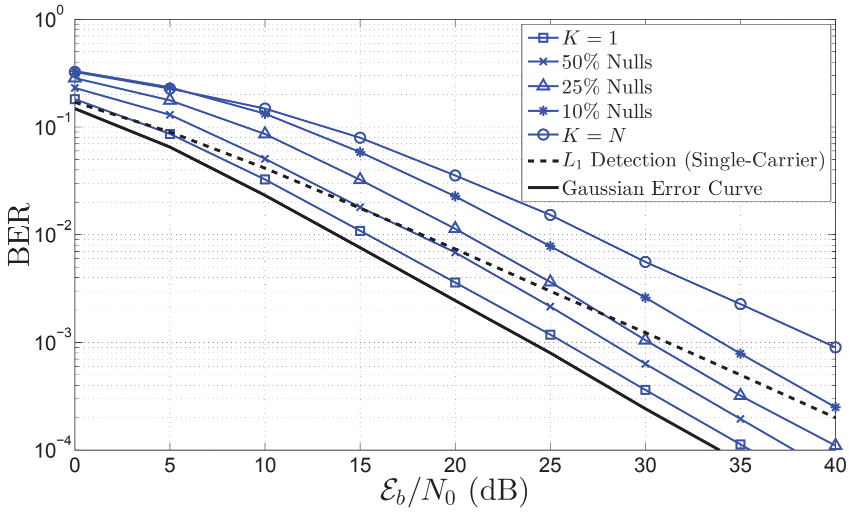


Figure 6.7: L_1-norm BER performance for BPSK-OFDM averaged over $\bar{\mathbf{H}}$ for $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$ and $\alpha = 1.5$. The curves are generated for $N = 256$.

combination of K and N. As anticipated, the detection performance improves with the number of null carriers. Using the same approach we also plot for $N = 256$ and $N = 512$ in Figs. 6.7 and 6.8, respectively, for $\alpha = 1.5$.

To see the range in which the BER results lie, we plot the best ($K = 1$) and

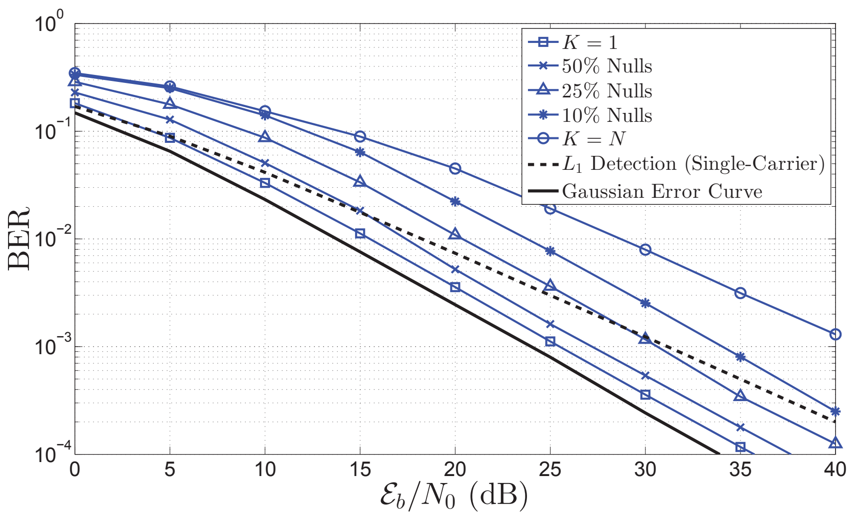


Figure 6.8: L_1-norm BER performance for BPSK-OFDM averaged over $\bar{\mathbf{H}}$ for $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$ and $\alpha = 1.5$. The curves are generated for $N = 512$.

the worst ($K = N$) cases as well. The $K = N$ scenario implies invoking the Euclidean detection rule in (6.11), while $K = 1$ corresponds to a single-carrier system with N samples. In all plots, we observe that the $K = N$ case worsens as N increases. Also, for 10%, 25% and 50% of null carriers, the BER remains the almost the same irrespective of N. For comparison, we have plotted the Gaussian error curve and the BER of a single-carrier system ($K = N = 1$) under L_1 detection in all figures.

Though the CS approach makes decoding the OFDM signal feasible, one can see that a certain number of nulls are required for the system to outperform its single-carrier counterpart. On the contrary, as seen in Fig. 6.1, ML detection outperforms a single-carrier system even for $K = N$. The trends seen for the $\alpha = 1.5$ and $\alpha = 1$ case may be extended intuitively to any $\alpha \neq 2$ as $\mathbf{z}$ is sparse in such scenarios.

6.5 Receiver Characteristics

Till now we have analyzed computationally-efficient techniques for robust detection of baseband OFDM signals in non-Gaussian AWSαSN. In this section, we highlight the design constraints that need to be considered to ensure $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$. We also show that linear passband-to-baseband conversion is actually *sub-optimal* in impulsive noise and reduces the operational SNR of the system. We propose a way around this in Section 6.6.

6.5.1 Passband-to-Baseband Conversion

The *continuous-time passband* transmit-receive equation for an OFDM signal is given by

$$r(t) = s(t) + w(t) \tag{6.32}$$

where $r(t)$, $s(t)$ and $w(t)$ are the received signal, the passband OFDM signal and a real AWSαSN process, respectively. The relationship of $s(t)$ with its baseband counterpart $\tilde{s}(t)$ is

$$s(t) = \Re\left\{\tilde{s}(t)e^{j2\pi f_c t}\right\} \tag{6.33}$$

where

$$\tilde{s}(t) = \sum_{k=-N/2}^{N/2-1} \sqrt{\frac{2}{T}} h_k x_k e^{\frac{j2\pi kt}{T}}, \tag{6.34}$$

T is the time period of the N-carrier OFDM symbol block and f_c is the carrier frequency. For simplicity of notation, N is assumed to be even. In (6.2), x_k is the symbol mapped onto the k^{th} sub-carrier $\forall\ k \in \{0, 1, \ldots, N-1\}$. To make this definition consistent with that in (6.34), we define $x_k = x_l$ and $h_k = h_l$ if $k \equiv l \pmod{N}\ \forall\ k, l \in \mathbb{Z}$, i.e., x_k and h_k are periodic in k with period N. As before, we assume K data carriers and $N - K$ nulls. Also, the passband signal energy per-symbol $\mathcal{E}_s$ is related to $\mathcal{E}_x$ as follows

$$\begin{aligned}
\mathcal{E}_s = \frac{1}{K}\int_0^T E[|s(t)|^2]\mathrm{d}t &= \frac{1}{2K}\int_0^T E[|\tilde{s}(t)|^2]\mathrm{d}t \\
&= \frac{1}{K}\sum_{k=-N/2}^{N/2-1} E[|h_k|^2]E[|x_k|^2] \\
&= \mathcal{E}_x\sigma_h^2. \qquad (6.35)
\end{aligned}$$

The transmitted signal in (6.2) can be obtained by scaling and sampling $\tilde{s}(t)$. From the properties of the IDFT, we have

$$\begin{aligned}
\mathbf{a}_n^{\mathsf{H}}\mathbf{H}\mathbf{x} &= \sum_{k=0}^{N-1} \frac{1}{\sqrt{N}} h_k x_k e^{\frac{j2\pi kn}{N}} \\
&= \sum_{k=-N/2}^{N/2-1} \frac{1}{\sqrt{N}} h_k x_k e^{\frac{j2\pi kn}{N}} \\
&= \sqrt{\frac{T}{2N}}\tilde{s}(nT/N) \qquad (6.36)
\end{aligned}$$

$\forall\ n \in \{0, 1, \ldots, N-1\}$. To allow downsampling via an integer factor, T should be restricted to a multiple of N and this is therefore implicitly assumed.

To ensure undistorted passband transmission, f_c has to be greater than the largest *absolute* frequency component in (6.34). This is the sum of $N/(2T)$ (the largest $|k|/T$ in (6.34)) and a factor proportional to the bandwidth per-carrier $(1/T)$. For simplicity, we choose the factor to be equal to 1. Therefore,

$$f_c > \left(\frac{N}{2} + 1\right)\frac{1}{T}. \tag{6.37}$$

The nulls in $\mathbf{x}$ are typically placed at the ends of the index set $k \in \{-N/2, \ldots, N/2-1\}$ [46]. The bound in (6.37) can be relaxed depending on the number of nulls. However, it does guarantee undistorted transmission for all $K \leq N$.

From Section 4.2, to attain $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$, the passband signal needs be discretized. This is a difficult task to do as f_c is large in practical wireless systems. However, in scenarios such as underwater acoustic communications, operational values of f_c are much lower and therefore this approach is feasible [4], [82]. We briefly discuss the design constraints required to get $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{IID}}$ and extend the results to the $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{ISO}}$ case.

Denoting the passband sampling frequency by $f_s = \lambda/T$, where $\lambda \in \mathbb{Z}^+$, the discrete-time equation corresponding to (6.32) can be written as

$$r[n] = s[n] + w[n] \tag{6.38}$$

$\forall\ n \in \{0, 1, \ldots, \lambda - 1\}$, where $w[n] \stackrel{d}{=} W \sim \mathcal{S}(\alpha, \delta_w)$. The abridged square bracket notation to denote a discrete signal, i.e., $r[n] = r(n/f_s)$. We also assume that the Nyquist criterion is met. Mathematically, this is given by

$$f_s > 2\left(\frac{\frac{N}{2}-1}{T} + \frac{1}{T} + f_c\right) = 2f_c + \frac{N}{T}$$

$$\Rightarrow \lambda > 2f_cT + N. \tag{6.39}$$

The discretized versions of (6.33) and (6.34) are

$$s[n] = \Re\left\{\tilde{s}[n]e^{\frac{j2\pi f_c n}{f_s}}\right\} \tag{6.40}$$

and

$$\tilde{s}[n] = \sum_{k=-N/2}^{N/2-1} \sqrt{\frac{2f_s}{\lambda}} h_k x_k e^{\frac{j2\pi kn}{\lambda}}, \tag{6.41}$$

respectively. As discussed in Section 4.1, to get $\tilde{s}[n]$ from $s[n]$, one needs to multiply the latter with a complex exponential, scale by a factor of 2 and pass the result through a low-pass filter. Precisely,

$$\tilde{s}[n] = 2v[n] * \left(s[n]e^{-\frac{j2\pi f_c n}{f_s}}\right). \tag{6.42}$$

where $v[n]\ \forall\ n \in \{0, 1, \ldots, L-1\}$ is the L-tap impulse response of the lowpass filter and $*$ is the linear convolution operator. Only the in-band information is retained, therefore the effective frequency response of $v[n]$ lies in $[-\frac{N}{2\lambda}, \frac{N}{2\lambda}]$. From (6.36), we have

$$\mathbf{a}_n^{\mathsf{H}}\mathbf{H}\mathbf{x} = \sqrt{\frac{\lambda}{2f_s N}}\tilde{s}[\lambda n/N] \tag{6.43}$$

$\forall\ n \in \{0, 1, \ldots, N-1\}$. To allow downsampling by an integer factor, λ needs to be a multiple of N, i.e., $\gcd(N, \lambda) = N$. This is implicitly assumed.

6.5.2 Design Constraints

As the passband-to-baseband conversion process is a linear system, analogous to (6.41) and (6.43), we get

$$\tilde{w}[n] = 2v[n] * \left(w[n]e^{-\frac{j2\pi f_c n}{f_s}}\right) \tag{6.44}$$

$\forall\ n \in \{0, 1, \ldots, \lambda-1\}$ and

$$z_n = \sqrt{\frac{\lambda}{2f_s N}}\tilde{w}[\lambda n/N] \tag{6.45}$$

$\forall\ n \in \{0, 1, \ldots, N-1\}$, respectively. To ensure z_n has IID real and imaginary components, from (6.45), it is sufficient that $\tilde{w}[n]$ has IID real and imaginary components. For this to hold, the condition $f_s = 4f_c$ *must* be met for $\alpha \neq 2$. This has been discussed in Section 4.2.3. We can see this by substituting $f_s = 4f_c$ in (6.44) to get

$$\tilde{w}[n] = 2v[n] * \left(w[n]e^{-\frac{j\pi n}{2}}\right) \tag{6.46}$$

and observing that

$$\Re\{\tilde{w}[n]\} = 2v[n] * w[n]\cos(\pi n/2) \quad \text{and} \tag{6.47}$$

$$\Im\{\tilde{w}[n]\} = -2v[n] * w[n]\sin(\pi n/2). \tag{6.48}$$

As $\cos(\pi n/2)$ is non-zero only when $\sin(\pi n/2) = 0 \; \forall \; n \in \mathbb{Z}$ and vice-versa, we note that the real and imaginary components of $\tilde{w}[n]$ are generated from two dissimilar sample sets of $w[n]$. Therefore, $\Re\{\tilde{w}[n]\}$ and $\Im\{\tilde{w}[n]\}$ are mutually *independent* $\forall \; n \in \{0, 1, \ldots, N-1\}$.

In Section 4.2.3, an intuitive argument was given to explain that the expressions in (6.47) and (6.48) are *statistically identical* for all n. We provide a proof in Appendix A.4 and show that

$$\Re\{\tilde{w}[n]\} \stackrel{d}{=} \Im\{\tilde{w}[n] \stackrel{d}{=} \frac{2}{2^{1/\alpha}} W \left(\sum_{m=0}^{L-1} |v[m]|^{\alpha} \right)^{1/\alpha} \tag{6.49}$$

for $f_s = 4f_c$.

Though $f_s = 4f_c$ ensures that the real and imaginary parts of z_n are IID, it does not guarantee independence within the components of $\mathbf{z}$. From (6.45), a sufficient condition for this to hold is the mutual independence of $\tilde{w}[\lambda n/N] \; \forall \; n \in \{0, 1, \ldots, N-1\}$. As $w[n]$ are samples of an AWSαSN process, from (4.35) and (6.45) we see that the condition is satisfied by constraining L to

$$L \leq \lambda / \gcd(N, \lambda) = \lambda/N. \tag{6.50}$$

Do note that the filter order and its cutoff both depend on λ/N.

As discussed in Section 6.2, for $\mathbf{z} \stackrel{d}{=} \mathbf{z}_{\text{ISO}}$, the passband-to-baseband conversion needs to be performed in the continuous-time domain. Also, from the discussion in Section 4.2.3, the $f_s = 4f_c$ constraint does not apply here. However, to attain independence within the components of $\mathbf{z}$ the impulse response $v(t)$

needs to be limited to the time interval $t \in [0, T/N]$. To attain this result, we note that the passband-to-baseband conversion block is a linear system. Therefore, from (6.36)

$$z_n = \sqrt{\frac{T}{2N}}\tilde{w}(nT/N) \tag{6.51}$$

Analogous to the relationship between (6.45) and (6.50), the constraint on $v(t)$ follows directly from (6.51).

6.5.3 SNR Degradation

The passband-to-baseband process is *lossy* in impulsive noise. Even if $\mathbf{z} \overset{d}{=} \mathbf{z}_{\text{IID}}$, the process is still sub-optimal. This can be quantified as SNR degradation. We can evaluate the distribution of z_n from (3.7), (6.45) and (6.49):

$$\Re\{z_n\} \overset{d}{=} \Im\{z_n\} \overset{d}{=} \frac{1}{2^{1/\alpha}}\sqrt{\frac{2\lambda}{f_s N}} W \left(\sum_{m=0}^{L-1} |v[m]|^\alpha\right)^{1/\alpha} \tag{6.52}$$

$\forall\ n \in \{0, 1, \ldots, N-1\}$. From (3.7) and (6.52) we have

$$\delta_z = \frac{1}{2^{1/\alpha}}\sqrt{\frac{2\lambda}{f_s N}} \delta_w \left(\sum_{m=0}^{L-1} |v[m]|^\alpha\right)^{1/\alpha} \tag{6.53}$$

Let $V(f)$ be the discrete-time Fourier transform (DTFT) of $v[n]$. As $V(f)$ is (effectively) non-zero (with unit magnitude) only in the interval $f \in [-\frac{N}{2\lambda}, \frac{N}{2\lambda}]$, we have from Parserval's theorem:

$$\sum_{n=0}^{L-1} |v[n]|^2 = \int_{-1/2}^{1/2} |V(f)|^2 \mathrm{d}f \approx \frac{N}{\lambda}$$
$$\Rightarrow \left(\sum_{n=0}^{L-1} |v[n]|^2\right)^{1/2} \approx \sqrt{\frac{N}{\lambda}}. \tag{6.54}$$

This allows us to express (6.53) as

$$\delta_z \approx \frac{1}{\sqrt{f_s}} \delta_w \frac{2^{1/2} \left(\sum_{m=0}^{L-1} |v[m]|^\alpha\right)^{1/\alpha}}{2^{1/\alpha} \left(\sum_{m=0}^{L-1} |v[m]|^2\right)^{1/2}} \tag{6.55}$$

or from (A.23) and (A.25),

$$\delta_z \approx \frac{1}{\sqrt{f_s}} \delta_w \frac{\left(\sum_{m=0}^{\lfloor \frac{L-1}{2} \rfloor} |v[2m]|^\alpha\right)^{1/\alpha}}{\left(\sum_{m=0}^{\lfloor \frac{L-1}{2} \rfloor} |v[2m]|^2\right)^{1/2}}. \tag{6.56}$$

Defining $\tilde{\mathbf{v}} = [v_1, v_2, \ldots, v_{\lfloor \frac{L-1}{2} \rfloor + 1}]^\mathsf{T}$ such that $v_{m+1} = v[2m] \ \forall \ m \in \{0, 1, \ldots, \lfloor \frac{L-1}{2} \rfloor\}$, we have

$$\delta_z \approx \frac{1}{\sqrt{f_s}} \delta_w \frac{\|\tilde{\mathbf{v}}\|_\alpha}{\|\tilde{\mathbf{v}}\|_2}. \tag{6.57}$$

Substituting this back into (6.16), we get

$$\frac{\mathcal{E}_b}{N_0} = \frac{\mathcal{E}_x \sigma_h^2 f_s}{4\delta_w^2 \log_2 M} \times \frac{\|\tilde{\mathbf{v}}\|_2}{\|\tilde{\mathbf{v}}\|_\alpha} = \mathrm{SNR} \times \frac{\|\tilde{\mathbf{v}}\|_2}{\|\tilde{\mathbf{v}}\|_\alpha} \tag{6.58}$$

or in dB scale

$$10 \log_{10} \frac{\mathcal{E}_b}{N_0} = \text{SNR (dB)} - 20 \log_{10} \frac{\|\tilde{\mathbf{v}}\|_\alpha}{\|\tilde{\mathbf{v}}\|_2}. \tag{6.59}$$

As $4\delta_w^2/f_s = N_0$ for the Gaussian case, we term $4\delta_w^2/f_s$ as the *pseudo*-PSD of the passband AWSαSN process. For a given pseudo-PSD, as observed in (6.58), $\mathcal{E}_b/N_0$ depends on $v[n]$, α and L. As $\|\tilde{\mathbf{v}}\|_\alpha \geq \|\tilde{\mathbf{v}}\|$ for any $\tilde{\mathbf{v}} \in \mathbb{R}^{\lfloor \frac{L-1}{2} \rfloor}$, the latter term in (6.59) is always positive and therefore causes reduction in the actual SNR, i.e., SNR $\geq \mathcal{E}_b/N_0$. Thus, the linear passband-to-baseband conversion process actually reduces the true SNR. This is analogous to the findings in the single-carrier case leading to the expression in (5.85). To visualize this effect, let $v[n] = \frac{1}{L}\ \forall\, n \in \{0, 1, \ldots, L-1\}$, i.e., the low-pass filter computes the average of the samples that fall into the convolution window. This results in

$$20 \log_{10} \frac{\|\tilde{\mathbf{v}}\|_\alpha}{\|\tilde{\mathbf{v}}\|_2} = 10 \left(\frac{2}{\alpha} - 1 \right) \log_{10} \left\lfloor \frac{L+1}{2} \right\rfloor. \tag{6.60}$$

We see that the SNR degradation varies logarithmically with $\lfloor \frac{L+1}{2} \rfloor$ and linearly with $2/\alpha - 1$. Table 6.1, lists down outcomes of (6.60) for various α and L. Even for α close to 2, there is at least a loss of 1 dB. On a final note, we observe that for $\alpha = 2$, the latter term in (6.59) is equal to zero for any $\tilde{\mathbf{v}}$. This signifies that the SNR depends only on the signal and noise powers in AWGN [1].

On the lines of the discussion on (5.84) and (5.85), the SNR term in (6.59) is a measure of $\mathcal{E}_b/N_0$. We call it SNR to differentiate between the definition in (6.16).

TABLE 6.1: TABULATED VALUES FOR (6.60).

α \ L	20	40	100	200
1	10.0	13.0	17.0	20.0
1.2	6.7	8.7	11.3	13.3
1.4	4.3	5.6	7.3	8.6
1.6	2.5	3.3	4.2	5.0
1.8	1.1	1.4	1.9	2.2

6.6 Passband Estimation and Detection

Instead of conversion to baseband, we can estimate soft-values of $\mathbf{x}_{(1)}$ directly from the passband samples. By doing so, we can completely avoid the SNR loss in passband-to-baseband conversion. Further still, the constraints that induce sparsity in $\mathbf{z}$ (discussed in Section 6.5.2) do not need to be enforced.

We define $x_k = x_l$ and $h_k = h_l$ if $k \equiv l \pmod{\lambda}$, i.e., h_k and x_k are periodic in k with sample period λ. As before, there are K data-carriers but now with $\lambda - K$ null-carriers. We fix the data location set $\mathcal{L}_{\mathbf{x}}$ to $\{-K/2, \ldots, K/2-1\}$ and $f_c = \xi/T$ where $\xi \in \mathbb{Z}^+$. From (6.40), (6.41) and the properties of the IDFT

$$\begin{aligned} s[n] &= \Re\left\{\sqrt{\frac{2f_s}{\lambda}} \sum_{k=-K/2}^{K/2-1} h_k x_k e^{\frac{j2\pi(k+\xi)n}{\lambda}}\right\} \\ &= \Re\left\{\sqrt{\frac{2f_s}{\lambda}} \sum_{k=-\lambda/2}^{\lambda/2-1} h_k x_k e^{\frac{j2\pi(k+\xi)n}{\lambda}}\right\} \\ &= \Re\left\{\sqrt{\frac{2f_s}{\lambda}} \sum_{k=0}^{\lambda-1} h_k x_k e^{\frac{j2\pi(k+\xi)n}{\lambda}}\right\}. \end{aligned} \tag{6.61}$$

On applying a change of variables from $k+\xi$ to k in (6.61), we have

$$s[n] = \Re\left\{\sqrt{\frac{2f_s}{\lambda}}\sum_{k=0}^{\lambda-1} h_{k-\xi}x_{k-\xi}e^{\frac{j2\pi kn}{\lambda}}\right\}. \tag{6.62}$$

As x_k, h_k and $e^{\frac{j2\pi kn}{\lambda}}$ are periodic in k with period λ,

$$\begin{aligned}\sum_{k=0}^{\lambda-1} h^*_{k-\xi}x^*_{k-\xi}e^{-\frac{j2\pi kn}{\lambda}} &= \sum_{k=0}^{\lambda-1} h^*_{k-\xi}x^*_{k-\xi}e^{\frac{j2\pi(\lambda-k)n}{\lambda}} \\ &= \sum_{k=0}^{\lambda-1} h^*_{\lambda-k-\xi}x^*_{\lambda-k-\xi}e^{\frac{j2\pi kn}{\lambda}}. \end{aligned} \tag{6.63}$$

Using (6.63), we can express (6.62) as

$$\begin{aligned} s[n] &= \sqrt{\frac{f_s}{2\lambda}}\sum_{k=0}^{\lambda-1}\left(h_{k-\xi}x_{k-\xi}e^{\frac{j2\pi kn}{\lambda}} + h^*_{k-\xi}x^*_{k-\xi}e^{-\frac{j2\pi kn}{\lambda}}\right) \\ &= \sqrt{\frac{f_s}{2\lambda}}\sum_{k=0}^{\lambda-1}\left(h_{k-\xi}x_{k-\xi} + h^*_{\lambda-k-\xi}x^*_{\lambda-k-\xi}\right)e^{\frac{j2\pi kn}{\lambda}}. \end{aligned} \tag{6.64}$$

Let $\mathbf{D}$ be the anti-diagonal $\lambda \times \lambda$ matrix with all non-zero elements equal to one.

Defining the block diagonal matrix

$$\mathbf{H}_\lambda = \begin{bmatrix} \mathbf{0}_{\xi-\frac{K}{2}} & 0 & \cdots & & 0 \\ 0 & \bar{\mathbf{H}} & \cdots & & 0 \\ \vdots & & \mathbf{0}_{\lambda-2\xi-K+1} & & \vdots \\ & & & \mathbf{D}\bar{\mathbf{H}}^*\mathbf{D} & \\ 0 & & \cdots & & \mathbf{0}_{\xi-\frac{K}{2}-1} \end{bmatrix} \tag{6.65}$$

and the $\lambda \times 1$ vector

$$\mathbf{x}_\lambda = \begin{bmatrix} \mathbf{0}_{(\xi-\frac{K}{2})\times 1} \\ \mathbf{x}_{(1)} \\ \mathbf{0}_{(\lambda-2\xi-K+1)\times 1} \\ \mathbf{D}\mathbf{x}^*_{(1)} \\ \mathbf{0}_{(\xi-\frac{K}{2}-1)\times 1} \end{bmatrix}, \tag{6.66}$$

we can represent (6.64) in the following vector form:

$$\mathbf{s} = \sqrt{\frac{f_s}{2}}\mathbf{A}_\lambda^{\mathsf{H}}\mathbf{H}_\lambda\mathbf{x}_\lambda. \tag{6.67}$$

Here $\mathbf{A}_\lambda$ is the *unitary* λ-point DFT matrix. Though $\mathbf{H}_\lambda$ and $\mathbf{x}_\lambda$ are complex, do note that $\mathbf{s} \in \mathbb{R}^\lambda$. Finally, using (6.67) we can write (6.38) as

$$\mathbf{r} = \sqrt{\frac{f_s}{2}}\mathbf{A}_\lambda^{\mathsf{H}}\mathbf{H}_\lambda\mathbf{x}_\lambda + \mathbf{w}, \tag{6.68}$$

where $r[n]$ and $w[n]$ $\forall\ n \in \{1, 2, \ldots, \lambda - 1\}$ are the n^{th} elements of $\mathbf{r}$ and $\mathbf{w}$, respectively. The problem in (6.68) is similar to that in (6.2). The difference lies in the inherent structure of $\mathbf{H}_\lambda$ and $\mathbf{x}_\lambda$. We also note that $\mathbf{r}, \mathbf{w} \in \mathbb{R}^\lambda$. Denoting the elements of $\mathbf{x}_\lambda$ by the λ-tuple $[x_{\lambda_0}, x_{\lambda_1}, \ldots, x_{\lambda_{\lambda-1}}]^\mathsf{T}$, we plot $|x_{\lambda_k}|$ against k in Fig. 6.9 for added clarity. From Fig. 6.9, we see that the constraints

$$\xi > \frac{K}{2} + 1 \tag{6.69}$$

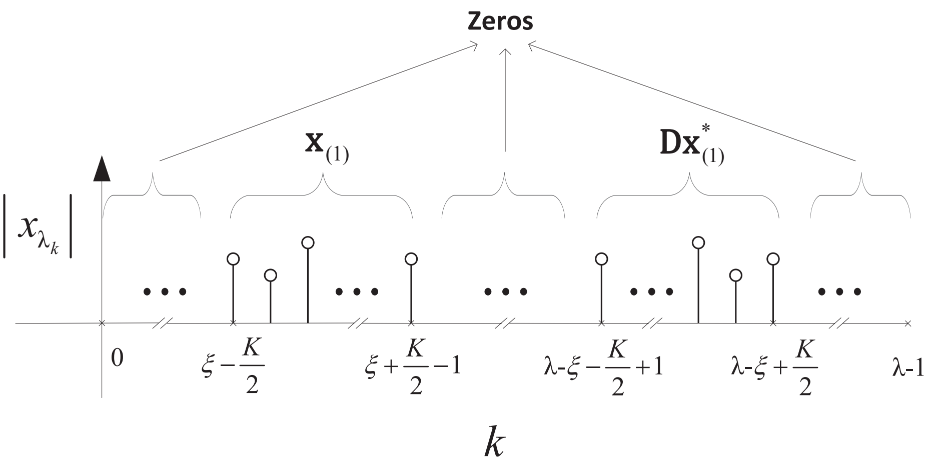


Figure 6.9: Placement of symbols and nulls in $\mathbf{x}_\lambda$.

and

$$\lambda > 2\xi + K, \tag{6.70}$$

ensure that the sidebands do not overlap and therefore need to be enforced to guarantee non-lossy transmission.

Analogous to (6.30), the CS estimate of $\mathbf{w}$ is given by

$$\begin{aligned} \hat{\mathbf{w}} = \quad & \underset{\boldsymbol{\gamma}\in\mathbb{R}^\lambda}{\texttt{arg}\,\texttt{min}} \quad \|\boldsymbol{\gamma}\|_1 \\ & \texttt{s.t} \qquad \bar{\bar{\mathbf{A}}}_\lambda \mathbf{r} = \bar{\bar{\mathbf{A}}}_\lambda \boldsymbol{\gamma}. \end{aligned} \tag{6.71}$$

where $\bar{\bar{\mathbf{A}}}_\lambda \in \mathbb{C}^{(\lambda-2K)\times\lambda}$ consists of the columns of $\mathbf{A}^*_\lambda$ corresponding to the locations of nulls in $\mathbf{x}_\lambda$. Given $\hat{\mathbf{w}}$, a modified passband equation may be constructed from (6.68)

$$\tilde{\mathbf{r}} = \sqrt{\frac{f_s}{2}}\mathbf{A}_\lambda^{\mathsf{H}}\mathbf{H}_\lambda\mathbf{x}_\lambda + (\mathbf{w} - \hat{\mathbf{w}}). \tag{6.72}$$

If $\lambda - 2K$ is greater than a certain threshold, the recovery of $\mathbf{w}$ via (6.71) will be good. Typically, $\lambda - 2K$ will be of large value. Following a similar line of reasoning as in Section 6.4.2, $\mathbf{w} - \hat{\mathbf{w}}$ can be approximated well by a Gaussian distribution with IID components. Thereafter, $\tilde{\mathbf{r}}$ may be passed through a linear passband-to-baseband conversion block to construct (6.2) with isotropic Gaussian $\mathbf{z}$. The ML detector in (6.11) may then be subsequently employed to generate hard-estimates of the transmitted symbols. Alternatively, (6.72) can be normalized by $\sqrt{\frac{f_s}{2}}$ and multiplied by $\mathbf{H}_\lambda^{-1}\mathbf{A}_\lambda$ from the left to form

$$\acute{\mathbf{r}} = \mathbf{H}_\lambda\mathbf{x}_\lambda + \acute{\mathbf{e}} \tag{6.73}$$

where $\acute{\mathbf{r}} = \sqrt{\frac{2}{f_s}}\mathbf{A}_\lambda\tilde{\mathbf{r}}$ and $\acute{\mathbf{e}} = \sqrt{\frac{2}{f_s}}\mathbf{A}_\lambda(\mathbf{w} - \hat{\mathbf{w}})$. Do note that $\acute{\mathbf{e}} \in \mathbb{C}^\lambda$ is approximately Gaussian as it is a linear transformation of $\mathbf{w} - \hat{\mathbf{w}}$. Further still, $\acute{\mathbf{e}}$ has IID components due to the orthonormal columns of $\mathbf{A}_\lambda$. Precisely, $\acute{\mathbf{e}} \sim \mathcal{CN}(\mathbf{0}_{\lambda\times 1}, 2\delta_{\acute{e}}^2\mathbf{I}_\lambda)$, where $2\delta_{\acute{e}}^2$ is the variance of each component of $\acute{\mathbf{e}}$. Therefore, we can remove the nulls and express (6.73) in terms of $\mathbf{x}_{(1)}$ and the corresponding observed and received components:

$$\begin{bmatrix} \acute{\mathbf{r}}_1 \\ \acute{\mathbf{r}}_2 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{H}}\mathbf{x}_{(1)} \\ \mathbf{D}\bar{\mathbf{H}}^*\mathbf{x}_{(1)}^* \end{bmatrix} + \begin{bmatrix} \acute{\mathbf{e}}_1 \\ \acute{\mathbf{e}}_2 \end{bmatrix}, \tag{6.74}$$

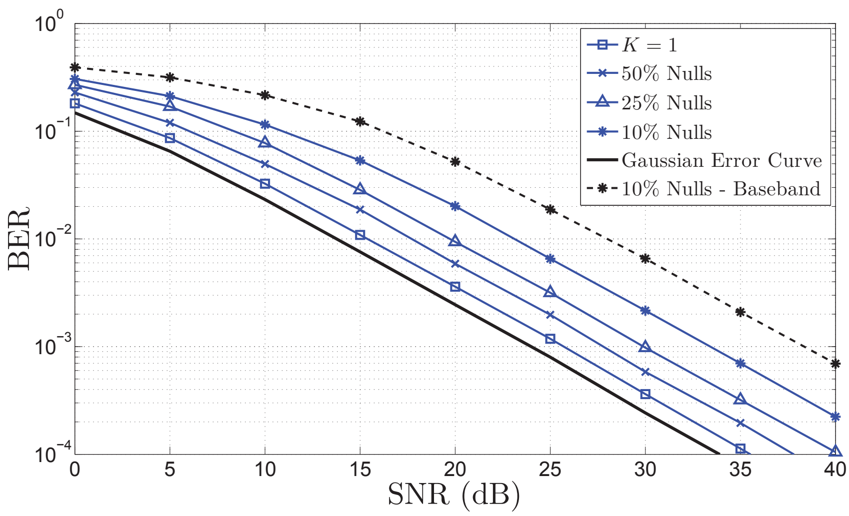


Figure 6.10: L_1-norm BER performance for BPSK-OFDM averaged over $\bar{\mathbf{H}}$ for $\alpha = 1.5$. The curves are generated for $\lambda = 256$ and decoding was performed directly on the passband samples.

where $\acute{\mathbf{r}}_1, \acute{\mathbf{r}}_2, \acute{\mathbf{e}}_1, \acute{\mathbf{e}}_2 \in \mathbb{C}^K$. Finally, the soft-estimate of $\mathbf{x}_{(1)}$ can be evaluated as

$$\begin{aligned}\hat{\mathbf{x}}_{(1)} &= \bar{\mathbf{H}}^{-1}\frac{\acute{\mathbf{r}}_1 + \mathbf{D}\acute{\mathbf{r}}_2^*}{2} \\ &= \mathbf{x}_{(1)} + \underbrace{\bar{\mathbf{H}}^{-1}\frac{\acute{\mathbf{e}}_1 + \mathbf{D}\acute{\mathbf{e}}_2^*}{2}}_{\mathbf{e}}\end{aligned} \tag{6.75}$$

where $\mathbf{e} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, 2\delta_{\acute{e}}^2(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})^{-1})$. This will be followed by the ML detection rule in (6.24).

In contrast to the baseband approach, there have been no assumptions about the relationship between the f_c and f_s. Further still, by performing operations in the passband, there is no SNR degradation due to linear passband-to-baseband conversion. Also, as λ is typically greater than $2K$, the CS algorithm will have more samples to work with and therefore $\mathbf{w}$ will be a better estimate. On the downside, passband sampling is difficult to perform when f_c is large. This is augmented by the fact that as λ increases, the DFT and CS operations increase in complexity as well. To sum up our discussion, we present the BER performance of BPSK-OFDM for $N = 256$ with varying nulls in Fig. 6.10. The trends are similar to those encountered in Fig. 6.7. If the measure in (6.59) is used, then the BER performance may be *increased* arbitrarily over its baseband counterpart by varying $\tilde{\mathbf{v}}$. Therefore, we plot against

$$\text{SNR (dB)} = 10 \log_{10} \frac{\mathcal{E}_x \sigma_h^2 f_s}{4\delta_w^2 \log_2 M}. \tag{6.76}$$

This measure is analogous to (5.87) in the single-carrier case. To highlight the increase in performance over a system that employs baseband conversion, we also plot the BER for the latter in Fig. 6.10 for $\lambda = 256$ and 10% nulls for $\alpha = 1.5$. We employ a 40-tap low pass filter with impulse response $v[n] = \frac{1}{40} \; \forall \, n \in \{0, 1, \ldots, 39\}$. From (6.60), the loss in SNR due to baseband conversion is approximately 4.3 dB. The advantage of passband processing can be clearly appreciated in Fig. 6.10.

6.7 Summary

We have investigated the performance of ML detection for uncoded OFDM in passband AWSαSN. It was assumed that the passband-to-baseband conversion scheme maintained independence between the components of the baseband noise vector. One novel result is that the error performance improves substantially by increasing the number of carriers in an OFDM system. This is depicted for channels with zero-Doppler and Rayleigh block fading. When the number of carriers is small, ML-detection can be performed with low computational cost.

This offers much better error performance over a single-carrier system. In a pure AWSαSN scenario, the rotation angle of the constellation has a significant influence in the single-carrier case. It is shown that the optimal rotation angles (per-carrier) need not to be evaluated for large carrier schemes. In fact, any *random* rotation may offer near-optimal error performance with high probability.

One significant disadvantage of ML detection is that the computational cost grows exponentially with the number of carriers. In recent years, CS theory has garnered much attention. One of its potential applications lies in the estimation and removal of impulses within an impulsive noise process as the latter is sparse. The relationship between the CS approach with the ML-estimation/detection problem has been discussed. We highlighted the pros and cons of this approach for OFDM in AWSαSN and compared the BER performance with the ML detection results. The constraints within a linear passband-to-baseband conversion block that guarantee sparsity at the baseband level has also been discussed. As in the single-carrier case, linear baseband conversion is shown to be a lossy process in AWSαSN and causes SNR degradation. This can be avoided by directly invoking a noise cancellation approach on the passband noise samples.

Chapter 7

Conclusions & Future Research

7.1 Conclusions

In some practical scenarios, impulsive noise is the dominant noise process in the available transmission spectrum. Gaussian-based techniques perform poorly in such environments. This is primarily due to the fact that impulsive noise distributions are heavy-tailed.

In our work, we have modeled the passband impulsive noise process as AWSαSN. If passed through a linear passband-to-baseband conversion block, the resultant complex noise samples have been shown to take a number of anisotropic but symmetric star-shaped bivariate SαS distributions. The exact statistics depend on the system parameters. Conventionally, passband-to-baseband conversion is performed in continuous-time and is optimized for AWGN. This results in decoding the I and Q components of the passband signal separately. Within this framework, we proposed a sampling rule that ensures independence between the I and Q channels for AWSαSN for all α. The resulting baseband distribution was shown to be anisotropic and offered the best error performance over all possible bivariate SαS noise configurations.

To harness the true potential of the aforementioned scheme, we have proposed new constellations and a corresponding design methodology. The advantage of the resulting system over a conventional continuous-time implementation of the passband-to-baseband conversion block is significant. With the exception of the Gaussian and Cauchy cases, the PDF of an SαS distribution does not exist in an analytic form. Though the PDF may be evaluated via efficient numerical techniques, a closed-form approximation is required in some instances. For a single-carrier system employing optimized constellations, we analyzed various closed-form detectors and compared them to ML detection. If the linearity in passband-to-baseband conversion is sacrificed, the error performance of a single-carrier system was shown to improve further. This was highlighted by evaluating soft-estimates of the transmitted symbol from the passband samples via various non-linear estimation schemes. Alternatively, one may perform joint-detection directly on the passband samples. Though more computationally intensive due to the high-frequency sampling, this provides a Linear passband-to-baseband conversion is found to be a lossy process. We have shown that there is significant SNR reduction in this case.

We have further presented ML detection results for a baseband OFDM system contaminated with SαS noise. Depending on the passband-to-baseband conversion process, the complex baseband noise vector can take up a number of statistical configurations. By ensuring the components of the vector are independent, the sparsity of the passband AWSαSN process is retained at the baseband level. We have proposed important rules within a linear passband-to-baseband process that ensure the noise vector is sparse. For this case, we have shown that ML detection performance improves as the number of carriers increases for a channel with zero-Doppler and Rayleigh block fading.

This has been observed even if all carriers are reserved for data. The advantage over a single-carrier system is apparent, even when the number of carriers is small. Further still, the dependence on the optimal constellation reduces as the number of carriers increases.

ML detection is feasible to implement when the number of carriers is small. However, the computational cost grows exponentially with the number of carriers. Therefore, one has to revert to sub-optimal techniques for large OFDM systems. Instead of employing joint-detection, we have evaluated soft-estimates of the transmitted OFDM symbol block followed by carrier-wise detection. We have discussed estimators under the framework of M-estimation theory and have shown CS estimation as a special case. Similar to the single-carrier case, linear passband-to-baseband conversion is shown to incur SNR reduction at the receiver. This can be avoided using an estimation plus detection scheme directly on the passband samples.

7.2 Future Research

One of the main challenges in engineering is to bring theoretical work to production. Though the motivation of using AWSαSN stems from the good statistical fit it provides to practical impulsive noise data, the schemes proposed in this thesis need to be tested in practical systems operating in impulsive noise. In Section 5.7, we briefly discussed a practical implementation for good PSK maps in impulsive noise. The proposed scheme rotated the constellations at the receiver by introducing a phase delay between the clocks at the transmitter and receiver. Likewise, future research could focus on the implementability of the schemes proposed in this thesis. This would be followed by experimental trials in impulsive noise environments.

The schemes that we have analyzed are all single-input single-output systems. In recent years, MIMO schemes have garnered much attention. This is due to the added performance they offer in terms of diversity gain (robustness against fading) and space-time coding. In the literature, MIMO systems are conventionally analyzed for Gaussian noise. Performance analysis is done in terms of second-order moments of the noise samples, which are infinite in non-Gaussian stable models. In Chapter 6, we have discussed OFDM performance in impulsive noise. Due to the correlation within the noise samples affecting an OFDM symbol block, ML detection is able to enhance error performance relative to its single-carrier counterpart. It would be interesting to see how both mechanisms could be combined to provide robustness against impulsive noise and fading.

Another logical step is to develop and analyze receivers operating in time-varying channels in the presence of impulsive noise. Such a channel distorts the transmitted signal by introducing a Doppler shift/spread [1], [2]. To compensate for this, the channel needs to be estimated periodically at the receiver. As highlighted many times in this thesis, techniques that have been optimized for Gaussian noise models perform poorly in the presence of impulsive noise. The channel estimates will be error-prone unless the receiver is specifically designed for such scenarios. Therefore, new robust mechanisms need to be employed at the receiver. The analysis and techniques discussed in this thesis will be helpful to explore this area.

Till now we have considered only uncoded schemes in this thesis. In the current literature, there is a lot of information on error correction codes. However, research on error control coding for impulsive noise channels is still in its initial phase. In OFDM, the DFT operation spreads a corrupting impulse across the carriers. This results in a colored noise vector that is heavy-tailed. A potential future direction could focus on developing error control codes that would take advantage of the correlation between the noise components. The concepts developed in this thesis would be fruitful in this regard.

Appendix A

Various Proofs

A.1 Noise Scale Parameters in the Conventional Receiver

Due to the linearity of the receiver, $z_I \stackrel{d}{=} z_Q \sim \mathcal{S}(\alpha, \delta_z)$ in AWSαSN. From (5.18), we have

$$z_I \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \int_0^{T/\xi} w(t) \cos(2\pi f_c t) \mathrm{d}t \tag{A.1}$$

By approximating the integration term with a limiting Riemann sum, we get

$$\begin{aligned} z_I &\stackrel{d}{=} \frac{1}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \sum_{n=0}^{\lfloor T f_s/\xi \rfloor - 1} w(n/f_s) \cos(2\pi f_c/f_s n) \\ &= \sum_{n=0}^{\lfloor T f_s/\xi \rfloor - 1} \left(\frac{1}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \cos(2\pi f_c/f_s n) \right) w(n/f_s) \end{aligned} \tag{A.2}$$

as $f_s \to +\infty$. In this formulation, f_s is the passband sampling frequency of the AWSαSN channel. As $w(n/f_s) \stackrel{d}{=} W \sim \mathcal{S}(\alpha, \delta_w) \ \forall \ n \in \{0, 1, \ldots, \lfloor T f_s/\xi \rfloor - 1\}$, then using (3.6), (3.7) and (A.2), we can express z_I as

$$z_I \stackrel{d}{=} \left(\sum_{n=0}^{\lfloor T f_s/\xi \rfloor - 1} \left| \frac{1}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \cos(2\pi f_c/f_s n) \right|^{\alpha} \right)^{1/\alpha} W \tag{A.3}$$

The scale parameter of z_I can be evaluated from (A.3) (3.7) and (A.2),

$$\begin{aligned}\delta_z &= \delta_w \left(\sum_{n=0}^{\lfloor Tf_s/\xi \rfloor - 1} \left| \frac{1}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \cos(2\pi f_c/f_s n) \right|^{\alpha} \right)^{1/\alpha} \\ &= \frac{\delta_w}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \left(\sum_{n=0}^{\lfloor Tf_s/\xi \rfloor - 1} |\cos(2\pi f_c/f_s n)|^{\alpha} \right)^{\frac{1}{\alpha}} \\ &\approx \frac{\delta_w}{f_s^{(1-\frac{1}{\alpha})}} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \left(\int_0^{\frac{T}{\xi}} |\cos(2\pi f_c t)|^{\alpha} \mathrm{d}t \right)^{\frac{1}{\alpha}} \end{aligned} \tag{A.4}$$

The approximation in (A.4) is justified for $f_s \to +\infty$. As

$$\int_0^{\frac{T}{\xi}} |\cos(2\pi f_c t)|^{\alpha} \mathrm{d}t = \frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})} \tag{A.5}$$

and $c(\alpha, \xi, g(t))$ is given in (5.17), we may express (A.4) as

$$\delta_z = \frac{\delta_w}{f_s^{(1-\frac{1}{\alpha})}} \sqrt{\frac{2}{\mathcal{E}_g}} \left(\int_0^T |g(t)|^{\alpha} \, \mathrm{d}t \right)^{\frac{1}{\alpha}} \left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})} \right)^{\frac{1}{\alpha}} \tag{A.6}$$

Using the same approach one can easily evaluate (A.6) from z_Q instead of z_I in (A.1).

A.2 SNR Derivation for the Conventional Receiver

As the reference SNR is defined in (5.76) for the discretized linear receiver, we need to express it in terms of δ_z in the conventional receiver. We slightly abuse notation by equating (A.6) to δ_c. We reserve δ_z for the baseband scale parameter in the discretized linear receiver. As ξ is assumed to be large, we may express

δ_z in (5.41) as

$$\begin{aligned}\delta_z &= \frac{\delta_w}{f_s}\left(\sum_{n=0}^{4\xi-1}|\ell_I[n]|^\alpha\right)^{\frac{1}{\alpha}} = \frac{\delta_w}{f_s}\sqrt{\frac{2}{\mathcal{E}_g}}\left(\sum_{n=0}^{2\xi-1}|g[2n]|^\alpha\right)^{\frac{1}{\alpha}} \\ &\approx \frac{\delta_w}{f_s}\sqrt{\frac{2}{\mathcal{E}_g}}\left(\frac{f_s}{2}\int_0^T|g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha} \\ &= \frac{\delta_w}{f_s^{(1-\frac{1}{\alpha})}}\frac{2^{(\frac{1}{2}-\frac{1}{\alpha})}}{\sqrt{\mathcal{E}_g}}\left(\int_0^T|g(t)|^\alpha \mathrm{d}t\right)^{1/\alpha}. \end{aligned} \tag{A.7}$$

On dividing (A.7) by (A.6) and simplifying, we get

$$\delta_z = \delta_c 2^{-\frac{1}{\alpha}}\left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})}\right)^{-1/\alpha}. \tag{A.8}$$

Finally, we substitute (A.8) in (5.76)

$$\frac{\mathcal{E}_b}{N_0} = \frac{E[\mathcal{E}_{s_i}]}{4\delta_c^2\log_2 M}\times 2^{\frac{2}{\alpha}}\left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})}\right)^{2/\alpha}. \tag{A.9}$$

A.3 Asymptotic Normality of e

Let us define

$$\breve{\mathbf{y}} = \begin{bmatrix}\Re\{\mathbf{y}\}\\ \Im\{\mathbf{y}\}\end{bmatrix}, \breve{\mathbf{x}} = \begin{bmatrix}\Re\{\mathbf{x}_{(1)}\}\\ \Im\{\mathbf{x}_{(1)}\}\end{bmatrix} \text{ and } \breve{\mathbf{e}} = \begin{bmatrix}\Re\{\mathbf{e}\}\\ \Im\{\mathbf{e}\}\end{bmatrix} \tag{A.10}$$

We can express (6.3) in terms of $\breve{\mathbf{y}}$ and $\breve{\mathbf{x}}$

$$\breve{\mathbf{y}} = \breve{\mathbf{A}}^{\mathsf{T}}\breve{\mathbf{H}}\breve{\mathbf{x}} + \breve{\mathbf{z}} \tag{A.11}$$

where

$$\breve{\mathbf{A}} = \begin{bmatrix} \Re\{\bar{\mathbf{A}}\} & -\Im\{\bar{\mathbf{A}}\} \\ \Im\{\bar{\mathbf{A}}\} & \Re\{\bar{\mathbf{A}}\} \end{bmatrix} \text{ and}$$

$$\breve{\mathbf{H}} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}\} & -\Im\{\bar{\mathbf{H}}\} \\ \Im\{\bar{\mathbf{H}}\} & \Re\{\bar{\mathbf{H}}\} \end{bmatrix}.$$

From the asymptotic normality property of ML estimation,

$$\breve{\mathbf{e}} \sim \mathcal{N}(\mathbf{0}_{2N\times 1}, \mathbf{\Sigma}^{-1}), \tag{A.12}$$

as $N \to \infty$, where $\mathbf{\Sigma}$ is the Fisher information matrix of $\breve{\mathbf{x}}$ with respect to the distribution $\tilde{f}_{\mathbf{z}}(\mathbf{y};\breve{\mathbf{x}}) = f_{\mathbf{z}}(\mathbf{y} - \bar{\mathbf{A}}^{\mathsf{H}}\bar{\mathbf{H}}\mathbf{x}_{(1)})$. Further still, as the model in (6.3) is that of linear regression, we have from [100], [102], Eq. 58

$$\begin{aligned} \mathbf{\Sigma} &= \frac{\mathcal{I}^{(0)}}{\delta_z^2}(\breve{\mathbf{A}}^{\mathsf{T}}\breve{\mathbf{H}})^{\mathsf{T}}\breve{\mathbf{A}}^{\mathsf{T}}\breve{\mathbf{H}} \\ &= \frac{\mathcal{I}^{(0)}}{\delta_z^2}\breve{\mathbf{H}}^{\mathsf{T}}\breve{\mathbf{A}}\breve{\mathbf{A}}^{\mathsf{T}}\breve{\mathbf{H}} = \frac{\mathcal{I}^{(0)}}{\delta_z^2}\breve{\mathbf{H}}^{\mathsf{T}}\breve{\mathbf{H}}, \end{aligned} \tag{A.13}$$

where $\mathcal{I}^{(0)}$ is the Fisher information of the location parameter provided by *one* real noise sample with distribution $\mathcal{S}(\alpha, 1)$ [100]. On substituting (A.13) into (A.12), we have

$$\breve{\mathbf{e}} \sim \mathcal{N}(\mathbf{0}_{2N\times 1}, \frac{\delta_z^2}{\mathcal{I}^{(0)}}(\breve{\mathbf{H}}^{\mathsf{T}}\breve{\mathbf{H}})^{-1}) \tag{A.14}$$

As

$$\breve{\mathbf{H}}^{\mathsf{T}}\breve{\mathbf{H}} = \begin{bmatrix} \bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} \end{bmatrix} \tag{A.15}$$

is a diagonal matrix, we can clearly see that the elements of $\breve{\mathbf{e}}$ are independent. Finally, taking advantage of the form in (A.15) and the fact that $\mathbf{e} = [\mathbf{I}_N \;\; j\mathbf{I}_N]\breve{\mathbf{e}}$, we have

$$\mathbf{e} \sim \mathcal{CN}(\mathbf{0}_{N\times 1}, \frac{2\delta_z^2}{\mathcal{I}^{(0)}}(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})^{-1}). \tag{A.16}$$

A.4 $\Re\{\tilde{w}[n]\}$ and $\Im\{\tilde{w}[n]\}$ are Statistically Identical for all n

The convolution operation in (6.47) can be written in its true form:

$$\Re\{\tilde{w}[n]\} = 2\sum_{l=0}^{L-1} v[l]w[n-l]\cos(\pi(n-l)/2) \tag{A.17}$$

$$= \sum_{l=0}^{L-1}\Big(2v[l]\cos(\pi(n-l)/2)\Big)w[n-l] \tag{A.18}$$

As $w[n] \overset{d}{=} W \sim \mathcal{S}(\alpha, \delta_w)$, we can use (3.6) to express (A.18) as

$$\Re\{\tilde{w}[n]\} \overset{d}{=} W\left(\sum_{l=0}^{L-1}|2v[l]\cos(\pi(n-l)/2)|^{\alpha}\right)^{1/\alpha} \tag{A.19}$$

$$\overset{d}{=} 2W\left(\sum_{l=0}^{L-1}|v[l]\cos(\pi(n-l)/2)|^{\alpha}\right)^{1/\alpha} \tag{A.20}$$

We know that $\cos(\pi(n-l)/2)$ is non-zero only for $l = 2m$ when n is even and $l = 2m+1$ when n is odd, where $m \in \mathbb{Z}$. Further still, the result will lie in the set $\{-1, +1\}$. As symmetric distributions are not influenced by the sign, we have

$$\Re\{\tilde{w}[n]\} \stackrel{d}{=} 2W\left(\sum_{m=0}^{\lfloor\frac{L-1}{2}\rfloor}|v[2m]|^{\alpha}\right)^{1/\alpha} \tag{A.21}$$

when n is even and

$$\Re\{\tilde{w}[n]\} \stackrel{d}{=} 2W\left(\sum_{m=0}^{\lfloor\frac{L}{2}\rfloor-1}|v[2m+1]|^{\alpha}\right)^{1/\alpha} \tag{A.22}$$

when n is odd. The expressions in (A.21) and (A.22) depend on the sums of the even and odd samples of $|v[n]|^{\alpha}$, respectively. We know that $v[n]$ is *effectively* band-limited to $[-\frac{N}{2\lambda}, \frac{N}{2\lambda}]$. Denoting the discrete-time Fourier transform (DTFT) of $|v[n]|^{\alpha}$ by $V_{\alpha}(f)$, we note that $V_{\alpha}(f)$ still retains characteristics of a low-pass filter, i.e., *most* of the energy of $|v[n]|^{\alpha}$ occupies the lower spectrum for finite L [79]. From the properties of the DTFT,

$$\begin{aligned} V_{\alpha}(0) &= \sum_{m=0}^{L-1}|v[m]|^{\alpha} && \text{(A.23)}\\ &= \sum_{m=0}^{\lfloor\frac{L-1}{2}\rfloor}|v[2m]|^{\alpha} + \sum_{m=0}^{\lfloor\frac{L}{2}\rfloor-1}|v[2m+1]|^{\alpha}. && \text{(A.24)} \end{aligned}$$

If $V_{\alpha}(f/2)$ is *truly* band-limited, the energy is divided equally amongst the two summation terms in (A.24). Therefore,

$$\frac{1}{2}V_{\alpha}(0) = \sum_{m=0}^{\lfloor\frac{L-1}{2}\rfloor}|v[2m]|^{\alpha} = \sum_{m=0}^{\lfloor\frac{L}{2}\rfloor-1}|v[2m+1]|^{\alpha}. \tag{A.25}$$

In practical filters, L is finite and therefore $V_{\alpha}(f/2)$ is not truly band-limited. However, (A.25) provides a good approximation for a large range of L as long as λ is at least a few multiples of N. Therefore, from (A.23) and (A.25), we can express (A.21) and (A.22) as

$$\begin{aligned}\Re\{\tilde{w}[n]\} &\stackrel{d}{=} 2W \times \left(\frac{1}{2}V_\alpha(0)\right)^{1/\alpha} \\ &\stackrel{d}{=} \frac{2}{2^{1/\alpha}}W\left(\sum_{m=0}^{L-1}|v[m]|^\alpha\right)^{1/\alpha}. \end{aligned} \tag{A.26}$$

Using a similar approach as in (A.18)-(A.26) we can evaluate the distribution of $\Im\{\tilde{w}[n]\}$ and observe that

$$\Re\{\tilde{w}[n]\} \stackrel{d}{=} \Im\{\tilde{w}[n]\}. \tag{A.27}$$

We note from (A.26) and (A.27), that the distribution of $\tilde{w}[n]$ is independent of n and therefore time-invariant.

Bibliography

[1] J. Proakis and M. Salehi, *Digital Communications*, ser. McGraw-Hill higher education. McGraw-Hill Education, 2007.

[2] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[3] M. Chitre, J. Potter, and S.-H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," *Oceanic Engineering, IEEE Journal of*, vol. 31, no. 2, pp. 497 –503, April 2006.

[4] M. Chitre, J. Potter, and O. S. Heng, "Underwater acoustic channel characterisation for medium-range shallow water communications," in *OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04*, vol. 1, Nov. 2004, pp. 40 –45.

[5] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 44, no. 1, pp. 249 –258, Feb 2002.

[6] T. Al-Naffouri, A. Quadeer, and G. Caire, "Impulsive noise estimation and cancellation in dsl using orthogonal clustering," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, 2011, pp. 2841–2845.

[7] J. P. Nolan, *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhauser, 2012, in progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

[8] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, 1994.

[9] C. L. Nikias and M. Shao, *Signal processing with Alpha-Stable Distributions and Applications*. New York: Chapman-Hall, 1996.

[10] M. Chitre, J. Potter, and S. Ong, "Viterbi decoding of convolutional codes in symmetric α-stable noise," *Communications, IEEE Transactions on*, vol. 55, no. 12, pp. 2230 –2233, Dec. 2007.

[11] A. Papoulis and U. S. Pillai, *Probability, Random Variables and Stochastic Processes*. Boston: McGraw-Hill, Dec 2001.

[12] M. Legg, A. Zaknich, A. Duncan, and M. Greening, "Analysis of impulsive biological noise due to snapping shrimp as a point process in time," in *OCEANS 2007 - Europe*, 2007, pp. 1–6.

[13] M. Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *The Journal of the Acoustical Society of America*, vol. 122, no. 5, pp. 2580–2586, 2007.

[14] M. W. Legg, *Non-Gaussian and non-homogeneous Poisson models of snapping shrimp noise*. Curtin University of Technology., 2009.

[15] B. G. Ferguson and J. L. Cleary, "In Situ Source Level and Source Position Estimates of Biological Transient Signals Produced by Snapping Shrimp in an Underwater Environment," *Journal of the Acoustical Society of America*, vol. 109, p. 3031, 2001.

[16] M. Nassar, K. Gulati, Y. Mortazavi, and B. Evans, "Statistical modeling of asynchronous impulsive noise in powerline communication networks," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, 2011, pp. 1–6.

[17] J. Lin, M. Nassar, and B. Evans, "Impulsive noise mitigation in powerline communications using sparse bayesian learning," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 7, pp. 1172–1183, 2013.

[18] F. H. Gregorio, J. E. Cousseau, and J. L. Figueroa, "Reducing impulsive noise in dsl systems - robustness and delay," in *Signal Processing Symposium, 2004. NORSIG 2004. Proceedings of the 6th Nordic*, 2004, pp. 224–227.

[19] D. Toumpakaris, W. Yu, J. Cioffi, D. Gardan, and M. Ouzzif, "A simple byte-erasure method for improved impulse immunity in dsl," in *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 4, 2003, pp. 2426–2430 vol.4.

[20] N. Nedev and S. McLaughlin, *Analysis of the Impact of Impulse Noise in Digital Subscriber Line Systems.* University of Edinburgh, 2003.

[21] T. Al-Naffouri, A. Quadeer, F. Al-Shaalan, and H. Hmida, "Impulsive noise estimation and cancellation in dsl using compressive sampling," in *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, 2011, pp. 2133–2136.

[22] X. Guo, H. Sun, and Y. Lu, "Joint suppression of radio frequency interference and lightning impulsive noise in hfswr," in *Radar Conference - Surveillance for a Safer World, 2009. RADAR. International*, 2009, pp. 1–6.

[23] F. Horner and J. Harwood, "An investigation of atmospheric radio noise at very low frequencies," *Proceedings of the IEE - Part B: Radio and Electronic Engineering*, vol. 103, no. 12, pp. 743–751, 1956.

[24] L. Ilzzo and L. Paura, "Error probability for fading cpsk signals in gaussian and impulsive atmospheric noise environments," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-17, no. 5, pp. 719–722, 1981.

[25] V. Jain and S. Gupta, "Optimum receiver for impulsive atmospheric radio noise," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 128, no. 4, pp. 189–192, 1981.

[26] K. T. Beng, T. E. Teck, M. Chitre, and J. Potter, "Estimating the spatial and temporal distribution of snapping shrimp using a portable, broadband 3-dimensional acoustic array," in *OCEANS 2003. Proceedings*, vol. 5, Sept 2003, pp. 2706–2713 Vol.5.

[27] D. Middleton and I. for Telecommunication Sciences, *Statistical-physical Models of Man-made and Natural Radio Noise*, ser. OT report. U.S. Department of Commerce, Office of Telecommunications, 1974, no. pt. 2.

[28] D. Middleton, "Statistical-physical models of electromagnetic interference," *Electromagnetic Compatibility, IEEE Transactions on*, vol. EMC-19, no. 3, pp. 106–127, 1977.

[29] ——, “Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models,” *Information Theory, IEEE Transactions on*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[30] H. Ishikawa, M. Itami, and K. Itoh, “A study on adaptive modulation of ofdm under middleton’s class-a impulsive noise model,” in *Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on*, 2007, pp. 1–2.

[31] V. Balakirsky and A. J. H. Vinck, “Limits on transmission rates over channels received from middleton’s impulsive noise model,” in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, 2003, pp. 265–265.

[32] A. Swami, “Non-gaussian mixture models for detection and estimation in heavy-tailed noise,” in *Acoustics, Speech, and Signal Processing, 2000. ICASSP ’00. Proceedings. 2000 IEEE International Conference on*, vol. 6, 2000, pp. 3802 –3805 vol.6.

[33] X. Li, J. Sun, S. Wang, L. Fan, and L. Chen, “Near-optimal detection with constant false alarm ratio in varying impulsive interference,” *Signal Processing, IET*, vol. 7, no. 9, pp. 824–832, 2013.

[34] L. Fan, X. Li, X. Lei, W. Li, and F. Gao, “On distribution of sas noise and its application in performance analysis for linear rake receivers,” *Communications Letters, IEEE*, vol. 16, no. 2, pp. 186–189, 2012.

[35] M. Ghosh, “Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems,” *Communications, IEEE Transactions on*, vol. 44, no. 2, pp. 145–147, Feb. 1996.

[36] G. Caire, T. Al-Naffouri, and A. Narayanan, “Impulse noise cancellation in OFDM: an application of compressed sensing,” in *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, July, pp. 1293–1297.

[37] F. Abdelkefi, P. Duhamel, and F. Alberge, “Impulsive noise cancellation in multicarrier transmission,” *Communications, IEEE Transactions on*, vol. 53, no. 1, pp. 94–106, Jan. 2005.

[38] ——, “A necessary condition on the location of pilot tones for maximizing the correction capacity in OFDM systems,” *Communications, IEEE Transactions on*, vol. 55, no. 2, pp. 356–366, Feb. 2007.

[39] S. Banerjee and M. Agrawal, “Underwater acoustic noise with generalized gaussian statistics: Effects on error performance,” in *OCEANS - Bergen, 2013 MTS/IEEE*, 2013, pp. 1–8.

[40] H. Soury, F. Yilmaz, and M.-S. Alouini, “Error rates of m-pam and m-qam in generalized fading and generalized gaussian noise environments,” *Communications Letters, IEEE*, vol. 17, no. 10, pp. 1932–1935, 2013.

[41] ——, “Average bit error probability of binary coherent signaling over generalized fading channels subject to additive generalized gaussian noise,” *Communications Letters, IEEE*, vol. 16, no. 6, pp. 785–788, 2012.

[42] J. H. McCulloch, *Numerical approximation of the symmetric stable distribution and density.* Cambridge, MA, USA: Birkhauser Boston Inc., 1998, pp. 489–499.

[43] J. P. Nolan, "Numerical calculation of stable densities and distribution functions," *Communications in Statistics Stochastic Models*, vol. 13, no. 4, pp. 759–774, 1997.

[44] C. Menn and S. T. Rachev, "Calibrated fft-based density approximations for α-stable distributions," *Computational Statistics & Data Analysis*, vol. 50, no. 8, pp. 1891 – 1904, 2006.

[45] J. P. Nolan. [Online]. Available: http://academic2.american.edu/~jpnolan/stable/stable.html

[46] H. Schulze and C. Lueders, *Theory and Applications of OFDM and CDMA: Wideband Wireless Communications*. Wiley, 2005.

[47] S. Herath, N. Tran, and T. Le-Ngoc, "On optimal input distribution and capacity limit of bernoulli-gaussian impulsive noise channels," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 3429–3433.

[48] H. V. Vu, N. H. Tran, T. V. Nguyen, and S. Hariharan, "On the capacity of bernoulli-gaussian impulsive noise channels in rayleigh fading," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, 2013, pp. 1281–1285.

[49] J. Wang, E. Kuruoglu, and T. Zhou, "Alpha-stable channel capacity," *Communications Letters, IEEE*, vol. 15, no. 10, pp. 1107–1109, 2011.

[50] J. Nolan, "Maximum likelihood estimation and diagnostics for stable distributions," in *Lvy Processes*, O. Barndorff-Nielsen, S. Resnick, and T. Mikosch, Eds. Birkhuser Boston, 2001, pp. 379–400. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-0197-7_17

[51] M. Shao and C. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, 1993.

[52] X. Ma and C. Nikias, "On blind channel identification for impulsive signal environments," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 3, 1995, pp. 1992–1995 vol.3.

[53] V. Chavali and C. da Silva, "Detection of digital amplitude-phase modulated signals in symmetric alpha-stable noise," *Communications, IEEE Transactions on*, vol. 60, no. 11, pp. 3365–3375, 2012.

[54] C. Brown and A. Zoubir, "On the estimation of the parameters of alpha;-stable distributions using linear regression in the characteristic function domain," in *Statistical Signal and Array Processing, 1998. Proceedings., Ninth IEEE SP Workshop on*, 1998, pp. 423–426.

[55] X. Guangrong and L. Xingzhao, "The improved algorithms to estimate alpha; of alpha stable distribution based on empirical characteristic function," in *Signal Processing, 2002 6th International Conference on*, vol. 1, 2002, pp. 182–185 vol.1.

[56] J. Park, G. Shevlyakov, and K. Kim, "Maximin distributed detection in the presence of impulsive alpha-stable noise," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 6, pp. 1687–1691, 2011.

[57] E. E. Kuruolu, P. J. Rayner, and W. J. Fitzgerald, "Least Lp-norm impulsive noise cancellation with polynomial filters," *Signal Processing*, vol. 69, no. 1, pp. 1 – 14, 1998.

[58] J. G. Gonzalez, "Robust techniques for wireless communications in non-gaussian environments," Ph.D. dissertation, University of Delaware, Newark, DE, USA, 1997.

[59] J. G. Gonzalez and G. R. Arce, "Statistically-efficient filtering in impulsive environments: weighted myriad filters," *EURASIP J. Appl. Signal Process.*, vol. 2002, pp. 4–20, January 2002.

[60] J. Gonzalez and G. Arce, "Weighted myriad filters: a robust filtering framework derived from alpha-stable distributions," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 5, May 1996, pp. 2833 –2836 vol. 5.

[61] ——, "Optimality of the myriad filter in practical impulsive-noise environments," *Signal Processing, IEEE Transactions on*, vol. 49, no. 2, pp. 438 –441, Feb 2001.

[62] G. Arce, *Nonlinear signal processing: a statistical approach*, ser. Wiley InterScience online books. Wiley-Interscience, 2005.

[63] R. E. Carrillo, T. Aysal, and K. Barner, "Generalized cauchy distribution based robust estimation," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008, pp. 3389–3392.

[64] T. Aysal and K. Barner, "Meridian filtering for robust signal processing," *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 3949–3962, 2007.

[65] P. Huber and E. Ronchetti, *Robust Statistics*, ser. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 2009.

[66] J. Gonzalez, J. Paredes, and G. Arce, "Zero-order statistics: A mathematical framework for the processing and characterization of very impulsive signals," *Signal Processing, IEEE Transactions on*, vol. 54, no. 10, pp. 3839 –3851, Oct. 2006.

[67] A. Rajan and C. Tepedelenlioglu, "Diversity combining over rayleigh fading channels with symmetric alpha-stable noise," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 9, pp. 2968–2976, 2010.

[68] J. Haring and H. Vinck, "Ofdm transmission corrupted by impulsive noise," in *Proceedings of the 2000 International Symposium on Power-Line Communications and its Applications,(ISPCL2000)*, 2000, pp. 5–7.

[69] K. Al-Mawali, F. Al-Qahtani, and Z. Hussain, "Adaptive power loading for ofdm-based power line communications impaired by impulsive noise," in *Power Line Communications and Its Applications (ISPLC), 2010 IEEE International Symposium on*, 2010, pp. 178–182.

[70] S. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for ofdm receivers," *Communications, IEEE Transactions on*, vol. 56, no. 1, pp. 5–9, 2008.

[71] K. M. Rabie and E. Alsusa, "Improving blanking/clipping based impulsive noise mitigation over powerline channels," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, 2013, pp. 3413–3417.

[72] M. Korki, N. Hosseinzadeh, and T. Moazzeni, "Performance evaluation of a narrowband power line communication for smart grid with noise reduction technique," *Consumer Electronics, IEEE Transactions on*, vol. 57, no. 4, pp. 1598–1606, 2011.

[73] D.-F. Tseng, R.-B. Yang, T.-R. Tsai, Y. Han, and W.-H. Mow, "Efficient clipping for broadband power line systems in impulsive noise environment," in *Power Line Communications and Its Applications (ISPLC), 2012 16th IEEE International Symposium on*, 2012, pp. 362–367.

[74] T. Li, W.-H. Mow, and M. Siu, "Joint erasure marking and viterbi decoding algorithm for unknown impulsive noise channels," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 9, pp. 3407–3416, 2008.

[75] C.-H. Yih, "Iterative interference cancellation for ofdm signals with blanking nonlinearity in impulsive noise channels," *Signal Processing Letters, IEEE*, vol. 19, no. 3, pp. 147–150, 2012.

[76] S. Sparrer, A. Schenk, and R. F. H. Fischer, "Communication over impulsive noise channels: Channel coding vs. compressed sensing," in *Systems, Communication and Coding (SCC), Proceedings of 2013 9th International ITG Conference on*, 2013, pp. 1–6.

[77] E. Candes and P. Randall, "Highly robust error correction byconvex programming," *Information Theory, IEEE Transactions on*, vol. 54, no. 7, pp. 2829–2840, July 2008.

[78] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York, NY, USA: Cambridge University Press, 2004.

[79] A. Mahmood, M. Chitre, and M. A. Armand, "PSK communication with passband additive symmetric α-stable noise," *Communications, IEEE Transactions on*, vol. 60, no. 10, pp. 2990 –3000, Oct. 2012.

[80] ——, "Improving PSK performance in snapping shrimp noise with rotated constellations," in *Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems*, ser. WUWNet '12, no. 12, New York, NY, USA, 2012, pp. 1–8.

[81] ——, "Baseband characterization of additive white symmetric α-stable noise," in *Global Telecommunications Conference (GLOBECOM 2012), 2012 IEEE*, Dec. 2012, pp. 3696–3701.

[82] M. Chitre, S. Ong, and J. Potter, "Performance of coded ofdm in very shallow water channels and snapping shrimp noise," in *OCEANS, 2005. Proceedings of MTS/IEEE*, 2005, pp. 996–1001 Vol. 2.

[83] A. Mahmood, M. Chitre, and M. A. Armand, "On receiver design in additive white symmetric alpha-stable noise," *Communications, IEEE Transactions on*, (under review).

[84] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition.* Wiley-Interscience, Jul. 2006.

[85] Z. D. Bai and J. C. Fu, "On the maximum-likelihood estimator for the location parameter of a cauchy distribution," *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, vol. 15, no. 2, pp. pp. 137–146, 1987. [Online]. Available: http://www.jstor.org/stable/3315202

[86] T. S. Ferguson, "Maximum likelihood estimates of the parameters of the cauchy distribution for samples of size 3 and 4," *Journal of the American Statistical Association*, vol. 73, no. 361, pp. pp. 211–213, 1978. [Online]. Available: http://www.jstor.org/stable/2286549

[87] R. Storn and K. Price, "Differential evolution a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[88] V. Černý, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*, vol. 45, no. 1, pp. 41–51, 1985.

[89] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[90] T. Chuah, B. Sharif, and O. Hinton, "Nonlinear decorrelator for multiuser detection in non-gaussian impulsive environments," *Electronics Letters*, vol. 36, no. 10, pp. 920 –922, May 2000.

[91] S. Yoon, I. Song, and S. Y. Kim, "Code acquisition for ds/ss communications in non-gaussian impulsive channels," *Communications, IEEE Transactions on*, vol. 52, no. 2, pp. 187 – 190, Feb. 2004.

[92] T. C. Chuah, "On reed-solomon coding for data communications over power-line channels," *Power Delivery, IEEE Transactions on*, vol. 24, no. 2, pp. 614 –620, April 2009.

[93] M. Shafieipour, H. S. Lim, , and T. C. Chuah, "Decoding of turbo codes in symmetric alpha-stable noise," *ISRN Signal Processing*, vol. 2011, Article ID 683972, 7 pages 2011.

[94] A. Mahmood, M. Chitre, and M. A. Armand, "Decoding uncoded OFDM signals in alpha-stable noise via convex programming," *Communications, IEEE Transactions on*, (under review).

[95] ——, "Maximum-likelihood detection performance of uncoded OFDM in impulsive noise," in *Global Telecommunications Conference (GLOBECOM 2013), 2013 IEEE*, Dec. 2013.

[96] G. Strang, *Linear Algebra and Its Applications*. Thomson Brooks/Cole Cengage learning, 2006.

[97] W. H. DuMouchel, "On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution," *The Annals of Statistics*, vol. 1, no. 5, 1973.

[98] ——, "Stable distributions in statistical inference: 2. information from stably distributed samples," *Journal of the American Statistical Association*, vol. 70, no. 350, pp. pp. 386–393, 1975.

[99] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, ser. Fundamentals of Statistical Signal Processing. Prentice-Hall PTR, 1998, no. v. 1.

[100] J. P. Nolan and D. Ojeda-Revah, "Linear and nonlinear regression with stable errors," *Journal of Econometrics*, vol. 172, no. 2, pp. 186 – 194, 2013.

[101] E. Candes and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[102] D. Ojeda, “Comparative study of stable parameter estimators and regression with stably distributed errors,” Ph.D. dissertation, American University, 2001.