

On Single-Carrier Communication in Additive White Symmetric Alpha-Stable Noise

Ahmed Mahmood, *Student Member, IEEE*, Mandar Chitre, *Senior Member, IEEE*,
and Marc Andre Armand, *Senior Member, IEEE*

Abstract—In this paper we analyze design aspects of a single-carrier digital communications receiver in the presence of impulsive noise. We use the additive white symmetric α -stable noise (AWS α SN) to model the channel noise. By introducing passband sampling, efficient constellations and suitable baseband detectors, we show that the uncoded error performance of the conventional (linear) receiver can be enhanced given the real and imaginary components of the transmitted symbol are decoded separately. The performance may be improved further 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 in a single-carrier communications system.

Index Terms—impulsive noise, AWS α SN, constellation, estimation, detection.

I. INTRODUCTION

GAUSSIAN distributions are typically used to model practical noise processes that affect digital communication systems [1]. However, these models are not appropriate to represent *impulsive noise* as the latter generates outliers with large probability. In certain communication scenarios, impulsive noise dominates the available transmission spectrum. Examples are the shallow underwater channel [2]–[4], atmospheric noise [5], communication over powerlines [6] and digital subscriber line transmission [7]. Models based on heavy-tailed distributions are much more effective in dealing with impulsive noise [8]–[11]. The additive white symmetric α -stable noise (AWS α SN) model is a good fit to impulsive noise processes with IID samples [11], [12]. Theoretical justification for using stable models stems from the *generalized central limit theorem* (GCLT) which states that the sum of independent and identically distributed (IID) copies of a random variable (vector) converges to a *stable* variable (vector) as the number of elements in the sum tends to infinity [8], [13]. The well-known central limit theorem (CLT) is, in essence, the GCLT with an added power constraint on the random variables

(vectors). This implies that the Gaussian probability density function (PDF) is also a member of the stable family.

In the literature, performance analysis is typically performed in the baseband [1]. This approach implicitly assumes the passband-to-baseband conversion scheme to be a linear system [1], [14]. In some scenarios, such as underwater acoustic communications, the system bandwidth is small enough to allow passband sampling [4], [15]. Motivated by this, we introduced uniform sampling before the passband-to-baseband conversion process in our previous works [12], [16]. In non-Gaussian AWS α SN, this mechanism outputs a complex baseband noise sample that takes on various symmetric bivariate PDF configurations as determined by the sampling and carrier frequencies. In general, the baseband noise PDF is anisotropic, i.e., not circularly symmetric. Of all possible statistical configurations, the case of independent real and imaginary components can be exploited to provide the best possible error performance under maximum-likelihood (ML) detection in a single-carrier digital communications scheme [12], [16]. To truly harness the potential of the new receiver, we employed rotated phase shift keying (PSK) constellations due to the anisotropy of the resulting baseband noise. Using simulations, we showed that the modified receiver performed very well.

The goal of this paper is to further develop the underlying theory of designing a robust communications receiver in impulsive noise modeled by AWS α SN. The contributions of this paper are as follows:

- Building on the works in [12] and [16], we provide more clearer and intuitive expressions for baseband noise statistics in a linear receiver for passband non-Gaussian AWS α SN. We focus on the scheme that outputs a baseband noise vector with IID components as this offers the best error performance amongst all linear receivers [12].
- In the literature, a number of robust detectors have been analyzed in the presence of AWS α SN [3], [17], [18]. We provide a thorough error performance comparison of various closed-form baseband detectors in conjunction with the aforementioned linear receiver (with passband sampling) in this paper.
- It is a well-known fact that linear systems are sub-optimal in non-Gaussian AWS α SN [8]. If linear passband-to-baseband conversion is employed at the receiver, we show that the information loss can be quantified as signal-to-noise ratio (SNR) degradation.
- We extend our analysis to accommodate non-linear and joint-detection on the passband samples. These schemes are shown to surpass the performance of linear receivers

A. Mahmood is with the ARL, Tropical Marine Science Institute, National University of Singapore, Singapore 119227, Singapore (e-mail: tmsahme@nus.edu.sg)

M. Chitre is with the Department of Electrical and Computer Engineering and the ARL, Tropical Marine Science Institute, National University of Singapore, Singapore 119227, Singapore (e-mail: mandar@nus.edu.sg)

M. A. Armand is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, Singapore (e-mail: eleama@nus.edu.sg)

Manuscript received April 9, 2014; revised July 4, 2014 and August 15, 2014.

significantly.

- As the baseband noise derived from passband AWS α SN may be anisotropic [12], [16], typical constellation design methodologies based on maximizing the minimum Euclidean distance between symbol points [1] are far from optimal. We discuss new design methodologies that ensure near-optimum symbol placement when the baseband noise samples are IID. This plays an important role in ascertaining system error performance.
- To appreciate the performance gain of all the aforementioned schemes in non-Gaussian AWS α SN, a comparison needs to be made with the conventional linear receiver (without passband sampling). An analysis of the baseband noise statistics and the resulting error performance is provided for the latter.

This paper is organized as follows: In Section II, we present concepts and notation related to our work. We discuss various passband-to-baseband conversion schemes and several closed-form baseband detectors in Section III that perform significantly better in non-Gaussian AWS α SN. The analysis is performed for receivers under the linear and non-linear frameworks. In Section IV, we discuss joint-detection schemes that may be employed directly on the passband samples. The potential of the aforementioned schemes can be truly harnessed only if efficient constellations are employed. We discuss good constellation design techniques in Section V. Finally, we compare the performance of all receiver structures in Section VI and wrap up the discussion by presenting conclusions in Section VII.

II. CONCEPTS AND NOTATION

A. Symmetric α -Stable Distributions

If X is a random variable with PDF $f_X(X = x)$, then X is *symmetric* if $f_X(x) = f_X(-x)$. Further still, if $X^{(i)} \forall i \in \{1, 2, \dots, N\}$ are independent and identically distributed (IID) copies of X and

$$\sum_{i=1}^N a_i X^{(i)} \stackrel{d}{=} cX \quad (1)$$

where $N \in \mathbb{Z}^+$ and $a_i, c \in \mathbb{R}$, then X is a symmetric α -stable (S α S) random variable [13]. This, in essence, is a direct consequence of the GCLT. The symbol $\stackrel{d}{=}$ implies equality in distribution. The relationship between the coefficients in (1) is

$$c^\alpha = \sum_{i=1}^N |a_i|^\alpha. \quad (2)$$

An S α S PDF is completely parametrized by the *characteristic exponent* $\alpha \in (0, 2]$ and the *scale parameter* $\delta \in (0, +\infty)$. We denote such a distribution by $\mathcal{S}(\alpha, \delta)$ [12], [19]. With the exception of the Gaussian ($\alpha = 2$) and Cauchy ($\alpha = 1$) cases, the PDF of a S α S random variable does not exist in closed form [13], [20]. For the Gaussian case, $\mathcal{S}(2, \delta)$ corresponds to $\mathcal{N}(0, 2\delta^2)$, i.e., $2\delta^2$ is the variance of the distribution.

Besides stability, a defining characteristic of non-Gaussian S α S distributions is that they have algebraic (heavy) tails [13]. The heaviness of these tails is characterized by α . As $\alpha \rightarrow 0$,

the tails become *increasingly* heavier. This effect can be seen from the asymptotic convergence of an S α S PDF for $\alpha \neq 2$ as $|x| \rightarrow +\infty$ [20]:

$$f_X(x) \sim \left(\frac{\alpha \delta^\alpha \sin(\pi\alpha/2) \Gamma(\alpha)}{\pi} \right) |x|^{-\alpha-1} \quad (3)$$

Here, $\Gamma(\cdot)$ denotes the gamma function. As $\int_\epsilon^\infty x^q dx$ is divergent for $q \geq -1$ for all $\epsilon > 0$, then from (3), we clearly observe that the p^{th} -order moments for $p > \alpha$ are infinite. Thus, second-order moments are infinite for all non-Gaussian S α S random variables [8], [13].

The definition in (1) can be extended to the multivariate case. A random vector \vec{X} is symmetric if its PDF $f_{\vec{X}}(\vec{X} = \vec{x})$ satisfies $f_{\vec{X}}(\vec{x}) = f_{\vec{X}}(-\vec{x})$. Also, if $\vec{X}^{(i)} \forall i \in \{1, 2, \dots, N\}$ are IID copies of \vec{X} , then

$$\sum_{i=1}^N a_i \vec{X}^{(i)} \stackrel{d}{=} c\vec{X} \quad (4)$$

where $a_i, c \in \mathbb{R}$, is true if and only if \vec{X} is a S α S vector. The equality in (2) also extends to the multivariate case.

B. Additive White Symmetric α -Stable Noise

In the literature, the AWS α SN channel provides a good fit to impulsive noise data recorded in practical scenarios [2], [4], [15]. By definition, if $X_i \forall i \in \{1, 2, \dots, N\}$ are samples of a *real* AWS α SN process, they are IID copies of $X \sim \mathcal{S}(\alpha, \delta)$. Consequently, this is a *stationary* process. Let \vec{X} denote the N -dimensional random vector with i^{th} element X_i . The joint-PDF of \vec{X} is then

$$f_{\vec{X}}(\vec{x}) = \prod_{i=1}^N f_X(x_i). \quad (5)$$

As the components of \vec{X} are IID S α S random variables, we deduce from (1) and (4) that \vec{X} is a S α S vector [13]. For $\alpha = 2$, the AWS α SN process reduces to additive white Gaussian noise (AWGN). As the second-order moments are infinite for all non-Gaussian S α S distributions, the term ‘white’ signifies the IID nature of the samples and not a flat power spectral density (PSD) when $\alpha \neq 2$ [12]. For practical impulsive noise, the AWS α SN model typically provides a good fit for $\alpha \geq 1$ [3] [21, Pg. 33]. Therefore, we compile results within this range of α .

C. Passband Transmission & Reception

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

$$r(t) = s_i(t) + w(t) \quad (6)$$

where $s_i(t) \forall i \in \{0, 1, \dots, M-1\}$ is the transmitted passband signal corresponding to the i^{th} symbol in the constellation of size M , $w(t)$ is a continuous-time AWS α SN process and $r(t)$ is the corresponding received signal. We denote the PDF of

any passband noise sample by $f_W(x)$ where $W \sim \mathcal{S}(\alpha, \delta)$. In an M -QAM signaling scheme, $s_i(t)$ is written as

$$\begin{aligned} s_i(t) &= \Re \left\{ \sqrt{\frac{2\mathcal{E}_{s_i}}{\mathcal{E}_g}} g(t) e^{j\phi_i} e^{j2\pi f_c t} \right\} \\ &= \sqrt{\frac{2\mathcal{E}_{s_i}}{\mathcal{E}_g}} g(t) \cos(2\pi f_c t + \phi_i) \end{aligned} \quad (7)$$

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], [22]. 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 [1], [22]. 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 $\xi > \beta/2$. In practical systems, ξ is typically set to a large number.

The energies of $s_i(t)$ and $g(t)$ over $t \in [0, T)$ are denoted by \mathcal{E}_{s_i} and \mathcal{E}_g , respectively. The baseband symbol $\sqrt{\mathcal{E}_{s_i}} e^{j\phi_i}$ is represented by the constellation point $(s_{I_i}, s_{Q_i}) = (\sqrt{\mathcal{E}_{s_i}} \cos(\phi_i), \sqrt{\mathcal{E}_{s_i}} \sin(\phi_i))$ in the complex plane. Conventionally, orthonormal signaling is used to represent the passband modulated signal [1]. This is highlighted below:

$$\begin{aligned} s_i(t) &= \sqrt{\mathcal{E}_{s_i}} \cos(\phi_i) \ell_I(t) + \sqrt{\mathcal{E}_{s_i}} \sin(\phi_i) \ell_Q(t) \\ &= s_{I_i} \ell_I(t) + s_{Q_i} \ell_Q(t). \end{aligned} \quad (8)$$

On comparison with (7),

$$\ell_I(t) = \sqrt{\frac{2}{\mathcal{E}_g}} g(t) \cos(2\pi f_c t) \quad \text{and} \quad (9)$$

$$\ell_Q(t) = -\sqrt{\frac{2}{\mathcal{E}_g}} g(t) \sin(2\pi f_c t). \quad (10)$$

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

$$\int_0^T \ell_I^2(t) dt = \int_0^T \ell_Q^2(t) dt = 1, \quad \int_0^T \ell_I(t) \ell_Q(t) dt = 0. \quad (11)$$

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 (6) 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 (8) is that one determines the I and Q components of the transmitted symbol by mere inspection. At the receiver, one may retrieve s_{I_i} and s_{Q_i} by multiplying (8) 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 (6), the resultant output can be expressed in the following form:

$$\vec{r} = \vec{s}_i + \vec{w} \quad (12)$$

where

$$\vec{r} = \begin{bmatrix} r_I \\ r_Q \end{bmatrix}, \quad \vec{s}_i = \begin{bmatrix} s_{I_i} \\ s_{Q_i} \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} w_I \\ w_Q \end{bmatrix}.$$

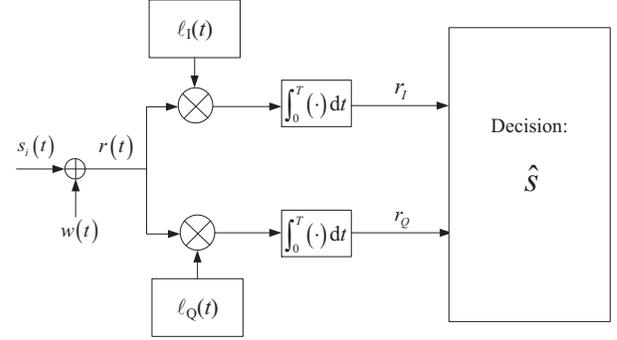


Fig. 1. Conventional continuous-time correlator-based receiver implementation.

Thus the continuous signal form in (6) is converted to the vector form in (12) which is termed as the baseband transmit-receive equation. This is then followed by detection in the complex plane to estimate the transmitted symbol. Given equiprobable symbols and (12), ML detection is optimal in reducing the error probability at the receiver. Mathematically, this is given by

$$\hat{s} = \arg \max_{\vec{\gamma} \in \mathbb{M}} f_{\vec{w}}(\vec{r} - \vec{\gamma}), \quad (13)$$

where $\vec{\gamma} = [\gamma_I, \gamma_Q]^T$, $f_{\vec{w}}(\vec{x})$ is the bivariate PDF of \vec{w} and \mathbb{M} is the set of all symbols in the constellation.

In Fig. 1 we present the receiver block structure based on the orthonormal signaling concepts discussed above. We term this schematic 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, \vec{w} is an isotropic Gaussian random vector. The components of \vec{w} are IID $\mathcal{N}(0, N_0/2)$ where $N_0/2$ is the two-sided PSD of the AWGN channel [1]. In non-Gaussian AWS α SN, \vec{w} is also an isotropic S α S vector (for a rigorous proof see Appendix A).

An isotropic S α S PDF $f_{\vec{w}}(\vec{x})$ has favorable geometric properties. Mathematically, $f_{\vec{w}}(\vec{x})$ is a function of $\|\vec{x}\|$, the Euclidean norm of \vec{x} [8]. Therefore, its equiprobable density contours are in the form of concentric circles around the origin. Further still, the marginal PDFs are *identical* [8]. In the Gaussian case, an isotropic distribution is only possible if \vec{w} has IID components [1]. However, for $\alpha \neq 2$, \vec{w} has identically distributed but *dependent* components [8], [13]. The conventional receiver performs poorly in non-Gaussian AWS α SN as it is a linear system and does not exploit the dependency between the I and Q components [8], [12]. It is therefore imperative that the receiver be designed more robust to impulsive noise.

In this paper, 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 (linear and non-linear, respectively) 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.

III. SOFT-ESTIMATES & BASEBAND DETECTION

In this section, we analyze various passband-to-baseband conversion schemes in non-Gaussian AWS α SN under the linear and non-linear frameworks. Given (6), the passband-to-baseband process outputs the soft-estimate \vec{r} of the transmitted symbol. This is subsequently passed through a detector. The optimal detection scheme depends on the statistics of the noise component \vec{w} , which in turn depends on the scheme adopted for passband-to-baseband conversion. Therefore, we will also discuss various detectors that may be used in conjunction with the proposed soft-estimation schemes.

A. Passband Sampling

To enhance the performance of the conventional receiver, the dependency between the I and Q noise components has to be exploited. Therefore, unlike the schematic in Fig. 1, the I and Q components of the transmitted symbol needs to be evaluated *jointly*. Alternatively, in [12] it was shown that by introducing *passband sampling* to the conventional receiver, \vec{w} could result in one of many anisotropic S α S distributions in non-Gaussian AWS α SN. If the I and Q channels are processed separately, the case of \vec{w} with independent components offers the best error performance over all possible statistical structures under ML detection [12], [16], [19]. We therefore focus on this particular case.

Denoting the sampling frequency by f_s , we observe that for $f_s = 4f_c$, (8) reduces to

$$s_i[n] = s_{I_i} \ell_I[n] + s_{Q_i} \ell_Q[n] \quad (14)$$

where

$$\ell_I[n] = \sqrt{\frac{2}{\mathcal{E}_g}} g[n] \cos(\pi n/2) \quad \text{and} \quad (15)$$

$$\ell_Q[n] = -\sqrt{\frac{2}{\mathcal{E}_g}} g[n] \sin(\pi n/2). \quad (16)$$

We use the square bracket notation to denote discrete-time signals sampled at $f_s = 4f_c$. The relationship between any discrete-time signal with its continuous-time counterpart is given by $q[n] = q(n/f_s)$. We note that only one of the functions in (15) and (16) 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 (14) into a superimposition of two distinct sequences:

$$s_i[n] = \begin{cases} s_{I_i} \ell_I[n] & \forall n \in \{0, 2, \dots, 4\xi - 2\} \\ s_{Q_i} \ell_Q[n] & \forall n \in \{1, 3, \dots, 4\xi - 1\} \end{cases} \quad (17)$$

By substituting variables in (17), we can separate the I and Q components completely:

$$\begin{aligned} s_i[2n] &= s_{I_i} \ell_I[2n] & \text{and} \\ s_i[2n+1] &= s_{Q_i} \ell_Q[2n+1] \end{aligned} \quad (18)$$

$\forall n \in \{0, 1, \dots, 2\xi - 1\}$. From (6), the sampled received signal is

$$r[n] = s_i[n] + w[n] \quad (19)$$

$\forall n \in \{0, 1, \dots, 4\xi - 1\}$. Following (17) and (18), the transmit-receive equation can be expressed as two parallel channels:

$$r_I[n] = r[2n] = s_i[2n] + w_I[n] \quad (20)$$

$$r_Q[n] = r[2n+1] = s_i[2n+1] + w_Q[n] \quad (21)$$

where

$$\begin{aligned} w_I[n] &= w[2n] & \text{and} \\ w_Q[n] &= w[2n+1] \end{aligned} \quad (22)$$

$\forall n \in \{0, 1, \dots, 2\xi - 1\}$. The expression in (22) can be thought of two independent (yet similar) AWS α SN 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 (20) and (21) is *sufficient* to ensure that \vec{w} will have independent components. In the remainder of this paper, we assume $f_s = 4f_c$ unless explicitly stated otherwise.

The arguments in this section may be extended to any noise process $w(t)$ that has IID samples. Do note, that for the sampling process to be non-lossy, the Nyquist criterion must be met for $s(t)$, i.e., $f_s > 2f_c + \beta/T$. As f_c is typically much larger than β/T [1], $f_s = 4f_c$ satisfies the Nyquist criterion.

B. Linear Baseband Conversion / Matched-Filter

Introducing passband sampling discretizes the processing at the receiver. Mathematically, the integrals in (11) reduce to the following sums:

$$\sum_{n=0}^{4\xi-1} \ell_I^2[n] = \sum_{n=0}^{4\xi-1} \ell_Q^2[n] = f_s \quad (23)$$

$$\sum_{n=0}^{4\xi-1} \ell_I[n] \ell_Q[n] = 0 \quad (24)$$

Eq. (24) is straightforward, we however prove (23) below:

$$\begin{aligned} \sum_{n=0}^{4\xi-1} \ell_I^2[n] &= \frac{2}{\mathcal{E}_g} \sum_{n=0}^{4\xi-1} g^2[n] \cos^2(\pi n/2) \\ &= \frac{2}{\mathcal{E}_g} \sum_{n=0}^{4\xi-1} \frac{g^2[n]}{2} (1 + \cos(\pi n)) \\ &= \frac{2}{\mathcal{E}_g} \sum_{n=0}^{4\xi-1} \frac{g^2[n]}{2}. \end{aligned} \quad (25)$$

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 $[-\beta/T, \beta/T]$. The Nyquist criteria for $g^2(t)$

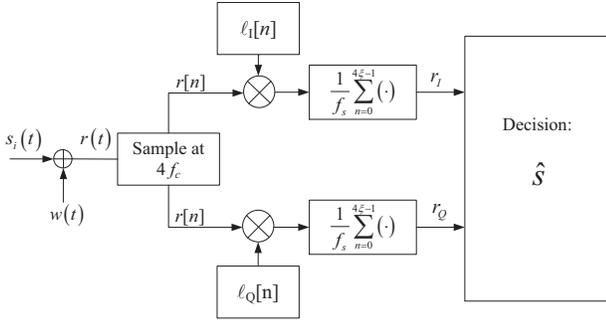


Fig. 2. Linear receiver schematic with passband sampling at $4f_c$.

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

$$\mathcal{E}_g = \frac{1}{f_s} \sum_{n=0}^{4\xi-1} g^2[n]. \quad (26)$$

Substituting this back into (25) gives us

$$\sum_{n=0}^{4\xi-1} \ell_I^2[n] = \frac{2}{\mathcal{E}_g} \times \frac{\mathcal{E}_g f_s}{2} = f_s. \quad (27)$$

Like its continuous counterpart, the goal of the linear discrete receiver is to re-acquire the I and Q components of the transmitted symbol. This is accomplished by initially estimating \vec{r} from $r[n]$. The properties of $\ell_I[n]$ and $\ell_Q[n]$ in (23) and (24) may be exploited to achieve this. We present the receiver structure in Fig. 2.

Following the arguments in Section III-A, \vec{w} will have independent components if the passband noise is AWS α SN. These are expressed as

$$w_I = \frac{1}{f_s} \sum_{n=0}^{4\xi-1} w[n] \ell_I[n], \quad w_Q = \frac{1}{f_s} \sum_{n=0}^{4\xi-1} w[n] \ell_Q[n]. \quad (28)$$

From (1), w_I and w_Q are each S α S random variables with characteristic exponent α . This in turn implies that \vec{w} is an S α S vector. Denoting the scale parameters of w_I and w_Q by δ_{w_I} and δ_{w_Q} , respectively, from (1) and (2) we can express w_I as

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

Therefore, we can express δ_{w_I} as function of δ (the scale parameter of the passband AWS α SN process):

$$\delta_{w_I} = \frac{\delta}{f_s} \left(\sum_{n=0}^{4\xi-1} |\ell_I[n]|^\alpha \right)^{\frac{1}{\alpha}}. \quad (29)$$

On substituting $\ell_I[n]$ with (15), we get

$$\begin{aligned} \delta_{w_I} &= \frac{\delta}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} \left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha \right)^{1/\alpha} \\ &\approx \frac{\delta}{f_s} \sqrt{\frac{2}{\frac{2}{f_s} \sum_{n=0}^{2\xi-1} g^2[2n]}} \left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha \right)^{1/\alpha} \end{aligned} \quad (30)$$

$$= \frac{\delta}{\sqrt{f_s}} \frac{\left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha \right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g^2[2n] \right)^{1/2}}. \quad (31)$$

We observe that $g[2n]$ results from sampling $g(t)$ at a rate of $f_s/2$. The approximation for \mathcal{E}_g in (30) is valid as long as $f_s/2 > 2\beta/T \Rightarrow \xi > \beta$. The form in (31) is intuitive as it depicts δ_{w_I} varying proportionally with $\delta/\sqrt{f_s}$ for all $\xi > \beta$. Similarly, δ_{w_Q} may also be evaluated from (16) and (28):

$$\delta_{w_Q} = \frac{\delta}{\sqrt{f_s}} \frac{\left(\sum_{n=0}^{2\xi-1} |g[2n+1]|^\alpha \right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g^2[2n+1] \right)^{1/2}}. \quad (32)$$

Given that the Nyquist criterion is also satisfied for $|g(t)|^\alpha$ with sampling rate $f_s/2$, we note that (31) and (32) are equivalent. Therefore, \vec{w} has IID components, each distributed by $\mathcal{S}(\alpha, \delta_w)$ where $\delta_w = \delta_{w_I} = \delta_{w_Q}$. Precisely, from (5), the bivariate PDF of \vec{w} is given by

$$f_{\vec{w}}(\vec{x}) = f_w(x_1) f_w(x_2) \quad (33)$$

where $f_w(\cdot)$ is the PDF of $w \sim \mathcal{S}(\alpha, \delta_w)$.

Specifically for the Gaussian case, $\delta_w = \delta/\sqrt{f_s}$. Therefore, w_I and w_Q are each $\mathcal{N}(0, 2\delta^2/f_s)$. 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 $2\delta^2 = N_0 f_s/2$, which implies that $w_I \stackrel{d}{=} w_Q \sim \mathcal{N}(0, N_0/2)$. Thus, \vec{w} is statistically equivalent to its counterpart in the conventional receiver in Fig. 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 s_{I_i} and s_{Q_i} in (19) are determined by the linear correlator [1]:

$$r_I = \frac{\sum_{n=0}^{4\xi-1} r[n] \ell_I[n]}{\sum_{n=0}^{4\xi-1} \ell_I^2[n]} = \frac{\sum_{n=0}^{4\xi-1} r[n] \ell_I[n]}{f_s} \quad (34)$$

$$r_Q = \frac{\sum_{n=0}^{4\xi-1} r[n] \ell_Q[n]}{\sum_{n=0}^{4\xi-1} \ell_Q^2[n]} = \frac{\sum_{n=0}^{4\xi-1} r[n] \ell_Q[n]}{f_s} \quad (35)$$

On comparing Fig. 2 and the expressions above, we note that the implementation is indeed ML-based.

In non-Gaussian AWS α SN, \vec{w} 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 with IID components in the latter. An instance of this is shown for the standard Cauchy case in Fig. 3. We note that the ‘tails’ are positioned along both the positive and negative directions of each axis in the complex plane. This attribute is shared amongst all $\alpha \neq 2$ cases. The sampling process keeps the transmitted information intact, yet

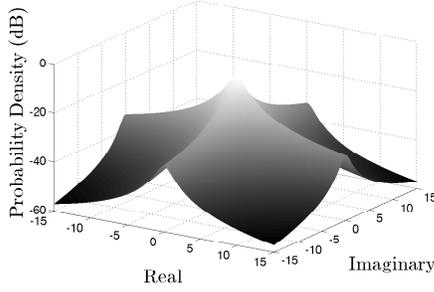


Fig. 3. The bivariate PDF of a standard Cauchy vector with IID components.

it statistically changes \vec{w} . Using the following argument we show why \vec{w} with independent components provides better performance given that the I and Q channels are processed individually.

We assume that the transmitted information is fully preserved in the conversion from $r(t)$ to \vec{r} . 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, \vec{s}_i can be fully recovered from $s_i(n/f_s) \forall n \in \{0, 1, \dots, \lfloor Tf_s \rfloor - 1\}$. If the accompanying noise samples $w(n/f_s) \forall n \in \{0, 1, \dots, \lfloor Tf_s \rfloor - 1\}$ are passed through the estimator, the information within \vec{w} 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 $H(w_I, w_Q)$ of the components of \vec{w} . This may be expressed as

$$H(w_I, w_Q) = H(w_I) + H(w_Q) - I(w_I; w_Q) \quad (36)$$

where $H(w_I)$ and $H(w_Q)$ are the self-entropies of w_I and w_Q , respectively, and $I(w_I; w_Q)$ is the information shared between them [23]. For any given unbiased estimator and Tf_s , $H(w_I, w_Q)$ will be *constant*. In the presence of non-zero $I(w_I; w_Q)$, the self-entropies increase correspondingly to maintain equality in (36). Therefore, receivers that process the I and Q channels separately perform better when $I(w_I; w_Q) = 0$ as $H(w_I)$ and $H(w_Q)$ 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 w_I and w_Q are independent for any IID noise process $w(t)$. This guarantees $H(w_I)$ and $H(w_Q)$ to be at their respective minimums. Therefore, we expect the receiver in Fig. 2 to generally perform better than that in Fig. 1.

C. Non-Linear Baseband Conversion

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

$$\vec{r} = \arg \min_{\vec{\mu}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} -\log \rho \left(r \left(\frac{n}{f_s} \right) - \mu_I \ell_I \left(\frac{n}{f_s} \right) - \mu_Q \ell_Q \left(\frac{n}{f_s} \right) \right) \quad (37)$$

where $\vec{\mu} = [\mu_I, \mu_Q]^T \in \mathbb{R}^2$, $\rho(x) \in \mathbb{R}^+ \forall x \in \mathbb{R}$ and $f_s > 2f_c + \beta/T$. The expression in (37) stems from robust generalized ML estimation (or M-estimator) theory [24]. If $\rho(x) = f_W(x)$, then \vec{r} is the ML estimate of $\vec{\mu}$. In the context of digital communications, \vec{r} is the soft-ML estimate of \vec{s}_i . Substituting $\rho(x)$ with the Gaussian PDF corresponding to $\mathcal{S}(2, \delta)$ in (37), results in two single-variable minimizations:

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

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

Therefore, the ML estimator of $\vec{\mu}$ *separately processes* the I and Q channels for all f_s in the Gaussian case and its form is similar to (34) and (35). However, (37) cannot be split into separate minimizations of μ_I and μ_Q if $w(t)$ is non-Gaussian AWS α SN.

Proposition 1: For $f_s = 4f_c$, the bivariate minimization in (37) is equivalent to individually evaluating

$$r_I = \arg \min_{\mu_I} \sum_{n=0}^{2\xi-1} -\log \rho(r_I[n] - \mu_I \ell_I[2n]) \quad \text{and} \quad (38)$$

$$r_Q = \arg \min_{\mu_Q} \sum_{n=0}^{2\xi-1} -\log \rho(r_Q[n] - \mu_Q \ell_Q[2n+1]) \quad (39)$$

for all $\rho(x)$.

Proof: For $f_s = 4f_c$, (37) becomes

$$\vec{r} = \arg \min_{\vec{\mu}} \sum_{n=0}^{4\xi-1} -\log \rho(r[n] - \mu_I \ell_I[n] - \mu_Q \ell_Q[n]) \quad (40)$$

From the discussion in Section III-A, we know that $\ell_I[n]$ is non-zero for $n \in \{0, 2, \dots, 4\xi - 2\}$ and $\ell_Q[n]$ for $n \in \{1, 3, \dots, 4\xi - 1\}$. Therefore, (40) may be rewritten as

$$\vec{r} = \arg \min_{\vec{\mu}} \left(\sum_{n=0}^{2\xi-1} -\log \rho(r[2n] - \mu_I \ell_I[2n]) + \sum_{n=0}^{2\xi-1} -\log \rho(r[2n+1] - \mu_Q \ell_Q[2n+1]) \right). \quad (41)$$

We observe that evaluating (41) is equivalent to *individually* minimizing over μ_I and μ_Q to get r_I and r_Q , respectively. Since no assumption was made about $\rho(x)$, (38) and (39) hold for all $\rho(x)$. \square

Corollary 1: Given that $f_s > 2f_c + \beta/T$, the separation in (41) is possible for all $\rho(x)$ if and only if $f_s = 4f_c$.

Corollary 1 is a direct consequence of the fact that $r(n/f_s)$ splits into (20) and (21) if and only if $f_s = 4f_c$ and the Nyquist criterion for $s(t)$ is met. Therefore, $f_s = 4f_c$ is a *sufficient* condition for any scheme to achieve the ML estimate of $\vec{\mu}$ in (37) if the estimation is done individually for the I and Q components. In Fig. 4, we present a general uncoded receiver schematic that optimizes error performance if the I and Q channels are processed separately.

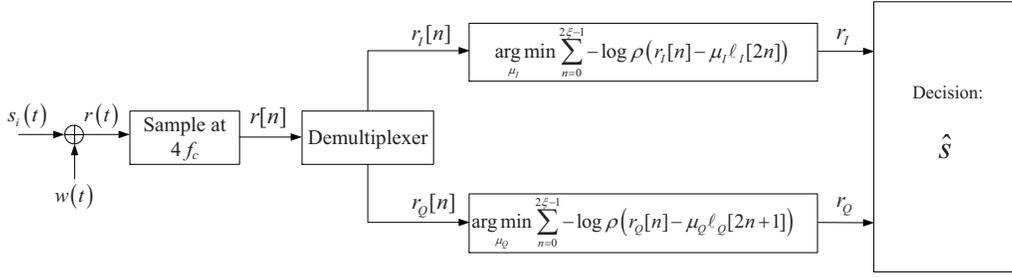


Fig. 4. General receiver schematic with passband sampling at $4f_c$.

Our focus for the remainder of this section will be on analyzing (38) and (39). As the expressions are similar, we drop the subscripts and deal with the general expression

$$y = \arg \min_{\mu} \sum_{n=0}^{2\xi-1} -\log \rho(x[n] - \mu \ell[n]). \quad (42)$$

From an implementation perspective, the ML estimator of μ in AWS α SN may not be desirable due to the lack of closed form S α S PDFs. The PDF of a Cauchy random variable $X \sim \mathcal{S}(1, \delta)$, however, is given by

$$f_W(x) = \frac{\delta}{\pi(\delta^2 + x^2)}. \quad (43)$$

On substituting (43) for $\rho(x)$ in (42) and simplifying, we get the ML Cauchy estimator for μ :

$$y = \arg \min_{\mu} \sum_{n=0}^{2\xi-1} \log(\delta^2 + (x[n] - \mu \ell[n])^2). \quad (44)$$

The cost function in (44) consists of multiple local minimas/maximas in μ . To observe this, we can rewrite (44) equivalently as

$$y = \arg \min_{\mu} \prod_{n=0}^{2\xi-1} (\delta^2 + (x[n] - \mu \ell[n])^2). \quad (45)$$

Clearly, (45) is a 4ξ order polynomial in μ . The ML Cauchy estimator and its variants have been employed vastly in the literature to combat impulsive noise [25]. 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 ML Cauchy estimate for μ in (44) may be seen in the light of a class of robust M-estimators; namely the *matched myriad filter* (MMyF) [25]–[27]. The MMyF estimate $y(K)$ of μ with *linearity parameter* $K \in \mathbb{R}^+$ is given by

$$y(K) = \arg \min_{\mu} \sum_{n=0}^{2\xi-1} \log \left(K^2 + \ell^2[n] \left(\frac{x[n]}{\ell[n]} - \mu \right)^2 \right). \quad (46)$$

From observation, (46) is equivalent to the cost function in (44) when $K = \delta$, i.e., $y(\delta)$ 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 [26]:

- 1) As $K \rightarrow +\infty$, the MMyF converges to the linear correlator, which is the optimum ML estimate in Gaussian noise, i.e., $\alpha = 2$.
- 2) As $K \rightarrow 0$, the MMyF becomes a mode-selector, i.e., the estimate is equal to the element in $\{x[n]/\ell[n]\} \forall n \in \{0, 1, \dots, 2\xi - 1\}$ 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., $\alpha \rightarrow 0$.

Thus by varying K , one may achieve ML optimality for three scenarios within the S α S 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 $K = K(\alpha, \delta)$ as a monotonically increasing function of α and δ that attains the three points of optimality: $K(2, \delta) = +\infty$, $K(1, \delta) = \delta$ and $K(0, \delta) = 0$. Further still, if $K(\alpha, \delta)$ offers the optimal estimate of μ for all $\alpha \in (0, 2]$ in (46), the scale parameter is separable, i.e. $K(\alpha, \delta) = K(\alpha)\delta$ [25, Eq. 31]. The MMyF estimate $y(K)$ can now be written as

$$y(K) = \arg \min_{\mu} \sum_{n=0}^{2\xi-1} \log \left(K^2(\alpha)\delta^2 + \ell^2[n] \left(\frac{x[n]}{\ell[n]} - \mu \right)^2 \right). \quad (47)$$

In the literature, a heuristic function has been proposed for $K(\alpha)$ that works well for all $\alpha \in (0, 2]$ [25]:

$$K(\alpha) = \sqrt{\frac{\alpha}{2 - \alpha}}. \quad (48)$$

The MMyF offers good near-optimal estimates of $\vec{\mu}$ in the general AWS α SN case. Like (44), the cost function in (47) will have at most 4ξ minimas/maximas in μ . As the number of samples in (47) increases, $y(K)$ converges to a normal distribution for all $\alpha \in (0, 2]$ [26], [28]. Keeping this in mind, it is intuitive to assume that \vec{r} will have an isotropic Gaussian distribution for large values of ξ . In this case Euclidean detection will be optimal.

Besides the Cauchy estimator and the MMyF, other functions known to perform well in impulsive noise may also be

used as $\rho(x)$ in (42). As the objective is to approximate ML estimation as close as possible, it is logical to find *analytic* functions $\bar{f}_W(x)$ that closely resemble $f_W(x)$ and substitute them for $\rho(x)$. An example is the PDF

$$\bar{f}_W(x) = \frac{d_1}{\delta} \exp(-d_2 \left| \frac{x}{\delta} \right|^p), \quad (49)$$

where $0 < p < \alpha$. Here, d_1 and d_2 are positive (normalizing) constants and δ is the scale parameter of the distribution. On substituting (49) for $\rho(x)$ in (42) and simplifying, we get

$$\begin{aligned} y &= \arg \min_{\mu} \sum_{n=0}^{2\xi-1} |x[n] - \mu \ell[n]|^p \\ &= \arg \min_{\mu} \|x[n] - \mu \ell[n]\|_p^p \end{aligned} \quad (50)$$

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 (49). The L_p -norm for $0 < p < \alpha$ is convergent in the ergodic sense and is known to perform very well in impulsive noise [8], [11], [29]. By changing p one can tweak the ‘tails’ of the general PDF in (49). For $p = 2$, (49) is a Gaussian PDF. On the other hand, as $p \rightarrow 0$, $\bar{f}_W(x)$ becomes a constant (zero), i.e., $\bar{f}_W(x) \rightarrow (d_1/\delta) \exp(-d_2)$. This implies that the tails of (49) become increasingly heavier as $p \rightarrow 0$. 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 .

Similarly, the asymptotic PDF expression in (3) may be employed as $\rho(x)$ to get

$$y = \arg \min_{\mu} \sum_{n=0}^{2\xi-1} \log |x[n] - \mu \ell[n]|. \quad (51)$$

We term this as the *log-norm estimator*. On comparison with (47), we note that (51) is the MMyF estimate with $K(\alpha) = 0$ which corresponds to a mode-type estimator. This is not surprising as (3) models each $w[n]$ as an impulse.

The cost functions in (47), (50) and (51) are in analytic forms. However, the estimators themselves cannot be represented in closed form. Therefore, the minimizations have to be numerically evaluated. The computational cost will depend on the algorithm used. 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 irrespective of the number of samples [30]. 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 \vec{r} is of the form in (12), the statistics of \vec{w} need to be known before the detection stage. The components of \vec{w} are independent as $f_s = 4f_c$. From (38) and (39), it is not hard to convince ourselves that $w_I \stackrel{d}{=} w_Q$, therefore \vec{w} has IID components. The list of near-optimal robust estimators

certainly does not exhaust here [8], [31], [32]. We have discussed popular schemes and due to the lack of space, we cease further discussion on non-linear estimators.

D. Baseband Detection

As of now, we have discussed how $r(t)$ can be processed to get \vec{r} . The next step is to detect the transmitted symbol within \mathbb{M} from \vec{r} . In the literature, a number of articles propose robust detectors in AWS α SN [3], [17], [18]. In essence, the detection problem is analogous to that of estimation but with the arguments restricted to a finite search space. We comment on the performance a few parametric & non-parametric detectors in conjunction with linear and non-linear estimation of $\vec{\mu}$ in non-Gaussian AWS α SN with $f_s = 4f_c$.

1) *Maximum-Likelihood Detection*: Using (33), we may rewrite the ML detector in (13) as

$$\hat{s} = \arg \max_{\tilde{\gamma} \in \mathbb{M}} f_w(r_I - \gamma_I) f_w(r_Q - \gamma_Q). \quad (52)$$

The statistics of \vec{w} need to be fully known to evaluate (52). In the case of the discretized linear receiver, the components of \vec{w} are IID $\mathcal{S}(\alpha, \delta_w)$. As $f_w(\cdot)$ does not generally exist in closed form, numerical evaluations such as those in [33], [34] are employed to evaluate (52). In the non-linear case, the statistics of \vec{w} 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 III-C \vec{w} should approximate a Gaussian vector with IID components on increasing ξ . Therefore \vec{w} should be near-isotropic.

2) *The Euclidean Detector*: The detection rule for this is

$$\hat{s} = \arg \min_{\tilde{\gamma} \in \mathbb{M}} \|(\vec{r} - \tilde{\gamma})\|^2 = \arg \min_{\tilde{\gamma} \in \mathbb{M}} (E[\|\tilde{\gamma}\|^2] - 2\vec{r}^T \tilde{\gamma}) \quad (53)$$

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

3) *The Myriad Detector*: We may invoke the myriad detector at the output of the linear estimator. The detection rule is

$$\begin{aligned} \hat{s} &= \arg \min_{\tilde{\gamma} \in \mathbb{M}} (\log |K(\alpha) \delta_w^2 + (r_I - \gamma_I)^2| \\ &\quad + \log |K(\alpha) \delta_w^2 + (r_Q - \gamma_Q)^2|). \end{aligned} \quad (54)$$

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

4) *The L_p -Norm Detector*: For the discretized linear receiver, the L_p -norm detector for $p < \alpha$ is defined as

$$\begin{aligned} \hat{s} &= \arg \min_{\tilde{\gamma} \in \mathbb{M}} \|(\vec{r} - \tilde{\gamma})\|_p^p \\ &= \arg \min_{\tilde{\gamma} \in \mathbb{M}} (|r_I - \gamma_I|^p + |r_Q - \gamma_Q|^p). \end{aligned} \quad (55)$$

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.

5) *The Log-Norm Detector*: Like its estimator counterpart, the asymptotic detector is based on (3). On substituting (3) in place of $f_w(x)$ in (52), we get

$$\hat{s} = \underset{\tilde{\gamma} \in \mathbb{M}}{\operatorname{arg\,min}} (|r_I - \gamma_I| |r_Q - \gamma_Q|) \quad (56)$$

or equivalently

$$\hat{s} = \underset{\tilde{\gamma} \in \mathbb{M}}{\operatorname{arg\,min}} (\log |r_I - \gamma_I| + \log |r_Q - \gamma_Q|). \quad (57)$$

We note that (57) is merely the logarithm of the cost function in (56). Either one may be used. The log-norm detector may be employed in the linear case as \vec{w} is S α S. The log-norm performs well in the medium-to-high SNR regime where errors are predominantly determined by impulses in w_I and/or w_Q . Like the L_p -norm and myriad detectors, it also offers near-optimal performance. However, it has the added advantage of not requiring any knowledge about α and δ_w .

IV. 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 \vec{s}_i directly from $r(n/f_s)$. Analogous to (37), the joint-detector is given by

$$\hat{s} = \underset{\tilde{\gamma} \in \mathbb{M}}{\operatorname{arg\,min}} \sum_{n=0}^{\lfloor Tf_s \rfloor - 1} -\log \rho \left(r \left(\frac{n}{f_s} \right) - \gamma_I \ell_I \left(\frac{n}{f_s} \right) - \gamma_Q \ell_Q \left(\frac{n}{f_s} \right) \right). \quad (58)$$

From the discussion in Section III-C, 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 $-\log \rho(x)$ by (49) and (3), respectively. Though cumbersome, one may also substitute $\rho(x)$ with $f_w(x)$ to implement ML joint-detection. In this case, the PDF will have to be numerically evaluated for each of its arguments in (58) for $\alpha \neq 1$.

From an implementation perspective, the joint-detector is preferred as evaluating μ in (42) for non-Gaussian AWS α SN requires a numerical technique even if $\rho(x)$ is in analytic form. Also, the $f_s = 4f_c$ constraint, which is required to simplify (37) into two single variable minimizations, does not significantly reduce the computational cost in joint-detection. This is due to the fact that it only nullifies a summation term in the argument of $\rho(x)$ in (58) $\forall n \in \{0, 1, \dots, \lfloor Tf_s \rfloor - 1\}$. Therefore, the constraint may be discarded. More importantly, converting $r(n/f_s)$ to the vector form in (12), even if \vec{r} is the ML estimate of $\vec{\mu}$, may result in loss of information (due to the local traps in the cost function) 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 $p \geq 1$ (convex problem) or the MMyF with large ξ . Even if there are no traps, there will be some loss at the

detection stage as the statistics of \vec{w} are not truly known and are assumed to be isotropic.

Though joint-detection is advantageous in both performance and implementation, it does not output soft-values. One way to ensure compatibility with soft-decoders is to use the costs in (58) to generate approximates to the log-likelihood ratios (LLRs) involved. For example, in the case of binary modulation, we have

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

The LLR is exact for $\rho(x) = f_w(x)$. 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 w_I and w_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 AWS α SN.

V. CONSTELLATION DESIGN

Signal constellations are conventionally designed for isotropic \vec{w} . This is reasonable as the passband noise process is typically modeled by AWGN and the receiver in Fig. 1 is employed. However, as highlighted previously, \vec{w} is S α S ($\alpha \neq 2$) with IID components if the discretized linear receiver is used in non-Gaussian AWS α SN. Similarly, if non-linear passband-to-baseband conversion is employed for small ξ , it is reasonable to assume that \vec{w} still retains some impulsiveness. In such a case $f_{\vec{w}}(\vec{x})$ will be anisotropic and of the form in Fig. 3. Statistically, the symmetry is given by

$$w_I + jw_Q \stackrel{d}{=} w_I - jw_Q \stackrel{d}{=} (w_I + jw_Q)e^{j\pi/2}. \quad (59)$$

If the constellation $s_i = s_{I_i} + js_{Q_i} \forall i \in \{0, 1, \dots, M-1\}$ has a certain error performance, then from (12) and (59), the symbol sets $s_i e^{jk\pi/2}$ and $s_i^* e^{jk\pi/2} \forall i \in \{0, 1, \dots, M-1\}$ offer similar performance for any $k \in \mathbb{Z}$. This has been exploited in [12], [16] by rotating constellations to achieve significant improvement for phase shift keying (PSK). However, the underlying theory for good symbol placement is still not established. Finding the constellation that offers globally optimal/near-optimal performance is a problem of interest, especially if the gains are large [12], [16]. We address these problems here.

Before we discuss constellation design, it is important that we define a suitable SNR measure for impulsive noise. In AWGN, error performance is typically analyzed against \mathcal{E}_b/N_0 , where $\mathcal{E}_b = E[\mathcal{E}_{s_i}]/\log_2 M$ is the average transmission energy per bit [1]. As per the discussion in Section III-B, $N_0 = 4\delta_w^2$ where δ_w is the scale parameter of $w_I \stackrel{d}{=} w_Q$ in the discretized linear receiver. Therefore, we employ the following SNR measure:

$$\frac{\mathcal{E}_b}{N_0} = \frac{E[\mathcal{E}_{s_i}]}{4\delta_w^2 \log_2 M}. \quad (60)$$

Given (12), the symbol error probability (SEP) is evaluated by

$$\text{SEP} = \frac{1}{M} \sum_{i=0}^{M-1} \int_{\vec{r} \notin \mathbb{S}_i} f_{\vec{w}}(\vec{r} - \vec{s}_i) d\vec{r},$$

where $\mathbb{S}_i \subseteq \mathbb{R}^2$ is the set of points (determined by the detection rule) that lie in the decision region of \vec{s}_i . Optimizing constellation corresponds to minimizing (61) with respect to $\vec{s}_i \forall i \in \{0, 1, \dots, M-1\}$. As the SEP is a function of \mathcal{E}_b/N_0 , a constraint also needs to be imposed on the ML decision regions are complex and cannot be expressed in closed form [16]. Intuitively, given \vec{s}_i is transmitted, one would want the tails of the PDF $f_{\vec{w}}(\vec{r} - \vec{s}_i)$ directed away from any other constellation point. This ensures that the impulses lie within the right decision region. Denoting the optimized constellation as \mathbb{M}^* , we propose minimizing a simpler constrained problem:

$$\begin{aligned} \mathbb{M}^* = \underset{\vec{s}_0, \dots, \vec{s}_{M-1}}{\text{argmin}} & \sum_{i=0}^{M-1} \sum_{\substack{k=0 \\ k \neq i}}^{M-1} f_{\vec{w}}(\vec{s}_k - \vec{s}_i) \\ \text{s. t.} & \mathcal{E}_b/N_0 \leq \tau \end{aligned} \quad (62)$$

or equivalently

$$\begin{aligned} \mathbb{M}^* = \underset{\vec{s}_0, \dots, \vec{s}_{M-1}}{\text{argmin}} & \sum_{i=0}^{M-1} \sum_{\substack{k=0 \\ k \neq i}}^{M-1} \log f_{\vec{w}}(\vec{s}_k - \vec{s}_i) \\ \text{s. t.} & \mathcal{E}_b/N_0 \leq \tau \end{aligned} \quad (63)$$

for some non-negative real number τ . The cost in (62) is merely (61) with \mathbb{S}_i restricted to only \vec{s}_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.

For $\alpha = 2$, optimizing (61) corresponds to maximizing the minimum Euclidean distance between all points. Extending this concept to non-Gaussian \vec{w} , it is reasonable to maximize a measure within the points of the constellation. This can be observed by splitting $f_{\vec{w}}(\vec{x})$ in (63) into a product of its marginals, as in (33), and substituting (3) in place of $f_w(x)$:

$$\begin{aligned} \mathbb{M}^* = \underset{\vec{s}_0, \dots, \vec{s}_{M-1}}{\text{argmin}} & \sum_{i=0}^{M-1} \sum_{\substack{k=0 \\ k \neq i}}^{M-1} -\log |s_{I_k} - s_{I_i}| \\ & -\log |s_{Q_k} - s_{Q_i}| \\ \text{s. t.} & \mathcal{E}_b/N_0 \leq \tau. \end{aligned} \quad (64)$$

Similarly we may use (49) (with $\delta = \delta_w$) for $f_w(x)$ to get

$$\begin{aligned} \mathbb{M}^* = \underset{\vec{s}_0, \dots, \vec{s}_{M-1}}{\text{argmin}} & \sum_{i=0}^{M-1} \sum_{\substack{k=0 \\ k \neq i}}^{M-1} -\|\vec{s}_k - \vec{s}_i\|_p^p \\ \text{s. t.} & \mathcal{E}_b/N_0 \leq \tau. \end{aligned} \quad (65)$$

Therefore, minimizing (63) can be interpreted as maximizing the combined log-norm or L_p -norm between the constellation points. As (64) and (65) are independent of α , the resultant constellation will be 'efficient' for all non-Gaussian \vec{w} .

Though the optimal constellation generally varies with SNR for a given M and $\alpha \neq 2$, it is almost constant in the medium-to-high regime where the errors are predominantly determined

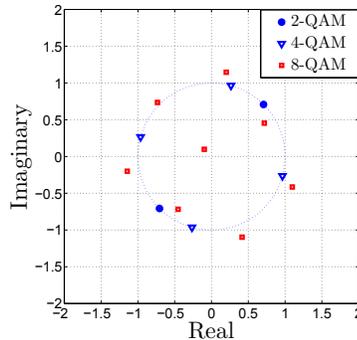


Fig. 5. Optimal Constellations for various M for medium-to-high \mathcal{E}_b/N_0 .

TABLE I
OPTIMAL SYMBOL PLACEMENT.

$M = 2$		$M = 4$		$M = 8$	
$\sqrt{\mathcal{E}_{s_i}}$	ϕ_i	$\sqrt{\mathcal{E}_{s_i}}$	ϕ_i	$\sqrt{\mathcal{E}_{s_i}}$	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

by the tails of $f_{\vec{w}}(\vec{x})$. From the discussion on (50) and (51), minimizing (63), (64) or (65) (for small p) in this regime will offer near-optimal results. The minimization has to be performed over $2M$ variables; $\{s_{I_i}, s_{Q_i}\} \forall i \in \{1, 2, \dots, M\}$. This may be accomplished via general search methods like Differential Evolution [35] or Simulated Annealing [36], [37].

In Fig. 5 we present constellations for various M that offer the best error performance for Cauchy \vec{w} with IID components for $\mathcal{E}_b/N_0 = 30$ dB and $E[\mathcal{E}_{s_i}] \leq 1$. Each constellation unit is generated by evaluating (63). The unit circle is also plotted for comparison. In complex form, the i^{th} point is $\sqrt{\mathcal{E}_{s_i}} e^{j\phi_i}$. In Table I, we have listed down $\sqrt{\mathcal{E}_{s_i}}$ and ϕ_i (in degrees) $\forall i \in \{0, 1, \dots, M-1\}$ for each of the constellations in Fig. 5. 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 [12]. We show this by presenting numerical results in the following section.

VI. RECEIVER PERFORMANCE

A. SNR Analysis

For the sake of fair comparison, the performance of each receiver needs to be analyzed for the passband AWS α SN process that amounts to (60) in the discretized linear receiver. Therefore, we need to evaluate \mathcal{E}_b/N_0 as a function of δ . From (31),

$$\delta_w = \frac{d(\alpha, \xi, g[n])}{\sqrt{f_s}} \delta, \quad (66)$$

where

$$d(\alpha, \xi, g[n]) = \frac{\left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha\right)^{1/\alpha}}{\left(\sum_{n=0}^{2\xi-1} g[2n]^2\right)^{1/2}}. \quad (67)$$

We note that (67) is a ratio of the L_α and L_2 norms of $g[2n] \forall n \in \{0, 1, \dots, 2\xi - 1\}$ and is solely a function of α , ξ and the *sampled* baseband shaping pulse. Finally, from (60) and (66),

$$\begin{aligned} \frac{\mathcal{E}_b}{N_0} \text{ (dB)} &= 10 \log_{10} \frac{\mathcal{E}_b}{N_0} \\ &= 10 \log_{10} \frac{E[\mathcal{E}_{s_i}] f_s}{4\delta^2 \log_2 M} - 20 \log_{10} d(\alpha, \xi, g[n]). \end{aligned} \quad (68)$$

As $\|\vec{x}\|_p \geq \|\vec{x}\|_2$ for all $\vec{x} \in \mathbb{R}^{2\xi}$ 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 $4\delta^2/f_s$ may be interpreted as the *pseudo*-PSD of the AWS α SN 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 operating \mathcal{E}_b/N_0 . In the special case of $g(t)$ being a rectangular pulse, (68) reduces to

$$\frac{\mathcal{E}_b}{N_0} \text{ (dB)} = 10 \log_{10} \frac{E[\mathcal{E}_{s_i}] f_s}{4\delta^2 \log_2 M} - 10 \left(\frac{2}{\alpha} - 1 \right) \log_{10} 2\xi. \quad (69)$$

We note that (69) decreases linearly with $10 \log_{10} 2\xi$ at a rate proportional to $\frac{2}{\alpha} - 1$. In essence, one can arbitrarily *reduce* the \mathcal{E}_b/N_0 by *increasing* ξ . For α close to 2, increasing ξ causes no significant effect in \mathcal{E}_b/N_0 . However, as α decreases, the reduction in \mathcal{E}_b/N_0 becomes apparent. Therefore, it is clear that linear passband-to-baseband is very sub-optimal in non-Gaussian AWS α SN.

B. Simulations & Results

In this section, we compare the symbol error rate (SER) performance of the different receiver schemes discussed in this paper. All results are generated via Monte Carlo simulations. The error curves are compiled for at least 5000 errors at high SER ($\geq 10^{-3}$) and at least 1000 errors at low SER ($< 10^{-3}$).

We plot the SER for *various baseband detectors* in the *discretized linear receiver* for $\alpha = 1$ and $\alpha = 1.5$ in Fig. 6 and Fig. 7, respectively. The 8-QAM constellation in Fig. 5 is employed. For comparison we have also plotted the SER for the *conventional receiver* with *Euclidean detection* for the same constellation. An analytical expression for \mathcal{E}_b/N_0 of the conventional receiver is given in Appendix B. The increase in performance due to sampling at $f_s = 4f_c$ and invoking optimized constellations and detectors over the conventional receiver is clear. The myriad, log-norm and L_p -norm (as $p \rightarrow 0$) detectors perform very well. We use (48) as the linearity parameter for the myriad detector. To emphasize the importance of constellation design, we have also plotted results of the myriad detector for the well-known 8-QAM rectangular map for both the Cauchy and $\alpha = 1.5$ cases. For Cauchy

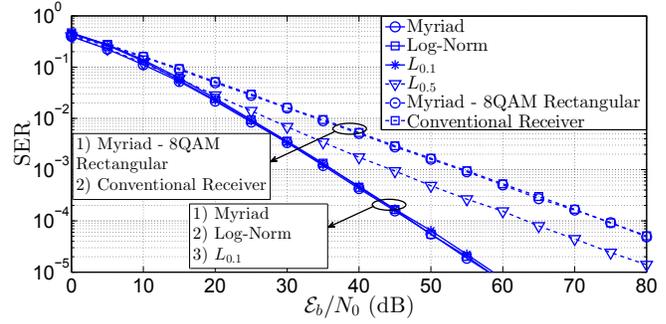


Fig. 6. SER for various receiver schemes in Cauchy AWS α SN for $M = 8$.

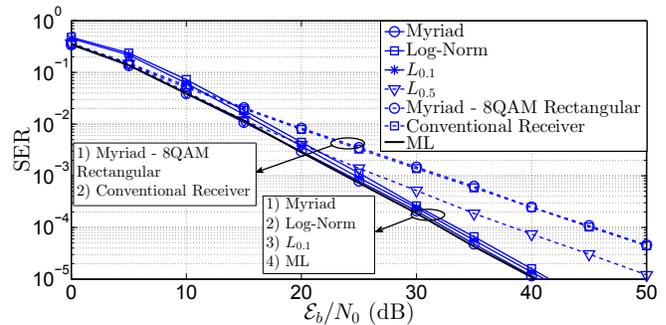


Fig. 7. SER for various receiver schemes in AWS α SN with $\alpha = 1.5$ and $M = 8$.

noise, myriad detection corresponds to ML detection. We have additionally presented the ML detector performance for the $\alpha = 1.5$ case in Fig. 7. As $\alpha \rightarrow 2$, the SER performance between the discretized linear receiver with ML detection and its conventional counterpart converge. In Figs. 6 and 7, we also note that the L_p -norm detector for small p performs well in the medium-to-high \mathcal{E}_b/N_0 regime, however, as predicted in Section, higher values of p offer better results when \mathcal{E}_b/N_0 is low.

We present the *joint-detector* performance for different non-linear receivers in Cauchy AWS α SN for $Tf_s = 40$ in Fig. 8 and $Tf_s = 400$ in Fig. 9. All plots are generated for $g(t)$ a rectangular pulse and the 8-QAM constellation in Fig. 5. For comparison, we have also plotted the Gaussian error curve for the same constellation in both figures. We use (48) as the linearity parameter for myriad detection. As the performance of joint-detection for a given \mathcal{E}_b/N_0 can be made arbitrarily better than that of the linear receiver by increasing ξ (and consequently reducing the pseudo-PSD), we therefore plot the SER against

$$\text{SNR (dB)} = 10 \log_{10} \frac{E[\mathcal{E}_{s_i}] f_s}{4\delta^2 \log_2 M}, \quad (70)$$

which is the first term on the right in (68). Though (70) is essentially \mathcal{E}_b/N_0 with N_0 written in terms of passband parameters, i.e., $N_0 = 4\delta^2/f_s$, we term it as SNR to differentiate it from the definition in (68). To compare the results in Figs. 8 and 9 with those of the conventional/discretized linear receivers, one needs only to shift the curves in Figs. 6 and 7 by $20 \log_{10} d(\alpha, \xi, g[n])$ to the right. One can clearly appreciate

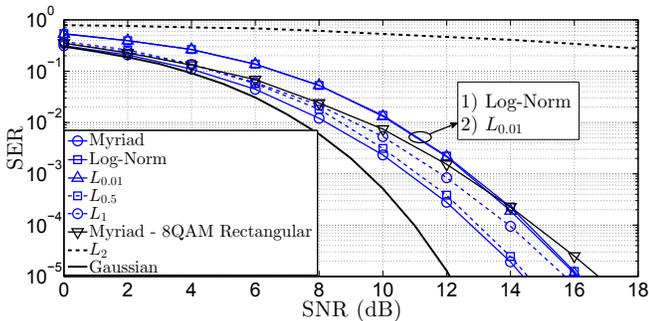


Fig. 8. SER for various joint-detection schemes in Cauchy AWS α SN with $Tf_s = 40$ and $M = 8$.

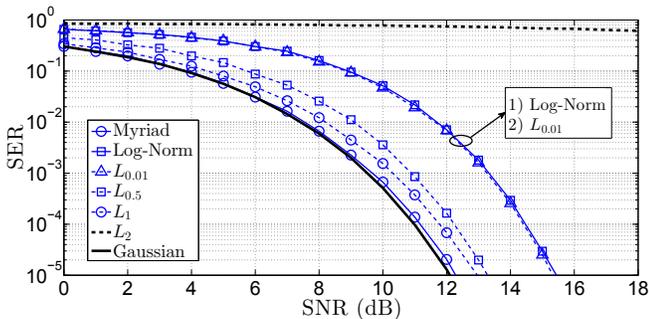


Fig. 9. SER for various joint-detection schemes in Cauchy AWS α SN with $Tf_s = 400$ and $M = 8$.

the performance of the non-linear receivers over their linear counterparts.

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 have tested this empirically for even larger values of Tf_s . Increasing Tf_s essentially expands the *available bandwidth* f_s relative to the symbol rate $1/T$. To show the effect of constellation design, we have plotted the joint-myriad detection performance for the 8-QAM rectangular map for $Tf_s = 40$ in Fig. 8. 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 rectangular counterpart.

In Fig. 10 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 constraint of $f_s = 4f_c$. These results can be compared to their joint-detector counterparts in Fig. 8 and Fig. 9. The isotropic assumption of \vec{w} 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, the performance for the $Tf_s = 400$ case is similar to the corresponding joint-detection scheme.

VII. DISCUSSION & CONCLUSION

In this paper we have discussed and analyzed features of a good communications receiver for single-carrier modulation

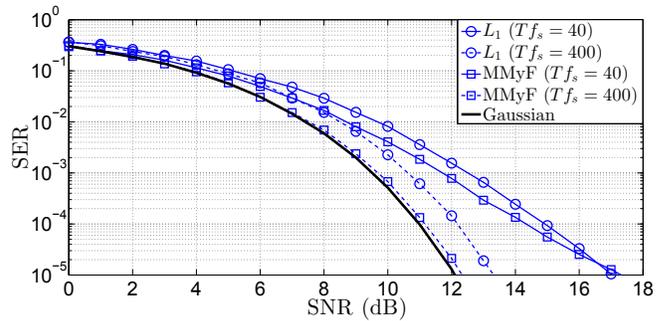


Fig. 10. SER for various receiver schemes in Cauchy AWS α SN with $M = 8$.

in impulsive noise. The analysis covers several schemes under the following design methodologies: Soft-Estimates followed by Baseband Detection and Joint-Detection.

The conventional (continuous-time) receiver was shown to perform poorly in non-Gaussian AWS α SN. 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 AWS α SN is that they are far from optimal [8]. We have categorized this mathematically as SNR degradation at the receiver. In the literature, impulsive noise is typically addressed at the baseband level. This approach implicitly assumes a linear passband-to-baseband conversion block which is sub-optimal (as indicated by the SNR loss it incurs). So no matter what is proposed at the baseband, there always will be a reduction in SNR due to linear baseband conversion.

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 latter 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 LLR for subsequent transmission.

From a practical perspective, implementing passband sampling may not be possible. For example, in wireless communications the passband signal is transmitted via radio frequency (RF) waves operating at high-frequencies (hundreds/thousands of MHz) [14]. Sampling at the Nyquist rate in such scenarios will result in costly hardware implementation. However, in underwater communications, signal transmission is done via acoustics as RF waves face severe attenuation [1]. Typical ranges for carrier frequencies run into tens of kHz, a significant difference from those adopted in wireless communications. Sampling the passband signal is therefore practically feasible and is employed in some underwater modems, where f_s can be set up to 500 kHz [3], [38]. In underwater modems, such as the one discussed in [39], carrier frequencies up to 78 kHz can be used with varying signal bandwidth. As long as synchronization is achieved, the $f_s = 4f_c$ condition, can be met. For example if $f_c = 50$ kHz, then $f_s = 200$ kHz satisfies $f_s = 4f_c$. A bandwidth of $1/T = 5$ kHz will satisfy the $Tf_s = 40$ kHz constraint.

APPENDIX A

A. Baseband Noise in the Conventional Receiver

If the schematic in Fig. 1 is employed in passband AWS α SN, then \vec{w} is a bivariate isotropic S α S vector. These properties are proven below:

Proposition 2: If $w(t)$ is a continuous-time real AWS α SN process and the conventional receiver is employed, then \vec{w} is isotropic in the limit $f_c \rightarrow \infty$.

Proof: We can express \vec{w} as

$$w_I + jw_Q = \int_0^T w(t)(\ell_I(t) + j\ell_Q(t))dt \quad (71)$$

$$= \sqrt{\frac{2}{\mathcal{E}_g}} \int_0^T w(t)g(t)e^{j2\pi f_c t} dt. \quad (72)$$

The integral in (72) may be divided into a sum of ξ integrals over adjacent intervals:

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int_{\lambda T/\xi}^{(\lambda+1)T/\xi} w(t)g(t)e^{j2\pi f_c t} dt. \quad (73)$$

We note that $\exp(j2\pi f_c t)$ is periodic in t with period T/ξ , thus we may express (73) as

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int_{\lambda T/\xi}^{(\lambda+1)T/\xi} w(t)g(t)e^{j2\pi f_c (t-\lambda T/\xi)} dt. \quad (74)$$

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

$$\sqrt{\frac{2}{\mathcal{E}_g}} \sum_{\lambda=0}^{\xi-1} \int_0^{T/\xi} w\left(t + \frac{\lambda T}{\xi}\right) g\left(t + \frac{\lambda T}{\xi}\right) e^{j2\pi f_c t} dt. \quad (75)$$

The integration and summation operations may be interchanged to get

$$\sqrt{\frac{2}{\mathcal{E}_g}} \int_0^{T/\xi} \sum_{\lambda=0}^{\xi-1} w\left(t + \frac{\lambda T}{\xi}\right) g\left(t + \frac{\lambda T}{\xi}\right) e^{j2\pi f_c t} dt. \quad (76)$$

From (1) and (2), we note that

$$\sum_{\lambda=0}^{\xi-1} w\left(t + \frac{\lambda T}{\xi}\right) g\left(t + \frac{\lambda T}{\xi}\right) \stackrel{d}{=} w(t)c(t; \alpha, \xi, g(t)), \quad (77)$$

where

$$c(t; \alpha, \xi, g(t)) = \left(\sum_{\lambda=0}^{\xi-1} \left| g\left(t + \frac{\lambda T}{\xi}\right) \right|^\alpha \right)^{1/\alpha}. \quad (78)$$

In (78), the expression within brackets corresponds to *sampling* $|g(t)|^\alpha$ at a rate of $f_c = \xi/T$ and summing the terms. Therefore, by using a Riemann sum argument, (78) converges to

$$c(t; \alpha, \xi, g(t)) \rightarrow c(\alpha, \xi, g(t)) = \left(\frac{\xi}{T} \int_0^T |g(t)|^\alpha dt \right)^{1/\alpha} \quad (79)$$

as $f_c \rightarrow \infty$. Finally, from (77) and (79), we may express (76) as

$$w_I + jw_Q \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \int_0^{T/\xi} w(t)e^{j2\pi f_c t} dt. \quad (80)$$

Now for \vec{w} to be isotropic,

$$w_I + jw_Q \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \int_0^{T/\xi} w(t)e^{j(2\pi f_c t + \phi)} dt \quad (81)$$

$\forall \phi \in \mathbb{R}$. We note that the phasors $\exp(j2\pi f_c t)$ in (80) and $\exp(j2\pi f_c t + j\phi)$ in (81) 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 (80) and (81) are statistically equivalent, thus \vec{w} is isotropic. \square

Though Proposition 1 is valid for $f_c \rightarrow \infty$, this condition can be somewhat relaxed for any band-limited $g(t)$. For example, if $g(t)$ is a rectangular pulse then (78) and (79) are both equal to $\xi^{1/\alpha}$ for all f_c . For general $g(t)$, f_c needs to be greater than the Nyquist rate of $|g(t)|^\alpha$ for (79) to hold. Two cases are highlighted below:

- For $\alpha = 2$, The Fourier transform of $g^2(t)$ is twice that of $g(t)$. When sampled at f_c , the Nyquist criterion is satisfied for $\xi > 2\beta$.
- As $\alpha \rightarrow 0$, $|g(t)|^\alpha$ tends to the unit-amplitude rectangular pulse and therefore (79) is exact for any $\xi \in \mathbb{Z}^+$.

Considering the above cases and the fact that $\alpha \in (0, 2]$, one would expect ξ to be only a few multiples greater than β for \vec{w} to be approximately isotropic. Increasing ξ any further would cause negligible change in the distribution of \vec{w} . To highlight how well (79) approximates (78), we plot $(c(t; \alpha, \xi, g(t))/c(\alpha, \xi, g(t)))^\alpha$ (on the dB scale) for various ξ and random $t \in [0, T/\xi]$ in Fig. 11. For this example, $g(t)$ is

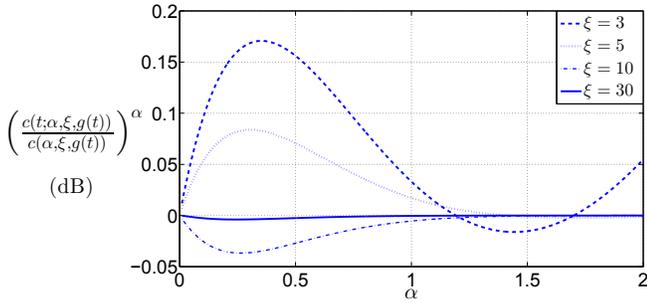


Fig. 11. A measure of similarity for (78) and (79) with $g(t)$ being a unit roll-off raised-cosine pulse.

a raised-cosine pulse with roll-off factor set to one. For (78) and (79) to be equivalent, the plotted measure should be equal to zero. One can clearly see that (79) offers an increasingly better approximation as ξ increases for any α .

Note that the formulation in (72)-(81) may be applied to any IID noise process (albeit with some modifications), irrespective of its samples following a symmetric or asymmetric distribution.

Proposition 3: If $w(t)$ is a continuous-time real AWS α SN process, then \vec{w} is a S α S vector in the conventional receiver.

Proof: We may write (71) in vector form:

$$\vec{w} = \begin{bmatrix} \int_0^T w(t) \ell_I(t) dt \\ \int_0^T w(t) \ell_Q(t) dt \end{bmatrix}. \quad (82)$$

Let $\vec{w}^{(i)} \forall i \in \{1, 2, \dots, N\}$ be IID copies of \vec{w} , then from (4) we have

$$\begin{aligned} \sum_{i=1}^N a_i \vec{w}^{(i)} &= \sum_{i=1}^N a_i \begin{bmatrix} \int_0^T w^{(i)}(t) \ell_I(t) dt \\ \int_0^T w^{(i)}(t) \ell_Q(t) dt \end{bmatrix} \\ &= \begin{bmatrix} \int_0^T \left(\sum_{i=1}^N a_i w^{(i)}(t) \right) \ell_I(t) dt \\ \int_0^T \left(\sum_{i=1}^N a_i w^{(i)}(t) \right) \ell_Q(t) dt \end{bmatrix}, \end{aligned} \quad (83)$$

where $w^{(i)}(t) \forall i \in \{1, 2, \dots, N\}$ are IID copies of $w(t) \sim \mathcal{S}(\alpha, \delta)$. On invoking (1), we get

$$\sum_{i=1}^N a_i \vec{w}^{(i)} \stackrel{d}{=} \begin{bmatrix} \int_0^T c w(t) \ell_I(t) dt \\ \int_0^T c w(t) \ell_Q(t) dt \end{bmatrix} = c \vec{w}, \quad (84)$$

where c may be evaluated from (2):

$$c = \left(\sum_{i=1}^N |a_i|^\alpha \right)^{1/\alpha}. \quad (85)$$

Therefore, we conclude that \vec{w} is S α S with characteristic exponent α and scale $c\delta$. \square

From the discussion in Section II-C, the components of isotropic \vec{w} for $\alpha \neq 2$ are identically distributed but dependent. We evaluate the scale parameters of w_I and w_Q as a function of the bandwidth next.

B. Scale parameters of \vec{w} in (80)

Due to the linearity of the receiver, $w_I \stackrel{d}{=} w_Q \sim \mathcal{S}(\alpha, \delta_w)$ in AWS α SN. From (80), we have

$$w_I \stackrel{d}{=} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \int_0^{T/\xi} w(t) \cos(2\pi f_c t) dt. \quad (86)$$

By approximating the integration term with a limiting Riemann sum, we get

$$w_I \stackrel{d}{=} \frac{1}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \sum_{n=0}^{\lfloor T f_s / \xi \rfloor - 1} w(n/f_s) \cos(2\pi f_c / f_s n) \quad (87)$$

as $f_s \rightarrow +\infty$. In this formulation, f_s is the passband sampling frequency of the AWS α SN channel. As $w(n/f_s) \sim \mathcal{S}(\alpha, \delta) \forall n \in \{0, 1, \dots, \lfloor T f_s / \xi \rfloor - 1\}$, then from (2) and (87),

$$\begin{aligned} \delta_w &= \frac{\delta}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \left(\sum_{n=0}^{\lfloor T f_s / \xi \rfloor - 1} |\cos(2\pi f_c / f_s n)|^\alpha \right)^{\frac{1}{\alpha}} \\ &\approx \frac{\delta}{f_s^{(1-\frac{1}{\alpha})}} \sqrt{\frac{2}{\mathcal{E}_g}} c(\alpha, \xi, g(t)) \left(\int_0^{\frac{T}{\xi}} |\cos(2\pi f_c t)|^\alpha dt \right)^{\frac{1}{\alpha}}. \end{aligned} \quad (88)$$

The approximation in (88) is justified as $f_s \rightarrow +\infty$. As

$$\int_0^{\frac{T}{\xi}} |\cos(2\pi f_c t)|^\alpha dt = \frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi} \Gamma(1 + \frac{\alpha}{2})} \quad (89)$$

and $c(\alpha, \xi, g(t))$ is given in (79), we may express (88) as

$$\delta_w = \frac{\delta}{f_s^{(1-\frac{1}{\alpha})}} \sqrt{\frac{2}{\mathcal{E}_g}} \left(\int_0^T |g(t)|^\alpha dt \right)^{\frac{1}{\alpha}} \left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi} \Gamma(1 + \frac{\alpha}{2})} \right)^{\frac{1}{\alpha}}. \quad (90)$$

Using the same approach one can easily evaluate (90) from w_Q instead of w_I in (86).

APPENDIX B

A. SNR Derivation for the Conventional Receiver

As the reference SNR is defined in (60) for the discretized linear receiver, we need to express it in terms of δ_w in the conventional receiver. We slightly abuse notation by equating (90) to δ_c . We reserve δ_w for the baseband scale parameter in the discretized linear receiver. As ξ is assumed to be large, we may express δ_w in (29) as

$$\begin{aligned} \delta_w &= \frac{\delta}{f_s} \left(\sum_{n=0}^{4\xi-1} |\ell_I[n]|^\alpha \right)^{\frac{1}{\alpha}} = \frac{\delta}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} \left(\sum_{n=0}^{2\xi-1} |g[2n]|^\alpha \right)^{\frac{1}{\alpha}} \\ &\approx \frac{\delta}{f_s} \sqrt{\frac{2}{\mathcal{E}_g}} \left(\frac{f_s}{2} \int_0^T |g(t)|^\alpha dt \right)^{1/\alpha} \\ &= \frac{\delta}{f_s^{(1-\frac{1}{\alpha})}} \frac{2^{(\frac{1}{2}-\frac{1}{\alpha})}}{\sqrt{\mathcal{E}_g}} \left(\int_0^T |g(t)|^\alpha dt \right)^{1/\alpha}. \end{aligned} \quad (91)$$

On dividing (91) by (90) and simplifying, we get

$$\delta_w = \delta_c 2^{-\frac{1}{\alpha}} \left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})} \right)^{-1/\alpha}. \quad (92)$$

Finally, we substitute (92) in (60)

$$\frac{\mathcal{E}_b}{N_0} = \frac{E[\mathcal{E}_{s_i}]}{4\delta_c^2 \log_2 M} \times 2^{\frac{2}{\alpha}} \left(\frac{\Gamma(\frac{1+\alpha}{2})}{\sqrt{\pi}\Gamma(1+\frac{\alpha}{2})} \right)^{2/\alpha}. \quad (93)$$

REFERENCES

- [1] J. Proakis and M. Salehi, *Digital Communications*, ser. McGraw-Hill higher education. McGraw-Hill Education, 2007.
- [2] M. Chitre, S. Kuselan, and V. Pallayil, "Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation," *The J. of the Acoustical Soc. of Amer.*, vol. 132, no. 2, 2012.
- [3] M. Chitre, J. Potter, and S.-H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," *IEEE J. Ocean. Eng.*, vol. 31, no. 2, pp. 497–503, April 2006.
- [4] M. Chitre, S. Ong, and J. Potter, "Performance of coded ofdm in very shallow water channels and snapping shrimp noise," in *OCEANS, 2005. Proc. of MTS/IEEE*, 2005, pp. 996–1001 Vol. 2.
- [5] F. Horner and J. Harwood, "An investigation of atmospheric radio noise at very low frequencies," *Proc. of the IEE - Part B: Radio and Electron. Eng.*, vol. 103, no. 12, pp. 743–751, 1956.
- [6] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Energy Convers.*, vol. 44, no. 1, pp. 249–258, Feb 2002.
- [7] N. Nedeve and S. McLaughlin, *Analysis of the Impact of Impulse Noise in Digital Subscriber Line Systems*. Univ. of Edinburgh, 2003.
- [8] C. L. Nikias and M. Shao, *Signal processing with Alpha-Stable Distributions and Applications*. New York: Chapman-Hall, 1996.
- [9] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.
- [10] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 145–147, Feb. 1996.
- [11] M. Chitre, J. Potter, and S. Ong, "Viterbi decoding of convolutional codes in symmetric α -stable noise," *IEEE Commun. Lett.*, vol. 55, no. 12, pp. 2230–2233, Dec. 2007.
- [12] A. Mahmood, M. Chitre, and M. A. Armand, "PSK communication with passband additive symmetric α -stable noise," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2990–3000, Oct. 2012.
- [13] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, 1994.
- [14] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [15] M. Chitre, J. Potter, and O. S. Heng, "Underwater acoustic channel characterisation for medium-range shallow water communications," in *OCEANS '04. MTS/IEEE TECHNO-OCEAN '04*, vol. 1, Nov. 2004, pp. 40–45.
- [16] A. Mahmood, M. Chitre, and M. A. Armand, "Improving PSK performance in snapping shrimp noise with rotated constellations," in *Proc. of the Seventh ACM Int. Conf. on Underwater Networks and Systems*, ser. WUWNet '12, no. 12, New York, NY, USA, 2012, pp. 1–8.
- [17] G. Tsihrintzis and C. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 904–914, 1995.
- [18] T. Saleh, I. Marsland, and M. El-Tanany, "Suboptimal detectors for alpha-stable noise: Simplifying design and improving performance," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2982–2989, 2012.
- [19] A. Mahmood, M. Chitre, and M. A. Armand, "Baseband characterization of additive white symmetric α -stable noise," in *IEEE Global Telecommunications Conf. (GLOBECOM)*, Dec. 2012, pp. 3696–3701.
- [20] J. P. Nolan, *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhauser, 2012, in progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- [21] M. W. Legg, *Non-Gaussian and non-homogeneous Poisson models of snapping shrimp noise*. Curtin Univ. of Technology., 2009.
- [22] B. Sklar, *Digital communications: fundamentals and applications*, ser. Prentice Hall Communications Engineering and Emerging Technologies Series. Prentice-Hall PTR, 2001.
- [23] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*. Wiley-Interscience, Jul. 2006.
- [24] P. Huber and E. Ronchetti, *Robust Statistics*, ser. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 2009.
- [25] 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.
- [26] J. G. Gonzalez, "Robust techniques for wireless communications in non-gaussian environments," Ph.D. dissertation, Univ. of Delaware, Newark, DE, USA, 1997.
- [27] J. Gonzalez and G. Arce, "Optimality of the myriad filter in practical impulsive-noise environments," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 438–441, Feb 2001.
- [28] V. Chavali and C. da Silva, "Detection of digital amplitude-phase modulated signals in symmetric alpha-stable noise," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3365–3375, 2012.
- [29] E. E. Kuruoglu, 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.
- [30] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
- [31] T. Aysal and K. Barner, "Meridian filtering for robust signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 3949–3962, 2007.
- [32] R. E. Carrillo, T. Aysal, and K. Barner, "Generalized cauchy distribution based robust estimation," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE Int. Conf. on*, 2008, pp. 3389–3392.
- [33] J. P. Nolan, "Numerical calculation of stable densities and distribution functions," *Communications in Stat. Stochastic Models*, vol. 13, no. 4, pp. 759–774, 1997.
- [34] J. H. McCulloch, *Numerical approximation of the symmetric stable distribution and density*. Cambridge, MA, USA: Birkhauser Boston Inc., 1998, pp. 489–499.
- [35] R. Storn and K. Price, "Differential evolution a simple and efficient heuristic for global optimization over continuous spaces," *J. of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [36] V. Černý, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *J. of Optimization Theory and Applications*, vol. 45, no. 1, pp. 41–51, 1985.
- [37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [38] 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. Proc.*, vol. 5, Sept 2003, pp. 2706–2713 Vol.5.
- [39] M. Chitre, I. Topor, and T. Koay, "The UNET-2 modem – an extensible tool for underwater networking research," in *OCEANS, 2012 - Yeosu*, May 2012, pp. 1–7.



Ahmed Mahmood (GM'12) received his B.Eng and M.Eng degrees in electrical engineering from the National University of Sciences and Technology (NUST), Pakistan, and a Ph.D. degree from the Department of Electrical & Computer engineering, National University of Singapore (NUS), Singapore. In 2014, Dr. Mahmood joined the Acoustic Research Laboratory (ARL) at NUS, where he is now a Research Fellow. He has served as a reviewer for the IEEE Transactions on Communications and IET Communications. His current research interests in-

volve underwater acoustic communications, robust signal processing, channel modelling, estimation & detection theory and error correction coding for impulsive noise.



Mandar Chitre (M'03–SM'12) received B.Eng. and M.Eng. degrees in electrical engineering from the National University of Singapore (NUS), Singapore, a M.Sc. degree in bioinformatics from the Nanyang Technological University (NTU), Singapore, and a Ph.D. degree from NUS. From 1997 to 1998, he worked with the ARL, NUS in Singapore. From 1998 to 2002, he headed the technology division of a regional telecommunications solutions company. In 2003, he rejoined ARL, initially as the Deputy Head (Research) and is now the Head of the laboratory.

Dr. Chitre also holds a joint appointment with the Department of Electrical & Computer Engineering at NUS as an Assistant Professor. His current research interests are underwater communications, autonomous underwater vehicles and acoustic signal processing.

Dr. Chitre has served on the technical program committees of the IEEE OCEANS, WUWNet, DTA and OTC conferences and has served as reviewer for numerous international journals. He was the chairman of the student poster committee for IEEE OCEANS06 in Singapore, and the chairman for the IEEE Singapore AUV Challenge 2013. He is currently the IEEE Ocean Engineering Society technology committee co-chair of underwater communication, navigation & positioning.



Marc A. Armand (M'00-SM'07) received his B.Eng and Ph.D. degree in electrical engineering from the University of Bristol, U.K., in 1995 and 1999, respectively. He is currently an Associate Professor at the Department of Electrical and Computer Engineering, National University of Singapore, which he joined in Dec 2000 as an Assistant Professor. He is a Senior Member of the IEEE and has published over 50 journal and conference papers in the area of coding theory and techniques, as well as detection, coding and channel modelling for

magnetic recording and underwater communications.