

# Optimal and Near-Optimal Detection in Bursty Impulsive Noise

Ahmed Mahmood, *Member, IEEE*, and Mandar Chitre, *Senior Member, IEEE*

**Abstract**—In many practical scenarios, the ambient noise process is known to be impulsive. To combat this, several robust measures have been proposed in the literature. Most of them assume white noise processes, i.e., the noise samples are independent and identically distributed heavy-tailed random variables. However, noise is seldom white in practice and therefore exhibits memory. For impulsive noise, dependency among samples results in outliers clustering together. The process is thus impulsive and bursty. In our work, we employ stationary  $\alpha$ -sub-Gaussian noise with memory order  $m$  ( $\alpha$ SGN( $m$ )) to model bursty impulsive noise. The model is based on the multivariate  $\alpha$ -sub-Gaussian ( $\alpha$ SG) distribution family and statistically characterizes adjacent samples from elliptical distributions. The latter assumption holds well for snapping shrimp noise found in warm shallow underwater channels. We investigate the performance of conventional robust detectors in  $\alpha$ SGN( $m$ ) and also propose novel near-optimal detectors. The Neyman–Pearson (NP) approach for binary hypothesis testing is considered and extensive simulation results for the aforementioned detectors are offered. For all instances, we employ an  $\alpha$ SGN( $m$ ) process whose parameters are tuned to snapping shrimp noise data sets. By incorporating good signal design rules, it is shown that there is a large performance gap between the new and conventional detectors in various impulsive regimes. Moreover, it is possible to derive a near-optimal detector if one only has information of the temporal statistics of the noise process.

**Index Terms**—Impulsive noise, bursty impulsive noise,  $\alpha$ SG,  $\alpha$ SGN( $m$ ), Neyman–Pearson formulation.

## I. INTRODUCTION

GAUSSIAN processes are typically used to model the ambient noise in a variety of scenarios [1], [2]. However, in several instances, the noise process is impulsive in nature [3]–[5]. Signal processing techniques optimized for Gaussian noise are remarkably inept to handle outliers and may drastically reduce system performance even when the noise process is slightly impulsive [6], [7]. Therefore, a significant amount of literature has been devoted to finding measures robust to impulsive noise [6], [8]–[10].

Manuscript received January 20, 2016; revised August 18, 2016; accepted August 22, 2016. Date of publication October 27, 2016; date of current version July 12, 2017.

Associate Editor: J. Potter.

A. Mahmood is with the Acoustic Research Laboratory (ARL), Tropical Marine Science Institute, National University of Singapore, Singapore 119222, Singapore (e-mail: ahmed.mahmood@u.nus.edu).

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

Digital Object Identifier 10.1109/JOE.2016.2603790

In warm shallow waters around the globe, snapping shrimp noise tends to dominate the high-frequency spectrum [4], [5], [11]. The noise process is known to be impulsive and its amplitude distribution is modeled well by the family of heavy-tailed symmetric  $\alpha$ -stable ( $S\alpha S$ ) distributions [5], [11]. The parameter  $\alpha$  controls the heaviness of the tails and allows tuning the model to exhibit various degrees of impulsiveness [12], [13]. The motivation for using  $S\alpha S$  distributions stems from the generalized central limit theorem (GCLT), which is essentially the central limit theorem (CLT) but with its power constraint removed [12]. In the literature, a typical approach in modeling impulsive noise is to assume it is white, i.e., the samples are independent and identically distributed (i.i.d.) random variables [6], [9]–[11]. If the samples are i.i.d.  $S\alpha S$  random variables, the process is that of white  $S\alpha S$  noise ( $WS\alpha SN$ ). The  $WS\alpha SN$  model has been commonly employed for snapping shrimp noise and various other noise/interference scenarios [5], [6], [14]–[16].

In practice, however, noise is seldom white. The white assumption only incorporates the amplitude distribution of the noise process and does not take into account the dependence between samples [6], [11]. If the noise is impulsive, the outliers tend to cluster together due to the implicit memory of the process, thus making received observations bursty (colored). Various works show the snapping shrimp noise process to be impulsive and bursty as well [5], [17], [18]. Therefore, techniques that take only the noise impulsiveness into account and not its memory will be suboptimal. It is, therefore, worth investigating the performance degradation caused by not taking the memory of an impulsive noise process into account.

In our work, we employ the stationary  $\alpha$ -sub-Gaussian noise with memory order  $m$  ( $\alpha$ SGN( $m$ )) model for bursty impulsive noise [18]. The aforementioned model is a recent attempt to characterize the dependency within adjacent samples of impulsive noise processes as elliptical distributions, while also constraining the amplitude distribution to be  $S\alpha S$ . The motivation for the  $\alpha$ SGN( $m$ ) model specifically arose from the analysis of snapping shrimp noise data sets, as they offer near-elliptical delay scatter plots [18]. Moreover,  $WS\alpha SN$  is a special case of  $\alpha$ SGN( $m$ ) ( $m = 0$ ), thus making the latter a more generic model to work with.

We now highlight the contributions of this paper. The performance of commonly employed conventional robust detectors in  $\alpha$ SGN( $m$ ) is investigated. We define a conventional (or white) detector as one that is optimized for i.i.d. noise samples. Results are compiled for the binary hypothesis problem within the

Neyman–Pearson (NP) framework. The white log-likelihood ratio (LLR) detector is analyzed along with its variants based on M-estimation theory, namely, the myriad detector (MyD) and the geometric mean detector (GMD) [8], [9]. For the  $\alpha$ SGN( $m$ ) model, we derive the LLR detector (optimal) and extend concepts of M-estimation theory to incorporate the dependence between samples. The novel detectors are shown to significantly outperform their white counterparts in  $\alpha$ SGN( $m$ ) tuned to snapping shrimp noise for varying degrees of impulsiveness. We also highlight good design rules for the transmitted signal. Extensive simulation results clearly show that near-optimal detection is possible by just compensating for the elliptical dependence between closely spaced samples. Similar trends are also observed when  $\alpha$ SGN( $m$ ) is replaced by actual snapping shrimp noise.

The layout of this paper is as follows. In Section II, we present fundamentals of the  $\alpha$ -stable framework that are essential to understand the  $\alpha$ SGN( $m$ ) process. In Section III, we discuss the  $\alpha$ SGN( $m$ ) model and the NP binary hypothesis problem. This is followed by a discussion on conventional robust measures in Section IV. We then derive the LLR detector and its near-optimal variants for general  $\alpha$ SGN( $m$ ) in Section V. Thereafter, a discussion on optimal and suboptimal signal design is done in Section VI. We wrap up by presenting our simulation results in Section VII and the conclusions in Section VIII.

## II. FUNDAMENTAL CONCEPTS

Stable random variables are unique in the sense that they alone exhibit the stability property [12], [13], which states that the distribution of any linear combination of i.i.d. stable random variables is conserved up until location and scale. Mathematically, if  $X^{(k)} \forall k \in \{1, 2, \dots, N\}$  are i.i.d. copies of a random variable  $X$ , then

$$\sum_{k=1}^N a_k X^{(k)} \stackrel{d}{=} cX + d \quad (1)$$

where  $a_k, c, d \in \mathbb{R}$  and  $N \in \mathbb{Z}^+$ , holds true if and only if  $X$  is stable [12], [13]. Symbol  $\stackrel{d}{=}$  implies equality in distribution. As the left-hand side of (1) is essentially a weighted sum of i.i.d. random variables, it is apparent that the well-known Gaussian distribution is also a member of the stable family [1], [12], [13]. With the exception of the former, all stable distributions are heavy tailed (algebraic tailed) and have been used in many instances to develop models for impulsive noise/interference scenarios [6], [11], [12].

The probability density function (pdf) of a stable random variable is completely parameterized by the characteristic exponent  $\alpha \in (0, 2]$ , skew  $\beta \in [-1, 1]$ , scale  $\delta \in \mathbb{R}^+$ , and location  $\mu \in \mathbb{R}$  [12], [13]. Consequently, a stable distribution may be denoted in terms of its parameters by  $\mathcal{S}(\alpha, \beta, \delta, \mu)$  [6]. Of these,  $\alpha$  controls the heaviness of the tails. Mathematically, if  $f_X(x)$  is the pdf of  $X$  (with outcome  $x$ ), then

$$f_X(x) \propto |x|^{-1-\alpha} \quad (2)$$

as  $|x| \rightarrow \infty$  for  $\alpha \neq 2$ . From (2), it is apparent that as  $\alpha \rightarrow 2$ , the tails get increasingly lighter. In fact, for  $\alpha = 2$ ,  $f_X(x)$  is no longer dependent on  $\beta$  and corresponds to the pdf of a

Gaussian random variable with mean  $\mu$  and variance  $2\delta^2$ , i.e.,  $\mathcal{N}(\mu, 2\delta^2)$  [12], [13]. If  $\beta$  and  $\mu$  are equated to zero, a stable distribution is termed S $\alpha$ S and may be denoted succinctly by  $\mathcal{S}(\alpha, \delta)$  [6]. As highlighted by its name, the pdf of an S $\alpha$ S random variable is symmetric in its argument. Specifically, if  $X$  is S $\alpha$ S, then  $f_X(x) = f_X(-x)$ . For the Gaussian case we note that  $\mathcal{S}(2, \delta) \stackrel{d}{=} \mathcal{N}(0, 2\delta^2)$ .

The stability property in (1) extends to stable random vectors as well and can be mathematically expressed in a similar form. If  $\vec{X} \in \mathbb{R}^m$  is a stable random vector and  $\vec{X}^{(k)} \forall i \in \{1, 2, \dots, N\}$  are its i.i.d. copies, then

$$\sum_{k=1}^N a_k \vec{X}^{(k)} \stackrel{d}{=} c\vec{X} + d \quad (3)$$

where  $a_k, c, d \in \mathbb{R}$ . The random vector  $\vec{X}$  is further termed S $\alpha$ S if its pdf  $f_{\vec{X}}(\mathbf{x})$  is also symmetric in its argument, i.e.,  $f_{\vec{X}}(\mathbf{x}) = f_{\vec{X}}(-\mathbf{x})$ , where  $\mathbf{x}$  is a sample outcome of  $\vec{X}$ . Analogous to the univariate case, if  $\vec{X}$  is Gaussian with location vector  $\boldsymbol{\mu} \in \mathbb{R}^m$  and covariance matrix  $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times m}$ , i.e.,  $\vec{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , then it satisfies (3) and is therefore a stable random vector [12], [13]. Furthermore, if  $\boldsymbol{\mu} = \mathbf{0}$  where  $\mathbf{0}$  is the all-zero vector, then  $\vec{X}$  is S $\alpha$ S as well.

The multivariate  $\alpha$ -sub-Gaussian ( $\alpha$ SG) distribution is a special subclass of the heavy-tailed ( $\alpha \neq 2$ ) stable family, which besides being S $\alpha$ S is also an elliptic distribution [19]. One notes that multivariate Gaussian distributions are elliptic as well [1, Ch. 7]. In fact, if  $\vec{X} \in \mathbb{R}^m$  is  $\alpha$ SG, then it may be expressed as

$$\vec{X} = A^{1/2} \vec{G} \quad (4)$$

where  $A \sim \mathcal{S}((\alpha/2), 1, 2(\cos((\pi\alpha)/4))^{2/\alpha}, 0)$  is a totally right-skewed heavy-tailed stable random variable and  $\vec{G} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$  where  $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times m}$  [12], [13], [19]. Both  $A$  and  $\vec{G}$  are statistically independent of each other. It can be shown that  $f_{\vec{X}}(\mathbf{x})$  depends on the statistic  $\|\boldsymbol{\Sigma}^{-1/2} \mathbf{x}\| = \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}$ , where  $\boldsymbol{\Sigma}^{1/2}$  stems from the Cholesky decomposition of  $\boldsymbol{\Sigma}$ , i.e.,  $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2} (\boldsymbol{\Sigma}^{1/2})^T$  [19]. More formally,  $f_{\vec{X}}(\mathbf{x}) = g(\|\boldsymbol{\Sigma}^{-1/2} \mathbf{x}\|)$  for some  $g(\cdot) \in \mathbb{R}^+$ . If  $\vec{Y}$  is a tuple of any  $n$  elements in  $\vec{X}$  where  $n \leq m$ , then  $\vec{Y}$  is also an  $\alpha$ SG vector [13], [19]. The underlying  $n \times n$  covariance matrix is then determined from the corresponding indices in  $\boldsymbol{\Sigma}$ . Finally, defining  $\boldsymbol{\Sigma} = [\sigma_{ij}]$ , if  $X_i$  is the  $i$ th component of  $\vec{X}$ , then  $X_i \sim \mathcal{S}(\alpha, \sqrt{\sigma_{ii}}) \forall i \in \{1, 2, \dots, m\}$  [19]. We should state that some texts such as [12] and [13] exclude the coefficient 2 from the scale parameter of  $A$ . With this definition,  $X_i \sim \mathcal{S}(\alpha, \sqrt{\sigma_{ii}/2})$ . However, for our work, we will use the former definition.

The concepts presented in this section are sufficient to explain the noise model employed in this paper, which is discussed next.

## III. NOISE MODEL AND PROBLEM FORMULATION

### A. The WS $\alpha$ SN and $\alpha$ SGN( $m$ ) Processes

The heavy-tailed WS $\alpha$ SN model has been used extensively in the literature to model impulsive noise [6], [12]. Analogous to white Gaussian noise (WGN), samples of the WS $\alpha$ SN process are i.i.d.  $\mathcal{S}(\alpha, \delta)$  random variables, with the latter offering impulsive noise realizations for  $\alpha \neq 2$ . Note that for  $\alpha = 2$ , the

$WS\alpha SN$  and WGN are statistically equivalent processes. Due to the i.i.d. assumption,  $WS\alpha SN$  offers a good fit to the amplitude distribution of an impulsive noise process [6]. Various works have highlighted the proximity of the  $WS\alpha SN$  model to snapping shrimp noise, the latter of which occurs in shallow tropical waters around the world [5], [11]. However, noise is seldom white in practice. Snapping shrimp noise is no exception to this, which besides being impulsive, is known to have strong temporal statistics [5], [18]. This results in the process having memory and therefore the corresponding realizations are bursty, i.e., impulses tend to cluster [18]. The  $WS\alpha SN$  fails to characterize the dependency between samples of snapping shrimp noise [5], [18]. Consequently, techniques designed to mitigate  $WS\alpha SN$  will surely not be optimal in practical deployments.

The relatively new  $\alpha SGN(m)$  model was developed to address the shortcomings of the  $WS\alpha SN$  process [18]. Not only does it ensure each sample to be  $S\alpha S$  but it also characterizes the memory in a bursty impulsive noise process [18]. The model is based on a sliding-window framework and constrains any immediately adjacent  $m + 1$  samples to be an  $\alpha SG$  vector. Precisely, let  $W_n \forall n \in \mathbb{Z}$  be the random samples of an  $\alpha SGN(m)$  process at index  $n$ . Then  $\vec{W}_{n,m} = [W_{n-m}, W_{n-m+1}, \dots, W_n]^T$  is an  $(m + 1)$ -dimensional  $\alpha SG$  random vector for all  $n \in \mathbb{Z}$ . From (4)

$$\vec{W}_{n,m} = A_n^{1/2} \vec{G}_{n,m} \quad (5)$$

where  $\vec{G}_{n,m} = [G_{n-m}, G_{n-m+1}, \dots, G_n]^T$  is Gaussian with distribution  $\mathcal{N}(\mathbf{0}, \mathbf{R}_m)$  and  $\mathbf{R}_m \in \mathbb{R}^{(m+1) \times (m+1)}$  [18]. As highlighted by its name, the  $\alpha SGN(m)$  process is stationary. Therefore, both the covariance matrix  $\mathbf{R}_m = [r_{ij}]$  and the distribution of  $A_n$  are independent of  $n$  [18]. Due to the aforementioned framework,  $\mathbf{R}_m$  is also symmetric Toeplitz besides being a positive-semidefinite matrix [18]. Consequently,  $r_{ii} = r_{jj} = \delta^2 \forall i, j \in \{1, 2, \dots, m\}$ , which results in  $W_n \sim \mathcal{S}(\alpha, \delta) \forall n \in \mathbb{Z}$ . We also note that the  $\alpha SGN(m)$  process depicts the Markov property as the conditional distribution of  $W_{n+1}$  given all previous samples is equivalent to that of  $W_{n+1}$  given  $\vec{W}_{n,m-1}$  [1, Ch. 16]. Thus, it is an  $m$  th-order Markov process. The corresponding state space may be defined as all possible outcomes of  $\vec{W}_{n,m-1}$ . Due to the adopted sliding-window framework, we note that the underlying  $G_n$  is essentially a Gaussian autoregressive process of order  $m$  ( $AR(m)$ ), which in turn is also a Markov process [1, Ch. 13], [18]. For the special case of  $m = 0$ , (5) reduces to

$$W_n = A_n^{1/2} G_n \quad (6)$$

where  $W_n \sim \mathcal{S}(\alpha, \delta)$  are i.i.d. samples for all  $n \in \mathbb{Z}$ . Thus, for  $m = 0$ , the  $\alpha SGN(m)$  process is a  $WS\alpha SN$  process [18]. Therefore, the  $\alpha SGN(m)$  model offers us a more generic framework to deal with impulsive noise. On a final note, we see that the model is parametrized by  $m + 3$  variables, namely,  $\alpha$ ,  $m$  and the  $m + 1$  elements of any row/column of  $\mathbf{R}_m$ . Alternatively, one can also parameterize the model with  $\delta$  and  $\hat{\mathbf{R}}_m$  instead of  $\mathbf{R}_m$ , where

$$\hat{\mathbf{R}}_m = \mathbf{R}_m / \delta^2 \quad (7)$$

is the normalized covariance matrix with 1 on its main diagonal.

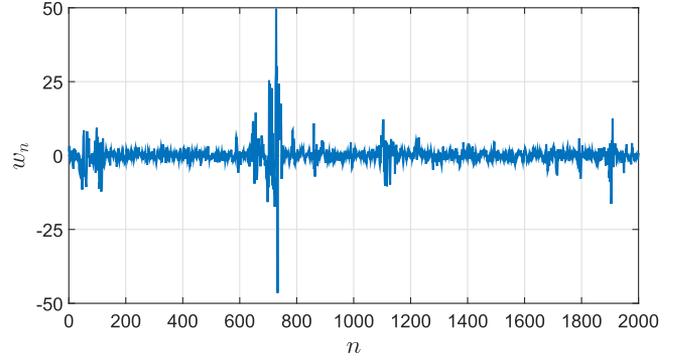


Fig. 1. A realization of  $\alpha SGN(4)$  for  $\alpha = 1.5$ ,  $\delta = 1$ , and  $\hat{\mathbf{R}}_m$  in (8).

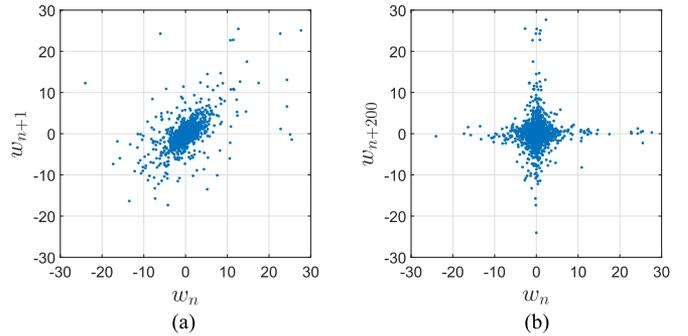


Fig. 2. Scatter plots for delays of (a) 1 and (b) 200, respectively, from the realization in Fig. 1.

As  $\vec{W}_{n,m}$  is an  $\alpha SG$  vector, it is elliptic as well [19]. Consequently, the  $\alpha SGN(m)$  process is particularly adept at characterizing impulsive data sets whose closely spaced samples (within a window of size  $m + 1$ ) have elliptic or near-elliptic distributions. As an example, we present a realization of  $\alpha SGN(4)$  in Fig. 1, denoted by  $w_n$ , and show its delay scatter plots for delays of 1 and 200 in Fig. 2(a) and (b), respectively. We employ  $\alpha = 1.5$ ,  $\delta = 1$ , and

$$\hat{\mathbf{R}}_4 = \begin{bmatrix} 1.0000 & 0.5804 & 0.2140 & 0.1444 & -0.0135 \\ 0.5804 & 1.0000 & 0.5804 & 0.2140 & 0.1444 \\ 0.2140 & 0.5804 & 1.0000 & 0.5804 & 0.2140 \\ 0.1444 & 0.2140 & 0.5804 & 1.0000 & 0.5804 \\ -0.0135 & 0.1444 & 0.2140 & 0.5804 & 1.0000 \end{bmatrix}. \quad (8)$$

In Fig. 2, the unit-delay scatter plot clearly highlights the elliptic projection in the  $W_n - W_{n+1}$  plane. With increasing delay, the projection converges to a four-tailed structure similar to that shown in Fig. 2, which in turn implies heavy-tailed i.i.d. components [6], [20]. The latter is observed for all possible delays when the noise process is white impulsive noise process [6]. For comparison, we have also plotted a realization of snapping shrimp data (normalized to its estimated scale) sampled at 180 kHz in Fig. 3 and its corresponding scatter plots in Fig. 4. Clearly, the realization is that of a bursty impulsive process with its unit delay scatter plot depicting a near-elliptic structure. Moreover, one can see how the  $\alpha SGN(m)$  process accurately models the dependencies between the snapping shrimp

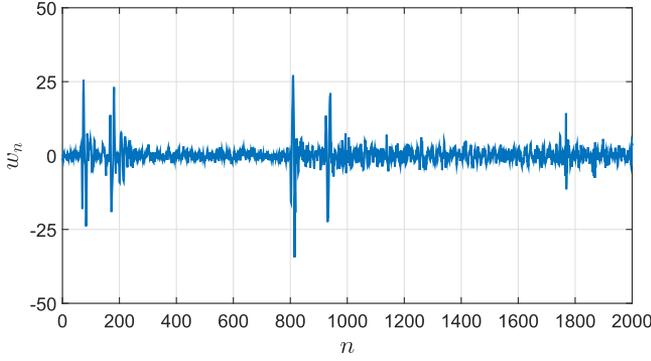


Fig. 3. A realization of snapping shrimp noise sampled at 180 kHz and normalized with its estimated scale.

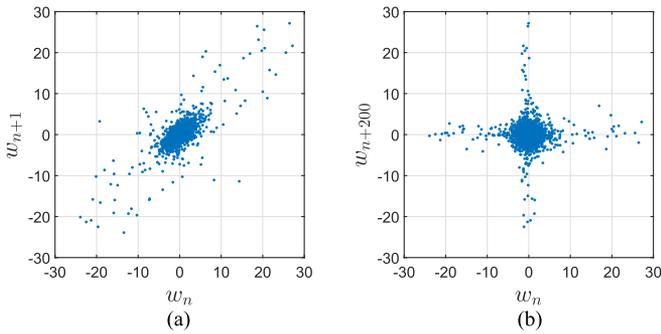


Fig. 4. Scatter plots for delays of (a) 1 and (b) 200, respectively, from the realization in Fig. 3.

noise samples for small and large delays. In fact, the  $\alpha\text{SGN}(m)$  model can be extended to any practical scenario where the ambient noise/interference is impulsive with elliptically distributed adjacent samples. The matrix in (8) is based on an estimate of  $10^6$  samples of snapping shrimp noise sampled at 180 kHz. Details of suitable robust estimators for  $\mathbf{R}_m$ ,  $\alpha$ , and  $\delta$  are presented in [18].

Though the plots in Figs. 1–4 offer compelling visuals that support the  $\alpha\text{SGN}(m)$  framework (for  $m \neq 0$ ) in snapping shrimp noise, a formal statistic needs to be presented that highlights memory in the latter. We revert to the sample covariation coefficient [12, p. 73], which can be evaluated by the fractional-order lower moment (FLOM) estimator

$$\lambda_p(\tau) = \frac{\sum_{i=1}^N W_i |W_{i+\tau}|^{p-1} \text{sign}(W_{i+\tau})}{\sum_{i=1}^N |W_{i+\tau}|^p} \quad (9)$$

where  $\tau \in \mathbb{Z}$  is the delay between samples,  $N$  is the number of considered samples, and  $p \in [1, \alpha)$  guarantees the statistic's robustness if  $W_n$  is an  $S\alpha S$  random variable [12], [18]. Note that if  $\alpha = 2$ , then  $\lambda_2(\tau)$  offers a robust estimate and is essentially the (normalized) sample autocorrelation function (ACF). We set  $p = 1$  to achieve the computationally efficient form

$$\lambda_1(\tau) = \frac{\sum_{i=1}^N W_i \text{sign}(W_{i+\tau})}{\sum_{i=1}^N |W_{i+\tau}|}. \quad (10)$$

The empirical amplitude distribution of snapping shrimp noise is tracked well by an  $S\alpha S$  distribution. Moreover,  $\alpha \geq 1.5$  typically characterizes the impulsiveness found in warm shal-

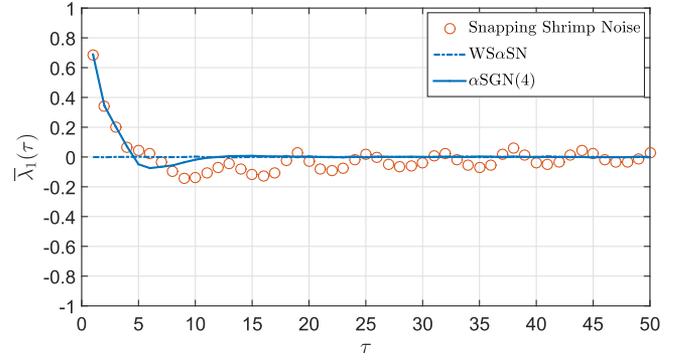


Fig. 5. Covariation versus delay of snapping shrimp noise and  $\alpha\text{SGN}(m)$  realizations for  $\alpha = 1.5$ ,  $\delta = 1$ , and varying  $m$ .

low waters [5]. Consequently, (10) is a robust statistic if  $W_n \forall n \in \mathbb{Z}$  are samples of such data. Within the  $\alpha$ -stable framework,  $\lambda_1(\tau) \neq 0$  if  $W_i$  and  $W_{i+\tau}$  are dependent. On the other hand,  $\lambda_1(\tau) = 0$  is a necessary condition for independence [12]. Consequently,  $\lambda_1(\tau)$  offers insight into the dependence or independence between samples at delay  $\tau$ .

Snapping shrimp noise is a nonstationary process [5]. Though it can be considered stationary within intervals of up to a few minutes, the corresponding  $\lambda_1(\tau)$  statistic is essentially time variant. As the  $\alpha\text{SGN}(m)$  process is stationary, it should be ideally tuned to a large snapping shrimp data set so that it fits the average temporal statistics over  $n$ . Therefore, the average sample covariation coefficient  $\bar{\lambda}_1(\tau)$  is a better metric in this regard. The latter is computed by evaluating  $\lambda_1(\tau)$  over several  $N$ -sample blocks and averaging the results.

In Fig. 5, we plot  $\bar{\lambda}_1(\tau)$  against  $\tau$  for a 600-s-long snapping shrimp noise data set sampled at 180 kHz and  $N = 10^6$ . There is clearly dependence among the samples with small  $\tau$ . For large  $\tau$ , snapping shrimp noise samples are effectively independent, i.e.,  $\bar{\lambda}_1(\tau)$  converges to zero with increasing  $\tau$ . This is accompanied by oscillatory behavior caused by *in situ* processing of the data (such as filtering) during recording. For comparison, in Fig. 5, we plot  $\bar{\lambda}_1(\tau)$  for  $\alpha\text{SGN}(m)$  realizations tuned to the snapping shrimp data for  $m \in \{0, 4\}$ . As expected,  $\bar{\lambda}_1(\tau) = 0$  for  $\tau \neq 0$  in  $\text{WS}\alpha\text{SN}$ . Clearly, the result deviates from that of snapping shrimp noise when  $\tau$  is small. On the other hand, the tuned  $\alpha\text{SGN}(4)$  realization offers exactly the same  $\bar{\lambda}_1(\tau)$  as the snapping shrimp data for immediately adjacent samples, i.e., for  $\tau \leq 4$ . The scatter plots in Figs. 2 and 4 along with the results in Fig. 5 highlight the effectiveness of the  $\alpha\text{SGN}(4)$  over its  $\text{WS}\alpha\text{SN}$  counterpart. The proximity of this model is further validated in Section VII where results in  $\alpha\text{SGN}(4)$  and snapping shrimp noise are shown to be almost identical.

## B. Binary Detection Problem

We are interested in testing the presence or absence of a known signal in  $\alpha\text{SGN}(m)$ . This is a binary detection problem and can be represented by two possible hypotheses [21]. Denoting the time samples of the received signal, transmitted signal, and noise process by  $x_n$ ,  $s_n$ , and  $w_n$  at index  $n$ , respectively,

the hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are expressed as

$$\left. \begin{array}{l} \mathcal{H}_0 : x_n = w_n \\ \mathcal{H}_1 : x_n = \theta s_n + w_n \end{array} \right\} \forall n \in \{1, 2, \dots, N\} \quad (11)$$

where  $\theta$  is a measure of signal strength and is known. We consider  $s_n$  to be of finite energy  $\mathcal{E}$ , i.e.,  $\sum_{n=1}^N s_n^2 = \mathcal{E}$ . One may vectorize the time-series expressions in (11) to get

$$\left. \begin{array}{l} \mathcal{H}_0 : \mathbf{x} = \mathbf{w} \\ \mathcal{H}_1 : \mathbf{x} = \theta \mathbf{s} + \mathbf{w} \end{array} \right\} \quad (12)$$

where  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ ,  $\mathbf{s} = [s_1, s_2, \dots, s_N]^T$ , and  $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$ . The energy constraint is then  $\|\mathbf{s}\|^2 = \mathbf{s}^T \mathbf{s} = \mathcal{E}$ , where  $\|\cdot\|$  denotes the Euclidean norm.

In a typical detection scenario, a test statistic  $T(\mathbf{x})$  is evaluated, which is then compared against a threshold  $\gamma \in \mathbb{R}$  to decide among the hypotheses [21]. Precisely,  $\mathcal{H}_1$  is selected as the correct hypothesis if  $T(\mathbf{x}) > \gamma$ , else it is  $\mathcal{H}_0$ . The probabilities of false alarm and detection are

$$\begin{aligned} \text{and} \quad P_{\text{FA}} &= P(T(\mathbf{x}) > \gamma; \mathcal{H}_0) \\ P_D &= P(T(\mathbf{x}) > \gamma; \mathcal{H}_1) \end{aligned}$$

respectively. For a given  $T(\mathbf{x})$ ,  $\gamma$  is determined for a given  $P_{\text{FA}}$ . The above formulation is the well-known NP approach to signal detection [21, Ch. 3]. Denoting  $\vec{W} = [W_1, W_2, \dots, W_N]^T$  as the random vector for which  $\mathbf{w}$  is a sample outcome and  $f_{\vec{W}}(\cdot)$  as its pdf, the LLR is defined as

$$L(\mathbf{x}) = \log \frac{f(\mathbf{x}; \mathcal{H}_1)}{f(\mathbf{x}; \mathcal{H}_0)} = \log \frac{f_{\vec{W}}(\mathbf{x} - \theta \mathbf{s})}{f_{\vec{W}}(\mathbf{x})} \quad (13)$$

where  $f(\mathbf{x}; \mathcal{H}_0) = f_{\vec{W}}(\mathbf{x})$  and  $f(\mathbf{x}; \mathcal{H}_1) = f_{\vec{W}}(\mathbf{x} - \theta \mathbf{s})$  are the pdfs of the received signal under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , respectively. The LLR detector employs  $T(\mathbf{x}) = L(\mathbf{x})$  and maximizes  $P_D$  for a given  $P_{\text{FA}}$ . This makes it the optimal detector as highlighted by the NP lemma [21, Ch. 3].

Unfortunately, it is not always possible to ascertain  $f_{\vec{W}}(\cdot)$  beforehand [21]. Further still, if  $W_n$  are samples of  $\alpha\text{SGN}(m)$ ,  $f_{\vec{W}}(\cdot)$  typically does not exist in closed form, thus requiring extensive numerical computation every time a detection event is to occur [19], [22]. The only exception to this is the Cauchy case ( $\alpha = 1$ ) [22], [23]. These constraints may make the LLR detector unfeasible for real-time implementation. It is therefore necessary to design  $T(\mathbf{x})$  that offers near-optimal detection performance. We will propose such measures in this paper and assume complete knowledge of the parameters  $\alpha$  and  $\mathbf{R}_m$  before detection. However, before we do so, we discuss a few commonly employed conventional (white) detectors that are robust to impulsive noise. The performance of these detectors will be later compared to the optimal and near-optimal detectors proposed in this paper.

#### IV. CONVENTIONAL ROBUST MEASURES

We term a robust detector as conventional or white, if it takes into account the assumption of i.i.d. noise samples. If so, the

LLR in (13) simplifies to

$$\begin{aligned} L(\mathbf{x}) &= \log \frac{\prod_{n=1}^N f_W(x_n - \theta s_n)}{\prod_{n=1}^N f_W(x_n)} \\ &= \sum_{n=1}^N \log f_W(x_n - \theta s_n) - \sum_{n=1}^N \log f_W(x_n) \end{aligned} \quad (14)$$

where  $f_W(\cdot)$  is the pdf of  $W_n \forall n \in \{1, 2, \dots, N\}$ . In the case of  $\alpha\text{SGN}(m)$ , we have  $W_n \sim \mathcal{S}(\alpha, \delta)$ . Note that the detector in (14) does not take into account the dependence between the components of  $\vec{W}$  when  $m > 0$ . Thus, it is optimal only when  $W_n$  are samples of  $\text{WS}\alpha\text{SN}$ , i.e., when  $m = 0$  [6], [12]. As highlighted previously,  $f_W(\cdot)$  does not exist in closed form for  $\mathcal{S}\alpha\text{S}$  random variables [12], [13]. To address this, one may replace (14) by suboptimal yet robust measures that try to approach its performance. In this context, researchers have looked toward generalized ML estimation (or M-estimation) theory to develop good detection schemes [8], [9]. The idea is to replace  $f_W(\cdot)$  in (14) with a general function  $\rho(\cdot) \in \mathbb{R}^+$ . More precisely, the statistic for this case is

$$T(\mathbf{x}) = \sum_{n=1}^N \log \rho(x_n - \theta s_n) - \sum_{n=1}^N \log \rho(x_n). \quad (15)$$

Ideally, one would want  $\rho(\cdot)$  to emulate  $f_W(\cdot)$  as much as possible. We term (15) as the white M-LLR. Many robust measures have been used for  $\rho(\cdot)$  in the literature [9], [12], [24]. We, however, consider only the following costs due to their common adoption.

##### 1) The Myriad

The test statistic of the myriad detector (MyD) is derived from a symmetric Cauchy distribution with i.i.d. components [10]. The MyD employs  $\log \rho(x) = -\log(K^2(\alpha)\delta^2 + x^2)$ , where  $K(\alpha)$  is the linearity parameter and is dependent on  $\alpha$  [10]. This results in

$$\begin{aligned} T(\mathbf{x}) &= - \sum_{n=1}^N \log(K^2(\alpha)\delta^2 + (x_n - \theta s_n)^2) \\ &\quad + \sum_{n=1}^N \log(K^2(\alpha)\delta^2 + x_n^2). \end{aligned} \quad (16)$$

We note that the MyD is parametric as  $\alpha$  and  $\delta$  need to be known before (16) is employed. Do note that the same information is required to invoke the LLR detector as well. However, the MyD offers the advantage of a closed-form  $T(\mathbf{x})$ . The term  $K(\alpha)$  is essentially an additional variable which may be tuned for better performance [10]. Though its optimal value may be numerically determined, the heuristic  $K(\alpha) = \sqrt{\alpha/2 - \alpha}$  is a good closed-form approximation that is typically employed in the literature [10]. With this heuristic, the MyD is optimal (equivalent to the LLR detector) when  $\alpha \rightarrow 0$ ,  $\alpha = 1$ , and  $\alpha = 2$  [10].

##### 2) The Geometric Mean

The geometric mean (or log-norm) detector (GMD) is essentially the MyD with  $K(\alpha) = 0$ . This corresponds to a very

impulsive scenario and may also be interpreted in terms of the zero-order statistics (ZOS) framework [25]. On setting  $K(\alpha) = 0$  in (16) and simplifying, we get

$$T(\mathbf{x}) = - \sum_{n=1}^N \log |x_n - \theta s_n| + \sum_{n=1}^N \log |x_n| \quad (17)$$

which corresponds to the measure  $\log \rho(x) = -\log |x|$ . We note that the GMD is nonparametric, i.e., it requires no information about the noise process and is optimal when  $\alpha \rightarrow 0$ . The ZOS framework offers a workable theory (finite moments) for almost all algebraic-tailed pdfs. Therefore, the GMD is robust in very impulsive data [25]. For the  $\alpha$ -stable framework, (17) does not diverge with increasing  $N$  for any  $\alpha$  and is thus a safe choice when  $\alpha$  is unknown. However, this leads to a substantial loss in performance when the actual value of  $\alpha$  deviates from zero, i.e., when the noise process is less impulsive [6], [20].

## V. ROBUST MEASURES IN $\alpha$ SGN( $m$ )

Though tuned to perform well in  $WS\alpha SN$ , the white detectors in Section IV do not take into account the memory of  $\alpha$ SGN( $m$ ) when  $m \neq 0$ . Given the popularity of these detectors, it is worth knowing how they fare in the latter case. Before we present these results, however, we derive the LLR detector for general  $\alpha$ SGN( $m$ ) and extend concepts in M-estimation theory to exploit the dependence between the noise samples.

### A. The LLR in $\alpha$ SGN( $m$ )

If  $W_n$  are samples of  $\alpha$ SGN( $m$ ), the LLR in (13) may be written in a more conducive form. Denoting the random vector  $\vec{W}_n = [W_1, W_2, \dots, W_n]^T$  and  $\mathbf{w}_n = [w_1, w_2, \dots, w_n]^T$  as its outcome, we can use the chain rule in probability theory to express  $f_{\vec{W}}(\cdot)$  as a product of conditional pdfs [1, p. 253]

$$f_{\vec{W}}(\mathbf{w}) = \prod_{n=1}^N f_{W_n | \vec{W}_{n-1}}(w_n | \mathbf{w}_{n-1}),$$

which may be split into

$$f_{\vec{W}}(\mathbf{w}) = f_{\vec{W}_m}(\mathbf{w}_m) \prod_{n=m+1}^N f_{W_n | \vec{W}_{n-1}}(w_n | \mathbf{w}_{n-1}). \quad (18)$$

As  $\alpha$ SGN( $m$ ) is a stationary Markov process of order  $m$ , we may express the conditional pdfs in (18) as

$$\begin{aligned} f_{W_n | \vec{W}_{n-1}}(w_n | \mathbf{w}_{n-1}) &= f_{W_n | \vec{W}_{n-1, m-1}}(w_n | \mathbf{w}_{n-1, m-1}) \\ &= f_{W_{m+1} | \vec{W}_m}(w_n | \mathbf{w}_{n-1, m-1}) \end{aligned}$$

where  $\mathbf{w}_{n, m} = [w_{n-m}, w_{n-m+1}, \dots, w_n]^T$ . Finally, substituting the above expression in (18) gives us

$$\begin{aligned} f_{\vec{W}}(\mathbf{w}) &= f_{\vec{W}_m}(\mathbf{w}_m) \prod_{n=m+1}^N f_{W_{m+1} | \vec{W}_m}(w_n | \mathbf{w}_{n-1, m-1}) \\ &= f_{\vec{W}_m}(\mathbf{w}_m) \prod_{n=m+1}^N \frac{f_{\vec{W}_{m+1}}(\mathbf{w}_{n, m})}{f_{\vec{W}_m}(\mathbf{w}_{n-1, m-1})}. \end{aligned} \quad (19)$$

On substituting (19) in (13), the LLR detector takes on the form

$$\begin{aligned} L(\mathbf{x}) &= \log \frac{f_{\vec{W}_m}(\mathbf{x}_m - \theta \mathbf{s}_m)}{f_{\vec{W}_m}(\mathbf{x}_m)} \\ &+ \log \frac{\prod_{n=m+1}^N \frac{f_{\vec{W}_{m+1}}(\mathbf{x}_{n, m} - \theta \mathbf{s}_{n, m})}{f_{\vec{W}_m}(\mathbf{x}_{n-1, m-1} - \theta \mathbf{s}_{n-1, m-1})}}{\prod_{n=m+1}^N \frac{f_{\vec{W}_{m+1}}(\mathbf{x}_{n, m})}{f_{\vec{W}_m}(\mathbf{x}_{n-1, m-1})}}. \end{aligned} \quad (20)$$

From a computational perspective, we note that (19) requires  $N - m$  evaluations of the  $(m + 1)$ -variate pdf  $f_{\vec{W}_{m+1}}(\cdot)$  and  $N - m + 1$  evaluations of the  $m$ -variate pdf  $f_{\vec{W}_m}(\cdot)$ . Moreover, as both  $\vec{W}_{m+1}$  and  $\vec{W}_m$  are  $\alpha$ SG random vectors, (20) requires a total of  $4(N - m) + 2$  numerical evaluations of multivariate  $\alpha$ SG pdfs. In comparison, (14) requires only  $2N$  evaluations of  $f_W(\cdot)$ , which is the univariate pdf of an  $\alpha$ SG random variable. Therefore, computing (20) is a more difficult task than the white LLR metric. On the bright side, the added computational complexity of evaluating a multivariate  $\alpha$ SG pdf does not increase significantly with its dimension as any such pdf can be expressed as an integral of a univariate heavy-tailed function [19]. Nevertheless, (20) is still difficult to evaluate and may not be feasible when real-time implementation is required.

### B. The Vectorized M-LLR

We can express  $\mathbf{R}_m$  in the block matrix form

$$\mathbf{R}_m = \begin{bmatrix} \mathbf{R}_{m-1} & \mathbf{r}_m \\ \mathbf{r}_m^T & r_{(m+1)(m+1)} \end{bmatrix} \quad (21)$$

where  $\mathbf{r}_m = [r_{1(m+1)}, r_{2(m+1)}, \dots, r_{m(m+1)}]^T$  and  $\mathbf{R}_{m-1} \in \mathbb{R}^{m \times m}$ . From the discussion in Section II, we note that  $f_{\vec{W}_m}(\mathbf{w}_m)$ ,  $f_{\vec{W}_{m+1}}(\mathbf{w}_{n, m})$ , and  $f_{\vec{W}_m}(\mathbf{w}_{n-1, m-1})$  can be determined completely from the statistics  $\|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_m\|$ ,  $\|\mathbf{R}_m^{-1/2} \mathbf{w}_{n, m}\|$ , and  $\|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_{n-1, m-1}\|$ , respectively. For the M-estimation framework to exploit the dependency between adjacent samples in  $\alpha$ SGN( $m$ ), we replace  $f_{\vec{W}_{m+1}}(\cdot)$  and  $f_{\vec{W}_m}(\cdot)$  in (19) by the general functions  $\rho_{m+1}(\cdot)$  and  $\rho_m(\cdot)$ , respectively. More precisely, we have

$$\begin{aligned} \varrho(\mathbf{w}; \mathbf{R}_m) &= \rho_m(\|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_m\|) \\ &\times \prod_{n=m+1}^N \frac{\rho_{m+1}(\|\mathbf{R}_m^{-1/2} \mathbf{w}_{n, m}\|)}{\rho_m(\|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_{n-1, m-1}\|)}. \end{aligned} \quad (22)$$

On substituting (22) in (13), we have the robust test statistic

$$T(\mathbf{x}) = \log \frac{\varrho(\mathbf{x} - \theta \mathbf{s}; \mathbf{R}_m)}{\varrho(\mathbf{x}; \mathbf{R}_m)}. \quad (23)$$

We note that as  $f_{\vec{W}_{m+1}}(\cdot)$  and  $f_{\vec{W}_m}(\cdot)$  are  $\alpha$ SG, they are symmetric in their respective arguments. This should extend to  $\rho_{m+1}(\cdot)$  and  $\rho_m(\cdot)$  as well and holds true as both  $\|\mathbf{R}_m^{-1/2} \mathbf{y}\| = \|\mathbf{R}_m^{-1/2} (-\mathbf{y})\| \forall \mathbf{y} \in \mathbb{R}^{m+1}$  and  $\|\mathbf{R}_{m-1}^{-1/2} \mathbf{y}\| = \|\mathbf{R}_{m-1}^{-1/2} (-\mathbf{y})\| \forall \mathbf{y} \in \mathbb{R}^m$ . We term (23) as the vectorized M-LLR.

### C. Modified Robust Measures

Now that we have derived the vectorized M-LLR, we need to derive suitable measures for  $\rho_{m+1}(\cdot)$  and  $\rho_m(\cdot)$ . We focus on vectorized versions of the white detectors discussed in Section IV.

1) *The Vector Myriad*: Analogous to the MyD, the vector myriad detector (vMyD) is derived from the symmetric multivariate Cauchy pdf [26]. The latter is  $\alpha$ SG and is given by

$$f(\mathbf{y}; d) = c_d \det(\Sigma)^{-1/2} (1 + \|\Sigma^{-1/2} \mathbf{y}\|^2)^{-(d+1)/2} \quad (24)$$

where  $d$  is the dimension of the distribution,  $\mathbf{y} \in \mathbb{R}^d$  is its argument,  $c_d \in \mathbb{R}^+$  depends on  $d$ ,  $\Sigma \in \mathbb{R}^{d \times d}$  is the covariance matrix of the underlying Gaussian random vector, and  $\det(\cdot)$  outputs the determinant of its argument [23], [27]. The vMyD is obtained by substituting

$$\bar{f}(\mathbf{y}; d) = c_d \det(\Sigma)^{-1/2} (K(\alpha) + \|\Sigma^{-1/2} \mathbf{y}\|^2)^{-(d+1)/2}$$

for  $f_{\bar{W}_m}(\cdot)$  and  $f_{\bar{W}_{m+1}}(\cdot)$  in (20) with  $\Sigma$  set to  $\mathbf{R}_{m-1}$  and  $\mathbf{R}_m$ , respectively. On simplifying and comparing the result to (23), we get

$$\varrho(\mathbf{w}; \mathbf{R}_m) = (K^2(\alpha) + \|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_m\|^2)^{-\frac{m+1}{2}} \prod_{n=m+1}^N \frac{(K^2(\alpha) + \|\mathbf{R}_m^{-1/2} \mathbf{w}_{n,m}\|^2)^{-\frac{m+2}{2}}}{(K^2(\alpha) + \|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_{n-1,m-1}\|^2)^{-\frac{m+1}{2}}} \quad (25)$$

and henceforth from (22),  $\rho_m(x) = (K^2(\alpha) + x^2)^{-((m+1)/2)}$ . We note that  $\rho_m(\cdot)$  is similar to  $\rho(\cdot)$  used for the MyD, but with the exception of  $\delta$ . The latter is omitted as it is incorporated in  $\mathbf{R}_m$  and  $\mathbf{R}_{m-1}$  in (25). As in the white noise case, we use  $K(\alpha) = \sqrt{(\alpha/(2-\alpha))}$  as it is optimal for  $\alpha \rightarrow 0$ ,  $\alpha = 1$ , and  $\alpha = 2$  [26]. Note that the vMyD requires the off-diagonal elements of  $\mathbf{R}_m$  in addition to  $\alpha$  and  $\delta$  as compared to its white counterpart.

2) *The Vector Geometric Mean*: The vector geometric mean detector (vGMD) is the vMyD with  $K(\alpha) = 0$ . From (25), we have

$$\varrho(\mathbf{w}; \mathbf{R}_m) = \|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_m\|^{-(m+1)} \prod_{n=m+1}^N \frac{\|\mathbf{R}_m^{-1/2} \mathbf{w}_{n,m}\|^{-(m+2)}}{\|\mathbf{R}_{m-1}^{-1/2} \mathbf{w}_{n-1,m-1}\|^{-(m+1)}}. \quad (26)$$

Consequently,  $\rho_m(x) = |x|^{-(m+1)}$ . The vGMD is derived by substituting (26) in (23). Do note that unlike its white counterpart, the vGMD requires  $\mathbf{R}_m$ , which characterizes the dependence within  $\mathbf{w}_n$ . On the other hand, it still does not require knowledge about  $\alpha$  and  $\delta$ . To see why the vGMD is independent of  $\alpha$ , we clearly observe that (26) and therefore (23) do not depend on  $\alpha$ . For the case of  $\delta$ , from (7) and (21), we have

$$\hat{\mathbf{R}}_m = \begin{bmatrix} \hat{\mathbf{R}}_{m-1} & \mathbf{r}_m / \delta^2 \\ \mathbf{r}_m^T / \delta^2 & 1 \end{bmatrix}. \quad (27)$$

On substituting  $\mathbf{R}_m = \hat{\mathbf{R}}_m \delta^2$  and  $\mathbf{R}_{m-1} = \hat{\mathbf{R}}_{m-1} \delta^2$  in (23), we find that  $\delta$  cancels out in  $T(\mathbf{x})$ . Therefore, the vGMD is a semiparametric detector, requiring only the normalized covariance matrices  $\hat{\mathbf{R}}_m$  and  $\hat{\mathbf{R}}_{m-1}$ .

Now that we have discussed robust detectors and their vectorized counterparts in  $\alpha$ SGN( $m$ ), we analyze how  $\mathbf{s}$  can influence the overall performance of the system.

## VI. SIGNAL SHAPE, BANDWIDTH, AND ENERGY CONSTRAINTS

### A. Optimal and Near-Optimal Signaling

If  $W_n$  are samples of  $\alpha$ SGN( $m$ ), then for a general positive-semidefinite  $\mathbf{R}_m$  and  $N$ , we note that  $f_{\bar{W}}(\cdot)$  is not rotationally symmetric (or rotationally invariant). The only exception to this arises from  $\mathbf{R}_m = \delta^2 \mathbf{I}_m$  and  $N \leq m$ , for which  $\bar{W}$  is an isotropic random vector [19]. In the latter scenario,  $\mathbf{s}$  (given the constraint  $\|\mathbf{s}\|^2 = \mathcal{E}$ ) does not have any effect on the performance of the LLR detector. However, for all other instances of  $\alpha$ SGN( $m$ ), the design of  $\mathbf{s}$  needs to be taken into account.

For the framework discussed in Section III-B, the optimal  $\mathbf{s}$  for an employed  $T(\mathbf{x})$  is such that it maximizes  $P_D$  for a given  $P_{FA}$  while satisfying the constraint  $\|\mathbf{s}\|^2 = \mathcal{E}$ . As

$$\begin{aligned} P_D &= \int_{\mathbf{x}: T(\mathbf{x}) \geq \gamma} f(\mathbf{x}; \mathcal{H}_1) d\mathbf{x} \\ &= \int_{\mathbf{x}: T(\mathbf{x}) \geq \gamma} f_{\bar{W}}(\mathbf{x} - \theta \mathbf{s}) d\mathbf{x} \end{aligned}$$

the optimization problem may be expressed as

$$\begin{aligned} \arg \max_{\mathbf{s}} & \int_{\mathbf{x}: T(\mathbf{x}) \geq \gamma} f_{\bar{W}}(\mathbf{x} - \theta \mathbf{s}) d\mathbf{x} \\ \text{s.t.} & \|\mathbf{s}\|^2 = \mathcal{E}. \end{aligned} \quad (28)$$

Solving (28) is not trivial for the general case of  $\alpha$ SGN( $m$ ). First, with the exception of the Cauchy case, the joint pdfs in (19) are not in closed form [22]. Therefore, evaluating  $f_{\bar{W}}(\cdot)$  becomes exceedingly taxing with increasing  $N$ . Second, the surface over which  $\mathbf{x}$  is integrated, as determined by  $T(\mathbf{x})$ , is complicated. This is highlighted for the WS $\alpha$ SN case in [6], [20] and intuitively extended to  $\alpha$ SGN( $m$ ) when  $N \gg m$ , as one would expect near independence between extreme samples in  $\bar{W}$ . Therefore, one may employ a suboptimal approach to address this problem.

We note that  $f_{\bar{W}}(\mathbf{w})$  is unimodal with its maximum at  $\mathbf{w} = \mathbf{0}$ . Intuitively, if  $f(\mathbf{x}; \mathcal{H}_1)$  is constrained to offer its minimum at the coordinate point at which  $f(\mathbf{x}; \mathcal{H}_0)$  is at its maximum (at  $\mathbf{x} = \mathbf{0}$ ), then both pdfs will be spaced out in  $\mathbb{R}^N$  in such a way that there is minimum overlap between the pdfs in regions where the random outcomes  $\mathbf{x}$  will most probably lie. This may then be exploited by the detector to achieve superior performance. Mathematically, the problem is expressed as

$$\begin{aligned} \hat{\mathbf{s}} &= \arg \min_{\mathbf{s}} f_{\bar{W}}(-\theta \mathbf{s}) \\ \text{s.t.} & \|\mathbf{s}\|^2 = \mathcal{E}. \end{aligned} \quad (29)$$

As  $f_{\bar{W}}(\cdot)$  is a symmetric function, we note that  $-\hat{\mathbf{s}}$  is also a solution of (29). By taking advantage of the unimodal and symmetric properties of  $f_{\bar{W}}(\cdot)$  and the monotonically increasing

logarithmic function, we can rewrite (29) as

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \log f_{\vec{W}}(\theta \mathbf{s}) \quad (30)$$

$$\text{s.t. } \|\mathbf{s}\|^2 \leq \mathcal{E}$$

without any loss in generality. One notes that (30) does not depend on the employed detector and offers the same  $\hat{\mathbf{s}}$  for any  $T(\mathbf{x})$ . As (30) spreads out  $f(\mathbf{x}; \mathcal{H}_0)$  and  $f(\mathbf{x}; \mathcal{H}_1)$  in  $\mathbb{R}^N$  by taking into account the structure of  $f_{\vec{W}}(\cdot)$ , it works well with the vectorized detectors presented in Section V. This observation stems from the fact that these detectors are good approximations of the LLR detector, the latter of which optimally exploits the structure of  $f_{\vec{W}}(\cdot)$ . From a computational point of view, we note that the problem in (30) is nonconvex. In our experience, using a random initialization point with a convex solver results in consistently good solutions when  $\theta$  is small. This regime is of interest in a typical detection scenario as it represents a weak received signal [21].

From another perspective, we note that the solution in (30) is equivalent to (28) when  $W_n$  is a WGN process and LLR (matched-filter) detection is employed [21, Ch. 4]. As  $f_{\vec{W}}(\cdot)$  is an isotropic pdf in this case [2, Ch. 2], the LLR detector offers no constraint on  $\mathbf{s}$  besides that it lies on the  $N$ -dimensional circle of radius  $\sqrt{\mathcal{E}}$ . This is a well-known result and is achieved by replacing  $f_{\vec{W}}(\cdot)$  in (30) by a product of  $N$  identical univariate Gaussian pdfs [21].

In applications such as sonar, various properties of  $\mathbf{s}$  are required to meet the systems objectives. The signal  $s_n \forall n \in \{1, 2, \dots, N\}$  is essentially bandlimited and is designed to achieve good Doppler/range resolutions [28]. Examples include the families of continuous and frequency modulated waveforms. We note that (30) does not take into account the bandwidth constraint of  $s_n$  and, as discussed later, leads to narrowband  $\hat{\mathbf{s}}$  in  $\alpha\text{SGN}(m)$  for  $m \neq 0$ . Therefore, we discuss both  $\hat{\mathbf{s}}$  and good bandlimited signals in  $\alpha\text{SGN}(m)$  next.

### B. Signal Design in $\alpha\text{SGN}(m)$

Before discussing signal design for general  $\alpha\text{SGN}(m)$ , we comment on  $\hat{\mathbf{s}}$  for non-Gaussian WS $\alpha$ SN. In this case,  $f_{\vec{W}}(\cdot)$  can be expressed as a product of its marginals. Therefore, (30) simplifies to

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \sum_{i=1}^N \log f_{W_i}(\theta s_i) \quad (31)$$

$$\text{s.t. } \|\mathbf{s}\|^2 \leq 1.$$

As  $\vec{W}$  is a heavy-tailed S $\alpha$ S random vector with i.i.d. components,  $f_{\vec{W}}(\cdot)$  is a symmetric multitailed pdf with tails directed along the positive and negative directions of each axis [6], [20]. One may use Lagrange multipliers to solve (31), presented in the Appendix, which leads to  $\hat{\mathbf{s}} = [\hat{s}_1, \hat{s}_2, \dots, \hat{s}_N]^T$  such that  $\hat{s}_n \in \{\pm \sqrt{\mathcal{E}/N}\} \forall n \in \{1, 2, \dots, N\}$ . This setting ensures that the tails of the pdfs under  $\mathcal{H}_1$  and  $\mathcal{H}_0$  do not point toward each other [6], [20]. If  $\mathbf{s}$  were to consist of any zeros, some of the tails would experience complete overlap resulting in performance loss under optimal detection. On the other hand, with increasing  $N$ , the tails become lighter and the performance gap under LLR detection between  $\hat{\mathbf{s}}$  and any  $\mathbf{s}$  that satisfies  $\|\mathbf{s}\|^2 = \mathcal{E}$  decreases. In fact,  $\mathbf{s}$  ceases to influence the performance for sufficiently

large  $N$  for any  $\|\mathbf{s}\|^2 = \mathcal{E}$ . This is substantiated by the fact that (14) is a sum of i.i.d. random variables and it converges to a Gaussian distribution under the CLT when  $N \rightarrow \infty$  whose parameters depend on  $\mathcal{E}$  [21, p. 390]. Consequently, narrowband or wideband  $s_n$  has no effect on the LLR detection performance in non-Gaussian WS $\alpha$ SN as its energy is distributed over a number of time samples for most practical purposes.

For general  $\alpha\text{SGN}(m)$ , investigating  $\hat{\mathbf{s}}$  is not as straightforward. For example, when  $N \leq m + 1$ ,  $\vec{W}$  is an elliptic distribution. Yet as  $N$  increases, analogous to the WS $\alpha$ SN case,  $f_{\vec{W}}(\cdot)$  starts to exhibit tails as extreme components in  $\vec{W}$  are decreasingly dependent on one another. This is highlighted by the scatter plots in Fig. 2. It is insightful to compare these observations with those of the underlying Gaussian AR( $m$ ) process, as optimal signal structures are already known for the latter [21, Ch. 4].

As  $W_n$  is a stationary process, so is the underlying  $G_n$ . For a stationary Gaussian AR( $m$ ) process, the joint pdf of any  $N$  samples is known to be Gaussian [21]. Therefore, we have  $\vec{G} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , where  $\Sigma \in \mathbb{R}^{N \times N}$  is a symmetric positive-semidefinite Toeplitz matrix for any  $N \in \mathbb{Z}^+$  [21], [29]. If  $N \leq m + 1$ , then  $\Sigma$  is equivalent to the top-left  $N \times N$  block matrix of  $\mathbf{R}_m$ , i.e.,  $\Sigma = \mathbf{R}_{N-1}$ . For  $N > m + 1$ , one may recursively evaluate the first row of  $\Sigma$  from that of  $\mathbf{R}_m$ . This is essentially the ACF of  $G_n$  up until a delay of  $N - 1$ . In turn, this may be used to generate the remaining components of  $\Sigma$  by exploiting the symmetric Toeplitz properties of the matrix. More precisely, if  $\Sigma = [\sigma_{ij}]$ , then

$$\sigma_{1k} = \begin{cases} r_{1k} & \text{for } 1 < k \leq m + 1 \\ \mathbf{r}_m^T \mathbf{R}_{m-1}^{-1} [\sigma_{1(k-1)}, \sigma_{1(k-2)}, \dots, \sigma_{1(k-m)}]^T \text{ o.w.} & \end{cases} \quad (32)$$

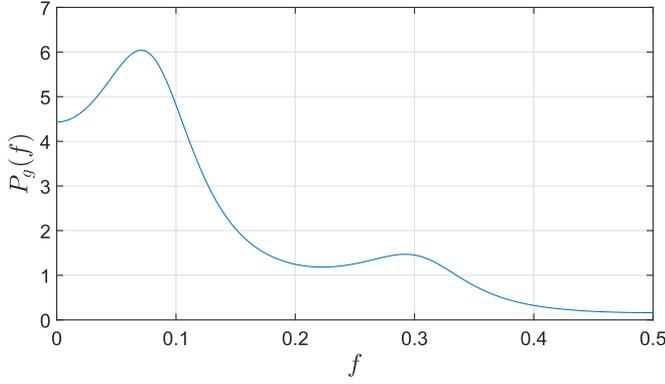
As  $\vec{G}$  is a Gaussian random vector, it has an  $N$ -dimensional elliptic distribution. As  $N$  increases, the correlation between any pair of extreme components in  $\vec{G}$  decreases accordingly. In fact, any two samples that have between them a delay more than the coherence time of the noise process are sufficiently independent and therefore offer jointly isotropic distributions [21].

If  $\mathbf{s}$  is immersed in  $\vec{G}$ , the optimal  $\mathbf{s}$  for LLR detection is given by the eigenvector corresponding to the minimum eigenvalue of  $\Sigma$  [21, Ch. 4]. This again is exactly what (30) offers when  $f_{\vec{W}}(\cdot)$  is replaced by the pdf of  $\vec{G}$ . In the probabilistic domain,  $\hat{\mathbf{s}}$  places the elliptic pdf under  $\mathcal{H}_1$  in such a way that there is minimum overlap with that corresponding to  $\mathcal{H}_0$ . In another interpretation, as  $N \rightarrow \infty$ ,  $\hat{\mathbf{s}}$  is such that its spectrum is concentrated where the power spectral density (PSD) of the noise is at its minimum [21]. In our setting, using the block matrix form in (21), we can express  $G_n$  as the difference equation

$$G_n = \mathbf{r}_m^T \mathbf{R}_{m-1}^{-1} \vec{G}_{n-1, m-1} + \sqrt{\kappa} Z_n$$

$$= \sum_{k=1}^m \psi_k G_{n-k} + \sqrt{\kappa} Z_n \quad (33)$$

where  $Z_n \sim \mathcal{N}(0, 1) \forall n \in \mathbb{Z}$ ,  $\kappa = \det \mathbf{R}_m / \det \mathbf{R}_{m-1}$ , and  $\psi_k$  is the  $(m - k + 1)$ th element of the  $m$ -dimensional vector  $(\mathbf{r}_m^T \mathbf{R}_{m-1}^{-1})^T = \mathbf{R}_{m-1}^{-1} \mathbf{r}_m$  [22]. Therefore, the one-sided PSD is


 Fig. 6. PSD of a Gaussian AR(4) process with  $\delta = 1$  and  $\hat{\mathbf{R}}_m$  in (8).

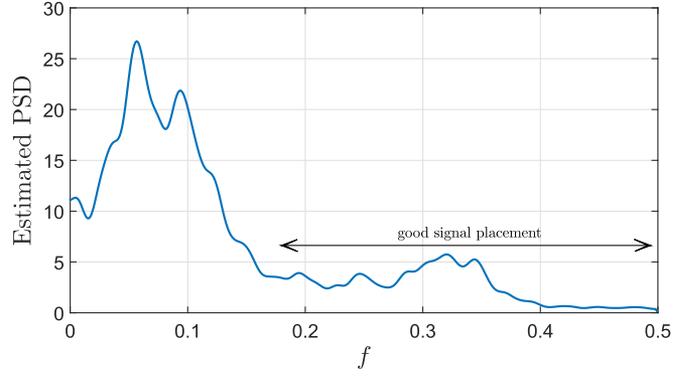
evaluated in closed form as [30, pp. 87–88]

$$P_g(f) = \frac{2\kappa}{|1 - \sum_{k=1}^m \psi_k e^{-j2\pi k f}|^2}, \quad \text{for } 0 \leq f < 1/2 \quad (34)$$

where  $\delta^2 = \int_0^{1/2} P_g(f) df$  and  $f$  denotes normalized frequency. For the covariance matrix in (8), we present the one-sided PSD of the underlying Gaussian AR(4) process in Fig. 6. For large  $N$ ,  $\hat{\mathbf{s}}$  is essentially a narrowband signal occupying  $f \approx 0.5$  where  $P_g(f)$  is at its minimum. A generalization to this argument is made if  $\mathbf{s}$  is to be bandlimited. Its spectrum is placed in the higher frequency range as the in-band noise power [area under the  $P_g(f)$  curve for the concerned bandwidth] is less than that in the lower spectra.

Going back to the  $\alpha$ SGN( $m$ ) case, for  $N \leq m + 1$ , (30) offers almost the same  $\hat{\mathbf{s}}$  for any  $\alpha$  and is approximately similar to that in the Gaussian AR( $m$ ) case. This is mainly because both  $\vec{W}$  and  $\vec{G}$  are elliptical distributions within this region. Moreover, the similarity extends even for  $N > m + 1$  as numerical evaluations of  $\hat{\mathbf{s}}$  show it to be approximately equivalent to the eigenvector of  $\Sigma$  corresponding to its minimum eigenvalue. Though this has an interpretation in the spectral domain for Gaussian AR( $m$ ) noise, the latter's PSD representation cannot be directly applied to  $W_n$  as its second-order moments do not exist [12]. However, the spectral arguments may be extended to  $W_n$  if the Gaussian PSD is defined as a robust estimate of its spectral shape, namely the pseudo-PSD [6]. To highlight this, we plot the estimated PSD from a realization of  $W_n$  in Fig. 7 generated for the covariance matrix in (8),  $\delta = 1$  and  $\alpha = 1.5$ . We use the conventional Welch method to estimate the power spectrum [31]. Clearly, the spectral shape follows that of the underlying  $G_n$  in Fig. 6, albeit with a different scale. As in the Gaussian case, evaluating (30) for any  $\alpha$  corresponds to  $\hat{\mathbf{s}}$  being a narrowband signal occupying  $f \approx 0.5$ . Similarly, good placement for a bandlimited signal corresponds to frequency regions with a lower noise spectra. This has been marked in Fig. 7.

On a final note, we see that  $f$  can be mapped onto the actual frequency  $\bar{f}$  by multiplying it with the employed sampling rate  $f_s$ , i.e.,  $\bar{f} = f_s f$ . As (8) is estimated for snapping shrimp noise sampled at 180 kHz, the actual frequency range in Figs. 6 and 7 is  $\bar{f} \in [0, 90]$  kHz.


 Fig. 7. Estimated PSD of an  $\alpha$ SGN(4) process with  $\delta = 1$  and  $\hat{\mathbf{R}}_m$  in (8).

### C. The SNR Measure

The entity  $\mathcal{E}\theta^2/(2\delta^2)$  has been employed as an SNR measure in the literature for WS $\alpha$ SN [20], [32]. Conventionally, the measure  $\mathcal{E}\theta^2/\sigma^2$  is adopted, where  $\sigma^2 = E[W_n^2]$  is the power of the noise process, as it completely determines the performance of the LLR detector (matched filter) in WGN [21, Ch. 4]. However, this does not carry any meaning in WS $\alpha$ SN as second-order moments of S $\alpha$ S random variables do not exist. A way around this is to note that the equality  $\sigma^2 = 2\delta^2$  holds for a Gaussian random variable as  $\mathcal{S}(2, \delta) \stackrel{d}{=} \mathcal{N}(0, 2\delta^2)$ . The scale has an interpretation within the  $\alpha$ -stable framework, which is why  $\mathcal{E}\theta^2/(2\delta^2)$  is employed. We employ the same measure in our work for  $\alpha$ SGN( $m$ ) as it signifies the effect of signal design on the detection performance of the system. To highlight this, we first express the SNR in terms of the spectral information of  $s_n$  and the noise. From Parseval's theorem, we have  $\mathcal{E} = \int_0^{1/2} |S(f)|^2 df$ , where  $|S(f)|^2$  is the one-sided energy spectral density (ESD) of  $s_n$ . As  $E[G_n^2] = \delta^2 \forall n \in \{1, 2, \dots, N\}$ , we can express the SNR as

$$\text{SNR} = \frac{\mathcal{E}\theta^2}{2 \int_0^{1/2} P_g(f) df} = \frac{\theta^2 \int_0^{1/2} |S(f)|^2 df}{2 \int_0^{1/2} P_g(f) df}. \quad (35)$$

On the other hand, under matched-filter detection and for large  $N$ , the performance for a wide-sense stationary (WSS) noise process with spectra  $2P_g(f)$  (and thus power  $\sigma^2 = 2 \int_0^{1/2} P_g(f) df = 2\delta^2$ ) is determined by [21, p. 112]

$$\overline{\text{SNR}} = \theta^2 \int_0^{1/2} \frac{|S(f)|^2}{P_g(f)} df. \quad (36)$$

We note that (36) essentially scales  $|S(f)|^2$  by the noise power spectra before averaging along the frequency axis. In contrast, (35) individually averages the energy and power spectra of the signal and noise. Both  $\overline{\text{SNR}}$  and SNR are equivalent if  $P_g(f)$  is flat (or white). However, this is not the case in  $\alpha$ SGN( $m$ ) for  $m \neq 0$ . For a narrowband signal present at  $f_0$  with bandwidth  $B \ll 0.5$ , the measures reduce to

$$\text{SNR} = \frac{\theta^2 |S(f_0)|^2}{2 \int_0^{1/2} P_g(f) df} B \quad (37)$$

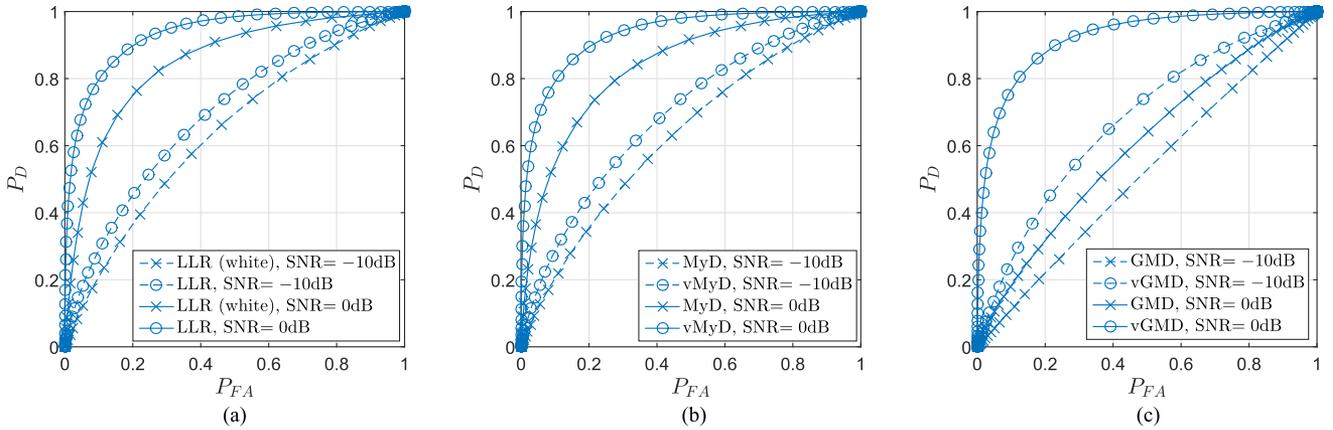


Fig. 8. ROC curves for  $\alpha = 1.5$ ,  $\text{SNR (dB)} \in \{-10, 0\}$  and  $N = 10$ . Performances of the LLR and LLR (white) detectors are shown in (a), the vMyD and MyD in (b), and the vGMD and GMD in (c).

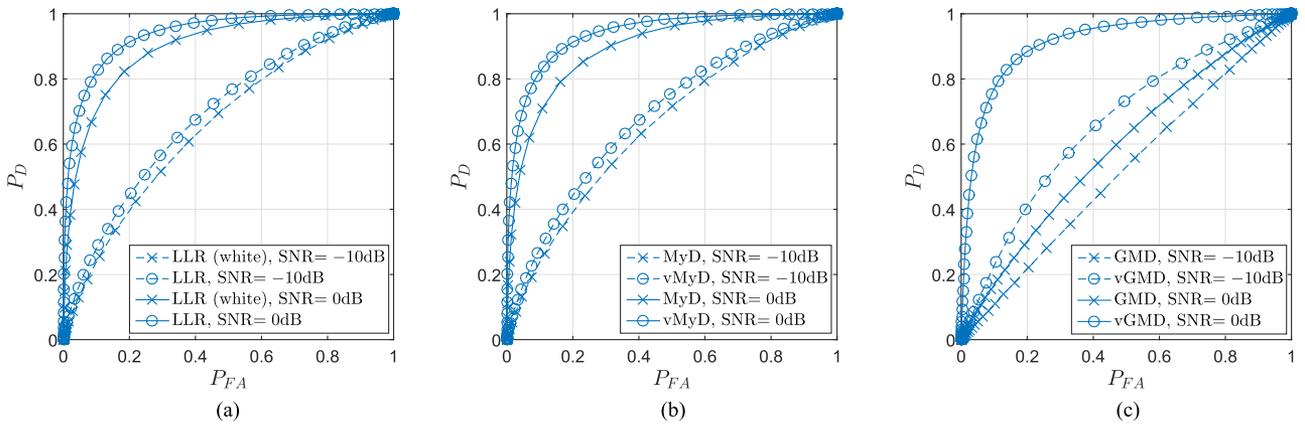


Fig. 9. ROC curves for  $\alpha = 1.9$ ,  $\text{SNR (dB)} \in \{-10, 0\}$  and  $N = 10$ . Performance of the LLR and LLR (white) detectors are shown in (a), the vMyD and MyD in (b), and the vGMD and GMD in (c).

and

$$\overline{\text{SNR}} = \frac{\theta^2 |S(f_0)|^2}{P_g(f_0)} B. \quad (38)$$

Similarly, for a signal bandlimited to  $B < 0.5$ , we have

$$\text{SNR} = \frac{\theta^2 \int_B |S(f)|^2 df}{2 \int_0^{1/2} P_g(f) df} \quad (39)$$

and

$$\overline{\text{SNR}} = \theta^2 \int_B \frac{|S(f)|^2}{P_g(f)} df. \quad (40)$$

In both cases, one can clearly see that for a given SNR,  $\overline{\text{SNR}}$  may be relatively better or worse by concentrating  $S(f)$  in regions where  $P_g(f)$  is lower or larger, respectively, than the average power of noise over the entire frequency band. In other words, matched-filter detection performance in a WSS colored noise process may be better or worse than that in white noise for the same SNR depending on the frequency band chosen. As the matched-filter detector is the LLR metric for the underlying Gaussian AR( $m$ ) process [21], the relation between SNR and  $\overline{\text{SNR}}$  extends to the LLR detector in  $\alpha\text{SGN}(m)$  as  $P_g(f)$  signifies the spectral shape of the latter.

## VII. SIMULATION RESULTS AND OBSERVATIONS

In all our simulations, we employ the  $\alpha\text{SGN}(4)$  process with  $\hat{\mathbf{R}}_4$  given in (8). As mentioned before, the motivation for this stems from the estimates one gets for snapping shrimp data sets sampled at 180 kHz in shallow tropical waters [18]. Typical values of  $\alpha$  have been shown to lie within  $\alpha \geq 1.5$  [5]. Therefore, we compile most results for  $\alpha = 1.5$  as it highlights the most impulsive practical noise scenario. We offer some insights for the  $\alpha = 1.9$  case as well. The results are compared to those observed in snapping shrimp noise and WS $\alpha$ SN. In all instances, we set

$$s_n = \frac{\sqrt{\mathcal{E}} \bar{s}_n}{\sqrt{\sum_{i=1}^N \bar{s}_n^2}} \quad (41)$$

where

$$\bar{s}_n = \Re\{\text{sinc}(n/10) e^{j2\pi f_c n}\}. \quad (42)$$

Here  $\text{sinc}(\cdot)$  is the normalized sinc function and  $f_c = 0.4$  is the carrier frequency. Thus,  $s_n$  is a continuous-wave bandlimited signal of bandwidth 0.2 centered at  $f = 0.4$ .

In Fig. 8, we present the receiver operating characteristic (ROC) curves for the  $\alpha = 1.5$  case for all discussed white and

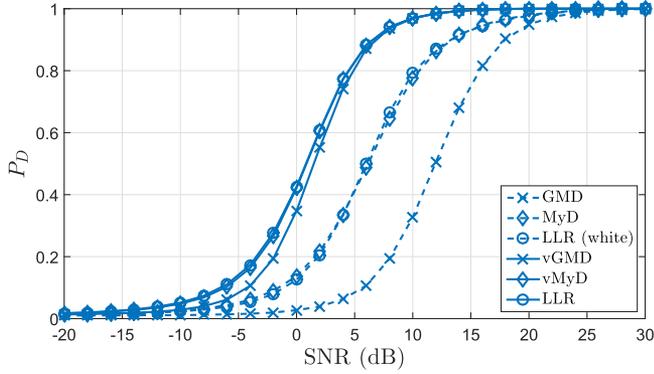


Fig. 10. Performance of white and vectorized detectors for  $P_{FA} = 10^{-2}$ ,  $N = 10$ , and  $\alpha = 1.5$ .

vectorized detectors. Results are plotted for SNRs of  $-10$  and  $0$  dB for  $N = 10$ . Each plot is constructed in such a way that we see the performance of a white detector with its vectorized counterpart. In all instances, it is observed that the vectorized detectors outperform the conventional robust detectors. Clearly, taking the dependence between samples under consideration offers superior performance in all cases. The ROC curve trends of the LLR (white) and the LLR detectors are almost similar with those of the MyD and vMyD, respectively. In the case of the GMD and the vGMD, the performance gap between them is far larger. Do note that the GMD and the vGMD do not take any information about  $\alpha$  and  $\delta$  under consideration, yet by just accommodating for  $\hat{\mathbf{R}}_4$ , the vGMD offers significant improvement in performance. We also note that the ROC curves of the vectorized detectors (circle markers) are almost similar in all cases. This highlights the fact that added information from  $\alpha$  and  $\delta$  offers only a little advantage over the case when  $\hat{\mathbf{R}}_4$  is already known and compensated for. In Fig. 9, we present ROC results for the  $\alpha = 1.9$  case. The remaining parameters are the same as those used to generate Fig. 8. The trends seen in the latter case extend here as well. However, the advantages offered by the vectorized LLR and myriad over their white counterparts are slightly reduced. The ROC curves of the GMD and the vGMD are almost similar to those in Fig. 8.

Though descriptive in their own right, the ROC curves do not highlight performance when  $P_{FA}$  is small. In Figs. 10 and 11, we plot  $P_D$  against the SNR (in decibels) for  $P_{FA} = 10^{-2}$  and  $P_{FA} = 10^{-4}$ , respectively, for  $N = 10$ . One can see that the SNR gains of the vectorized detectors over the conventional detectors are substantial. As expected, the GMD offers the worst performance as it does not take any of the noise parameters into account. Further still, as was previously depicted by the ROC curves, the LLR detector, the vMyD, and the vGMD all offer comparable performance. We also plot performance curves for the  $\alpha = 1.9$  case in Figs. 12 and 13 for  $P_{FA} = 10^{-2}$  and  $P_{FA} = 10^{-4}$ , respectively, for  $N = 10$ . Do note that the curves for  $P_{FA} = 10^{-2}$  are notably different from those for the  $\alpha = 1.5$  case in Fig. 10. The performance of the LLR (white) detector and MyD is close to that of the vectorized detectors. In fact, they even outperform the vGMD at low SNR. However, as  $P_{FA}$  is reduced, their performance deteriorates much faster as seen in Fig. 13.

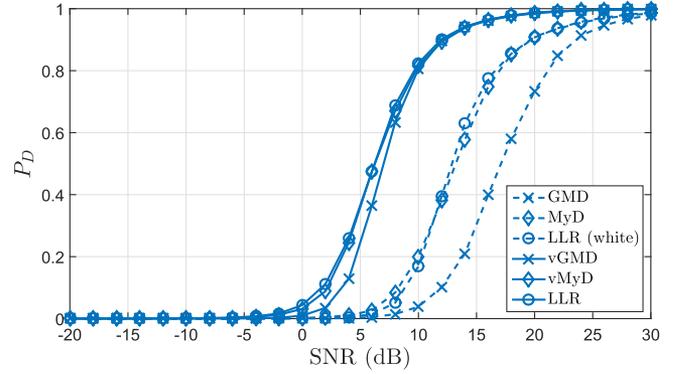


Fig. 11. Performance of white and vectorized detectors for  $P_{FA} = 10^{-4}$ ,  $N = 10$ , and  $\alpha = 1.5$ .

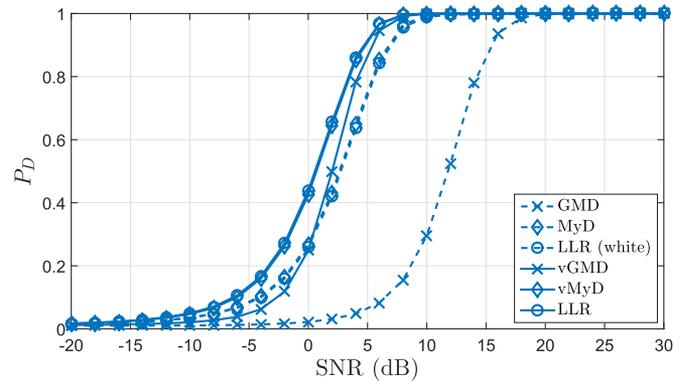


Fig. 12. Performance of white and vectorized detectors for  $P_{FA} = 10^{-2}$ ,  $N = 10$ , and  $\alpha = 1.9$ .

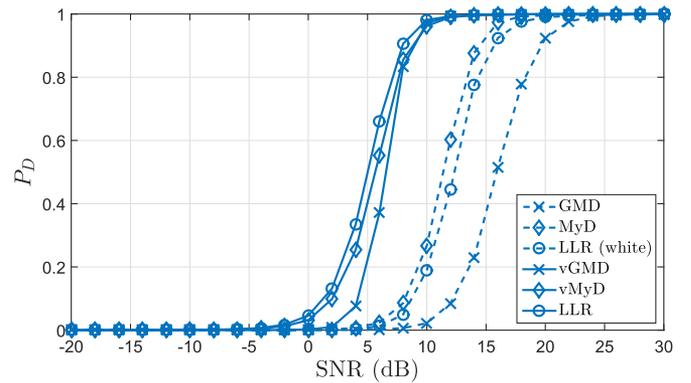


Fig. 13. Performance of white and vectorized detectors for  $P_{FA} = 10^{-4}$ ,  $N = 10$ , and  $\alpha = 1.9$ .

We note that the vGMD is slightly worse than the LLR detector and the vMyD in Figs. 10–13. This is a little more pronounced in the figures corresponding to the  $\alpha = 1.9$  case. The reason for this is that the vGMD is robust for even the most impulsive scenarios within the  $\alpha$ -stable framework and is optimal as  $\alpha \rightarrow 0$ . However, this makes it inefficient when the true value of  $\alpha$  is large. As the vGMD is essentially the vMyD with  $\alpha = 0$ , a good idea is to invoke the vMyD with  $\alpha$  set to the most impulsive case in the environment of interest. As  $\alpha \geq 1.5$  is the typical estimated range for practical snapping shrimp data

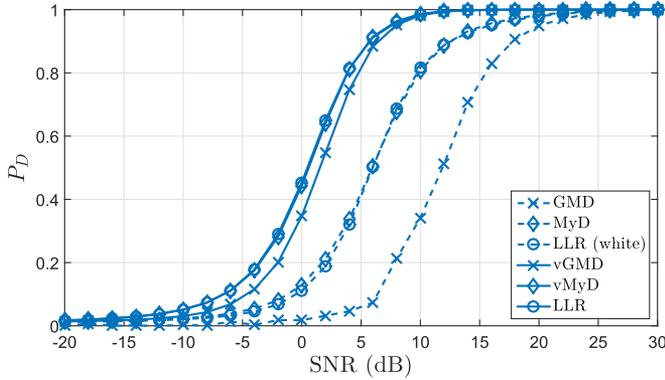


Fig. 14. Performance of white and vectorized detectors in snapping shrimp noise for  $P_{FA} = 10^{-2}$  and  $N = 10$ .

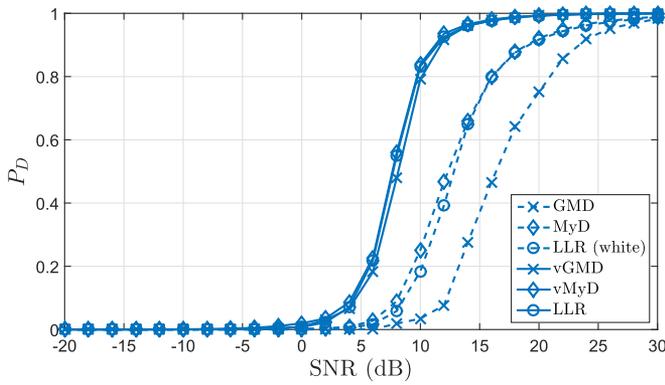


Fig. 15. Performance of white and vectorized detectors in snapping shrimp noise for  $P_{FA} = 10^{-4}$  and  $N = 10$ .

sets [5], invoking the vMyD with  $\alpha = 1.5$  is robust for this entire range and offers superior performance than the vGMD. However, unlike the vGMD, this detector requires additional information of  $\delta$  besides  $\hat{\mathbf{R}}_m$ .

If  $\alpha\text{SGN}(m)$  is indeed an effective model for snapping shrimp noise, then detection performance in either case should be almost similar. To highlight this, we synthesize data by immersing our signal in snapping shrimp noise and employ numerical simulations to evaluate  $P_D$  as a function of SNR for  $N = 10$ . Results are plotted for  $P_{FA} = 10^{-2}$  and  $P_{FA} = 10^{-4}$  in Figs. 14 and 15, respectively. The estimated parameters for this data set (with obvious notation) are  $\hat{\alpha} = 1.54$  and  $\hat{\mathbf{R}}_4$  in (8). It thus exhibits severe snapping shrimp noise. Clearly the trends are very similar to those observed in Figs. 10 and 11, which are compiled for  $\alpha\text{SGN}(4)$  with  $\alpha = 1.5$  and  $\hat{\mathbf{R}}_4$  in (8). As the vectorized detectors are optimized for the  $\alpha\text{SGN}(m)$  framework, their effectiveness in snapping shrimp noise clearly highlights the proximity of both processes. Harnessing the memory within closely spaced samples offers several decibels worth of gain over the best conventional detector. As the  $\text{WS}\alpha\text{SN}$  model has been used predominantly in the literature to model snapping shrimp noise [6], [11], we also plot the detection performance of the conventional detectors in  $\text{WS}\alpha\text{SN}$  for  $\alpha = 1.5$ ,  $N = 10$  and  $P_{FA} \in \{10^{-2}, 10^{-4}\}$  in Fig. 16. The results are slightly skewed in comparison to those observed in Figs. 14 and 15. We

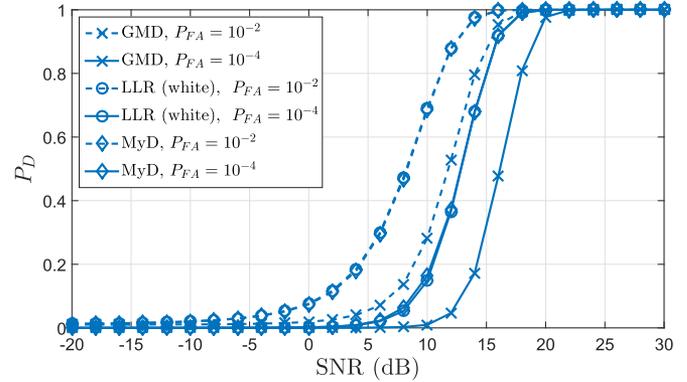


Fig. 16. Performance of white detectors in  $\text{WS}\alpha\text{SN}$  for  $N = 10$ ,  $P_{FA} \in \{10^{-2}, 10^{-4}\}$  and  $\alpha = 1.5$ .

note that though the LLR (white) detector in  $\text{WS}\alpha\text{SN}$  is optimal, its performance is far worse than what can actually be achieved in practice by robust memory-exploiting detectors. Moreover, as its samples are i.i.d. random variables, the  $\text{WS}\alpha\text{SN}$  model does not offer us the mathematical framework to derive such detectors.

Until now we have kept  $N = 10$  constant in our simulations. It is interesting to see how the detectors fare when  $N$  is varied for a given SNR. Though this would not matter in WGN, the latter option offers a more attractive prospect in  $\text{WS}\alpha\text{SN}$ , Gaussian  $\text{AR}(m)$  noise, and  $\alpha\text{SGN}(m)$ , as a larger  $N$  offers more degrees of freedom to exploit the anisotropic pdfs associated with the random processes. To highlight this, ROC curves are shown for the white and vectorized detectors in Fig. 17 for the  $\alpha = 1.5$  case. The plots are generated for  $\text{SNR} = -10$  dB and comparisons are made for  $N = 10$  and  $N = 100$ . There is slight improvement in performance when  $N$  is increased. Though not presented here, similar trends are seen in the  $\alpha = 1.9$  case as well.

Perhaps a better visualization is offered when  $P_D$  is plotted against the SNR (in decibels). In Figs. 18 and 19, we present such curves for the  $\alpha = 1.5$  case for  $P_{FA} = 10^{-2}$  and  $P_{FA} = 10^{-4}$ , respectively. As the trends are mostly the same, we only show results for the LLR and LLR (white) detectors. Comparisons are made for the  $N = 10$  and  $N = 100$  cases. One can clearly see the improvement in performance among all detectors when  $N$  is large. Interestingly, the LLR (white) detector also performs increasingly well. This is expected as  $f_{\bar{W}}(\cdot)$  starts developing tails when  $N$  is large. This is exactly what robust white detectors are designed to mitigate. We have also highlighted the SNR gains between the LLR and LLR (white) detectors for both instances of  $N$  to show how the performance gap increases when  $N$  is large. One also notes that there is a noticeable difference between SNR gains for different  $P_D$ . In fact, both Figs. 18 and 19 portray the  $N = 100$  curves to rise more sharply than their  $N = 10$  counterparts. This can be explained via a limiting argument based on the central limit theorem (CLT). To highlight this, we compare the LLR detector performance for the  $\alpha = 1.5$  case with that for a Gaussian  $\text{AR}(4)$  noise process for  $N \in \{10, 100\}$  and  $P_{FA} = 10^{-4}$  in Fig. 20. For the Gaussian case, we employ  $\hat{\mathbf{R}}_m$  in (8). The corresponding  $P_D$  can be expressed in closed

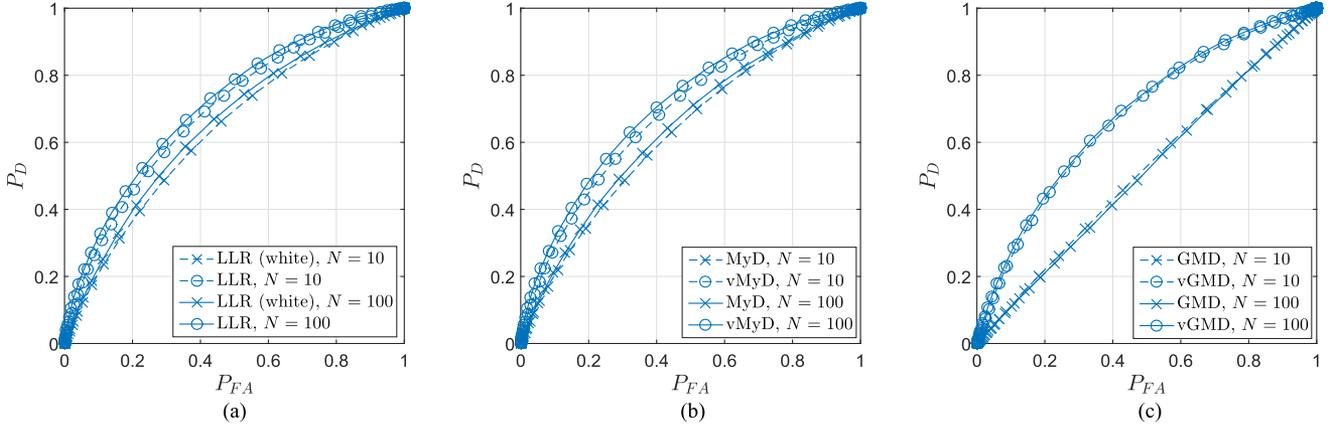


Fig. 17. ROC curves for  $\alpha = 1.5$ ,  $\text{SNR} = -10$  dB, and  $N \in \{10, 100\}$ . Performances of the LLR and LLR (white) detectors are shown in (a), the vMyD and MyD in (b), and the vGMD and GMD in (c).

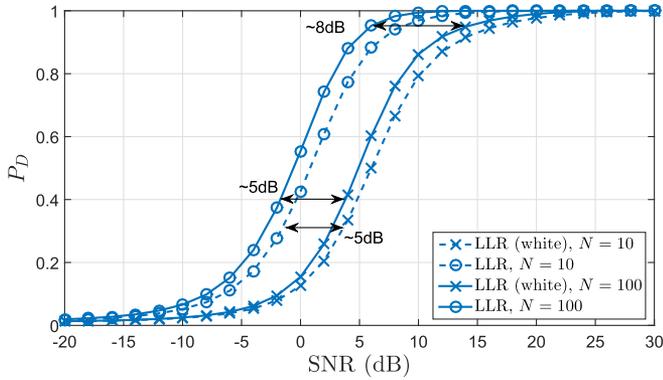


Fig. 18. Performance comparison between the LLR and LLR (white) detectors for  $P_{FA} = 10^{-2}$ ,  $N \in \{10, 100\}$ , and  $\alpha = 1.5$ .

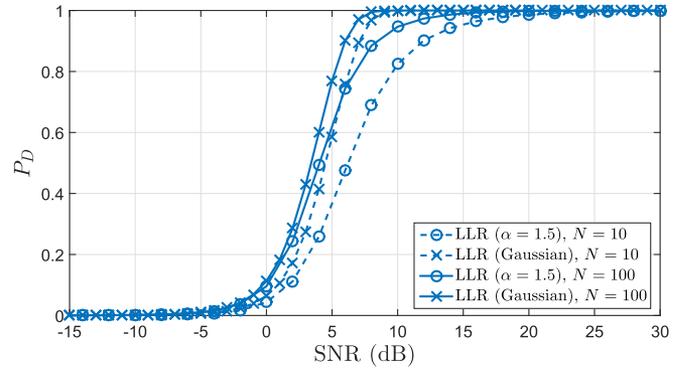


Fig. 20. Performance comparison of LLR detection in Gaussian AR(4) noise and  $\alpha\text{SGN}(4)$  for  $\alpha = 1.5$  and  $P_{FA} = 10^{-4}$ .

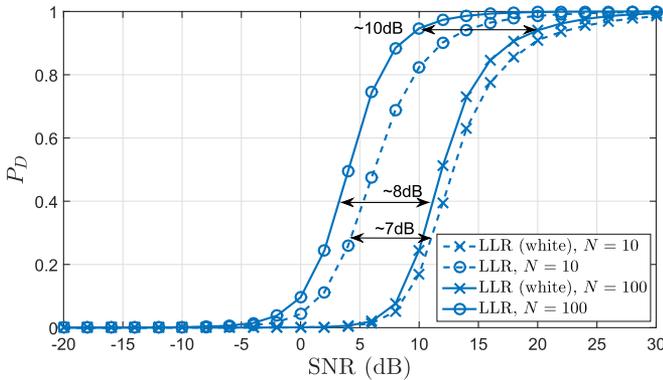


Fig. 19. Performance comparison between the LLR and LLR (white) detectors for  $P_{FA} = 10^{-4}$ ,  $N \in \{10, 100\}$ , and  $\alpha = 1.5$ .

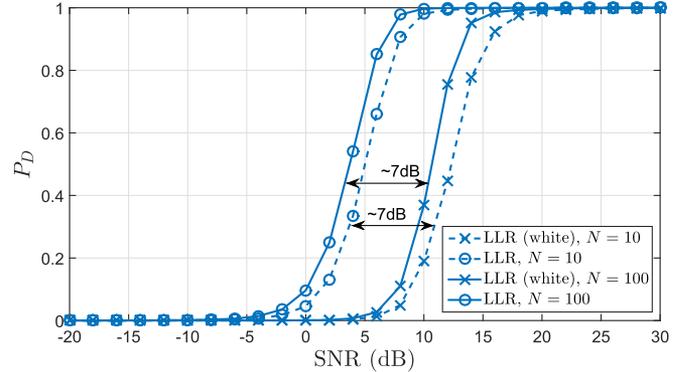


Fig. 21. Performance comparison between the LLR and LLR (white) detectors for  $P_{FA} = 10^{-4}$ ,  $N \in \{10, 100\}$ , and  $\alpha = 1.9$ .

form

$$P_D = Q \left( Q^{-1}(P_{FA}) - \frac{\mathcal{E}}{2\delta^2} \mathbf{s}^T (2\Sigma/\delta^2)^{-1} \mathbf{s} \right) \quad (43)$$

where  $Q(\cdot)$  is the tail probability of the standard Gaussian distribution and  $\Sigma$  is determined via (32) [21]. In Fig. 20, one can clearly see that the performance curves for  $\alpha = 1.5$  approach that of the Gaussian case as  $N$  increases. We also present the performance of the LLR detector for the  $\alpha = 1.9$  case, for

$N \in \{10, 100\}$ , and  $P_{FA} = 10^{-4}$  in Fig. 21. Most of the observations made in Fig. 19 extend to this case as well. However, we also note that the SNR gain is not observably dependent on the SNR. This can be attributed to the fact that for  $\alpha = 1.9$  the noise process is less impulsive. Therefore, a smaller  $N$  suffices to produce performance curves similar to those for the Gaussian case in Fig. 20.

To wrap up, we highlight the degradation in performance when the signal transmitted is in the lower half of the

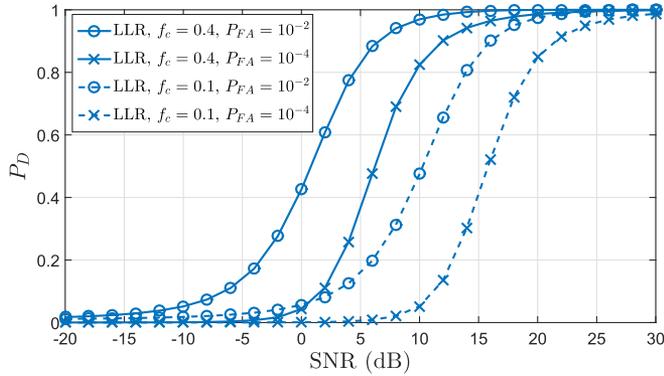


Fig. 22. Performance comparison between LLR detectors for  $f_c \in \{0.1, 0.4\}$ ,  $P_{FA} \in \{10^{-2}, 10^{-4}\}$ , and  $\alpha = 1.5$ .

spectrum. In Fig. 22, we plot the LLR detector performance for the  $\alpha = 1.5$  case for  $f_c \in \{0.1, 0.4\}$  and  $P_{FA} \in \{10^{-2}, 10^{-4}\}$ . The SNR gains between the respective curves due to good carrier placement are clearly significant.

## VIII. CONCLUSION

In our work, we considered the  $\alpha$ SGN( $m$ ) model, which characterizes not only the amplitude distribution of snapping shrimp noise in the warm shallow underwater channel but also the near-elliptical dependencies between closely spaced samples. We proposed and analyzed several novel robust detectors for the binary detection problem in  $\alpha$ SGN( $m$ ) and compared their performances with conventional robust detectors. Parameters of the noise model were tuned to estimates found for snapping shrimp noise. Detectors that exploited the dependency between adjacent samples performed significantly better than conventional robust detectors. In fact, by just taking the dependencies under consideration, detection performance was shown to be at par with that of the optimum detector. Similar trends were observed when  $\alpha$ SGN( $m$ ) was replaced by actual snapping shrimp noise. Moreover, signal design rules were proposed that should be used with the robust methods to truly harness the dependence within the noise samples.

## APPENDIX OPTIMIZING THE SIGNAL IN WS $\alpha$ SN

Reverting the inequality constraint in (31) to the original equality constraint, the corresponding problem may be expressed as the maximization

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\arg \max} - \underbrace{\sum_{i=1}^N \log f_W(\theta s_i)}_{h(\mathbf{s}, \lambda)} + \lambda(\|\mathbf{s}\|^2 - \mathcal{E}) \quad (44)$$

where  $\lambda$  is the Lagrange multiplier. Differentiating  $h(\mathbf{s}, \lambda)$  with respect to  $s_n$  and  $\lambda$  gives us

$$\frac{\partial h(\mathbf{s}, \lambda)}{\partial s_n} = -\frac{\theta f'_W(\theta s_n)}{f_W(\theta s_n)} + 2\lambda s_n$$

for all  $n \in \{1, 2, \dots, N\}$  and

$$\frac{\partial h(\mathbf{s}, \lambda)}{\partial \lambda} = \sum_{i=1}^N s_i^2 - \mathcal{E},$$

respectively, where  $f'_W(\cdot)$  is the first-order derivative of  $f_W(\cdot)$ . On equating the above equations to zero and simplifying, we get

$$\frac{s_n^2}{\mathcal{E}} \sum_{i=1}^N \left( \frac{f'_W(\theta s_i)}{f_W(\theta s_i)} \right)^2 = \left( \frac{f'_W(\theta s_n)}{f_W(\theta s_n)} \right)^2 \quad (45)$$

for all  $n \in \{1, 2, \dots, N\}$ . Clearly,  $s_n \in \{\pm\sqrt{\mathcal{E}/N}\} \forall n \in \{1, 2, \dots, N\}$  is a solution of (45). Moreover, as  $f_W(\cdot)$  is symmetric,  $f'_W(0) = 0$ . Thus, setting any  $K < N$  components of  $\mathbf{s}$  to zero and the remaining  $N - K$  components to  $\pm\sqrt{\mathcal{E}/(N - K)}$  are solutions as well. However, the latter implies overlap of a tail between the pdfs under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , which increases the cost in (31). Therefore,  $\hat{s}_n \in \{\pm\sqrt{\mathcal{E}/N}\} \forall n \in \{1, 2, \dots, N\}$  is the optimal solution.

## REFERENCES

- [1] A. Papoulis and U. S. Pillai, *Probability, Random Variables and Stochastic Processes*, Boston, MA, USA: McGraw-Hill, Dec. 2001.
- [2] J. Proakis and M. Salehi, *Digital Communications*, New York, NY, USA: McGraw-Hill, 2007.
- [3] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Energy Conv.*, vol. 44, no. 1, pp. 249–258, Feb. 2002.
- [4] W. W. L. Au and K. Banks, "The acoustics of the snapping shrimp *Squilla palpestris* in Kaneohe Bay," *J. Acoust. Soc. Amer.*, vol. 103, no. 1, pp. 41–47, 1998.
- [5] M. W. Legg, "Non-Gaussian and non-homogeneous poisson models of snapping shrimp noise," Ph.D. dissertation, Curtin Univ. Technol., Bentley, W.A., Australia, 2009.
- [6] A. Mahmood, "Digital communications in additive white symmetric alpha-stable noise," Ph.D. dissertation, Nat. Univ. Singapore, Singapore, Jun. 2014.
- [7] M. Chitre, J. Potter, and S. Ong, "Viterbi decoding of convolutional codes in symmetric  $\alpha$ -stable noise," *IEEE Commun. Lett.*, vol. 55, no. 12, pp. 2230–2233, Dec. 2007.
- [8] P. Huber and E. Ronchetti, *Robust Statistics*, ser. Probability and Statistics, New York, NY, USA: Wiley, 2009.
- [9] G. Arce, *Nonlinear Signal Processing: A Statistical Approach*, ser. Online Books, New York, NY, USA: Wiley-Interscience, 2005.
- [10] 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, Jan. 2002.
- [11] 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, Apr. 2006.
- [12] C. L. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*, New York, NY, USA: Chapman-Hall, 1996.
- [13] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*, New York, NY, USA: Chapman & Hall, 1994.
- [14] M. Chitre, J. Potter, and O. S. Heng, "Underwater acoustic channel characterisation for medium-range shallow water communications," in *Proc. MTT/IEEE TECHNO-OCEAN Conf.*, vol. 1, Nov. 2004, pp. 40–45.
- [15] J. Park, G. Shevlyakov, and K. Kim, "Maximin distributed detection in the presence of impulsive alpha-stable noise," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1687–1691, 2011.
- [16] A. Rajan and C. Tepedelenlioglu, "Diversity combining over Rayleigh fading channels with symmetric alpha-stable noise," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2968–2976, 2010.

- [17] 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 *Proc. OCEANS Conf.—Europe*, 2007, DOI: 10.1109/OCEANSE.2007.4302279.
- [18] A. Mahmood and M. Chitre, "Modeling colored impulsive noise by Markov chains and alpha-stable processes," in *Proc. Oceans Conf.*, May 2015, DOI: 10.1109/OCEANS-Genova.2015.7271550.
- [19] J. P. Nolan, "Multivariate elliptically contoured stable distributions: Theory and estimation," *Comput. Stat.*, vol. 28, no. 5, pp. 2067–2089, 2013.
- [20] A. Mahmood, M. Chitre, and M. A. Armand, "On single-carrier communication in additive white symmetric alpha-stable noise," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3584–3599, Oct. 2014.
- [21] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, ser. Signal Processing, Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [22] A. Mahmood and M. Chitre, "Generating random variates for stable sub-Gaussian processes with memory," *Signal Process.*, to be published.
- [23] M. Roth, "On the multivariate t distribution," Div. Autom. Control, Linköping Univ., Linköping, Sweden, *Tech. Rep.*, Apr. 2013.
- [24] T. Aysal and K. Barner, "Meridian filtering for robust signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 3949–3962, 2007.
- [25] J. Gonzalez, J. Paredes, and G. Arce, "Zero-order statistics: A mathematical framework for the processing and characterization of very impulsive signals," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3839–3851, Oct. 2006.
- [26] J. G. Gonzalez, "Robust techniques for wireless communications in non-gaussian environments," Ph.D. dissertation, Univ. Delaware, Newark, DE, USA, 1997.
- [27] S. Kotz and S. Nadarajah, *Multivariate t-Distributions and Their Applications*, London, U.K.: Cambridge Univ. Press, 2004.
- [28] M. Noemm and P. A. Hoehner, "CutFM sonar signal design," *Appl. Acoust.*, vol. 90, pp. 95–110, 2015.
- [29] J. Hamilton, *Time Series Analysis*, Princeton, NJ, USA: Princeton Univ. Press, 1994.
- [30] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Upper Saddle River, NJ, USA: Pearson/Prentice-Hall, 2005, vol. 452.
- [31] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.
- [32] A. Mahmood, M. Chitre, and M. A. Armand, "PSK communication with passband additive symmetric  $\alpha$ -stable noise," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2990–3000, Oct. 2012.



**Ahmed Mahmood** (S'11–M'14) received the B.Eng. and M.Eng. degrees in electrical engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan and the Ph.D. degree from the Department of Electrical & Computer Engineering, National University of Singapore (NUS), Singapore.

In 2014, he joined the Acoustic Research Laboratory (ARL), NUS, where he is now a Research Fellow. His current research interests involve underwater acoustic communications, robust signal processing, channel modeling, estimation and detection theory, and error-correction coding for impulsive noise.

Dr. Mahmood has served as a reviewer for several technical journals, including the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE JOURNAL OF OCEANIC ENGINEERING, IEEE COMMUNICATIONS LETTERS, and IEEE SIGNAL PROCESSING LETTERS. He was the Technical Program Committee Chair for the 2016 Workshop on Underwater Networks (WUWNet'16).



**Mandar Chitre** (S'04–M'05–SM'11) received the B.Eng. and M.Eng. degrees in electrical engineering from the National University of Singapore (NUS), Singapore, in 1997 and 2000, respectively, the M.Sc. degree in bioinformatics from the Nanyang Technological University (NTU), Singapore, in 2004, and the Ph.D. degree from NUS, in 2006.

From 1997 to 1998, he worked with the Acoustic Research Laboratory (ARL), NUS. 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. He also holds a joint appointment with the Department of Electrical and Computer Engineering at NUS as an Associate Professor. His current research interests are underwater communications, underwater acoustic signal processing, and marine robotics.

Dr. Chitre has served on the technical program committees (TPCs) 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 OCEANS'06 in Singapore, the chairman for the IEEE Singapore AUV Challenge 2013, and the TPC Chair for the WUWNet'16 conference. He is currently the IEEE Ocean Engineering Society Technology Committee Cochair of Underwater Communication, Navigation, and Positioning.