Uncoded Acoustic Communication in Shallow Waters with Bursty Impulsive Noise

Ahmed Mahmood and Mandar Chitre
Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore
e-mail: {tmsahme, mandar}@nus.edu.sg

Abstract—The shallow underwater acoustic channel offers a challenging environment. Besides long delay spreads caused by multiple surface-bottom reflections, the channel is time variant as well. In tropical waters, the problem is compounded further by impulsive noise created by snapping shrimp. Conventionally, the noise process is modeled by white impulsive noise. However, in reality, snapping shrimp noise depicts memory and is thus bursty as well. We investigate the performance of a high-rate uncoded acoustic communication system operating in tropical shallow waters. The stationary α-sub-Gaussian noise with memory order m (αSGN(m)) model is employed as it characterizes both the temporal and amplitude statistics of snapping shrimp noise. We show that there is a stark deviation from the expected performance of the white maximum-likelihood (ML) detector when the ambient noise is αSGN(m) (m > 0). Moreover, we derive the ML detector for the shallow water acoustic channel with additive αSGN(m) and compare its performance to that of its white counterpart.

I. INTRODUCTION

The soundscape in warm shallow waters is impulsive due to the presence of snapping shrimp [1]–[3]. These crustaceans live in large colonies and are able to create sharp snaps (or impulses) by cavitating bubbles [4]. Snapping shrimp noise tends to be the dominant source of ambient noise at frequencies greater than 2kHz [3] and is thus an unavoidable impairment for acoustic communications in shallow tropical waters [5]. The process exhibits non-Gaussian statistics and is modeled effectively by heavy-tailed distributions [1]. In fact, the amplitude statistics of snapping shrimp noise are known to be tracked very well by the non-Gaussian symmetric α-stable (SαS) distribution family [1], [6]. Besides being impulsive, closely-spaced samples of the snapping shrimp noise process are dependent [1], [7]. The implicit memory causes outliers to cluster together which results in bursty impulsive noise.

In the literature, white noise models are typically employed to model impulsive noise [8], [9]. An example is the white SαS noise (WSαSN) process, which has been used extensively to model snapping shrimp noise [6], [8]–[11]. As it constrains samples to be independent and identically distributed (IID) SαS random variables, WSαSN can only model amplitude statistics and is unable to characterize memory [1], [7]. Consequently, communication schemes optimized for WSαSN are sub-optimal in snapping shrimp noise. Though such schemes are robust to impulses and significantly outperform conventional detectors in WSαSN [8], [10], their performance will deviate from theoretical or simulation results in practice.

To highlight this deviation, we employ the stationary α-sub-Gaussian noise with memory order m (αSGN(m)) model [7] in our work. Recent results show the dependence between adjacent samples of snapping shrimp noise to exhibit near-elliptic structures [7]. The αSGN(m) process constrains m + 1 adjacent samples to be a multivariate α-sub-Gaussian (αSG) distribution. As αSG distributions are elliptic, they model the dependence very well. Moreover, αSG distributions are a special class of the heavy-tailed SαS family and thus any sample of αSGN(m) is a SαS stable random variable [7], [12]. Consequently, the model also tracks the amplitude statistics of the snapping shrimp noise process. In fact for m = 0, αSGN(m) is equivalent to WSαSN and effectively highlights the scenario when snapping shrimp noise samples are far apart.

The objective of this paper is to investigate the uncoded error performance of an acoustic communication scheme operating in shallow waters with αSGN(m) (m > 0). Based on existing literature, we model the shallow underwater channel as a sparse Rayleigh block fading channel [5], [13]. We consider three eigenrays corresponding to the direct arrival, surface reflection and bottom bounce. We also take into account the variability in delays of the eigenrays and tune them to empirical estimates recorded in [5]. Moreover, the αSGN(m) model is also tuned to snapping shrimp noise data. A high-rate single-carrier binary phase shift keying (BPSK) scheme is considered and a decision feedback equalizer (DFE) is invoked to neutralize intersymbol interference (ISI). The DFE is used in conjunction with a maximum-likelihood (ML) symbol detector. We investigate the error performance of the white ML detector (optimal in WSαSN) in our simulations and highlight its deviation in error performance from the WSαSN case. Further still, the ML detector for general additive αSGN(m) is derived and its performance compared with the former. Results are compiled for severely and slightly impulsive scenarios. In either case, the gain achieved by taking the noise memory into account is significant.

This paper is organized as follows: In Section II we present the system model and the αSGN(m) process. In Section III we discuss the ML detectors for both the WSαSN and αSGN(m) cases. We wrap up by presenting our results in Section IV.

II. SYSTEM MODEL

A. Signal Transmission

Due to hardware and channel constraints, the transmitted passband signal in underwater acoustic communication is
typicaly limited to a few tens/hundreds of kHz [14], [15]. As this results in a low Nyquist rate, one may directly sample the passband signal rather than convert it via linear analog circuits to baseband form [15]. The former approach is preferable in warm shallow waters as linear systems are very suboptimal in impulsive noise. The passband samples may then be processed robustly via suitable non-linear techniques to great effect [8], [10]. We therefore work directly with the sampled passband received signal.

Let \( x_k \in \{ \pm 1 \} \) denote the transmitted BPSK symbol in the \( k^{th} \) signaling interval. Also, let \( f_c = \xi/T \) and \( f_s = N/T \) denote the carrier and passband sampling frequencies for some \( \xi, N \in \mathbb{Z}^+ \), respectively, where \( 1/T \) is the bit rate. Then the discretized passband signal is given by

\[
s[n] = \sum_k x_k \phi[n - kN] u_N[n - kN],
\]

(1)

where \( \phi[n] = \sqrt{\frac{2}{T \varphi}} g[n] \cos(2\pi \xi/Nn) \) and

\[
u_N[n] = \begin{cases} 1 & \text{for } 0 \leq n \leq N - 1, \\ 0 & \text{o.w.} \end{cases}
\]

Moreover, \( g[n] \) is the discrete baseband pulse shaping signal and \( \mathcal{E}_g \) is the energy of \( g[n] \), i.e., \( \mathcal{E}_g = \sum_{n=0}^{N-1} g[n]^2 \) [16]. As \( f_s = N/T \), there are \( N \) samples per transmitted symbol. Thus, the \( k^{th} \) signaling interval, i.e., samples consisting solely of \( x_k \), is given by \( s_{k,i} = s[kN + i] \forall i \in \{0, 1, \ldots, N - 1\} \). To satisfy the Nyquist criterion, the condition

\[
f_s > 2 \left( f_c + \frac{\beta}{T} \right) \Rightarrow N > 2(\xi + \beta)
\]

(2)

must be met, where \( \beta \in \mathbb{R}^+ \) is an excess bandwidth parameter that depends on \( g[n] \).

**B. The Passband Channel Model**

In an underwater acoustic communication scheme, the packet length is typically larger than the coherence time of the channel. Consequently, the channel changes within a packet but is approximately constant over a few symbols. Coherence times run into hundreds of milliseconds and are rarely less than 100 ms [14]. Though the acoustic channel is typically modeled as a linear time variant system, it can be effectively modeled as linear time invariant for a symbol block spanning less than its coherence time. For each such block, the received passband signal is

\[
r[n] = h[n] * s[n] + w[n],
\]

(3)

where \( h[n] \) is the \( L \)-tap (causal) channel impulse response, \( w[n] \) is the ambient noise of the channel and \( * \) is the linear convolution operator. If the first tap of \( h[n] \) corresponds to the direct arrival, then the \( p^{th} \) signaling interval is given by

\[
r_{p,i} = r[pN + i] \forall i \in \{0, 1, \ldots, N - 1\}.
\]

More precisely,

\[
r_{p,i} = \sum_{\tau=0}^{L-1} h[\tau] s[pN + i - \tau] + w_{p,i},
\]

(4)

On substituting (1) in (4) and simplifying, the convolution term may be expressed as

\[
\sum_{\tau=0}^{L-1} h[\tau] s[pN + i - \tau] = x_p \sum_{\tau=0}^{L-1} h[\tau] \phi[i - \tau] u_N[i - \tau]
\]

\[
+ \sum_{k, k \neq 0} x_{p-k} \sum_{\tau=0}^{L-1} h[\tau] \phi[kN + i - \tau] u_N[kN + i - \tau].
\]

Without any loss in generality we assume \( L \geq N \). This is valid for a high-rate communication scheme in shallow waters as the delay spread \( L \) is large and causes severe ISI. Moreover, for the case of \( L < N \), one may append zeros to \( s[n] \) so that it satisfies the assumption. Using this and the structure of \( u_N[n] \), we have

\[
\sum_{\tau=0}^{L-1} h[\tau] s[pN + i - \tau] = x_p \sum_{\tau=0}^{L-1} h[\tau] \phi[i - \tau]
\]

\[
+ \sum_{k=1}^{\lfloor \frac{L-2}{N} \rfloor + 1} x_{p-k} \sum_{\tau=0}^{L-1} h[\tau] \phi[kN + i - \tau],
\]

(5)

which allows us to express (4) as

\[
r_{p,i} = r_{p,i} - \gamma_{p,i} = x_p \ell_i + w_{p,i} + \epsilon_{p,i},
\]

(6)

where \( \gamma_{p,i} \) is the ISI term at index \( n = pN + i \).

One notes that the decision \( \hat{x}_p \) for \( x_p \) is made only after the samples \( r_{p,i} \forall i \in \{0, 1, \ldots, N - 1\} \) are received. By employing a DFE, the ISI is estimated by passing the immediately previous \( \lfloor \frac{L-2}{N} \rfloor + 1 \) decisions through a filter. This is then subtracted from \( r[n] \) to get

\[
\hat{r}_{p,i} = r_{p,i} - \hat{\gamma}_{p,i} = x_p \ell_i + w_{p,i} + \epsilon_{p,i},
\]

(7)

and \( \epsilon_{p,i} \) is a possible residual error term due to at least one of the previous \( \lfloor \frac{L-2}{N} \rfloor \) + 1 terms being in error. The receiver schematic depicting the aforementioned scheme is presented in Fig. 1. The corresponding error performance depends greatly on the symbol-by-symbol detector, which in turn should be robust to the noise process \( w[n] \). If not optimized, the receiver may suffer from error propagation (non-zero \( \epsilon_{p,i} \)) even at high signal-to-noise ratio (SNR) due to repeated incorrect decisions made by the detector [16].

In the literature, typical implementations of a DFE includes a feedforward filter [16]–[18]. As the objective of this paper is to highlight the detector’s impact on receiver performance in \( \alpha\text{SGN}(m) \) for a given equalization scheme, we intentionally employ the minimalistic receiver structure in Fig. 1. In this case, the feedforward filter is essentially a unit-weight one-tap filter. Moreover, for any corrected \( \lfloor \frac{L-2}{N} \rfloor + 1 \) adjacent symbols,
the DFE completely removes ISI in a no noise scenario for all future symbols. This is highlighted by (5) & (6).

C. The αSGN(m) Process

The αSGN(m) model characterizes both the amplitude and temporal statistics of the snapping noise process [7]. The model is based on a sliding window framework and constrains any immediately adjacent $m + 1$ samples to be αSG [7], [12], thus allowing individual noise samples to be heavy-tailed SoS and those spaced less than or equal to $m + 1$ samples apart jointly elliptic. Mathematically, if $\hat{W}_{n,m} = [W_{n-m}, W_{n-m+1}, \ldots, W_n]^T$ is a random vector consisting of the current sample (at index $n$) and $m$ immediately previous samples of αSGN(m), then

$$\hat{W}_{n,m} \overset{d}{=} A_n^{1/2} \tilde{G}_{n,m} \tag{8}$$

where $d$ denotes equality in distribution [7], [12]. The random variable $A_n$ is totally right-skewed heavy-tailed stable and $\tilde{G}_{n,m} = [G_{n-m}, G_{n-m+1}, \ldots, G_n]^T$ is a zero-mean $(m + 1)$-dimensional Gaussian random vector with covariance matrix $C_m = [c_{ij}], i.e., \tilde{G}_{n,m} \sim \mathcal{N}(0, C_m) \forall n \in \mathbb{Z}$.

As the αSGN(m) process is stationary, $C_m$ and the statistics of $A_n$ do not vary with time [7]. Moreover, due to the sliding window framework $C_m$ is a symmetric Toeplitz matrix [7].

The distribution of a stable random variable depends on four parameters and may be denoted succinctly by $S(\alpha, \beta, \delta, \mu)$, where $\alpha \in (0, 2]$ is the characteristic exponent, $\beta \in [-1, 1]$ is the skew parameter, $\delta \in (0, \infty)$ is the scale and $\mu \in (-\infty, \infty)$ is the location of the distribution [19], [20]. Moreover, if $\beta = \mu = 0$, a stable random variable is SoS and thus may be represented as $S(\alpha, \delta)$ [8], [19]. As each sample of αSGN(m) is an SoS random variable, we have $W_n \sim S(\alpha, \delta_w) \forall n \in \mathbb{Z}$. One way to ensure this is to set $A_n \sim S(\frac{\alpha}{2}, 1, 2(\cos(\frac{\pi \alpha}{2}))^{1/\alpha}, 0)$ and $c_{ij} = c_{ji} = \delta_{ij}^2 \forall i, j \in \{1, 2, \ldots, m + 1\}$ [7], [12]. We adhere to the above parameterizations in our text.

On a final note, we see that the sliding window framework ensures that any current sample depends solely on the previous $m$ samples [7]. Thus the αSGN(m) process is Markovian of order $m$. A special case is that of $m = 0$, where αSGN(0) simplifies to a WSoSN process. In such a case, the samples $W_n \sim S(\alpha, \delta_w)$ are IID random variables. As we see next, the Markov and stationarity properties of αSGN(m) allow mathematical tractability and a simpler form of the optimal detector.

III. ROBUST DETECTION IN αSGN(m)

If $w[n]$ is assumed to be a WSoSN process, then given (6), the ML detector in the $p^\text{th}$ signaling interval is

$$\hat{x}_p = \arg \min_{\zeta \in \{\pm 1\}} \sum_{i=0}^{N-1} -\log f_{W_i}(\hat{r}_{p,i} - \zeta \ell_i), \tag{9}$$

where $f_{W_i}(\cdot)$ is the probability density function (PDF) corresponding to $S(\alpha, \delta_w)$. We term (9) as the white ML (wML) detector and note that it does not take the dependence between noise samples into account. It is therefore sub-optimal in αSGN(m) for $m > 0$.

For general αSGN(m), it is preferable to express the PDF of $N$ consecutive noise samples in suitable form before we derive the ML detector. Let $\hat{W}_N = [W_0, W_1, \ldots, W_{N-1}]^T$ be a random vector of $N$ such samples and $f_{\hat{W}_N}(\cdot)$ its PDF. Moreover, let $w_N = [w_0, w_1, \ldots, w_{N-1}]^T$ be a sample outcome. Then from the chain rule of probability [21],

$$f_{\hat{W}_N}(w_N) = \prod_{i=0}^{N-1} f_{W_i|\hat{W}_{i-1}}(w_i|w_{i-1}) \tag{10}$$

As $W_n$ is an αSGN(m) process, it is Markov of order $m$. Therefore, (10) may be simplified to

$$f_{\hat{W}_N}(w_N) = \prod_{i=0}^{m-1} f_{W_i|\hat{W}_i}(w_i|w_{i-1}) \times \prod_{i=m}^{N-1} f_{W_i|\hat{W}_{i-1,m-1}}(w_i|w_{i-1,m-1}), \tag{11}$$

where $w_{i,m} = [w_{i-m}, w_{i-m+1}, \ldots, w_i]^T$ is a vector of $m + 1$ immediately adjacent noise samples at index $i$. Finally, as the process is stationary, we get

$$f_{\hat{W}_N}(w_N) = \prod_{i=0}^{m-1} f_{W_i|\hat{W}_i}(w_i|w_{i-1}) \times \prod_{i=m}^{N-1} f_{W_i|\hat{W}_{i-1,m-1}}(w_i|w_{i-1,m-1}). \tag{12}$$

Note that $f_{W_m|\hat{W}_{m-1}}(\cdot)$ can be expressed in terms of $f_{W_{m+1}}(\cdot)$ and $f_{\hat{W}_{m-1}}(\cdot)$, both of which are multivariate αSG distributions [12], [22]. Similarly, $f_{W_i|\hat{W}_{i}}(\cdot)$ for $i \in \{0, 1, \ldots, m - 1\}$ can be expressed in terms of $f_{\hat{W}_{i+1}}(\cdot)$ and $f_{W_i}(\cdot)$ which again are multivariate αSG. Letting $w_i = w_{p,i} \forall i \in \{0, 1, \ldots, N-1\}$, the ML detector in the $p^{th}$ signaling interval in αSGN(m) is

$$\hat{x}_p = \arg \min_{\zeta \in \{\pm 1\}} \left\{ \sum_{i=0}^{m-1} f_{W_i|\hat{W}_i}(w_{p,i}|w_i) + \sum_{i=m}^{N-1} \log f_{W_i|\hat{W}_{i-1,m-1}}(w_{p,i}|w_{i-1,m-1}) \right\} \tag{13}$$

where from (6) we have $w_{p,i} = \hat{r}_{p,i} - \zeta \ell_i$.

Now that we have discussed the wML and ML detectors, we compare their respective error performance in the warm shallow underwater channel. The simulation setup is discussed next and the model is tuned to empirical results.
IV. RESULTS & DISCUSSION

A. Experimental Setup

For our simulations, we consider an \( \alpha \text{SGN}(4) \) process with the normalized covariance matrix \( \hat{C}_4 = C_4 / \delta_w^2 \) given by

\[
\hat{C}_4 = \begin{bmatrix}
0.0000 & 0.5804 & 0.2140 & 0.1444 & -0.0135 \\
0.5804 & 1.0000 & 0.5804 & 0.2140 & 0.1444 \\
0.2140 & 0.5804 & 1.0000 & 0.5804 & 0.2140 \\
0.1444 & 0.2140 & 0.5804 & 1.0000 & 0.5804 \\
-0.0135 & 0.1444 & 0.2140 & 0.5804 & 1.0000
\end{bmatrix}.
\]

This setting for \( m \) and \( \hat{C}_m \) is based on estimates from snapping shrimp data sets observed at \( f_s = 180 \text{kHz} \) [7].

We consider a scenario where the transmitter and receiver are placed 1000m apart and 5m below the surface. The sea floor is assumed to be at a constant 20m depth from the water surface. For the shallow underwater channel, \( h[n] \) is known to be sparse. We consider three eigenpaths, namely the direct arrival, surface reflection and bottom bounce. Assuming acoustic velocity in water to be 1500m/s and \( f_s = 180 \text{kHz} \), the three non-zero taps of \( h[n] \) can be evaluated from the geometry of the channel. These are located at \( n \in \{0, 6, 54\} \). Based on the empirically supported formulation in [5], [13], each of these taps undergo independent Rayleigh fading. Moreover, the time-varying channel also cause tap migration, with eigenrays at larger delays more prone to its effects [5]. Following the arrival spread results stated in [5, pg. 6], we assume no variability of the direct arrival between different blocks. However, the second and third arrivals may excite any one of the taps \( n \in \{5, 6, 7\} \) and \( n \in \{47, 49, \ldots, 61\} \), respectively. The probability of a particular eigenray exciting a tap follows a Gaussian distribution [5].

As highlighted in Section II-B, the channel is constant over the transmission of a few symbols (defined as a symbol block). By defining the block length as less than half the coherence time, the DFE should be able to update its taps effectively after every block transmission to process the subsequent block. In such a way it is able to process long packets (consisting of several blocks). We employ \( N = 10 \) and a block length of 1000 symbols. With \( f_s = 180 \text{kHz} \), this results in a block length of 55.56ms. As highlighted in [14], coherence times of underwater acoustic channels are rarely less than 100ms. Thus our choice of block length is sufficient for our cause.

On another note, we see that errors in the channel estimate will negatively impact the receiver performance in Fig. 1, even in a no noise scenario. Though such analysis exists in the literature, we are interested in investigating the effect of \( \alpha \text{SGN}(m) \) and the detector as an isolated phenomenon in our work. Therefore, for our preliminary analysis, we assume complete knowledge of \( h[n] \) for a symbol block. Robust channel estimation in snapping shrimp noise can be tackled as an independent problem [11], [23] and incorporated as part of the receiver later on. The channel taps are assumed to fade/tap migrate independently after each symbol block. On a final note, we employ \( \xi = 2 \) and a raised cosine pulse for \( g[n] \).

This allows a potential symbol rate of \( 1/T = 18\text{kHz} \) and ensures (2) is satisfied.

B. Simulation Results

To quantify our results we employ the signal-to-noise ratio (SNR) measure \( \mathcal{E}_b/(4\delta_w^2) \), where \( \mathcal{E}_b \) is the average energy per received bit [8], [10]. The latter is evaluated by \( \mathcal{E}_b = N \int_{-1/2}^{1/2} S_a(f)S_h(f)df \), where \( S_a(f) \) and \( S_h(f) \) are the two-sided power spectral densities (PSDs) of the cyclostationary signal \( s[n] \) and the random channel \( h[n] \), respectively.

In Fig. 2, we plot the bit error rate (BER) curves of the \( w\text{ML} \) and ML detectors for a single-carrier BPSK system in \( \alpha \text{SGN}(4) \) with \( \alpha = 1.5 \). This value of \( \alpha \) illustrates severe snapping noise [1], [6]. For comparison, the performance of the DFE receiver in WSnoSN is also shown. Though robust to outliers, one can clearly see the expected BER of the \( w\text{ML} \) detector to deviate significantly if the ambient noise process depicts memory. Moreover, with increasing SNR, the difference between both error curves increases. On the other hand, as expected, the ML detector offers better performance than the \( w\text{ML} \) detector. Further still, its BER decreases at a sharper rate than that of its white counterpart. At a BER of \( 10^{-4} \), the gain is \( \sim 8 \text{dB} \) and is highlighted in the figure.

In Fig. 3, we again present the BER performance of the \( w\text{ML} \) and ML detectors but for \( \alpha = 1.9 \). This represents a channel with mild snapping shrimp noise [9]. The trends are more or less similar to those in Fig. 2, but with the \( w\text{ML} \) detector performance deviating from its expected BER curve at medium-to-high SNR, i.e., \( \text{SNR} > 15\text{dB} \). We also note that the ML detector offers \( \sim 6.2\text{dB} \) gain over its white counterpart at BER = \( 10^{-4} \), thus signifying the advantage of taking the noise memory in consideration.
REFERENCES