LOS and NLOS Classification for Underwater Acoustic Localization

Roee Diamant*, Hwee-Pink Tan† and Lutz Lampe*

*Electrical and Computer Engineering, The University of British Columbia
†Institute for Infocomm Research, Singapore

Email: {roeed, lampe}@ece.ubc.ca and hptan@i2r.a-star.edu.sg

Abstract

The low sound speed in water makes propagation delay (PD) based range estimation attractive for underwater acoustic localization (UWAL). However, due to the long channel impulse response and the existence of reflecting objects, PD-based UWAL suffers from significant degradation when PD measurements of non-line-of-sight (NLOS) communication links are falsely identified as having originated from line-of-sight (LOS) communication links. In this paper, we consider this problem and present a two-step algorithm to classify PD measurements into LOS and NLOS links, relying on the availability of propagation delay and received signal strength measurements for a single transmitter-receiver pair. In the first step, by comparing signal strength-based and PD-based range measurements, we identify object-related NLOS (ONLOS) links, where signals are reflected from objects with high reflection loss, e.g., ships hull, docks, rocks, etc. In the second step, excluding PD measurements related to ONLOS links, we use a constrained expectation-maximization algorithm to classify PD measurements into two classes: LOS and sea-related NLOS (SNLOS), and to estimate the statistical parameters of each class. Since our classifier relies on models for the underwater acoustic channel, which are often simplified, alongside simulation results, we validate the performance of our classifier based on measurements from three sea trials. Both our simulation and sea trial results demonstrate a high detection rate of ONLOS links, and accurate classification of PD measurements into LOS and SNLOS.

Index Terms


Parts of this work have been presented at the IEEE Oceans Conference, Sep. 2010, Seattle, USA.
I. INTRODUCTION

Underwater acoustic communication networks (UWAN) are envisaged to fulfill the needs of a multitude of applications such as navigation aids, early warning systems for natural disasters, ecosystem monitoring and military surveillance [1]. The data derived from UWAN is typically interpreted with reference to a node’s location, e.g., reporting an event occurrence, tracking a moving object or monitoring a region’s physical conditions. However, localization for underwater nodes is non-trivial. Since GPS signals do not propagate through water, localization of unlocalized nodes is often based on underwater acoustic communication and triangulation using a set of anchor nodes with known locations. This underwater acoustic localization (UWAL) typically employs propagation delay (PD) measurements for range estimation, i.e., time of arrival (ToA) or time difference of arrival (TDoA) of received signals [2], since angle of arrival methods would require multiple hydrophones and signal-strength based methods would fail due to inaccurate propagation models.

Existing UWAL schemes, e.g., [3], [4], [5], implicitly assume that PD measurements correspond to the line-of-sight (LOS) link between the transmitter and receiver. However, signals can arrive from non-LOS (NLOS) communication links in several ways, as illustrated in Figure 1. For the node pairs \((u, a_2)\) and \((u, a_3)\), sea surface and bottom reflections links (referred to as sea-related NLOS (SNLOS)) exist, respectively, in addition to an LOS link. For \((u, a_1)\), the signal arrives from the reflection off a rock (referred to as object-related NLOS (ONLOS)). Lastly, between nodes \(u\) and \(a_2\), there is also an ONLOS link due to a ship. While it is expected that power attenuation in the LOS link is smaller than in NLOS links, it is common that the LOS signal is not the strongest. This is because, as shown in multipath models [6] as well as measurements [7], the underwater acoustic channel (UWAC) consists of groups of NLOS links with small path delay, but significant phase differences, often resulting in negative superposition with the LOS link (if delay differences between the LOS and NLOS links are smaller than the system resolution for path separation) as well as positive superposition between NLOS links. If PD measurements of NLOS links are mistakenly treated as corresponding to delay in the LOS link, e.g., in node pairs \((u, a_2)\) and \((u, a_3)\), localization accuracy will significantly be degraded.

In this paper, we propose a two-step algorithm to classify a vector of PD measurements for a single transmitter-receiver distance into three classes: LOS, SNLOS and ONLOS, which is
a problem that has not been treated in previous literature. Such a classification can improve the accuracy of UWAL by either rejecting NLOS-related PD measurements, correcting them, or using them to bound range estimation. We first identify ONLOS-related PD-measurements by comparing PD-based range estimations with range estimations obtained from received-signal-strength (RSS) measurements. Considering the difficulty in acquiring an accurate attenuation model, our algorithm requires only a lower bound for RSS-based distance estimations. After excluding PD measurements related to ONLOS, we apply a constrained expectation-maximization (EM) algorithm to further classify the remaining PD measurements into LOS and SNLOS, and estimate the statistical parameters of both classes to improve the accuracy of UWAL. Results from extensive simulations and three sea trial experiments in different areas of the world demonstrate the efficacy of our approach through achieving a high detection rate for ONLOS links and good classification of non-ONLOS related PD measurements into LOS and SNLOS.

It should be noted that while our approach can also be adapted to other types of fading channels, it is particularly suited for UWAL for the following reasons. First, our algorithm relies on significant power absorption due to reflection loss in ONLOS links, which are typical in the underwater environment. Second, we assume that the difference in propagation delay between signals traveling through SNLOS and LOS links is noticeable, which is acceptable in the UWAC due to the low sound speed in water (approximately 1500 m/sec). Third, our algorithm is particularly beneficial in cases where NLOS paths are often mistaken for the LOS path, which occurs in UWAL, where the LOS path is frequently either not the strongest or non-existent. Last, we assume that the variance of PD measurements originating from SNLOS links is greater than that of measurements originating from LOS, which fits channels with long delay spread such as the UWAC.

The remainder of this paper is organized as follows. Related work on PD-based underwater localization is described in Section II. Our system model and assumptions are introduced in Section III. In Section IV, we present our approach to identify ONLOS links. Next, in Section V, we formalize the EM algorithm to classify non-ONLOS related PD measurements into LOS and SNLOS. Section VI includes performance results of our two-step algorithm obtained from synthetic UWAC environments (Section VI-A) and from three different sea trials (Section VI-B). Finally, conclusions are offered in Section VII. The key notations used in this paper are summarized in Table I.
TABLE I: List of key notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i$</td>
<td>PD measurements</td>
</tr>
<tr>
<td>$X$</td>
<td>vector of PD measurements $x_i$ of the same communication link</td>
</tr>
<tr>
<td>$d$</td>
<td>transmission distance</td>
</tr>
<tr>
<td>$d^{PD}$</td>
<td>PD-based range estimation</td>
</tr>
<tr>
<td>$d_{RSS\text{,min}}$</td>
<td>lower bound of RSS-based range measurement</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>propagation loss factor</td>
</tr>
<tr>
<td>$\gamma_{\text{max}}$</td>
<td>upper bound for $\gamma$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>absorption loss factor</td>
</tr>
<tr>
<td>$M$</td>
<td>assumed number of classes in PD model</td>
</tr>
<tr>
<td>$k_m$</td>
<td>weight of the $m$th distribution in the mixture distribution model</td>
</tr>
<tr>
<td>$\omega_m$</td>
<td>vector of parameters of the $m$th distribution</td>
</tr>
<tr>
<td>$\theta$</td>
<td>distribution parameters of $X$</td>
</tr>
<tr>
<td>$\beta_m$</td>
<td>power coefficient of $\omega_m$</td>
</tr>
<tr>
<td>$\sigma_m$</td>
<td>scale coefficient of $\omega_m$</td>
</tr>
<tr>
<td>$\mu_m$</td>
<td>mean coefficient of $\omega_m$</td>
</tr>
<tr>
<td>$T_{\text{LIR}}$</td>
<td>upper bound on the length of the channel impulse response</td>
</tr>
<tr>
<td>$c$</td>
<td>propagation speed in the channel</td>
</tr>
<tr>
<td>$x_{\text{LOS}}$, $d_{\text{LOS}}$</td>
<td>delay and distance in the LOS link, respectively</td>
</tr>
<tr>
<td>$d_{\text{ONLOS}}$</td>
<td>distance of the ONLOS link</td>
</tr>
<tr>
<td>RL</td>
<td>reflection loss in ONLOS link</td>
</tr>
<tr>
<td>$x_l$</td>
<td>group of $x_i$ measurements with the same distribution $\omega_m$</td>
</tr>
<tr>
<td>$\lambda_l$</td>
<td>classifier for group $x_l$</td>
</tr>
<tr>
<td>$y_i$</td>
<td>classifier for $x_i$</td>
</tr>
</tbody>
</table>

II. RELATED WORK

PD measurements for range estimation can be obtained (i) from the symbols of a received data packet or (ii) from multiple impulse-type signals transmitted in a short period of time. The former is a standard in many ultra short baseline systems (e.g., [8]) and involves inspecting the output of an energy detector [9]. The latter involves inspecting the estimated channel impulse response by performing a matched filter operation at the receiver [7], or by performing a phase-only correlation and using the kurtosis metric to mitigate channel enhanced noise [10]. The PD is then estimated by setting a detection threshold to identify the arrival of the first path. In [11], a fixed threshold is set based on the channel noise level and a target false alarm probability.
In [12], an adaptive threshold is used based on the energy level of the strongest path. A good overview of practical PD estimators is given in [9].

Mistaking NLOS links for the LOS link gives rise to ranging errors which are usually regarded as part of the measurement noise [13]. In [14], direct sequence spread spectrum (DSSS) signals, which have narrow auto-correlation, are transmitted to allow better separation of paths in the estimated channel response. Following this approach, curve fitting of ToA measurements based on DSSS was suggested in [15]. Averaging ToA measurements from different signals is suggested in [16], where results show considerable reduction in measurement errors. In [3], NLOS-related noise for UWAC is modeled using the Ultra Wideband Saleh-Valenzuela (UWB-SV) model [17], and a method for mitigating multipath noise for a given multipath model was introduced in [18].

Several works suggested methods to compensate for location ambiguities such as flips and rotations that arise due to NLOS-related range estimation errors. In [19], additional anchor nodes were used to resolve such ambiguities. In [20], a three-phase protocol is suggested for this problem. First, an ambiguity-free sub-tree of nodes is determined. Then, localization based on triangulation is performed where the node is first assumed to be located in the center of a rectangular area. Finally, a refinement phase is performed using a Kalman filter to mitigate noise arising from ranging. A robust protocol for mitigating localization ambiguities is suggested in [21] by rejecting measurements leading to ambiguities, e.g., when there are insufficient anchor nodes or when the location of anchor nodes is almost collinear. The problem of localization when all measurements are obtained from NLOS links is considered in [22], where the relationship between anchor node distances and NLOS factor is used to improve localization. However, these protocols are only applicable when a large number of anchor nodes are available.

Associating PD measurements with LOS or NLOS can improve localization accuracy. In [23], measurements which increase the global variance are rejected, assuming that NLOS-based measurements have larger variance than LOS-based measurements. In [24], localization accuracy is improved by selecting ToA measurements based on minimal statistical mode (i.e., minimal variance and mean). Alternatively, the authors in [25] suggested a method for reducing the effect of NLOS-based noise by assigning each measurement with a weight inversely proportional to the difference between the measured and expected distances from previous localization. In [26], an NLOS factor (i.e., the difference between the arrival times of the NLOS and LOS-based signals) is estimated using a maximum likelihood estimator based on an attenuation model, and
NLOS-based measurements are incorporated after a factor correction instead of being rejected. However, to the best of our knowledge, no prior work considered NLOS and LOS classification of PD measurements for the special characteristics of the UWAC.

III. SYSTEM SETUP AND ASSUMPTIONS

Referring to Figure 1, our system comprises of one or more transmitter-receiver pairs, \((u, a_j)\), exchanging a single communication packet of \(N\) symbols or impulse signals, from which a vector \(X = [x_1, \ldots, x_N]\) of PD measurements \(x_i\), is obtained using detectors such as in, e.g., [8], [7], [10]. We model \(x_i\) such that

\[
x_i = x_{\text{LOS}} + n_i ,
\]

where \(x_{\text{LOS}}\) is the transmitter-receiver PD in the LOS link, assumed to be fixed during the time \(X\) is obtained\(^1\), and \(n_i\) is zero-mean (for LOS links) or non-zero-mean (for SNLOS or ONLOS links) measurement noise. We assume signals are separated by guard intervals such that \(n_i\) are i.i.d. Each measurement \(x_i\) corresponds to a measured time \(t_i\), and a PD-based estimate, \(d_i^{\text{PD}}\), is obtained by multiplying \(x_i\) with an assumed propagation speed, \(c\). In addition, based on an attenuation model for an LOS link, we obtain RSS-based range estimates, \(d_i^{\text{RSS}}, i = 1, \ldots, N\), from the received signals.

In the following, we introduce our system model for obtaining RSS-based range measurements as well as the assumed probability density function (PDF) for PD measurements.

A. RSS-Based Range Measurements

Let \(d_{\text{LOS}}\) denote the distance corresponding to \(x_{\text{LOS}}\), i.e., \(d_{\text{LOS}} = x_{\text{LOS}}c\). For the purpose of obtaining RSS-based range measurements, we use the popular model [27]

\[
\text{TL}_{\text{LOS}}(d_{\text{LOS}}) = \text{PL}(d_{\text{LOS}}) + \text{AL}(d_{\text{LOS}}) + \epsilon ,
\]

where \(\text{PL}(d_{\text{LOS}}) = \gamma \log_{10}(d_{\text{LOS}})\) is the propagation loss, \(\text{AL}(d_{\text{LOS}}) = \alpha \frac{d_{\text{LOS}}}{1000}\) is the absorption loss, \(\gamma\) and \(\alpha\) are the propagation and absorption coefficients, respectively, and \(\epsilon\) is the model noise assumed to be Gaussian distributed with zero mean and variance \(\phi\). Considering the simplicity of the model in (2), we do not directly estimate \(d_i^{\text{RSS}}\) but rather estimate a lower bound \(d_i^{\text{RSS,min}}\),

\(^1\)A relaxation of this assumption is presented further below
for which we apply upper bounds for $\gamma$ and $\alpha$ in (2) according to the expected underwater environment.

For an ONLOS link with distance $d_{ONLOS} = d_{ONLOS,1} + d_{ONLOS,2}$, where $d_{ONLOS,1}$ and $d_{ONLOS,2}$ are the distance from source to reflector and from reflector to receiver, respectively\(^2\), we assume that the power attenuation in logarithmic scale is given by [27]

$$TL_{ONLOS}(d_{ONLOS}) = TL_{LOS}(d_{ONLOS,1}) + TL_{LOS}(d_{ONLOS,2}) + RL,$$

where $RL$ is the reflection loss of the reflecting object. We further assume that $RL$, which depends on the material and structure of the object and the carrier frequency of the transmitted signals, is sufficiently large such that

$$TL_{ONLOS}(d_{ONLOS}) \gg TL_{LOS}(d_{ONLOS}).$$

\(B. \) PDF for PD Measurements

We model the PDF of the noisy measurement $x_i$ as a mixture of $M = 3$ distributions, corresponding to LOS, SNLOS, and ONLOS links, such that (assuming independent measurement noise samples in (1))

$$P(X|\theta) = \prod_{x_i \in X} \sum_{m=1}^{M} k_m P(x_i|\omega_m),$$

where $\theta = \{\omega_1, k_1, \ldots, \omega_M, k_M\}$, $\omega_m$ are the parameters of the $m$th distribution, and $k_m$ ($\sum_{m=1}^{M} k_m = 1$) is the a-priori probability of the $m$th distribution. Clearly, $P(X|\theta)$ depends on both the UWAC and the detector used to estimate $x_i$. While recent works used the Gaussian distribution for $P(x_i|\omega_m)$ (cf., [28] and [29]), we take a more general approach and model it according to the generalized Gaussian PDF [30], such that

$$P(x_i|\omega_m) = \frac{\beta_m}{2\sigma_m \Gamma\left(\frac{1}{\beta_m}\right)} e^{-\left|\frac{x_i - v_m}{\sigma_m}\right|^{\beta_m}}$$

with parameters $\omega_m = \{\beta_m, v_m, \sigma_m\}$. We associate the parameters $\omega_1$, $\omega_2$, and $\omega_3$ with distributions corresponding to the LOS, SNLOS, and ONLOS links, respectively. Thus, by (1), $v_1 = x_{LOS}$. The use of parameter $\beta_m$ in (6) gives our model a desired flexibility, with $\beta_m = 1$.

\(^2\)Referring to the ONLOS link between node pair $(u, a_2)$ in Figure 1, $d_{ONLOS,1} = d_{21}$ and $d_{ONLOS,2} = d_{22}$.
\( \beta_m = 2 \), and \( \beta_m \to \infty \) corresponding to Laplace, Gaussian, and uniform distribution, respectively, and we verify model (6) in sea trials further below.

Following [23] and [24], we assume that PD measurements of NLOS links increase the variance of the elements of \( X \). Thus, if \( \varsigma_1, \varsigma_2, \) and \( \varsigma_3 \) are the respective variances of measurements related to the LOS, SNLOS, and ONLOS links, then we have

\[
\varsigma_1 < \varsigma_2, \varsigma_3.
\]

Since, for the PDF (6),

\[
\varsigma_m = (\sigma_m)^2 \frac{\Gamma\left(\frac{3}{\beta_m}\right)}{\Gamma\left(\frac{1}{\beta_m}\right)},
\]

and by (8), \( \varsigma_m \) does not change much with \( \beta_m \), constraint (7) can be modified to

\[
\sigma_1 < \sigma_2, \sigma_3.
\]

Furthermore, let \( T_{\text{LIR}} \) be the assumed length of the UWAC impulse response, which is an upper bound on the time difference between the arrivals of the last and first paths. Then, since \( \varsigma_1, \varsigma_2, \) and \( \varsigma_3 \) in (8) capture the spread of measurements related to the LOS, SNLOS, and ONLOS links respectively,

\[
\sqrt{\varsigma_m} < T_{\text{LIR}}, \ m = 1, 2, 3.
\]

Moreover, the propagation delay through the LOS link is almost always shorter than any NLOS link\(^3\). Hence, we have

\[
v_1 < v_2, v_3 < v_1 + T_{\text{LIR}},
\]

Parameters \( v_1 \) and \( v_3 \) are determined by the location of nodes and obstacles in the channel, parameters \( \sigma_1 \) and \( \sigma_3 \) are determined by the variance of the channel noise, while the other variables in \( \theta \) are random. Accounting for constraints (7)-(11) we assume \( P(v_2|v_1) \) is uniform between \( v_1 \) and \( v_1 + T_{\text{LIR}} \), \( P(\sigma_2|\sigma_1, \beta_2) \) is uniform between \( \sigma_1 \) and \( T_{\text{LIR}} \frac{\Gamma\left(\frac{1}{\beta_2}\right)}{\Gamma\left(\frac{3}{\beta_2}\right)} \), \( P(\sigma_1|\beta_1) \) is uniform between 0 and \( T_{\text{LIR}} \frac{\Gamma\left(\frac{1}{\beta_1}\right)}{\Gamma\left(\frac{3}{\beta_1}\right)} \), and \( P(\beta_m), \ m = 1, 2, 3, \) is uniform between 1 and a deterministic parameter, \( G \).

\(^3\)We note that in some UWACs, a signal can propagate through a soft ocean bottom, in which case SNLOS signals may arrive before the LOS signal [27]. However, such scenarios are not considered in this work.
C. Remark on Algorithm Structure

We offer a two-step approach to classify PD measurements into LOS, SNLOS, and ONLOS. First, assuming large attenuation in an ONLOS link, we compare PD-based and RSS-based range estimates to differentiate between ONLOS and non-ONLOS links. Then, assuming PDF (6) for PD measurements, we further classify non-ONLOS links into LOS and SNLOS links. We thus exploit in the first step that \( \hat{d}_{i}^{\text{RSS}} \) is significantly different for ONLOS compared to LOS and SNLOS links. This in turn simplifies the second step, as we do not need to jointly classify LOS, SNLOS, and ONLOS in a single step. In the following sections, we describe our two-step approach for classifying PD measurements starting from identifying ONLOS links, and followed by classifying non-ONLOS related elements of \( X \) into LOS and SNLOS through estimating the statistical parameters of both classes.

IV. Step One: Identifying ONLOS Links

Considering (4), we identify whether measurement \( x_{i} \in X \) is ONLOS-related based on three basic steps as follows:

- **Estimation of \( d_{i}^{\text{PD}} \)**

  We first obtain the PD-based range estimation as \( d_{i}^{\text{PD}} = c \cdot x_{i} \).

- **Estimation of \( d_{i}^{\text{RSS, min}} \)**

  Next, assuming knowledge of the transmitted power level, we measure the RSS for the \( i^{th} \) received signal/symbol, and estimate \( d_{i}^{\text{RSS, min}} \) based on (2), replacing \( \gamma \) and \( \alpha \) with upper bounds \( \gamma_{\text{max}} \) and \( \alpha_{\text{max}} \), respectively.

- **Thresholding**

  Finally, we compare \( d_{i}^{\text{PD}} \) with \( d_{i}^{\text{RSS, min}} \). If \( d_{i}^{\text{RSS, min}} > d_{i}^{\text{PD}} \), then \( x_{i} \) is classified as ONLOS. Otherwise, it is determined as non-ONLOS.

Next, we analyze the expected performance of the above ONLOS link identification algorithm in terms of (i) detection probability of non-ONLOS links, \( \Pr_{d, \text{non-ONLOS}} \), and (ii) detection probability of ONLOS links, \( \Pr_{d, \text{ONLOS}} \). To this end, since explicit expression for \( d_{\text{LOS}} \) cannot be obtained from (2), in the following, we use upper bound \( \hat{d}_{\text{RSS, min}} \) such that

\[
\log_{10}(\hat{d}_{\text{RSS, min}}) = \frac{\text{TL}}{\gamma_{\text{max}}}. \tag{12}
\]
We note that (12) is a tight bound when the carrier frequency is low or when the transmission distance is small.

A. Classification of non-ONLOS links

For non-ONLOS links, we expect $d_{i,\text{RSS},\min} \leq d_{\text{LOS}}$. Thus, since by bound (12), $\Pr(d_{i,\text{RSS},\min} \leq d_{\text{LOS}}) \geq \Pr(d_{i,\text{RSS},\min} \leq d_{\text{LOS}})$, and substituting (2) in (12), we get

$$\Pr_{d,\text{non-ONLOS}} \geq 1 - Q\left(\frac{(\gamma_{\text{max}} - \gamma) \log_{10}(d_{\text{LOS}}) - \alpha \frac{d_{\text{LOS}}}{1000}}{\phi}\right),$$

(13)

where $Q(x)$ is the Gaussian Q-function.

B. Classification of ONLOS links

When the link is ONLOS, we expect $d_{i,\text{RSS},\min} \geq d_{\text{ONLOS}}$. Then, substituting (3) in (12), and since $\Pr(d_{i,\text{RSS},\min} \geq d_{\text{ONLOS}}) \leq \Pr(d_{i,\text{RSS},\min} \geq d_{\text{ONLOS}})$, we get

$$\Pr_{d,\text{ONLOS}} \leq Q\left(\frac{\gamma_{\text{max}} \log_{10}(d_{\text{ONLOS}}) - \gamma \log_{10}(d_{\text{ONLOS},1}d_{\text{ONLOS},2}) - \alpha \frac{d_{\text{ONLOS}}}{1000} - RL}{\phi}\right).$$

(14)

Next, we continue with classifying non-ONLOS links into LOS and SNLOS links.

V. STEP 2: CLASSIFYING LOS AND SNLOS LINKS

After excluding ONLOS-related PD measurements in Step 1, the remaining elements, organized in vector $X^{\text{ex}}$, are further classified into LOS ($m = 1$) and SNLOS ($m = 2$) links and their statistical distribution parameters, $\omega_m$, are estimated.

Recall that estimations $x_i$ correspond to measurement times $t_i$. Assuming that the channel impulse response is constant within a coherence time, $T_c$, and that for transmitted signal bandwidth, $B$, system resolution (i.e., quantization noise) is limited by $\Delta T = \frac{1}{B}$, we can set equivalence constraints (denoted by operation $\Leftrightarrow$) such that closely spaced measurements are classified into the same class. PD measurements satisfying equivalence constraints are collected into vectors $x_l$, whose elements are classified into the same class. Therefore, $x_l$ have distinct elements. To formalize this, we determine $x_i \Leftrightarrow x_j$ if

$$|t_i - t_j| \leq T_c$$

(15a)

$$|x_i - x_j| \leq \Delta T.$$
Furthermore, while we assume in (1) that $x_{\text{LOS}}$ is constant for the time period during which vector $X$ is obtained, nodes may actually slightly move during that time\(^4\), and such motion can affect the distribution of PD measurements of the same class. To illustrate this, let $x_i$, $x_j$, and $x_n$ correspond to the same class (either LOS or NLOS), such that $x_i \leftrightarrow x_j$ and $x_j \leftrightarrow x_n$. Due to node motions, condition (15b) might not be hold for the pair $x_i$ and $x_n$. Accounting for such motions, we construct vectors $x_l$ such that if $x_i \leftrightarrow x_j$ and $x_j \leftrightarrow x_n$, it follows that $x_i$ and $x_n$ should also be classified to the same state. That is, vectors $x_p$ and $x_q$ are merged if they have a common element. To form vectors $x_l$, $l = 1, \ldots, L$, we begin with $|X^\text{ex}|$ ($|x|$ symbolizes the number of elements in vector $x$) initial vectors of single PD measurements, and iteratively merge vectors. This process continues until no two vectors can be merged. As a result, we reduce the problem of classifying $x_i \in X^\text{ex}$ into classifying $x_l$, which account for resolution limitations and node drifting.

While classification of measurement samples into two distinct distributions is a common problem solved by the expectation maximization (EM) algorithm (cf. [31]), here classification should also satisfy constraints (10), (9), and (11), where the latter two constraints introduce dependencies between $\omega_1$ and $\omega_2$. We start by formulating the log-likelihood function $L(\theta|\theta^p)$, where $\theta^p$ is the vector of distribution parameters, $\theta$, estimated in the $p^{\text{th}}$ iteration of the EM algorithm. Next, we formulate a constrained optimization problem to estimate parameters $k_m$, $\nu_m$, $\sigma_m$ and $\beta_m$ that maximize $L(\theta|\theta^p)$, and offer a heuristic approach to efficiently solve it. Finally, given an estimation for $\theta$, we calculate the probability of $x_l$ belonging to class $m$, and classify the elements of $X^\text{ex}$ accordingly.

A. Formalizing the Log-Likelihood Function

Let the random variable $\lambda_l$ be the classifier of $x_l$, such that if $x_l$ is associated with class $m$, $\lambda_l = m$. Also let $\lambda = [\lambda_1, \ldots, \lambda_L]$. Since elements in $X^\text{ex}$ are assumed independent,

$$
\Pr(\lambda_l = m|x_l, \theta^p) = \frac{k_m \prod_{x_i \in x_l} \Pr(x_i|\omega_m^p)}{\sum_{j=1}^{2} k_j \prod_{x_i \in x_l} \Pr(x_i|\omega_j^p)}. \tag{16}
$$

\(^4\)For example, an anchored node often moves around the location of its anchor.
Then, we can write the expectation of the log-likelihood function with respect to the conditional distribution of $\lambda$ given $X^{\text{ex}}$ and the current estimate $\theta^p$ as

$$L(\theta|\theta^p) = E[\ln (\Pr(X^{\text{ex}},\lambda|\theta)) | X^{\text{ex}}, \theta^p] = 
\sum_{m=1}^{2} \sum_{l=1}^{L} \Pr(\lambda_l = m|x_l, \theta^p) \ln P(x_i|\omega_m) + \sum_{l=1}^{L} \Pr(\lambda_l = m|x_l, \theta^p) \ln k_m,$$

where $\ln x = \log_e x$ is the natural logarithmic function.

Assuming knowledge of $\theta^p$, $\theta^{p+1}$ is estimated as the vector of distribution parameters that maximizes (17) while satisfying constraints (9), (10) and (11). This procedure is repeated for $P_{\text{last}}$ iterations, and the convergence of (17) to a local maxima is proven [31]. Then, we calculate $\Pr(\lambda_l = m|x_l, \theta^{P_{\text{last}}})$ using (16), and associate vector $x_l$ with the LOS path if

$$\Pr(\lambda_l = 1|x_l, \theta^{P_{\text{last}}}) > \Pr(\lambda_l = 2|x_l, \theta^{P_{\text{last}}}) ,$$

or with SNLOS otherwise. Estimation $\theta^{P_{\text{last}}}$ and classifications $\lambda_l$ can then be used to further increase the accuracy of UWAL, e.g., [23], [26].

We observe that the two terms on the right-hand side of (17) can be separately maximized, i.e., given $\theta^p$, we can obtain $\omega_m^{p+1}$ from maximizing the first term, and $k_m^{p+1}$ from maximizing the second term. Thus (see details in [31]),

$$k_m^{p+1} = \frac{1}{L} \sum_{l=1}^{L} \Pr(\lambda_l = m|x_l, \theta^p) .$$

In the following, we describe the details of our classification procedure for the estimation of $\omega_m$, followed by a heuristic approach for the initial estimates $\theta^0$.

B. Estimating the Distribution Parameters $w_1$ and $w_2$

To estimate $\omega_m$, we consider only the first term on the right-hand side of (17), which for the PDF (6) is given by

$$f(u_m, \sigma_m, \beta_m) = \sum_{l=1}^{L} \sum_{x_i \in x_l} \Pr(\lambda_l = m|x_l, \theta^p) \left[ \ln \beta_m - \ln(2\sigma_m) - \ln \Gamma\left(\frac{1}{\beta_m}\right) - \left(\frac{|x_i - u_m|}{\sigma_m}\right)^{\beta_m} \right] .$$

(19)
Then, considering constraints (9), (10) and (11), we find \( \omega_{m}^{p+1} \) by solving the following optimization problem:

\[
\omega_{1}^{p+1}, \omega_{2}^{p+1} = \arg\min_{\omega_{1}, \omega_{2}} - \sum_{m=1}^{2} f(v_{m}, \sigma_{m}, \beta_{m})
\]  

(20a)

s.t. : \( v_{1} \leq v_{2} \leq v_{1} + T_{LIR} \)  

(20b)

\[
\sigma_{m} \sqrt{\frac{\Gamma \left( \frac{3}{\beta_{m}} \right)}{\Gamma \left( \frac{1}{\beta_{m}} \right)}} - T_{LIR} \leq 0, \ m = 1, 2
\]  

(20c)

\[
\sigma_{1} - \sigma_{2} \leq 0
\]  

(20d)

We observe that convexity of \( f(v_{m}, \sigma_{m}, \beta_{m}) \) depends on \( \beta_{m} \). In Appendix A, we present an alternating optimization approach (cf. [32]) to efficiently solve (20).

Next, we present an algorithm to obtain the initial estimation, \( \theta^{0} \), whose accuracy affects the above refinement as well as the convergence rate of the EM algorithm.

C. Forming Initial Estimation \( \theta^{0} \)

Our algorithm to estimate \( \theta^{0} \) is based on identifying a single group, \( x_{l^{*}} \), whose elements belong to the LOS class with high probability, i.e., \( \Pr (\lambda_{l^{*}} = 1) \approx 1 \). This group is then used as a starting point for the K-means clustering algorithm [31], resulting in an initial classification \( \lambda_{l} \) for \( x_{l}, \ l = 1, \ldots, L \), to form two classified sets \( X_{m}^{ex} \), \( m = 1, 2 \). Finally, we evaluate the mean, variance, and kurtosis of the elements in vector \( X_{m}^{ex} \), denoted as \( E[X_{m}^{ex}], \text{Var}[X_{m}^{ex}], \text{and Kurtosis}[X_{m}^{ex}] \), respectively, to estimate \( \theta^{0} \) using the following properties for distribution (6):

\[
\begin{align*}
\frac{|X_{m}^{ex}|}{|X^{ex}|} &= k_{m}, \\
E[X_{m}^{ex}] &= v_{m}, \\
\text{Var}[X_{m}^{ex}] &= \sigma_{m}^{2} \frac{\Gamma \left( \frac{3}{\beta_{m}} \right)}{\Gamma \left( \frac{1}{\beta_{m}} \right)}, \\
\text{Kurtosis}[X_{m}^{ex}] &= \frac{\Gamma \left( \frac{5}{\beta_{m}} \right) \Gamma \left( \frac{1}{\beta_{m}} \right)}{\Gamma \left( \frac{3}{\beta_{m}} \right)}^{2} - 3.
\end{align*}
\]  

(21a)

(21b)

(21c)

(21d)

Since we assume that \( \sigma_{1} < \sigma_{2} \) (see (9)), we expect small differences between measurements of the LOS link, compared to those of SNLOS links. We use this attribute to identify group \( x_{l^{*}} \).
by filtering $X^{\text{ex}}$ and calculating the first derivative of the sorted filtered elements. Group $x_l$ corresponds to the smallest filtered derivative.

D. Discussion

We note that the constraints in (20) do not set bounds on the values of $\omega_1$ and $\omega_2$, but rather determine the dependencies between them. This is because, apart from the value of $T_{\text{LIR}}$ and distribution (6), we do not assume a-priori knowledge about the values of $k_m$ and $\omega_m$, $m = 1, 2$. In a scenario where the LOS path is always the strongest and PD measurements are all LOS-related, i.e., all elements of $X^{\text{ex}}$ belong to one class, our classifier might still estimate both $k_1$ and $k_2$ to be non-zero, resulting in wrong classification into two classes. In this case, a simple statistical evaluation of the mean of $X^{\text{ex}}$ might give a better estimation of $d_{\text{LOS}}$ than $\nu_1$.

To limit this shortcoming of our classifier, we assume that $\nu_1$ and $\nu_2$ are distinct if $X^{\text{ex}}$ is indeed a mixture of two distributions. To this end, in the last iteration, $P_{\text{last}}$, we classify $X^{\text{ex}}$ as a single class (of unknown type) if the difference $|\nu_{1}^{P_{\text{last}}} - \nu_{2}^{P_{\text{last}}}|$ is smaller than a threshold value, $\Delta \nu$ (determined by the system resolution for distinct paths). Then, if required, we find the distribution parameters of the (single) class by solving a relaxed version of (20), setting $k_1 = 1$ and $k_2 = 0$.

Nevertheless, we motivate relevance of our classifier in Section VI-B by showing that scenarios in which $X^{\text{ex}}$ is indeed a mixture of two distributions are not rare in real sea environments.

E. Summarizing the Operation of the Classifier

We now summarize the operation of our classification algorithm, whose pseudo-code is presented in Algorithm 1. First, we evaluate $d_i^{\text{PD}}$ and $d_i^{\text{RSS}, \text{min}}$ (lines 1-2). If $d_i^{\text{RSS}, \text{min}} > d_i^{\text{PD}}$, we classify $x_i$ as ONLOS; otherwise, we classify it as non-ONLOS (lines 3-6) and form the vector of non-ONLOS PD measurements, $X^{\text{ex}}$, and groups $x_l$, $l = 1, \ldots, L$ (line 7-8). Next, we form the initial solution, $\theta^0_m$ (line 9), and run the EM algorithm for $P_{\text{last}}$ iterations (lines 10-18). The procedure starts with estimating $k^0_m$ (line 11), followed by an iterative procedure to estimate $\omega^0_m$ for a pre-defined number of repetitions $N_{\text{repeat}}$ (lines 12-17). After iteration $P_{\text{last}}$, we check if vector $X^{\text{ex}}$ consists of two classes (line 20), and determine classifiers $\lambda_l$, $l = 1, \ldots, L$ (lines 21-22); otherwise $X^{\text{ex}}$ is classified as a single class (of unknown type), and, if estimating
\( \omega_m \) is required, we repeat the above procedure while setting \( k_1 = 1, k_2 = 0 \) (lines 23-24). The implementation code of the above algorithm can be downloaded from [33].

The EM algorithm, as well as the alternating optimization process in Appendix A, are proven to converge to a local maximum of the log-likelihood function (17). In the following, we provide the posterior hybrid Cramér-Rao bound (PHCRB) for classifying \( \omega_m \) as a benchmark for our classifier.

**F. Posterior Hybrid Cramér-Rao Bound (PHCRB)**

Consider the vector of measurements \( X^\text{ex} \) whose elements are drawn from distributions (6). Our classifier estimates the vector \( \theta = \left[ \upsilon_1, \sigma_1, \beta_1, k_1, \upsilon_2, \sigma_2, \beta_2, k_2 \right] = \left[ \theta_1, \ldots, \theta_8 \right] \), with \( \theta_1 \) being a deterministic parameter (determined by the LOS distance), and \( \theta_r = \left[ \theta_2, \ldots, \theta_8 \right] \) being a vector of random variables. To lower bound the estimation of \( \theta_1 \), we assume classification is given and for the combination of deterministic and random variables we use the PHCRB, such that

\[
E_{X^\text{ex}, \theta_r | \theta_1} \left[ \left( \hat{\theta} - \theta \right) \left( \hat{\theta} - \theta \right)^T \right] \geq H^{-1}(\theta_1, \lambda),
\]

where \( H(\theta_1, \lambda) \in \mathbb{R}^{8 \times 8} \) is the hybrid Fischer information matrix\(^5\) (HFIM). We observe that constraints (11), (9), and (10), introduce dependencies between pairs \((\upsilon_2, \upsilon_1)\), \((\sigma_1, \sigma_2)\), and \((\sigma_m, \beta_m)\), \( m = 1, 2 \), respectively. In addition, since \( k_2 + k_1 = 1 \), \( k_1 \) and \( k_2 \) are dependent. Thus, for \( y_n \) being the classifier of \( x_i \) (i.e., \( y_i = \lambda_l \) if \( x_i \in x_l \)), following [34] the \((j,q)\)th element of the HFIM is

\[
H(\theta_1, \lambda)_{j,q} = E_{\theta_r | \theta_1} [F(\theta_r, \theta_1, \lambda)] + E_{\theta_r | \theta_1} \left[ -\frac{\partial^2}{\partial \theta_j \partial \theta_q} \log P(\theta_r | \theta_1) \right],
\]

where

\[
F(\theta_r, \theta_1, \lambda)_{j,q} = E_{X^\text{ex} | \theta_r, \theta_1} \left[ -\sum_{i=1}^{N} \frac{\partial^2}{\partial \theta_j \partial \theta_q} \log k_{y_i} P(x_i | \omega_{y_i}) \right].
\]

To calculate \( P(\theta_r | \theta_1) = P(k_2 | k_1)P(k_1)P(\upsilon_2 | \upsilon_1)P(\sigma_2 | \sigma_1, \beta_2)P(\sigma_1 | \beta_1)P(\beta_1)P(\beta_2) \), we use the uniform distributions presented in Section III-B, and assume \( P(k_1) \) is uniform between 0 and 1 and \( P(k_2 | k_1) \) is the dirac delta function located at \( 1 - k_1 \). Exact expressions for (23) are given in Appendix B. In the following, we evaluate the PHCRB through Monte-Carlo simulations.

\(^5\)Note that while the EM algorithm works on vectors \( x_l \), the actual inputs to our classifier are PD measurements. Thus, in forming the PHCRB, we use \( x_i \) rather than \( x_l \).
VI. PERFORMANCE EVALUATION

In this section, we present results from both computer simulations and sea trials to demonstrate the performance of our classification algorithm. The results are presented in terms of detection probabilities of LOS, SNLOS, and ONLOS links. In addition, we measure estimation errors $|v^p_m - v_m|$, $|\sigma^p_m - \sigma_m|$, and $|\beta^p_m - \beta_m|$. We compare our results to the PHCRB presented in Section V-F, as well as to several benchmark methods. The purpose of the simulations is to evaluate the performance of our classifier in a controlled environment, while results of sea trials reflect performance in actual environment.

A. Simulations

Our simulation setting includes a Monte-Carlo set of 10000 channel realizations, where two time-synchronized nodes, uniformly placed into a square area of 1 km, exchange packets. The setting includes two horizontal and two vertical obstacles of length 20 m, also uniformly placed into the square area, such that a LOS always exists between the two nodes. In each simulation, we consider a packet of 200 symbols of duration $T_s = 10$ msec and bandwidth $B = 6$ kHz transmitted at a propagation speed of $c = 1500$ m/sec. To model movement in the channel (dealt with by forming groups $x_i$), during packet reception the two nodes move away from each other at constant relative speed of 1 m/sec, and $d_{\text{LOS}}$ is considered as the LOS distance between the nodes when the hundredth symbol arrives.

In our simulations, we use model (1) to obtain set $X$ as follows. For each channel realization and nodes time-varying positions, we find the LOS distance between the two nodes, and determine $v_1 = x_{\text{LOS}}$. Based on the position of nodes and obstacles, we identify ONLOS links as single reflections from obstacles and determine $v_3$ as the average delay of the found ONLOS links. We use $T_{\text{LIR}} = 0.1$ sec and base on constraint (11), we randomize $v_2$ according to a uniform distribution between $v_1$ and $v_1 + T_{\text{LIR}}$. For the other distribution parameters $\theta$, we determine $\beta_m$, $m = 1, 2, 3$ as an integer between 1 and 6 (i.e., for the assumed uniform distribution of $\beta_m$, $G = 6$) with equal probability, and $\sigma_m$, $m = 1, 2, 3$ according to (8) with $\varsigma_m$ uniformly distributed between 0 and $(T_{\text{LIR}})^2$, preserving $\varsigma_1 < \varsigma_2$. Based on model (5), we then randomize $x_i$, $i = 1, \ldots, 200$ using distribution (6) and a uniformly distributed $k_m$, $m = 1, 2, 3$ between 0 and 1, while keeping $\sum_{m=1}^3 k_m = 1$ and setting $k_3 = 0$ if no ONLOS link is identified.
Considering the discussion in Section V-D, we use $\Delta v = \frac{1}{c}$ as a detection threshold to check if measurements in vector $X^{ex}$ correspond to a single link. Additionally, for forming groups $x_i$ (see (15)), we use an assumed coherence time $\tilde{T}_c = a \cdot T_s$, where $a = \{1, 5, 10\}$, and an quantization threshold $\Delta T = 0.16$ msec based on the bandwidth of the transmitted signals. Note that since distance between nodes changed by 2 m during reception of the 200 symbols, when $a > \frac{2}{\Delta T_c} \approx 8.3$ we include all similar measurements in one vector $x_i$, whereas $a = 1$ corresponds to single element vectors $x_i$.

To simulate channel attenuation (2), we use $\gamma = 15$, $\alpha = 1.5$ dB/km (considering a carrier frequency of 15 kHz [27]), and set $\epsilon$ to be zero-mean Gaussian with variance $5/dB^2/\mu Pa@1m$. We use a source power level of 100 dB/$\mu Pa@1m$ and a zero-mean Gaussian ambient noise with power 20 dB/$\mu Pa@1m$, such that the signal-to-noise ratio (SNR) at the output of the channel is high. Attenuation in LOS and SNLOS links is determined based on (2), while for ONLOS links we use (3) and set RL = 10 dB/$\mu Pa@1m$. To obtain the lower bound on RSS-based distance, $d_{i_{\text{RSS, min}}, i = 1, \ldots, 200}$, we use the attenuation model in (2) with $\gamma_{\text{max}} = 20$ and $\alpha_{\text{max}} = 2$ dB/km. An implementation of the simulation environment can be downloaded from [33].

First, in Figure 2 we show empirical detection probabilities for ONLOS and non-ONLOS links as a function of $\gamma_{\text{max}}$, as well as corresponding results using bounds (14) and (13). We observe a good match between the empirical results and the analytical bound for $P_{d^{\text{ONLOS}}}$, and that $P_{d^{\text{ONLOS}}}$ is hardly affected by $\gamma_{\text{max}}$. However, $P_{d^{\text{non-ONLOS}}}$ increases dramatically with $\gamma_{\text{max}}$, and the bound in (13) is not tight. This is because while for $P_{d^{\text{non-ONLOS}}}$ choosing $\gamma_{\text{max}} < \gamma$ might lead to $d_{i_{\text{RSS, min}}} > d_{\text{LOS}}$ and neglectance of $\alpha$ in (13) renders analytical inaccuracies, for $P_{d^{\text{ONLOS}}}$ the large RL compensates on analytical inaccuracies and is more significant than the affect of $\gamma_{\text{max}}$.

In Figure 3, we show the empirical complimentary cumulative distribution function (C-CDF) of $\rho_{\text{err}} = |c\hat{x} - d_{\text{LOS}}|$, where $\hat{x}$ is i) $\hat{\upsilon}_1$, estimated using our classifier, ii) the mean of the elements in $X$ (Mean), iii) the minimum of $X$ (Min), or iv) taken as the average value of $X$ after removal of outliers, as suggested in [23] (Outlier). Results for $\hat{x} = \hat{\upsilon}_1$ are shown for $\tilde{T}_c = \{T_s, 5T_s, 10T_s\}$. The results in Figure 3 are also compared with the square of the PHCRB presented in Section V-F (PHCRB). We observe that the Outlier method outperforms the naive approaches of using the average or minimum value of $X$, where the latter performs extremely poorly for large values of
\( \rho_{\text{err}} \). However, using our classifier, i.e., \( \hat{x} = \hat{\psi}_{1} \), results improve significantly. For example, using our approach \( \rho_{\text{err}} \approx 7 \) m in 90\% of the cases, compared to 11.2 m when using the Outlier method, and results are close to the square of the PHCRB. Such an improvement immediately translates into better localization performance as PD estimation errors significantly decrease. Comparing results for different values of \( \tilde{T}_{c} \), we observe that using equivalence constraints (i.e., \( \tilde{T}_{c} > T_{s} \)), performance improves compared to the case of \( \tilde{T}_{c} = T_{s} \). However, a tradeoff is observed as results for \( \tilde{T}_{c} = 5T_{s} \) are slightly better than for \( \tilde{T}_{c} = 10T_{s} \). This is because over-estimated coherence time, \( \tilde{T}_{c} \), renders possibly wrong assignment to vectors \( x_{l} \).

Convergence of the EM iterative procedure is demonstrated in Figure 4, where we show average estimation errors of the distribution parameters of the LOS class as a function of the EM iteration step number. We observe that convergence is reached after 10 iteration steps. While improvement compared to the initialization process (see Section V-C) is shown for all estimations, the impact of the EM algorithm is mostly demonstrated by the improvement in estimation \( \hat{k}_{1} \), which greatly affect classification performance. This improvement is also observed in Figure 5, where we show empirical detection probabilities\(^6\) for LOS (LOS (EM)) and SNLOS (SNLOS (EM)) links, as well as the total detection probability (ALL (EM)), which is calculated as the rate of correct classification (of any link). Also shown are classification performance using only the initialization process (init), i.e., before the EM algorithm is employed. We observe that the constraint EM algorithm achieves a significant performance gain compared to the K-means, used in the initialization process. Furthermore, results show that for the former, the detection rate is more than 92\% in almost all cases.

Next, we present classification results based on real data collected from sea trials.

B. Sea Trials

While our simulations demonstrate good classification performance for our algorithm, it relies on the distribution model (6), and upper bound on transmission loss models (2) and (3), which might not be accurate in practice. Thus, to show the resilience of our classifier to different sea environments, we conducted a series of three sea trials in Israel and in Singapore. In the

\(^{6}\)We note that detection probabilities are calculated only when vector \( X \) consists of both LOS and SNLOS related PD measurements; classification cannot be made otherwise.
following, we describe the experimental setup and present the offline classification results of the recorded data.

To acquire PD measurements, we used a matched filter (MF), as well as the phase-only-correlator (POC) as described in [10]. The MF estimation method assumes an impulse-like autocorrelation of the transmitted symbols, which is not required for the POC method. However, the latter introduces some degree of noise enhancement [10]. For the $i$th received signal, $x_i$ is estimated as the first peak at the output of the POC or MF that passes a detection threshold\(^7\).

1) **Classifying ONLOS links:**

In this section, we show the performance of ONLOS link identification in an experiment conducted at the Haifa harbor, Israel, in May 2009. The experiment included four vessels, each representing an individual node in the network. In each vessel, a transceiver was deployed at a fixed depth of 3 m. The four nodes were time synchronized using GPS and transmitted with equal transmission power at a carrier frequency of $15 \, kHz$. Referring to Figure 6, node 2 was placed at a fixed location $2A$, while nodes 1, 3 and 4 sent packets to node 2 while moving between various locations, creating a controlled environment of five non-ONLOS and four ONLOS communication links with maximum transmission distance of 1500 m. For each link, $(2, j)$, $j \in \{1,3,4\}$, we evaluated (i) $d_{PD}$ as the product of an assumed propagation speed of $1550 \, m/sec$ and the position of the first peak of the POC for the synchronization signal of each received packet, and (ii) $d_{RSS,min}$, employing an energy detector over the synchronization signal and using (2) for $\alpha_{max} = 2 \, db/km$ and $\gamma_{max} = 20$. We note that results only changed slightly when alternative methods for obtaining $d_{RSS,min}$ and $d_{PD}$ were applied.

In Table II, we present values of $d_{RSS,min}$ and $d_{PD}$ for each of the 9 communication links. Applying our proposed ONLOS link identification method, all four ONLOS links were correctly classified and there was no false classification of non-ONLOS links. In particular, we observe that for all ONLOS links, $d_{PD}$ is much lower than $d_{RSS,min}$, validating our assumption that the reflection loss of the reflecting objects (which could have been harbor docks, ship hulls, etc.) are sufficiently high to satisfy assumption (4).

2) **Classifying non-ONLOS links:** Next, we present results from two separate experiments conducted in open sea: (i) the first along the shores of Haifa, Israel, in August 2010 and

---

\(^7\)The detection threshold is chosen according to the Neyman-Pearson criterion [35] for a false alarm probability of $10^{-4}$. 
(ii) the second in the Singapore straits in November 2011, with water depths of 40 m and 15 m respectively. This is done to demonstrate our classifier’s performance in different sea environments and for two different \( X \) configurations (one from symbols in packets, and the other from independently transmitted signals). As communication links were all non-ONLOS links in both experiments, \( X^{\text{ex}} = X \) and we only present results for LOS/SNLOS classification.

The first experiment included three vessels, representing three mobile nodes, which drifted with the ocean current at a maximum speed of 1 m/sec, and were time-synchronized using a method described in [36]. Throughout the experiment, the node locations were measured using GPS receivers, and the propagation speed was measured to be \( c_1 = 1550 \) m/sec with deviations of no more than 2 m/sec across the water column. Each node was equipped with a transceiver, deployed at 10 meters depth, and transmitted more than 100 data packets which were received by the other two nodes. Each packet consisted of 200 direct-sequence-spread-sequence (DSSS) symbols of duration \( T_s = 10 \) msec and a spreading sequence of 63 chips was used. From each packet, vector \( X \) was obtained by applying (i) the MF, and (ii) the POC for the \( i \)th DSSS symbol.

As discussed in Section V-D, we can classify \( x_i \in X^{\text{ex}} \) to LOS or SNLOS only if \( X^{\text{ex}} \) comprises both classes. Thus, to motivate the use of our classifier, we evaluate the likelihood

### TABLE II: Harbor trial results for ONLOS link classification.

<table>
<thead>
<tr>
<th>Link</th>
<th>( d_{\text{RSS}, \text{min}} ) [m]</th>
<th>( d_{\text{PD}} ) [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2A,3B)</td>
<td>579</td>
<td>780</td>
</tr>
<tr>
<td>(2A,4A)</td>
<td>179</td>
<td>242</td>
</tr>
<tr>
<td>(2A,4B)</td>
<td>343</td>
<td>415</td>
</tr>
<tr>
<td>(2A,1A)</td>
<td>428</td>
<td>610</td>
</tr>
<tr>
<td>(2A,3A)</td>
<td>647</td>
<td>817</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Link</th>
<th>( d_{\text{RSS}, \text{min}} ) [m]</th>
<th>( d_{\text{PD}} ) [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2A,1B)</td>
<td>1557</td>
<td>1155</td>
</tr>
<tr>
<td>(2A,3C)</td>
<td>1639</td>
<td>740</td>
</tr>
<tr>
<td>(2A,1D)</td>
<td>1549</td>
<td>1254</td>
</tr>
<tr>
<td>(2A,1C)</td>
<td>1816</td>
<td>950</td>
</tr>
</tbody>
</table>
for the occurrence of such a condition by measuring the difference $\rho_{\text{diff}} = c_1(v_2 - v_1)$, assumed to be limited by node motions if $X^{\text{ex}}$ consists only of one class. For a given parameter $T_w$ and a maximum node velocity $v$ ($v = 1$ m/sec in our simulations), by taking the first $\left\lfloor \frac{T_w}{T_s} \right\rfloor$ PD measurements from each packet, the case of $\rho_{\text{diff}} > 2v \cdot T_w$ indicates that $X^{\text{ex}}$ comprises two classes with high probability. In Figure 7 we show the empirical C-CDF of $\rho_{\text{diff}}$ for $T_w = 1$ sec and $T_w = 0.5$ sec, where $\tilde{T}_e = 10T_s$. We observe that the number of cases where $\rho_{\text{diff}} > 2T_w$ is greater when estimating the channel impulse response using POC than when using MF. This is because auto-correlation of the used DSSS symbols is not sufficiently narrow, which decreases the path separation in MF compared to POC. From Figure 7, we observe that as expected, $\rho_{\text{diff}}$ decreases with $T_w$. However, even for $T_w = 0.5$ sec, using both MF and POC, $\rho_{\text{diff}} > 1$ in more than 20% of the received packets, which motivates the use of our classifier in real sea environment.

An estimation parameter of interest is $\beta_m$, which determines the type of distribution of the $m^{th}$ class. We consider only packets received for which $\rho_{\text{diff}} > 2T_w$, i.e., vector $X$ consists of two classes, and show results only for the POC method (while noting that similar results are obtained using the MF method) as it produces more measurements of interest. In Figure 8, for $T_w = 1$ sec, we show the empirical histogram of $\beta_1$ and $\beta_2$ for different values of $\tilde{T}_e$. As the results are similar for both $\tilde{T}_e = T_s$ and $\tilde{T}_e = 10T_s$, this implies that small changes in values of PD measurements do not affect the estimated type of distribution. We also observe that the LOS class seems to lean towards $\beta_1 = 6$, which implies a uniform distribution, while SNLOS focuses on $\beta_2 = 2$, which is the normal distribution. However, as both estimated parameters take almost any value in the considered range, we conclude that no prior assumption on the type of distribution of PD measurements can be made in this experiment.

In Figure 9, using $\tilde{T}_e = 10T_s$, we show the empirical C-CDF of $\rho_i^e = |c_1\hat{x} - E(d_{\text{LOS}})|$, where $E(d_{\text{LOS}})$ is the mean of the GPS-based transmitter-receiver distance during the reception of each packet, $\hat{x}$ is either $v_1$, $E(X)$, or $\min(X)$, and elements $x_i \in X$ were estimated using the POC method. Assuming GPS location uncertainties of 5 m, we require $\rho_i^e$ to be below 6 m. Results show that $\rho_i^e$ for $\hat{x} = \min(X)$ is lower than for $\hat{x} = E(X)$. However, for $\hat{x} = v_1$, $\rho_i^e$ is always the lowest, and is smaller than 6 m in more than 90% of the cases (compared to 55% of the cases for $\hat{x} = \min(X)$).

The second sea trial included two underwater acoustic modems, manufactured by Evologics
GmbH, which were deployed at a depth of 5 m. One of them was deployed from a static platform and the other from a boat anchored to the sea bottom. Throughout the experiment, the boat changed its location, resulting in three different transmitter-receiver distances which were monitored using GPS measurements. Measurements \( x_i \in X \) were obtained at each node using four-way packet exchange in a TDoA fashion, resulting in a rate of one two-way PD measurement every 6 sec. For each transmission distance, the boat remained static for 20 min, allowing around 200 measurements \( x_i \) at each node. In this experiment, a propagation speed of \( c_2 = 1540 \text{ m/sec} \), as measured throughout the year in the Singapore straits [28], was considered.

In Figure 10, we show the empirical histogram of \( \rho^e_i = |c_2 x_i - d_i| \) for a single vector \( X \), where \( d_i \) is the GPS-based transmitter-receiver distance measured at time \( t_i \) (i.e., when \( x_i \) is measured), with mean and variance of \( E(d) = 324.1 \text{ m} \) and \( \text{Var}(d) = 3 \text{ m}^2 \), respectively. We also plotted \( \rho^e_i \) by replacing \( d_i \) with \( E(d) \) and \( x_i \) with \( E(X) \), and \( \min(X) \), as well as the theoretical PDFs shifted by \( E(d) \) for distribution estimated parameters \( \omega_1 \) and \( \omega_2 \). Vector \( X \) was classified into LOS and SNLOS classes with \( v_1 \) and \( v_2 \) corresponding to \( \rho^e_2 = 0.1 \text{ m} \) and 18.4 m, respectively, indicating the long channel impulse response. The estimated factors for each class were \( \beta_1 = 1 \) and \( \beta_2 = 6 \), where the former matches the narrow peak distribution observed for the LOS class, and the latter matches the near uniform distribution observed for the SNLOS class. We note the good fit between the shape of the theoretical PDF and the empirical histogram for both classes. In addition, we observe that \( \rho^e_2 \) for estimation \( v_1 \) gives much better results than the naive approach of taking the average or minimum value of \( X \).

In Figure 11 we plot the ratio \( \frac{\rho^e_2}{E(d)} \) for the three locations of the boat in the sea trial, averaged for the two nodes. We observe that the minimum value of \( X \) usually, but not always, results in better propagation delay estimation than the mean value, which in turn always results in better estimation than \( v_2 \), as expected. However, \( v_1 \) yields the best results, with average estimation error of 0.7 m compared to more than 10 m for the other methods.

Thus, based on the results obtained from both sea trials, we conclude that our classifier significantly improves PD-based range estimations in different sea environments compared to the often used, naive, approach.
VII. Conclusions

In this paper, we considered the problem of classifying propagation delay (PD) measurements in the underwater acoustic channel into three classes: line-of-sight (LOS), sea surface- or bottom-based reflections (SNLOS), and object-based reflections (ONLOS), which is important for reducing possible errors in PD-based range estimation for underwater acoustic localization (UWAL). We presented a two-step classifier which first compares PD-based and received signal strength based ranging to identify ONLOS links, and then, for non-ONLOS links, classifies PD measurements into LOS and SNLOS paths, using a constrained expectation maximization algorithm. We also offered a heuristic approach to efficiently maximize the log-likelihood function, and formalized the Cramér-Rao Bound to validate the performance of our method using numerical evaluation. As our classifier relies on the use of simplified models, alongside simulations, we presented results from three sea trials conducted in different sea environments. Both our simulation and sea trial results confirmed that our classifier can successfully distinguish between ONLOS and non-ONLOS links, and is able to accurately classify PD measurements into LOS and SNLOS paths. Further work will include using these classifications to improve the accuracy of UWAL.

Appendix A

Alternating Optimization Approach for Solving (20)

In alternating optimization, a multivariate maximization problem is iteratively solved through alternating restricted maximization over individual subsets of variables [32]. Using the alternating optimization approach, we iteratively alternate parameters $\omega_1$ and $\omega_2$ to solve (20).

The process is initialized by setting $\nu_m^{p+1,0} = \nu_m^p$, $\sigma_m^{p+1,0} = \sigma_m^p$, and $\beta_m^{p+1,0} = \beta_m^p$, and ends after $N_{\text{repeat}}$ iterations by setting $\nu_m^{p+1} = \nu_m^{p+1,N_{\text{repeat}}}$, $\sigma_m^{p+1} = \sigma_m^{p+1,N_{\text{repeat}}}$, and $\beta_m^{p+1} = \beta_m^{p+1,N_{\text{repeat}}}$, $m = 1, 2$. We first find a set of possible solutions for $\nu_m^{p+1,n+1}$ and $\sigma_m^{p+1,n+1}$, by deriving the first
derivative of (19) with respect to \( \nu_m \) and \( \sigma_m \), respectively, and setting it to zero, i.e.,
\[
\frac{\partial}{\partial \nu_m} f(v_m, \sigma_m^{p+1,n}, \beta_m^{p+1,n}) = \sum_{l=1}^{L} \text{Pr}(\lambda_l = m|x_l, \theta^p) \sum_{x_i \in x_l} \frac{|x_i - v_m^{p+1,n} - 1 - \beta_m^{p+1,n}}{(\sigma_m^{p+1,n})^{\beta_m^{p+1,n}}} \cdot S_gn(x_i - v_m^{p+1,n}) = 0
\]
(25a)
\[
\frac{\partial}{\partial \sigma_m} f(v_m^{p+1,n}, \sigma_m, \beta_m^{p+1,n}) = \sum_{l=1}^{L} \text{Pr}(\lambda_l = m|x_l, \theta^p) \sum_{x_i \in x_l} -1 + \frac{\beta_m^{p+1,n} |x_i - v_m^{p+1,n} + 1|^{\beta_m^{p+1,n}}}{(\sigma_m)^{\beta_m^{p+1,n}+1}} = 0 ,
\]
(25b)
where \( S_gn(x) \) is the algebraic sign of \( x \). Next, we set
\[
v_m^{p+1,n+1} = \arg\max_{v_m} f(v_m, \sigma_m^{p+1,n}, \beta_m^{p+1,n})
\]
(26a)
\[
\sigma_m^{p+1,n+1} = \arg\max_{\sigma_m} f(v_m^{p+1,n}, \sigma_m, \beta_m^{p+1,n})
\]
(26b)
\[
\beta_m^{p+1,n+1} = \arg\max_{\beta_m} f(v_m^{p+1,n}, \sigma_m^{p+1,n+1}, \beta_m)
\]
(26c)
where \( v_m \) and \( \sigma_m \) in (26a) and (26b), respectively, are limited to real solutions of (25) which satisfy the following modified constraints of (20)
\[
v_2^{p+1,n} - T_{LIR} \leq \nu_1 \leq v_2^{p+1,n}
\]
(27a)
\[
v_1^{p+1,n+1} \leq \nu_2 \leq v_1^{p+1,n+1} + T_{LIR}
\]
(27b)
\[
0 \leq \sigma_1 \leq \min \left( \sigma_2^{p+1,n}, T_{LIR} \sqrt{\frac{\Gamma(\frac{1}{\beta_1^{p+1,n}})}{\Gamma(\frac{3}{\beta_1^{p+1,n}})}} \right)
\]
(27c)
\[
0 \leq \sigma_2 \leq \min \left( \sigma_3^{p+1,n+1}, T_{LIR} \sqrt{\frac{\Gamma(\frac{1}{\beta_2^{p+1,n+1}})}{\Gamma(\frac{3}{\beta_2^{p+1,n+1}})}} \right),
\]
(27d)
and \( \beta_m^{p+1,n+1} \) in (26c) is obtained by numerically solving
\[
\frac{\partial}{\partial \beta_m} f(v_m^{p+1,n+1}, \sigma_m^{p+1,n+1}, \beta_m) = \sum_{l=1}^{L} \text{Pr}(\lambda_l = m|x_l, \theta^p) \sum_{x_i \in x_l} \frac{1}{\beta_m} + \frac{1}{\beta_m^2} \psi \left( \frac{1}{\beta_m} \right) - \frac{|x_i - v_m^{p+1,n+1}|}{(\sigma_m^{p+1,n+1})^{\beta_m}} \log \left( \frac{|x_i - v_m^{p+1,n+1}|}{\sigma_m^{p+1,n+1}} \right) = 0 ,
\]
(28)
where \( \psi(\cdot) \) is the digamma function, i.e., the first derivative of \( \log \Gamma(\cdot) \). We note that (25b) is analytically tractable for integer \( \beta_m^{p+1,n} \leq 5 \). Additionally, (25a) is also tractable if we replace
$S_{gn}(x_i - v_m)$ in (25a) with $S_{gn}(x_i - v^{p+1,n}_m)$. Finally, the complexity of numerically evaluating (26c) can be greatly reduced by starting the numerical search for $\beta_m$ from $\beta_m^p$.

In [32] it was proven that alternating maximization converge if for each alternation, problem constraints (in our case (27)) are handled internally. This convergence has been demonstrated by means of numerical simulations in Section VI-A.

APPENDIX B

EXPRESSIONS FOR THE PHCRB

We observe that $F(\theta_r, \theta_1, \lambda)_{j,q}$ in (24) is non-zero only when $j, q = 1, \ldots, 4$ and $m = 1$, or when $j, q = 5, \ldots, 8$ and $m = 2$, and the expressions are similar. Thus, denoting $b_{i,m} = x_i - v_m$, we give a closed form expression of $F(\theta_r, \theta_1, \lambda)_{j,q}$ for $j, q = R + 1, \ldots, R + 4$, where $R = 4 \cdot (m - 1)$, as follows.

$$F(\theta_r, \theta_1, \lambda)_{j,q} = \begin{cases} -\frac{\beta_m}{\sigma_m} |b_{i,m}|^{\beta_m-2} [\beta_m - 1 + 2 |b_{i,m}| \delta(b_{i,m})], & j = R + 1, q = R + 1; \\ -\beta_m^2 S_{gn}(b_{i,m}) |b_{i,m}|^{\beta_m-1} \sigma_m^{-1}, & j = R + 1, q = R + 2; \\ S_{gn}(b_{i,m}) |\sigma_m^{-1} b_{i,m}|^{\beta_m-1} + \beta_m \sigma_m^{-1} \log(\frac{1}{\sigma_m}) |b_{i,m}|^{\beta_m-1} + \\ k_m S_{gn}(b_{i,m}) \beta_m \sigma_m^{-1} b_{i,m} |b_{i,m}|^{\beta_m-1} \log |b_{i,m}|, & j = R + 1, q = R + 3; \\ -\beta_m^2 \sigma_m^{-1} b_{i,m} |b_{i,m}|^{\beta_m-1} S_{gn}(b_{i,m}), & j = R + 2, q = R + 1; \\ \frac{1}{\sigma_m^2} - (\beta_m + 1) \beta_m |b_{i,m}|^{\beta_m-1} \sigma_m^{-2}, & j = R + 2, q = R + 2; \\ |b_{i,m}|^{\beta_m} \sigma_m^{-1} - \beta_m \sigma_m^{-1} \log |b_{i,m}|, & j = R + 2, q = R + 3; \\ |b_{i,m}|^{\beta_m-1} S_{gn}(b_{i,m}) (\frac{1}{\sigma_m} \beta_m [\beta_m \log(\frac{b_{i,m}}{\sigma_m}) + \sigma_m], & j = R + 3, q = R + 1; \\ |b_{i,m}|^{\beta_m-1} \sigma_m^{-1} [\beta_m \log(\frac{b_{i,m}}{\sigma_m}) + |b_{i,m}|^{\beta_m-1}], & j = R + 3, q = R + 2; \\ -\frac{1}{\beta_m} - \frac{1}{\beta_m} \psi'(\frac{1}{\beta_m}) - (\frac{|b_{i,m}|}{\sigma_m})^{\beta_m} (\log(\frac{|b_{i,m}|}{\sigma_m}))^2 - \frac{2}{\beta_m^2} \psi'(\frac{1}{\beta_m}), & j = R + 3, q = R + 3; \\ -\frac{1}{k_m}, & j = R + 4, q = R + 4; \\ 0, & \text{otherwise,} \end{cases}$$

where $\delta(\cdot)$ is the Dirac delta function, and $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions, i.e., the first and second derivative of $\log \Gamma$, respectively.

For the second term of (23), we observe that $\frac{\partial^2}{\partial \theta_r \partial \theta_1} \log P(\theta_r | \theta_1)$ is non-zero for $j = 2, q = 2, 7, j = 3, q = 3$, and $j = 7, q = 2, 7$. Thus, defining $a_{1,m} = \Gamma\left(\frac{1}{\beta_m}\right)$, $a_{2,m} = \Gamma\left(\frac{3}{\beta_m}\right)$, and
\[ a_{3,m} = \frac{a_{1,2}}{a_{2,2}}. \]

\[
\frac{\partial^2}{\partial \theta_j \partial \theta_q} \log P(\theta_j | \theta_1) = \begin{cases} 
\frac{1}{(T_{LIR} \sqrt{a_{3,2} - \sigma_1})^2} \cdot \frac{-0.5(\beta_2 - 2) T_{LIR} a_{3,2}^0.5}{(T_{LIR} \sqrt{a_{3,2} - \sigma_1})^2} \left( 3a_{3,2} \frac{\Gamma'(\frac{3}{\beta_2})}{\beta_2} + \frac{\Gamma'(\frac{1}{\beta_1})}{\beta_1} \right), & j = 2, q = 2; \\
\frac{1}{2\beta_1} \left( -\frac{1}{\beta_1} \psi'(\frac{1}{\beta_1}) + 9 \psi''(\frac{1}{\beta_1}) - \frac{1}{\beta_1} \left( \psi'(\frac{1}{\beta_1}) - 3 \psi(\frac{3}{\beta_1}) \right) \right), & j = 3, q = 3; \\
\frac{-0.5(\beta_2 - 2) T_{LIR} a_{3,2}^0.5}{(T_{LIR} \sqrt{a_{3,2} - \sigma_1})^2} \left( 3a_{3,2} \frac{\Gamma'(\frac{3}{\beta_2})}{\beta_2} + \frac{\Gamma'(\frac{1}{\beta_1})}{\beta_1} \right), & j = 7, q = 2; \\
\frac{\partial}{\partial \beta_2} \left[ \frac{0.5 T_{LIR} (\beta_2 - 2)(a_{3,2})^{-0.5}}{T_{LIR} \sqrt{a_{3,2} - \sigma_1}} \left( 3a_{3,2} \frac{\Gamma'(\frac{3}{\beta_2})}{\beta_2} + \frac{\Gamma'(\frac{1}{\beta_1})}{\beta_1} \right) \right], & j = 7, q = 7; \\
0, & \text{otherwise},
\end{cases}
\]

where \( \Gamma'(\cdot) \) is the first derivative of \( \Gamma(\cdot) \).

REFERENCES


Unlocalized node, $u$

Anchor node

Sea surface

Sea bottom

ONLOSS

SNLOS

ONLOS

Ship

rocks

Fig. 1: Illustration of various types of communication links: LOS, SNLOS and ONLOS links.

Fig. 2: $Pr_{d,\text{non-ONLOS}}$ and $Pr_{d,\text{ONLOS}}$ vs. $\gamma_{\text{max}}$. $\gamma = 15$, $\alpha = 1.5$.


Fig. 3: Empirical C-CDF of $\rho_{\text{err}}$.

Fig. 4: Estimation error of LOS and SNLOS distribution parameters as a function of EM iteration number.
Algorithm 1 Classifying ONLOS link

1: \( d^{\text{PD}}_i =: c 	imes x_i \)

2: Calculate \( d^{\text{RSS, min}}_i \) using RSS measurements, \( \gamma_{\text{max}}, \alpha_{\text{max}} \) and model (2)

3: if \( d^{\text{RSS, min}}_i > d^{\text{PD}}_i \) then

4: Classify \( x_i \) as ONLOS link

else

5: Classify \( x_i \) as non-ONLOS link

6: Exclude ONLOS measurements to form vector \( X^\text{ex} \)

7: Form groups \( x_l \) satisfying (15)

8: Estimate \( \theta^0 \)

9: for \( p := 2 \) to \( P_{\text{last}} \) do

10: Calculate \( k^p_m, m = 1, 2 \) using (16) and (V-A)

11: for \( i := 1 \) to \( N_{\text{repeat}} \) do

12: Calculate possible solutions for \( \nu^p_m \) and \( \sigma^p_m, m = 1, 2 \) using (25)

13: Eliminate complex solutions and values outside bounds (27)

14: Estimate \( \omega^p_m, m = 1, 2 \) from (26)

15: \( m = 1, 2: \nu^{p-1}_m =: \nu^p_m, \sigma^{p-1}_m =: \sigma^p_m, \beta^{p-1}_m =: \beta^p_m \)

16: end for

17: end for

18: if \( |v^1_{\text{last}} - v^2_{\text{last}}| > \Delta v \) then

19: Calculate \( \Pr(\lambda_l = m|x_l, \theta^{\text{last}}), m = 1, 2 \) using (16)

20: Set \( \lambda_l, l = 1, \ldots, L \) according to (18)

21: else

22: Vector \( X^\text{ex} \) consists of a single class

23: Repeat steps 10-18 for \( k_1 = 1, k_2 = 0 \)

24: end if

25: end if
Fig. 5: Classification Detection Rate with and without constraint EM.

Fig. 6: Satellite picture of the sea trial location (picture taken from Google maps on September 29, 2009.).
Fig. 7: Empirical C-CDF of $\rho_{\text{diff}}$, using both POC and MF. $\bar{T}_c = 10T_s$. Israel experiment.

Fig. 8: Empirical histogram of estimation $\beta_m$, $m = 1, 2$, using POC. $T_w = 1$ sec. Israel experiment.
Fig. 9: Empirical C-CDF of $\rho_{e1}$, using POC. $\tilde{T}_c = 10T_s$, $T_w = 1$ sec. Israel experiment.

Fig. 10: Empirical histogram of $\rho_{e2}$ from Singapore experiment. Bin width 0.3 m, $E(d) = 324.1$ m.
Fig. 11: $\frac{\rho_s}{E(d)}$ averaged over results from the two nodes. Singapore experiment.