3-2015

A Two-hydrophone Range and Bearing Localization with Performance Analysis

John Thomas Gebbie
Portland State University, john.gebbie@gmail.com

Martin Siderius
Portland State University, siderius@pdx.edu

John S. Allen III
University of Hawaii at Manoa

Citation Details

Gebbie, John, Martin Siderius, and John S. Allen III. "A two-hydrophone range and bearing localization algorithm with performance analysis." The Journal of the Acoustical Society of America 137.3 (2015): 1586-1597.

A two-hydrophone range and bearing localization algorithm with performance analysis

John Gebbie and Martin Siderius[a)]
Northwest Electromagnetics and Acoustics Research Laboratory, Department of Electrical and Computer Engineering, Portland State University, 1900 SW 4th Avenue, Suite 160, Portland, Oregon 97201

John S. Allen III
Department of Mechanical Engineering, University of Hawai'i-Manoa, 2540 Dole Street, Holmes Hall 302, Honolulu, Hawaii 96822

(Received 18 April 2014; revised 7 October 2014; accepted 5 January 2015)

An automated, passive algorithm for detecting and localizing small boats using two hydrophones mounted on the seabed is outlined. This extends previous work by Gebbie *et al.* [(2013). J. Acoust. Soc. Am. **134**, EL77 – EL83] in which a similar two-hydrophone approach is used to produce an ambiguity surface of likely target locations leveraging multipath analysis and knowledge of the local bathymetry. The work presented here improves upon the prior approach using particle filtering to automate detection and localization processing. A detailed analysis has also been conducted to determine the conditions and limits under which the improved approach can be expected to yield accurate range and unambiguous bearing information. Experimental results in 12 m of water allow for a comparison of different separation distances between hydrophones, and the Bayesian Cramér-Rao lower bound is used to extrapolate the performance expected in 120 m water. This work demonstrates the conditions under which a low cost, passive, sparse array of hydrophones can provide a meaningful small boat detection and localization capability.
 [http://dx.doi.org/10.1121/1.4906835]

[SED] Pages: 1586–1597

I. INTRODUCTION

Passive acoustic methods have been shown to be a viable approach for small boat localization (Bruno *et al.*, 2010). These methods have some advantages over shore-based RADAR, infrared, and optical systems in that they are more robust to inclement weather and have the ability to be deployed in remote locations. They also have less impact on marine life than active sonar methods, and are also less susceptible to the effects of clutter. It was recently shown in Gebbie *et al.* (2013) that two bottom-mounted hydrophones constitute a simple yet favorable geometry for small boat localization, in that bathymetric variations can be leveraged to improve range localization and break the left–right ambiguity. This article describes an improved ambiguity surface function, and also provides a performance analysis to determine the conditions and limits under which these capabilities can be realized. Additional experimental data is used to assess the effect of hydrophone separation distance, and to validate the performance analysis. The manual processing of correlogram striations employed in the prior approach is replaced with an automated particle filter design, without any loss of overall algorithm generality.

The small boat localization technique described in this article is based on measuring and processing multipath time delays. This approach has been the subject of recent work in the field of marine mammal bioacoustics. It has been shown that certain species of whales, which use broadband clicks for echolocation, can be localized by measuring the time delay between the direct and bottom or surface-reflected multipaths to estimate range and depth (Aubauer *et al.*, 2000; Nosal and Frazer, 2006). Multipath arrivals are the sequence of echoes of a target's radiated sound, a result of waves reflecting from the surface and seabed in different sequences before reaching the receiver. Since whale clicks are typically noisy pulses of a short duration (Weirathmueller *et al.*, 2013), the delay between the direct and multipath arrivals can be discerned directly from the received time series (Thode, 2004). Tiemann *et al.* (2006) showed that by using bathymetry to pre-compute expected arrival times, full three dimensional (3D) localization could be performed from a single hydrophone. Each arrival corresponds to an eigenray, which is a unique acoustic path between a source and a receiver. For natural variations in bathymetry, if the animal is at a certain range and depth along one bearing, then the relative travel times of the eigenrays differ from those that would be observed if the animal were at the same range and depth along another bearing.

Unlike whale clicks, boat noise is distributed continuously in time, so the same measurement methodology does not directly apply. With boat noise, multipath arrivals overlap in time obscuring individual arrivals and their relative arrival times. The generalized correlation algorithm (Carter and Knapp, 1976) is a method of gathering the energy in broadband noise and compressing it into a single broadband pulse. The difficulty that arises with using noise correlation is that it produces peaks for all combinations of multipath

[a)]Author to whom correspondence should be addressed. Electronic mail: siderius@pdx.edu

arrivals. The problem becomes more tractable if the relative delays between the multipath arrivals are large, which can be realized by placing the hydrophones directly on the seabed.

The method outlined in Gebbie *et al.* (2013) was a proof of concept that employed a manual analysis of the correlogram. With a particle filter (Ristic *et al.*, 2004) that automatically processes the correlogram, systematic improvements are sought that allow for greater automation. The experimental data allowed for different hydrophone separations, and this is used to validate the performance characterization of the left−right disambiguation capability.

The rest of this article is organized in the following manner. The physics of multipath in shallow water, as it pertains to this work, is described in Sec. II. The processing algorithm is presented in Sec. III along with an explanation of how the left−right disambiguation capability operates. Experimental results from roughly 12 m of water are then reported in Sec. IV. Last, a discussion of the possible performance bounds that might be expected in deeper water is given in Sec. V, and this is computed for different hydrophone separation distances giving theoretical insight into how that parameter affects performance.

II. MULTIPATH STRUCTURE IN SHALLOW WATER

This section outlines how boat noise propagation through a shallow water waveguide produces the passive acoustic observations that serve as the input to the localization algorithm described later. The geometry of this problem is illustrated in Figs. 1(a) and 1(b). It shows two bottom-mounted hydrophones and a source (target) on the surface. The top two diagrammatic plots in Fig. 2 show the received time series, $r_1(t)$ and $r_2(t)$. The bottom plot shows their cross-correlation, $\mathbf{\Gamma}_{1,2}(\tau)$. The source waveform is represented as a single pulse for illustration purposes, however a boat's signature is continuous broadband noise. Thus $r_1(t)$ and $r_2(t)$ signals would appear as noise, obscuring the clean separation between multipath arrivals. However, this does not affect $\mathbf{\Gamma}_{1,2}(\tau)$ because the noise is compressed into a short-duration pulse through the cross-correlation operation. Environmental factors such as bottom loss and rough-surface scattering serve to decorrelate high-order eigenrays, whereas low-order eigenrays often retain enough coherence to appear as stable features in $\mathbf{\Gamma}_{1,2}(\tau)$; therefore, only the first-order multipath arrival is used.

Since receiver 1 is farther from the source than receiver 2, both peaks in $r_1(t)$ are shifted later in time to account for the additional travel time. The term, τ_0 denotes the time delay between direct arrivals, and $\tau_\pm$ are the delays between the direct arrival at one receiver with the multipath arrival at the other receiver. In $\mathbf{\Gamma}_{1,2}(\tau)$, the strongest peak is in the center, with an absolute offset at τ_0, and is used to constrain the ambiguity surface to a hyperbola on the surface. A correlation peak between just the multipath arrivals is not shown in $\mathbf{\Gamma}_{1,2}(\tau)$ because it shows up on the time axis close to τ_0, and is often hidden by the direct correlation. This makes measuring that time delay difficult, but information contained in the multipath-only correlation is included in the Bayesian

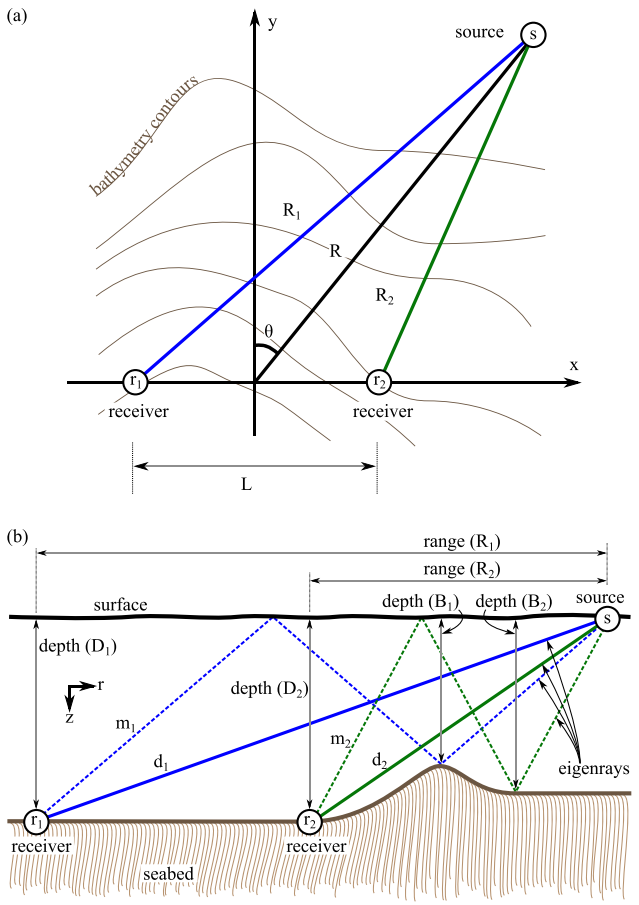

FIG. 1. (Color online) (a) A top-down view of the geometry. (b) A side view of the geometry.

Cramér-Rao lower bound (CRLB) calculation, described later in Sec. V. As the target initially moves into the far field of the hydrophone pair, τ_- and τ_+ start to converge but are sufficiently large that the flanking peaks are distinct from the center peak. This separation eventually vanishes in the distant far field. Bathymetric variations affect only τ_+ and τ_- because the lengths of reflected eigenrays depend on the depth at the location of the bottom reflection.

Let $t_{n,d}$ and $t_{n,m}$ be the travel times along each eigenray to the nth receiver for direct and multipath rays, respectively. The absolute time delays are defined as

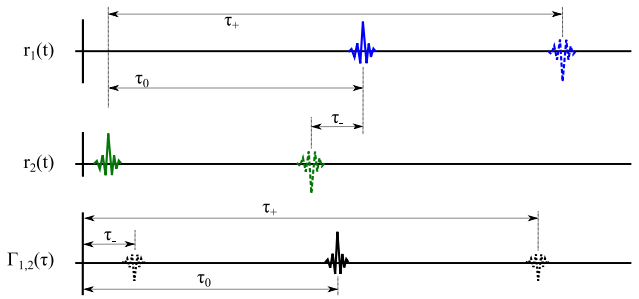

FIG. 2. (Color online) Diagrammatic plots of the relative multipath arrival times in Fig. 1. The top two plots are the time series at each receiver, and the bottom is their cross-correlation. For a source of continuous boat noise (instead of the pulse) the $r_n(t)$ would appear as noise, but the $\Gamma(\tau)$ would still exhibit the same peaks due to correlations between arrivals.

$$\tau_0 = t_{1,d} - t_{2,d}, \tag{1}$$

$$\tau_+ = t_{1,m} - t_{2,d}, \tag{2}$$

$$\tau_- = t_{1,d} - t_{2,m}. \tag{3}$$

Assuming a constant sound speed c, the ray travel times to the nth receiver can be written in terms of the target range R_n, receiver depth D_n, and seabed reflection depth B_n as

$$t_{n,d} = \frac{\sqrt{R_n^2 + D_n^2}}{c}, \tag{4}$$

$$t_{n,m} = \frac{\sqrt{R_n^2 + (D_n + 2B_n)^2}}{c}. \tag{5}$$

The ranges to each hydrophone R_n can be written in terms of the range R to the midpoint between the receivers, the spacing between the hydrophones L, and the target bearing θ using the law of cosines

$$R_1 = \sqrt{R^2 + (L/2)^2 + RL\sin\theta}, \tag{6}$$

$$R_2 = \sqrt{R^2 + (L/2)^2 - RL\sin\theta}. \tag{7}$$

These equations describe the relationship between the measured time delays and the range and bearing of the target, which are used in the CRLB simulation presented in Sec. V.

A. Cross-correlation and cross-spectrum

The time delays between arrivals contain the information needed for target localization, but these quantities are only observable in the second order statistics of the received signals. While time delay information is contained in both the auto- and cross-spectra, only the cross-spectra is used as an observation since it contains the time delay of the direct arrivals, which is strongly informative of bearing. However, both auto- and cross- spectra are used in Sec. V to compute the Bayesian CRLB.

Only the first two arrivals at each receiver are needed to estimate the range and bearing. Although later arrivals also contain this information, they are limited by lower levels and increased decorrelation due to additional rough surface reflections and greater path lengths. As such, the received time series at each receiver is modeled with the first two arrivals and additive uncorrelated Gaussian noise, $n(t)$. Multipaths beyond the first two arrivals are treated as uncorrelated noise at each receiver that contributes to $n(t)$. The source waveform is modeled as Gaussian noise represented by $s(t)$. The received waveforms are thus defined as

$$r_1(t) = s(t - t_{1,d}) + \beta s(t - t_{1,m}) + n_1(t), \tag{8}$$

$$r_2(t) = s(t - t_{2,d}) + \beta s(t - t_{2,m}) + n_2(t), \tag{9}$$

in which β represents the additional amount of propagation loss relative to the first arrival, is real-valued in the range $(-1,1)$, and does not vary with time. The same β is used in both Eqs. (8) and (9) based on the assumptions that the hydrophone spacing is small relative to the target range implying that the ray grazing angles on the seabed are similar, and that the composition of the seabed is locally homogeneous in the vicinity of the two reflection points. Phase changes due to boundary reflections are not modeled as this information is later discarded using the envelope operation. The Fourier transform at each receiver is

$$Y_1(\omega) = S(\omega)e^{-iwt_{1,d}} + \beta S(\omega)e^{-i\omega t_{1,m}} + N_1(\omega), \tag{10}$$

$$Y_2(\omega) = S(\omega)e^{-i\omega t_{2,d}} + \beta S(\omega)e^{-i\omega t_{2,m}} + N_2(\omega). \tag{11}$$

Henceforth, we drop the explicit dependence on ω.

Cross-correlation consists of holding one signal constant while sliding the other signal on the time axis, multiplying, and integrating. The fixed signal is referred to here as the "correlated" signal, and the sliding signal as the "reference" signal. Let r_1 be the correlated signal and r_2 be the reference signal. The cross-spectrum is computed by multiplying the correlated spectrum by the complex conjugate of the reference spectrum,

$$\begin{aligned} C_{1,2} &= E[Y_1 Y_2^*] \\ &= |S|^2 e^{-i\omega(t_{1,d} - t_{2,d})} + |S|^2 \beta^2 e^{-i\omega(t_{1,m} - t_{2,m})} \\ &\quad + |S|^2 \beta e^{-i\omega(t_{1,m} - t_{2,d})} + |S|^2 \beta e^{-i\omega(t_{1,d} - t_{2,m})}, \end{aligned} \tag{12}$$

in which $E[]$ is the expected value operator, and * denotes complex conjugation. The second term is the correlation of multipath arrivals on each receiver, which is effectively obscured by the correlation of direct arrivals (first term). This is due to the same time delay for the second term having a lower amplitude due to β^2.

For a finite bandwidth signal with a flat spectrum, each exponential term will appear in the time domain as a sinc function. For a baseband signal of bandwidth b, the sinc function is defined as $\sin(2\pi bt)/(2\pi bt)$. If β were complex and had a constant phase with respect to frequency, as might be the case with sub-critical bottom reflections from a single-layer seabed, this would cause skewing of the sinc function such that the maximal value would not align exactly with the time delay. The envelope operation is thus applied to produce positive valued peaks having a maximal value at the time delay, which simplifies time delay estimation. The envelope of an arbitrary signal $x(t)$ is computed by $|x(t) + \mathcal{H}[x(t)]|$, in which $\mathcal{H}$ is the Hilbert transform. This is equivalent to computing the inverse Fourier transform of only the positive frequencies, followed by taking an absolute value and multiplying by a factor of 2.

The β^2 term in Eq. (12) is often hidden by the first term (correlation of direct arrivals) such that the envelope of the cross correlation time series consists of three pulses. A useful way to approximate these pulses is with a Gaussian function, for which the width and offset are easily modulated, and is defined as

$$g(x, a, \sigma) \triangleq e^{-(1/2)(x-a/\sigma)^2}. \tag{13}$$

The cross correlation time series is then approximated as

$$\Gamma_{1,2}(\tau) \approx g(\tau, \tau_0, \sigma) + \beta g(\tau, \tau_-, \sigma) + \beta g(\tau, \tau_+, \sigma) \quad (14)$$

in which the time axis is $x = \tau$, and the pulse offsets occur at $a \in \{\tau_0, \tau_-, \tau_+\}$. The width of each pulse depends on the signal bandwidth b, and since the variable σ specifies the half-width of the Gaussian, it is defined as

$$\sigma \triangleq \frac{1}{2b}. \quad (15)$$

Equation (14) is an important relationship in that the time offset of the three peaks contain information about the target range and bearing.

The auto-spectrum for each receiver is defined as

$$C_{1,1} = E[Y_1 Y_1^*] = |S|^2\{1 + |\beta|^2 + 2|\beta| \cos[\omega(t_{1,m} - t_{1,d}) - \theta_\beta]\} + |N_1|^2 \quad (16)$$

with a similar definition for $C_{2,2}$. This is required later in the calculation of the range CRLB in Sec. V. Note that unlike the cross spectrum, the auto spectrum includes the noise term. Section III describes the methodology for processing the acoustic data into a localization.

III. LOCALIZATION ALGORITHM

The localization algorithm consists of the following parts: the extraction of multipath time delays, the ray model, and the ambiguity function. The first step processes the acoustic data, the second step runs the model at discretized points on the surface, and the third step combines them into an ambiguity surface that yields the localization.

A. Acoustic processing

The measured cross-correlation function, $\mathbf{\Gamma}(\tau)$, is obtained from raw data by the following procedure. A snapshot of data is obtained from the channels; the time span of this data constitutes the total averaging time and should be short enough so that the effects of target motion are negligible. Overlapping segments are formed within the snapshot using the weighted overlapping segment averaging method (Carter *et al.*, 1980). Segments are windowed using a Hann function, then zero-padded to twice the original length to avoid wrapping effects of the discrete Fourier transform (DFT). Time windowing provides better spectral estimates at the expense of a small amount of frequency resolution. The cross-spectrum is computed for each segment by multiplying the spectra of one channel with the complex-conjugate of the other, followed by a bandpass filter. The average for the snapshot is computed across all segments. The resulting cross-spectrum is pre-whitened. Pre-whitening preserves phase information while enforcing a flat power spectrum and is defined as $X(\omega)/|X(\omega)|$ for an input spectrum $X(\omega)$, and is based on the phase transform (PHAT) algorithm (Knapp and Carter, 1976). This prevents loud tonal components that may stand out in the target's acoustic signature from corrupting $\mathbf{\Gamma}(\tau)$. Last, this is brought back into the time domain with the inverse DFT while simultaneously taking the envelope.

A correlogram is then formed by stacking $\mathbf{\Gamma}(\tau)$ from each snapshot vertically such that τ is on the horizontal axis and the absolute time of each snapshot is on the vertical axis. Viewed in this manner, striation lines appear that correspond to the correlations of individual arrivals as they evolve over time. The center striation, which has an offset of τ_0 at a given snapshot, is often the strongest as it corresponds to correlations of the direct arrivals on each hydrophone. The nearest flanking striations are the multipath arrival at one hydrophone correlating with the direct arrival at the other hydrophone, and have offsets $\tau_\pm$ for a given snapshot.

A sequential importance resampling (SIR) particle filter (Ristic *et al.*, 2004) is then used to extract the striation offsets from the correlogram. The output of the tracker are the measured values τ_0, τ_-, and τ_+ at each snapshot. The details of the tracker construction are given in the Appendix.

B. Ray model processing

In the second stage of processing, predictions of eigenray propagation-time differences, $\bar{\tau}_0(\boldsymbol{x})$, $\bar{\tau}_-(\boldsymbol{x})$, and $\bar{\tau}_+(\boldsymbol{x})$, are computed for each possible (Easting, Westing) target position, $\boldsymbol{x}$. A ray tracer, such as described in Jensen *et al.* (2011), can be used to compute the travel time of the direct and bottom-surface eigenrays between each receiver and each point $\boldsymbol{x}$ on the surface. An eigenray between a point on the surface and a receiver on the seabed in a region of varying bathymetry could involve a path that is not confined to a vertical plane due to three dimensional seabed variations. While this could be done in three dimensions, a simpler N-by-2D approach is used here. A N-by-2D ray model invokes a 2D ray tracer separately along N radial lines for each receiver, and the bathymetry is interpolated along each radial line. This flattens the problem to two dimensions, range and depth, and a single ray fan starting at the receiver can be used to determine the eigenrays for all points along the radial which greatly reduces computational complexity. The $\bar{\tau}_0(\boldsymbol{x})$, $\bar{\tau}_-(\boldsymbol{x})$, and $\bar{\tau}_+(\boldsymbol{x})$ at each radial point can then be gridded using a technique such as Delaunay triangulation (de Berg, 2008). A constant water sound speed profile (SSP) is adopted in this analysis, but the technique can be readily applied to environments having a varying SSP. The choice of constant SSP is based on the experiment described in Sec. IV, which was close to isovelocity.

C. Ambiguity surface for target location

In the last stage of processing, the measured time delays are matched against modeled time delays computed at regular points on the surface. An ideal ambiguity function would be a delta function at the true target location, but in practice the aim is to maximize the value at the true target location relative to all other locations. At the true target location, the error between the predicted time delays and the measured time delays will be minimized, provided the model is sufficiently accurate. It is important to point out that *all* three delay errors that are minimized together at this point: $|\tau_0 - \bar{\tau}_0(\boldsymbol{x})|$, $|\tau_+ - \bar{\tau}_+(\boldsymbol{x})|$, and $|\tau_- - \bar{\tau}_-(\boldsymbol{x})|$. When $|\tau_0 - \bar{\tau}_0(\boldsymbol{x})|$ is at a

minimum, this corresponds to a hyperbola on the $\boldsymbol{x}$ plane. When the other two are also at a minimum, this corresponds to a range at some distance down one leg of the hyperbola. However, since the hyperbola has two legs, this also corresponds to a "false" target position on the other leg. The false target position may retain a greater amount of overall error due to bathymetric variations.

Thus, in order to constrain the ambiguity function to produce a large value at the true target position, it is constructed as the product of three Gaussians. Using Eq. (13),

$$\Phi(\boldsymbol{x}) \triangleq g[\tau_0, \bar{\tau}_0(\boldsymbol{x}), \sigma] \times g[\tau_+, \bar{\tau}_+(\boldsymbol{x}), \sigma] \times g[\tau_-, \bar{\tau}_-(\boldsymbol{x}), \sigma]. \quad (17)$$

The first term constrains the target location to the hyperbola determined by the direct arrivals. The latter two terms are maximized at the true target range, and possibly the false target range. Values of $\Phi(\boldsymbol{x})$ are in the range (0, 1] and are interpreted as yielding information about the relative certainty of the target being at a particular location, $\boldsymbol{x}$, on the water surface.

D. Left–right disambiguation

An interesting feature of the ambiguity surface defined in Eq. (17) is that in addition to estimating the target range and bearing, it can also predict (with some uncertainty) which "side" of the array the target is on. This predictive capability is manifest as a taller peak in the ambiguity surface on the "true" side of the array (the side of the array where the target is actually located). This happens naturally in some situations by virtue of the construction of Eq. (17), and in those situations no further processing steps are required.

Consider a single snapshot of acoustic data that is short enough to reasonably ignore target motion. The bathymetry is varied such that depths on one side of the array differ from the corresponding "mirror" points on the other side of the array. This difference affects propagation which causes the multipath travel times to differ. Equation (17) will reveal an ambiguity surface with two peaks, one on "true" side and one on the "false" side. The height of these peaks depends on error between modeled and measured time delays for each set of arrivals on each side. These errors are random variables based on the inherent uncertainty in measuring time delays with the generalized cross correlator (Carter and Knapp, 1976). Errors due to environmental uncertainty are ignored in this section, and the environment is assumed to be perfectly characterized.

On the true side, there exists a possible target position on the surface for which these random variables will be normally distributed with zero mean, i.e., the true target position. However, such a position does not necessarily exist on the false side, implying that some of the means will deviate from zero. The first term of Eq. (17) constrains the ambiguity surface to a hyperbola based on the delay between direct arrivals. This is oriented with a leg extending onto each side of the array. Conceptually, sliding along the hyperbola on the true side, at some point the multipath delays will exactly match the modeled predictions (no environmental uncertainty) at the true target range. However, on the false side, the multipath delays may become close, but do not perfectly match the modeled predictions at the false target range. This small amount of error, if it exists, is amplified by Eq. (17) and is often sufficient to produce a taller peak on the true side.

Let point $\boldsymbol{x}_t$ be the location of the peak on the true side, and $\boldsymbol{x}_f$ be the tallest peak on the false side. Ultimately, the question being asked is what is the probability that the peak on the true side is taller than the peak on the false side? This quantity is expressed as

$$P_t \triangleq P[\Phi(\boldsymbol{x}_t) > \Phi(\boldsymbol{x}_f)]. \quad (18)$$

Let

$$\alpha_b^a \triangleq [\tau_a - \bar{\tau}_a(\boldsymbol{x}_b)]^2, \quad \begin{matrix} a \in \{0, -, +\} \\ b \in \{t, f\}. \end{matrix} \quad (19)$$

Combining the terms of Eq. (17), it can be shown that Eq. (18) is equivalent to

$$P_t = P[\alpha_t^0 + \alpha_t^+ + \alpha_t^- < \alpha_f^0 + \alpha_f^+ + \alpha_f^-]. \quad (20)$$

On each side of this inequality note that the measured values (τ_a) are normally distributed random variables. On the true side, their means equal the corresponding predictions $\bar{\tau}_a$, so $\tau_a - \bar{\tau}_a(\boldsymbol{x}_t)$ are zero-mean normally distributed random variables. However, on the false side, there may be some error between the mean of τ_a and the prediction $\bar{\tau}_a$, leading to $\tau_a - \bar{\tau}_a(\boldsymbol{x}_f)$ being normally distributed random variables but possibly having a non-zero mean.

Since each α_t^a is the square of a normal random variable, they are each χ_k^2 distributed with $k = 1$ degrees of freedom. Added together, the left side of the inequality (the three α_t^a's) form a χ_k^2 random variable with $k = 3$ degrees of freedom. On the right side, the α_f^a random variables are the product of normal random variables that have the same variance, but do not necessarily have zero mean, so these have non-central $\chi_{k,\kappa}^2$ distributions (Zelen and Severo, 1972). Since

$$\alpha_t^0 + \alpha_t^+ + \alpha_t^- \sim \chi_k^2 \quad \text{and} \quad \alpha_f^0 + \alpha_f^+ + \alpha_f^- \sim \chi_{k,\kappa}^2, \quad (21)$$

then

$$P_t = P[\chi_3^2 < \chi_{3,\kappa}^2]. \quad (22)$$

The parameter κ is the non-centrality parameter defined as

$$\kappa = \sum_{\forall a} \left(\frac{\tau_a - \bar{\tau}_a(\boldsymbol{x}_f)}{\sigma_{\tau_a}} \right)^2, \quad (23)$$

so the probability P_t depends entirely on κ. The parameter σ_{τ_a} is the standard deviation of τ_a, and depends on the accuracy with which the time delay can be measured. This relationship between central and non-central χ^2 distributions is shown in Fig. 3. Equation (22) can be rearranged as the distribution of the difference between two random variables,

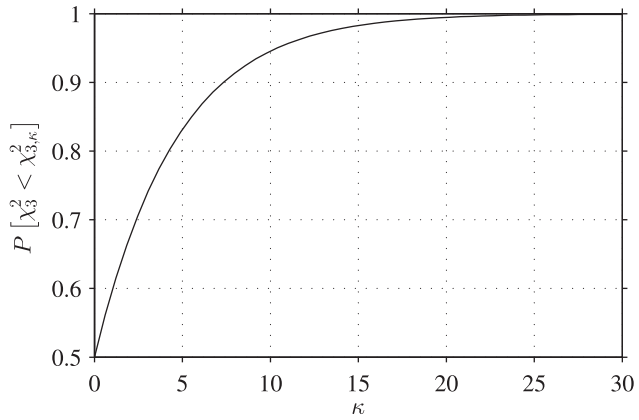

FIG. 3. Probability that χ_k^2 is less than a non-central $\chi_{k,\kappa}^2$ as a function of κ for $k=3$.

which can be computed by correlating their respective distribution functions, or alternatively conjugate-multiplying their characteristic functions. When $\kappa=0$, the non-central $\chi_{k,\kappa}^2$ converges to a central χ_k^2 and the probability that one is less than the other is 0.5, which would imply there is no time delay error on the false side making it indistinguishable from the true side. Note that Eq. (23) indicates that the accumulation of normalized squared time delay errors on the false side determine the array side prediction performance.

1. Simulation

Consider an environment with a planar but not flat bathymetry. The plane has a constant slope such that one side of the array is deeper than the other. The ability to break the left−right ambiguity is based on one side having different multipath characteristics than the other, so the gradient vector cannot be perfectly parallel to array. Rather, it is the component of the gradient that is perpendicular to the array that is relevant. This is important because it says that to determine which side of the array the target is located, sensors deployed on a slope should be placed along bathymetric contours as opposed to an uphill−downhill arrangement. With an uphill−downhill arrangement, the bathymetry along the true and false legs of the hyperbola are identical, which eliminates differences in the multipath that are critical to this capability. In this simulation, the gradient vector is perpendicular to the array, and the sensors are on the bottom in 120 m of water. The simulation places the target on the deeper side at 1 km range at a bearing of 45°. The sound speed in the water is assumed to be constant with boundary reflections that are lossless.

In this case the environment is completely known, so the measured time delays are solely the result of correlating Gaussian noise signals in the presence of additive Gaussian noise. Hahn and Tretter (1973) derived the CRLB for the variance of the time delay in this scenario. Knapp and Carter (1976) further showed that this bound is reached with the generalized cross correlation algorithm, which is the same method used here. If the signal and noise have flat spectra in the band $[f_{\min}, f_{\max}]$ (using positive frequencies only), then this bound is specified as

$$\sigma^2 = \frac{3f_s[1+2(\mathrm{SNR})]}{4\pi N\left(f_{\max}^3 - f_{\min}^3\right)(\mathrm{SNR})^2}, \tag{24}$$

where N is the number of samples, f_s is the sample rate, and SNR is the signal to noise ratio. The values used in the simulation are $f_s = 102\,400$ Hz, $N=4096$, $f_{\min}=0$ Hz, $f_{\max}=3000$ Hz, and SNR = 10 dB.

The simulation is based on image theory. Since the seabed is a plane, the target is reflected over it to its image position. The receivers are then reflected over the surface to their image positions. Multipath time delays between the source and each receiver are computed using the distance between the image source and the image receivers. The ambiguity surface, Eq. (17), is computed on the shallow side to determine the location of false target. The false target position corresponds to a set of time delays that are close to, but do not necessarily match those of the true target position. These time delay errors exist for both direct-with-direct and direct-with-multipath correlations. From these errors, Eq. (23) is invoked along with Eq. (24) to determine κ, which maps directly to a probability value that specifies whether the peak at the true target position will be taller. The results of this simulation are shown in Fig. 4 for variations in the seabed slope and receiver separation.

As the seabed becomes more sloped, the left−right disambiguation capability increases. Conversely, a flat seabed eliminates the unique propagation characteristics from each side of the array, which removes the left−right disambiguation capability. Consequently, the probability of choosing

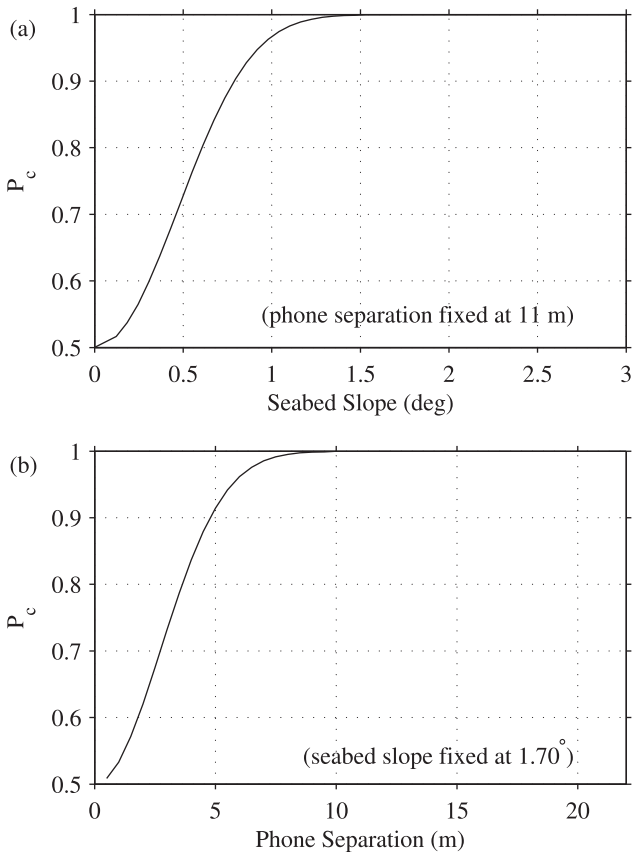

FIG. 4. (a) Probability of choosing the true side for different perpendicular seabed slopes. (b) Same method for varying phone separation.

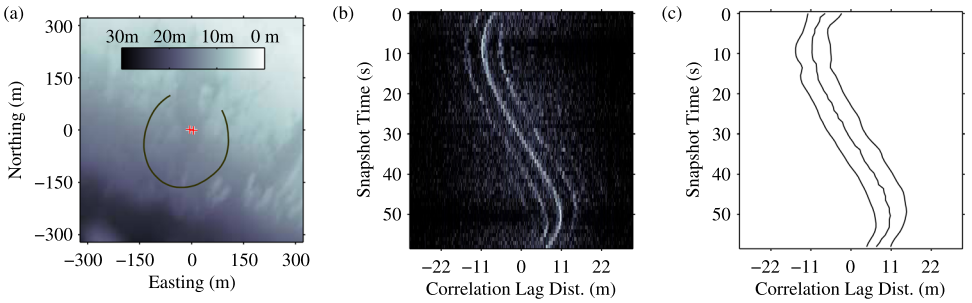

FIG. 5. (Color online) (a) Bathymetry and GPS boat track. The solid line shows the track of the small boat with a counter-clockwise trajectory. The "+" annotations indicate array element locations. (b) Correlogram showing 10 $\log_{10}|\Gamma(\tau)|^2$ for consecutive snapshots, plotted with 30 dB of dynamic range, and with τ converted to wave travel distance. (c) Striation lines for τ_0 and $\tau_\pm$ output by the SIR particle filter.

the correct side reduces to 0.5. The same behavior is observed for the separation distance between receivers. As the receivers become closer together, even on a sloped seabed, the multipath rays from the target to each receiver become increasingly similar. Namely, the points on the seabed from which reflections occur start to converge, and the relative difference between pairs of rays on each side of the array correspondingly decreases. This suppresses the left−right disambiguation capability.

IV. SHALLOW WATER EXPERIMENT

Passive acoustic signals were collected in August 2011 by a moored horizontal line array (HLA) at the Kilo Nalu Nearshore Reef Observatory (Gebbie *et al.*, 2011). The observatory provided power and ethernet connectivity via an undersea cable running approximately 0.4 km from shore to a fixed underwater station deployed in roughly 12 m of water. The station was located about 1 km southeast of Honolulu Harbor, a commercial port. Only two elements of the HLA, spaced 11 m apart, were used in this study. The array was configured with a sample rate of 102.4 kHz, 24-bit dynamic range, 300 Hz low-cut filter, and 110 dB anti-aliasing filter set at 46.4 kHz. The hydrophones (HTI-92-WB) had a sensitivity of −160 dB re 1 V/μPa.

A rigid-hulled small boat with a single outboard engine was used as a target, and a handheld GPS device recorded its location. Time stamps in the GPS data and recorded acoustic data allowed for coarse-grained synchronization (on the order of 1 s) between the two sets of data. The boat executed several different maneuvers including driving in circles around the array deployment site. Spectral analysis indicated that the boat radiated noise in the 0−10 kHz band with the bulk of the energy below 3 kHz.

Bathymetry information for the local area was obtained from the SHOALS LIDAR bathymetry database at the University of Hawai'i (University of Hawai'i at Manoa, Costal Geology Group, 2012) which was ungridded data having roughly 1 m resolution. Grab samples near the deployment site indicated the seabed was composed of medium/coarse sand. The bathymetry, hydrophone locations, and track of the boat are shown in Fig. 5(a). The bathymetry and hydrophone locations were used to compute the direct and bottom-surface eigenrays for a grid of points on the surface.

A. Results

A correlogram is shown in Fig. 5(b) in which multipath effects are evident. The strong, center striation is the correlation of direct arrivals. This is supported by the fact that as the target circles around the array, this striation stays between ±11 m, which are the limits for the correlation lag distance for the configured hydrophone spacing of 11 m. The multipath-with-direct correlations are visible as "shadow" striations that run adjacent to the main striation. All these striation lines were automatically traced using a SIR particle filter (see the Appendix) and are shown in Fig. 5(c). Shadow striations from higher-order eigenrays are also faintly visible throughout the entire run, but are not used in this processing.

A comparison of localization using only the direct arrivals [corresponding to using just the first term of Eq. (17)] with all three arrivals [all terms of Eq. (17)] is shown in Figs. 6(a) and 6(b). The effects of using a flat seabed versus

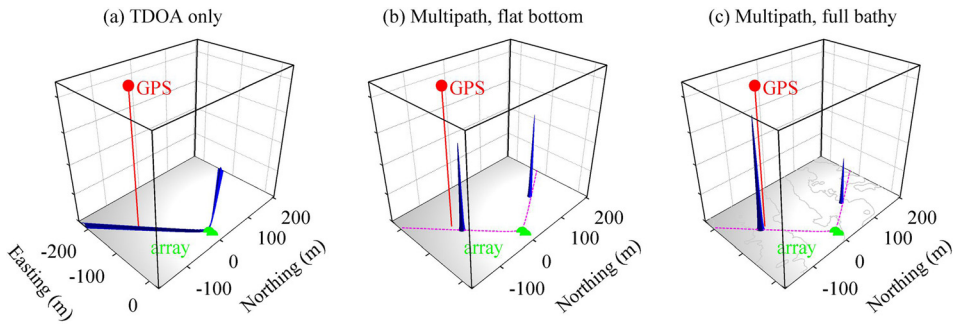

FIG. 6. (Color online) Comparison of ambiguity surfaces, $\Phi(\boldsymbol{x})$, for a single snapshot showing the effect of utilizing multipath and bathymetry information. Plots are normalized to unit volume to show the relative concentration of target location certainty. (a) A hyperbolic ambiguity is associated with only using the first term of Eq. (17). (b) Inclusion of multipath [the latter two terms of Eq. (17)] and the assumption of a flat seabed cause the hyperbola to collapse to a single range, but a left−right ambiguity remains. (c) Using actual bathymetry improves the range estimate of the peak near the true target and also allocates a greater amount of target location certainty. Contour lines are shown at 2.5 m intervals. In (b) and (c), the hyperbola defined by τ_0 is shown.

actual bathymetry are shown in Figs. 6(b) and 6(c), in which the peak near true target location is amplified with respect to the false target location. The full ambiguity function, $\Phi(\boldsymbol{x})$, is shown for several snapshots throughout the boat track in Fig. 7(a) for an 11 m spacing. The same processing was applied to phones separated by 2 m for the same times, and the results are shown in Fig. 7(b). These surfaces are the summation of individual ambiguity functions over a sequence of snapshots. Two things are evident with the shorter separation: Each peak is less sharply defined in bearing due to the smaller aperture, and the false peaks on the near side are more prominent indicating a decrease in the left–right disambiguation capability. Poorer bearing resolution for shorter array lengths is consistent with the theoretical behavior described in Sec. V. The image data in Figs. 6 and 7 were post-processed with a 2-D Hann filter to aid visualization of narrow features. Range errors at the start of the tracks in Figs. 7(a) and 7(b) correspond to the SIR particle filter locking onto the striation lines.

V. DISCUSSION

This section examines the theoretical limits on how accurately the target range can be determined in water depths that are on the order of hundreds of meters, typical of many continental shelves. It also addresses the issue of range estimation sensitivity to hydrophone separation distance. Over the past several decades, many studies have investigated the performance limits of algorithms for localizing noise-producing targets. This has mainly been done by deriving and calculating the CRLB for the target location. The CRLB represents the lowest achievable variance of an unbiased estimator, and conceptually operates by computing the amount of information passed from a set of hidden parameters (i.e., the source location) to the set of observed parameters (i.e., received waveforms). These bounds hold under the assumption that only the hidden parameters are random, and all other parameters are deterministic. Friedlander (1988) used this to compute the range and depth accuracy of a submerged source from two vertically aligned receivers using the direct arrival and a single multipath arrival.

Van Trees (1968) expanded on the CRLB to also handle non-deterministic parameters, in a formulation often referred to as the Bayesian CRLB. The term "Bayesian" is used to indicate that prior information about these random parameters is being used. This was the basis for the study by Hamilton and Schultheiss (1993) which examined the performance impact of having imperfect knowledge of the bathymetry by treating the depth of reflections from the seabed as normally distributed random variables. In the analysis presented here, a two-hydrophone geometry is adopted, similar to Friedlander (1988), but the Bayesian CRLB approach of Hamilton and Schultheiss (1993) is applied while treating several additional parameters as non-deterministic. This more accurately represents the imperfect knowledge typically available in a real deployment scenario.

The experiment described in Sec. IV was in 12 m of water, however, it is of interest to evaluate whether sufficient information exists for the effective use of this technique in deeper water. This is addressed later in this section via a simulation. Further, the CRLB formulation provides a useful means of evaluating performance sensitivity to individual system parameters, and this is used to estimate performance sensitivity to the effects of hydrophone separation distance, again via simulation.

A. CRLB problem formulation

In a localization problem, the CRLB places a lower bound on how accurately the location can be estimated based on how much information about the location exists in the observed data. As the amount of information about a hidden location parameters increases, the variance with which they can be estimated decreases, and vice-versa. In fact, the definition of the CRLB is that it is the inverse of the Fisher information matrix. The i, j element of the Fisher information matrix is defined as

$$J_{D_{i,j}} = E\left[\frac{\partial \log p(y|\xi_i)}{\partial \xi_i}\frac{\partial \log p(y|\xi_j)}{\partial \xi_j}\right], \tag{25}$$

in which $p(y|\xi)$ is the conditional probability density function of an observed quantity y given a hidden parameter ξ.

The basis for this analysis starts with an article by Friedlander (1988), in which the target range and depth CRLB were derived. In that scenario, Gaussian noise is emitted by a submerged source and travels along two paths to a pair of vertically arranged receivers. One path is direct and the other reflected. Propagation delays are specified as a

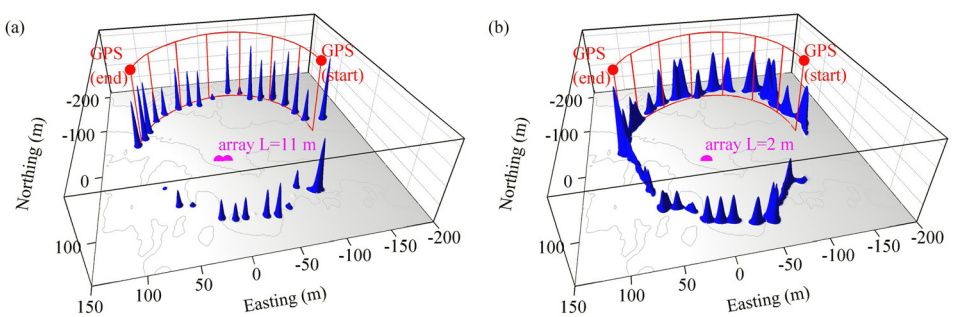

FIG. 7. (Color online) Summation of a ambiguity surfaces, $\Phi(\boldsymbol{x})$, over a sequence of snapshots comparing two array lengths, (a) $L = 11$ m, (b) $L = 2$ m. The boat is moving in a circle around the array, and the data shown here is when it is on the south side. The camera is pointed mainly southward. The longer array length leads to better left–right disambiguation and better resolution in bearing. Within a few cycles, the particle filter locks onto the striation lines.

function of source range and depth, and are considered the hidden parameters. Instead of applying Eq. (25) directly, Whittle's theorem [Whittle (1953), Theorem 9] is invoked to compute the Fisher information matrix by integrating a kernel based on the cross spectral density matrix and its derivatives over a band of frequencies. In that formulation, the i,j element of the Fisher information matrix is

$$J_{D_{i,j}} = \frac{N}{2\pi f_s}\int_{\omega_{\min}}^{\omega_{\max}} \mathrm{tr}\left[\frac{\partial C}{\partial \xi_i}C^{-1}\frac{\partial C}{\partial \xi_j}C^{-1}\right]d\omega, \tag{26}$$

in which tr is the trace operator, $[\omega_{\min}, \omega_{\max}]$ is the signal frequency band, ξ_i and ξ_j are parameters (e.g., range, bearing), N is the number of samples, and f_s is the sample rate.

The term C is the cross spectral density matrix for which $C_{1,1}$ and $C_{2,2}$ are described by Eq. (16), and $C_{1,2}$ is defined by Eq. (12). Note that $C_{2,1} = C_{1,2}^*$. All the parameters (the ξ's) are assumed to be hidden. Parameters that are deterministic are simply used in the calculation of C and the derivatives of C with respect to the hidden parameters. In Friedlander (1988), the environment is deterministic, and the only hidden parameters are the target range and depth.

The Fisher information divides parameters into two categories: those for which no prior information is available and those that are known exactly. Obviously, not all problems can be defined in this way. However, Van Trees (1968) proposed a "Bayesian" version of the CRLB, often referred to simply as the Bayesian CRLB, that allowed a hidden parameter to have some (but not necessarily perfect) prior information. Hamilton and Schultheiss (1993) used this formulation to determine the CRLB for a target's range, but treated the depth of the multipath reflection from the seabed as a hidden parameter having a finite prior variance.

The Bayesian "prior" information about the variances of the parameters is specified in the matrix

$$J_P = \mathrm{diag}[\mathrm{var}(\xi_1)^{-1}, \mathrm{var}(\xi_2)^{-1}, \ldots, \mathrm{var}(\xi_N)^{-1}], \tag{27}$$

in which diag puts elements on the main diagonal, and var(ξ) is the variance of ξ. The total Fisher information matrix is then

$$J_T = J_D + J_P, \tag{28}$$

and the lower bounds on the variances of the individual parameters fall on the main diagonal of

$$\text{Bayesian CRLB} = J_T^{-1}. \tag{29}$$

At the extremes, an infinite prior variance implies no prior information about that parameter exists, which drives that element of J_P to zero thus increasing the CRLB. Conversely, a small variance for the prior implies accurate knowledge of a parameter, thus increasing the value in J_P and decreasing the CRLB. Intuitively, as more prior information about a parameter is included (corresponding to a larger value somewhere on the diagonal of J_P), that parameter, and potentially other parameters as well, can be estimated with a smaller variance corresponding to an overall smaller CRLB.

Whittle's theorem, Eq. (26), requires second-order derivatives of the cross spectral density matrix with respect to the parameters. For this reason, constant sound speeds are assumed throughout much of the literature since relationships between the water sound velocity profile and the observed relative ray travel times are difficult to determine analytically. For simplicity, a constant sound velocity profile is adopted here and used to show what affect treating other parameters as uncertain has on range localization accuracy. The derivatives of C are taken for all pairs of the parameters R, θ, L, D_1, D_2, B_1, B_2, β, and c, such that J_D is a 9×9 matrix. These derivatives are computed analytically using a symbolic math engine (MathWorks, 2013) and used directly in the simulation described in Sec. V B. The individual equations that make up J_D are large, and the details are not particularly germane so are omitted here. Many other works have presented explicit derivations [e.g., Friedlander (1988); Hamilton and Schultheiss (1993)], and the method presented here follows essentially the same procedure.

B. Simulation in 120 m water

In this simulation, the Bayesian CRLB for range (specified as σ_R) and bearing (specified as σ_θ) are computed as a function of the true target range. This is done for long and short hydrophone separation distances in order to examine performance sensitivity to this parameter. The hydrophone separation is also modeled to have a small amount of error, as would be the case if the hydrophones were mounted on a taut cable. The seabed is assumed to be roughly flat with some minor variations in composition, and the depth of the receivers also has a slight amount of error. The water has a constant sound speed, but there is some uncertainty in the exact speed. In Eq. (26), $N = 4096$, $f_s = 102.4$ kHz, $[\omega_{\min}, \omega_{\max}] = 2\pi[0, 3000]$. The source level emits 130 dB and the noise level is 60 dB. As the target range increases, the signal to noise ratio (SNR) at the receiver decreases. The parameter values (μ_ξ) for the simulation are given in Table I, along with the Bayesian priors (σ_ξ). Range (R) and bearing (θ) are the parameters under investigation, and so their infinite priors cause their entries in J_P to be zero.

The Bayesian CRLB for range (R) and bearing (θ) are shown in Figs. 8(a) and 8(b), respectively, for both hydrophone separations (L). As the target moves farther from the array, the lower bound on range estimation variance increases, which corresponds to the direct and multipath

TABLE I. Hidden parameters for CRLB calculation.

Parameter (ξ)	Value (μ_ξ)	Bayesian prior (σ_ξ)
R	100 – 1500 m	∞
θ	45°	∞
L	11, 2 m	0.02 m
D_1	120 m	2 m
D_2	120 m	2 m
B_1	120 m	10 m
B_2	120 m	10 m
β	−0.9	0.01
c	1530 m/s	5 m/s

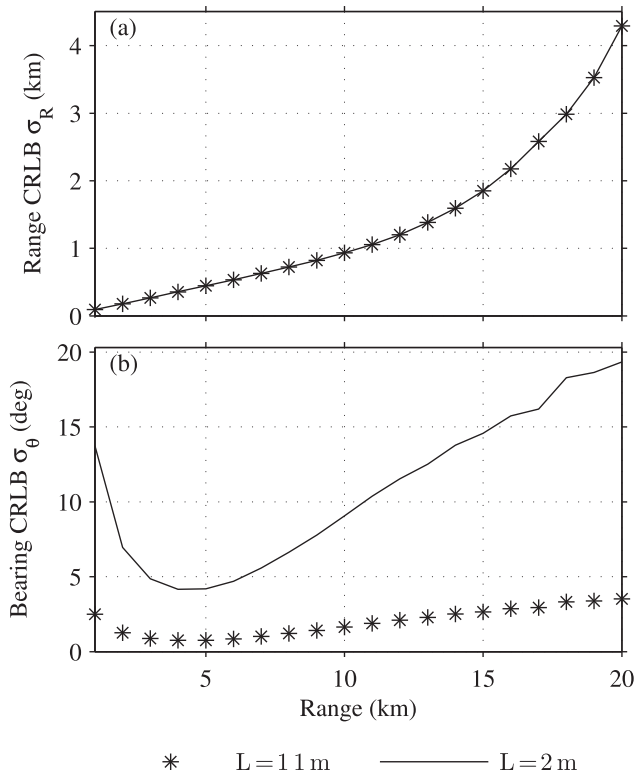

FIG. 8. (a) Bayesian CRLB for range estimation as a function of true range. Dashed and solid lines representing two phone separations essentially overlap indicating that this factor does not impact range resolution. (b) For comparison, the same metric is applied to bearing estimation, in which the dashed and solid lines show significant deviations, which indicates bearing resolution decreases as the phone separation decreases.

eigenray travel times converging, and therefore containing less information about the target range. The lines for the 2 m and 11 m hydrophone separations essentially overlap, indicating that hydrophone separation does not affect the lower bound on range variance. However, it does have a significant impact on the bearing estimation, σ_θ, shown in Fig. 8(b), as expected. The increase in bearing CRLB at closer ranges are due to the uncertainty in the receiver depths.

VI. CONCLUSION

This article presents a simple technique for localizing a small boat using multipath arrivals recorded on two bottom-mounted hydrophones. The conditions under which this approach can be expected to perform well are also determined. Wider hydrophone separation distances provide better bearing accuracy, but yield no improvement in range accuracy. Bathymetric variations affect multipath arrival times, a fact that can be exploited to break the left−right ambiguity on sloped seabeds. To expose this capability, array elements should be placed along contours of the bathymetry, rather than in an uphill-downhill arrangement. It was also determined that left−right disambiguation improves with wider hydrophone separation distances, and with greater seabed slope. Placing the hydrophones on the seabed has the advantage of maximizing the time separation between multipath arrivals, which facilitates automated extraction from a correlogram. Experimental results demonstrate operation in 12 m of water, but simulations suggest that the technique may work in deeper water on the order of hundreds of meters.

One application for this technique might be monitoring small boat traffic in shallow water regions with relatively minimal deployment complexity and equipment. A factor that is likely to impact performance of this technique is error in the vertical water sound speed profile. This is because refraction effects become more pronounced at greater ranges, leading to greater error in the time delay prediction, and therefore decreased ranging accuracy. If uncertainty in the sound velocity profile is included in the Fisher information matrix, this would lead to a larger Bayesian CRLB. One possible method of evaluating this might be to replace the analytic partial derivatives with numeric methods, but this approach may require substantially greater computational resources.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support for this research by the Office of Naval Research Ocean Acoustics Program. J.S.A. acknowledges support from the Department of Homeland Security under award 2008-ST-061-ML0001. We thank Portland State University for helping fund the equipment. We also thank Tom Monroe of High Tech, Inc., Reid McCargar, Grant Pusey, and Bill Stevens of Metron, Inc.

APPENDIX: STRIATION EXTRACTION PARTICLE FILTERING

This appendix describes the sequential Bayesian filtering methodology used to extract striation lines from a correlogram. The striation lines correspond to correlations of different pairs of arrivals at opposite hydrophones. As can be seen in the correlogram in Fig. 5(b), a prominent center striation line is flanked by two weaker striations. The center striation is caused by the direct arrival correlating with the direct arrival at the opposite hydrophone, and the flanking striations are due to a direct arrival at one hydrophone correlating with a multipath arrival at the opposite hydrophone. The offset of all three striations on the time delay (horizontal) axis at each time step (vertical axis) is the desired output. The extraction is divided into two trackers; the first tracks the center striation, which is then fed to the second tracker that tracks both flanking striations.

Given their ability to handle nonlinear problems, particle filters have been applied previously to the problem of acoustically tracking objects in space using multipath (Ward *et al.,* 2003), and to directly analyze features of received data (Jain and Michalopoulou, 2011; Michalopoulou and Jain, 2012). The approach taken here is more akin to the latter by focusing on extracting time delays from the passive acoustic data. Sequential Importance Resampling (SIR) is among the simplest of the particle filter formulations (Ristic *et al.,* 2004). In the tracking context, a sequence of observations are fed to the tracker to estimate hidden state variables at discrete time steps, denoted by k. State estimates at the

previous time step, $k-1$, are used to refine the probability at the current time step, giving this formulation its "Bayesian" nature.

In this context, the observations are the cross correlation at each time step, $\mathbf{\Gamma}_k = \Gamma_k(\tau)$ [Eq. (14) shows a noise-free measurement], and the state variables are the time delays, $\boldsymbol{\tau}_k$. Bold typeface indicates variables may be multidimensional. The prediction of the current state's ($\boldsymbol{\tau}_k$) probability distribution from a previous state's ($\boldsymbol{\tau}_{k\text{-}1}$) probability distribution is computed according to

$$p(\boldsymbol{\tau}_k|\mathbf{\Gamma}_{1:k-1}) = \int p(\boldsymbol{\tau}_k|\boldsymbol{\tau}_{k-1}) p(\boldsymbol{\tau}_{k-1}|\mathbf{\Gamma}_{1:k-1}) \mathrm{d}\boldsymbol{\tau}_{k-1} \tag{A1}$$

in which $\mathbf{\Gamma}_{1:k\text{-}1} = \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \ldots, \mathbf{\Gamma}_{k\text{-}1}$. Conceptually, this is based on a Markov state-space model in which the probability of the next state depends only on the previous state,

$$p(\boldsymbol{\tau}_k|\boldsymbol{\tau}_{1:k-1}) = p(\boldsymbol{\tau}_k|\boldsymbol{\tau}_{k-1}). \tag{A2}$$

To incorporate a new measurement $\mathbf{\Gamma}_k$, the state is updated as

$$p(\boldsymbol{\tau}_k|\mathbf{\Gamma}_{1:k}) \propto p(\mathbf{\Gamma}_k|\boldsymbol{\tau}_k) p(\boldsymbol{\tau}_k|\mathbf{\Gamma}_{1:k-1}), \tag{A3}$$

for which the right hand side is then normalized to integrate to one. It is useful to think of $p(\mathbf{\Gamma}_k \mathbf{\Gamma} \boldsymbol{\tau}_k)$ as a function of $\boldsymbol{\tau}_k$ that is parameterized by an actual (i.e., non-random) observation $\mathbf{\Gamma}_k$. Due to the proportionality relationship in Eq. (A3), this expression can be represented as a likelihood function that need not integrate to one

$$L(\boldsymbol{\tau}_k|\mathbf{\Gamma}_k) \propto p(\mathbf{\Gamma}_k|\boldsymbol{\tau}_k). \tag{A4}$$

Particle filters approximate the probability distributions as a sum of weighted delta functions,

$$p(\boldsymbol{\tau}_k|\mathbf{\Gamma}_{1:k}) \approx \sum_{i=1}^{N} w^i \delta(\boldsymbol{\tau}_k - \boldsymbol{\tau}_k^i), \tag{A5}$$

for which $\boldsymbol{\tau}_k^i$ is the ith particle and w^i is its weight, and N is the number of particles. The weights collectively sum to one, and provide a means to draw samples from a distribution other than $p(\boldsymbol{\tau}_k|\mathbf{\Gamma}_k)$, for which there is often no closed form solution. This other distribution is called the importance density, and in the SIR algorithm is defined as the posterior probability distribution at the previous time step, $p(\boldsymbol{\tau}_{k\text{-}1}|\mathbf{\Gamma}_{1:k\text{-}1})$.

Particle degeneracy happens when a significant portion of the particle weights become very small. These particles are essentially lost as they do not contribute meaningfully to the probability distribution. Resampling is used to re-draw the samples from regions of state space having greater probability density. This method follows from Ristic *et al.* (2004).

The first particle filter extracts the center striation using a random-walk state model defined as $\tau_{0_k} = \tau_{0_{k-1}} + \upsilon$, in which $\upsilon \sim \mathcal{N}(0, \sigma_{\tau_0}^2)$. It utilizes the one Gaussian from Eq. (17) that corresponds to the center striation. The log-likelihood function is thus defined as

$$\log L(\tau_{0_k}|\mathbf{\Gamma}_k) = \int_{-\infty}^{\infty} \Gamma_k(\tau) g(\tau, \tau_{0_k}, \sigma) \mathrm{d}\tau. \tag{A6}$$

Taking advantage of the fact that the center striation is often the strongest, all the particles are initialized at the first time step to the offset of the tallest peak in the first observation.

The second particle filter outputs τ_{-_k} and τ_{+_k} by making use of the two remaining Gaussians in Eq. (17) that correspond to the flanking striations. Tracking the flanking striations is a different problem than tracking the center striation because the multipath physics cause them to behave differently. The center striation is based primarily on the target bearing, but the flanking striations are highly dependent on each other and on the center striation. Specifically, the flanking striations are roughly parallel to each other and to the center striation, and each is offset from the center striation by opposite but roughly equal amounts. To exploit this behavior, the estimate of τ_{0_k} that is output by the first tracker is taken as an input to the second tracker. Instead of tracking τ_{-_k} and τ_{+_k} directly, the distance between the center and left striation is tracked with one auxiliary variable (ζ_1) and the additional amount of distance on the opposite side of the center striation is tracked with another (ζ_2).

The output at time k is then

$$\tau_{-_k} = \tau_{0_k} - \zeta_{1_k}, \qquad \tau_{+_k} = \tau_{0_k} + \zeta_{1_k} + \zeta_{2_k}. \tag{A7}$$

The advantage of this construction is it enforces the behavioral pattern while minimizing the tracker complexity by employing a simple random-walk model for ζ_1 and ζ_2. Note that the quantity ζ_{2_k} captures the slight multipath path differences between each hydrophone, so it essentially contains the information that allows for left – right disambiguation. The state update is $[\zeta_{n_k} = \zeta_{n_{k-1}} + w_n]_{n\in\{1,2\}}$, in which $[w_n \sim \mathcal{N}(0, \sigma_{\zeta_n}^2)]_{n\in\{1,2\}}$ such that $\sigma_{\zeta_2}^2 < \sigma_{\zeta_1}^2$. The log-likelihood function is defined as

$$\log L(\tau_{-_k}, \tau_{+_k}|\mathbf{\Gamma}_k, \tau_{0_k}) = \int_{-\infty}^{\infty} \Gamma_k(\tau)[g(\tau, \tau_{-_k}, \sigma) + g(\tau, \tau_{+_k}, \sigma)]\mathrm{d}\tau. \tag{A8}$$

Particles in this tracker are initialized to be uniformly distributed within the possible range of multipath delay values thereby not making any assumptions about the starting range of the target. The upper bound of ζ_1 occurs when the target is directly overhead and the lower bound is based on the target bandwidth [see Eq. (15)]. The lower bound of ζ_2 is zero, and the upper bound needs to be large enough to capture the possible difference between the depths of the two seabed reflection points, and is generally smaller than ζ_1.

To obtain point estimates from each posterior probability distribution, a Gaussian kernel smoother is applied to the particles to estimate the maximum *a priori* (MAP) value. The output of the tracker after both filters are the measured values τ_{0_k}, τ_{-k}, and τ_{+k} at each time interval k. In Sec. IV, the parameters were set to $N = 500$, $\sigma_{\tau_0} = 0.413$ ms, $\sigma_{\zeta_1} = 0.098$ ms, and $\sigma_{\zeta_2} = 0.001$ ms.

Aubauer, R., Lammers, M. O., and Au, W. W. L. (**2000**). "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," J. Acoust. Soc. Am. **107**, 2744–2749.
Bruno, M., Chung, K. W., Salloum, H., Sedunov, A., Sedunov, N., Sutin, A., Graber, H., and Mallas, P. (**2010**). "Concurrent use of satellite imaging and passive acoustics for maritime domain awareness," in *International Waterside Security Conference (WSS)*, pp. 1–8.
Carter, G., and Knapp, C. (**1976**). "Time delay estimation," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP1976*, Volume 1, pp. 357–360.
Carter, G. C., Nuttall, A. H., and Yuen, C. K. (**1980**). "On the weighted overlapped segment averaging method for power spectral estimation," Proc. IEEE **68**, 1352–1354.
de Berg, M. (**2008**). "Delaunay triangulations," in *Computational Geometry Algorithms and Applications*, 3rd ed. (Springer, Berlin), Chap. 9, pp. 191–218.
Friedlander, B. (**1988**). "Accuracy of source localization using multipath delays," IEEE Trans. Aerospace Electron. Syst. **24**, 346–359.
Gebbie, J., Siderius, M., and Allen, J. S. (**2011**). "Passive acoustic array harbor security applications," Mar. Technol. Soc. J. **45**, 103–110.
Gebbie, J., Siderius, M., McCargar, R., Allen, J. S., and Pusey, G. (**2013**). "Localization of a noisy broadband surface target using time differences of multipath arrivals," J. Acoust. Soc. Am. **134**, EL77–EL83.
Hahn, W., and Tretter, S. (**1973**). "Optimum processing for delay-vector estimation in passive signal arrays," IEEE Trans. Inf. Theory **19**, 608–614.
Hamilton, M., and Schultheiss, P. M. (**1993**). "Passive ranging in multipath dominant environments: Part II-unknown multipath parameters," IEEE Trans. Signal Process. **41**, 1–12.
Jain, R., and Michalopoulou, Z.-H. (**2011**). "A particle filtering approach for spatial arrival time tracking in ocean acoustics," J. Acoust. Soc. Am. **129**, EL236–EL241.
Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (**2011**). "Ray methods," in *Computational Ocean Acoustics*, 2nd ed., edited by W. M. Hartmann (Springer, New York), Chap. 3, pp. 115–232.
Knapp, C., and Carter, G. (**1976**). "The generalized correlation method for estimation of time delay," IEEE Trans. Acoust., Speech Signal Process. **24**, 320–327.
MathWorks (**2013**). *Symbolic Math Toolbox version 5.11*, The MathWorks Inc., Natick, MA, matlab r2013b v8.2.0.701 edition.
Michalopoulou, Z.-H., and Jain, R. (**2012**). "Particle filtering for arrival time tracking in space and source localization," J. Acoust. Soc. Am. **132**, 3041–3052.
Nosal, E., and Frazer, L. N. (**2006**). "Track of a sperm whale from delays between direct and surface-reflected clicks," Appl. Acoust. **67**, 1187–1201.
Ristic, B., Arulampalam, S., and Gordon, N. (**2004**). "A tutorial on particle filters," in *Beyond the Kalman Filter: Particle Filters for Tracking Applications* (Artech House, Boston), Chap. 3, pp. 35–66.
Thode, A. (**2004**). "Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array," J. Acoust. Soc. Am. **116**, 245–253.
Tiemann, C. O., Thode, A. M., Straley, J., O'Connell, V., and Folkert, K. (**2006**). "Three-dimensional localization of sperm whales using a single hydrophone," J. Acoust. Soc. Am. **120**, 2355–2365.
University of Hawai'i at Manoa, Costal Geology Group (**2012**). *SHOALS LIDAR Bathymetry Database*, www.soest.hawaii.edu (Last viewed April 23, 2014).
Van Trees, H. L. (**1968**). "Multiple parameter estimation," in *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York), Chap. 2.4.3, pp. 74–85.
Ward, D. B., Lehmann, E. A., and Williamson, R. C. (**2003**). "Particle filtering algorithms for tracking an acoustic source in a reverberant environment," IEEE Trans. Speech Audio Process. **11**, 826–836.
Weirathmueller, M. J., Wilcock, W. S. D., and Soule, D. C. (**2013**). "Source levels of fin whale 20 Hz pulses measured in the northeast pacific ocean," J. Acoust. Soc. Am. **133**, 741–749.
Whittle, P. (**1953**). "The analysis of multiple stationary time series," J. R. Statist. Soc. Ser. B (Methodological) **15**, 125–139.
Zelen, M., and Severo, N. C. (**1972**). "Probability functions," in *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed., edited by M. Abramowitz and I. A. Stegun (U.S. Government Printing Office, Washington, DC), Vol. 55, pp. 925–964.