

# Localizing Snapping Shrimp Noise Using a Small-Aperture Array

Yuen Min Too, *Member, IEEE*, Mandar Chitre, *Senior Member, IEEE*, George Barbastathis, *Member, IEEE*, and Venugopalan Pallayil, *Senior Member, IEEE*

**Abstract**—In warm shallow waters, one can consistently hear the crackling sound due to snapping shrimp. An individual sound or “snap” is highly impulsive. Knowing the source locations of these snaps facilitates the development of some unique passive sensing applications. One common approach in localizing these snaps is to calculate the time difference of arrival (TDoA) of snaps across sensors of an array. For a snap originating from a short distance away, in terms of a few multiples of the array aperture size, the signal from the snap forms a curved wavefront, such that both the direction of arrival (DoA) and the range of the snap can be obtained by studying the TDoA of the wavefront. However, this method is infeasible for long-range 3-D localization due to the need for large-aperture 3-D arrays. Smaller aperture arrays allow estimation of DoA for far-field sounds, but not the range. Since the ocean surface acts like an acoustic mirror that reflects the snaps, we can estimate the range of the snaps by measuring the TDoA between the direct arrivals and their surface reflections, even if the snaps occur in the far-field of the array. Given a set of perfectly associated direct and surface-reflected snaps, we derive a range estimator parameterized by nominal receiver depth and receiver orientation. In practice, we know these parameters only approximately. The problem is further complicated by the fact that snaps are acoustically similar to each other, and so associating a snap with its surface reflection can be difficult. We propose an algorithm that solves the problem of estimation of snap locations, snap associations, and receiver parameters, jointly. We verify the method through numerical simulation and through experimental results.

**Index Terms**—Association and estimation, geometric model, localization, optimization, snapping shrimp.

## I. INTRODUCTION

**S**NAPPING shrimp, also known as pistol shrimp, belong to the Alpheidae family of crustaceans. Snapping shrimp thrive in shallow waters less than 60 m in depth, and in

Manuscript received January 10, 2017; revised August 20, 2017; accepted November 19, 2017. This work was supported by the National Research Foundation (NRF), Prime Minister’s Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) program. The Center for Environmental Sensing and Modeling (CENSAM) is an interdisciplinary research group (IRG) of the Singapore MIT Alliance for the Research and Technology (SMART) center. (*Corresponding author: Yuen Min Too.*)

**Associate Editor: K. Foote.**

Y. M. Too, M. Chitre, and V. Pallayil are with the Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore, Singapore 119222 (e-mail: ymtoo@arl.nus.edu.sg; mandar@arl.nus.edu.sg; venu@arl.nus.edu.sg).

G. Barbastathis is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: gbarb@mit.edu).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JOE.2017.2777718

water temperatures warmer than 11° [1]. These creatures form colonies in debris, coral, and man-made underwater structures. They create episodic transient impulsive signals, known as snaps, via the collapse of cavitation bubbles generated by rapid closure of their claws [2]. Millions of snaps are heard as a persistent background crackle, and can be a nuisance to sonar and underwater acoustic communication systems [3]–[5]. On the other hand, there is evidence that snapping shrimp noise can provide natural insonification for passive sensing. Snaps reflected by an otherwise silent underwater object can be used to image that object [6]. Even passive ranging of silent submerged objects is feasible given the rough location of the shrimp [7]. Coral reefs are home to a multitude of living creatures some of which are extremely noisy like the snapping shrimp. Studies have suggested that by observing the spatial and temporal distribution of the sounds of the reef, we can obtain ecological information of coral reefs using either single sensor [8], [9] or array of sensors [10], [11]. Images revealing the form of submerged structures can be obtained based on the noise generated by the shrimp inhabiting these structures [12], [13]. A snap reflected off the seabed contains information about the sediment. Listening to a large number of snaps therefore allows us to estimate seabed sediment properties [14]. All of these applications rely on the ability to accurately estimate the location of individual snaps.

The location of a snap is defined by its direction of arrival (DoA) and range. Consider a passive receiver in the form of an array of sensors. The wavefront curvature method is widely adopted to estimate the location of snaps by measuring the time difference of arrival (TDoA) of snaps across sensors. A three-sensor uniform linear array with interelement spacing of 9.7 m was deployed in Sydney Harbour [12]. This experiment demonstrates the reliability of passive localization of snapping shrimp noise by showing that the predicted spatial distribution of snapping shrimp agrees with the physical structure of the wharf. A bilinear array of 6-m diameter was able to localize biological sources, including snapping shrimp, on or inside the reef structure [11]. The primary limitation of the wavefront curvature method is that the sources have to be in the near-field of the receiver which restricts the range of the sources to within approximately 10× the aperture size of the receiver. Although large-aperture receivers have been used to locate shrimp over small ranges, deploying an extremely large 2-D receiver is not practical in shallow waters. This limits the ability to locate shrimp in three dimensions.

Small-aperture arrays can only estimate the DoA but not the range of the snaps in the far-field. Since the ocean surface acts like an acoustic mirror, we can estimate the range of the snaps by measuring the TDoA between the direct arrival and surface-reflected snaps. This idea is based on creating a 2-D geometric model of the associated multipath arrivals of the snaps, and has been experimentally demonstrated previously [13]. However, the model relies on two key assumptions: the water surface must be completely flat, and the orientation of the receiver must be accurately known with respect to the water surface. Some of the arrivals may not completely fit into this simple geometric model. Other passive localization attempts have also been proposed. For instance, colonies of shrimp can be localized using triangulation from a moving small-aperture receiver [15]. While this may be feasible in some scenarios, it requires a mobile platform, long-duration measurements, and the noise environment to be unchanged over the measurement duration. Another approach is to assume that snapping shrimp live on a flat seabed [7]. This may not be generally valid, and moreover, an exact knowledge of the local bathymetry is not always available. Quite a number of encouraging results have been obtained from localizing cetaceans using the direct and surface-reflected signals [16]–[18]. Only a small number of signals are observed during the observation period and hence the direct and surface-reflected signals can be easily identified. This is not true for snapping shrimp noise as the receiver observes a large number of snaps in a short period of time and all the arrivals are acoustically the same.

In this study, we formulate a more general 3-D ray model in Section II. We construct a range estimator for each snap, parameterized by the nominal receiver depth and the receiver orientation. We discuss the sensitivity of these parameters to the estimated range in Section III. Given that we do not have the perfect knowledge of the parameters or the association of direct and surface-reflected snaps, we introduce an algorithm to solve the practical localization problem in Sections IV and V. Subsequently, experimental results are shown to verify the algorithm in Section VI.

Bold lower case symbols denote vectors. Bold upper case symbols denote matrices. We define  $\mathbb{U}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}| = \hat{\mathbf{x}}$  as an operator generating a unit vector from vector  $\mathbf{x}$ . The cardinality of a set  $\mathcal{S}$  is denoted as  $|\mathcal{S}|$ . Superscript  $\mathsf{T}$  denotes the transpose.

## II. THREE-DIMENSIONAL GEOMETRIC MODEL

An array of receiver observes the direct arrival and reflections of each snap in its field of view. We assume an isovelocity shallow water channel and a geometric ray model, which is valid for the high-frequency acoustic signal such as the snaps [1], [10]. Since the water-to-air interface reflects an acoustic wave almost perfectly [19], only the reflection at the water surface is considered in our model. The bottom reflection is ignored due to the nearly identical time of arrival (ToA) between direct arrival and bottom reflection as snapping shrimp often live close to the seabed. Given an acoustic pressure recording of the receiver, we can detect the direction of arrival and the time of arrival (DoA-ToA) of the impulsive transient signals such as the snaps based on the methods used in [7], [11], [12], [20]–[22]. By

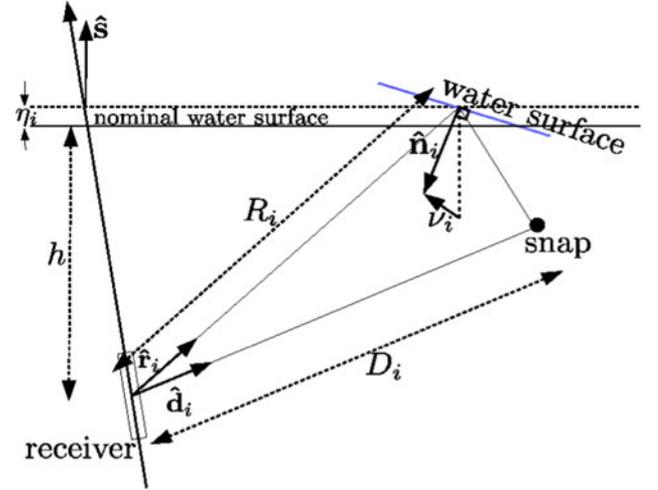


Fig. 1. Three-dimensional geometric illustration of a direct and surface-reflected arrival pair.  $\hat{\mathbf{d}}_i$  is a unit vector of the direct arrival.  $\hat{\mathbf{r}}_i$  is a unit vector of the surface-reflected arrival.  $D_i$  is distance to the location of snap.  $R_i$  is distance to the point of reflection of the snap on the ocean surface.  $h$  is nominal depth of the receiver.  $\eta_i$  is depth deviation from nominal value due to wave motion.  $\hat{\mathbf{n}}_i$  is a unit vector normal to the surface (pointing downward) at the point of reflection of snap.  $\hat{\mathbf{s}}_i$  is orientation of the receiver, and  $\mathbf{v}_i$  refers to local orientation of the sea surface of snap. The subscripts of the variables denote a particular snap indexed by  $i$ .

formulating the detection problem as an inverse problem of the underdetermined linear system with sparsity constraint, the technique described in [22] is able to perform high-resolution detection using small-aperture array, hence it is used throughout the study. We initially assume that we can perfectly associate direct snaps and their surface-reflected arrivals; this assumption will be relaxed later. We set up a right-handed coordinate system with its origin at the acoustic center of the receiver and the  $x$ -axis pointing along the broadside direction. Let  $\hat{\mathbf{d}}_i$  and  $\hat{\mathbf{r}}_i$  be the unit vectors of the direct and reflected pair  $i$ , respectively. The unit vectors depict the reversed propagating directions of the direct and reflected arrivals. Let  $\tau_i$  be the time difference between the direct and reflected arrivals. If  $c$  denotes the speed of sound,  $\delta_i = c\tau_i$  is the difference in path length of the direct and reflected paths. We consider  $\hat{\mathbf{d}}_i$ ,  $\hat{\mathbf{r}}_i$ , and  $\delta_i$  to be the measured quantities for each snap  $i$ .

Let  $D_i$  be the distance to the location of snap  $i$ , i.e.,  $D_i\hat{\mathbf{d}}_i$  is the position vector of that snap. Let  $R_i\hat{\mathbf{r}}_i$  be the position vector of the point of reflection in the ocean surface, and  $\hat{\mathbf{n}}_i$  be the unit vector normal to the surface (pointing downward) at that point. We assume the undisturbed ocean surface to be a plane given by the equation  $\mathbf{x}^T\hat{\mathbf{s}} = h + \eta_i$  where  $\mathbf{x}$  is any point on the surface,  $\hat{\mathbf{s}}$  is normal to the surface (pointing upwards),  $h$  is the nominal depth of the receiver, and  $\eta_i$  is the depth deviation from nominal value due to wave motion. Since the yaw of the receiver does not affect the propagating path of the direct and surface-reflected snap on the same azimuth plane,  $\hat{\mathbf{s}}$  effectively captures the exact orientation of the receiver on the seafloor. We assume that  $h$  and  $\hat{\mathbf{s}}$  do not change over the observation period, but are unknown. In practice, we may know them approximately. Fig. 1 illustrates the geometry of a direct and surface-reflected arrival pair.

The path length difference  $\delta_i$  is given by

$$\delta_i = R_i + |D_i\hat{\mathbf{d}}_i - R_i\hat{\mathbf{r}}_i| - D_i. \quad (1)$$

Rearranging and squaring both sides, we get

$$R_i = \frac{\delta_i^2/2 + D_i \delta_i}{D_i + \delta_i - D_i \hat{\mathbf{d}}_i^T \hat{\mathbf{r}}_i}. \quad (2)$$

Since the point of reflection must lie on the ocean surface, the perturbed nominal receiver depth can be written as

$$R_i \hat{\mathbf{r}}_i^T \hat{\mathbf{s}} = h + \eta_i \quad (3)$$

which is the projection of  $R_i \hat{\mathbf{r}}$  onto  $\hat{\mathbf{s}}$ . The normal vector of the local water surface  $\hat{\mathbf{n}}_i$ , which may not equal to  $\hat{\mathbf{s}}$ , is given by

$$\hat{\mathbf{n}}_i = \mathbb{U}(\mathbb{U}(D_i \hat{\mathbf{d}}_i - R_i \hat{\mathbf{r}}_i) - \hat{\mathbf{r}}_i). \quad (4)$$

Since  $\hat{\mathbf{r}}_i$ ,  $\hat{\mathbf{n}}_i$ , and  $\hat{\mathbf{v}}_i = \mathbb{U}(D_i \hat{\mathbf{d}}_i - R_i \hat{\mathbf{r}}_i)$  are coplanar, there exists  $\beta_1$  and  $\beta_2$  such that  $\hat{\mathbf{n}}_i = \beta_1 \hat{\mathbf{v}}_i - \beta_2 \hat{\mathbf{r}}_i$ . Law of reflection ensures  $-\hat{\mathbf{r}}_i^T \hat{\mathbf{n}}_i = \hat{\mathbf{v}}_i^T \hat{\mathbf{n}}_i$  so that  $-\beta_1 \hat{\mathbf{r}}_i^T \hat{\mathbf{v}}_i + \beta_2 = \beta_1 - \beta_2 \hat{\mathbf{r}}_i^T \hat{\mathbf{v}}_i$ , which leads to  $\beta_1 = \beta_2$ . Since  $\hat{\mathbf{n}}_i = \beta_1 (\hat{\mathbf{v}}_i - \hat{\mathbf{r}}_i)$  and  $\hat{\mathbf{n}}_i$  is a unit vector,  $\beta_1 = 1/|\hat{\mathbf{v}}_i - \hat{\mathbf{r}}_i|$  and, thus,  $\hat{\mathbf{n}}_i = \mathbb{U}(\hat{\mathbf{v}}_i - \hat{\mathbf{r}}_i)$ . This confirms the correctness of (4).

In the case of a calm unperturbed water surface,  $\hat{\mathbf{n}}_i = -\hat{\mathbf{s}}$  and  $\eta_i = 0$ . In the presence of waves,  $\hat{\mathbf{n}}_i = -\hat{\mathbf{s}} + \mathbf{v}_i$  where  $\mathbf{v}_i$  indicates the local roughness of the sea surface. Substituting (4), we get

$$\mathbb{U}(\mathbb{U}(D_i \hat{\mathbf{d}}_i - R_i \hat{\mathbf{r}}_i) - \hat{\mathbf{r}}_i) = -\hat{\mathbf{s}} + \mathbf{v}_i. \quad (5)$$

The left-hand side of (3) and (5) depicts the indirect noisy measurements of  $h$  and  $\hat{\mathbf{s}}$ , respectively. If  $h$  and  $\hat{\mathbf{s}}$  are known and  $\eta_i$  is ignored, the range of snap  $i$  can be estimated based on

$$\tilde{D}_i = \frac{\delta_i^2/2 - \frac{h}{\hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} \delta_i}{\frac{h}{\hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} - \frac{h}{\hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} \hat{\mathbf{d}}_i^T \hat{\mathbf{r}}_i - \delta_i} \quad (6)$$

where tilde denotes the estimate. The estimated location of snap  $i$  is

$$\tilde{\Lambda}_i = \tilde{D}_i (\mathbf{T} \hat{\mathbf{d}}_i) \quad (7)$$

where  $\mathbf{T}$  is the transformation of the receiver axes to the nominal water surface axes. Since  $\hat{\mathbf{s}} = [\sin(\alpha^\circ), \sin(\rho^\circ) \cos(\alpha^\circ), \cos(\rho^\circ) \cos(\alpha^\circ)]^T$  where  $\alpha^\circ$  and  $\rho^\circ$  are the small angle of pitch rotation and roll rotation of the normal vector of nominal water surface, we can approximate the transformation matrix as

$$\mathbf{T} \approx \begin{bmatrix} \cos(-\alpha^\circ) & 0 & \sin(-\alpha^\circ) \\ 0 & 1 & 0 \\ -\sin(-\alpha^\circ) & 0 & \cos(-\alpha^\circ) \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\rho^\circ) & -\sin(\rho^\circ) \\ 0 & \sin(\rho^\circ) & \cos(\rho^\circ) \end{bmatrix} \quad (8)$$

where the first matrix describes the pitch at the  $y$ -axis and second matrix is the roll at the  $x$ -axis according to the right-handed rule. For convenience, the range estimator is developed based on the assumption of the unperturbed nominal receiver depth  $\eta_i = 0$  for any  $i$ . In reality,  $\eta_i$  is nonzero and contributes to the range estimation error, of which the first-order approximation can be

written as

$$D_i - \tilde{D}_i \approx \frac{\delta_i^2 + \delta_i^2 \hat{\mathbf{d}}_i^T \hat{\mathbf{r}}_i}{2 \left( \delta_i - \frac{h}{\hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} + \frac{h}{\hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} \hat{\mathbf{d}}_i^T \hat{\mathbf{r}}_i \right)^2 \hat{\mathbf{r}}_i^T \hat{\mathbf{s}}} \eta_i. \quad (9)$$

Let  $\eta_i$  be an independent identically distributed (IID) zero-mean random variable for all  $i$ . If we know the originating sources of the snaps, we can improve the estimated range by calculating the mean of an ensemble of estimated range values from the same snap source. This exact knowledge is not known in practice, and it can be approximated by accumulating the snaps that have approximately the same DoA as the same snap source. This is because the DoA is 3-D, and that the snaps have a large amount of energy at high frequency, diffraction is limited, and hence a geometrical ray model with a structure occluding snaps from behind it is reasonable. Through the experimental data, the reasoning and verification of the improvement process is discussed in Section VI.

### III. SENSITIVITY ANALYSIS

For ease of notation, index  $i$  is omitted in this section. Let  $\hat{\mathbf{d}}$ ,  $\hat{\mathbf{r}}$ , and  $\delta$  be known exactly but not  $h$  and  $\hat{\mathbf{s}}$ . The range estimator in (6) is a function of the prior knowledge of the parameters such as

$$\tilde{D}(h + \epsilon_h, \hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}}) = \frac{\frac{\delta^2}{2} - \frac{h + \epsilon_h}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} \delta}{\frac{h + \epsilon_h}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} - \frac{h + \epsilon_h}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} \hat{\mathbf{d}}^T \hat{\mathbf{r}} - \delta} \quad (10)$$

where the *a priori* parameters ( $h + \epsilon_h$ ,  $\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}}$ ) are defined by the summation of the actual value of the parameters ( $h$ ,  $\hat{\mathbf{s}}$ ) and the errors ( $\epsilon_h$ ,  $\epsilon_{\hat{\mathbf{s}}}$ ).

To simplify the notation, we define  $\tilde{R} = h/(\hat{\mathbf{r}}^T \hat{\mathbf{s}})$ , which linearizes the parameter errors such that  $\tilde{R} + \epsilon_{\tilde{R}} = (h + \epsilon_h)/(\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}}))$ . Hence, we can write

$$\tilde{D}(\tilde{R} + \epsilon_{\tilde{R}}) = \frac{\frac{\delta^2}{2} - (\tilde{R} + \epsilon_{\tilde{R}}) \delta}{(\tilde{R} + \epsilon_{\tilde{R}}) - (\tilde{R} + \epsilon_{\tilde{R}}) \hat{\mathbf{d}}^T \hat{\mathbf{r}} - \delta} \quad (11)$$

where  $\epsilon_{\tilde{R}}$  is related to  $\epsilon_h$  and  $\epsilon_{\hat{\mathbf{s}}}$ . The range estimation error, which is the difference between the estimated range based on actual and estimated parameters, can be written as

$$\tilde{D}(\tilde{R} + \epsilon_{\tilde{R}}) - \tilde{D}(\tilde{R}) = \tilde{D}(\tilde{R}) \epsilon_{\tilde{R}} Q \quad (12)$$

where

$$Q = \frac{(1 + \hat{\mathbf{d}}^T \hat{\mathbf{r}}) \frac{\delta^2}{2}}{((1 - \hat{\mathbf{d}}^T \hat{\mathbf{r}}) (\tilde{R} + \epsilon_{\tilde{R}}) - \delta) (\frac{\delta^2}{2} - \tilde{R} \delta)}.$$

The magnitude of the estimation error turns out to be

$$|\tilde{D}(\tilde{R} + \epsilon_{\tilde{R}}) - \tilde{D}(\tilde{R})| \propto \tilde{D}(\tilde{R}). \quad (13)$$

We notice that the error is linearly proportional to the range of the snap.

We are also interested in identifying which parameter is more significant in generating larger range estimation error. The previous paragraph has shown the direct relation between  $|\epsilon_{\tilde{R}}|$  and the range estimation error. Thus, it is sufficient to just present the relation between  $|\epsilon_{\tilde{R}}|$  and the parameters. Let  $\epsilon_{\hat{\mathbf{s}}} = \mathbf{0}$ . We

have

$$|\epsilon_{\tilde{R}_1}| = \frac{1}{\hat{\mathbf{r}}^T \hat{\mathbf{s}}} |\epsilon_h|. \quad (14)$$

Let  $h = 0$ . We have

$$\begin{aligned} |\epsilon_{\tilde{R}_2}| &= \frac{h}{\hat{\mathbf{r}}^T \hat{\mathbf{s}}} \left( \frac{1}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} \right) |\hat{\mathbf{r}}^T \epsilon_{\hat{\mathbf{s}}}| \\ &\leq \frac{h}{\hat{\mathbf{r}}^T \hat{\mathbf{s}}} \left( \frac{1}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} \right) |\epsilon_{\hat{\mathbf{s}}}|. \end{aligned} \quad (15)$$

The upper bound is due to Cauchy–Schwarz inequality and  $|\hat{\mathbf{r}}| = 1$ . For comparison, we replace the magnitude of the errors by the same fractional perturbation, denoted by  $0 \leq e \leq 1$ , with respect to the magnitude of the parameters. This can be written as  $|\epsilon_h| = eh$  and  $|\epsilon_{\hat{\mathbf{s}}}| = e|\hat{\mathbf{s}}| = e$ . Equations (14) and (15), respectively, become

$$|\epsilon_{\tilde{R}_1}| = \frac{h}{\hat{\mathbf{r}}^T \hat{\mathbf{s}}} e \quad (16)$$

$$|\epsilon_{\tilde{R}_2}| \leq \frac{h}{\hat{\mathbf{r}}^T \hat{\mathbf{s}}} \left( \frac{1}{\hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}})} \right) e. \quad (17)$$

Snaps are located in far-field and the receiver orientation is close to  $[0, 0, 1]^T$ , and hence  $0 < \hat{\mathbf{r}}^T (\hat{\mathbf{s}} + \epsilon_{\hat{\mathbf{s}}}) < 1$ . When snaps are farther away, the value approaches 0, and when snaps are nearer, the value approaches 1. In general, the error of the receiver orientation is more significant than the error of the nominal receiver depth for range estimation error in the worst case scenario. For distant snaps, range estimation error due to inaccurate knowledge of receiver orientation is larger as compared to the error due to inaccurate knowledge of receiver depth.

#### IV. PARAMETER ESTIMATION

We showed that the estimated location of snap is a function of parameters  $h$  and  $\hat{\mathbf{s}}$ . However, by assuming prior knowledge of the parameters, we may not be able to obtain accurate range estimation. In this section, we propose a method to improve the estimation of the parameters, with the assumption that these parameters remain constant over the observation period. We assume that we know the associated direct and surface-reflected snaps so that we can obtain an estimator for the parameters. Through numerical simulations, we will show that the derived method is capable of estimating the parameters and, thus, a better estimation of the locations of snaps can be achieved.

##### A. Formulation

Estimating  $h$  and  $\hat{\mathbf{s}}$  from one associated direct and surface reflection arrival of snap is not possible since there are more unknowns than number of nonlinear equations. Let  $P$  be the number of associated direct and surface-reflected snaps collected during the observation period. The small variation over the nominal receiver depth  $h$  is represented by a set of IID random variables  $\{\eta_i\}_{i=1}^P$ , while the small variation in the water surface normal vector is represented by a set of IID random vectors  $\{\nu_i\}_{i=1}^P$ . Fig. 2 depicts the 3-D geometric model, showing multiple snaps reflected from different water surface conditions. Let  $\tilde{h}$  and  $\tilde{\mathbf{s}}$  be the estimates of  $h$  and  $\hat{\mathbf{s}}$ . Let  $\{D'_i\}_{i=1}^P$  be bounded

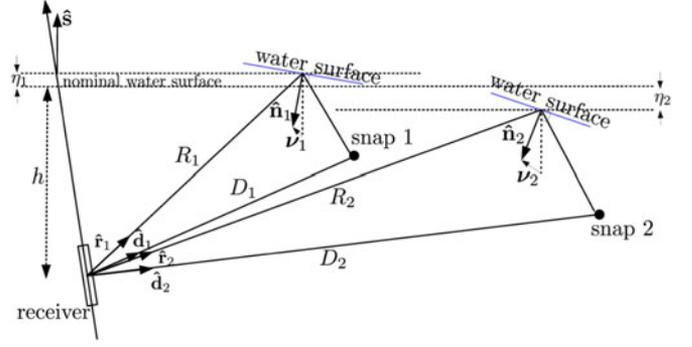


Fig. 2. Geometric illustration of two direct and surface-reflected snaps.

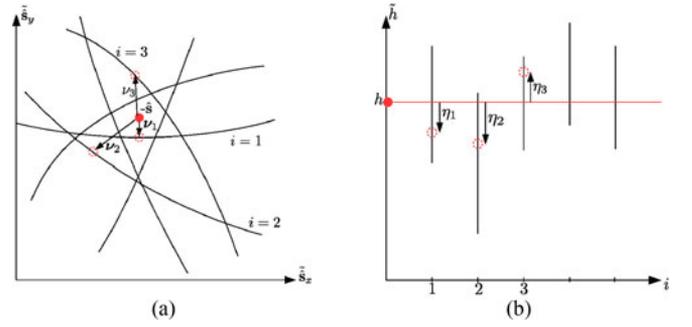


Fig. 3. Pictorial explanation of the effect of unknowns in (5) and (3). (a) A set of curves containing the shifted  $\hat{\mathbf{s}}$  (ordinate), denoted by dashed circles, is plotted based on the upper and lower bound of  $D_i$  (abscissa) for all  $i$ . Only variables in  $x$  and  $y$  axes of  $\hat{\mathbf{s}}$  are considered for the ease of visualization. (b) A set of vertical lines containing the shifted  $h$  (ordinate), denoted by dashed circles, is plotted based on the upper and lower bound of  $D_i$  (abscissa) for all  $i$  as well as the uncertainty in  $\hat{\mathbf{s}}$ .

by  $\{\{l_{D_i}, u_{D_i}\}_{i=1}^P\}$ , respectively, and considering  $R_i$  as a function of  $D_i$ , then (5) for  $i = 1, 2, \dots, P$  can be illustrated as a set of curves on a unit ball by varying  $\{D_i\}_{i=1}^P$ . All the curves should pass close to  $\hat{\mathbf{s}}$  since they are shifted by  $\{\nu_i\}_{i=1}^P$  from  $\hat{\mathbf{s}}$ . According to (3), varying  $\{D_i\}_{i=1}^P$  and  $\hat{\mathbf{s}}$  creates lines on  $h$  shifted by  $\{\eta_i\}_{i=1}^P$ . All the lines should lie near  $h$ . Fig. 3 gives a pictorial illustration of effect of unknowns in (5) and (3). Estimating  $h$  and  $\hat{\mathbf{s}}$  reduces to finding  $\tilde{h}$  and  $\tilde{\mathbf{s}}$  that is closest to all the curves and the lines. The closeness can be measured by the distance between the estimates and the points on the curves and the lines.

We can write this as an optimization problem

$$\begin{aligned} \tilde{h}, \tilde{\mathbf{s}} = \arg \min_{h', \hat{\mathbf{s}}': \|\hat{\mathbf{s}}'\|_2=1} \sum_{i=1}^P \min_{l_{D_i} \leq D'_i \leq u_{D_i}} \left( \|\mathbf{R}'_i \hat{\mathbf{r}}_i^T \hat{\mathbf{s}}' - h'\|_2^2 + \lambda \|\hat{\mathbf{s}}'\|_2^2 \right) \\ + \mathbb{U}(\mathbb{U}(D'_i \hat{\mathbf{d}}_i - \mathbf{R}'_i \hat{\mathbf{r}}_i) - \hat{\mathbf{r}}_i) \|^2 \end{aligned} \quad (18)$$

where  $D'_i$  defines a point on the respective curve and line which has the shortest Euclidean distance from  $h'$  and  $\hat{\mathbf{s}}'$ , and  $\mathbf{R}'_i$  is a function of  $D'_i$ .  $\lambda > 0$  is a tuning parameter controlling the relative importance of the closeness in the curves and closeness in the lines. Note that  $\{D'_i\}_{i=1}^P$  is not the estimate of  $\{D_i\}_{i=1}^P$ . It is used solely to describe the distribution of the curves and the lines.

Estimating  $h$  and  $\hat{\mathbf{s}}$  involves solving the nonconvex optimization problem in (18). Considering a smooth bathymetry and calm sea state, we can initialize  $h^{(0)}$  to the average nominal receiver depth at the receiver deployment location and  $\hat{\mathbf{s}}^{(0)} = [0, 0, 1]^T$ , both values of which should be close to the actual  $h$  and  $\hat{\mathbf{s}}$ . For  $k = 1$  iteration, the single-variable inner minimization of (18) can be efficiently solved for  $\{D_i^{(k)}\}_{i=1}^P$  given  $h^{(k-1)}$  and  $\hat{\mathbf{s}}^{(k-1)}$ . Subsequently, we solve the outer minimization of (18) for  $h^{(k)}$  and  $\hat{\mathbf{s}}^{(k)}$  given  $\{D_i^{(k)}\}_{i=1}^P$ . The two-step minimization is repeated for the next iteration until the objective function value converges. Let  $f$  be the objective function of (18). The decrease of the objective function value over iterations can be shown as follows:

$$\begin{aligned} f(D_i^{(k)}, h^{(k)}, \hat{\mathbf{s}}^{(k)}) &\leq f(D_i^{(k)}, h^{(k-1)}, \hat{\mathbf{s}}^{(k-1)}) \\ &\leq f(D_i^{(k-1)}, h^{(k-1)}, \hat{\mathbf{s}}^{(k-1)}) \end{aligned} \quad (19)$$

for  $k = 1, 2, \dots$ . By fixing  $D_i^{(k)}$ , minimizing  $f(D_i^{(k)}, h^{(k)}, \hat{\mathbf{s}}^{(k)})$  with respect to  $h^{(k)}$  and  $\hat{\mathbf{s}}^{(k)}$  is a convex problem, and the global minimum can be obtained, which leads to the first inequality in (19). By fixing  $h^{(k-1)}$  and  $\hat{\mathbf{s}}^{(k-1)}$ , minimizing  $f(D_i^{(k)}, h^{(k-1)}, \hat{\mathbf{s}}^{(k-1)})$  with respect to bounded  $D_i^{(k)}$  is a single-variable nonconvex optimization problem, and the global minimum can also be obtained which yields the second inequality in (19). This shows that locally optimal estimates can be achieved.

The objective function in (18) can be divided into two cost functions. The first cost is the distance between the estimate and the points on the lines, while the second cost is the distance between the estimate and the points on the curves. For  $\lambda \rightarrow 0$ , the first cost dominates and  $D_i^{(1)}$  can always be found such that the objective function value is close to zero regardless of  $h^{(0)}$  and  $\hat{\mathbf{s}}^{(0)}$  if the bounds  $l_{D_i}$  and  $u_{D_i}$  are not tight. The locally optimal estimates are simply the prior knowledge of the parameters. For  $\lambda \rightarrow \infty$ , the second cost dominates but it is less likely that we can find  $D_i'$  such that the objective function value is zero over the iterations. However, this might lead to overfitting in  $\hat{\mathbf{s}}$  and, consequently, yield large error in the estimate of  $h$ .

## B. Numerical Simulations

We verify the parameter estimation performance of the proposed method by performing two different simulations. The first simulation features longer range and randomly distributed snaps, while the second simulation has shorter range and structurally distributed snaps. About 2000 associated direct and surface-reflected snaps were generated based on  $\eta_i \sim \mathcal{N}(0, 0.04)$  in units of meter and

$$\begin{aligned} \mathbf{v}_i &= \hat{\mathbf{n}}_i + \hat{\mathbf{s}} \\ &= - \begin{bmatrix} \sin(\alpha_i + \alpha^o) \\ \sin(\rho_i + \rho^o) \cos(\alpha_i + \alpha^o) \\ \cos(\rho_i + \rho^o) \cos(\alpha_i + \alpha^o) \end{bmatrix} + \begin{bmatrix} \sin(\alpha^o) \\ \sin(\rho^o) \cos(\alpha^o) \\ \cos(\rho^o) \cos(\alpha^o) \end{bmatrix} \end{aligned} \quad (20)$$

for  $i = 1, 2, \dots, 1000$  where  $\alpha_i \sim \mathcal{N}(0, 25)$  and  $\rho_i \sim \mathcal{N}(0, 25)$  in units of degree are the pitch and roll of the local water surface,

respectively, whereas  $\alpha^o$  and  $\rho^o$  describe the orientation of the receiver. We set  $l_{D_i} = 0$  m and  $u_{D_i} = 300$  m for all  $i$ . The parameters were estimated using (18). We compared the estimated locations of snaps based on the actual and estimated parameters for comparison. The locations were modified such that the origin of the  $z$ -axis is set on the water surface for convenient illustration.

In simulation 1, we set  $h = 15$  m,  $\alpha^o = -7^\circ$ , and  $\rho^o = -5^\circ$ . The sources were uniformly distributed across  $[-60^\circ, +60^\circ]$  in azimuth angle,  $[-5^\circ, +10^\circ]$  in elevation angle, and  $[100$  m,  $200$  m] in range. Let the average nominal receiver depth be  $\tilde{h}^{(0)} = 14$  m. When  $\lambda = 0$ , we obtain  $\tilde{h} = h^{(0)}$  and  $\tilde{\mathbf{s}} = \hat{\mathbf{s}}^{(0)}$ , which yield estimated range error of approximately 100 m. This shows that using the prior knowledge parameters is not a reliable method for range estimation. The lowest estimation errors of  $\tilde{h}$  and  $\tilde{\mathbf{s}}$  are attained at  $\lambda = 5000$  and  $\lambda = 7000$ , respectively. The estimator with larger  $\lambda$  tends to overfit the curves which determine  $\hat{\mathbf{s}}$ . When  $\lambda$  is gradually reduced, more effort is given to minimizing the error in  $h$ . The calculated range of the snaps based on the estimated parameters is more accurate for large  $\lambda$  and better  $\hat{\mathbf{s}}$  estimate, as shown in Fig. 4(c) and (d). This agrees with the previous result stating that the error in  $\tilde{\mathbf{s}}$  is more significant than the error in  $\tilde{h}$  for snaps that are farther apart from the receiver.

In simulation 2, we examine the capability of the method to estimate the parameters given that snaps are structurally distributed in space. We set  $h = 5$  m,  $\alpha^o = 7^\circ$ , and  $\rho^o = -5^\circ$ . The snaps were uniformly generated within a rectangular space defined by  $[10$  m,  $20$  m] in the  $x$ -axis,  $[-20$  m,  $20$  m] in the  $y$ -axis,  $[-1$  m,  $0$  m] in the  $z$ -axis, and two vertical columns, both sharing the same interval  $[-1$  m,  $2$  m] in the  $z$ -axis but with one column at  $10$  m in the  $x$ -axis,  $[9$  m,  $10$  m] in the  $y$ -axis and the other column at  $[15$  m] in the  $x$ -axis,  $[-20$  m,  $-19$  m]. Let the average nominal receiver depth be  $\tilde{h}^{(0)} = 3$  m. The accuracy of  $\tilde{D}$  is less dependent on  $\tilde{\mathbf{s}}$  because this parameter becomes less significant when the range of the snap decreases. Similarly, there is a large estimation error for the range of the snaps parameterized by the prior knowledge of the parameters.

In short, a large  $\lambda$  (several thousands) seems to be reasonable for the two-step minimization. This is because the increment in the parameter estimation error is small for large  $\lambda$ . Even though a large  $\lambda$  does not produce optimal estimation for shorter range snaps such as those in simulation 2, the parameter error is relatively smaller than those using the underestimated  $\lambda$ .

## V. ASSOCIATION AND ESTIMATION

To determine the snapping shrimp locations from DoA-ToA estimates, there are two crucial pieces of information that need to be known in practice. One is the parameters of the range estimator like nominal receiver depth and receiver orientation. The other is the association of direct and surface-reflected snaps from multiple arrivals. If many snaps and reflections arrive at the array concurrently, associating a snap with its reflection can be a hard problem. Given the DoA-ToA of all arrivals, we discuss an algorithm to solve the problem of snapping shrimp noise localization. We present the idea starting with the coarse pairing

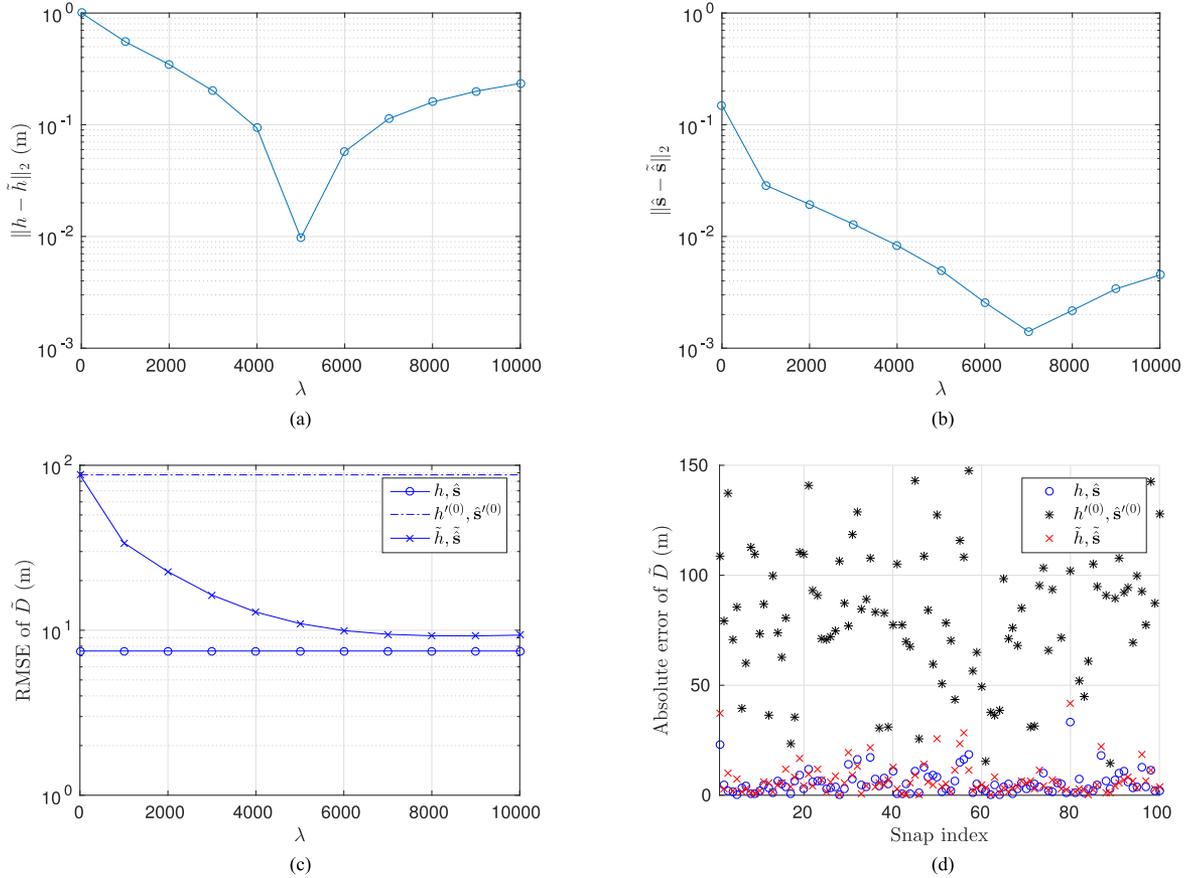


Fig. 4. Accuracy of  $\tilde{h}$  and  $\tilde{s}$  and the performance of the range estimator of snap using these parameters in simulation 1. Given  $\lambda$  (abscissa), we solve the optimization problem in (18) for  $\tilde{h}$  and  $\tilde{s}$  and then calculate the respective errors (ordinate) denoted by the y-axis in (a)–(c) plots. (a) Euclidean error of the estimated nominal receiver depth over different  $\lambda$ .  $h$  is 15 m. (b) Euclidean error of the estimated receiver orientation over different  $\lambda$ .  $\hat{s}$  is  $[-0.1219, -0.0865, 0.9888]^T$ . (c) Root-mean-square error (RMSE) of the estimated range of the snaps based on the actual, the prior knowledge, and estimated parameters over different  $\lambda$ . (d) Absolute error of the range based on the prior knowledge and the estimated parameters, which were calculated using  $\lambda = 8000$  as minimum RMSE of  $\tilde{D}$  is achieved. Only estimated range of 100 snaps is plotted for the ease of visualization.

procedure that eliminates obvious wrong associations among all the arrivals, followed by an algorithm to associate the arrivals and at the same time estimate the parameters.

#### A. Coarse Pairing

If  $N$  arrivals are observed, we can form, at most,  $N(N - 1) \approx N^2$  (for large  $N$ ) associated direct and surface-reflected snaps. For brevity, we refer to the associated direct and surface-reflected snaps simply as pairs. These pairs include a large proportion of wrong associations. In fact, by removing the number of obviously wrong pairs, we can reduce the number of pairings from  $N^2$  to  $P'$ . The  $P'$  pairs denoted by  $\{\hat{\mathbf{d}}_i, \hat{\mathbf{r}}_i, \delta_i\}_{i=1}^{P'}$ , where  $\hat{\mathbf{d}}_i = [\cos(\phi_i^d) \cos(\theta_i^d), \sin(\phi_i^d) \cos(\theta_i^d), \sin(\theta_i^d)]^T$  and  $\hat{\mathbf{r}}_i = [\cos(\phi_i^r) \cos(\theta_i^r), \sin(\phi_i^r) \cos(\theta_i^r), \sin(\theta_i^r)]^T$ , and where superscript d and r represent direct arrival and surface reflection, can be judiciously formed based on the physical properties of surface reflection of a calm water surface. Pairs must satisfy the following constraints: 1)  $|\phi_i^d - \phi_i^r| \leq \varepsilon_\phi$ ; 2)  $\theta_i^r > |\theta_i^d|$ ; 3)  $0 \leq \delta_i \leq 2h_u$ , for  $i = 1, 2, \dots, P'$ , where  $h_u$  is the maximum water depth. The first constraint indicates that the azimuth angle of the reflection has to be within a small deviation from the direct arrival of the snap. The second constraint requires the unit

vector of the reflection to be above the direct arrival. Finally, the difference in path length is a positive real value, bounded by constraint 3. The problem size has been extensively reduced to  $N$  arrivals and  $P'$  pairs where  $P' \ll N^2$ .

#### B. Joint Association and Estimation

Let the  $N$  arrivals be the vertices, and the  $P'$  pairs of the arrivals be the edges of a Graph. We define a weighted Graph with incidence matrix  $\mathbf{G}$ . The weight of the edge  $i$ , denoted by  $w_i$ , indicates the likeliness of the associated direct and surface-reflected pair  $i$ . Snaps have similar acoustic signature which is transient in time. Hence, identifying the pairs from multiple nearly identical and transient acoustic signals is nontrivial. We suggest that the weights be defined by the fitness of the pairs with respect to the geometric model. Any pair that conforms to the geometrical constraints will have a large weight and vice versa. However, the geometric model contains unknown parameters and we can only define the weight as a function of the parameters such as the nominal receiver depth and receiver orientation.  $\mathbf{G}$  is an  $N$  by  $P$  binary matrix with the “1” elements representing pairing on the vertices, which means each column of  $\mathbf{G}$  contains two “1” elements.

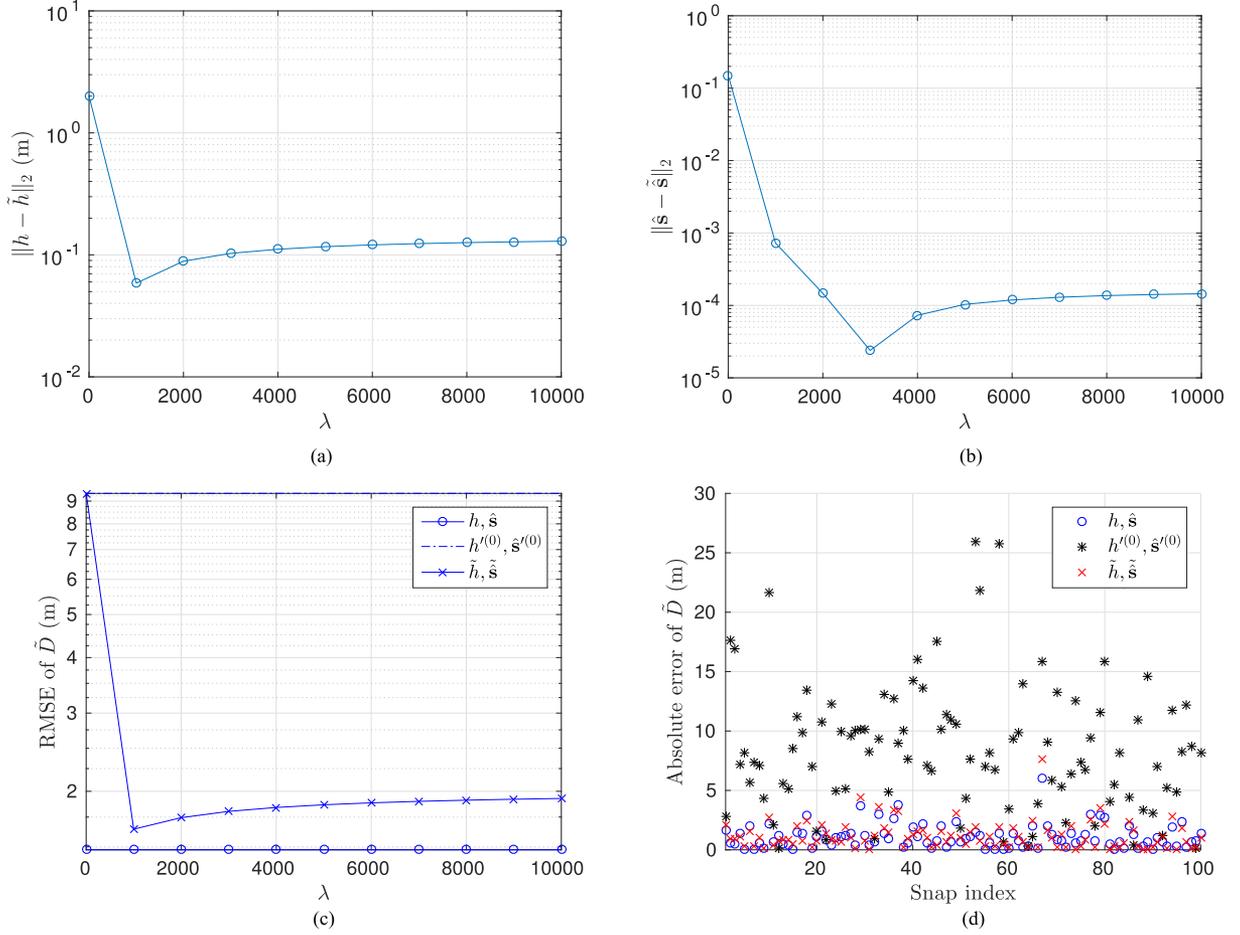


Fig. 5. Accuracy of  $\tilde{h}$  and  $\tilde{s}$  and the performance of the range estimator of snap using these parameters in simulation 2. Given  $\lambda$  (abscissa), we solve the optimization problem in (18) for  $\tilde{h}$  and  $\tilde{s}$  and then calculate the respective errors (ordinate) denoted by the y-axis in (a)–(c) plots. (a) Euclidean error of the estimated nominal receiver depth over different  $\lambda$ .  $h$  is 5 m. (b) Euclidean error of the estimated receiver orientation over different  $\lambda$ .  $\hat{s}$  is  $[0.1219, -0.0865, 0.9888]^T$ . (c) RMSE of the estimated range of the snaps based on the actual, the prior knowledge, and estimated parameters over different  $\lambda$ . (d) Absolute error of the range based on the prior knowledge and the estimated parameters which were calculated using  $\lambda = 1000$  as minimum RMSE of  $\tilde{D}$  is achieved. Only estimated range of 100 snaps are plotted for the ease of visualization.

The challenge of this problem is that the pairing depends on the unknown parameters but the parameter estimation requires a good set of pairs. To localize the snaps, we have to jointly associate the pairs and estimate the unknown parameter to find the set of pairs and parameters that maximizes the sum of the weights of edges of the Graph and minimize the objective function of (18). Formally, this can be written as

$$\begin{aligned} & \arg \max_{\substack{\mathbf{x} \in \{0,1\}^{P'} \\ \tilde{h}, \tilde{s}, \|\tilde{s}\|_2 = 1}} \mathbf{w}^T \mathbf{x} - \mu f(\mathbf{x}, \tilde{h}, \tilde{s}) \\ & \text{s.t. } \mathbf{G}\mathbf{x} \leq \mathbf{1} \\ & w_i = \begin{cases} \frac{1}{\|\hat{\mathbf{s}} + \hat{\mathbf{n}}_i(\tilde{h}, \tilde{s})\|_2^2}, & (\hat{\mathbf{n}}_i(\tilde{h}, \tilde{s}))_z \leq \kappa \\ 0, & \text{otherwise} \end{cases} \\ & \text{for } i = 1, 2, \dots, P' \end{aligned} \quad (21)$$

where  $\mathbf{x}$  is a  $P'$ -dimensional binary column vector with “1” elements indicating the existence of the pairs,  $\hat{\mathbf{n}}_i(\tilde{h}, \tilde{s}) = \mathbb{U}(\mathbb{U}(\tilde{D}_i(\tilde{h}, \tilde{s})\hat{\mathbf{d}}_i - \tilde{h}/(\hat{\mathbf{r}}_i^T \tilde{\mathbf{s}})\hat{\mathbf{r}}) - \hat{\mathbf{r}}_i)$  for  $\tilde{D}_i(\tilde{h}, \tilde{s})$  as given in (6)

parameterized by  $\tilde{h}$  and  $\tilde{s}$ .  $f(\mathbf{x}, \tilde{h}, \tilde{s})$  is the objective function of (18) with respect to variables  $\mathbf{x}$ ,  $\tilde{h}$ , and  $\tilde{s}$ .  $\mu > 0$  indicates the importance between arrival association and parameter estimation. The first constraint (inequality) of the optimization problem is that every arrival can only be associated once. The second constraint defines the weight as a function of the parameters. Let  $(\hat{\mathbf{n}}_i(\tilde{h}, \tilde{s}))_z$  be the  $z$ -axis element of the normal vector  $\hat{\mathbf{n}}_i(\tilde{h}, \tilde{s})$ .  $\kappa$  is a real value greater and close to  $-1$ . The hard threshold  $(\hat{\mathbf{n}}_i(\tilde{h}, \tilde{s}))_z \leq \kappa$  constrains the normal vector of the local water surface to point approximately downward, representing a calm sea state. The optimization problem is complicated and there is no obvious algorithm to solve it optimally. We next propose an algorithm to solve this problem approximately.

### C. Alternating Association and Estimation

Instead of solving the joint association and estimation problem, the location of the snaps can be estimated by alternately associating the pairs with fixed parameters and then estimating the parameters with fixed selected pairs. Given the prior

---

**Algorithm 1:** Alternating association and estimation for snap localization.

---

**Require:**  $\{(\hat{\mathbf{d}}_i, \hat{\mathbf{r}}_i, \delta_i)\}_{i=1}^{P'}$ ,  $\mathbf{G}$ ,

$\mathbf{x}^{(0)} \leftarrow \mathbf{0}$ ,

$\tilde{h}^{(0)} \leftarrow$  average water depth,

$\tilde{\mathbf{s}}^{(0)} \leftarrow [0, 0, 1]^\top$ ,

$k \leftarrow 0$ ,

maxIter  $\leftarrow$  maximum number of iterations,

$\epsilon_h \leftarrow$  positive small value,

$\epsilon_{\tilde{\mathbf{s}}} \leftarrow$  positive small value

1: Given

$$w_i = \begin{cases} \frac{1}{\|\tilde{\mathbf{s}}^{(k)} + \hat{\mathbf{n}}_i(\tilde{h}^{(k)}, \tilde{\mathbf{s}}^{(k)})\|_2^2}, & (\hat{\mathbf{n}}_i(\tilde{h}^{(k)}, \tilde{\mathbf{s}}^{(k)}))_z \leq \kappa \\ 0, & \text{otherwise} \end{cases} \quad \text{for}$$

$i = 1, 2, \dots, P'$ , solve

$$\arg \max_{\mathbf{x} \in \{0, 1\}^{P'}} \mathbf{w}^\top \mathbf{x} \quad \text{s.t. } \mathbf{G}\mathbf{x} \leq \mathbf{1}$$

for  $\mathbf{x}^{(k+1)}$ .

2: Given  $\mathbf{x}$ , we solve (18) for  $\tilde{h}^{(k+1)}$  and  $\tilde{\mathbf{s}}^{(k+1)}$ .

3: **if**  $k \leq$  maxIter OR  $(\|\tilde{h}^{(k)} - \tilde{h}^{(k+1)}\|_2 > \epsilon_h$  AND  $\|\tilde{\mathbf{s}}^{(k)} - \tilde{\mathbf{s}}^{(k+1)}\|_2 > \epsilon_{\tilde{\mathbf{s}}}$  AND  $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\|_2 > 0)$  **then**

4:  $k \leftarrow k + 1$

5: Go to 1

6: **end if**

7: **return**  $\mathbf{x}$ ,  $\tilde{h} \leftarrow \tilde{h}^{(k)}$ ,  $\tilde{\mathbf{s}} \leftarrow \tilde{\mathbf{s}}^{(k)}$

---

knowledge of the parameters, we can compute the weights of the edges and select a set of pairs that maximizes the sum of the weights (association). Next, based on the pairing, we can improve the prior knowledge of the parameters for the range estimator (estimation). The association and estimation are repeated in an alternating manner until some criteria are fulfilled. We summarize the method in Algorithm 1. A shortcoming of the proposed algorithm is the lack of convergence proof as the association does not guarantee the reduction of the objective function in estimation compared to the previous iteration. However, in practice, given a good initialization of  $h$  and  $\hat{\mathbf{s}}$ , the algorithm converges in a few iterations.

After execution of the algorithm, we can further refine the remaining pairs by deciding on a threshold with respect to the amplitudes of  $\mathbf{w}$  to separate the pairs into two clusters. Correct pairings fall into the high-amplitude cluster, while nuisance pairings fall into the low-amplitude cluster. The association and estimation algorithm selects a distinct set of pairs that maximize the weights of the edges of the pairs, even if the pairs have very small weights, which most probably mean that they are nuisance pairs. However, the existence of a small number of nuisance pairs does not significantly affect the performance of parameter estimation. The distribution of the weights is most likely bimodal with one small-amplitude peak representing the nuisance pairs, and one large-amplitude peak representing the correct pairs. A threshold is set such that two weight clusters, which have minimum intracluster variance, are formed. Subsequently, we can remove one of the weight clusters that belongs to the nuisance pairs.

## D. Numerical Simulations

We use the same data generated in simulations 1 and 2 in the previous section, but do not assume perfect association of the direct and surface-reflected snaps. In addition, 100 direct arrivals and 100 reflections are independently discarded to create a 10% nuisance arrival noise and hence the maximum number of correct pairs is 1800. For coarse pairing in both simulations, we set  $\epsilon_\phi = 20^\circ$  and  $h_u = 20$  m. We use  $\lambda = 8000$  to compute the parameter estimation and  $\kappa = -0.9$  to compute the weights for simulation 1 and simulation 2. The estimated locations of the snaps  $\tilde{\Lambda}$  based on these parameters are shown in Fig. 6. We only depict the location of snaps at  $x$ - and  $y$ -axes in simulation 1 for simpler visualization. We present the estimated receiver orientation  $\tilde{\mathbf{s}}$  in the form of pitch  $\tilde{\alpha}$  and roll  $\tilde{\beta}$  for the ease of comparison with the actual simulated receiver orientation in  $\alpha^o$  and  $\rho^o$ . The proposed method is able to recover most of the direct and surface-reflected snaps in both cases. For snaps that are farther from the receiver, the accuracy of the estimated parameters is slightly degraded. Fig. 6(b) appears to show a more power-law-like range error when compared to Fig. 6(a), which is the actual location of snaps. This is expected as the proposed range estimator does not consider the error introduced by the disturbed ocean surface such as  $\eta$  and  $\nu$ . The simulated data were contaminated by the disturbed ocean surface and hence the error (difference between the actual and the estimated location) is not only linearly proportional to the range of the snap.

## VI. EXPERIMENTAL RESULTS

An experiment was conducted at St. John Island, Singapore, in August 2014. ROMANIS, a broadband planar array comprising 508 sensors and measuring 1.3 m in diameter, was deployed to collect the acoustic pressure recording of ambient noise, which is dominated by snapping shrimp noise [20], [23]. The frequency band of ROMANIS is 25–85 kHz, which is also the frequency band that contains the more significant amount of energy from snapping shrimp. The array was located at  $1^\circ 13.027' \text{ N}$ ,  $103^\circ 51.106' \text{ E}$  with average water depth of 5 m. The experimental location is considered as shallow water and the assumption of isovelocity channel is true for shallow Singapore waters [24]. Our own recent measurements using CTD in Singapore waters also agreed that the channel is isovelocity. It was surrounded by man-made structures such as the jetty, watergate, and a fish farm. Large amount of snaps were observed from these areas as they provide suitable shelter for snapping shrimp. ROMANIS was positioned approximately 8 m away facing the jetty where a lot of snaps can be found. Fig. 7 shows the photographs of the experiment and a labeled map indicating the position of the surrounding man-made structures.

Four data sets containing snapping shrimp noise from the jetty were recorded at different time slots. Each of the data set is a 300-s acoustic pressure recording. The sea state during the recording was calm. One of the data sets was collected on August 12, 2014 at 15:49:43 local time. Fig. 8 shows a 10-s clip of ambient noise recording during the experiment, which is dominated by snapping shrimp noise. As can be observed,

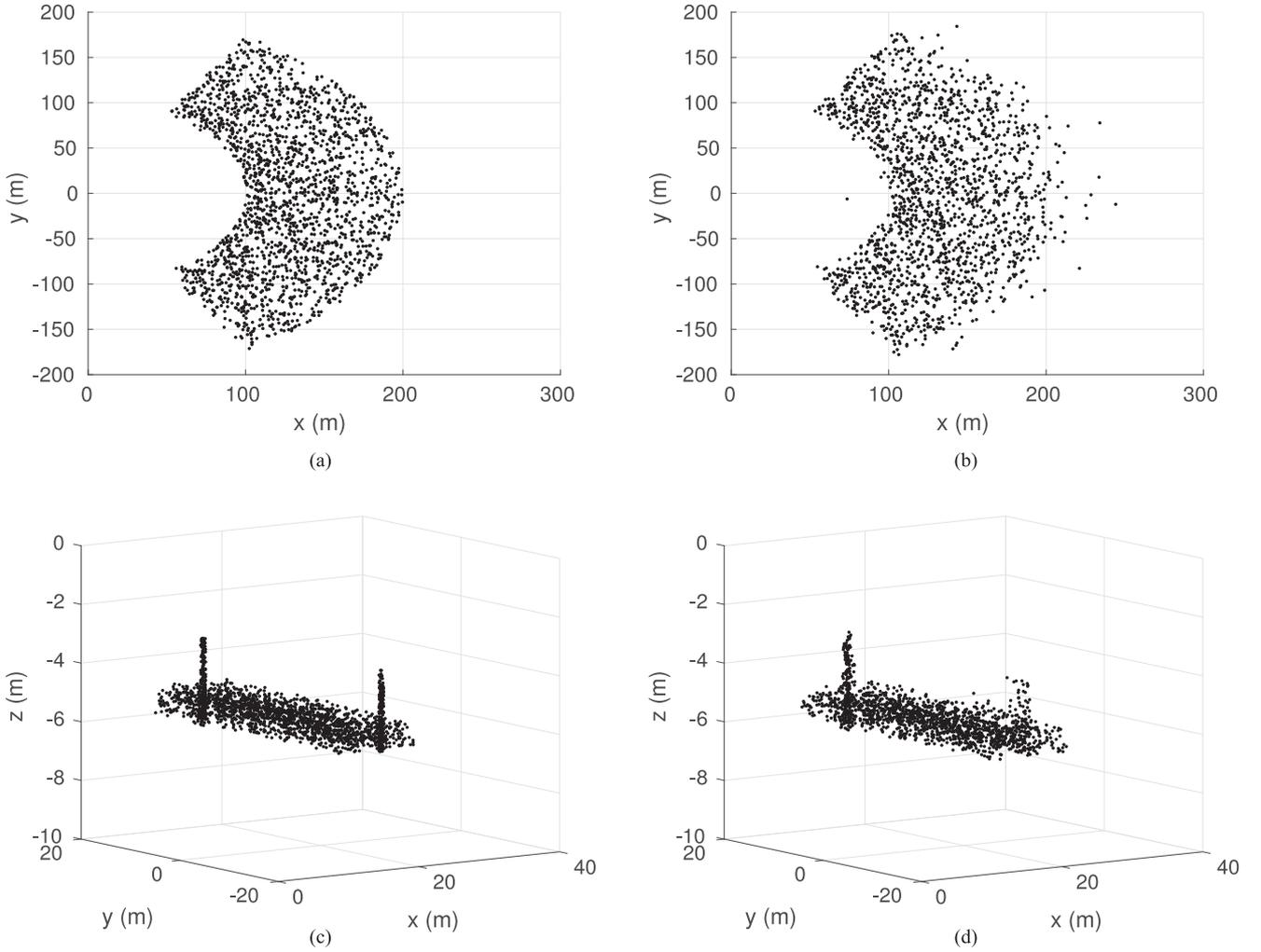


Fig. 6. Location of the snaps based on alternating association and estimation. The actual location of the snaps is presented for comparison. (a) Actual location of snaps at the  $x$ - and  $y$ -axes in simulation 1. (b) Estimated location of snaps at the  $x$ - and  $y$ -axes in simulation 1.  $\hat{h} = 15.1689$  m,  $\hat{\rho} = -4.9510^\circ$ ,  $\hat{\alpha} = -6.9403^\circ$ , and the number of correct pairs is 1795. (c) Actual location of snaps in simulation 2. (d) Estimated location of snaps in simulation 2.  $\hat{h} = 5.1205$  m,  $\hat{\rho} = -4.9840^\circ$ ,  $\hat{\alpha} = 7.0444^\circ$ , and the number of correct pairs is 1787.

the recorded signal comprises a large number of impulsive transient signals. These signals are the direct arrival of the snaps and the respective multipath propagations. From the data set, 34 531 arrivals, which comprises direct arrival and multipath reflection of snapping shrimp noise, were detected [22] and, subsequently, the location of these snaps were estimated. We set  $\epsilon_\phi = 20^\circ$  and  $h_u = 20$  m for the coarse pairing. We use  $\lambda = 5000$  for the parameter estimation and  $\kappa = -0.9$  for  $\mathbf{w}$ . We present two results regarding the estimated location of snaps. The first is purely the estimated location of snaps, while the second is the spatially smoothed estimated location of snaps. This postprocessing step (spatial smoothing) is to improve the estimated range by calculating the mean of estimated ranges of snap sources having approximately the same DoA. The reason is that we usually observe snapping shrimp living close together in colonies whether on coral reefs or man-made structures. So the direct snap arrivals from the same DoA over time are most likely to have originated from the same shrimp colony, and not from different colonies at different ranges. Also, it is likely that these shrimp colonies reside on solid structures that

form a barrier for direct arrivals propagation from other shrimp sources farther away. Noted that all the aforementioned assumptions might not be valid for more complex snap propagations which are affected by diffraction. Let  $\mathcal{B}$  be the discrete set of all possible DoA. Then, we define  $\mathcal{B}'_{(\phi_i, \theta_i)} = \{(\phi, \theta) | |\phi - \phi_i| \leq \epsilon_\phi, |\theta - \theta_i| \leq \epsilon_\theta, (\phi, \theta) \in \mathcal{B}\}$ , where  $\epsilon_\phi$  and  $\epsilon_\theta$  are some small angles. We can now rectify the estimated range of the snap  $i$  such that

$$\tilde{D}'_i = \frac{1}{|\mathcal{B}'_{(\phi_i, \theta_i)}|} \sum_{j: (\phi_j, \theta_j) \in \mathcal{B}'_{(\phi_i, \theta_i)}} \tilde{D}_j. \quad (22)$$

Fig. 9 displays the estimation results of the shrimp locations for comparison. Based on  $\epsilon_\phi = 1^\circ$  and  $\epsilon_\theta = 1^\circ$ , the estimated snap location with spatial smoothing reveals some of the pillars of the jetty. In Fig. 9(a), it can be seen that without spatial smoothing, the pillars of the jetty are not so clearly revealed, because there is a bigger spread of estimated range of snaps from the same colony, possibly due to vertical wave motion. Comparing the layout of the jetty in Fig. 10 and the estimated



(a)



(b)



(c)

Fig. 7. Experiment at St. John Island, Singapore. (a) Deploying ROMANIS at the jetty. (b) Jetty. (c) Position of ROMANIS in the labeled Google map.

locations of snaps in Fig. 9(b), we see that the vertical patterns formed by the estimated snap locations match the pillars of the jetty. This suggests that the patterns of snaps probably originated from the snapping shrimp lodging on the pillars. In particular, the measured distance between pillars A and B is about 6 m, while the estimated distance in the corresponding structures in Fig. 9(b) is about 5.87 m. The distance between pillars A and C is more than 4 m (since the pillars are not truly vertical but are inclined outward such that the top of the pillars are 4 m apart but the base of the pillars are more than 4 m apart). The measured distance between the two vertical structures corresponding to pillars A and C in Fig. 9(b) is 6.70 m. The estimated sources of

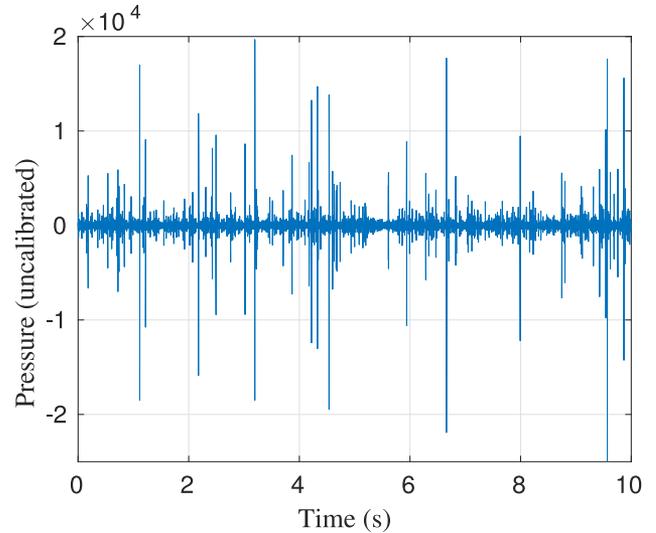
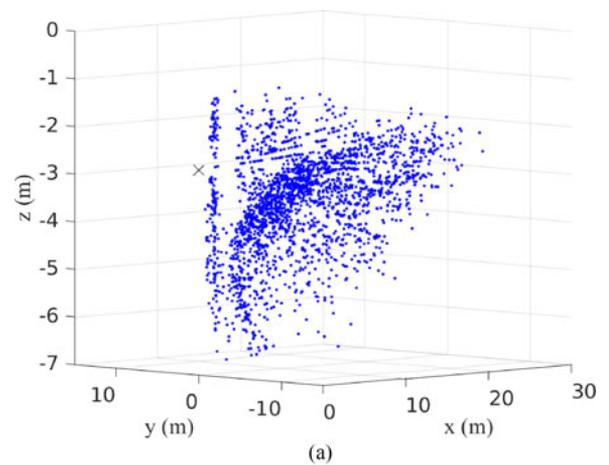
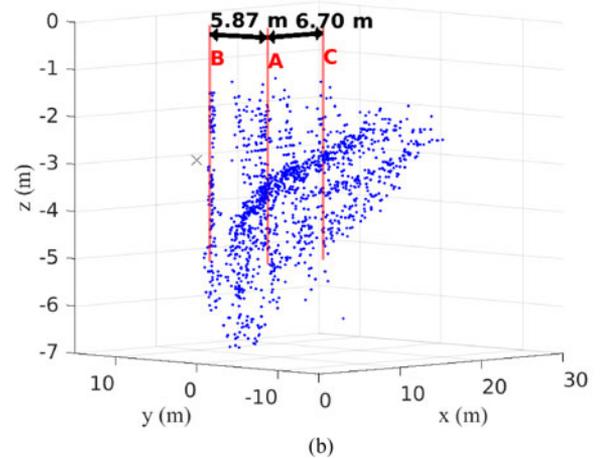


Fig. 8. Ten-second acoustic pressure recording of ambient noise in Singapore waters dominated by snapping shrimp noise.



(a)



(b)

Fig. 9. Estimated locations of snaps on the jetty. (a) Dots show the estimated locations of snaps without spatial smoothing. The origin of ROMANIS is denoted by the black cross. (b) Dots show the estimated locations of snaps with spatial smoothing and vertical lines illustrate the  $x$ - $y$  position of the pillars of the jetty. The origin of ROMANIS is denoted by the black cross. The videos of the plots are available as supplementary downloadable material available at <http://ieeexplore.ieee.org>.

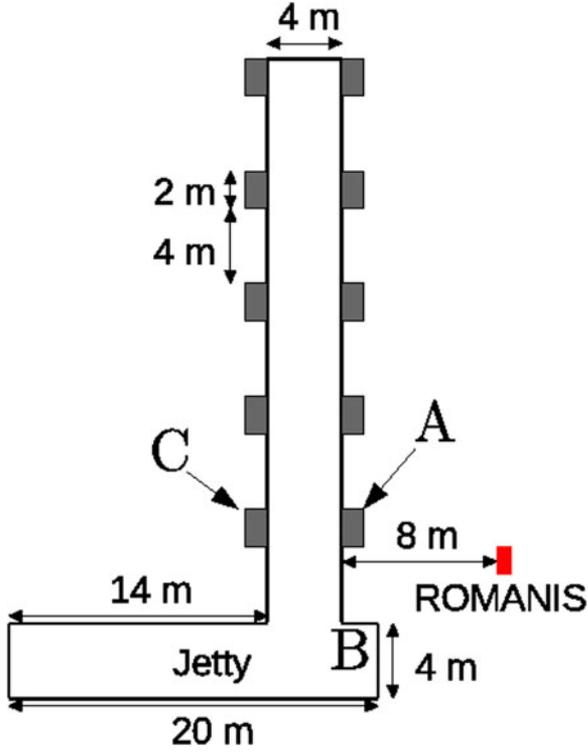


Fig. 10. Layout of the jetty. The average water depth around ROMANIS, denoted by red rectangle, was 5 m.

snaps trace out a slope extending from the seabed to the shore, which is as expected from the local bathymetry.

Most of the estimated snaps from pillar A are approximately 8–30 m from the receiver, which agrees with the actual position of the jetty. Note that by reducing the detection threshold, the DoA-ToA of impulsive signal detection method allows slightly curved wavefront to be detected but it may suffer from higher false positive detection rate. This is the reason that the snaps within the near-field of the receiver, which is 8–13 m, are in the plot. A large portion of the snaps are still beyond the near-field of the array. The remaining data sets were collected at 16:06:46, 16:44:16, and 16:51:25 local time, respectively. Combining with the aforementioned data sets, the estimated nominal depth and receiver orientation over four data sets collected in different time slots are shown in Fig. 11. On August 12, 2014, a high tide was reported at 12:45 followed by a low tide at 18:25 with the tidal height dropping gradually in between. According to the Singapore Tide Table 2014, the tidal height on April 12, 2014 at Tanjong Pagar, the closest point to St. John Island, measured 1.5 m at 16:00:00 and 1.1 m at 17:00:00 [25]. The difference between these tidal heights is 0.4 m. This agrees with our observation regarding the changes of the estimated nominal receiver depth. The rate of reduction of  $\tilde{h}$  is the highest between data set 16:06:46 and 16:44:16 as they have the largest time difference. The rate is lower for data sets that are smaller in time difference. The estimated receiver orientation obtained using snapping shrimp noise is consistent over data sets. The receiver orientation is shown to be slightly tilted with respect to the sea level.

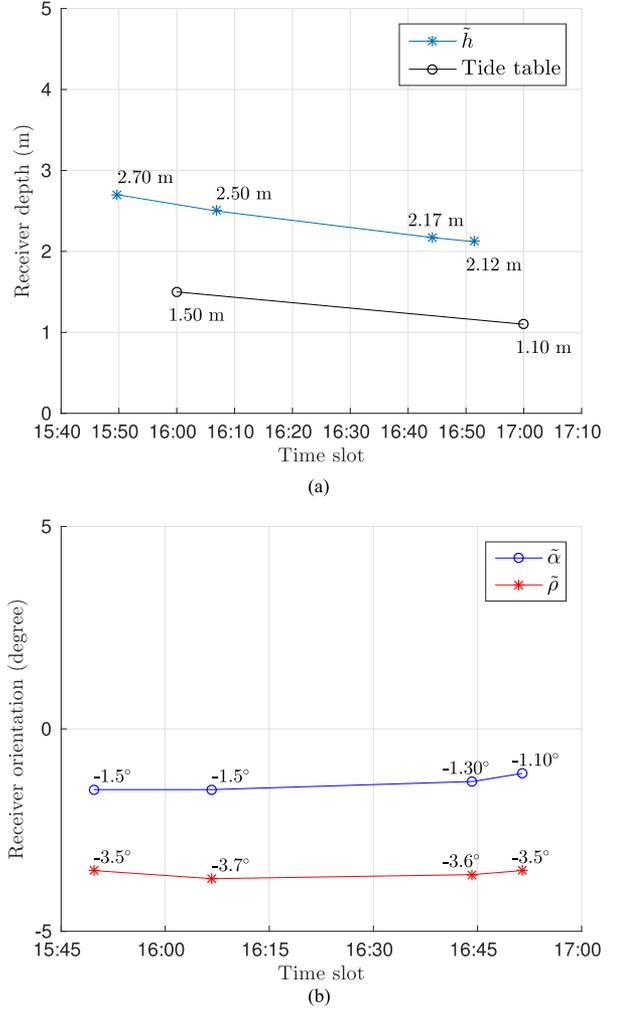


Fig. 11. Estimated parameters based on different data sets. (a) Estimated nominal receiver depth  $\tilde{h}$ . The asterisk is the estimated nominal receiver depth and the circle denotes the tidal height. (b) Estimated receiver orientation defined by roll  $\tilde{\rho}$  and pitch  $\tilde{\alpha}$ .

## VII. CONCLUSION

The success in localizing snapping shrimp noise potentially lays the foundation for a wide variety of passive sensing applications with snapping shrimp noise. Our work utilizes the knowledge of acoustic wave propagation as well as mathematical optimization to allow far-field localization of snapping shrimp noise using a small-aperture array. For instance, the ability to localize snapping shrimp in a 3-D space using a small-aperture receiver facilitates coral reef monitoring. The idea of using snapping shrimp noise for coral reef monitoring either by means of large-aperture receivers or merely one-sensor receivers has been investigated in [9] and [11]. The former is capable of covering a large region of interest but is inefficient for long-term monitoring. The latter is easy to implement but is limited by the area of (monitoring) coverage. A large-area monitoring system using a small-aperture sensor array would be a viable approach that fills the gap between the two approaches. It is important to stress that the proposed algorithm relies on direct and surface-reflected snaps. In the experimental data, even though multipath

reflections such as bottom-reflected snaps are detected, the number of these reflections is generally insignificant compared to the direct and surface-reflected snaps. The proposed algorithm is robust to the possible small number of wrong associations. However, the performance of the algorithm may degrade if a large number of bottom-reflected snaps are observed. Through the numerical and experimental data sets, the alternating association and estimation algorithm converged after a number of iterations given the prior knowledge of the parameters. Further analysis, for example, on the rate of convergence and the correctness of the algorithm is essential to examine this behavior. Overall, by using a small-aperture receiver, we were able to localize far-field snapping shrimp in a 3-D space based on our method. This paves the way for efficient, long-term, and portable passive sensing with snapping shrimp noise.

#### ACKNOWLEDGMENT

The authors would like to thank their colleagues at Acoustic Research Laboratory for their support in the development of ROMANIS and during the field experiments. They would also like to thank L. L. Ong for proofreading the paper.

#### REFERENCES

- [1] D. H. Cato and M. J. Bell, "Ultrasonic ambient noise in Australian shallow waters at frequencies up to 200 kHz," Mater. Res. Lab., Maribyrrong, Vic. Australia, Tech. Rep. MRL-TR-91-23, 1992.
- [2] M. Versluis, B. Schmitz, A. von der Heydt, and D. Lohse, "How snapping shrimp snap: Through cavitating bubbles," *Science*, vol. 289, no. 5487, pp. 2114–2117, 2000.
- [3] M. W. Legg, "Non-Gaussian and non-homogeneous Poisson models of snapping shrimp noise," Ph.D. dissertation, Dept. Imag. Appl. Phys., Curtin Univ. Technol., Bentley, WA, Australia, 2010.
- [4] M. Chitre, S. H. Ong, and J. Potter, "Performance of coded OFDM in very shallow water channels and snapping shrimp noise," in *Proc. MTS/IEEE OCEANS Conf.*, Washington, DC, USA, 2005, pp. 996–1001.
- [5] A. Mahmood, M. Chitre, and M. A. Armand, "Improving PSK performance in snapping shrimp noise with rotated constellations," in *Proc. 7th ACM Int. Conf. Underwater Netw. Syst.*, 2012, pp. 1–8.
- [6] C. Epifanio, J. Potter, G. Deane, M. Readhead, and M. Buckingham, "Imaging in the ocean with ambient noise: the ORB experiments," *J. Acoust. Soc. Amer.*, vol. 106, no. 6, pp. 3211–3225, 1999.
- [7] M. Chitre, S. Kuselan, and V. Pallayil, "Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation," *J. Acoust. Soc. Amer.*, vol. 132, no. 2, pp. 838–847, 2012.
- [8] S. D. Simpson, M. Meekan, J. Montgomery, R. McCauley, and A. Jeffs, "Homeward sound," *Science*, vol. 308, no. 5719, pp. 221–221, 2005.
- [9] M. O. Lammers, R. E. Brainard, W. W. Au, T. A. Mooney, and K. B. Wong, "An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats," *J. Acoust. Soc. Amer.*, vol. 123, no. 3, pp. 1720–1728, 2008.
- [10] M. Chitre, M. Legg, and T. B. Koay, "Snapping shrimp dominated natural soundscape in Singapore waters," *Contrib. Mar. Sci., Nat. Univ. Singapore*, 2012.
- [11] S. E. Freeman, M. J. Buckingham, L. A. Freeman, M. O. Lammers, and L. Gerald, "Cross-correlation, triangulation, and curved-wavefront focusing of coral reef sound using a bi-linear hydrophone array," *J. Acoust. Soc. Amer.*, vol. 137, no. 1, pp. 30–41, 2015.
- [12] B. G. Ferguson and J. L. Cleary, "In situ source level and source position estimates of biological transient signals produced by snapping shrimp in an underwater environment," *J. Acoust. Soc. Amer.*, vol. 109, pp. 3031–3037, 2001.
- [13] Y. M. Too and M. Chitre, "Localization of impulsive sources in the ocean using the method of images," in *Proc. MTS/IEEE OCEANS Conf.*, St. John's, NL, Canada, 2014, pp. 1–6.
- [14] M. Chitre, T. B. Koay, and J. Potter, "Origins of directionality in snapping shrimp sounds and its potential applications," in *Proc. IEEE OCEANS Conf.*, San Diego, CA, USA, 2003, vol. 2, pp. 889–896.
- [15] M. Olivieri, S. A. L. Glegg, and R. K. Coulson, "Measurements of snapping shrimp colonies using a wideband mobile passive sonar," *J. Acoust. Soc. Amer.*, vol. 103, no. 5, pp. 3000–3000, 1998.
- [16] R. Aubauer, M. O. Lammers, and W. W. Au, "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," *J. Acoust. Soc. Amer.*, vol. 107, no. 5, pp. 2744–2749, 2000.
- [17] A. Širović, J. A. Hildebrand, and S. M. Wiggins, "Blue and fin whale call source levels and propagation range in the Southern Ocean," *J. Acoust. Soc. Amer.*, vol. 122, no. 2, pp. 1208–1215, 2007.
- [18] W. M. Zimmer, "Range estimation of cetaceans with compact volumetric arrays," *J. Acoust. Soc. Amer.*, vol. 134, no. 3, pp. 2610–2618, 2013.
- [19] L. M. Brekhovskikh, *Fundamentals of Ocean Acoustics*. New York, NY, USA: Springer-Verlag, 2003.
- [20] T. B. Koay, E. T. Tan, M. Chitre, and J. R. Potter, "Estimating the spatial and temporal distribution of snapping shrimp using a portable, broadband 3-dimensional acoustic array," in *Proc. IEEE OCEANS Conf.*, San Diego, CA, USA, 2003, vol. 5, pp. 2706–2713.
- [21] S. Kuselan, M. Chitre, and V. Pallayil, "Ambient noise imaging through joint source localization," in *Proc. MTS/IEEE OCEANS Conf.*, Waikoloa, HI, USA, 2011, pp. 1–5.
- [22] Y. M. Too, M. Chitre, and V. Pallayil, "Detecting the direction of arrival and time of arrival of impulsive transient signals," in *Proc. MTS/IEEE OCEANS Conf.*, Monterey, CA, USA, 2016, pp. 1–8.
- [23] J. R. Potter and T. B. Koay, "Do snapping shrimp chorus in time or cluster in space? Temporal-spatial studies of high-frequency ambient noise in Singapore waters," in *Proc. Eur. Conf. Underwater Acoust.*, Lynos, France, 2000, pp. 497–502.
- [24] H. Song *et al.*, "Numerical modeling of a time reversal experiment in shallow Singapore waters," in *Proc. IEEE OCEANS Conf.*, Singapore, 2007, pp. 1–6.
- [25] "Singapore tide table year 2014," Hydrographic Dept., Maritime Port Authority Singapore, 2014.



**Yuen Min Too** (M'15) received the B.Eng. degree in biomedical from the Universiti Teknologi Malaysia, Johor Bahru, Malaysia, in 2010 and the Ph.D. degree in underwater acoustics and signal processing from the Department of Electrical and Computer Engineering, National University of Singapore (NUS), Singapore, in 2017.

Since 2015, he has been with the Acoustic Research Laboratory, NUS, where he was a Research Engineer, and is currently a Research Fellow. His research interests include underwater acoustics and signal processing, specifically on compressed sensing and array signal processing in complex underwater noise environments. He is also interested in developing machine learning techniques for underwater applications.



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

From 1997 to 1998, he was with the Acoustic Research Laboratory (ARL), NUS, Singapore. From 1998 to 2002, he headed the technology division of a regional telecommunications solutions company. In 2003, he rejoined ARL, initially as the Deputy Head (Research) and is currently the Head of the laboratory. He also holds a joint appointment with the Department of Electrical and Computer Engineering, NUS, as an Associate Professor. His current research interests include underwater communications, autonomous underwater vehicles, model-based signal processing, and modeling of complex dynamic systems.

Dr. Chitre has served on the technical program committees of the IEEE OCEANS, WUWNet, DTA, and OTC conferences and has served as a reviewer for numerous international journals. He serves as an Associate Editor for the IEEE JOURNAL OF OCEANIC ENGINEERING, and has also served on editorial boards of the IEEE COMMUNICATIONS MAGAZINE and the IEEE ICRA conference. He was the Chairman of the Student Poster Committee for IEEE OCEANS'06 in Singapore, and the Chairman for the IEEE Singapore AUV Challenge 2013. He is currently the IEEE Ocean Engineering Society Technology Committee Co-Chair of underwater communication, navigation, and positioning.



**George Barbastathis** (M'03) received the B.S. degree in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 1993 and the M.Sc. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1994 and 1997, respectively. He completed his Ph.D. work at the University of Illinois at Urbana-Champaign, Champaign, IL, USA.

Since 1999, he has been a member of the Faculty of Mechanical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, where he is currently the Singapore Research Professor of Optics and a Professor of Mechanical Engineering. He is also with the Institute for Soldier Nanotechnologies, MIT. From 2006 to 2007, he was a Visiting Scholar at the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Since 2008, he has been a Research Scientist at the Singapore-MIT Alliance for Research and Technology Centre, Singapore. His research interests include the physics and engineering of 3-D optical imaging systems based on volume holography, digital holography, and dielectric nanostructures with variable subwavelength period in the pass-band regime. His target applications are in optical imaging and metrology for biological, environmental, and energy related research.

Prof. Barbastathis has served as a Topical Editor for the *Journal of the Optical Society of America A* and serves as an Editor for the new journal *Light: Science and Applications*. In 2011, he was elected as a Fellow of the Optical Society of America.



**Venugopalan Pallayil** (SM'07) received the M.Sc. degree in physics and the Ph.D. degree in microwave electronics from Cochin University of Science and Technology, Kerala, India, in 1981 and 1992, respectively.

From 1986 to 1998, he was an R&D Scientist at the Defence Research and Development Organisation, India, contributing to the development of airborne ASW sensors. In late 1998, he joined the Acoustic Research Laboratory (ARL), Tropical Marine Science Institute (TMSI), National University of Singapore, Singapore, as a Research Fellow. He is currently a Senior Research Fellow and the Deputy Head at ARL and also supporting TMSI as Operations Manager. At ARL, he has been actively involved and led many projects, including underwater ambient noise imaging, time reversal mirror experiments, and digital thin line array development. He has a strong interest in the development of AUV-based sensors and applications.

Dr. Pallayil was one of the leading members of the team which won the prestigious Defence Technology Prize in 2004 for the project ROMANIS on ambient noise imaging. As an active IEEE member, he has served the local Oceanic Engineering Society (OES) Chapter in many capacities, including the Chairmanship. He was also the Chair for Finance for the OCEANS'06 Asia Pacific Conference in Singapore and has recently been elected to serve on the IEEE OES Administrative Committee for 2018–2020. He has been actively involved with the organization of the Singapore AUV Challenge, an IEEE OES student outreach activity, since 2013. He is a member of the Acoustical Society of America and the Society of Acoustics, Singapore.