IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity

PERRY NAUGHTON[1], PHILIPPE ROUX[2], CURT SCHURGERS[1], RYAN KASTNER[3], JULES S. JAFFE[4] and PAUL L. D. ROBERTS,[4]

[1]Department of Electrical and Computer Engineering, University of California, San Diego, San Diego, CA 92093 USA (e-mail: {pnaughto,cschurgers}@ucsd.edu)
[2]ISTerre, Université Grenoble Alpes, CNRS, Grenoble, France 38000 USA (e-mail: philippe.roux@univ-grenoble-alpes.fr)
[3]Department of Computer Science and Engineering, University of California, San Diego, San Diego, CA 92093 USA (e-mail: kastner@ucsd.edu)
[4]Scripps Institution of Oceanography, University of California, San Diego, San Diego CA 92093 USA (e-mail: {jjaffe,plroberts}@ucsd.edu)

Corresponding author: Perry Naughton (e-mail: pnaughto@ucsd.edu)

This work is part of "INSPIRE Track I: Distributed Sensing Collective to Capture 3D Soundscapes" supported by the NSF under Grant No. 1344291. This material is also based upon work supported by NSF Grant No. OCE 09- 27449, the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144086, the NSF Graduate Research Opportunities Worldwide program, the Chateaubriand STEM Fellowship, the San Diego chapter of the ARCS Foundation, and the Friends of the International Center Scholarship at UCSD. ISTerre is part of LabEx OSUG@2020.

ABSTRACT There is increasing interest in deploying swarms of underwater vehicles for marine surveys. One of the main challenges when designing these systems is coming up with an appropriate way to localize each vehicle in relation to one another. This work considers the self-localization of a deforming swarm of subsurface floating vehicles using impulsive sources of opportunity, such as the sounds of snapping shrimp that are present in warm coastal waters. Impulsive sound sources provide high intensity, broadband signals that facilitate accurate arrival time detections across each vehicle. This makes them useful references for a self-localization solution. However, the similarity between different signals presents a significant correspondence problem, which must be solved to provide accurate estimates of the changing geometry of the swarm. A geometric solution to this correspondence problem is shown and an optimization procedure is proposed to track the geometry of a swarm as it changes. The method is verified using a swarm of 17 self-ballasting subsurface floats that independently drifted with currents off of the coast of San Diego, California. The changing geometry of the floats was estimated using both an acoustic localization system and the proposed approach. The two estimates show good agreement, validating our method. We believe that this new localization strategy is useful for high endurance, low power, multi-vehicle surveys.

INDEX TERMS Ambient Noise, Self-localization, Underwater Acoustics, Underwater Vehicles

I. INTRODUCTION

UNDERWATER vehicles and sensors have the potential to expand our scientific, commercial, and naval capability. They have already proven useful in visually mapping benthic habitats [1], inspecting the hulls of ships [2], and in underwater mine detection [3]. In order to increase the efficacy of marine surveys, there is increasing interest in designing multi-node systems [4]–[6]. An important consideration in all marine surveys is how to localize each node so that it can interact with its environment and provide a spatial reference to the data it collects. These methods are still an active area of research [7], [8].

The implementation of a specific underwater localization protocol dictates important trade-offs in a survey including: the cost of the survey, the power consumption, duration, and the area covered. Almost all underwater localization relies on some form of acoustics, the difference is where the acoustic signal is generated and how it is processed. The most accurate systems use dedicated infrastructure to base their position from so that each measurement is an independent reference to a global coordinate frame [5], [9]. While effective, these systems require significant effort to set up and limit the range of a deployment. Other systems leverage the return of a signal they generate (e.g. Doppler shifts [10], [11]) to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

measure the movement of each node in their environment. Successive measurements can be integrated into a trajectory estimate. These systems are less restricted in their movement and require less infrastructure, but they force the signal to be generated by each node, making this method power intensive. Additionally, the integration over successive time samples means this method suffers from drift. In multi-node systems, positioning information can be exchanged between nodes through acoustic communication [12]–[14] to increase the localization accuracy of each individual node. These systems can also suffer from large power consumption, poor localization performance, or both.

A promising alternative to active localization systems is to leverage ambient noise in the ocean to estimate the changing geometry of a swarm. By focusing only on passive signals, we mitigate many of the trade-offs of using active sources: there could be greater coverage, less set up time, and longer deployments. These advantages would improve the capability of many ocean surveys that rely on estimating the relative location of vehicles, acoustic signals, or both by allowing longer, infrastructure free deployments. Examples of such surveys include measuring submesoscale ocean dynamics [5] and detecting the presence of marine vessels [15]. These improvements come at the cost of performing the localization offline and only being able to recover the swarm geometry in a relative coordinate frame because the vehicles cannot communicate and the locations of the noise sources are unknown, respectively. While promising, localizing mobile vehicles using ambient signals presents many challenges. These challenges stem from the fact that we cannot use signals we design, we must use what is already present in the ambient soundscape.

The first step of estimating a relative geometry is to derive spatial constraints between vehicles by detecting signals at each vehicle and computing timing information between vehicles. This becomes difficult when considering moving vehicles for two reasons. The first reason is that it can be difficult to match the arrival time of an individual signal across the recordings of each vehicle. This happens when signals have similar structure and occur frequently in time. For example, consider trying to find matching pairs between two identical sound sources across two vehicles. If the sounds are transmitted far apart in time, then the signals with the closest arrival times on the recordings of each vehicle yield the matching pairs. However, if they transmit at the same time, it is difficult to determine the matching pairs without knowing the geometry between the vehicles and sound sources. This problem, typically known as a correspondence problem, must be solved before further processing can proceed. A second challenge is that not all vehicles recover the same information. A self-localization solution must incorporate a sparse set of constraints to estimate the whole swarm geometry.

This work looks at the self-localization of underwater vehicles using impulsive noise signals from snapping shrimp. These signals dominate the higher frequencies of the ambient soundscape in warm coastal waters [16], [17], have large source levels that can be heard from 100s of meters away, and occur frequently in time from many locations [18]. This makes it difficult to pick out a single "snap" across multiple vehicles. We propose a method to distinguish the correspondence between signals that is based on a geometric argument. With the known correspondences, we formulate a self-localization procedure that can handle a sparse set of constraints. As a demonstration, we utilize snapping shrimp signals to estimate the relative geometry of a subsurface floating swarm that was deployed off of the coast of San Diego, California. We believe that other environmental signals are also appropriate for our solution, including sounds from other marine animals. Our self-localization results are verified by independent measurements from an accurate high frequency pinging system.

II. PROBLEM DESCRIPTION AND BACKGROUND

We consider a set of moving vehicles (shown by red, green, and blue spheres in Figure 1 (a)) that are recording ambient sounds in the water column. Most of the sounds are from the far field, and some sounds are salient in each recording. Our goal is to track the deforming geometry of the vehicles using the salient signals in the ambient acoustic noise, such as the shrimp pictured in Figure 1 (a). It is easy to determine accurate arrival times for salient signals on each recording (shown by the black arrows in Figure 1 (b)); we discuss our approach in Section III-A. However, it is difficult to determine the correspondence between signals across vehicles. Consider the two shrimp in Figure 1 (a), a yellow one and a purple one. If they produce similar sounds, at similar times, from dissimilar positions, the order of arrival of each may be recorded differently at each vehicle. This is shown in Figure 1 (c), where the order of arrival times between the purple and yellow shrimp is different for one of the vehicles. This makes picking out the same signal in all recordings difficult. Additionally, choosing correspondences based on metrics such as correlation fail because of the similarity between the signals [19]. We propose a geometric argument to estimate correspondences between recordings in Section II-A, and discuss our implementation and challenges in Section III-B and III-B1, respectively. The difference of arrival times (τ_1 and τ_2 from Figure 1 (c)) place constraints on the final geometry of the vehicles (they must be separated by $c\tau_1$ and $c\tau_2$ along the direction of travel from the purple and yellow sound sources, respectively, where c is the propagation speed of the signal) shown in Figure 1 (d). Not all constraints are shown so the figure does not become overloaded. Using these constraints to estimate the geometry of the vehicles and locations of the noise sources is discussed in Section II-B and implementation details are described in Section III-C.

We can only estimate the relative geometry because we do not assume global information for either the source or vehicle positions. This means that our solution is only unique up to a rigid rotation, translation, and mirroring. We are concerned about estimating only the 2D position of the vehicles because the depth of the vehicles can be accurately estimated using

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

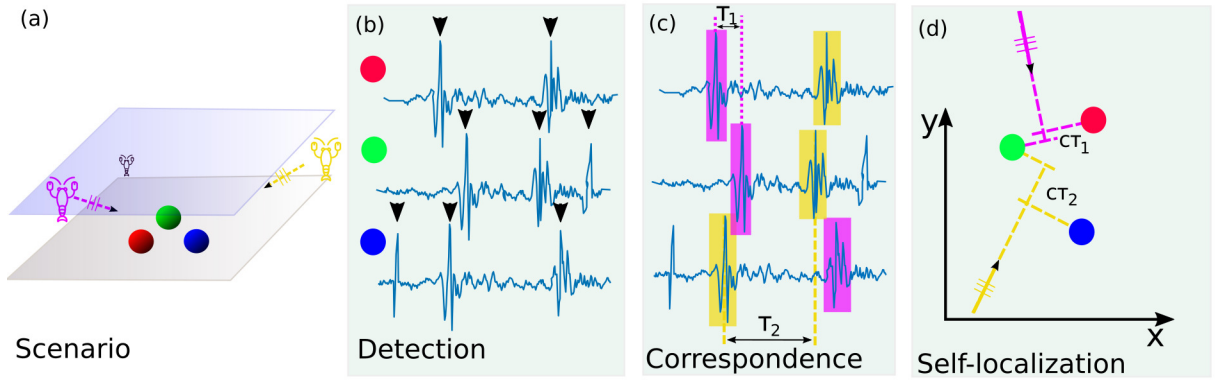

FIGURE 1: Overview: (a) We are trying to track the changing geometry of moving vehicles (shown by red, green, and blue spheres) using ambient sounds in the ocean, such as the pictured snapping shrimp. (b) Our first step is to determine accurate arrival times for these signals on each recording (shown by the black arrows); we discuss our approach in Section III-A. (c) Our next step is to determine the appropriate correspondence between the detected signals. This is a difficult problem that we examine in Section II-A, and discuss our implementation and challenges in Section III-B and III-B1, respectively. (d) The final step is to perform the self-localization estimate using the difference of arrival times (τ_1 and τ_2 from panel (c)). This is discussed in Section II-B and implementation details are described in Section III-C.

a pressure sensor. We denote the 2 dimensional position of vehicle i, $R_i = (R_{xi}, R_{yi})$, and the two-dimensional position of the sound source j, $S_j = (S_{xj}, S_{yj})$. Throughout this work we also denote the $N \times 2$ dimensional vector $\mathbf{R} = [R_{x1}, R_{y1}; R_{x2}, R_{y2}; \ldots R_{xN}, R_{yN}]$ which is the collection of the N vehicle positions. We define a $M \times 2$ dimensional vector $\mathbf{S} = [S_{x1}, S_{y1}; S_{x2}, S_{y2}; \ldots S_{xM}, S_{yM}]$ for the positions of M different sound sources. Our estimation makes use of the time each sound, j, is received at each vehicle, i, denoted $t_{i,j}$. Since we do not know the emission time of each sound source, we will work with time difference of arrival (TDOA) measurements to negate the unknown emission time. The TDOA is defined as:

$$\tau_{i,h} = \frac{\|R_i - S_j\|_2 - \|R_h - S_j\|_2}{c} \tag{1}$$

where $\tau_{i,h} = t_{i,j} - t_{h,j}$, $\|\cdot\|_2$ is the Euclidean distance, and c is the speed of sound in water. This gives us a system of equations that we can use to estimate the unknown positions of the swarm, $\mathbf{R}$, as well as the unknown positions of the sound sources, $\mathbf{S}$:

$$\arg\min_{\mathbf{R},\mathbf{S}} \sum_{h=1}^{N} \sum_{i=1, i\neq h}^{N} \sum_{j=1}^{M} \upsilon_{h,i,j} \left(\|R_i - S_j\|_2 - \|R_h - S_j\|_2 - c(\tau_{i,h}) \right) \tag{2}$$

where $\upsilon_{h,i,j}$ is an indicator variable that takes the value of 1 if noise source j is detected by both vehicles i and h and 0 if such a correspondence cannot be made. This minimization problem is sensitive to two main pitfalls that need to be addressed to find an accurate estimate for $\mathbf{R}$, which are described throughout the rest of this section. First, correspondences must be correct for this minimization function to minimize a meaningful objective function. Finding appropriate values for both $\tau_{i,h}$ and $\upsilon_{h,i,j}$ by looking at minimal sets of vehicles is described in Section II-A. Another difficulty with equation 2 is that it is prone to get stuck in local minima. Developing a robust solution to solving equation 2 is discussed in Section II-B.

A. CORRESPONDENCE SOLUTION

The first challenge is to determine correct correspondences of noise sources across the vehicles of the swarm. This challenge is illustrated in Figure 1 (a) and (c). If two noise events happening at similar times are recorded by two vehicles, they may not be recorded in the same order due to the geometry of the vehicles and noise sources. Furthermore, it is known that correlation based metrics for determining the correspondence between impulsive noises fail because different impulsive noises have similar signal characteristics [19]. This problem needs to be solved to estimate an accurate geometry of our swarm; incorrect correspondences will degrade our geometry estimation.

In order to solve this correspondence problem, we make three simplifying assumptions. The first assumption is on the maximum inter vehicle distance within the swarm, which can either be defined in a deployment, estimated or set conservatively. This assumption limits the number of possible correspondences by defining a window, δ_w, in time that we allow correspondences to be accepted from. The second assumption is that we can model our localization problem as a 2D problem, meaning that we are considering vehicles that are mostly planar. Generally, the depth of the vehicles can be cheaply and accurately determined in the ocean by measuring pressure. Our assumption here also extends to parametrizing the noise source locations. This assumption means that the horizontal distance between a vehicle and sound source is a good approximation for the total distance traveled. In other words, the vertical distance between the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

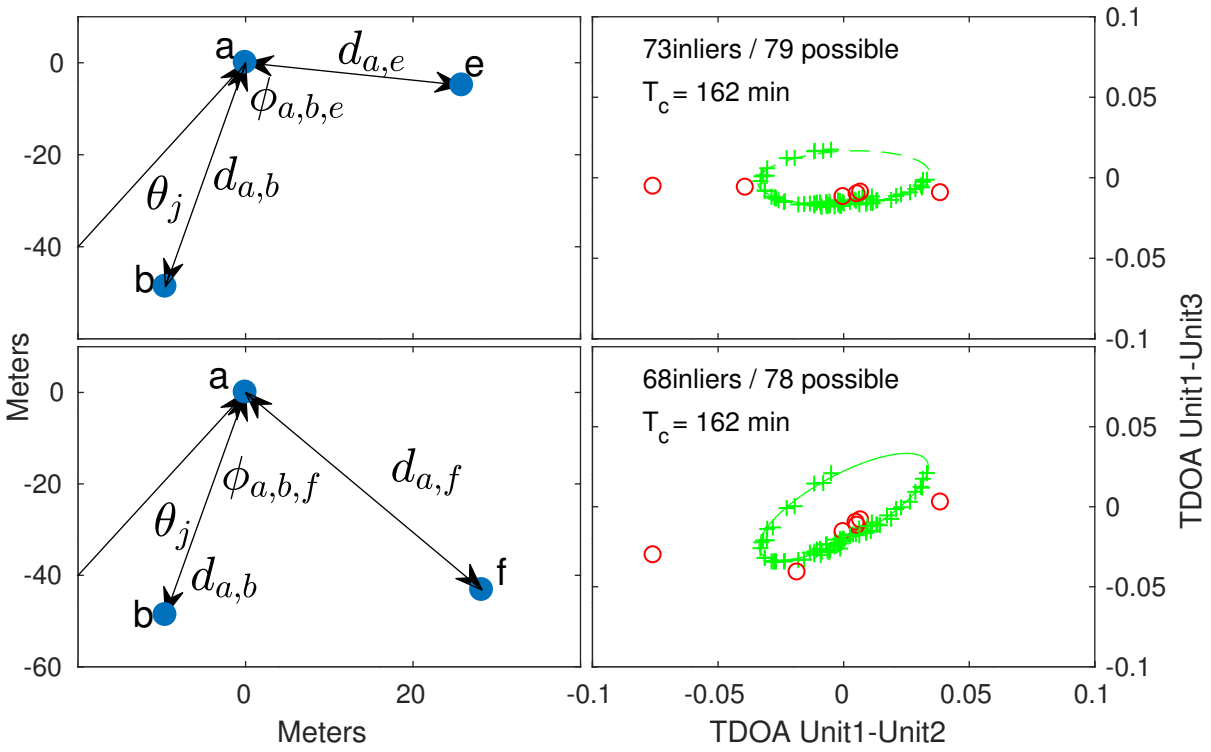

FIGURE 2: Finding correct correspondences: (left) the relative geometry of two different vehicle triplets at $T_c = 162$ min. (right) the corresponding output to the RANSAC framework for the geometry on the left. We can see that the geometry of the vehicles puts constraints on the TDOA measurements between the arrivals of the impulses. For correct correspondences, the TDOA measurements lie on an ellipse parameterized by the angle of arrival, θ_j, of each noise source. Comparing the top and bottom panels shows that changing the geometry of the vehicles changes the geometry of the ellipse. We can also see that the RANSAC framework does a good job at finding an optimal inlier set (green '+'), while rejecting outliers (red 'o'). The correspondences that lie on each ellipse are accepted for further processing while the correspondences in red are labeled as incorrect and rejected. Some outliers are not shown to focus more on the ellipse ($\delta_w = 0.2s$).

sound source and vehicle is far smaller than the horizontal distance. The final assumption is that the noise locations are in the far field. This is related to the previous assumption. The second and third assumptions allow us to parameterize the noise source locations by an angle instead of a three-dimensional coordinate. The validity of this assumption is based on the ratio between the inter-vehicle spacing and the distances between the vehicles and noise sources. Typically, for the far field assumption to hold, the vehicle separation should be 2-4 times smaller than the distance of the noise sources to the swarm [20], [21]. This assumption is justified for shallow water cases where the bathymetry is far smaller than the horizontal distance that the signal must travel. These assumptions minimize the number of vehicles that need to be considered to solve the correspondence problem. This minimal set is discussed in the following paragraphs and is important because the correspondence problem quickly becomes intractable as we consider more vehicles. The addition of another vehicle increases the combinations of possible correspondences that must be considered. While it is possible to consider a solution that is not planar, it would require considering more vehicles at a time in our correspondence solution. Our results in Section IV show experimental justification for these assumptions.

A key insight in solving the correspondence problem is that there are constraints on the TDOA measurements that are imposed by the geometry of the vehicles. With the far field assumption we can approach the correspondence problem by looking at two TDOA measurements between three vehicles. Under a far field assumption these TDOA measurements must lie on an ellipse [20]. We can see this mathematically by examining the TDOA between two pairs of vehicles: pair a and e, and pair a and b (this configuration is shown in top left panel of Figure 2, which will be discussed in full in Section III-B):

$$\tau_{a,b,j} = \frac{d_{a,b}}{c} cos(\theta_j) \tag{3}$$

$$\tau_{a,e,j} = \frac{d_{a,e}}{c} cos(\theta_j + \phi_{a,b,e}) \tag{4}$$

Here $d_{a,b}$ is the distance between vehicle b and a. θ_j is the angle of arrival of sound source j in relation to the line intersecting vehicles a and b and $\theta_j + \phi_{a,b,e}$ is the angle of arrival in relation to vehicle a and e. $\tau_{a,b}$ and $\tau_{a,e}$ trace

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

an ellipse, centered at ($\tau_{a,b} = 0, \tau_{a,e} = 0$) parameterized by the angle of arrival of the sound source, θ_j. There are two interesting things about this result. The first is that with the knowledge of the ellipse parameters we can estimate the 2D geometry of the 3 vehicle swarm (e.g. parameters $d_{a,e}$, $d_{a,b}$ and $\phi_{a,b,e}$). Likewise, given the geometry of the swarm we can also estimate the parameters of the ellipse. Most importantly, given that the sound sources are coming from the far field, they must lie on this ellipse [20].

Our approach is to evaluate correspondences three vehicles at a time. We determine the correct ones by computing the two TDOA measurements of each correspondence and examining which TDOA values lie on the ellipse. Correspondences that result in a point that lies off of the ellipse we reject as incorrect and we retain the correspondences on the ellipse for further processing. The problem with this approach is that the ellipse is not known. Unlike Wendeberg et al. [20] which assumed perfect correspondence, we are using the ellipse result to determine the correct correspondences without knowledge of what the ellipse should be. We do this by determining the ellipse with the most "votes" in a Random Sampling and Consensus (RANSAC) framework [22]. The RANSAC algorithm estimates an ellipse using a minimal set (in our case it takes 3 points to estimate an ellipse with a known center), counting how many candidate correspondences "agree" with that minimal set and iterating until we find the set with the most "agreement". Typically, the set with the most agreement is called the "inlier set". Here agreement is defined as being within a threshold distance from the ellipse, ϵ_t. The RANSAC algorithm is summarized below for a collection of candidate correspondences, $\chi_{a,b,e}$.

1) Initialize maxVotes = 0.
2) Select a minimum set from $\chi_{a,b,e}$ (in this case 3 entries) at random and estimate an ellipse from these points.
3) From the estimated ellipse, measure the distance to each point in $\chi_{a,b,e}$. Count the number of points that are within ϵ_t ms from the ellipse.
4) If the number of points is more than the current value of maxVotes then update the current ellipse model as the correct model and the points within ϵ_t as the new inlier set.
5) Repeat until all possible combinations of 3 from $\chi_{a,b,e}$ are examined or a predetermined maximum number of iterations is reached.
6) The resulting values of the ellipse model and inlier set are accepted.

The output of the RANSAC algorithm gives us information on the correct correspondences which we can use to estimate the geometry of the vehicles.

B. ESTIMATION

Once we have a collection of arrival times for each vehicle and we know which arrivals from one vehicle correspond to the arrivals in another vehicle we can use this information to estimate a geometry for the swarm.

Given a swarm, $\mathbf{R}_\rho$, and noise sources, $\mathbf{S}_\sigma$, where every vehicle has a known time of arrival for each noise source, we define two new variables:

$$\overline{\mathbf{R}}_\rho = \begin{bmatrix} x_{\rho_2} - x_{\rho_1} & y_{\rho_2} - y_{\rho_1} \\ x_{\rho_3} - x_{\rho_1} & y_{\rho_3} - y_{\rho_1} \\ \vdots & \vdots \\ x_{\rho_n} - x_{\rho_1} & y_{\rho_n} - y_{\rho_1} \end{bmatrix}$$

$$\overline{\mathbf{D}}_{\rho,\sigma} = \begin{bmatrix} t_{\rho_2,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_2,\sigma_2} - t_{\rho_1,\sigma_2} & \dots & t_{\rho_2,\sigma_m} - t_{\rho_1,\sigma_m} \\ t_{\rho_3,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_3,\sigma_2} - t_{\rho_1,\sigma_2} & \dots & t_{\rho_2,\sigma_m} - t_{\rho_1,\sigma_m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{\rho_n,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_n,\sigma_2} - t_{\rho_1,\sigma_2} & \dots & t_{\rho_n,\sigma_m} - t_{\rho_1,\sigma_m} \end{bmatrix}$$

Matrix $\overline{\mathbf{R}}_\rho$ is a relative distance matrix to vehicle R_{ρ_1} and matrix $\overline{\mathbf{D}}_{\rho,\sigma}$ is a matrix of the TDOA of each vehicle compared to vehicle R_{ρ_1}. Again, we are not able to estimate the global location of the vehicles because we do not know the locations of any sound sources or vehicles. Making our parameterization relative to the location of one unit does not change our result because we are only able to estimate the geometry of the swarm in a relative sense. With this in mind, we can arbitrarily set $R_{\rho_1} = (0, 0)$ so that $\overline{\mathbf{R}}_\rho = \mathbf{R}_\rho$ and drop the $\overline{\mathbf{R}}_\rho$ notation.

With the above parameterizations, the angle of arrival of each sound source and the relative geometry of the vehicles can be estimated by minimizing the following equation [21]:

$$\underset{\mathbf{R}_\rho, \mathbf{\Lambda}_\sigma}{\arg\min} \|\mathbf{R}_\rho \mathbf{\Lambda}_\sigma - c\overline{\mathbf{D}}_{\rho,\sigma}\| \tag{5}$$

Here $\mathbf{\Lambda}_\sigma$ is a parameterization of the matrix $\mathbf{S}$ in angles instead of 2D coordinates:

$$\mathbf{\Lambda}_\sigma = \begin{bmatrix} cos(\alpha_{\sigma_1}) & cos(\alpha_{\sigma_2}) & \dots & cos(\alpha_{\sigma_m}) \\ sin(\alpha_{\sigma_1}) & sin(\alpha_{\sigma_2}) & \dots & sin(\alpha_{\sigma_m}) \end{bmatrix}$$

Equation 5 can be trivially solved in an affine space using a singular value decomposition of the matrix $\overline{\mathbf{D}}_{\rho,\sigma}$, meaning that the constraint $cos(\alpha_i)^2 + sin(\alpha_i)^2 = 1$ is not upheld. We refer the reader to Thrun [21] for these details. A Euclidean solution can then be estimated by solving a nonlinear optimization problem with P^2 variables, where P is the dimension of the space. In our case P is 2. This number is independent of the number of vehicles, N, and noise sources, M, that we are considering.

In the general case, every vehicle in $\mathbf{R}_\rho$ may not know every arrival time in $\mathbf{S}_\sigma$. In this case, our approach is to estimate different subsets of the swarm in which all vehicles do have an arrival time for every sound, then merge and refine these results using equation 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

III. EXPERIMENT AND IMPLEMENTATION

Our experiment consisted of 17 subsurface floats, which we call Autonomous Underwater Explorers (AUEs). Jaffe et al. [5] and Naughton et al. [15] have more details about the deployment, including both the AUE specifications and environmental considerations. Some key points are summarized here. The AUEs were programmed to hold a depth of 10 meters and drifted with the subsea currents for 5 hours while collecting acoustic data. The vehicles were spaced such that neighboring vehicles were within 10s of meters of each other, with the whole swarm spanning approximately 300m in each direction. The trajectories of the AUEs were recovered using 5 surface pingers that collected timing and positioning information from onboard GPS receivers and transmitted linearly frequency modulated chirps (7-15kHz) that were received by the AUEs. These trajectories are used to validate our estimates of the swarm geometry from the impulsive noises. In addition to the chirps from the surface pingers, the AUEs recorded impulsive noise from snapping shrimp. The shrimp produced a signal from an unknown location, but are known to spend most of their time on the seafloor, which was approximately 50 meters below the oceans surface during this deployment. The clocks of the AUEs are corrected by comparing GPS measurements at the beginning and end of the AUE deployments from onboard GPS sensors. The acoustic environment of the AUEs during the deployment was not ideal for the propagation of the noises around the elements of the AUE swarm. The AUE's depth target, 10m, was on the boundary of the mixed layer and a steep thermocline, providing a strong downward refracting profile. The shallow depth target allowed surface reflections of the noise field to interfere with the direct path (i.e. the environment is a Lloyd's mirror) cf Figure 3 in Naughton et al. [15]. Therefore, shadow zones are expected between the vehicles depending on vehicle distance. Between the shadow zones and the hydrophone dropouts due to the buoyancy adjustment (dropouts account for about a ten percent loss of audio), we cannot expect all vehicles to hear all sources.

To estimate the changing geometry of the vehicles we assume that the geometry of the swarm is unchanged for a period of time, $\delta_s = 4$ min, centered at a center time, T_c. This time period was chosen empirically and is based on the expected relative movement between the vehicles. Notice that the relative movement is different from the absolute motion of the vehicles. For the slowly drifting vehicles in our experiment, the relative movement between the receivers, driven by variations in the current experienced at each vehicle, is smaller than the rate that the swarm moves through the ocean, which is driven by the mean flow of the current. This allows δ_s to be large compared to the average current speed. The center time of each window, T_c, is chosen so that each window overlaps by $\delta_s/2$ with the next and previous window. Our approach is to solve for the relative geometry independently for each time window by performing the following steps (which are outlined in Figure 1 (b-d)):

1) For each T_c in the deployment, select a window of data δ_s seconds long, centered at T_c, from each vehicle.
2) Detect the impulsive sources of opportunity for each vehicle (Figure 1 (b)).
3) Determine the correct correspondences between vehicles using the developments of Section II-A (Figure 1 (c)).
4) Jointly solve for the vehicle positions, $\mathbf{R}$, and noise positions, $\mathbf{S}$, using the development of Section II-B (Figure 1 (d)).

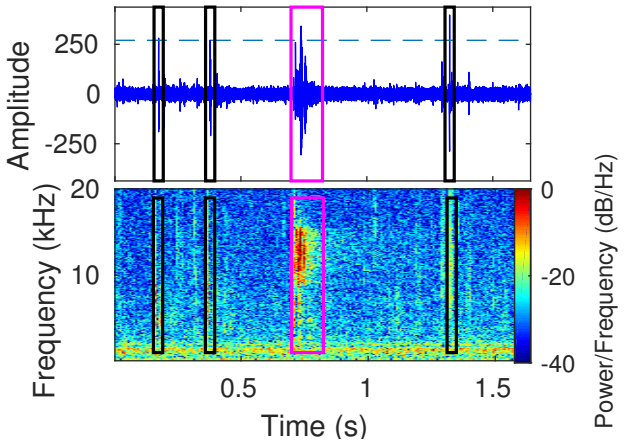

FIGURE 3: The signals: The frequency content and time series are plotted to show the two types of signals utilized in this work. The high energy signal enclosed by a pink box is transmitted by 5 different pingers with known GPS positions and transmit times. We use these signals to estimate the ground truth geometry of our moving swarm. The high amplitude impulses enclosed by black boxes are assumed to be snapping shrimp whose positions and emission time are unknown. We estimate the geometry of our swarm using these noise sources and compare our solution to the more controlled measurements from the pingers. Also, shown is the amplitude threshold that the snap must exceed to be considered for further processing.

A. DETECTION

The first step in our pipeline is to extract the arrival time of signals that could be detected across vehicles of the swarm. For this demonstration we focus on impulsive noise from snapping shrimp, a common signal in the soundscape of warm coastal waters [16]. There are three main considerations for our snapping shrimp detector: we only classify an 'impulse' as something we can accurately determine the arrival time of, we do not classify signals from the GPS pingers as snaps (the pings are also a broadband pulse), and we only classify the first arrival of a snap. To ensure the first requirement, we only consider impulses that are larger than 10 standard deviations of the acoustic time series. From these candidate snaps, we examine the spectrum of the signal to make sure that there is adequate energy in the bands 4-17 kHz. In order to make sure that we do not classify pings from the buoys as a snap, we detect the pings using a matched

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

filter and define a dead zone in time around each ping where we do not allow snaps to be accepted. The window of the dead zone is 25 ms before the detected ping and 100 ms after the detected ping. The width of this dead zone was chosen empirically by looking at the received reflections from the pinger and is conservative. Similarly, to make sure that we only detect the first arrival of the impulsive sound we define a dead zone around detected snaps so that we do not allow another snap to be detected 10 ms after one of the snaps. Figure 3 shows examples of the impulsive snaps and the surface pingers that we use to estimate the trajectories of the vehicles.

B. CORRESPONDENCE SOLUTION

The first step in determining the correct correspondences is to determine all the possible correspondences. We start with the vehicle that will be computed in both TDOA measurements, vehicle a. For each detected noise event, $t_{a,j}$, from a we look for corresponding events in both vehicle b and e that are within $\pm\delta_w = 0.2$s away from the arrival time, $t_{a,j}$. There can be more than one noise event that corresponds to $t_{a,j}$ from either vehicle b or e. The outcome of this first pass is a collection of L different possible correspondences between the three vehicles for the given time window centered at T_c. We call this collection χ_{a,b,e,T_c}. To make the processing easier, we throw away entries in χ_{a,b,e,T_c} where there are ambiguous correspondences. Meaning if there are entries in χ_{a,b,e,T_c} that contradict each other, we discard them.

We then run the candidate correspondences through the RANSAC procedure described in Section II-A. Figure 2 shows two examples of the swarm geometry of a vehicle triangle (left) and the corresponding output of our RANSAC algorithm (right). The comparison between the top pair and the bottom pair shows that changing geometry of the vehicles changes the geometry of the ellipse, where each correspondence on the ellipse is related to the angle of arrival of a sound source. The right panels show that the algorithm does a good job accepting the candidate matches (green '+') and rejecting the outlier correspondences (red 'o'). This analysis is performed for all possible combinations of a, b, and e.

1) Failure Modes

Unfortunately there are configurations when the ellipse estimation degenerates and becomes unstable. The first case is when the far field assumption is violated. For our model, this is a function of two different physical components of the environment: the inter-vehicle distances, $d_{a,b}$ and $d_{a,e}$ as well as the source positions, $\mathbf{S}$. This means our system will only find correspondences for subsets of the swarm where the vehicle triplets are close together. Another failure mode is when the vehicle triplet becomes linear, i.e. $\phi_{a,b,e}$ in equation 4 approaches 0. In this case, $\tau_{a,b,j}$ and $\tau_{a,e,j}$ trace a line instead of an ellipse and our estimation will become unstable. We throw away the information χ_{a,b,e,T_c} that produces a degenerate ellipse.

These two failure modes mean that our system only accepts correspondences that are local compared to the size of the swarm. It also means that for certain choices of a, b, and e, χ_{a,b,e,T_c} will not produce a valid ellipse and some vehicle triplets will not have any matching correspondences. Our next step is to combine the information from the good vehicle triplets to construct correspondences across as much of the swarm as possible.

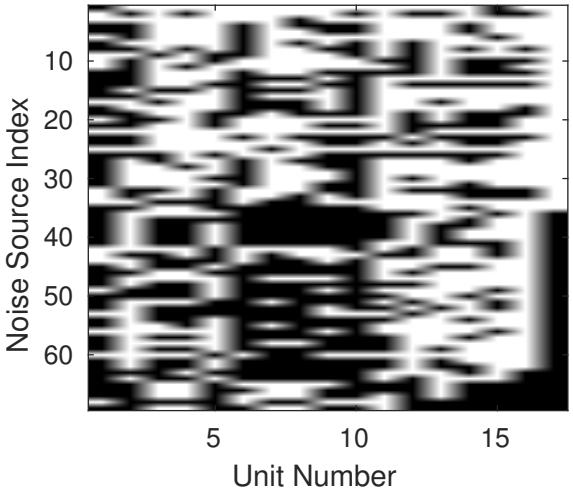

FIGURE 4: Sparsity of constraints: The x-axis labels the 17 vehicles and the y-axis the noise indices that were detected across a minimum of 4 vehicles for a specific choice of T_c. A white box indicates that the time of arrival is known for the vehicle on the x-axis and the noise index on the y-axis. These time of arrivals are used to solve our final optimization problem for the location of the vehicles and sound sources. The structure of this problem makes the optimization difficult.

2) Merging Correspondences

Once we have all the information from the sets of vehicle triplets, we combine the information to look for a full correspondence set across the entire swarm. We do this by going back to correspondence matches between pairs of vehicles instead of the vehicle triplets that we were working with before. Notice that there is redundant information when trying to find correspondences between vehicle pairs when we are using information from vehicle triplets. That is, when considering correspondences between vehicle a and b, the ellipse estimation can fail for many choices of e and we can still have some idea of correct correspondences between a and b from the values of e in χ_{a,b,e,T_c} that produced valid results.

In order to accept a correspondence between a pair of vehicles we ensure that it was accepted in at least 70 percent of the valid ellipses that considered the correspondence, and we ensure that there were at least 2 ellipses that were possible between the pairs. This protects us from an incorrect single ellipse. Finally, if there is any contradiction between any

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

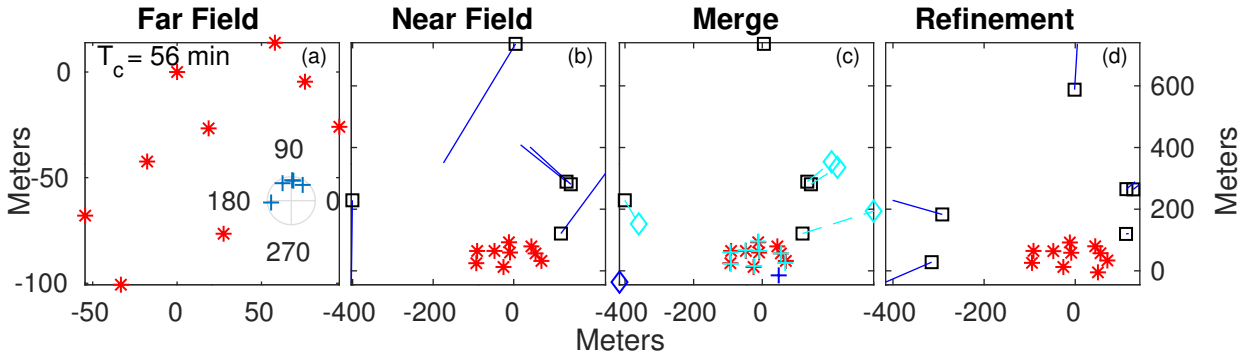

FIGURE 5: Steps of our self-localization solution: (a) A far field approximation is used to solve for an initial geometry of a subset of the swarm, $\mathbf{R}_\rho$, shown by red '*' as well as the angle of arrival of a subset of the noise sources, $\mathbf{\Lambda}_\sigma$. The angle of arrival of the noise sources relative to the swarm geometry is shown in the inset. (b) The solution is refined by relaxing the far field approximation. Noise sources that were parameterized by an angle are now parameterized by a 2D position, so we are relaxing the assumption that the noise sources are in the far field. Lines indicate the initial and final locations of each entry in $\mathbf{R}_\rho$ and $\mathbf{S}_\sigma$. (c) A new subset is chosen with different, but overlapping, values of ρ and σ. The steps in (a) are computed and the solution is fit to the previous solution. In (c), vehicles that overlap with the original subset are shown by cyan '+' whereas new ones are a dark blue '+'. Likewise, new sound sources are shown by a dark blue diamond and the corresponding noise sources are shown by a cyan diamond. (d) The initial guess of step (c) is refined with equation 2 to get a new, refined value of the swarm geometry and the noise sources. Steps (c)-(d) are iterated until there are no more values of ρ and σ.

pairs of correspondences, we throw away all of those possible correspondences.

Once the pairwise correspondences are made we assume that they are associative so that correspondences can be drawn that were not accepted by the ellipse estimation. This may accept correspondences that could violate the far field assumption. We address this in Section III-C2 by allowing noise sources to be parameterized by a two-dimensional coordinate instead of only an angle. Our last constraint is that we throw away noise events that we cannot track across a minimum of 4 vehicles. An example of correspondences for a specific choice of T_c is shown in Figure 4, where a white box indicates that the time of arrival is known for the unit indicated on the x-axis and the snap index located on the y-axis. We can see the sparsity of the optimization problem we are left with, a single noise source is rarely tracked across the entire swarm. This makes finding a global solution to the optimization problem difficult.

C. SELF-LOCALIZATION

After determining the correct correspondences, $t_{i,j}$, we construct a joint estimate of the geometry of the swarm, $\mathbf{R}$, and the relative locations of the sound sources, $\mathbf{S}$. This involves solving the minimization problem of equation 2. We do this by breaking the swarm into subsets, estimating each subset independently, then incrementally piecing subsets together to estimate a solution for the entire swarm. We choose a subset of the swarm $\rho \subseteq \{1, 2, ..., N\}$ of length k and a subset of the noise sources $\sigma \subseteq \{1, 2, ..., M\}$ of length l so that every vehicle in ρ has an arrival time for every sound in σ. Subsets are chosen so that $k, l \geq 5$. For each noise source, we choose a collection of 5 noise sources that have the greatest overlap in detected arrival times for the same vehicles (a minimum of 5 vehicles for the subset to be kept). This criterion ensures every valid noise source is included in at least one subset, it also creates redundancy in the collection of subsets. In the next section, we evaluate the stability of different subsets and discard some subsets if they do not meet specific criterion, so this redundancy helps reduce the number of noise sources that are excluded. Examples of two different subsets of noise sources and vehicles that were chosen by these criteria are shown in Figure 5. At the end of collecting subsets, we verify that every vehicle is included in at least one subset. If there are vehicles that cannot be included, we decrease the constraint of $k, l \geq 5$ to 4 and 3 but we add these subsets to the global solution last, and if there are still vehicles that cannot be included, they are not included in the final result for that choice of T_c. For each subset we start with a far field assumption and leverage the developments of Thrun [21]. From the far field approximation, we refine the positions of the sound sources by allowing the entires in $\mathbf{S}$ to be parameterized by (x, y) coordinates instead of angles. Once we have a solution for each subset, we piece them together and run through the optimization of equation 2 again. The details of each step are explained in the following subsections.

1) Far field approximation

Given a subset of the swarm, ρ, and noise sources, σ, we estimate their far field approximations $\mathbf{R}_\rho$ and $\mathbf{\Lambda}_\sigma$ solving the minimization procedure described in Section II-B. Each time we compute a solution to equation 5 we evaluate its stability. In order to keep this estimate for further processing it must meet two criteria. First we make sure that there is an adequate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

variety of angles that produced the result. If all the angles of arrival are similar then a small change in matrix $\overline{\mathbf{D}}_{\rho,\sigma}$ can produce a very different estimate of $\mathbf{R}_\rho$. We also check to make sure that the residuals of equation 5 are reasonable before we accept the result. An example of a solution to equation 5 can be found in Figure 5 (a) where the estimates of the geometry of a subset are shown by red asterisks and the angle of arrival is shown by the polar plot inset.

2) Near field refinement

Once we have estimates for $\mathbf{R}_\rho$ and $\mathbf{\Lambda}_\sigma$ we use these values as an initial estimate for equation 2 to refine the positions of the noise sources, $\mathbf{S}_\sigma$, and make small modifications to the geometry of the swarm $\mathbf{R}_\rho$, by allowing the positions of the noise sources to be parameterized by a two-dimensional position instead of an angle. While not always necessary, this can improve our estimate of the swarm geometry if the detected noise sources are not accurately parameterized by an angle after the merging correspondences step of Section III-B2. For our initial estimate of $\mathbf{S}_\sigma$ we choose $\mathbf{S}_\sigma = r\mathbf{\Lambda}_\sigma$ where r is a constant that is an estimate of how far the noise sources are from the swarm (we choose $r = 400m$). This refinement step is shown in Figure 5 (b) where the blue lines indicate the difference between the initial and final estimates of the noise sources (black squares) and the geometry of the swarm (red '*'). The refinement is done using the Levenberg-Marquardt algorithm.

3) Adding different subsets to the estimate

From a refined estimate of the swarm, our next step is to add more vehicles to the estimate. We choose a different value of ρ and σ that has the most vehicles in common with our current estimate of the swarm geometry. We then perform the far field approximation and refinement on this new subset of data. Next we find an optimal rigid body transformation [23] between the overlapping units of the new subset and current estimate to position the two estimates in the same coordinate frame. If there is not enough overlap in vehicles to determine an appropriate transformation, we use the estimated location of the noise sources to help with the rotation estimate (we do not add subsets that do not have enough combined overlap between the vehicles and noise sources). This step is shown in Figure 5 (c) where we are adding a new subset of vehicles (dark blue '+') and noise sources (dark blue diamonds) by using the overlapping vehicle estimates (shown by the red '*' and cyan '+'). Once everything is in the same coordinate frame we use this new estimate as an initial estimate for further refinement in equation 2 (Figure 5 (d)). We iterate this step until all of predetermined subsets have been added.

IV. RESULTS

The final results of our optimization procedure are shown in Figure 6 for four different times during the deployment. In this figure, the estimated geometry from the high frequency pinging system is shown by a blue circle. The self-localization estimate, $\mathbf{R}$, from the impulsive noises is shown by red asterisks. Additionally, some noise source locations,

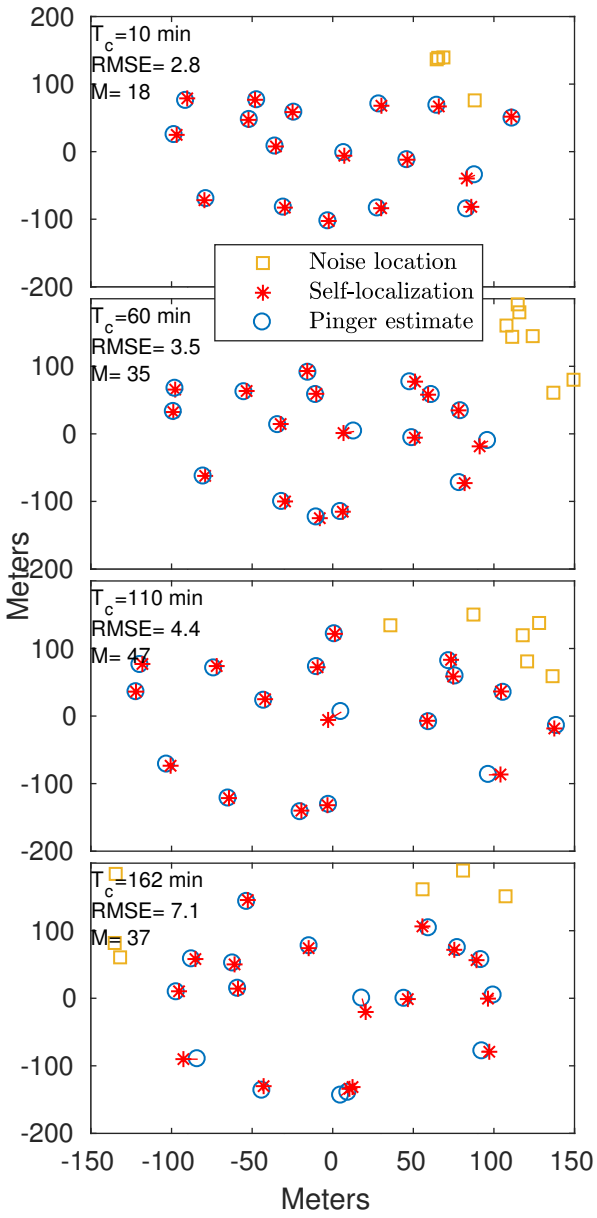

FIGURE 6: Full Solution: The full solution of our method is shown for different values of T_c (a: $T_c = 10$ min b: $T_c = 60$ min c: $T_c = 110$ min d: $T_c = 162$ min). The blue circles represent the geometry estimate from the high frequency GPS pingers and the red '*' represents our estimate using the impulsive noise sources. Yellow squares represent the estimated location of noise sources that are close to the swarm. Many noise source locations are out of frame to focus on the swarm. The number of noise sources that are estimated, M, is also shown. Also reported is the RMSE in meters between the swarm geometry estimated by the impulsive noise sources and the acoustic localization system. This Figure shows good agreement between our self-localization solution and the estimate provided by the acoustic localization system.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2779835, IEEE Access

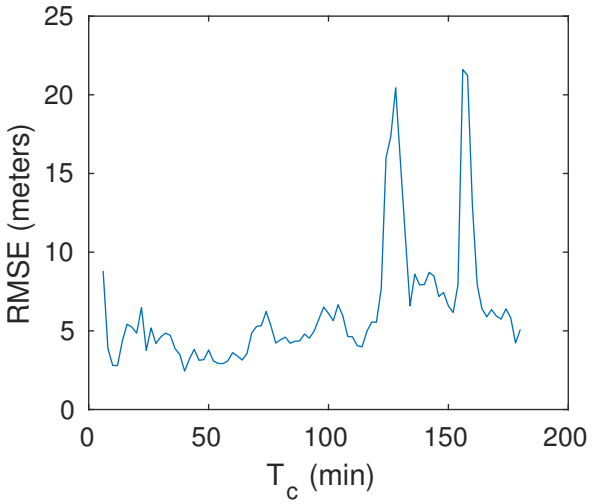

FIGURE 7: Results: RMSE, in meters, as a function of center time T_c, between a median filtered version of the self-localization solution and the estimates from the acoustic localization system. The RMSE measures the discrepancy between the two solutions. The RMSE shows good agreement for most of the deployment, validating our proposed method.

S, are shown by a yellow box, although most are outside of the frame to focus on the swarm estimate. The plots show the overlay after an optimal rigid transformation is applied between the two estimates [23], since we cannot place the estimate from the noise sources in a global coordinate frame. Also reported is the number of snaps that were used in the estimate and the root mean square error (RMSE) in meters, which measures the discrepancy between the geometry of the two estimates. For many cases, all the units are estimated and the estimate matches what is reported by the pingers. When the estimate of the swarm has a higher RMSE (see Figure 6 b-d) it is typically the same units that do not match the estimate provided by the pingers. An example is the unit that is closest to the origin, (0,0). This result may point to an inaccurate estimate from the pinging system (e.g. that a reflected arrival was detected as the direct arrival for a ping).

Figure 7 shows the RMSE, in meters, as a function of deployment time for the first 3 hours of the deployment after a median filter of length 5 is applied to the position estimates. Additionally, a video comparing the positions is provided in the supplementary material. The median filter ensures that there is an estimate of the vehicle locations for every time step and smooths solutions that arrived at a suboptimal solution. Only the first three hours are shown. After the third hour, the swarm drifted over the impulsive noises and the bottom section of the swarm moved away from the top section. These events significantly violated our far field model, described in Section III-B and III-C1, degrading the results. The RMSE is usually around 5 meters, and for periods of time it stays below 5 meters. When the RMSE does jump, it usually results from weak connections between the bottom four vehicles that are far away from the rest of the units. The "ground truth" that we are comparing our measurements to are also estimates. The expected accuracy of the pinger estimates are on the order of a few meters [5]. In this frame of reference, the results presented show the effectiveness of the proposed method.

V. RELATED WORK

In the ocean, array self-localization was demonstrated using low frequency (350-700Hz) ambient noise on a stationary array [24]. This application is built on the result that the cross correlation of ambient acoustic noise in the ocean converges to an estimate of the acoustic Time Domain Greens function (TDGF), also known as the acoustic impulse response, between two receivers [25], [26]. Using estimated inter vehicle distances from the TDGF, the authors formulate a nonlinear least squares problem to solve for both the relative clock offsets and positions of each receiver in a bottom mounted hydrophone array. Extending this solution to the mobile case is non-trivial because inter-vehicle motion changes the TDGF, making it difficult to estimate. The emergence of the TDGF theoretically requires an isotropic noise field [26], [27]. Implementations rely on long time windows to increase the coherent contributions along the endfire beam. In a sense, the longer time windows can mitigate anisotropy through averaging. Unfortunately, movement restricts the time that one can correlate over because the TDGF must be stationary throughout the duration of the window, and it is not clear that the TDGF can consistently emerge from the noise floor for mobile vehicles [15]. The variance of the noise floor was shown to be inversely proportional to the time-bandwidth product under the assumption of an isotropic noise distribution [28]. On the other hand, the signal part depends on the distribution of noise sources in the ocean and is a time varying process [29], [30], meaning that the trade-off between the noise floor and signal is environmentally dependent. There has been some work on enhancing the signal part of the noise correlation through a stochastic search using a genetic algorithm [31], which may be able to help track the changing TDGF using short time windows. This method is still susceptible to loud anisotropies in the noise field that can significantly bias the results. Other work has shown that while the TDGF may be able to be recovered for some times during the deployment, it seems unlikely that the TDGF will always be able to be recovered for all orientations with these anisotropies [15]. More expensive vector sensors can be used to increase the emergence of the TDGF by adding more directionality in the processing [32] but these sensors are expensive, difficult to mount and there may be environmental cases that still render the TDGF unusable.

In terrestrial wireless sensor networks, studies have focused on ad hoc self-localization of microphone arrays. Many scenarios, constraints and sensor packages have been considered and a review can be found in Plinge et al. [33]. The studies that are most similar to this work consider only one receiver per node, assume no knowledge of the ambient

sound positions or their emittance times and assume that the receivers are time synchronized. Some of these are similar to the work of Sabra et al. [28] in that they either assume a diffuse noise field [34] or rely on sources being in the endfire beam [35] so that they can directly estimate the distance between receivers. These works suffer from the problems discussed in the last paragraph; with short time windows the probability of either a diffuse noise field or the probability that a noise source passes through the endfire beam of all receiver pairs is low. Under a far field assumption, the work of Thrun [21] showed that the relative geometry of an ad hoc microphone array could be robustly estimated in an affine space and later upgraded to a euclidean space when the noise sources are parameterized by an angle. This method only requires a nonlinear optimization on a space that is P^2 variables, where P is the dimension that the receivers occupy. Therefore, this method is less prone to local minima than solutions that try to solve a least squares problem over all sound and receiver positions. The work of Thrun [21] was extended to a three-dimensional case [36]. Typically, these works assume that the correspondence problem is solved, meaning that there are unambiguous correspondences of events between audio tracks of the receivers. In a real life ocean environment, this assumption is impractical for most frequency bands. The correspondence problem is exasperated the more nodes one considers at the same time. An approach to solving the correspondence problem is to look at minimal problems, where the correspondence ambiguity is the smallest. In Wendeberg et al. [20], the authors consider a two-dimensional localization problem and describe a receiver triangle. Under a far field assumption, they show that the TDOA measurements between the three receivers with one receiver as anchor form an ellipse. They use this information to extract distance and angle measurements from the receiver triangle and estimate the geometry of an array of receivers. For the three-dimensional case, minimal cases have been examined both in the far field, [37] and without a far field assumption [38]. These approaches look at a minimum of 4 and 5 receivers respectively.

Using low frequency noise is beneficial because it is ubiquitous in the ocean, but the drawback is that there can be large anisotropies that can bias the result of the TDGF and prevent accurate localization estimates. The impulses from snapping shrimp are a noise source that dominate higher frequencies in warm coastal waters [16] and have been leveraged for remote sensing underwater [17]. Most of the work using impulsive noise has been done in acoustic daylight imaging using special purpose hardware like the Acoustic Daylight Ocean Noise Imaging System (ADONIS) [39] and the second generation Remotely Operated Mobile Ambient Noise Imaging System (ROMANIS) [19]. The acoustic daylight imaging literature is interesting to us because they also must solve a correspondence problem between impulsive noise events. However, the array geometry is known and the element spacing is close, on the order of a wavelength, simplifying this correspondence problem.

VI. CONCLUSION

In this work we looked at the self-localization problem of an untethered deforming swarm of subsurface vehicles. We assumed that the swarm was stationary for short periods of time and estimated the relative geometry at each time step. Our sources of opportunity were impulsive, giving us good time resolution but providing a significant correspondence problem. We proposed a method to solve this correspondence problem by assuming that the noise sources could be modeled by a plane wave and examining the correspondences 3 vehicles at a time. We accepted correspondences that agreed with a geometric model. From these correspondences, an incremental solution was developed that first solved local subproblems and then merged those subproblems into a global estimate of the swarm geometry. Our approach was able to accurately track the deformation of the swarm as a function of time, verified by an independent, accurate localization procedure.

VII. ACKNOWLEDGEMENTS

The authors would like to thank the members of the Jaffe lab that were part of the data collection and AUE tests, and the reviewers for their feedback.

REFERENCES

[1] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," Journal of Field Robotics, vol. 27, no. 1, pp. 21–51, 2010. [Online]. Available: http://dx.doi.org/10.1002/rob.20324

[2] A. Kim and R. M. Eustice, "Real-time visual slam for autonomous underwater hull inspection using visual saliency," IEEE Transactions on Robotics, vol. 29, no. 3, pp. 719–733, June 2013.

[3] R. Stokey, T. Austin, B. Allen, N. Forrester, E. Gifford, R. Goldsborough, G. Packard, M. Purcell, and C. von Alt, "Very shallow water mine countermeasures using the remus auv: a practical approach yielding accurate results," in OCEANS, 2001. MTS/IEEE Conference and Exhibition, vol. 1. IEEE, 2001, pp. 149–156.

[4] A. Hackbarth, E. Kreuzer, and E. Solowjow, "Hippocampus: A micro underwater vehicle for swarm applications," in Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. IEEE, 2015, pp. 2258–2263.

[5] J. S. Jaffe, P. J. S. Franks, P. L. D. Roberts, D. Mirza, C. Schurgers, R. Kastner, and A. Boch, "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics," Nature Communications, vol. 8, no. 14189, 2017, article. [Online]. Available: http://dx.doi.org/10.1038/ncomms14189

[6] T. Schmickl, R. Thenius, C. Moslinger, J. Timmis, A. Tyrrell, M. Read, J. Hilder, J. Halloy, A. Campo, C. Stefanini, L. Manfredi, S. Orofino, S. Kernbach, T. Dipper, and D. Sutantyo, "Cocoro – the self-aware underwater swarm," in 2011 Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops, Oct 2011, pp. 120–126.

[7] H.-P. Tan, R. Diamant, W. K. Seah, and M. Waldmeyer, "A survey of techniques and challenges in underwater localization," Ocean Engineering, vol. 38, no. 14, pp. 1663 – 1676, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0029801811001624

[8] J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb, "A survey of underwater vehicle navigation: Recent advances and new challenges," in IFAC Conference of Manoeuvering and Control of Marine Craft, vol. 88, 2006, pp. 1–12.

[9] M. M. Hunt, W. M. Marquet, D. A. Moller, K. R. Peal, W. K. Smith, and R. C. Spindel, "An acoustic navigation system," Woods Hole Oceanographic Institution, Tech. Rep., 1974.

[10] R. McEwen, H. Thomas, D. Weber, and F. Psota, "Performance of an auv navigation system at arctic latitudes," IEEE Journal of Oceanic Engineering, vol. 30, no. 2, pp. 443–454, April 2005.

[11] B. Allen, R. Stokey, T. Austin, N. Forrester, R. Goldsborough, M. Purcell, and C. von Alt, "Remus: a small, low cost auv; system description, field trials and performance results," in OCEANS '97. MTS/IEEE Conference Proceedings, vol. 2, Oct 1997, pp. 994–1000 vol.2.

[12] L. Freitag, M. Grund, S. Singh, J. Partan, P. Koski, and K. Ball, "The whoi micro-modem: an acoustic communications and navigation system for multiple platforms," in Proceedings of OCEANS 2005 MTS/IEEE, Sept 2005, pp. 1086–1092 Vol. 2.

[13] L. E. Freitag, M. Grund, J. Partan, K. Ball, S. Singh, and P. Koski, "Multi-band acoustic modem for the communications and navigation aid auv," in Proceedings of OCEANS 2005 MTS/IEEE, Sept 2005, pp. 1080–1085 Vol. 2.

[14] D. Mirza, P. Naughton, C. Schurgers, and R. Kastner, "Real-time collaborative tracking for underwater networked systems," Ad Hoc Networks, 2014.

[15] P. Naughton, P. Roux, R. Yeakle, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, "Ambient noise correlations on a mobile, deformable array," The Journal of the Acoustical Society of America, vol. 140, no. 6, pp. 4260–4270, 2016. [Online]. Available: http://dx.doi.org/10.1121/1.4971172

[16] F. A. Everest, R. W. Young, and M. W. Johnson, "Acoustical characteristics of noise produced by snapping shrimp," The Journal of the Acoustical Society of America, vol. 20, no. 2, pp. 137–142, 1948. [Online]. Available: http://dx.doi.org/10.1121/1.1906355

[17] J. R. Potter, T. Lim, and M. Chitre, "Ambient noise environments in shallow tropical seas and the implications for acoustic sensing," in Oceanology International, vol. 97, 1997, pp. 2114–2117.

[18] M. W. Legg, A. J. Duncan, A. Zaknich, and M. V. Greening, "Analysis of impulsive biological noise due to snapping shrimp as a point process in time," in OCEANS 2007-Europe. IEEE, 2007, pp. 1–6.

[19] M. Chitre, S. Kuselan, and V. Pallayil, "Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation," The Journal of the Acoustical Society of America, vol. 132, no. 2, pp. 838–847, 2012. [Online]. Available: http://dx.doi.org/10.1121/1.4733553

[20] J. Wendeberg, T. Janson, and C. Schindelhauer, "Self-localization based on ambient signals," Theoretical Computer Science, vol. 453, pp. 98–109, 2012.

[21] S. Thrun, "Affine structure from sound," in Advances in Neural Information Processing Systems 18 Neural Information Processing Systems, 2005, pp. 1353–1360. [Online]. Available: http://papers.nips.cc/paper/2770-affine-structure-from-sound

[22] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM, vol. 24, no. 6, pp. 381–395, Jun. 1981. [Online]. Available: http://doi.acm.org/10.1145/358669.358692

[23] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-9, no. 5, pp. 698–700, Sept 1987.

[24] K. G. Sabra, P. Roux, A. M. Thode, G. L. D'Spain, W. S. Hodgkiss, and W. A. Kuperman, "Using ocean ambient noise for array self-localization and self-synchronization," IEEE Journal of Oceanic Engineering, vol. 30, no. 2, pp. 338–347, April 2005.

[25] P. Roux, W. A. Kuperman, and the NPAL Group, "Extracting coherent wave fronts from acoustic ambient noise in the ocean," The Journal of the Acoustical Society of America, vol. 116, no. 4, pp. 1995–2003, 2004. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/116/4/10.1121/1.1797754

[26] P. Roux, K. G. Sabra, W. A. Kuperman, and A. Roux, "Ambient noise cross correlation in free space: Theoretical approach," Journal of the Acoustical Society of America, vol. 117, no. 1, pp. 79–84, Jan 2005. [Online]. Available: <Go to ISI>://WOS:000226334800007

[27] P. Roux, W. Kuperman, and K. G. Sabra, "Ocean acoustic noise and passive coherent array processing," Comptes rendus Geoscience, vol. 343, no. 8, pp. 533–547, 2011.

[28] K. G. Sabra, P. Roux, and W. A. Kuperman, "Emergence rate of the time-domain green's function from the ambient noise cross-correlation function," The Journal of the Acoustical Society of America, vol. 118, no. 6, pp. 3524–3531, 2005. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/118/6/10.1121/1.2109059

[29] C. Leroy, S. Lani, K. G. Sabra, W. S. Hodgkiss, W. A. Kuperman, and P. Roux, "Enhancing the emergence rate of coherent wavefronts from ocean ambient noise correlations using spatio-temporal filters," The Journal of the Acoustical Society of America, vol. 132, no. 2, pp. 883–893, 2012. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/132/2/10.1121/1.4731231

[30] S. E. Fried, W. A. Kuperman, K. G. Sabra, and P. Roux, "Extracting the local green's function on a horizontal array from ambient ocean noise," The Journal of the Acoustical Society of America, vol. 124, no. 4, pp. EL183–EL188, 2008. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/124/4/10.1121/1.2960937

[31] K. F. Woolfe, K. G. Sabra, and W. Kuperman, "Optimized extraction of coherent arrivals from ambient noise correlations in a rapidly fluctuating medium," The Journal of the Acoustical Society of America, vol. 138, no. 4, pp. EL375–EL381, 2015.

[32] B. Nichols and K. G. Sabra, "Cross-coherent vector sensor processing for spatially distributed glider networks," The Journal of the Acoustical Society of America, vol. 138, no. 3, pp. EL329–EL335, 2015. [Online]. Available: http://dx.doi.org/10.1121/1.4929615

[33] A. Plinge, F. Jacob, R. Haeb-Umbach, and G. A. Fink, "Acoustic microphone geometry calibration: An overview and experimental evaluation of state-of-the-art algorithms," IEEE Signal Processing Magazine, vol. 33, no. 4, pp. 14–29, July 2016.

[34] I. McCowan, M. Lincoln, and I. Himawan, "Microphone array shape calibration in diffuse noise fields," IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 3, pp. 666–670, March 2008.

[35] R. Biswas and S. Thrun, "A distributed approach to passive localization for sensor networks," in In Proceedings of the National Conference on Artificial Intelligence (Vol. 20, No. 3, p. 1248)., 2005.

[36] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, March 2008, pp. 2445–2448.

[37] S. Burgess, Y. Kuang, J. Wendeberg, K. Åström, and C. Schindelhauer, "Minimal solvers for unsynchronized tdoa sensor network calibration," in International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics. Springer, 2013, pp. 95–110.

[38] S. Zhayida, S. S. Rex, Y. Kuang, F. Andersson, and K. Åström, "An automatic system for acoustic microphone geometry calibration based on minimal solvers," arXiv preprint arXiv:1610.02392, 2016.

[39] J. R. Potter and M. Chitre, "Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms," The Journal of the Acoustical Society of America, vol. 106, no. 6, pp. 3201–3210, 1999. [Online]. Available: http://dx.doi.org/10.1121/1.428174

Perry Naughton is a Ph.D. candidate in the Electrical and Computer Engineering department at the University of California, San Diego where he also earned his Master's degree (2014) and his Bachelor of Science (2012)

He was a visiting Ph.D. student at the Institut des Sciences de la Terre, Université Grenoble-Alpes, for the 2017 academic year. His research interests include underwater acoustics, localization, computer vision and signal processing. He is a recipient of the National Science Foundation Graduate Research Fellowship.

Philippe Roux received the Ph.D. degree in the application of time reversal to ultrasounds from the University of Paris, Paris, France, in 1997.

He is a Physicist with a strong background in ultrasonics and underwater acoustics. Between 2002 and 2005, he was a Research Associate in the Marine Physical Laboratory, Scripps Institution of Oceanography, San Diego, CA, USA. He is now a full-time Research Director at the Institut des Sciences de la Terre, Centre National de la Recherche Scientifique, Université Grenoble-Alpes, France, where he develops small-scale laboratory experiments in geophysics. He is a Fellow of the Acoustical Society of America. He received the Medwin Prize in Acoustical Oceanography in 2013.

Curt Schurgers received his Ph.D. from UCLA in Integrated Circuits and Systems in 2002, and his M.S. from the Katholieke Universiteit Leuven (Belgium) in 1997.

He has also worked as a researcher at IMEC, Belgium (1997-1999), and at MIT (2003). Currently, he is a professor at the Electrical and Computer Engineering Department at the University of California, San Diego.

Ryan Kastner is a professor in the Department of Computer Science and Engineering at the University of California, San Diego. He received a PhD in Computer Science (2002) at UCLA, a Masters (2000) and Bachelors (1999) in both Electrical Engineering and Computer Engineering, all from Northwestern University. He spent the first five years after his PhD as a professor in the Department of Electrical and Computer Engineering at the University of California, Santa Barbara.

His current research interests fall into three areas: hardware acceleration, hardware security, and remote sensing. He is the co-director of the Wireless Embedded Systems Master of Advanced Studies Program. He also co-directs the Engineers for Exploration Program. He has published over 150 technical articles, and has authored three books, "Synthesis Techniques and Optimizations for Reconfigurable Systems", "Arithmetic Optimizations for Polynomial Expressions and Linear Systems", and "Handbook on FPGA Design Security". He has served as member of numerous conference technical committees spanning topics like reconfigurable computing (ISFPGA, FPL, FPT), hardware design (DAC, ICCAD, DATE), hardware security (HOST), and underwater networking (WUWNet).

Jules Jaffe received the B.A. degree in physics from the State University of New York at Buffalo, Buffalo, NY, USA, in 1973, the M.S. degree in biomedical information science from the Georgia Institute of Technology, Atlanta, GA, USA, in 1974, and the Ph.D. degree in biophysics from the University of California, Berkeley in 1982.

He subsequently spent several years in Silicon Valley as an image-processing consultant. In 1984, he joined the Woods Hole Oceanographic Institution, Woods Hole, MA, USA, where he worked as an Assistant and Associate Scientist. In 1988, he joined the Scripps Institution of Oceanography, University of California at San Diego, as an Assistant Research Oceanographer. He was subsequently promoted to Associate and then Full Oceanographer, which is his current position. His lab specializes in the development of a variety of underwater technology that ranges from miniature vehicles to underwater optical and acoustical instruments. In 2003, he was a H. Burr Steinbuck Visiting Scholar at the Woods Hole Oceanographic Institution. In 2006, he was a Visiting Miller Professor at the University of California, Berkeley. Dr. Jaffe is a Fellow of the Acoustical Society and a past recipient of a National Science Foundation Creativity Award. In 2012, he won a "best paper award" at the International Ocean Optics Meeting. He is currently an Associate Editor of the IEEE Journal of Oceanic Engineering, and a past Editor-in-Chief of Methods in Oceanography (Elsevier).

Paul Roberts received a B.S. in Computer Engineering (2002), M.S. in Electrical Engineering (2004), and Ph.D. in Electrical Engineering (Applied Ocean Science) (2009) all from the University of California, San Diego.

He has worked as a development engineer in the Jaffe Lab since 2010. His main interests are underwater microscopy, underwater imaging, oceanographic instrument design, signal and image processing, computer vision, and machine learning.