

Self-localization of a mobile swarm using noise correlations with local sources of opportunity

Perry Naughton^{a)}

*Department of Electrical and Computer Engineering, University of California, San Diego,
9500 Gilman Drive, La Jolla, California 92093, USA*

Philippe Roux

Institut des Sciences de la Terre, UMR 5275, Université Grenoble-Alpes, 38000 Grenoble, France

Curt Schurgers

*Department of Electrical and Computer Engineering, University of California, San Diego,
9500 Gilman Drive, La Jolla, California 92093, USA*

Ryan Kastner

*Department of Computer Science and Engineering, University of California, San Diego, 9500 Gilman Drive,
La Jolla, California 92093, USA*

Jules S. Jaffe and Paul L. D. Roberts

*Marine Physical Lab., Scripps Oceanography, UC San Diego, 9500 Gilman Drive, La Jolla, California 92093,
USA*

(Received 9 May 2018; revised 26 September 2018; accepted 17 October 2018; published online 16 November 2018)

Groups of coordinated underwater vehicles or sensors are powerful tools for monitoring the ocean. A requirement of many coordinated surveys is to determine a spatial reference between each node in a swarm. This work considers the self-localization of a swarm of independently moving vehicles using acoustic noise from a dominating incoherent source recorded by a single hydrophone onboard each vehicle. This method provides an inexpensive and infrastructure-free spatial reference between vehicles. Movement between the vehicles changes the swarm geometry and a self-localization estimate must be generated from data collected on short time scales. This challenges past self-localization approaches for acoustic arrays. To overcome this challenge, the proposed self-localization algorithm jointly estimates the vehicle geometry and the directionality of the ambient noise field, without prior knowledge of either estimate. To demonstrate this method, experimental results are provided when a boat is the main dominating source. The results demonstrate the ability to both estimate the direction of arrival of the boat and the relative positions of the vehicles in the swarm. The approach in this paper is not limited to moving vessels. Simulations are provided to examine three different factors that affect the proposed solution: inter-vehicle motion, vehicle geometry, and the azimuthal variance of the noise field. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5070154>

[KGS]

Pages: 2811–2823

I. INTRODUCTION

Swarms of underwater vehicles provide a new modality for passive sensing in ocean environments.¹ Swarms can measure the spatial distribution of different chemicals or physical processes in the ocean,² freely drifting vehicles can measure three dimensional water flows to study phenomena such as internal waves in physical oceanography,¹ and using multiple vehicles has been proposed to optically image the seafloor.³ Like underwater arrays, the received structure of a signal across elements in a swarm can be used to determine information about sound sources and environments.^{4,5} Swarms of vehicles have the added advantages of being mobile, easy to deploy, and robust since every vehicle is its own independent system.

Estimating the individual location of each node in a swarm is necessary for most deployments. Localizing

underwater vehicles is usually based on active acoustics, which contain trade-offs in power consumption, cost, and required infrastructure.^{6,7} The localization could be performed onboard each vehicle by using Doppler shifts^{8,9} and inertial navigation systems, or by acoustic communication.¹⁰ Performing the localization onboard each vehicle is power intensive and requires expensive equipment. This reduces the endurance and increases the price point of each vehicle. Alternatively, surveys can set up dedicated acoustic infrastructure in a global coordinate frame to send signals to each vehicle.^{1,10–12} Travel time measurements from the infrastructure to the vehicle provide distance estimates, and multiple distance estimates can be combined to estimate the position of each vehicle. In these applications, the acoustic infrastructure can require significant effort to deploy and forces the vehicles to remain within its range, limiting the area the vehicles can survey. Using active acoustic systems for localization is restrictive in either cost, time, endurance, or mobility.

^{a)}Electronic mail: pnaughto@ucsd.edu

A promising alternative to active localization systems is to use the ambient acoustic noise in the ocean. Ambient acoustic noise is ubiquitous,¹³ free, and does not require an active source; this makes localization independent of infrastructure, inexpensive, and low power. *This article provides a self-localization estimate for moving vehicles using ambient acoustic noise from dominating noise sources in the ocean recorded by a single hydrophone onboard each vehicle.* Movement between vehicles forces the self-localization solution to use constraints collected on short time scales (on the order of seconds), which complicates previously proposed self-localization methods focused on ambient noise. The proposed solution assumes that the vehicle geometry is stationary for short periods of time, and that there are dominant incoherent sound sources in the soundscape, to jointly estimate the geometry of the vehicles along with the direction of the dominant sources. This approach reduces the time needed for a self-localization solution by utilizing strong anisotropies in the noise field and can accommodate many different environments. These strong anisotropies are usually considered a nuisance when trying to extract information from ambient noise.⁵

There has been a significant amount of work on processing low frequency (100–1000 Hz) acoustic noise in the ocean because it is omnipresent and suffers the least from attenuation.¹⁴ In this frequency band, an array element self-localization solution was proposed.^{15,16} This self-localization solution requires contributions of the noise field to come from all directions, which allows an estimate of the acoustic impulse response to be recovered between all pairs of receivers by correlating the noise recorded at each receiver.^{17–20} This “isotropic” assumption is usually upheld by considering long time windows of the ambient noise field to accumulate as many different noise sources as possible, hopefully coming from different directions.^{21–23} In order for this accumulation to work, the propagation path must remain stationary between the receivers during the averaging process. Traditionally, the time evolution of the propagation medium in the ocean limits the averaging time to tens of minutes or hours.^{15,21} When considering arrays, the length of time needed can be shortened by leveraging coherent array processing through spatio-temporal filters.^{22,23} This is not applicable to the scenario we are considering. The *ad hoc* array created by mobile vehicles is sparse with uncertain and changing receiver positions and independent, possibly unsynchronized clocks. These characteristics impede coherent processing. In the proposed work, the movement between vehicles is the dominant change in the propagation path and is more restrictive on the length of noise record than what has previously been considered for recovering the acoustic impulse response (on the order of seconds, instead of tens of minutes).⁵ With this restriction, it is not clear that the acoustic impulse response can be estimated between moving vehicles, and self-localization approaches that rely on these estimates may not be robust.

Despite the tough restrictions on the length of time that can be used, noise correlations across mobile vehicles do exhibit arrivals from different noise directions.⁵ Without noise coming from all directions, or when there are significant incoherent dominant sources, this arrival structure is

biased from the dominant noise directions in the ocean environment. Consequently, accurate distance estimates between individual vehicles cannot be estimated and only the receiver separation projected onto a dominant noise direction can be retrieved. This projected information is used in the proposed self-localization solution because it does not require the long time averages required by an isotropic distribution. Similar self-localization problems have been examined in terrestrial wireless sensor networks to recover an array geometry using projected distances from many different sources.¹⁹ Under a far field assumption, the work of Thrun²⁵ 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 noise sources are parameterized by an angle. Similar works provided solutions for near-field and three-dimensional cases,^{24,26–28} which are more appropriate for indoor applications where noise sources and microphones are in the same vicinity.

For mobile underwater vehicles, a self-localization solution is required that jointly estimates the vehicle positions and the dominant noise directions of the ambient noise field because recovering an estimate of the acoustic impulse response using short time windows may be unrealistic for some environments. We propose a robust method to perform this estimation and evaluate our method using data collected from a multi vehicle survey where each vehicle independently drifts with subsea currents. We show that we are able to accurately estimate the vehicle geometry as well as the direction of arrival of a moving boat, which provides a high signal-to-noise ratio (SNR) incoherent source that dominates the ambient soundscape. Simulations are provided to examine three different factors that affect the proposed solution: inter-vehicle motion, vehicle geometry, and the azimuthal variance of the noise field.

The rest of this article is organized as follows. In Sec. II, the theory behind our self-localization solution is described. Section III describes the at sea experiment, and Sec. IV describes the data analysis, including the detection of dominant sources using noise correlations, which is the input to the processing presented in Sec. II. The proposed method is tested on data that was collected off of the coast of San Diego, CA, where low frequency noise was introduced by a small boat recording its global positioning system (GPS) location. The results provided in Sec. V show that the estimation is able to accurately recover the relative positions of slowly moving vehicles and describe the angle of arrival of the moving vessel. Simulations are provided to examine the effects of different relevant parameters (e.g., inter-vehicle movement, vehicle geometry, noise directionality) on the estimation result in Sec. VI. This article is concluded in Sec. VII.

II. METHODS

A. Problem description

We consider a group of underwater vehicles that have a single hydrophone on board each vehicle and are close enough to each other that they record common noises in the ambient soundscape. The deployments considered in Refs. 1, 4, and 5 fit this requisite. We are able to determine their

depth using onboard pressure sensors and are interested in estimating their horizontal positions *relative* to one another. We can only solve for the positions of the vehicle up to a rigid transformation (i.e., a translation and a rotation) because we do not assume to know the locations of any noise sources or vehicle positions. This is useful for deployments that aim to study spatial data in a relative sense (e.g., chemicals, currents, or optical images), or to measure the time difference of arrivals (TDOAs) of acoustic signals that are received by the elements of the swarm. Additionally, relative localization solutions can be augmented with a global reference when it is available, such as a GPS location when the vehicle surfaces.

We estimate the relative geometry of the vehicle swarm using cross-correlations of low frequency noise in the ocean and assume that this noise is coming from far away dominant incoherent sources. In our experiment, the low frequency noise is dominated by a moving vessel, so the developments in this section reflect a single moving sound source. However, our approach is not limited to this scenario and can be modified to include multiple dominating sources. Examples of sources that could be used are low frequency noise from shore or noise from biological sources.¹⁵ Naughton *et al.*⁵ shows examples of noise correlations that result from multiple dominating sources (e.g., shipping noise and noise from the coast), that are common in coastal environments.

The outline for our procedure is shown in Fig. 1. We start with acoustic data from independent vehicles and use noise correlations to extract TDOA constraints between vehicles. Our experiment and TDOA extraction methods are reported in Secs. III and IV, respectively. With these constraints, the

estimation described in Secs. IIC and IID is performed and then refined by the procedure described in Sec. IIE.

B. Definitions

We define the two dimensional position of vehicle i , at time t , as $R_{i,t}$, the collection of vehicle positions at time t as a matrix, \mathbf{R}_t , and the two dimensional position of the dominant sound source at time t as $S_t = [S_{x,t}, S_{y,t}]$. We can arbitrarily assign the x, y position of any vehicle (in this case the first vehicle), to $R_{1,t} = [0,0]$ because the solution can only be solved in a relative coordinate frame. Throughout the text we will refer to the vehicle defined at the origin as the anchor vehicle. With this assignment, we define a relative position matrix for n vehicles as

$$\bar{\mathbf{R}}_t = \begin{bmatrix} x_{2,t} - x_{1,t} & y_{2,t} - y_{1,t} \\ x_{3,t} - x_{1,t} & y_{3,t} - y_{1,t} \\ \vdots & \vdots \\ x_{n,t} - x_{1,t} & y_{n,t} - y_{1,t} \end{bmatrix} = \begin{bmatrix} x_{2,t} & y_{2,t} \\ x_{3,t} & y_{3,t} \\ \vdots & \vdots \\ x_{n,t} & y_{n,t} \end{bmatrix}.$$

We assume that S_t is far from the vehicle swarm so that the propagation of noise from S_t to every vehicle $R_{i,t}$ can be modeled by a plane wave and parameterized by a single angle, α_t . Typically, this assumption holds as long as the sound source is separated from the swarm by about 2–3 times the swarm’s aperture.²⁵ We also assume that S_t is moving so that $S_t \neq S_{t+\epsilon}$ for some small value of ϵ . This gives us two collections of variables we wish to estimate, the time evolution of the angles from the dominant noise source, α_t , and the time evolution of the relative position matrix, $\bar{\mathbf{R}}_t$.

C. Solution

The estimation is based on TDOA measurements. The time difference of arrival between vehicle i and j for the noise source at time t is defined as

$$\delta_{i,j,t} = \frac{\|R_{i,t} - S_t\|_2}{c} - \frac{\|R_{j,t} - S_t\|_2}{c}. \quad (1)$$

Here, $\|\cdot\|_2$ denotes Euclidean distance, and c is the speed of sound underwater. By only looking at the cases where $j = 1$, remembering that we set $R_{1,t} = [0,0]$ and incorporating the assumption that the source can be modeled by an angle, Eq. (1) can be rewritten as

$$\delta_{i,1,t} \approx \frac{[\bar{R}_{i,t}^x \ \bar{R}_{i,t}^y] [\cos(\alpha_t) \sin(\alpha_t)]^T}{c}. \quad (2)$$

Equation (2) describes a constraint on the relative geometry, $\bar{\mathbf{R}}_t$, but it is only a constraint along the angle α_t . Therefore, estimating the relative geometry in two dimensions is underdetermined when considering only a constraint along a single direction, α_t . In order to estimate the direction perpendicular to α_t , the estimation must consider angles that describe the component perpendicular to α_t . In other words, the estimation must include at least one more angle that is sufficiently different from α_t to provide constraints along the

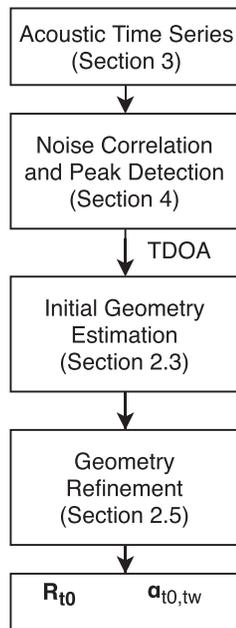


FIG. 1. Outline of our processing pipeline: We start with raw acoustic data from each individual vehicle. The vehicles and the deployment are described in Sec. III. We compute noise correlations between the vehicles and extract TDOA values from the noise correlations, as described in Sec. IV. Finally, we perform the estimation described in Sec. II by first estimating the relative positions using the formulation of Secs. IIC and IID and then refining the positions using the procedure described in Sec. IIE.

component perpendicular to α_t . This requires assumptions on both the motion of the sound source, S_t , and the relative positions of the vehicles, $\bar{\mathbf{R}}_t$. These assumptions ensure an overdetermined set of constraints by measuring the relative geometry of the swarm from sufficiently diverse angles. The first assumption is that the relative geometry remains unchanged ($\bar{\mathbf{R}}_{t_0} \approx \bar{\mathbf{R}}_{t_0 \pm t_w}$) for some time window, t_w , so that only the relative positions of the swarm at the center time, $\bar{\mathbf{R}}_{t_0}$, need to be estimated. This is an assumption on the *relative* movement between the vehicles, which is usually much smaller than the vehicle movement on a global scale, especially for vehicles that are drifting with currents or moving in a formation. We have seen this assumption on the vehicle geometry holds true for the presented experiment as well as in other deployments.^{4,5} The second assumption is that the source, S_t , will provide a collection of angles $\alpha_{t_0-t_w:t_0+t_w}$ that are sufficiently different from the center angle, α_{t_0} , so that

the geometry of the swarm can be interrogated by different directions. Together, these two assumptions provide an overdetermined system of constraints for a self-localization solution.

This defines the approach. Given a time window, $t = t_0 - t_w: t_0 + t_w$, the estimation will produce an estimate of the vehicle geometry at the center time, $\bar{\mathbf{R}}_{t_0}$ as well as the collection of angles, $\alpha_{t_0-t_w:t_0+t_w}$, which we will denote α_{t_0, t_w} for simplicity. The constraints of Eq. (2) set up a minimization problem to estimate the angle of arrival of the source at each time, and the relative geometry of the vehicles²⁵

$$\operatorname{argmin}_{\bar{\mathbf{R}}_{t_0}, \Lambda_{t_0, t_w}} \|\bar{\mathbf{R}}_{t_0} \Lambda_{t_0, t_w} - c \mathbf{D}_{t_0, t_w}\|^2. \quad (3)$$

Here, Λ_{t_0, t_w} describes the constraints placed by the angles, α_t ,

$$\Lambda_{t_0, t_w} = \begin{bmatrix} \cos(\alpha_{t_0-t_w}) & \cos(\alpha_{t_0-t_w+1}) & \cdots & \cos(\alpha_{t_0}) & \cdots & \cos(\alpha_{t_0+t_w-1}) & \cos(\alpha_{t_0+t_w}) \\ \sin(\alpha_{t_0-t_w}) & \sin(\alpha_{t_0-t_w+1}) & \cdots & \sin(\alpha_{t_0}) & \cdots & \sin(\alpha_{t_0+t_w-1}) & \sin(\alpha_{t_0+t_w}) \end{bmatrix},$$

and \mathbf{D}_{t_0, t_w} is a collection of the TDOAs compared to vehicle 1, $\delta_{i,1,t}$. Since we are comparing to the same vehicle, we drop the subscript of 1 (i.e. $\delta_{i,1,t} = \delta_{i,t}$) and define \mathbf{D}_{t_0, t_w} as

$$\mathbf{D}_{t_0, t_w} = \begin{bmatrix} \delta_{2,t_0-t_w} & \delta_{2,t_0-t_w+1} & \cdots & \delta_{2,t_0} & \cdots & \delta_{2,t_0+t_w} \\ \delta_{3,t_0-t_w} & \delta_{3,t_0-t_w+1} & \cdots & \delta_{3,t_0} & \cdots & \delta_{3,t_0+t_w} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \delta_{n,t_0-t_w} & \delta_{n,t_0-t_w+1} & \cdots & \delta_{n,t_0} & \cdots & \delta_{n,t_0+t_w} \end{bmatrix}.$$

In the noise free case, \mathbf{D}_{t_0, t_w} should be at most rank 2 because it is the product of two rank 2 matrices ($\bar{\mathbf{R}}_{t_0}^*$, Λ_{t_0, t_w}^*). Therefore, an affine solution to Eq. (3) can be found using a singular value decomposition of the matrix \mathbf{D}_{t_0, t_w} ,

$$UVW^T = \mathbf{D}_{t_0, t_w}, \quad (4)$$

and only considering the first two singular values and vectors for the estimation of $\bar{\mathbf{R}}_{t_0}$ and Λ_{t_0, t_w} ,

$$\bar{\mathbf{R}}_{t_0} = U'V' \quad \text{and} \quad \Lambda_{t_0, t_w} = W'^T. \quad (5)$$

In other words, U' is an $(n-1) \times 2$ matrix consisting of the first two left singular vectors (where n is the number of vehicles), W' is an $(2 * t_w) \times 2$ matrix consisting of the first two right singular vectors (where t_w controls the number of time samples used in the estimation) and V' is a diagonal matrix consisting of the two largest singular values. The product of $\bar{\mathbf{R}}_{t_0}$ and Λ_{t_0, t_w} from Eq. (5) gives an optimal rank 2 approximation to \mathbf{D}_{t_0, t_w} , and an optimal affine solution to the minimization problem of Eq. (3). However, this solution is not unique and given any invertible matrix, C ,

$$\bar{\mathbf{R}}_{t_0}^* = U'V'C^{-1} \quad \text{and} \quad \Lambda_{t_0, t_w}^* = CW'^T, \quad (6)$$

is also a solution to Eq. (3). In general, an affine solution will not uphold the Euclidean constraint that $\cos(\alpha_i)^2 + \sin(\alpha_i)^2 = 1$, so we choose C such that this discrepancy is minimized²⁵

$$C^* = \operatorname{arg} \min_C \|[1 \ 1] \Lambda_{t_0, t_w}^* \cdot \Lambda_{t_0, t_w}^* - [1 \ 1 \ \cdots \ 1]\|^2, \quad (7)$$

where \cdot denotes the dot product. Once an optimal value of C is computed, the optimal value of the relative distance matrix, $\bar{\mathbf{R}}_{t_0}^*$, and the optimal angles of arrival, Λ_{t_0, t_w}^* , are given by Eq. (6). The nonlinear least squares problem of Eq. (7) is only a four-parameter estimation problem and can be estimated using methods like gradient descent (our implementation uses Levenberg-Marquardt). Solving Eqs. (6) and (7) is more robust to local minima than trying to solve a nonlinear optimization problem over the two dimensional positions of all vehicles and the angle of arrival of all sources.

A variety of angles are needed in order for this estimation procedure to be well conditioned. When all α_{t_0, t_w} are similar, the position of the vehicles will not be constrained along the dimension that is perpendicular to the angles α_{t_0, t_w} , and the optimization procedure of Eq. (3) will be unstable. This introduces a fundamental trade-off when considering a single moving source: increasing the time window, t_w , increases the opportunity for more azimuthal variety, but it also means that the assumption of a constant swarm geometry is less likely to hold, $\bar{\mathbf{R}}_{t_0} = \bar{\mathbf{R}}_{t_0 \pm t_w}$.

D. Condition number of C

C provides a ‘‘Euclidean upgrade’’ to the affine solution based on the angles of arrival computed in Eq. (6).²⁵ When the angles of arrival are similar, the estimation of Eq. (6)

becomes ill-conditioned and will be more sensitive to noise on the input values, \mathbf{D}_{t_o, t_w} . A way to measure this is to look at the condition number of C after it is estimated by Eq. (7), which estimates the worst case loss of precision from the Euclidean upgrade. If the condition number of C is large, the final solution to Eq. (6) will be more sensitive to noise on the input, \mathbf{D}_{t_o, t_w} , and the solution is less likely to produce a good estimate.

The choice of the anchor unit when computing $\bar{\mathbf{R}}_t$ affects the condition number of C . Since the swarm presented in Sec. III is relatively small, Eq. (6) is computed for all possible values of the anchor unit and the solution that yields the smallest condition number for C is chosen. This chooses a solution that is least sensitive to input noise.

E. Refinement

The problem with the solution outlined in the last section is that it does not utilize all of the information between all vehicle pairs. Since the relative distance matrix is relative to an arbitrary anchor unit, we are only considering the TDOA values compared to that anchor unit, i.e., $\delta_{i,1,t}$. This means that the solution is susceptible to poor detections on the anchor unit and that there is information received by other pairs that could be useful for the self-localization solution.

To overcome this, the solution to Eq. (6) is refined by minimizing the following nonlinear optimization problem using the Levenberg-Marquardt algorithm:

$$\begin{aligned} \bar{\mathbf{R}}_{t_o}^*, \alpha_{t_o, t_w}^* = \arg \min_{\mathbf{R}_{t_o}, \alpha_{t_o, t_w}} \sum_{h=1}^N \sum_{i=1, i \neq h}^N \sum_{t=t_o-t_w}^{t_o+t_w} & ([R_{h,t_o} - R_{i,t_o}] \\ & \times [\cos(\alpha_t) \sin(\alpha_t)]^T - c\delta_{h,i,t})^2. \end{aligned} \quad (8)$$

We use the solution computed in Eq. (6) as the starting point for this optimization problem. Equation (8) refines the solution of Eq. (6) to include the information from new pairs of vehicles. An important thing to note about our estimation is that given a center time, t_o , the only parameter that needs to be chosen is the length of the time window, t_w . Choosing t_w will be discussed in Secs. V and VI.

III. EXPERIMENT

We described the core estimation of the relative vehicles positions in Sec. II, and we will describe how we extract the constraints for our estimation in Sec. IV. In this section, we describe our experiment, which will make the description of how we extract constraints between vehicles in Sec. IV more clear.

Our self-localization approach was tested using data from a multi-vehicle deployment. Acoustic data was collected using Autonomous Underwater Explorers (AUEs), designed and built at the Scripps Institution of Oceanography.¹ Each AUE is a self-contained sensor that collects acoustic data from an HTI-96-MIN hydrophone, in addition to accelerometer, temperature, and pressure data. The AUEs track a depth in the ocean's water column by adjusting their buoyancy. When the motor is on, the hydrophones are saturated. For an individual AUE, the motors ran

at least 10% of the duration of the experiment and are independent of the other AUEs. The AUEs have no actuation in either horizontal direction and currents move the AUEs substantially while deployed in the ocean. The clocks of the AUEs are synchronized at the beginning and end of the experiment using a GPS receiver on each AUE, and a linear drift correction is applied to each AUE clock before any processing on the acoustic data is performed. The AUEs measure their depth using an onboard pressure sensor. In order to provide a "Ground Truth" estimate for our vehicle geometry, an acoustic triangulation system is set up to act similar to a GPS system. Five acoustic buoys are positioned on the surface of the ocean to send a linearly modulated chirp (8–15 kHz) to be received by the hydrophone of each AUE. Each of five buoys takes turns pinging and each packet of five pings occurs every ten seconds. The AUEs have limited on-board processing and no communication infrastructure, so all localization is done offline after they are retrieved. We believe that these localization estimates are accurate to about 1–3 meters.^{1,4,5}

The experiment was performed off of the coast of the Scripps Institution of Oceanography and is described in Fig. 2. Eight AUEs and the five pinger localization system were deployed close to shore, in water that was approximately 20 meters deep. The pingers were separated between 0.5 and 1 km apart. The AUEs were programmed to hold a depth of 6 meters and were accurate to ± 1 meter. While the AUEs were deployed, a small boat circled the acoustic pinging array twice while recording its GPS position at 1 Hz. In the first loop around the array, the boat traveled approximately 11 m/s and circled the array in approximately 400 s. During the second loop, the boat was slowed to 4 m/s and traveled around the array for approximately 700 s. Both the bearing and range of the boat from the center of the swarm are shown in Fig. 3. In this figure, we can see the change in speed of the boat in the bearing plot. The range plot also shows that we can model the boat as a single angle in relation to the AUE swarm.

IV. DATA ANALYSIS

In order to provide physical constraints from the moving boat, arrival information must be extracted between AUEs. This is done by correlating low frequency noise between pairs of vehicles to extract TDOA information from dominant noise directions [i.e., Eq. (1)]. These TDOA measurements provide the input to the self-localization solution described in Sec. II. This section describes implementation details regarding these low frequency correlations.

There are conditions that make the boat difficult to detect for some deployment times. First, there is missing data in the acoustic recordings resulting from the vehicles running their motor to adjust its buoyancy (described in Sec. III). The boat can also be difficult to detect when it is broadside to the vehicles, and when there are other noise sources in the vicinity of the swarm broadcasting in the same frequency band. To combat these issues, the acoustic data is preprocessed before correlating each pair of audio tracks. The preprocessing is described in Sec. IV A. Temporal

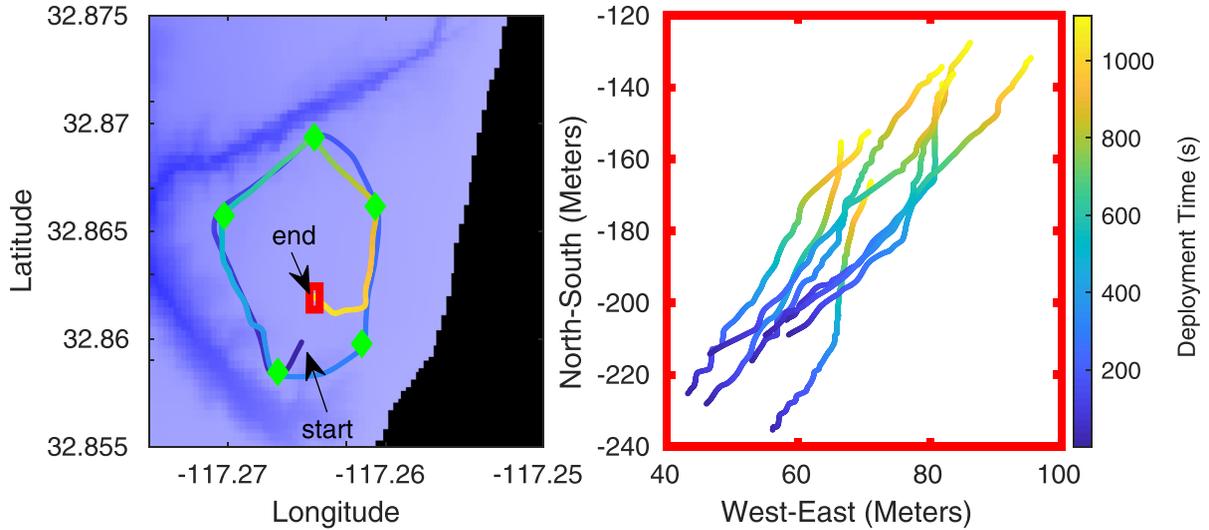


FIG. 2. (Color online) Experiment setup: (left) Five acoustic pingers, shown by green diamonds, are moored along the perimeter of the deployment, which took place off of the coast of La Jolla. The bathymetry as well as the shoreline are shown by colors (darker blue is deeper, black denotes the shoreline). These pingers collect GPS data and take turns transmitting a linearly modulated chirp (8–15 kHz). Eight AUEs, deployed inside the red box, hold a depth of 6 meters and float with the subsea currents. The trajectories of the AUEs are estimated using time of flight measurements from the five acoustic pingers. While the AUEs were deployed, a boat circled the swarm twice (once at approximately 11 m/s and once at approximately 4 m/s) and the GPS trajectory of the boat is shown with the start and end positions indicated. The right panel shows a close up of the AUE trajectories where the red bounding box matches the box on the left panel. Deployment times for both the boat and AUE trajectories are shown by the colorbar on the right.

smoothing is also employed to reject spurious arrivals caused by missing data or other noise sources in the vicinity of the array, described in Sec. IV B.

A. Preprocessing and noise correlations

Each AUE, indexed by i , records a pressure signal that we denote as $\tilde{p}_i(t)$. Before correlations are performed between different vehicles, the noise is preprocessed similar to work focused on estimating the acoustic impulse response between receivers.^{21,23,29,30} These preprocessing steps help suppress spurious arrivals in the noise correlation function, allowing the noise correlations to focus on the dominant source. The first pre-processing step is to whiten the signal's frequency spectrum (the absolute whitening described in Brooks *et al.*²⁹ was chosen) as well as defining the

bandwidth, B_ω , that will be correlated over. The whitening weighs each frequency band equally across the defined bandwidth, B_ω , in the computation of the noise correlations instead of having a few high energy bands dominate. Next, time series values with high amplitudes are truncated to the fourth standard deviation of each $\tilde{p}_i(t)$ signal. More insight into these preprocessing steps can be found in the literature.^{21,23,29,30} We define the preprocessed version of $\tilde{p}_i(t)$ as $p_i(t)$ and compute the normalized noise correlation between two signals, $p_i(t)$ and $p_j(t)$ as

$$C_{i,j,t_c}(\tau) = \frac{\int_{t_c-(T_r/2)}^{t_c+(T_r/2)} p_i(t)p_j(t+\tau)dt}{\sqrt{\int_{t_c-(T_r/2)}^{t_c+(T_r/2)} p_i(t)^2 dt} \sqrt{\int_{t_c-(T_r/2)}^{t_c+(T_r/2)} p_j(t)^2 dt}}, \quad (9)$$

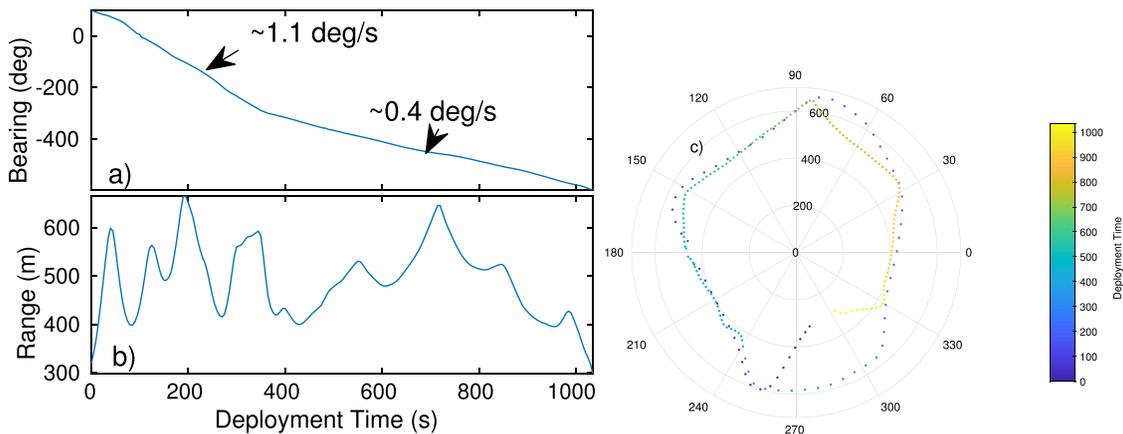


FIG. 3. (Color online) Range and bearing of the boat: (a) The bearing of the boat in relation to the center of the AUE swarm. (b) The distance between the boat and the center of the AUE swarm. The bearing plot shows a change in the derivative of the bearing around 400s into the deployment as the boat slows its speed [from 1.1 deg/s (11 m/s) to 0.4 deg/s (4 m/s)]. (b) The range plot shows that the boat is in the far field of the AUE swarm for the entire deployment (e.g., the boat is 2–3 times further away than the aperture of the swarm). (c) The same information is shown in polar form.

where t_c controls the center time of the signal and T_r the length of the correlation window. The normalized cross-correlation allows us to make comparisons between different correlation times by removing the total energy at each time step. The noise correlations effectively have two user defined parameters: T_r and B_ω . For our experiments, we took $T_r = 3$ s and $B_\omega = [0.1-1.0$ kHz].

The envelope of the noise correlation between two different pairs is shown by the colorbar in Fig. 4. Each row is a short time noise correlation centered at the center time, t_c , indicated on the y-axis. The center times are spaced one second apart. As expected, the noise correlations are dominated by the circling boat. The noise correlation changes as the boat moves around the vehicles.

B. Temporal smoothing

While the SNR of the boat is large for most correlations, the short time windows and hydrophone dropouts from buoyancy adjustment create peaks in the noise correlations that are not from the boat or eliminate peaks in the correlations entirely. To overcome some of these difficulties and improve the estimation, the detections are smoothed along the deployment time axis to enforce the constraint that the peak in the noise correlation must vary continuously in time. To do this, the Viterbi algorithm³¹ was implemented to estimate the sequence of TDOA values, $\hat{\delta}_{i,j,t}$ for $t_c = [1, 2, \dots, t_{end}]$. The implementation is heuristic in that the choices of probability distributions are not rigorously justified but were chosen so that the estimates visually tracked the peaks in the correlation well. Note that this temporal smoothing may not

be needed in cases where the detection is strong throughout the duration of other deployments.

The Viterbi algorithm requires two probabilities to be defined, one describing transition probabilities between $\hat{\delta}_{i,j,t-1}$ and $\hat{\delta}_{i,j,t}$, and the other the probability of the noise correlation having the observed structure given the true TDOA value, $\delta_{i,j,t-1}$. We define the transition probabilities as Gaussians centered around the last state

$$P(\delta_{i,j,t} = \beta | \delta_{i,j,t-1} = \gamma) = \mathcal{N}(\gamma, \sigma^2). \quad (10)$$

For the observation probability, we use a heuristic that the probability of the noise correlation observation given that the TDOA measurement is equal to β is related to the envelope of the normalized noise correlation evaluated at β ,

$$P(C_{i,j,t} | \delta_{i,j,t} = \beta) \propto \text{env}(C_{i,j,t}(\beta)), \quad (11)$$

where $\text{env}(\cdot)$ denotes the envelope operator.

For our purposes, this algorithm tries to accumulate the most correlation energy while only allowing small variation in the time value of consecutive correlation peaks. This temporal constraint on the peaks of the correlation function can be adjusted by the standard deviation choice of the transition probability. We chose $\sigma = 0.3$ ms, which visually tracked the peaks well for the pairs examined. In our case, this variable was not optimized, but could be chosen given *a priori* knowledge of expected movement of a source in relation to the vehicles. Given these probabilities, the Viterbi algorithm produces the most likely sequence of TDOA values, $\hat{\delta}_{i,j,t}$, based on the time evolution of the noise correlations. This smoothing is performed for noise correlations between all pairs of vehicles.

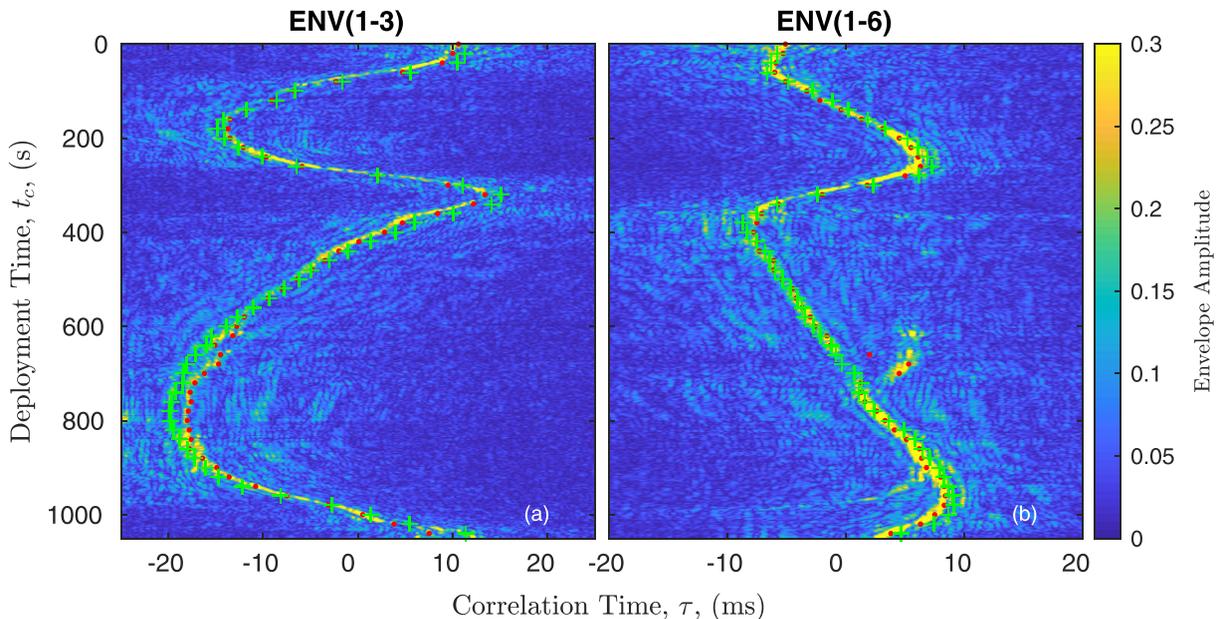


FIG. 4. (Color online) Noise correlations: (a), (b) The envelope of the noise correlations are shown between two different pairs of AUEs while the boat is circling the array. Each row is the envelope of a short time cross-correlation with the amplitude shown by the colorbar. The center time of the cross-correlation is shown by the deployment time on the y-axis. The expected peaks of each noise correlation are shown by green “+” and were computed using the GPS position from the boat and the estimated locations of the AUEs from the five acoustic pingers. The red dots show the estimated location of the peaks, $\hat{\delta}_{i,j,t}$, for each deployment time after the smoothing operation of Sec. IV B is performed. The detected peaks in the noise correlation match the expected peaks, meaning that the boat is the main dominating source in the noise correlation. This provides confidence that the estimated signal shown by dots can be used to self-localize the swarm. Parameters: $T_r = 3$ s and $B_\omega = [0.1-1.0$ kHz].

The accuracy of the detection for every pair of vehicles is summarized in Fig. 5. For this figure, the value of t_c is shown on the x axis and multiple statistics about the discrepancy between $\hat{\delta}_{i,j,t}$ and $\delta_{i,j,t}$ (i.e., the “+” and “-” of Fig. 4) are shown for all pairs. Also shown is the expected uncertainty of this difference, which is based on the bandwidth of the correlation process (uncertainty of 0.5 ms), the uncertainty in the sound speed (0.5 ms assuming a sound speed uncertainty of 20 m/s and vehicle separation of 50 meters) and the uncertainty of the AUE positions (2 ms) for a total of 3 ms. Figure 5 shows that the residuals do not have a mean of 0 for much of the deployment time, but this bias is usually under 1 ms. The extrema of the discrepancy are typically under the expected uncertainty with some exceptions. The spike that was visible in the correlations given in Fig. 4 (see the right panel at $t_c = 700$ s) is also visible in Fig. 5. What is difficult to deduce in Fig. 5 but is visible in Fig. 4 is that the errors are not distributed around the expectation (plotted by “+” in Fig. 4). Figure 4 shows that the detected correlation peaks frequently have a constant offset from the estimated correlation peaks. This could be a result of slight inaccuracies in the “ground truth” positions of the AUEs, or incorrect assumptions about the sound speed structure, both of which would effect the value of $\delta_{i,j,t}$. This is important because this error is clearly not a result of a white Gaussian process and cannot be modeled as such when we try to understand the robustness of the solution using simulations in Sec. VI. In fact, the results in Sec. V have better outcomes than what would be expected if the error is modeled as a Gaussian process with a standard deviation as high as what is reported in

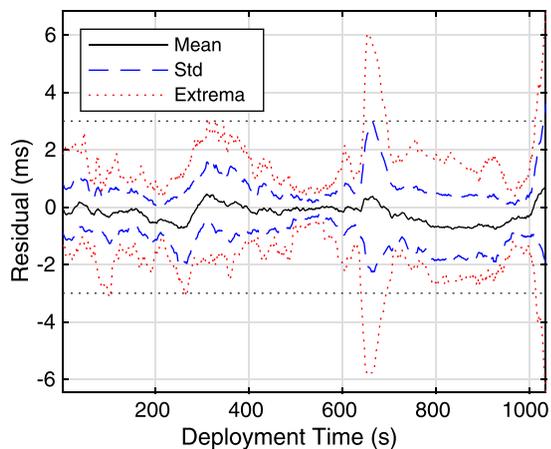


FIG. 5. (Color online) Discrepancy between the detected, $\hat{\delta}_{i,j,t}$, and expected noise correlation peaks, $\delta_{i,j,t}$, for all pairs: The mean (black solid), standard deviation (blue dashed) and the maximum and minimum (red dotted) discrepancies are shown as a function of deployment time for all combination of pairs of vehicles. Also shown (black dotted) is the uncertainty of the computed discrepancy based on the uncertainty of the AUE locations (2 ms), the uncertainty in the sound speed profile (0.5 ms), and the bandwidth of the correlation (0.5 ms) for a total of 3 ms. The expected noise correlation peaks do not perfectly match the measured correlation peaks, and there is at least one pair that is off by more than 1 ms at every deployment time. There is a spike close to deployment time 700 s where many of the noise correlations jumped to another peak, shown in Fig. 4(b). These discrepancies may adversely affect the comparison between the estimated relative trajectories from the boat noise and the measured ones from the high frequency pinging system. However, for most deployment times the discrepancies are within the estimated uncertainty.

Fig. 5. This indicates that while the estimates of the TDOA, $\hat{\delta}_{i,j,t}$, do not exactly match the expected TDOA, $\delta_{i,j,t}$, they still provide a consistent model that generates an answer similar to what we expect from the “ground truth” locations of the AUEs.

V. RESULTS

The estimation procedure described by Eqs. (6) and (8) was applied to the data described in Fig. 4 and Fig. 5. Examples of two solutions with different center times, $t_o = 160$ and $t_o = 945$, and a time window of 30 s (i.e., $t_w = 15$ s), are shown in Fig. 6. The left panel of Fig. 6 compares the estimate of the vehicle geometry (shown by “*” markers) to the ground truth estimate (shown by “○” markers for the center time). The estimated ground truth motion of each AUE around the center time is shown by black dots. Since we only estimate the geometry in a relative sense, the comparison is given after an optimal rigid transformation is computed between the estimates using the algorithm described in Arun *et al.*³² The root mean square error (RMSE) in meters is reported after this fit is applied. The results show good agreement, on the order of one or two meters for the two cases shown. The results show promise for the method proposed in Sec. II; without any knowledge of the position of the noise source, or any prior information on the relative geometry of the vehicles, we are able to estimate the relative geometry, $\hat{\mathbf{R}}_{t_o}$, as well as the angle of arrivals of the noise source α_{t_o,t_w} .

Another way to evaluate the estimation is to compare the estimated angles of arrival with the true angles of arrival. The RMSE is reported, in degrees, for the discrepancy between the estimated angles of arrival and the known angles of arrival from the GPS measurements of the boat. This RMSE is reported after the rotation from the rigid transformation was applied to match coordinate frames. We can see that the angle estimates match well to what they are expected to be, with an RMSE of 2.9 degrees and 0.48 degrees for the top and bottom figures, respectively. In this frame of reference, the estimation recovers the geometric interpretation of the soundscape, which can be useful for passive sensing in ocean environments.

It is important to put the results in Fig. 6 into the context of Fig. 5, which showed that the detection results do not perfectly agree with their expectation. In fact, Fig. 5 shows that a 2 ms error between the expected time value of the noise correlation function and the measured time value is not uncommon. This could translate to an RMSE on the order of a few meters. The source of the discrepancies in Fig. 5 are difficult to understand. As previously mentioned, the positions of the AUEs from the high frequency pinging system are also estimates of the AUE positions and are expected to have an accuracy around 1–3 meters. A static sound speed of $c = 1500$ m/s is assumed in the calculation of Eqs. (3) and (8) and this may have a slight effect on the accuracy. In reality, most of the error probably results from the small angles that are used for the estimation. We can see that the error is perpendicular to the angle of arrivals, especially in the bottom panel of Fig. 6. This is the direction that is the most difficult to resolve given the angle of arrivals of that deployment time.

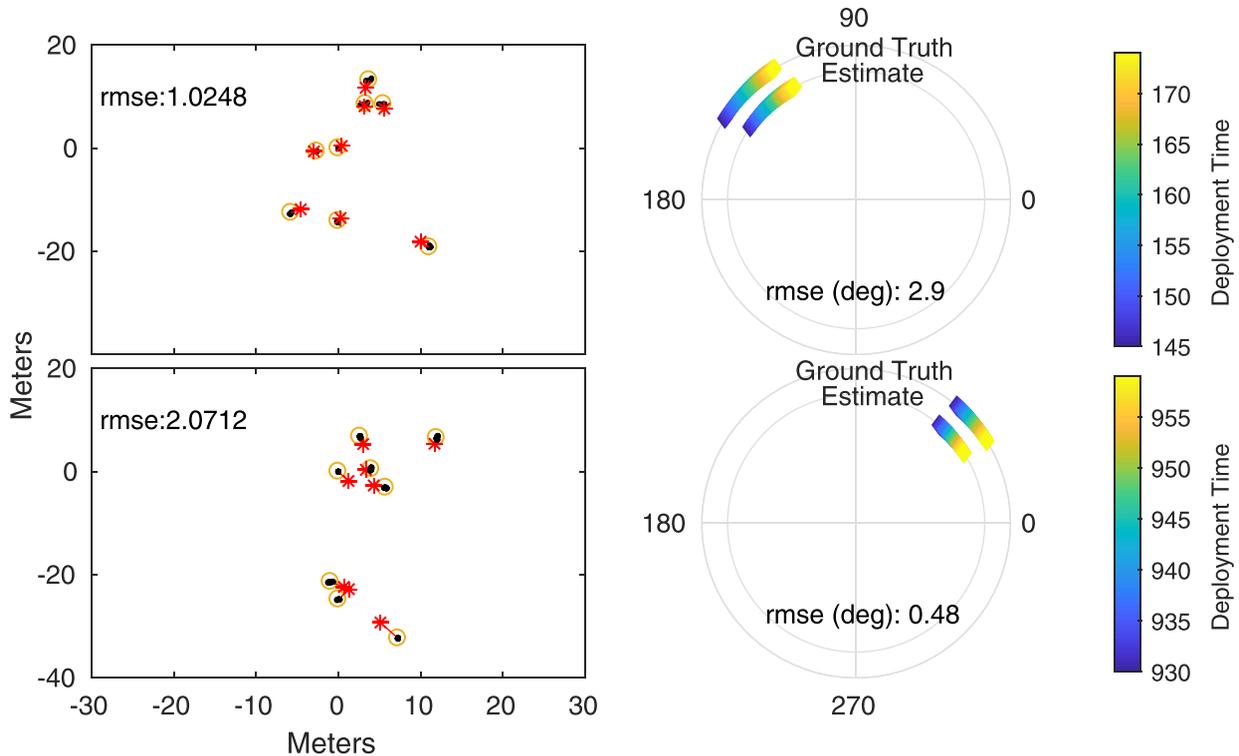


FIG. 6. (Color online) Example solutions: The solutions to Eq. (8) are shown for two different deployment times, $t_o = 160$ s and $t_o = 945$ s. In the left panels, small black dots represent the GPS pinger estimated positions of each vehicle for each second in the chosen 30 s window ($t_w = 15$ s), the positions at the center time, t_o , are marked by an “O.” The swarm geometry estimated from the noise is shown by “*” after an optimal translation and rotation is applied to put the estimate in the same coordinate frame as the ground truth estimates. The RMSE in meters is reported between the position at the center time and the estimate after the fit is performed. On the right panels, the angles of arrival are reported for both the ground truth (on the outside ring) and the angle of arrival estimates (on the inside ring) after the optimal rotation is applied. The RMSE error is shown, in degrees, for the angles at each time step. From these two examples, we can see that the method is accurately able to estimate the relative geometry of each vehicle and the angle of arrival of the dominant source.

The estimation results of Eq. (8) are not similar across the entire deployment time. Figure 7 shows a time ($t_o = 110$ s) that resulted in a solution that was not close to the geometry of the swarm. In the top solution of Fig. 7, there is greater variation in the true angles of arrival than what we see in Fig. 6. Despite this extra time window (here $t_w = 30$ s) and azimuthal variation, the solution provided in the top panel of Fig. 7 is not a good estimate, both in terms of estimating the relative geometry of the swarm and also the angle of arrival of the noise. However, as we increase the window length to add more variation to the angle of arrivals, we eventually get a solution that is more comparable to the solutions reported in Fig. 6. This example shows that increasing the angle of arrivals provides more stability even if it increases the total magnitude of the vehicle motion during the time window (the vehicle motion for the time window is shown by black dots). It also highlights the difficulty of understanding the interplay of all the different factors that could influence the performance the estimation. These factors are examined through simulations in Sec. VI.

VI. SIMULATIONS

A. Simulation description

This section aims to decouple some of the effects that influence the results shown in Figs. 6 and 7 through simulations. Simulations are first shown for two different vehicle geometries to demonstrate that the vehicle geometry

influences the solution and are later shown for random realizations of the vehicle geometry to show that some results generalize to many geometries.

The first set of simulations use two different geometries for the swarm, \mathbf{R} , which are shown later in Figs. 9(a) and 9(d). In order to understand the effects of vehicle motion on our solution, simulations when the stationary assumption is exact ($\mathbf{R}_{t_o} = \mathbf{R}_{t_o, t_w}$) are compared to simulations when the stationary assumption is only approximate ($\mathbf{R}_{t_o} \approx \mathbf{R}_{t_o, t_w}$). In Figs. 9(a) and 9(d), the total movement for the vehicles is shown by small black dots for each of 60 time instances (an arbitrarily number of time instances, but similar to the examples shown in Figs. 7). The center time is chosen as the position of the vehicles when the stationary assumption is exact, which is also considered the ground truth position when the vehicles have motion. Figures 9(b) and 9(e) show results when there is movement and Figs. 9(c) and 9(f) show times when the vehicles are completely stationary. Sixty source positions, S_i , are defined in polar coordinates with $r = 1000$ m and various values of α_{t_o, t_w} .

We define α_{t_o, t_w} as 60 source positions evenly spaced in angle from $\alpha_{t_o - t_w}$ to $\alpha_{t_o + t_w}$, and describe α_{t_o, t_w} using two parameters: the center angle, α_{t_o} , and the “span” of the angles, $(\alpha_{t_o + t_w} - \alpha_{t_o - t_w})$. Figure 8 shows a visual explanation of these values. The span refers to the total azimuthal range in degrees and is a measure of the amount of azimuthal information that is given by the source. For example, a span of 40° with a center angle of $\alpha_{t_o} = 0^\circ$ means that the 60

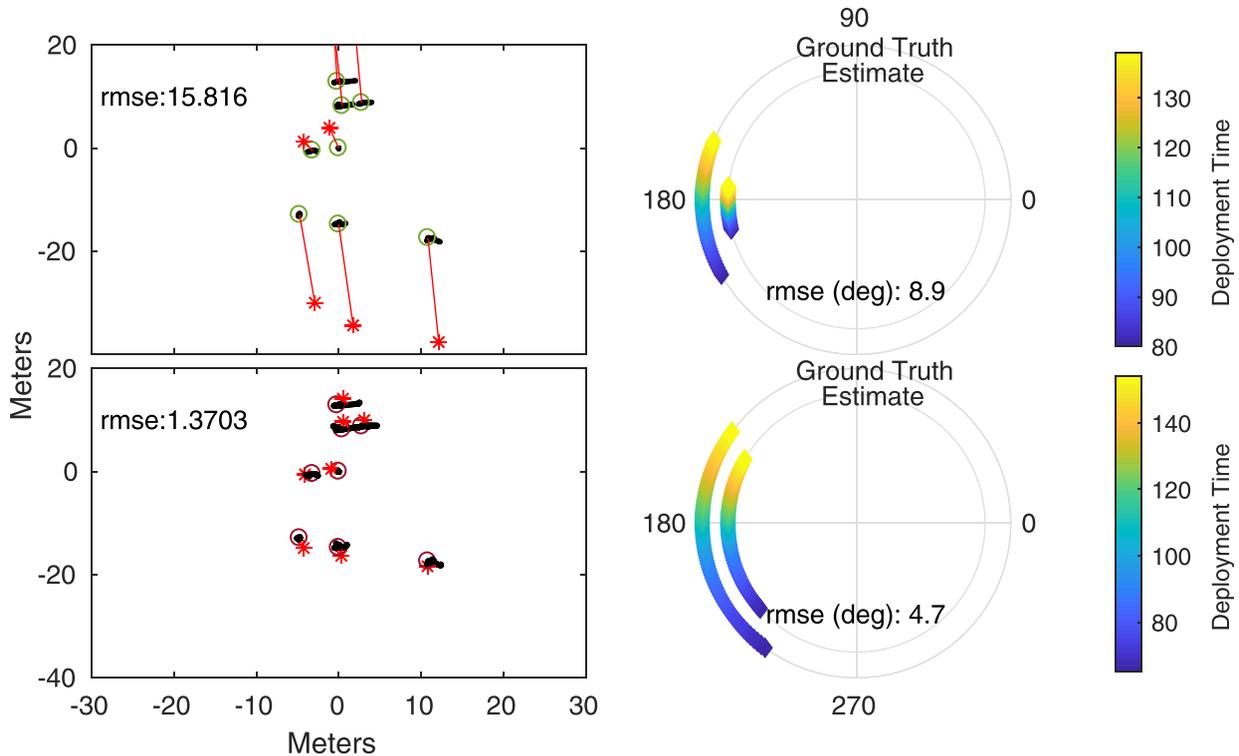


FIG. 7. (Color online) Challenges: Similar plots to Fig. 6, but here the time window length is varied (top $t_w = 30$ s, bottom $t_w = 45$ s) for the same deployment time, $t_o = 110$ s. This figure demonstrates that some angles of arrival are more challenging to the estimation procedure. The estimation is poor in the top panel using a 60 s time window because of the lack of noise sources North or South of the swarm. Therefore, it is hard to estimate distances between the AUEs along this direction. By increasing the time window to 90 s ($t_o = 45$ s), the levels of accuracy that were reported in Fig. 6 are achieved because the angles of arrival now span the North-South direction. More insight into the interplay between the angle of arrival, vehicle motion, and vehicle geometry is investigated through simulations in Sec. VI.

source positions are evenly spaced from -20° to 20° . A large azimuthal span means that there are adequate angles for the estimation problem to be well conditioned. A small azimuthal span means that all the angles are similar, and the estimation will be challenged. In order to understand the interaction between the swarm geometry, \mathbf{R} , and the angles of arrival, α_{t_0, t_w} , the choices of α_{t_0, t_w} are varied while keeping the relative motion constant. For each of the lines in Figs. 9(b), 9(c), 9(e), and 9(f) the center angle, α_{t_0} , is shown for increments of

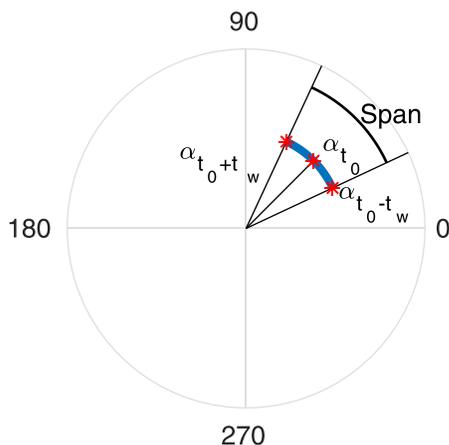


FIG. 8. (Color online) Simulation explanation: A visual definition of the parameters used in the simulations (Figs. 9 and 10). The span dictates the magnitude of the angles used, while α_{t_0} describes the orientation of angles in relation to the vehicles.

45° as indicated by the legend. The center angles are defined with respect to the vehicle geometry in Figs. 9(a) and 9(d). The x axis describes the span of angles (i.e., $\alpha_{t_0 + t_w} - \alpha_{t_0 - t_w}$). With the defined parameterizations of the source positions and vehicle geometry, TDOA values are calculated using Eq. (1) ($c = 1500$ m/s). The relative geometry of the vehicles was estimated using these TDOA values in Eqs. (3) and (8) after zero mean Gaussian noise with a standard deviation of 0.8 meters was added to the $c\Delta$ terms of Eqs. (3) and (8). This process was repeated 200 times with different noise realizations, and the median results are shown.

B. Simulation analysis

There are some emerging trends from the simulation results. The first trend is that movement degrades the self-localization estimate. This is especially apparent when the span of α_{t_0, t_w} is small. The median RMSE values of the solution are large when there is small azimuthal variation and movement [Figs. 9(b) and 9(e)], but are small when the vehicles are stationary, Figs. 9(c) and 9(f). More subtly, the stationary cases always have a smaller median value of the RMSE than the cases with movement. This result is expected, as the model assumes that the relative geometry of the vehicles is stationary.

We can also see that the angle of arrival with respect to the geometry of the swarm plays a role in the results from the different curves shown in Fig. 9. Some of the curves

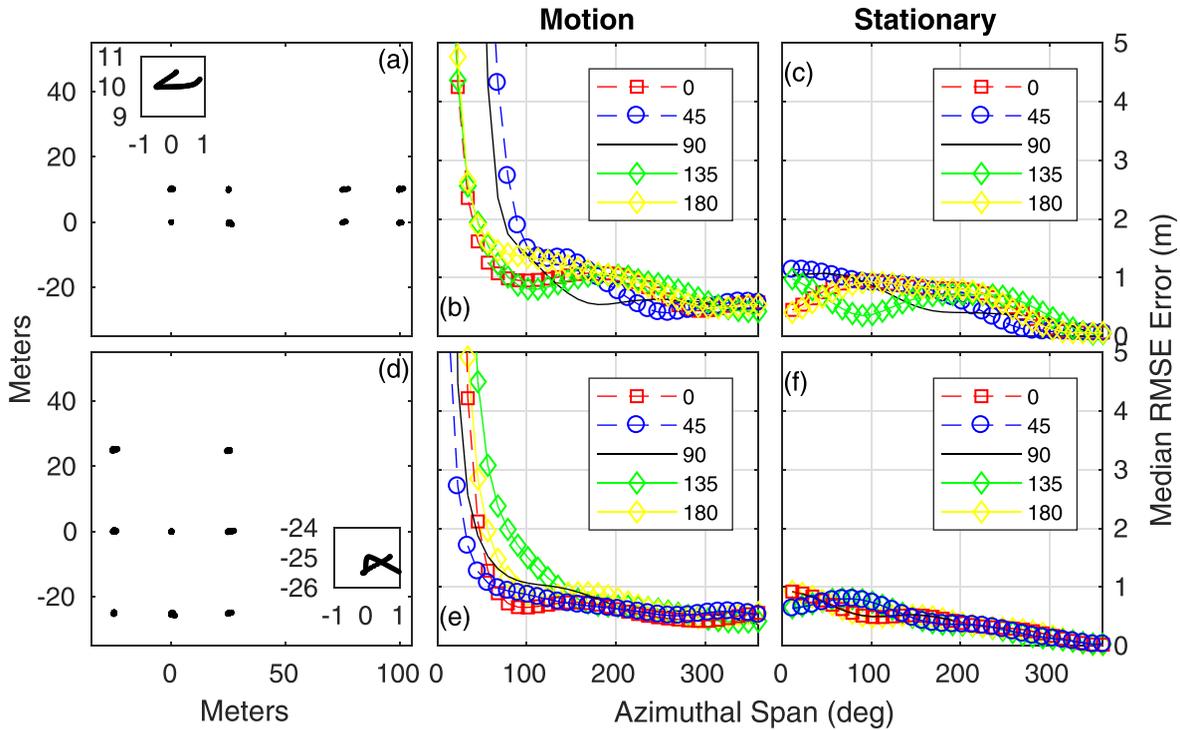


FIG. 9. (Color online) Simulations are provided for two different swarm geometries, (a) and (d). (a) and (d) show the position of the floats for each of 60 s with black dots. A close up of one of the vehicles is provided in the inset to show an example of the vehicle movement. The vehicle movement is similar to what is seen in the real data. The simulations are computed for vehicles that are moving, (b) and (e), as well as for vehicles that are stationary, (c) and (f). The stationary case uses the center time of the defined trajectory for each vehicle. The simulations are computed similar to what was shown in Figs. 6 and 7 where the swarm geometry and angle of arrival are computed using noisy TDOA measurements. In (b) and (c) and (e) and (f), the center angle of arrival, α_{t_0} , for each line is given in the legend. The total variation of the angles is shown on the x axis and the median RMSE of 200 iterations between the ground truth and the estimated geometry are shown. These simulations show the effects of swarm geometry, motion, and angle of arrival on the total accuracy. More details are described in the text.

have lower RMSE for the same azimuthal span [compare $\alpha_{t_0} = 0^\circ$ which is better than $\alpha_{t_0} = 45^\circ$ in Fig. 9(b)]. Again, this is especially apparent in the simulations with movement, where we can see that there is a separation in the median RMSE curves as a function of the center angle of arrival. The angle of arrivals that result in the best RMSE curve are not consistent across the different swarm geometries. Additionally, the RMSE curve is also not symmetric in angles even though the swarm geometries are symmetric. This last fact is likely due to the small scale relative motion in the vehicles, which is constant across the noise

realizations and geometries—the relative motion favors some angles of arrival more than others.

In order to show that the trends of Figs. 9(b) and 9(e) generalize to geometries that are not shown, Fig. 10 is given. Figure 10 shows similar information to Figs. 9(b) and 9(e) but this time the vehicle positions are chosen randomly in a box ($0 < x < 50$, $0 < y < 50$). For each realization of the vehicle geometry, the same motion vectors and noise levels are assigned that were given in Figs. 9(a) and 9(d). Figure 10 shows that the trends are similar to what was reported in Fig. 9 except the effects of the center angle α_{t_0} are not as apparent

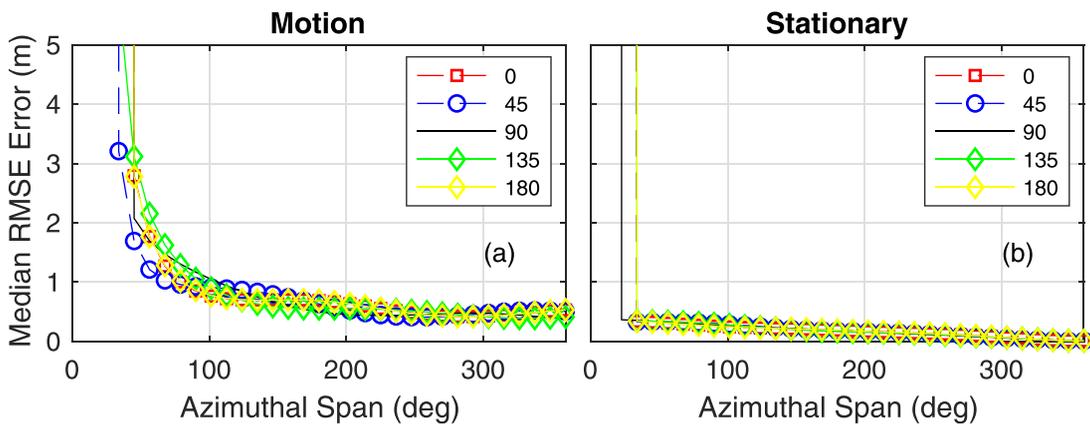


FIG. 10. (Color online) Simulation with random positions: In order to show that the trends in Fig. 9 are typical of most deployments, this figure shows the same setup as Fig. 9 except that for each of 200 iterations, a random geometry is chosen in the box.

after averaging over the different geometries. In summary, the simulations shown in Figs. 9 and 10 demonstrate three points: (1) motion degrades the localization estimate, (2) this degradation can be overcome by adding more azimuthal span, and (3) the center angle and the vehicle geometry plays some role in the results. These points are also consistent with the results shown when examining real data in Figs. 6 and 7.

C. Comparing the simulations and data

The simulations of Fig. 9 support the findings that were presented in the data analysis (Figs. 6 and 7). Figures 9(b) and 9(e) show that when the relative geometry of the vehicles is not stationary, there are certain angles of arrival that produce worse results [e.g., $\alpha_{t_0} = 45^\circ$ for Fig. 9(b) and $\alpha_{t_0} = 135^\circ$ for Fig. 9(e)]. The angles that perform worse are dictated by the relative geometry of the swarm and the motion of the vehicles. This result is also seen in the experimental results. Figure 7 shows an example of a challenged angles of arrival that are coming from approximately 180° . For these angles of arrival, the solution required a larger span of angles to obtain a good solution than the solutions in Fig. 6. This larger span of angles was needed in order to accurately estimate distances between the AUEs in two dimensions (e.g., sound sources needed to be located North or South of the swarm before the distances of the AUEs were accurately estimated in the North-South direction). Conversely, the simulations show that for some center angles, only a small span of angles is needed to consistently recover a solution (e.g., 0° for both plots), which is consistent with the data presented in Fig. 6. The simulations also show that when there is no movement between the vehicles, small angles can be used to successfully recover the swarm geometry. Inter-vehicle movement could be a factor in the difference between the results presented in Figs. 6 and 7, especially since the vehicle movement in the simulations is not large (it does not stray further than one meter away from the position at the center time).

VII. CONCLUSION

This work demonstrated a self-localization solution for moving vehicles using low frequency noise with dominant noise directions. The movement between vehicles changes the propagation path and forces the solution to consider short time scales. A solution was proposed that was able to jointly estimate the array geometry as well as the angle of arrival of the noise sources using short time windows. Experimental validation was provided using a moving boat with known location to dominate the ambient soundscape, which allowed us to verify that our solution was able to produce accurate direction of arrival estimates without any *a priori* knowledge of the source location. Simulations were provided that showed that small azimuthal variation, coupled with vehicle movement degrades the solution. Therefore, having azimuthal variety is important to the stability of the solution. The stability of the solution was also dependent on the interaction between array geometry and the angles of arrival of the source; different center angles resulted in different

accuracies for the solution in both data and simulations. Ultimately, the self-localization solution provides a new and passive way to track the movement of a swarm of underwater vehicles, providing a spatial estimate to the data they collect. This is useful for many coordinated surveys, especially those that require high endurance missions.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation grant “INSPIRE Track I: Distributed Sensing Collective to Capture 3D Soundscapes” supported by Grant No. 1344291. This material is also based upon work supported by NSF Grant No. 09-27449, the National Science Foundation Graduate Research Fellowship under Grant No. 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.

- ¹J. S. Jaffe, P. J. S. Franks, P. L. D. Roberts, D. Mirza, C. Schurgers, R. Kastner, and Adrien Boch, “A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics,” *Nature Commun.* **8**, 14189 (2017).
- ²K. S. Johnson, J. N. Plant, L. J. Coletti, H. W. Jannasch, C. M. Sakamoto, S. C. Riser, D. D. Swift, N. L. Williams, E. Boss, N. Haëntjens, L. D. Talley, and J. L. Sarmiento, “Biogeochemical sensor performance in the soccom profiling float array,” *J. Geophys. Res. Oceans* **122**(8), 6416–6436 (2017).
- ³J. S. Jaffe, “Multi autonomous underwater vehicle optical imaging for extended performance,” in *Proceedings of OCEANS 2007—Europe*, Aberdeen, Scotland (June 18–21, 2007), pp. 1–4.
- ⁴P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, “Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity,” *IEEE Access* **6**, 1635–1646 (2018).
- ⁵P. 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,” *J. Acoust. Soc. Am.* **140**(6), 4260–4270 (2016).
- ⁶J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb, “A survey of underwater vehicle navigation: Recent advances and new challenges,” in *Proceedings of the IFAC Conference of Manoeuvring and Control of Marine Craft*, Lisbon, Portugal (September 20–22, 2006), pp. 1–12.
- ⁷H.-P. Tan, R. Diamant, W. K. G. Seah, and M. Waldmeyer, “A survey of techniques and challenges in underwater localization,” *Ocean Eng.* **38**(14), 1663–1676 (2011).
- ⁸B. 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 *Proceedings of the OCEANS’97. MTS/IEEE Conference*, Halifax, Nova Scotia (October 6–9, 1997), pp. 994–1000.
- ⁹R. McEwen, H. Thomas, D. Weber, and F. Psota, “Performance of an AUV navigation system at arctic latitudes,” *IEEE J. Ocean. Eng.* **30**(2), 443–454 (2005).
- ¹⁰D. Mirza, P. Naughton, C. Schurgers, and R. Kastner, “Real-time collaborative tracking for underwater networked systems,” *Ad Hoc Netw.* **34**, 196–210 (2014).
- ¹¹M. M. Hunt, W. M. Marquet, D. A. Moller, K. R. Peal, W. K. Smith, and R. C. Spindel, “An acoustic navigation system,” Technical Report, Woods Hole Oceanographic Institution, Woods Hole, MA (1974).
- ¹²B. J. Sotirin and J. A. Hildebrand, “Acoustic navigation of a large aperture array,” *J. Acoust. Soc. Am.* **87**(1), 154–167 (1990).
- ¹³G. M. Wenz, “Acoustic ambient noise in the ocean: Spectra and sources,” *J. Acoust. Soc. Am.* **34**(12), 1936–1956 (1962).
- ¹⁴P. Roux, W. A. Kuperman, and K. G. Sabra, “Ocean acoustic noise and passive coherent array processing,” *C. R. Geosci.* **343**(8), 533–547 (2011).

- ¹⁵K. 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 J. Ocean. Eng.* **30**(2), 338–347 (2005).
- ¹⁶A. M. Thode, P. Gerstoft, W. C. Burgess, K. G. Sabra, M. Guerra, M. D. Stokes, M. Noad, and D. H. Cato, "A portable matched-field processing system using passive acoustic time synchronization," *IEEE J. Ocean. Eng.* **31**(3), 696–710 (2006).
- ¹⁷O. A. Godin, "Emergence of deterministic Green's functions from noise generated by finite random sources," *Phys. Rev. E* **80**(6), 066605 (2009).
- ¹⁸P. Roux, W. A. Kuperman, and the NPAL Group, "Extracting coherent wave fronts from acoustic ambient noise in the ocean," *J. Acoust. Soc. Am.* **116**(4), 1995–2003 (2004).
- ¹⁹P. Roux, K. G. Sabra, W. A. Kuperman, and A. Roux, "Ambient noise cross correlation in free space: Theoretical approach," *J. Acoust. Soc. Am.* **117**(1), 79–84 (2005).
- ²⁰K. G. Sabra, P. Roux, and W. A. Kuperman, "Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide," *J. Acoust. Soc. Am.* **117**(1), 164–174 (2005).
- ²¹S. 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," *J. Acoust. Soc. Am.* **124**(4), EL183–EL188 (2008).
- ²²S. W. Lani, K. G. Sabra, W. S. Hodgkiss, W. A. Kuperman, and P. Roux, "Coherent processing of shipping noise for ocean monitoring," *J. Acoust. Soc. Am.* **133**(2), EL108–EL113 (2013).
- ²³C. 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," *J. Acoust. Soc. Am.* **132**(2), 883–893 (2012).
- ²⁴A. 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 Process. Mag.* **33**(4), 14–29 (2016).
- ²⁵S. Thrun, "Affine structure from sound," in *Proceedings of the Advances in Neural Information Processing Systems 18 Neural Information Processing Systems Conference*, Vancouver, BC (December 5–8, 2005), pp. 1353–1360.
- ²⁶S. Burgess, Y. Kuang, J. Wendeborg, K. Åström, and C. Schindelbauer, "Minimal solvers for unsynchronized TDOA sensor network calibration," in *Proceedings of the International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics*, Sophia Antipolis, France (September 5–6, 2013), pp. 95–110.
- ²⁷M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV (March 31–April 4, 2008), pp. 2445–2448.
- ²⁸S. Zhayida, S. S. Rex, Y. Kuang, F. Andersson, and K. Åström, "An automatic system for acoustic microphone geometry calibration based on minimal solvers," [arXiv:1610.02392](https://arxiv.org/abs/1610.02392) (2016).
- ²⁹L. A. Brooks and P. Gerstoft, "Green's function approximation from cross-correlations of 20-100 Hz noise during a tropical storm," *J. Acoust. Soc. Am.* **125**(2), 723–734 (2009).
- ³⁰K. G. Sabra, P. Roux, and W. A. Kuperman, "Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function," *J. Acoust. Soc. Am.* **118**(6), 3524–3531 (2005).
- ³¹A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory* **13**(2), 260–269 (1967).
- ³²K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(5), 698–700 (1987).