Title

Automated two-dimensional localization of underwater acoustic transient impulses using vector sensor image processing (vector sensor localization)

Journal

The Journal of the Acoustical Society of America, 149(2)

ISSN

0001-4966

Authors

Thode, Aaron M
Conrad, Alexander S
Ozanich, Emma
et al.

Publication Date

2021-02-01

DOI

10.1121/10.0003382

Peer reviewed

JASA ARTICLE

Automated two-dimensional localization of underwater acoustic transient impulses using vector sensor image processing (vector sensor localization)

Aaron M. Thode,[1,a)] Alexander S. Conrad,[2] Emma Ozanich,[1] Rylan King,[3] Simon E. Freeman,[3] Lauren A. Freeman,[3] Brian Zgliczynski,[1] Peter Gerstoft,[1,b)] and Katherine H. Kim[2]

[1]*Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093-0238, USA*

[2]*Greeneridge Sciences, Inc., 90 Arnold Place, Suite D, Santa Barbara, California 93117, USA*

[3]*Naval Undersea Warfare Center, Newport, Rhode Island 02841, USA*

ABSTRACT:

Detecting acoustic transients by signal-to-noise ratio (SNR) becomes problematic in nonstationary ambient noise environments characteristic of coral reefs. An alternate approach presented here uses signal directionality to automatically detect and localize transient impulsive sounds collected on underwater vector sensors spaced tens of meters apart. The procedure, which does not require precise time synchronization, first constructs time-frequency representations of both the squared acoustic pressure (spectrogram) and dominant directionality of the active intensity (azigram) on each sensor. Within each azigram, sets of time-frequency cells associated with transient energy arriving from a consistent azimuthal sector are identified. Binary image processing techniques then link sets that share similar duration and bandwidth between different sensors, after which the algorithm triangulates the source location. Unlike most passive acoustic detectors, the threshold criterion for this algorithm is bandwidth instead of pressure magnitude. Data collected from shallow coral reef environments demonstrate the algorithm's ability to detect SCUBA bubble plumes and consistent spatial distributions of somniferous fish activity. Analytical estimates and direct evaluations both yield false transient localization rates from 3% to 6% in a coral reef environment. The SNR distribution of localized pulses off Hawaii has a median of 7.7 dB and interquartile range of 7.1 dB.

 . https://doi.org/10.1121/10.0003382

(Received 29 August 2020; revised 16 December 2020; accepted 6 January 2021; published online 2 February 2021)

[Editor: Karim G. Sabra] Pages: 770–787

I. INTRODUCTION

Coral reefs are amongst the most biodiverse ecosystems in the ocean and host the greatest density of marine life in the photic zone. Coral reef ecosystems are also under direct threat from both direct human impacts (overfishing, pollution, tourism, and overuse) (Pandolfi *et al.*, 2005) and indirect human impacts (sea surface temperature warming and ocean acidification from increased atmospheric carbon dioxide) (Glynn 1993; Brown 1997; Kleypas and Yates, 2009). Scientists using SCUBA or snorkeling equipment have studied coral reef ecology extensively, typically by means of direct sampling and experiments (i.e. Preskitt *et al.*, 2004). In addition to understanding reefs, timely monitoring is an essential component of coral reef conservation to quantify impacts and allocate mitigation resources (Pandolfi *et al.*, 2005). Examining ambient coral reef sounds offers the opportunity to understand reef characteristics in a non-invasive fashion and is of particular interest at night when visual methods are not possible without adding light and likely altering behavior of native biota. Previous work has shown linkages between coral reef acoustic signatures and predominant frequencies and relative reef health (Freeman and Freeman, 1993; Bertucci *et al.*, 2016), a consistent evening chorus after sunset found on reefs around the world (McWilliams *et al.*, 2017; Freeman *et al.*, 2014), and that reef sound provides a critical cue for pelagic coral (Vermeij *et al.*, 2010) and fish larvae (Gordon *et al.*, 2019) to settle. The latter suggests that sound is not just a byproduct of reef activities, but a critical component of a healthy reef that results in continued recruitment of juvenile corals and fish to maintain the density of life expected of such an ecosystem (i.e., Gordon, 2020). As such, understanding the variation and expected features of coral reef soundscapes, including the heterogeneity of sound on a single reef, are a critical component to understand coral reef ecosystems as a whole. To date, however, the vast majority of acoustic research into coral reef environments has focused on detection, and not localization, of ambient sound sources in this complex ecosystem. Passive acoustic detection of underwater biologic signals generally relies on applying a threshold to the acoustic pressure time series measured by a hydrophone (Zimmer, 2011). This threshold can be expressed in the time or frequency domain and defined either in terms of an absolute pressure or a signal-to-noise ratio (SNR), where the pressure is normalized relative to an estimate of the background noise spectrum. In the latter case, the threshold

[a)]Electronic mail: athode@ucsd.edu

[b)]ORCID: 0000-0002-0471-062X.

https://doi.org/10.1121/10.0003382

does not have to be adjusted in response to variations in ambient sound levels, a method defined as the "clutter map constant false alarm rate (CFAR)" processor in the radar literature (Nitzberg, 1986). All standard commercial bioacoustics software use CFAR amplitude thresholding as their fundamental detection strategy (Gillespie *et al.*, 2008). Detected sounds are then typically localized by cross-correlating time series from the same event detected on different hydrophones in order to determine relative time-of-arrival (RTOA) of a transient signal, which can then be transformed into a location estimate. Both whales (Wilcock, 2012) and fish (Putland *et al.*, 2018) have been localized by these approaches.

The CFAR approach works well whenever background noise levels are statistically stationary, and the number of transients per unit time is relatively low. However, in certain environments such as coral reefs, biological transient signals comprise a significant fraction of the total ambient sound. In these circumstances, the ambient levels are statistically non-stationary, and establishing a reference background spectrum becomes difficult. Even when the SNR threshold can be established, the detector will often reject numerous low-SNR transients, which comprise a significant portion of the total transients available. For example, Fig. 1, which shows the cumulative density function for 965 manually-identified fish transients on a Hawaiian coral reef (discussed in Sec. III), illustrates that at least 50% of valid localizable transients have SNR levels below 7 dB, and over a quarter have SNR values below 3 dB. A 3 dB CFAR detector threshold would be overwhelmed with spurious detections.

Coral reefs are also challenging environments for implementing RTOA localization methods because the low-frequency sounds characteristic of fish activity experience significant distortion, absorption, and interference when propagating through shallow waters over the complex bathymetry of a coral reef. As a result, the correlation (and thus RTOA estimate) between signals arising from the same event is often low, even when detected on sensors separated by only a few tens of meters. Figure 1(b) illustrates an example of this situation for the same 965 fish sounds in Fig. 1(a) by plotting the cumulative probability distribution of the Pearson's correlation coefficient between manually-matched sounds recorded on sensors 19 m apart. Another (dashed) distribution is also plotted for sound samples randomly correlated between the sensors. The similar distributions between the random and correct associations indicate that fish sounds detected on separate sensors are often significantly decorrelated to the point where it becomes difficult to obtain precise RTOA measurements between conventional hydrophones.

This paper presents an alternative approach to acoustic transient detection and localization that uses signal azimuth, and not SNR, through the use of vector sensors instead of hydrophones. Passive underwater acoustic vector sensors have the ability to estimate both acoustic pressure and vector particle velocity from a single point (D'Spain *et al.*, 1991; Nehorai and Paldi, 1994; Deal, 2018). Two-dimensional (2-D) particle velocity sensors have been incorporated inside Directional Low Frequency Analysis and Recording (DIFAR) sonobuoys for decades (Holler, 2014), as well as commercial autonomous recording packages such

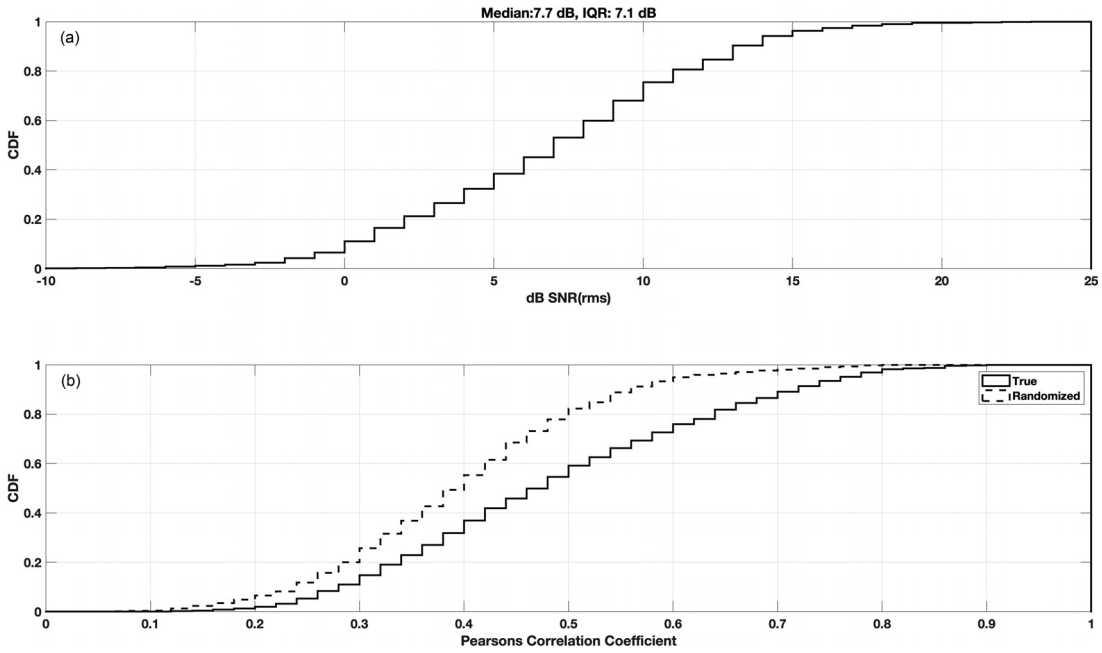

FIG. 1. Data collected on a Hawaiian coral reef illustrate motivation for an alternative to an SNR-based detector. (a) Cumulative density function of SNR for 965 manually-identified fish sounds collected over three non-contiguous five-minute intervals on February 29, 2020. (b) Cumulative density function for Pearson's correlation coefficient between signals measured on two sensors 19 m apart. The solid line represents correctly-matched signals between sensors; the dashed line represents correlations between randomly-matched signals that occur within 150 s of each other. Details on the manual analysis are provided in Sec. III C 3.

as the Directional Autonomous Seafloor Acoustic Recorder (DASAR) (Greene *et al.*, 2004). More recent versions of these sensors exhibit increased monitoring bandwidth (Shipps and Deng, 2003) and sensitivity along three dimensions. Vector sensors have advantages over conventional hydrophones in terms of array gain (Cray and Nuttall, 2001; D'Spain *et al.*, 2006), directional noise suppression (Thode *et al.*, 2016), geoacoustic inversion (Shi *et al.*, 2019; Dahl and Dall'Osto, 2020), and resolution of left/right ambiguity in towed passive acoustic arrays (Thode *et al.*, 2010). They also have advantages when localizing signals, in that only three widely spatially-separated sensors are required to successfully triangulate signals in a 2-D horizontal plane, and the sensors do not need to be precisely time-aligned (Hawkes and Nehorai, 2003). Conventional hydrophones, by contrast, require precise time-alignment in order to measure the relative arrival time of a signal between sensors, and also require at least four sensors to reliably locate a source in a horizontal plane *via* hyperbolic localization methods. The use of triangulation vs cross-correlation for signal localization also simplifies the automated processing of acoustic transients (Thode *et al.*, 2012).

Here, we demonstrate how vector sensor processing can further simplify automated processing and localization in nonstationary ambient sound environments, whenever the transients of interest are restricted to impulsive signals collected on sensors separated from each other by roughly 60 meters or less. For our purposes, we define a transient as a signal that displays an instantaneous bandwidth equal to the total signal bandwidth during any point of its reception. We thus exclude frequency-modulated tones and sweeps from the current discussion, although extending the algorithm described here to these signals is conceptually straightforward.

Section II describes the algorithm in detail, and then Sec. III displays the results of the algorithm applied to DASAR data collected in a shallow coral reef environment. The acoustic ambient sound environment in this biome is highly non-stationary and dominated by numerous transient pulses from a variety of natural and artificial sources. Bulk localization results are presented for SCUBA diving operations and vocalizing fish, where independent position data is available for the former. An Appendix analyses the false alarm rates produced by the method in the presence of a diffuse ambient sound field with a statistically stationary directionality.

II. METHODS

A. Definitions and linking algorithm

Figure 2 illustrates the geometry between two vector sensors, labeled A and B. The x and y axes lie parallel to the ocean bottom, and the angles are azimuths that follow the compass convention (increasing clockwise from the y axis). Any vertical dimensions, angular elevations, or vertical particle velocities are ignored in this paper.

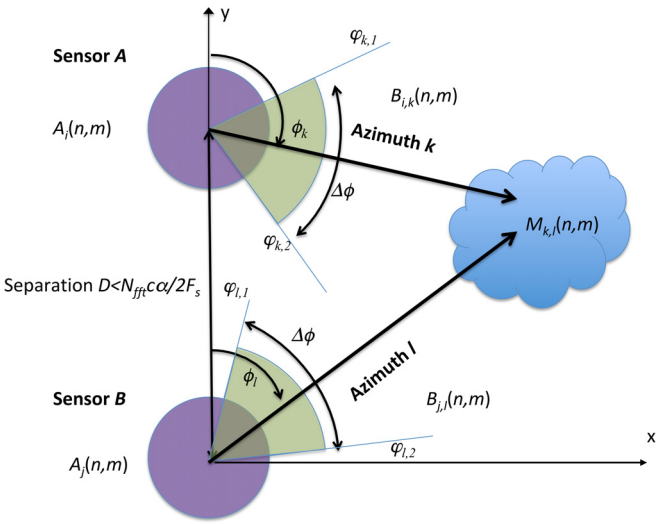

FIG. 2. (Color online) Definition of terms used in localization method. The vertical axis is aligned with true north.

The pressure and two orthogonal particle velocity times series are designated as $p_i(t)$, $v_{i,x}(t)$, and $v_{i,y}(t)$, where the subscript i represents sensor A or B. Each time series is divided into N potentially-overlapped partitions, each with N_{fft} samples. After applying a fast Fourier transform (FFT) of length N_{fft} to the nth time chunk, one obtains a frequency domain estimate of the pressure $P_i(n,m)$ and particle velocities $V_{x,i}(n,m)$ and $V_{y,i}(n,m)$, where time index n represents a given time chunk, and m is a frequency index that ranges from 0 to $N_{fft}/2$.

Two matrices can be defined from these measurements: a spectrogram $S_i(n,m) = |P_i(n,m)|^2$, which measures the square magnitude of the measured acoustic pressure versus time and frequency, and an "azigram" (Thode *et al.*, 2019) $A_i(n,m)$, which is constructed by defining the "active intensity" of the acoustic field at sensor I,

$$I_{i,q}(n,m) = Re\left[P_i(n,m)V^*_{q,i}(n,m)\right], \tag{1}$$

where the index q represents either the x or y direction, and the symbol "*" indicates complex conjugation. The active intensity measures the in-phase product of p and v_q, which indicates the portion of the acoustic field that is actively transporting acoustic energy through the measurement point along axis q.

The azigram $A_i(n,m)$ at sensor i is then defined as

$$A_i(n,m) = \tan^{-1}\left[I_{i,x}(n,m)/I_{i,y}(n,m)\right], \tag{2}$$

and represents the dominant azimuth from which acoustic energy at time chunk n and frequency index m is propagating from along the horizontal plane. The use of the form $tan{-}1(x/y)$ produces an azimuth that follows the compass convention (clockwise azimuth from y axis). Both I and A can be treated as 2-D images and manipulated by binary image processing techniques, including thresholding and segmentation (Gonzalez and Woods, 2002). For example, at

https://doi.org/10.1121/10.0003382

sensor i, the azimuthal sector ϕ_k is defined by the azimuthal angular range $[\varphi_k - \Delta\varphi/2, \varphi_k - \Delta\varphi/2]$, with $\Delta\phi$ defining the width of the sector. Sectors can be defined to overlap, i.e., $(|\phi_{k+1} - \phi_k| < \Delta\phi)$. One can thus construct a binary image $B_{i,k}(n,m)$ that indicates which time/frequency cells in the azigram $A_i(n,m)$ arrive near an azimuth ϕ_k via

$$B_{i,k}(n,m) = \mathbb{I}\big(\varphi_k - \Delta\varphi/2 < A_i(n,m) < \varphi_k + \Delta\varphi/2\big), \tag{3}$$

where $\mathbb{I}$ is the indicator function. The binary image thus represents all time-frequency cells where energy is propagating through sensor i from an azimuthal sector centered around ϕ_k. All time-frequency components associated with a transient signal from a discrete source (as opposed to a distributed source) would be contained in the same binary image.

Next, we define a "bidirectional mask" $M_{k,l}$ as the intersection of two binary images, evaluated for every time-frequency cell (n,m),

$$M_{k,l}(n,m) = B_{A,k}(n,m) \cap B_{B,l}(n,m). \tag{4}$$

A given mask $M_{k,l}$ is activated at time-frequency cell (n,m) if energy is detected arriving from sector k on sensor A and arriving from sector l on sensor B. Figure 2 illustrates the various definitions of these 2-D matrices, and also shows that each mask represents a hypothesis that a signal is arriving from a particular 2-D region in the environment. Thus, a mask tests whether energy in particular time-frequency cells in both sensors' spectrograms represents a potentially *physically localizable* source, regardless of the spectrogram magnitude corresponding to that cell. As one example distant signals will be present in the mask $M_{k,k}(n,m)$, where active azimuthal sectors on the two sensors are identical $(k = l)$. For N_{az} azimuthal sectors, there exist N_{az} binary images (B) for each sensor and $N_{az}{}^2$ mask images (M) for every pairwise combination of sensors.

Once a set of masks $M_{k,l}(n,m)$ has been generated, a detection function D is derived from each mask using $D_{k,l}(n) = \Delta f \sum_{m1}^{m2} M_{k,l}(n,m)$, where m_1 and m_2 are the upper and lower frequency indices that define a bandwidth of interest, and Δf is the spacing between adjacent frequency bins in Hz. Thus, D is an estimate of the instantaneous bandwidth of a signal arriving from the 2-D region associated with the mask. For a given time output n, the $D_{k,l}$ value is checked to determine that it represents a local maximum with respect to time and arrival angle by confirming the following:

$$D_{k,l}(n) > \max\big\{D_{k+1,l}(n), D_{k-1,l}(n), D_{k,l-1}(n), D_{k,l+1}(n)\big\}, \tag{5a}$$

and

$$D_{k,l}(n) > \max\big\{D_{k,l}(n-1), D_{k,l}(n+1)\big\}. \tag{5b}$$

The mask associated with the peak value of D is then selected for subsequent processing of the detected transient.

A simple summation across the frequency band is used when defining D because acoustic transients can display "notches" in active intensity resulting from destructive interference with multipath arrivals, and thus a transient appearing in a bidirectional mask may be divided into several disconnected regions along the frequency axis. To date, attempts to use morphological convolution or other image processing techniques (Gonzalez and Woods, 2002) to both connect regions and remove isolated small time-frequency regions seems less robust than a simple summation across bandwidth. The fact that D is a measure of the instantaneous bandwidth at a single moment implies that this crude detector performs best for short-duration, broadband signals, i.e., impulsive transients, and not tones or frequency-modulated signals. Detectors for these latter calls might be created using more sophisticated image segmentation techniques.

Analogous to pressure-only acoustic detectors (Zimmer, 2011), a threshold D_{thresh} is defined such that the origin and duration of a valid transient are flagged when D exceeds D_{thresh}. However, in contrast to standard detectors, which set a threshold in terms of the *magnitude* of acoustic pressure or its SNR, D_{thresh} has units of frequency and can be interpreted as the minimum required instantaneous bandwidth of potential transient signal. Both thresholds are intuitively related because as the SNR of an impulsive signal decreases, its peak instantaneous bandwidth also tends to shrink, and thus a lower value of D_{thresh} becomes required for detection. A lower SNR will also increase the spread of the azimuthal measurements because the precision of the active intensity measurements decreases with decreasing SNR. Thus, increasing the sector width $\Delta\varphi$ and decreasing D_{thresh} permits the detection of signals with lower SNR, but at the cost of increasing the possibility of spurious matches between random azimuthal values produced by a diffuse, perhaps directional, noise background. The Appendix estimates the false detection rate of the two-sensor scenario as a function of the degree of directionality of the background ambient sound for various combinations of sector width and D_{thresh}.

B. Localization algorithm

The procedure in the previous subsection works for identifying the presence of a localizable impulse, but the angular width of the sector $\Delta\phi$ is generally so large that a more precise azimuthal estimate is needed at both sensors in order to obtain a viable position. To that end, once a time interval containing a pulse has been identified (from FFT sample n_1 to n_2), a time subset of the bidirectional mask can be combined with the original two azigrams to yield a precise and robust estimate of each sensor's azimuth. For example, for sensor i,

$$\varphi_i = \text{circular median}\big\{A_i(n_1 : n_2, m) \cap M_{k,l}(n_1 : n_2, m)\big\}, \tag{6}$$

which can be interpreted as taking the circular median of azigram time-frequency cells that satisfy the criteria of a given mask.

To this point, the discussion has focused on two sensors. In principle, masks could be generated that combine three or more different azimuthal sectors from three or more sensors, but the resulting combinatorial complexity is inefficient. To save computational effort, a mask derived from two sensors can be directly applied to any additional sensors in the array, sidestepping the combinatorial issue.

Taking the circular median of the azimuthal estimates of the azigram time-frequency cells is crucial for making the localization process robust. For example, Eq. (6) could be reformulated to use the weighted average of the azigram samples to generate the final azimuth, where the weight applied to each time-frequency cell is the dB magnitude of the active intensity of that cell. However, this approach is vulnerable to azimuthal outliers when a mask is applied to additional sensors beyond the two originally used to derive it because typically several time-frequency cells in a third azigram will have azimuths that lie outside the target sectors, and these outliers can strongly bias an averaged bearing estimate.

Once the azimuths associated with the detection have been refined at each sensor, the 2-D localization is straightforward. A maximum-likelihood triangulation method is provided by Lenth (1981). This method, extensively described in that reference, permits the uncertainty of the resulting position to be quantified when three or more sensors are used in the triangulation.

C. Practical implementation issues

A fundamental requirement of this processing approach is that each transient must be present over the same time indices (n_1:n_2) on both sensor azigrams. Thus, the azigrams must be time-aligned to within an FFT window. Whenever the ambient sound environment is highly non-stationary, and has several transient pulses embedded over the course of a single spectrogram, then a simple spectrogram correlation (Mellinger and Clark, 2000) over a 60 s interval is sufficient to ensure this alignment.

Even with this spectrogram correlation step, a practical limit on sensor separation remains, in order to ensure that spectrograms from two sensors will display the same relative timing (FFT snapshots) between groups of pulses arriving from multiple directions. For two sensors separated by a distance L, a transient arriving from the endfire direction on the more distant sensor will experience a maximum time delay of $+L/c$ relative to the closer sensor, where c is the local sound speed. A signal arriving from the opposite direction will display a $-L/c$ time shift between sensors. If the sampling rate is F_s samples per second, then the number of samples in a FFT snapshot, N_{fft}, needs to be greater than $2\,F_sL/c$, the maximum number of samples a transient arriving from any direction on the second sensor can shift relative to the first sensor. The exact amount of overlap between successive FFT samples is not crucial but should be at least 50% in order to ensure that an impulse will arrive on both sensors within the same time window. This reasoning yields an expression for the maximum valid separation between sensors for a given FFT size,

$$D \leq \frac{N_{fft}c}{2\alpha F_s} = \frac{\Delta T c}{2\alpha}, \tag{7}$$

where ΔT is the FFT sampling length in seconds, and α is a safety factor that represents the number of multiples of the two-way travel time used to derive the FFT length. For example, if $c = 1500$ m/s, $F_s = 1$ kHz, $N_{fft} = 256$, and $\alpha = 3$, then $D = 64$ m. Equation (7) provides an upper bound on the sensor separation required to ensure that the relative arrival times of all transients on two sensor spectrograms/azimuths are unaffected by their location. In the data examples that follow, the sensor separation is at most 40 m, so Eq. (7) is satisfied.

When the sensor packages are deployed on the sea floor, their relative spacing needs to be determined to within a meter, and their azimuths relative to each other need to be determined to within a few degrees in order to localize accurately out to five times the sensor separation (Lenth, 1981). The azimuth measurement requirement is generally not needed for conventional hydrophones.

D. Parameter values

Table I lists the various parameters used in the particular application of the algorithm presented here. All azigrams were constructed using a 256 point FFT (0.256 s duration for a 1 kHz sampling rate) with 90% overlap between samples ($\Delta f = 3.9$ Hz). Detections were monitored between 100 and 450 Hz (bandwidth B of 350 Hz), because in shallow coral reef environments flow noise is often present below 100 Hz.

The key parameters for the algorithm are the azimuthal sector width $\Delta\phi$ and the detection threshold D_{thresh}. The detailed analytical analysis in the Appendix finds multiple combinations of these parameters yield the same estimated false detection and localization rates, given intrinsic ambient sound directionalities that can be modeled as a von Mises (VM) distribution (Mardia and Jupp, 2009) with a concentration parameter κ. For example, when $\kappa = 0.8$ (representative of the noise directionality observed off Hawaii Island), sector width/threshold combinations of 90°/120 Hz and

TABLE I. Baseline parameter values used for the algorithm in this paper. For SCUBA air plume detection, a 250 Hz detection threshold was used. Some analyses also used a sector width of 45° and a detection threshold of 50 Hz.

Parameter	Value
FFT size, overlap	256, 90
Sector width $\Delta\phi$ (deg)	90
Sector separation (deg)	45
Threshold D_{thresh} (Hz)	120
Frequency range of interest B (Hz)	100–450
Max Detection Time (s)	2
Min separation between detections (s)	0.06

https://doi.org/10.1121/10.0003382

45°/50 Hz are predicted to generate relatively low false detection rates of 2–3 per minute. Table I thus sets the baseline threshold to 120 Hz and the baseline sector width to 90°, with sectors overlapping by 50% [$(\phi_{k+1} - \phi_k)/\Delta\phi = 0.5$]. For situations where very high SNR signals are to be detected (e.g., SCUBA plumes in Sec. III B), the threshold was increased to 250 Hz. Detections were limited to 2 s maximum duration, and subsequent detections needed to occur at least 60 ms after the end of the previous detection.

III. ILLUSTRATIVE EXAMPLES

A. Equipment and deployment configurations

Experimental data to demonstrate this localization procedure were collected at a shallow reef environment off the coast of Kona, Hawaii, between February 22 and March 13, 2020. The data were collected at two locations spaced 233 m apart, with three DASAR packages deployed at each site.

DASARs are autonomous acoustic recording packages equipped with an omnidirectional acoustic pressure sensor (sensitivity of −149 dB re V/ 1 μPa) and two horizontal directional sensors capable of measuring the north-south and east-west components of acoustic particle velocity (Greene *et al.*, 2004). These packages have been used extensively for research on bowhead whale and ambient sound studies in the arctic (Blackwell *et al.*, 2015; Thode *et al.*, 2020). A DASAR samples each time series at 1 kHz with a maximum usable acoustic frequency of 450 Hz due to anti-aliasing filter roll-off.

The first site, dubbed "Block2020," was located west of a shallow ridge at 19.76856 °N, 156.05261 °W between February 22 and March 1, while the second "Shark2020" site lay in a sandy "bay" nearly surrounded by coral reef at 19.77016 °N, 156.05222 °W and was deployed between February 22 and March 13. The Block site deployed three DASARs along a rough N-S line parallel to the reef ridge at a depth of 21 m (70 ft), in order to allow the largest triangulation baseline relative to the reef, while the Shark site arranged the DASARs in a rough triangle at 20 m depth (66 feet), in order to permit localization in all directions (Fig. 3).

The DASARs were placed using SCUBA operations, and their separations were measured directly underwater. Their azimuths relative to each other were measured from each DASAR using a line-of-sight magnetic compass. Unlike traditional RTOA localization methods, triangulation methods also require the azimuths (with respect to true north) of all lines connecting the DASARs to be measured. "Structure from motion" photo-mosaic methods (Ullman, 1979; Gracias *et al.*, 2003; Lirman *et al.*, 2007; Edwards *et al.*, 2017) were used to map out the position of instruments on the ocean floor relative to reference ceramic plate markers. While the primary purpose of the plates was to aid in photomosaic reconstruction, divers also generated SCUBA bubble plumes over one plate in order to generate test signals at a known location (Fig. 3, bottom).

The sensor orientations on the ocean floor were determined by driving a small boat in a 100-m radius circle

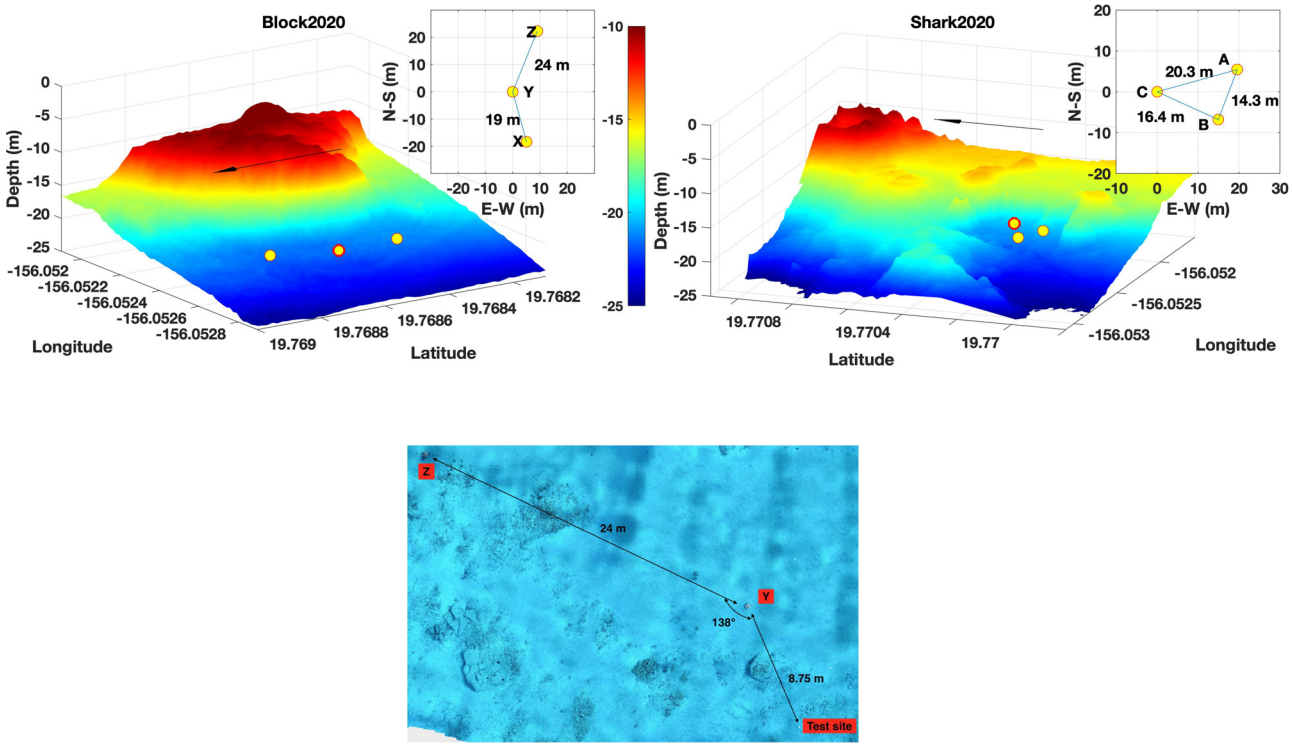

FIG. 3. (Color online) Locations of passive acoustic deployments off Kona, Hawaii. Top row shows LIDAR[1] bathymetry surrounding Block2020 (top left) and Shark2020 (top right) deployments. Black arrows show direction of true north, and insets show detailed layouts of sensors at both locations. Bottom: Photomosaic of ocean floor at Block2020 site, showing DASARs Z, Y, and a ceramic plate (Plate 61) used as a test site for generating acoustic pulses from SCUBA air plumes.

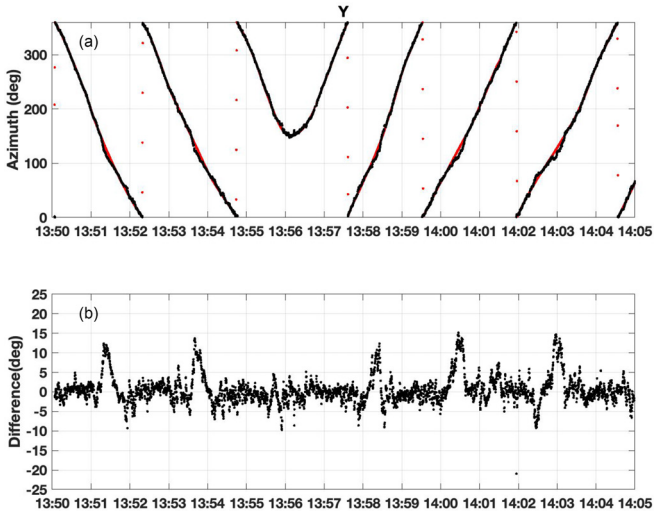

FIG. 4. (Color online) Results of DASAR orientation calibration using boat tracks over 15 min at Block site, 22 February 2020. Top: Interpolated GPS azimuth of boat relative to DASAR Y (red) overlain with derived acoustic azimuth from DASAR after calibration (black). Bottom: Difference between acoustic and GPS azimuth vs time. Azimuths are measured with respect to true north (0°).

around the deployment, as measured from the deployment center, and logging the global positioning system (GPS) position every second (Fig. 4). The boat performed three clockwise circles and three counterclockwise circles and then used a brute-force optimization to fit the relative bearings to the boat, measured by each sensor from the integrated active intensity between 75 and 250 Hz. The optimization solved the relative clock offset between the internal DASAR clock and GPS time, the 2-D sensor position on the ocean floor, and sensor orientation. Performing both clockwise and counterclockwise circles permitted clock offset and sensor orientation to be distinguished without ambiguity.

B. Demonstration with air release at a controlled location

At 8:20:28 local time on 24 February 2020, SCUBA divers generated six controlled air bubble plumes at the Block2020 site over a ceramic tile (61) placed 8.75 m from DASAR Y at 21 m (73 feet) depth (Fig. 3, bottom). Figure 5 displays the spectrograms and azigrams for each DASAR over a seven-second window centered on the resulting bubble plumes. The different colors on each azigram demonstrate how the bearings to the plumes differ on each DASAR.

The initial linking algorithm was performed between DASAR Y and Z. Figure 6 displays two example binary images B, the resulting bi-directional mask M, and the detection function D for two angular sectors. A 90° sector width with 50% overlap generated eight binary

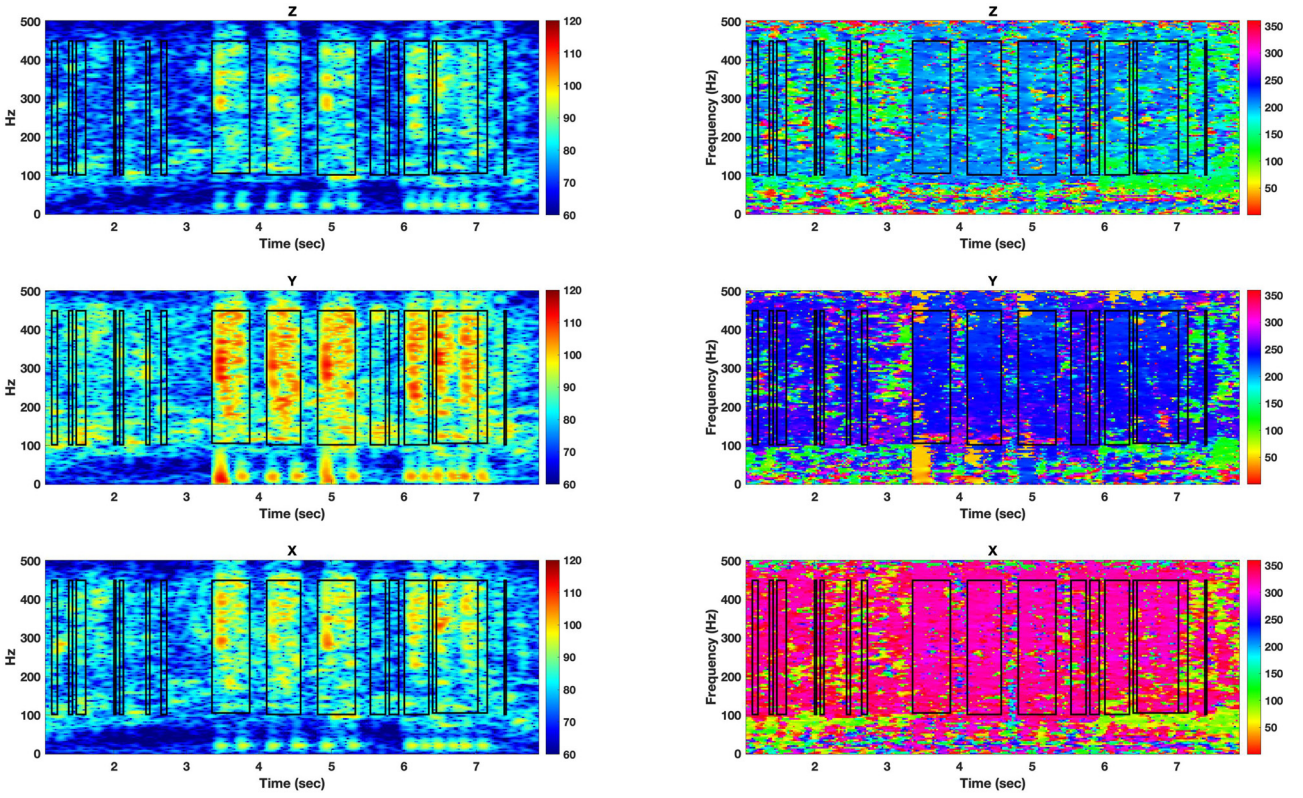

FIG. 5. (Color online) Six controlled SCUBA bubble plumes released at the Block2020 test site. Each row represents a different DASAR, with the left column showing conventional spectrograms (color scale in units of dB re 1 uPa^2/Hz) and the right column showing corresponding azigrams, with the HSV scale displaying degrees from true north). The black bounding boxes are the detector outputs using the parameters in Table I, but with $D_{thresh} = 200$ Hz. Detection masks M were obtained by comparing DASARs Y and Z (top and middle rows).

https://doi.org/10.1121/10.0003382

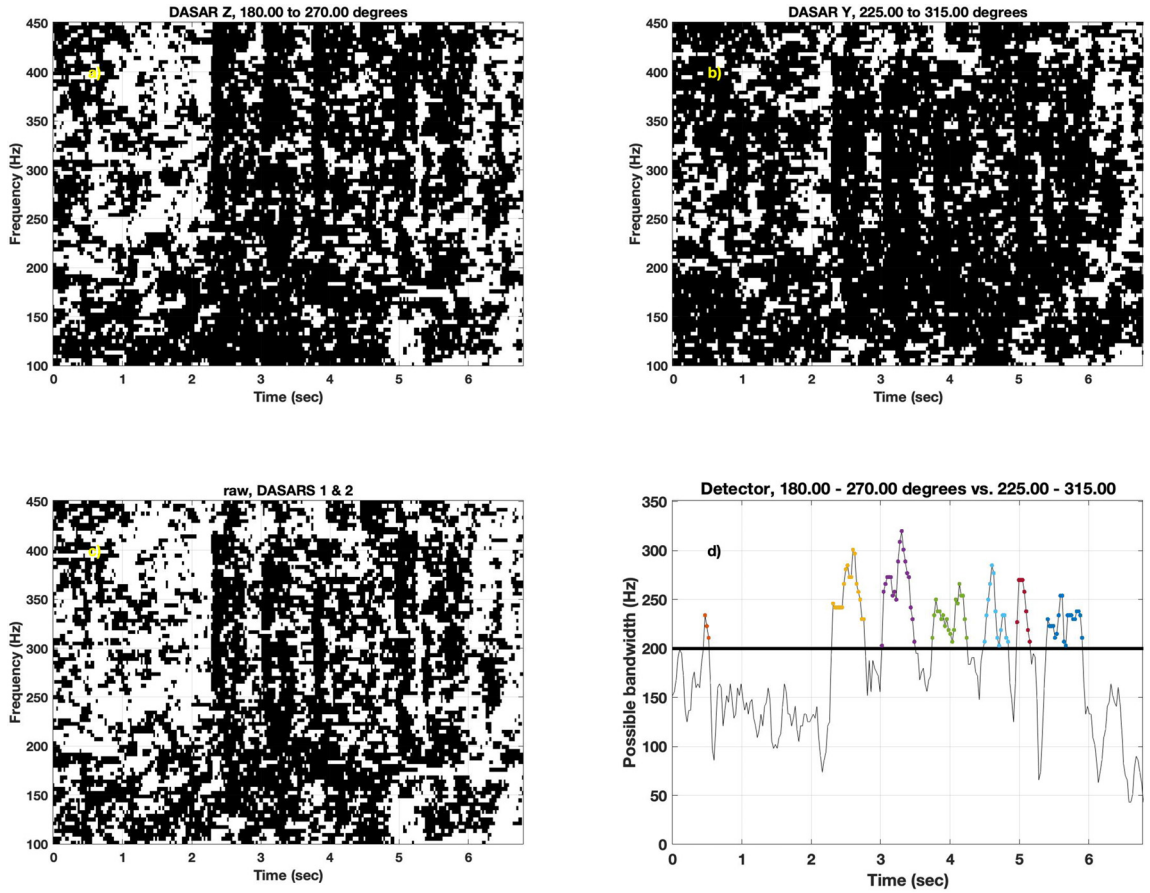

FIG. 6. (Color online) Example of binary images B, mask M, and detection function D generated from the azigrams in Fig. 5 for DASARs Z and Y. (a) Binary image $B_{Z,225^\circ}$; (b) binary image $B_{Y,270^\circ}$; (c) bidirectional mask $M_{225^\circ,270^\circ}$; (d) detection function $D_{225^\circ,270^\circ}$ summed over each column of (c). Horizontal line indicates a 200 Hz threshold, and each color indicates a different transient flagged by the detector.

images per DASAR, resulting in 64 different masks. These masks were then applied to DASAR X to generate a third bearing estimate, and the three bearings were then used to triangulate both a position and the 90% confidence ellipse.

Figure 7 displays the localization results for the six air pulses, along with the actual position indicated by the photo mosaic. The acoustically localized positions are within 1–2 m of the true position, but the positional uncertainty spans from roughly 10 cm to several meters, depending on how well the bearings intersect.

C. Fish activity

The algorithm was applied to the complete data sets at both the Block and Shark sites in order to detect low-frequency pulses characteristic of coral reef fish (Tricas and Boyle, 2014), using different pairs of sensors to generate the initial masks. Both the 90°/120 Hz and 45°/50 Hz parameter combinations were tested, with the former being the default.

1. Raw counts

Table II summarizes the number of raw linkages and successful localizations of the various combinations, when measured over 24 h from midnight 28 February through midnight 29 February. The table also lists the number of "close" localizations occurring within 50 m from the array center, as well as the number of these close localizations that had a 90% confidence ellipse smaller than 5 meters.

2. Spatial distributions

Figure 8 shows 2-D distributions of pulse locations for several configurations at the Block site, while Fig. 9 shows the associated runs at the Shark site. Both figures use a 50×50 meter grid with a 5 m grid cell size. Acoustic activity at the Block site was concentrated at one location at the southern end of the reef, while the Shark site had a more complex spatial distribution. Figures 10 and 11 show direct comparisons between the spatial distributions produced by the two parameter combinations, along with the percentage change in pulse detections that occurs when only high-precision positions are retained. These figures show how the fraction of high-precision localizations in a grid cell decreases with increasing distance from the DASAR array center.

3. False localization evaluation

The false localization rate of the system at the Block site was directly evaluated for the ZYX 90°/120 Hz

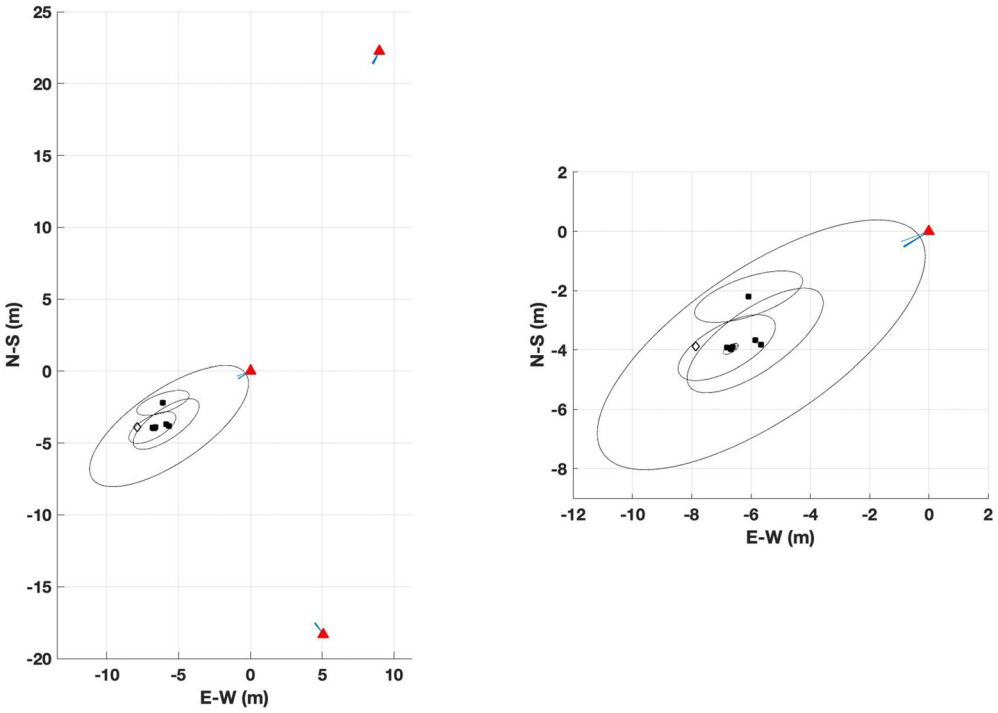

FIG. 7. (Color online) Acoustic localizations of six SCUBA bubble plumes (filled squares), along with the location determined by photomosaic (open diamond). Ellipses indicate 90% confidence interval of localizations. Triangles represent DASAR locations. Right plot is expanded version of left figure.

parameter combination across three different 5-min intervals on February 29, which yielded 1452 successful localizations. The first interval (370 samples) began at local midnight, the second interval (581 samples) began at 12:02 local time (a time of minimal transient production at the reef), and the third interval (501 samples) began at 19:30 local time (a time of peak transient production at the reef). The second interval had the highest number of localized samples due to the close presence of a singing humpback whale to the north of the deployment. Humpback whale song mostly consists of frequency-modulated signals, but the SNR of the signals were so strong that the instantaneous bandwidth of the signals often exceeded the detection threshold of 120 Hz.

To evaluate each localization, the pressure and particle velocity time series from each sensor was first extracted, all centered at the time where the peak value of D occurred at that sensor. The length of the time series was the duration of the detection, buffered by an additional 25 ms of data before and after the detection window. All time series were bandpass filtered over the bandwidth B listed in Sec. II D. In order to enhance the SNR of

TABLE II. Raw transient counts over 24 h on 29 February 2020 at Block Site, for various detector configurations. All configurations use a default sector width of 90 degrees and 120 Hz threshold, unless otherwise noted. The first two DASARs listed in a given combination produce the azimuthal masks that are then applied to all DASARs.

DASAR combination	Raw linkages	Successful localizations	Close localizations (<50 m range from array center)	Close localizations with <5 m uncertainty
Block				
ZY	147 369	107 518 (73%)	69 736 (65%)	NA
ZY (45°, 50 Hz)	322 101	207 608 (64%)	123 609 (60%)	NA
XY	153 984	114 571 (74%)	53 663 (47%)	NA
ZXY	113 473	113 473	60 348 (53%)	32 842 (54%)
ZYX	108 501	108 501	55 732 (51%)	29 131 (52%)
ZYX (45°, 50 Hz)	212 138	212 138	95 549 (45%)	40 879 (43%)
Shark				
CA	170 745	114 820 (67%)	47 970 (42%)	NA
CB	160 366	98 386 (61%)	54 932 (56%)	NA
CAB	115 131	115 131	46 268 (40%)	14 557 (31%)
CAB (45°,50 Hz)	247 780	247 780	119 701 (48%)	31 767 (26%)
BAC	103 765	103 765	44 876 (43%)	14 169 (32%)

https://doi.org/10.1121/10.0003382

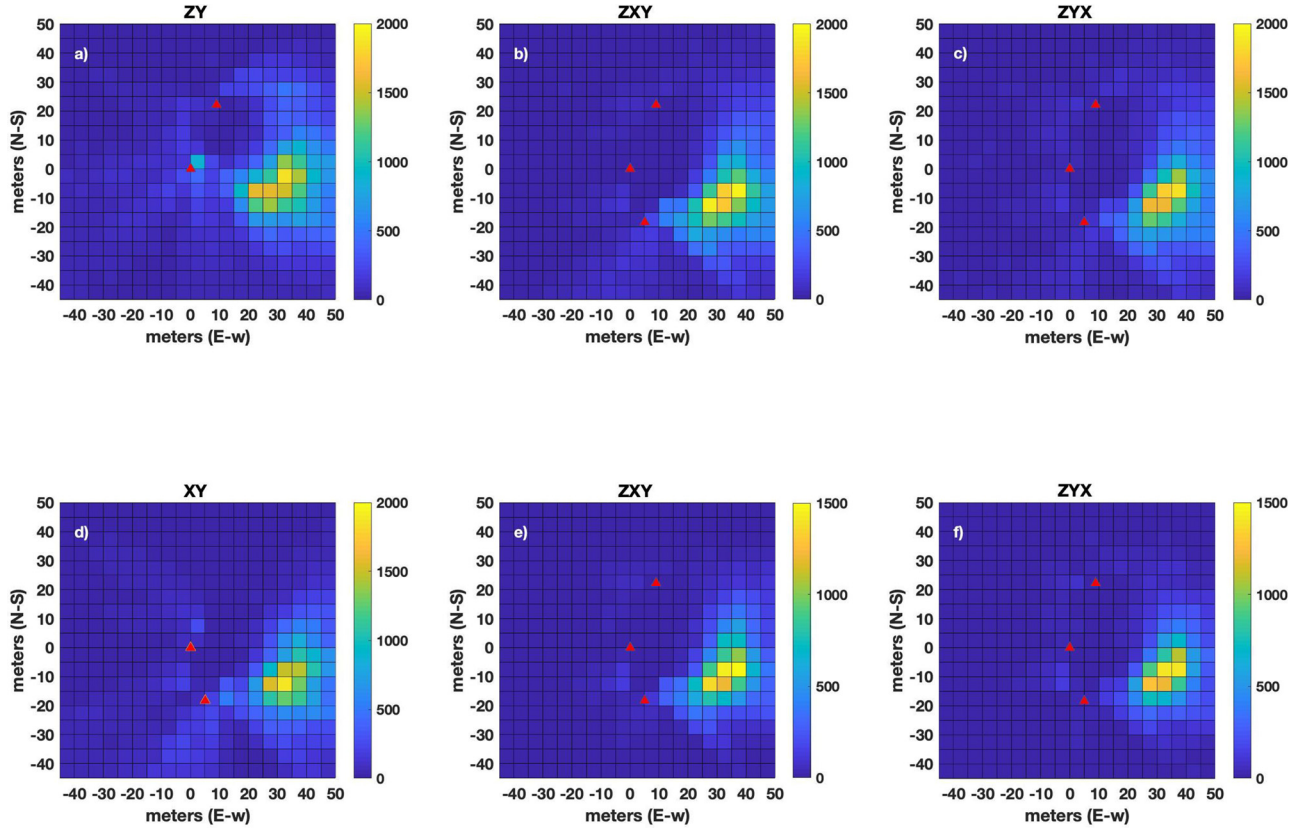

FIG. 8. (Color online) 2-D distributions of biological pulses around Block2020 site during 29 February 2020, using the default parameters for different combinations of DASARs to link and localize. Each grid cell covers 5 × 5 m. Distributions are shown for (a) Z and Y only; (b) Z and X used to generate the mask, which is applied to Y for additional bearing; (c) Z and Y used to generate a mask, applied to X; (d) X and Y only. Subplots (e) and (f) display scenarios (b) and (c), using only localizations with 90% confidence intervals that are less than 5 m.

signals arriving from the measured azimuth ϕ_i [Eq. (6)], the time series at sensor i were combined using additive beamforming,

$$x_{i,beam} = p_i(t) + Z_0\left[v_{i,x}(t)\sin\dot{\varphi}_i + v_{i,y}(t)\cos\dot{\varphi}_i\right], \tag{8}$$

where $\dot{\varphi}_i$ is the angle from Eq. (6) after rotating from the geographic coordinate axes to the sensor reference axes.

The three beamformed time series (one for each sensor) were then visualized through a high-resolution continuous wavelet transform (Torrence and Compo, 1998) that employed the analytic Morse wavelet (Lilly and Olhede, 2012), along with an azigram derived from the same wavelet decomposition. A manual inspection of these images determined whether the signals detected across the three sensors arose from a similar source.

When analyzed at these high time resolutions, the original FFT outputs of the directional detector were often resolved into multiple pulses. All such pulses arriving from the same direction (as determined from the azigram) were counted as one signal sample. The same signal had to be present on all three sensors to be classified as a valid detection.

Figure 12 provides three examples of manual evaluations of these often low-SNR signals, with two valid localizations shown in the first four columns, and an example of a false result shown in the last two columns. The manual evaluation was not straightforward; the propagation of these low-frequency (<200 Hz) signals in shallow water over a complex bathymetry created complex multipath interference effects that often modified the structure of a signal at a particular sensor. For example, the left two columns of Fig. 12 show how the peak frequency of the same single-pulse transient can vary by up to 100 Hz between two sensors. Despite these frequency variations, the presence of the same signal (true detection) could often be identified by the relative timings and durations of multiple pulses comprising the detected signal. For example, the middle two columns show an example of how four irregularly-timed pulses could be flagged between the sensors, despite the presence of other extraneous pulses on some sensors. Finally, the last two columns on the right show a false localization where three signals are arriving roughly from the same direction but share no temporal or spectral features in common between the sensors.

This manual review procedure, when applied to all 1452 localizations, found false localization rates of 6.4, 3.2, and 5.2% for the midnight, noon, and evening samples, respectively. The Pearson correlation coefficient was also computed between the time series, to determine whether false detections between sensors could be predicted by a low correlation coefficient. All false localizations had

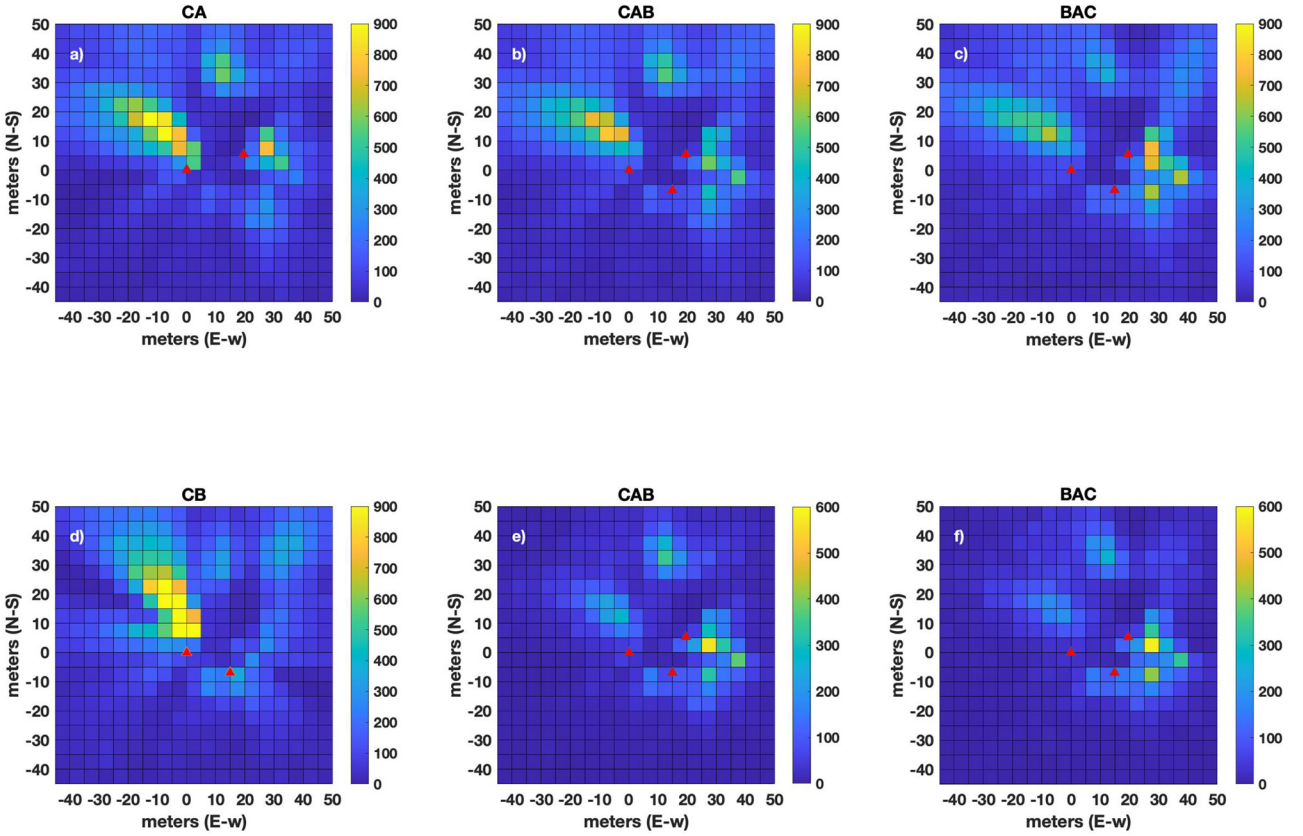

FIG. 9. (Color online) 2-D distributions of biological pulses around Shark2020 site during 29 February 2020 using different combinations of DASARs to link and localize. Each grid cell covers 5 × 5 m. Figure follows the same organization as Fig. 8.

coefficients less than 0.5; however, 50%–70% of all valid localizations also had correlation coefficients less than 0.5 [e.g., middle columns of Fig. 12; Fig. 1(b)], presumably due to the propagation effects mentioned above.

Figure 13 displays distributions of the detection duration and the root-mean-square (rms) SNR for the manually-analyzed samples, with noise estimated using the rms value of data in the time series buffer. False localizations tend to have shorter durations: 58% of false detections had durations less than 25.6 ms (one spectrogram time bin), while only 24% of true localizations were that short in duration. The median SNR for true localizations is only 3 dB larger than that for the false localizations, and the mode SNR is 9.5 and 6 dB, respectively. However, SNR values for true localizations exist below 0 dB. One reason that these small SNR values exist is that extraneous transients appear in the buffer, but the manual review does reveal that the method does successfully flag many low-SNR signals.

IV. DISCUSSION

The tabulation of impulsive sounds localized by various iterations of the algorithm in Table II shows that over a given day over 100 000 localizable sounds are measured at each site, which translates into a little over one per second. At both sites from roughly 2/3 to 3/4 of the sounds that pass the detector are localizable, a number dubbed the "localized fraction" from this point on.

Both the analytical analysis in the Appendix and the manual evaluations of Sec. III C 3 yield similar low false alarm rates for localizable signals at the Block site. For example, the analytical analysis of the 90°/120 Hz parameter combination predicts three false localizations per minute, or 4320 per day, if the coral reef background noise directionality is characterized by a VM distribution with a $\kappa = 0.8$. Table II shows that the ZYX run for that same parameter combination produced 108 501 successful localizations, which suggests that only 4.0% of the total are spurious detections. The direct evaluations of the false rate found similar values between 3% and 6%. The lower bound was from a time period with many humpback whale sounds, which were easy to detect and localize; the false rate during times when only fish activity exists was between 5% and 6%. The manual review also found that no simple metric exists for flagging spurious localizations: the Pearson correlation coefficient, duration, and SNR distributions substantially overlap between true and false locations. We hypothesize that the complex propagation environment often reduces the correlation of a low-frequency transient between sensors.

The Block site also shows that the particular order in which DASARs are used to generate the masks has only a mild effect on both the total detections and the localized fraction: the ZY, XY, ZXY, and ZYX combinations produce similar results. The addition of a third

https://doi.org/10.1121/10.0003382

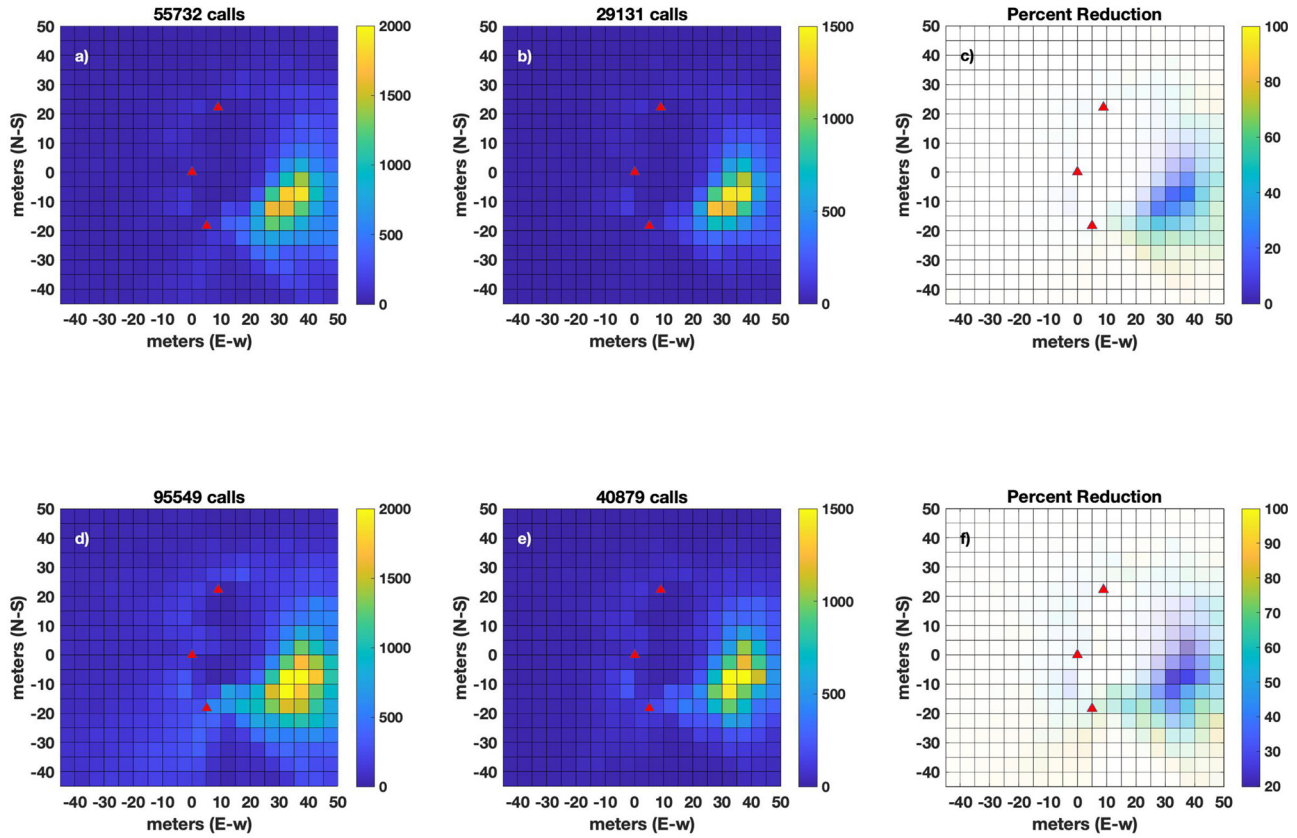

FIG. 10. (Color online) Comparison between two threshold levels at Block site on 29 February 2020, "ZYX" arrangement. The top row shows 2-D pulse distribution arising from applying a sector width of 90° and detection threshold of 120 Hz. The bottom row shows the results from applying 45° and 50 Hz. The left column shows all locations, the middle column shows locations with a 90% confidence interval (CI) less than 5 m, and the right column is the percentage difference between the two, with the opacity of a grid cell proportional to the call count within the cell.

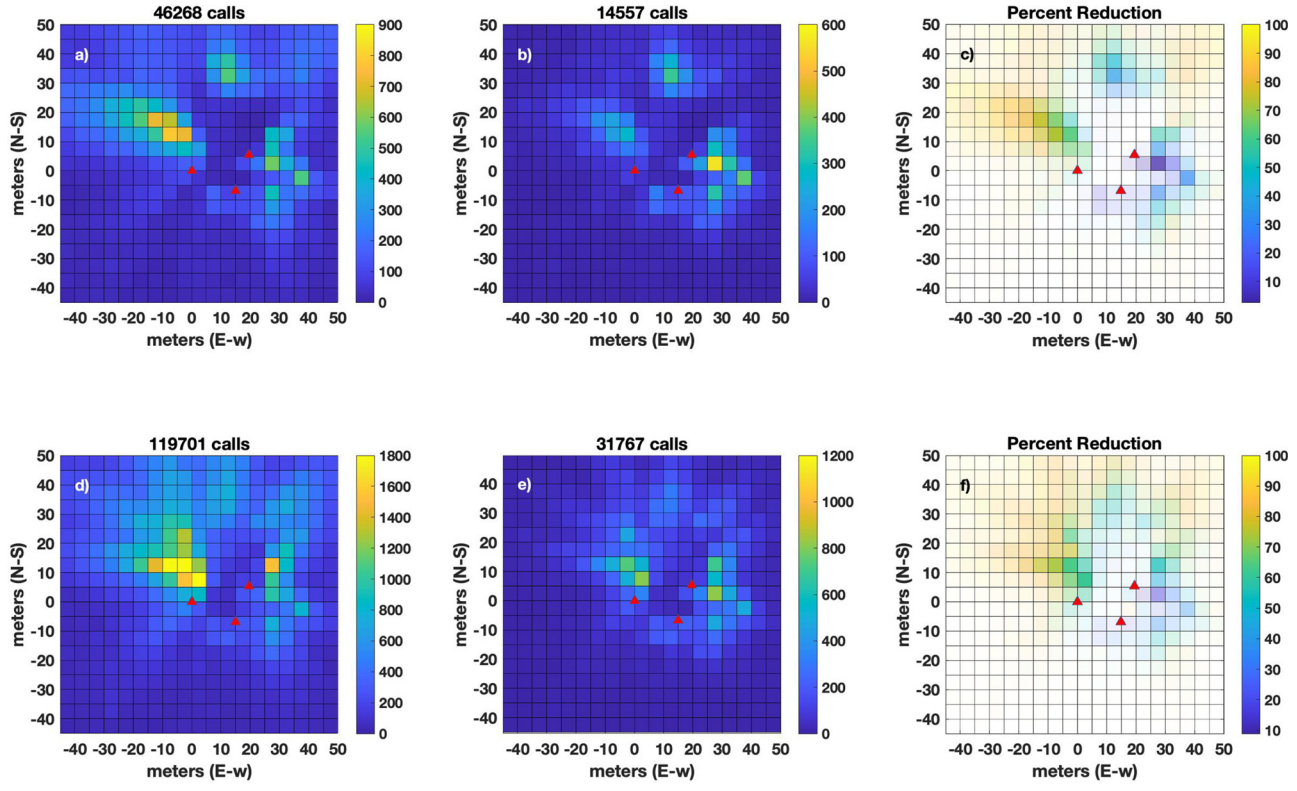

FIG. 11. (Color online) Same as Fig. 10, except applied to the Shark2020 site, "CAB" arrangement.

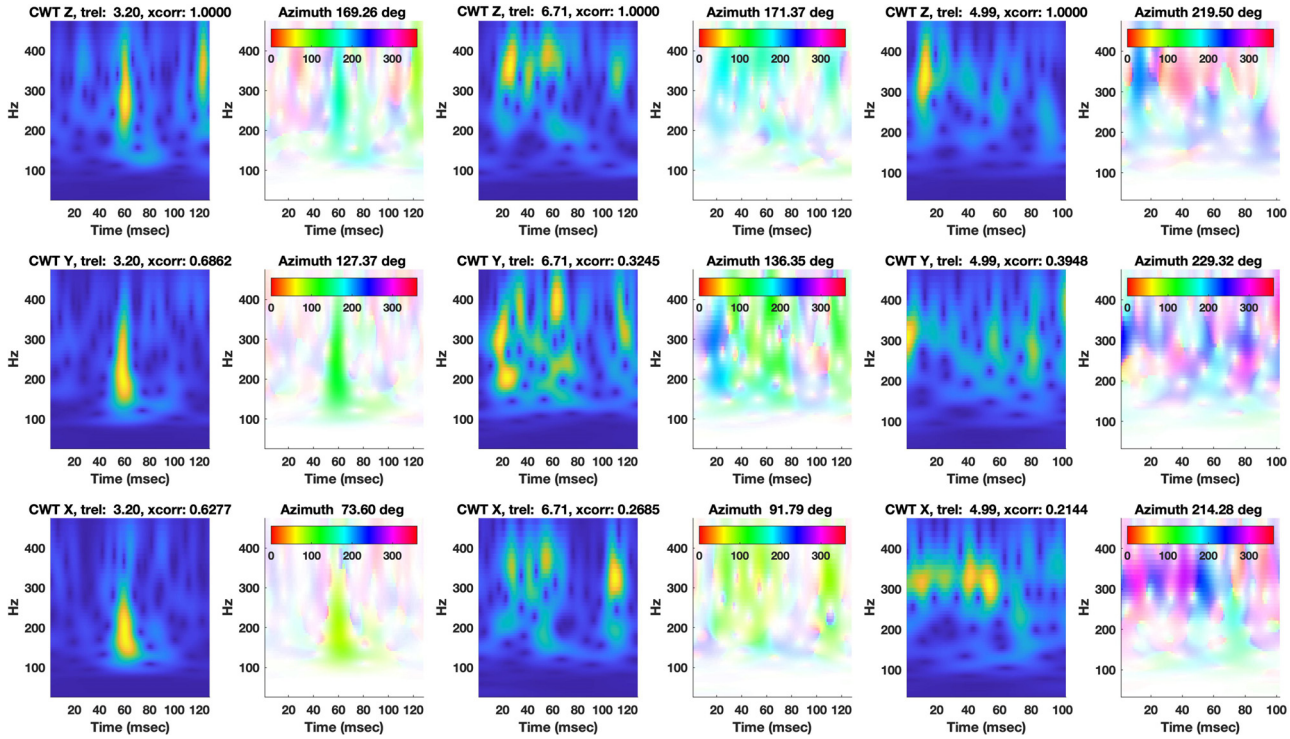

FIG. 12. (Color online) Examples of manual review of three outputs of the directional detector. Each row represents a different DASAR. Columns 1, 3, and 5 are continuous wavelet transforms (CWT) of the detected signal, centered on the peak detection time, with the Pearson's correlation "xcorr" displayed relative to the top sensor. Columns 2, 4, and 6 are corresponding azigrams computed from wavelet transforms of the particle velocity channels, with the opacity of a pixel proportional to the value of the CWT. Columns 1 and 2 display a valid single-pulse localization, columns 3 and 4 display a valid irregular four-pulse localization, and columns 5 and 6 display a false localization.

DASAR at the Block site also seems to have little impact on the pulse count, localization fraction, or spatial distribution, but the third sensor does permit the uncertainty of the positions to be estimated: roughly half of the sounds localized within 50 m have sufficient precision to be localized within the 5 m × 5 m grid blocks shown in the spatial distributions.

The spatial distributions of the pulses shown in Fig. 8 are also unaffected by the sensor order. A little under half of the localizable pulses lie within 50 m of the array site, and thus are presumably within the range resolution of the triangulating array. The vast majority of the localized pulses are concentrated in a 20 × 20 m region on the south side of the ridge, which happens to be the side of the reef facing the incoming current. There is some anecdotal evidence among fishermen that pelagic reef fish prefer to forage on the portion of the reef facing the current (Johannes, 1981; Lieske and Myers, 2002).

In comparison with the Block Site, only 1/3 of the Shark site localizations within 50 m range achieve a 5 m × 5 m precision, and Fig. 9 suggests that the choice of DASAR used for the detector can change the relative balance of sounds localized to the east and west of the site. For example, when DASAR B is used in the detector, Fig. 9(c) shows a higher percentage of sounds localized to the east. Based on boat calibration measurements (shared in the supplemental material), we hypothesize that DASAR B was placed close to a vertical reef wall, resulting in interfering reflections that biased the azimuthal estimates and reduced the localization quality.[2]

Table II displays that using a lower threshold of 50 Hz and a narrower angular sector nearly doubles the number of pulses detected at both sites, while shifting the localization fraction by only around 5% and decreasing the fraction of high-precision localizations by only 5 or 10% points, depending upon the site. Figures 10 and 11 show how the resulting spatial distributions for both parameter sets are similar, suggesting that both parameter sets have similar low false alarm rates, as predicted by the analysis in the Appendix.

Figures 10 and 11 also show that the spatial distributions at both sites change relatively little when samples are limited to only high-precision results (second column). However, the third column of both figures show that the relative fraction of high-precision localizations in a grid cell decreases with increasing distance from the array center, a situation to be expected from an array with a fixed baseline. These results, along with the percentages shown in the final column of Table II, provide a conclusion similar to that from the SCUBA bubble plume result with only six samples: while around half of the localizations have low precision, most of the localizations have low bias.

V. CONCLUSION

An automated method has been derived for detecting impulsive sounds on vector sensors that exploits only the direction, and not the amplitude, of transient sounds,

https://doi.org/10.1121/10.0003382

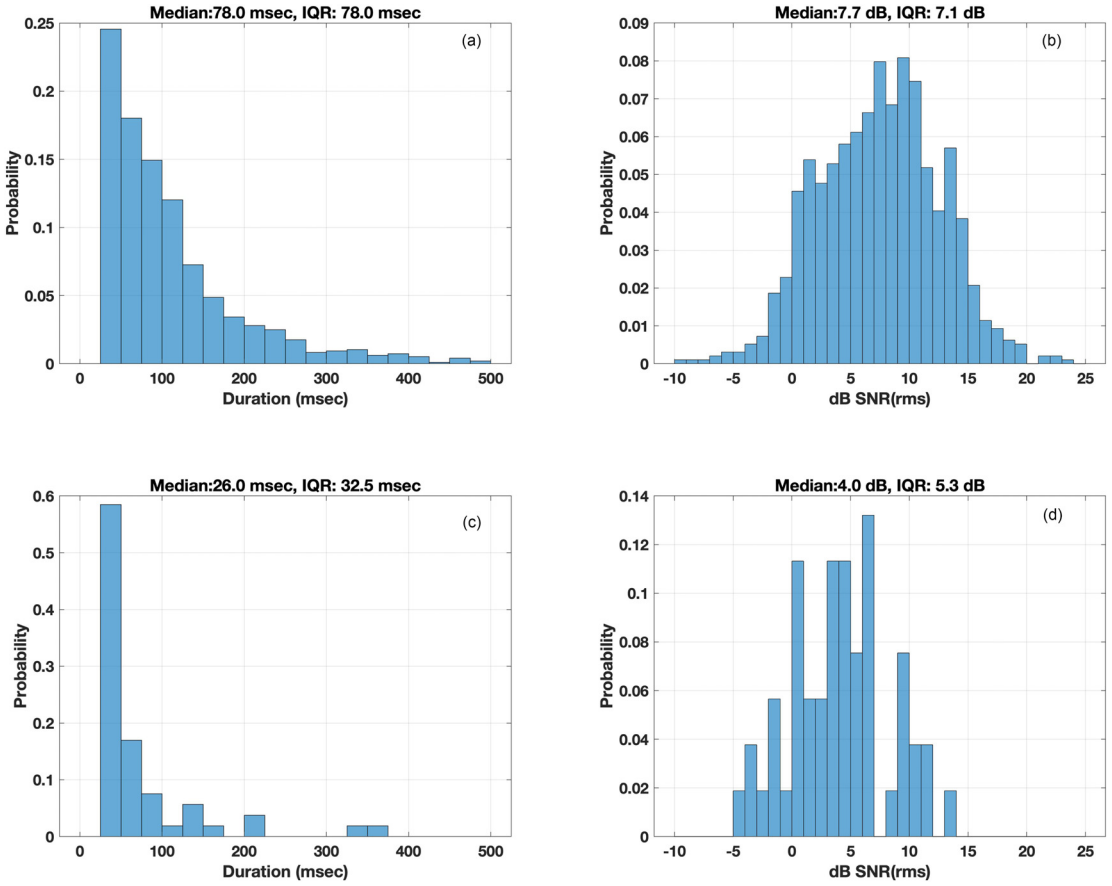

FIG. 13. (Color online) Statistics from manual evaluation of transient detections for ZYX configuration, 90°/120 Hz parameter set, for all three time intervals combined. The top row displays true localization distributions for (a) detection duration in ms, and (b) SNR (rms), while the bottom row shows the false localization distributions for (c) duration and (d) SNR.

making the method a subset of the "localize before detect" paradigm. An analytical analysis of the method (Appendix) shows that it should work even if the baseline diffuse ambient field is moderately directional, provided that appropriate sector widths and bandwidth thresholds are chosen. The present formulation of the algorithm requires that the sensors be placed close enough together to satisfy the conditions of Eq. (7), and that the signals should be impulsive (have a large instantaneous bandwidth). The requirement for a large instantaneous bandwidth is essential for reducing the false detection rate in directional noise backgrounds, but it may be possible that frequency-modulated sweeps and other signals with sufficiently high time-bandwidth product could be tracked with a modified procedure.

Data collected from both SCUBA operations and bioacoustic activity on a coral reef demonstrate that the algorithm works, and that the resulting 2-D spatial distributions of bioacoustic activity are robust to the choice of sensors, parameters used, and resolution criteria. The results indicate that although half of the localizations have high uncertainty, the biases of most estimates are small, and direct manual evaluations of the resulting localizations found only a 3%–6% false localization rate, even when the background sound field was intrinsically directional.

This work may have a variety of bioacoustic applications, including mapping diel and seasonal cycles in bioacoustic activity, identifying the sources and purposes of bioacoustic impulses, and even underwater navigation.

ACKNOWLEDGMENTS

The authors would like to thank Rich Walsh and Christian McDonald of the Scripps Institution of Oceanography (SIO) Scientific Diving Program for SCUBA operational support for this effort. We also enthusiastically thank the 100 Island Challenge team at SIO (Stuart Sandin, Dr. Brian Zgliczynski, and Nicole Pedersen) for allowing us to use their photomosaic equipment, and providing the structure-from-motion reconstruction of the Block2020 site. The data under this paper has been collected as part of the PADRES effort under the DARPA Persistent Aquatic Living Sensors (PALS) program.

APPENDIX

1. Analytical expressions for false detection and localization rates created by diffuse directional background sound field

Transient acoustic pulses are generally embedded within a more diffuse ambient sound field. If this field has an intrinsic directionality of its own, then the algorithm in Sec. II could generate false positions by triangulating on random bearings

generated from noise samples that coincidently yield intersecting azimuths. This Appendix derives false detection and localization rates for the algorithm described in Sec. II, for two sensors with noise samples drawn from two (potentially different) directional distributions. For most practical cases sensors tens of meters apart would be expected to have identical background noise distributions.

Figure 1 illustrates the notation for this analysis, using the compass convention for defining angles (increasing clockwise from the vertical y axis toward the horizontal x axis). The time-frequency ambient sound directionality is modeled as a set of VM distributions, with noise samples in adjacent frequency bins assumed statistically independent. The VM probability density function, an analytic continuous distribution around a circle that is analogous to the Gaussian distribution, is defined as (Mardia and Jupp, 2009)

$$p_i(\varphi_k|\mu_A,\kappa_A) = \frac{e^{\kappa_A \cos(\varphi_k-\mu_A)}}{2\pi I_{0(\kappa_A)}}, \tag{A1}$$

where φ_k is the azimuth measured at sensor A, μ_A is the location parameter, κ_A is the concentration parameter, and $I_o(\kappa)$ is the modified Bessel function of order 0. For large κ_A, the variance becomes $\sigma^2 = 1/\kappa_A$. Whenever κ_A is zero the angular distribution is uniform. In the examples that follow, μ is set to 90°, so that the noise directionality peaks broadside of the sensor array (along the positive x axis in Fig. 1). The sector widths and boundaries are defined such that the peak directionality is centered in the middle of a sector, which places an upper bound on the false detection rate.

The probability that a given time-frequency cell of an azigram for sensor A yields an azimuth from sector k (bounded by angles $\varphi_{k,1}$ and $\varphi_{k,2}$) is given by the cumulative density distribution for Eq. (A1):

$$\Phi(\varphi_{k,1},\varphi_{k,2},\mu_A,\kappa_A) = \Phi_{Ak} = \int_{\varphi_{k,1}}^{\varphi_{k,2}} p(\varphi|\mu_A,\kappa_A)d\varphi. \tag{A2}$$

Thus, the joint probability $p_{AB}(k,l)$ that a single time-frequency bin has a dominant direction from sector k on sensor A and from sector l on sensor B becomes $p_{AB}(k,l) = \Phi_{Ak}\Phi_{Bl}$.

Assuming that simultaneous measurements of the noise field at different frequency bins are statistically independent, the probability that at least N_B $(D_{thresh}/\Delta f)$ out of N_F $(B/\Delta f)$ total frequency bins will satisfy the detection criterion is provided by the cumulative binomial distribution,

$$p_{AB}(false\,detect|k,l) = \sum_{n=N_B}^{N_F} \frac{N_F!}{n!(N_F-n)!} p_{AB}(k,l)^n \times (1-p_{AB}(k,l))^{N_F-n}. \tag{A3}$$

Equation (A3) thus represents the probability of a false detection on two directional sensors monitoring sectors k and l. The mean number of false detections per minute is $p_{AB}(false\,detect|k,l)(60F_s)/N_{fft}(1-ovlap)$, with $ovlap$ being the fractional overlap between FFT samples. As the number of false detections per unit time increases, adjacent detections are more likely to be merged into a single event, thereby also increasing the average false detection duration.

If a broadband false match takes place between sectors k and l, what is the probability that the resulting azimuths will generate a successful localization? Although the localization algorithm provided by Lenth (1981) is a complex maximum-likelihood procedure, the fundamental technique can be approximated by triangulation. Assume sensor B has detected an event at azimuth φ_l. As can be inferred from Fig. 1, the spread of overlapping azimuths on sensor A that would produce a valid triangulation lie betwee $\varphi'_{kl,1}$ and $\varphi'_{kl,2}$, where

$$\begin{aligned} \varphi'_{kl,1} &= \max(\varphi_{k,1},\varphi_l) \\ \varphi'_{kl,2} &= \max(\varphi_{k,2},\varphi_l) \end{aligned}, \tag{A4}$$

when $\sin(\varphi_l) > 0$ and

$$\begin{aligned} \varphi'_{kl,1} &= \min(\varphi_{k,1},\varphi_l) \\ \varphi'_{kl,2} &= \min(\varphi_{k,2},\varphi_l), \end{aligned} \tag{A5}$$

when $\sin(\varphi_l) < 0$. Equations (A4) and (A5) state that a valid triangulation requires that if a ray is launched along azimuth φ_l from sensor B, and is increasing along the x axis (traveling "east"), then φ_k must be greater than φ_l. Conversely, φ_k must be less than φ_l whenever the launched ray is decreasing along the x axis (traveling "west"). The values $\varphi_{k,1}$ and $\varphi_{k,2}$ are the bounds for the portion of sector k on sensor A that produce a legitimate triangulation.

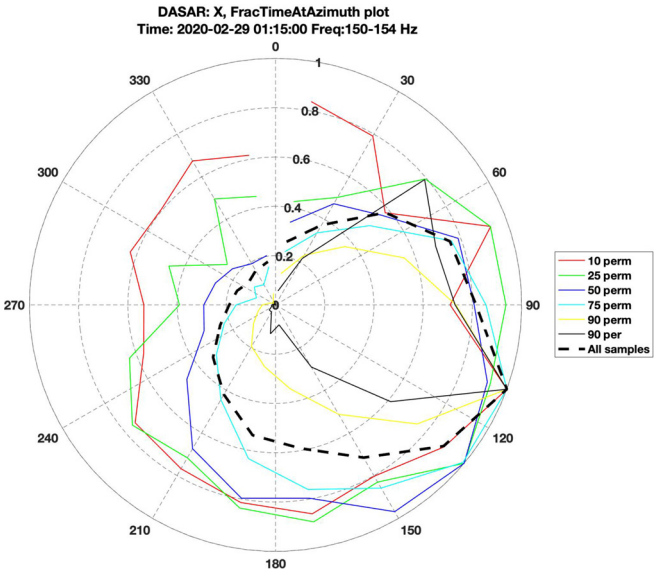

FIG. 14. (Color online) Short-term azimuthal distribution of ambient noise percentiles at DASAR X, Block2020 site, using 2343 samples accumulated over 10 min from single 4 Hz FFT frequency bin (150–154 Hz). For each percentile the direction of maximum directionality has been normalized to one, and the compass convention is used for plotting angles.

https://doi.org/10.1121/10.0003382

The conditional probability $p_A(false\,loc|k,l,\varphi_l)$ that sensor A will generate a localizable result within sector k given an azimuth φ_l on sensor B is $\Phi(\varphi'_{kl,1},\varphi'_{kl,2},\mu_A,\kappa_A)/\Phi(\varphi_{k,1},\varphi_{k,2},\mu_A,\kappa_A)$. Likewise, the conditional probability $p_B(\varphi_l|\mu_B,\kappa_B,l)$ of measuring azimuth φ_l on sensor B, given that it is within sector l, becomes $p_B(\varphi_l|\mu_B,\kappa_B)/\Phi(\varphi_{l,1},\varphi_{l,2},\mu_B,\kappa_B)$. Thus, the conditional probability $p_{AB}(false\,loc|k,l)$ of generating a localization, given random azimuths from sectors k and l becomes

$$p_{AB}(false\,loc|k,l)=\int_{\varphi_{l,1}}^{\varphi_{l,2}} p_A(loc|k,l,\varphi_l)p_B(\varphi_l|\mu_B,\kappa_B,l)d\varphi_l$$
$$=\frac{1}{p_{AB}(k,l)}\int_{\varphi_{l,1}}^{\varphi_{l,2}}\Phi\left(\varphi'_{kl,1},\varphi'_{kl,2},\mu_A,\kappa_A\right)p_B\times(\varphi_l|\mu_B,\kappa_B)d\varphi_l. \quad \text{(A6)}$$

The total probability $p_{AB}(false\,loc,k,l)$ of localizing a false detection at sectors k and l can be determined from Eqs. (A3) and (A6),

$$p_{AB}(false\,loc|k,l)=p_{AB}(false\,loc|k,l)p_{AB}(false\,detect|k,l). \quad \text{(A7)}$$

The expected false localizations per minute from all sector combinations becomes

$$\frac{60F_s}{N_{fft}(1-ovlap)}\sum_{l,k}p_{AB}(k,l)p_{AB}(false\,loc|k,l)$$
$$=\frac{60F_s}{N_{fft}(1-ovlap)}\sum_{l,k}p_{AB}(false\,detect|k,l)$$
$$\times\int_{\varphi_{l,1}}^{\varphi_{l,2}}\Phi\left(\varphi'_{kl,1},\varphi'_{kl,2},\mu_A,\kappa_A\right)p_B(\varphi_l|\mu_B,\kappa_B)d\varphi_l\,. \quad \text{(A8)}$$

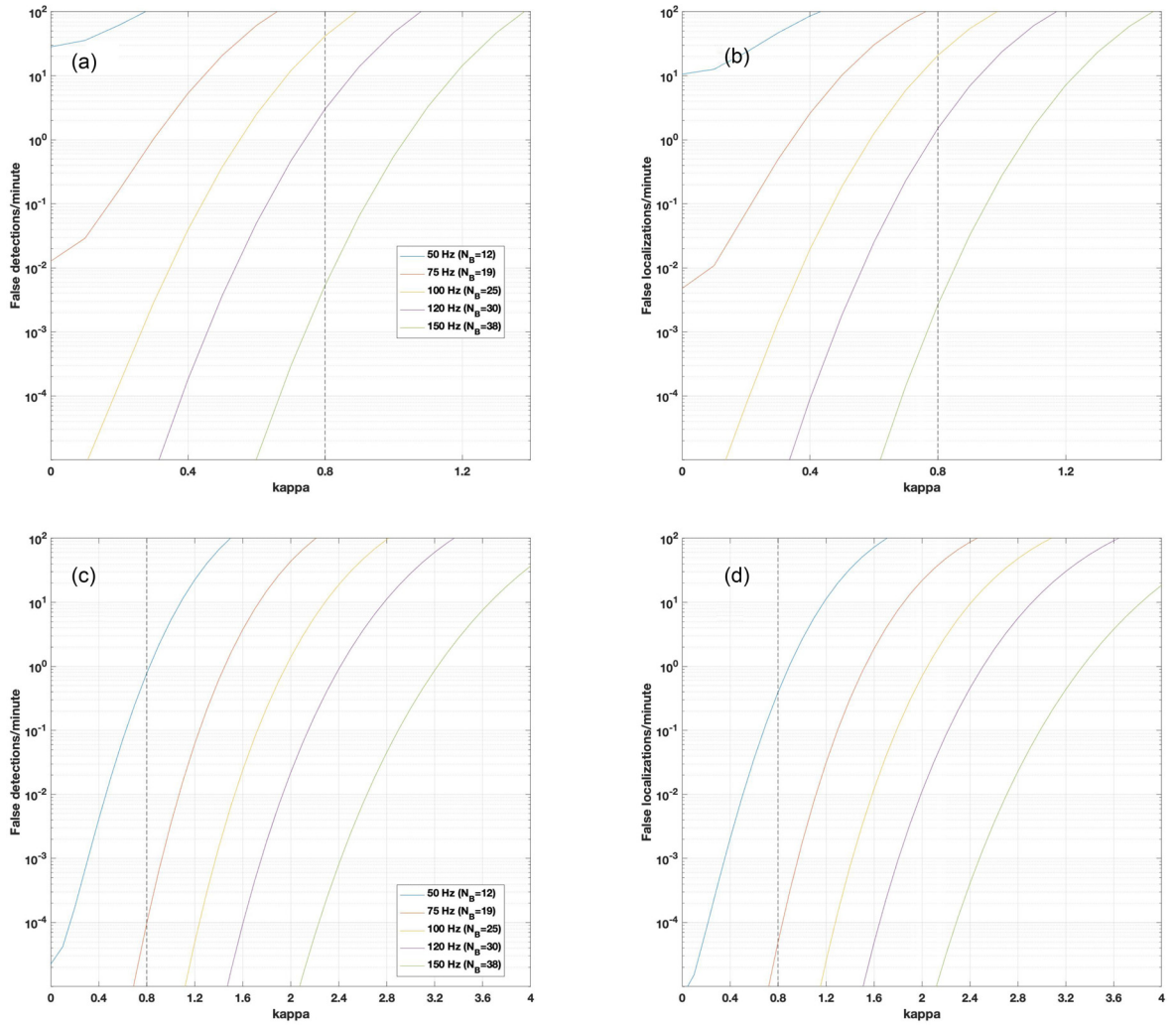

FIG. 15. (Color online) False localization rate analysis of a von Mises ambient noise distribution: $\mu_A=\mu_B=90°$, sector width $\Delta\varphi=90°$, and sector boundaries at 45° and 135°. Each curve represents a different detection threshold D_{thresh}. (a) Probability of a broadband joint detection $p(false\,detect|\,k,l)$ [Eq. (A3)], versus concentration parameter $\kappa=\kappa_A=\kappa_B$, with *NF=103* frequency bins. (b) Mean number of false triangulations per minute [Eq. (A8)] versus κ. Vertical dashed line marks $\kappa=0.8$, which is the estimated fit to the median percentile azimuthal distribution in Fig. 13. Subplots (c) and (d) show similar formats, but for $\Delta\varphi=45°$ and sector boundaries at 67.5° and 112.5°.

The use of a third sensor should reduce the false localization rate even further, but the maximum likelihood algorithm used by the study makes estimating this contribution analytically difficult, because the third azimuth does not need to intersect the other two in order to yield a solution. Note that Eq. (A8) is only the false alarm probability for a single column of an azigram; computing the probability of a longer-duration false detection and localization would require an additional binomial factor.

Establishing a representative value of κ for the coral reef background noise is not straightforward, because the noise distribution is nonstationary except on short timescales ($\sim$1 s) because transient pulses are a dominant component of the Hawaiian coral reef noise. Figure 14 shows the directionality of the ambient noise background for different percentiles (sorted by pressure amplitude), accumulated over 10 min and between 150 and 154 Hz (roughly one FFT frequency bin).

In general, the higher the noise percentile, the more directional the field becomes, since high-SNR transient signals dominate the higher percentiles. At the 50th percentile, which is generally not sampling transient signals, a noise sample is five times more likely to arrive from the southeast than from the northwest (0.2 relative likelihood for a detection to the northwest, relative to the maximum likelihood to the southeast). Equation (A1) shows that the ratio between minimum and maximum probability reaches 0.2 when κ reaches 0.8. The directionality of the 50th percentile seems centered at $\mu = 130°$.

Figures 15(a) and 15(b) show the output of the false broadband detection rate [Eq. (A3)] and localization rate [Eq. (A8)] per minute of operation as a function of $\kappa_A = \kappa_B$, a sector width of 90°, and monitoring bandwidth B equal to that in Table I. The sector boundaries are defined so that the peak of the VM distribution at $\mu = 90°$ is centered within a sector (and thus one sector boundary starts at 45°, another at 135°, etc.) The figure also assumes $\mu_A = \mu_B$, so the ambient sound directionality for both sensors peaks in the direction of the positive x axis, or broadside of the two sensors. For reference, a VM κ value of 0.8 is marked by a dotted line. Figures 15(c) and 15(d) show the same quantities, but using a sector width of 45°, with corresponding sector boundary adjustments in order to keep the sector centered around the VM peak. While not shown here, the false detection and localization rates increase by roughly a factor of 2 if the peak of the VM arrives from endfire, $\mu_A = \mu_B = 0°$.

Both cases show that the false localization rate is dominated by the false detection rate, Eq. (A3). Once two azimuthal regions have been linked by the algorithm, it is relatively likely to be localized on two sensors. As the angular distribution of the noise gets more concentrated around a particular azimuth, the false detection rate increases, and either the bandwidth threshold D_{thresh} must increase or the sector width $\Delta\varphi$ must decrease in order to reduce the rate. The κ value of 0.8, which is representative of the coral reef data, estimates three false localizations per minute (4320 per day) with a 120 Hz threshold when the sector width is 90°; a similar rate is maintained when the sector width is halved to 45° and the detection threshold decreased to 50 Hz. Given that roughly 100 000 detections a day are logged in Hawaii (Table II), this analysis suggests that the parameters in Table I should limit false localizations to 5% or less of the dataset.

This narrower sector width should also permit adequate performance in more directional noise fields. For example, Fig. 15(d) shows that when $\kappa = 2.4$, the original threshold of 120 Hz is sufficient to restrict false localizations to about 1 per minute.

[1]LIDAR data collected from OCM Partners, 2020: 2013 USACE NCMP Topobathy Lidar: Big Island (HI) - LMSL from 2010-06-15 to 2010-08-15. NOAA National Centers for Environmental Information, https://inport.nmfs.noaa.gov/inport/item/49744.

[2]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0003382 for further details on the relative degradation in the calibration quality of DASAR B.

Bertucci, F., Parmentier, E., Lecellier, G., Hawkins, A. D., and Lecchini, D. (**2016**). "Acoustic indices provide information on the status of coral reefs: An example from Moorea Island in the South Pacific," Scientific Reports **6**, 33326.

Blackwell, S. B., Nations, C. S., McDonald, T. L., Thode, A. M., Mathias, D., Kim, K. H., Greene, C. R., Jr., and Macrander, A. M. (**2015**). "The effects of airgun sounds on bowhead whale calling rates: Evidence for two behavioral thresholds," PLoS One **10**, e0125720.

Brown, B. (**1997**). "Coral bleaching: Causes and consequences," Coral Reefs **16**, 129–138..

Cray, B. A., and Nuttall, A. H. (**2001**). "Directivity factors for linear arrays of velocity sensors," J. Acoust. Soc. Am. **110**, 324–331.

D'Spain, G. L., Hodgkiss, W. S., and Edmonds, G. L. (**1991**). "Energetics of the deep ocean's infrasonic sound field," J. Acoust. Soc. Am. **89**, 1134–1158.

D'Spain, G. L., Luby, J. C., Wilson, G. R., and Gramann, R. A. (**2006**). "Vector sensors and vector sensor line arrays: Comments on optimal array gain and detection," J. Acoust. Soc. Am. **120**, 171–185.

Dahl, P. H., and Dall'Osto, D. R. (**2020**). "Estimation of seabed properties and range from vector acoustic observations of underwater ship noise," J. Acoust. Soc. Am. **147**, EL345–EL350.

Deal, T. J. (**2018**). "Vector sensor cross-spectral density for surface noise in a stratified ocean—Formulas for arbitrary sensor geometries," J. Acoust. Soc. Am. **143**, 605–615.

Edwards, C. B., Eynaud, Y., Williams, G. J., Pedersen, N. E., Zgliczynski, B. J., Gleason, A. C., Smith, J. E., and Sandin, S. A. (**2017**). "Large-area imaging reveals biologically driven non-random spatial patterns of corals at a remote reef," Coral Reefs **36**, 1291–1305.

Freeman, S. E., Rohwer, F. L., D'Spain, G. L., Friedlander, A. M., Gregg, A. K., Sandin, S. A., and Buckingham, M. J. (**2014**). "The origins of ambient biological sound from coral reef ecosystems in the Line Islands archipelago," J. Acoust. Soc. Am. **135**: 1775–1788.

Freeman, L. A., and Freeman, S. E. (**1993**). "Rapidly obtained ecological indicators in coral reef soundscapes," Mar Ecol Prog Ser. **501**, 69–82.

Gillespie, D., Mellinger, D., Gordon, J., Mclaren, D., Redmond, P., McHugh, R., Trinder, P., Deng, X., and Thode, A. (**2008**). "PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans," J. Acoust. Soc. Am. **30**, 54–62.

Glynn, P. W. (**1993**). "Coral Reef Bleaching: Ecological," Perspectives - Coral Reefs **12**, 1–17.

Gonzalez, R., and Woods, R. (**2002**). *Digital Image Processing* (Prentice-Hall, Upper Saddle River, NJ).

Gordon, T. A., Radford, A. N., Davidson, I. K., Barnes, K., McCloskey, K., Nedelec, S. L., Meekan, M. G., McCormick, M. I., and Simpson, S. D. (**2019**). "Acoustic enrichment can enhance fish community development on degraded coral reef habitat," Nature Commun. **10**(1), 1–7.

Gordon, T. A. (**2020**). "The changing song of the sea: Soundscapes as indicators and drivers of ecosystem transition on tropical coral reefs," Ph.D. thesis, University of Exeter.
Gracias, N. R., van der Zwaan, S., Bernardino, A., and Santos-Victor, J. (**2003**). "Mosaic-based navigation for autonomous underwater vehicles," IEEE J. Oceanic Eng. **28**, 609–624.
Greene, C. R., McLennan, M. W., Norman, R. G., McDonald, T. L., Jakubczak, R. S., and Richardson, W. J. (**2004**). "Directional frequency and recording (DIFAR) sensors in seafloor recorders to locate calling bowhead whales during their fall migration," J. Acoust. Soc. Am. **116**, 799–813.
Hawkes, M., and Nehorai, A. (**2003**). "Wideband source localization using a distributed acoustic vector-sensor array," IEEE Trans. Signal Process. **51**, 1479–1491.
Holler, R. A. (**2014**). "The evolution of the sonobuoy from World War II to the Cold War," U.S. Navy J. Underwater Acoust. **62**, 322–347.
Johannes, R. E. (**1981**). *Words of the Lagoon: Fishing and Marine Lore in the Palau District of Micronesia* (University of California Press, Los Angeles, CA).
Kennedy, E. V., Holderied, M. W., Mair, J. M., Guzman, H. M., and Simpson, S. D. (**2010**). "Spatial patterns in reef-generated noise relate to habitats and communities: Evidence from a Panamanian case study," J. Exp Mar Biol Ecol. **395**(1), 85–92.
Kleypas, J. A., and Yates, K. K. (**2009**). "Coral reefs and ocean acidification," Oceanography **22**, 4.
Lenth, R. V. (**1981**). "On finding the source of a signal," Technometrics **23**, 149–154.
Lieske, E., and Myers, R. (**2002**). *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem* (Academic Press, New York).
Lilly, J. M., and Olhede, S. C. (**2012**). "Generalized Morse wavelets as a superfamily of analytic wavelets," IEEE Trans. Signal Process. **60**, 6036–6041.
Lirman, D., Gracias, N. R., Gintert, B. E., Gleason, A. C. R., Reid, R. P., Negahdaripour, S., and Kramer, P. (**2007**). "Development and application of a video-mosaic survey technology to document the status of coral reef communities," Environ. Monitor. Assess. **125**, 59–73.
Mardia, K. V., and Jupp, P. E. (**2009**). *Directional Statistics* (John Wiley & Sons, New York).
McWilliam, J. N., McCauley, R. D., Erbe, C., and Parsons, M. J. (**2017**). "Patterns of biophonic periodicity on coral reefs in the Great Barrier Reef," Scientific Reports **7**(1), 1–13.
Mellinger, D. K., and Clark, C. W. (**2000**). "Recognizing transient low-frequency whale sounds by spectrogram correlation," J. Acoust. Soc. Am. **107**, 3518–3529.
Nehorai, A., and Paldi, E. (**1994**). "Vector-sensor array processing for electromagnetic source localization," IEEE Trans. Signal Process. **42**, 376–398.
Nitzberg, R. (**1986**). "Clutter Map CFAR Analysis," IEEE Trans. Aerospace and Electronic Systems **AES-22**, 419- 421.
Pandolfi, J. M., Jackson, J. B. C., Baron, N., Bradbury, R. H., Guzman, H. M., and Hughes, T. P. (**2005**). "Are US coral reefs on the slippery slope to slime?" Science **307**(5716), 1725–1726.
Preskitt, L. V., Vroom, P. S., and Smith, C. M., "A rapid ecological assessment (REA) quantitative survey method for benthic algae using photoquadrats with SCUBA," (**2004**). Pacific Science **58**, 201–209.
Putland, R., Mackiewicz, A., and Mensinger, A. (**2018**). "Localizing individual soniferous fish using passive acoustic monitoring," Ecolog. Inf. **48**, 60–68.
Shi, J., Dosso, S. E., Sun, D., and Liu, Q. (**2019**). "Geoacoustic inversion of the acoustic-pressure vertical phase gradient from a single vector sensor," J. Acoust. Soc. Am. **146**, 3159–3173.
Shipps, J. C., and Deng, K. (**2003**). "A miniature vector sensor for line array applications," in *Proceedings of IEEE Oceans*, September 22–26, San Diego, CA, pp. 2367–2370.
Thode, A., Skinner, J., Scott, P., Roswell, J., Straley, J., and Folkert, K. (**2010**). "Tracking sperm whales with a towed acoustic vector sensor," J. Acoust. Soc. Am. **128**, 2681–2694.
Thode, A. M., Blackwell, S. B., Conrad, A. S., Kim, K. H., Marques, T., Thomas, L., Oedekoven, C. S., Harris, D., and Bröker, K. (**2020**). "Roaring and repetition: How bowhead whales adjust their call density and source level (Lombard effect) in the presence of natural and seismic airgun survey noise," J. Acoust. Soc. Am. **147**, 2061–2080.
Thode, A. M., Kim, K. H., Blackwell, S. B., Greene, C. R., and Macrander, M. A. (**2012**). "Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys," J. Acoust. Soc. Am. **131**, 3726–3747.
Thode, A. M., Kim, K. H., Norman, R. G., Blackwell, S. B., and Greene, C. R., Jr. (**2016**). "Acoustic vector sensor beamforming reduces masking from underwater industrial noise during passive monitoring," J. Acoust. Soc. Am. **139**, EL105–EL111.
Thode, A. M., Sakai, T., Michalec, J., Rankin, S., Soldevilla, M. S., Martin, B., and Kim, K. H. (**2019**). "Displaying bioacoustic directional information from sonobuoys using 'azigrams,'" J. Acoust. Soc. Am. **146**, 95–102.
Torrence, C., and Compo, G. P. (**1998**). "A practical guide to wavelet analysis," Bull. Am. Meteorol. Soc. **79**, 61–78.
Tricas, T. C., and Boyle, K. S. (**2014**). "Acoustic behaviors in Hawaiian coral reef fish communities," Mar. Ecol. Prog. Ser. **511**, 1–16.
Ullman, S. (**1979**). "The interpretation of structure from motion," Proc. R. Soc. Lond. Ser. B. **203**, 405–426.
Vermeij, M. J., Marhaver, K. L., Huijbers, C. M., Nagelkerken, I., and Simpson, S. D. (**2010**). "Coral larvae move toward reef sounds," PloS one **5**(5), e10660.
Wilcock, W. (**2012**). "Tracking fin whales in the northeast Pacific Ocean with a seafloor seismic network," J. Acoust. Soc. Am. **133**, 2503–2503.
Zimmer, W. M. X. (**2011**). *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, Cambridge, UK).