Simultaneous localization of multiple broadband non-impulsive acoustic sources in an ocean waveguide using the array invariant

Citation: Gong, Zheng, Purnima Ratilal, and Nicholas C. Makris. "Simultaneous Localization of Multiple Broadband Non-Impulsive Acoustic Sources in an Ocean Waveguide Using the Array Invariant." The Journal of the Acoustical Society of America 138, no. 5 (November 2015): 2649–2667.

As Published: http://dx.doi.org/10.1121/1.4932547

Publisher: Acoustical Society of America (ASA)

Persistent URL: http://hdl.handle.net/1721.1/107751

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Simultaneous localization of multiple broadband non-impulsive acoustic sources in an ocean waveguide using the array invariant

Zheng GongPurnima RatilalNicholas C. MakrisZHM

Citation: The Journal of the Acoustical Society of America **138**, 2649 (2015); doi: 10.1121/1.4932547
View online: http://dx.doi.org/10.1121/1.4932547
View Table of Contents: http://asa.scitation.org/toc/jas/138/5
Published by the Acoustical Society of America

Simultaneous localization of multiple broadband non-impulsive acoustic sources in an ocean waveguide using the array invariant

Zheng Gong[a)]
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Purnima Ratilal
Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02115, USA

Nicholas C. Makris
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 17 October 2014; revised 15 September 2015; accepted 22 September 2015; published online 2 November 2015)

The *array invariant* method, previously derived for instantaneous range and bearing estimation of a single broadband impulsive source in a horizontally stratified ocean waveguide, can be generalized to simultaneously localize multiple uncorrelated broadband noise sources that are not necessarily impulsive in the time domain by introducing temporal pulse compression and an image processing technique similar to the Radon transform. This can be done by estimating the range and bearing of broadband non-impulsive sources from measured beam-time migration lines of modal arrivals along a horizontal array arising from differences in modal group velocity and modal polar angle for each propagating mode. The generalized array invariant approach is used to estimate the range of a vertical source array and vocalizing humpback whales over wide areas from measurements made by a towed horizontal receiver array during the Gulf of Maine 2006 Experiment. The localization results are shown to have roughly 12% root-mean-squared errors from Global Positioning System measured ground truth positions for controlled source transmissions and less than 10% discrepancy from those obtained independently via moving array triangulation for vocalizing humpbacks, respectively. © *2015 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4932547]

[ZHM] Pages: 2649–2667

I. INTRODUCTION

In free-space with no dispersion, source localization with an array of acoustic sensors is possible in the near field of the array by a variety of methods, including focusing[1] and triangulation.[2,3] In ocean waveguides, modal dispersion enables source range and bearing to be instantaneously determined via the array invariant method[4] from measured beam-time migration lines arising from range and bearing dependent differences between modal group speeds along a horizontal line array even in the far field of a sensor array, if there is at least one propagating mode whose cutoff frequency is sufficiently lower than that of the source.[4,5] The ability to make such far field range estimates in an ocean waveguide is a significant improvement over what is possible in free-space, and has been experimentally demonstrated to enable source range and bearing estimation for source ranges extending to tens of kilometers in continental-shelf waveguides.[4,6]

The array invariant method has been introduced by Lee and Makris[4,7] for instantaneous range and bearing estimation of a single broadband impulsive source in a horizontally stratified ocean waveguide from passive beam-time intensity data received on an array with horizontal or vertical aperture. Søstrand[8] showed that small explosive charges dropped from the endfire direction of a bottom-mounted horizontal array at a 10 to 115 km range can be localized in range with an approach that requires knowledge of the waveguide invariant parameter $\beta(R)$ and multiple propagating modes, using a formulation similar to that of Yang's endfire case [Eq. (10)].[9,10] The array invariant approach, used here, does not require multiple propagating modes, nor knowledge of the range-dependent parameter β, and is not limited to the endfire direction[4–7] or ranges greater than 10 km.

In this paper, it is shown that the array invariant approach can be generalized to instantaneously and simultaneously localize multiple uncorrelated broadband noise sources in horizontal stratified ocean waveguides that are not necessarily impulsive in the time domain by introducing a temporal pulse compression via matched filtering and an image transform technique that is similar to the Radon transform, following the theory derived in Ref. 5. This includes localizing sources at different ranges but on the same beam, and sources distributed widely in azimuth. Here, a source is classified as non-impulsive, when its signal duration T is much larger than the width of correlation peak (e-folding time scale τ_e) of the nonlinear matched filter (NMF) output. Experimental source range estimates obtained from the generalized array invariant method, first with nonlinear matched

a)Electronic mail: zgong@mit.edu

filtering and then without any matched filtering, show that nonlinear matched filtering is necessary to make source localization possible in cases where temporal overlap between modal arrivals cannot be otherwise resolved. This is done by localizing a vertical source array, with known ground truth positions measured by a Global Positioning System (GPS), which transmitted a 0.6- or 1-s broadband linear frequency modulated (LFM) non-impulsive signal with 50 Hz bandwidth centered at various frequencies in the 300 to 1200 Hz range during the Gulf of Maine 2006 Experiment (GOME'06) conducted north of Georges Bank. For all transmissions of this waveform (over 2500), the generalized array invariant method which includes nonlinear matched filtering is empirically shown to enable source localization with roughly 12% root-mean-squared (rms) error in a range-dependent ocean environment with variation of approximately 50% in water depth, at ranges between 1 and 20 km. Without any matched filtering, the estimation error was on average roughly 800% of the true source range, highlighting the need for pulse compression of the kind enabled by nonlinear matched filtering. In addition, for the same source array at ranges between 6 and 20 km with 32 controlled broadband (275 to 760 Hz) LFM non-impulsive source transmissions of 2.4-s signal duration, we also show that the generalized array invariant method with nonlinear matched filtering and without any matched filtering yields on average a roughly 9% and more than 10 000% rms error of the true source range, respectively, again highlighting the need for pulse compression of the kind enabled by nonlinear matched filtering. Apparently, Søstrand's method was not effective at ranges less than 10 km because at these shorter ranges there are too many modes and insufficient time separation for an approach which does not use matched filtering to resolve the modal arrivals. After mode stripping and time separation at long ranges this becomes less of a problem.

Passive acoustic survey methods employing acoustic sensors at fixed locations[11–17] or mobile platforms[12,18,19] have been widely used to detect, localize, and track passive sound sources. In this paper, it is shown that instantaneous localization of multiple vocalizing whales can be made without ambiguity from simultaneous vocalization measurements made by a single towed horizontal receiver array using the generalized array invariant method. A modified image transform technique, similar to that used in Ref. 5 but adapted for whale vocalizations with arbitrary time-dependence and intensity levels, is applied to automatically determine the estimates of source range and bearing from beam-time intensity data obtained after time-domain beamforming and nonlinear matched filtering of acoustic array measurements. Source range and bearing estimates obtained from the generalized array invariant method are shown to be highly consistent with the position estimates independently obtained by the Moving Array Triangulation (MAT) method with less than 10% discrepany.[6]

The generalized array invariant approach is explained in Sec. II and Appendix A. The generalized array invariant method is applied to localize a vertical source array and simultaneously localize multiple vocalizing humpback whales from data acquired near Georges Bank during GOME'06 in Sec. III. The resulting source range and bearing estimates of the vertical source array and vocalizing humpback whales are then compared with the ground truth positions measured by GPS and the independent source localization estimates obtained by the MAT method, respectively.

II. THE GENERALIZED ARRAY INVARIANT METHOD

A. Theory

In free space with no dispersion, array measurements of an impulsive acoustic source in the beam-time domain exhibit one single global maximum when the array is located in the far field of the source. In an ocean waveguide, however, multi-modal dispersion leads to a beam-time migration of modal arrivals along the array arising from differences in modal group velocity and modal polar angle associated with each mode n. This beam-time migration falls onto a single curve, which follows a unique dependence on the source range and bearing and is invariant to the ocean environments, so long as the bandwidth of the source signal is sufficiently large and there is at least one propagating mode whose modal cutoff frequency is sufficiently lower than that of the source. By detecting this curve in the measured beam-time data, source range and bearing can be directly obtained. For simplicity, the theory is first explained in an ideal waveguide, and later is shown to be approximately valid in a horizontally stratified waveguide.

In an ideal waveguide with isovelocity water column sound speed and perfectly reflecting boundaries, an exact solution of the acoustic field can be expressed as an infinite sum of discrete modes via modal decomposition.[20] Each mode n has a deterministic modal polar angle ϕ_n that can be directly calculated using the water column depth D, source frequency f, and mode number n.[20] The geometry of the wavenumber coordinate is shown in Fig. 1. Multi-modal dispersion in an ideal waveguide causes modes to travel at different group velocities v_{gn} due to different modal polar angles. For example, at frequencies significantly above the cutoff frequency of each propagating mode, low order modes, which propagate almost horizontally due to their near $\pi/2$ polar angles, travel faster than high order modes, which propagate more vertically due to their smaller polar angles. With a coherent horizontal line array as a receiver,

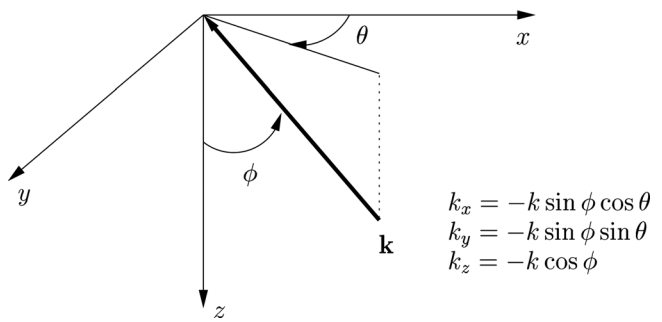

FIG. 1. Definition of the polar angle ϕ and the bearing θ of plane waves. The angles are defined in the "coming from" direction (Refs. 4 and 5). This figure is reproduced from Fig. 2-2 of Ref. 5, with permission.

this modal dispersion leads to a time lag τ in modal arrivals across the receiver array. Decreasing modal polar angle with increasing mode number n also leads to a shift in array scan angle θ from directions at or near source bearing θ_o toward array broadside direction $\theta = 0$ for sources not at array broadside. This time lag and array scan angle shift lead to be a migration of modal arrivals in the beam-time domain along a curve that only depends on the source range and bearing. The migration for each mode n may appear as a single point along this curve for an acoustic source transmitting signals at single frequency. This is because of the discrete nature of the waveguide modes. However, as the bandwidth of the source signal increases, the gaps between these discrete points fill in, because both the group velocity v_{gn} and polar angle ϕ_n of each mode n have changed with frequency, either shifting toward an earlier segment with a greater array scan angle or a later segment with a smaller array scan angle depending on whether the frequency is increased or decreased. If the source bandwidth is sufficiently large, all gaps will be filled and all these discrete line segments merge to a single migration curve for all mode numbers. For a sufficiently impulsive broadband source, this migration in the beam-time domain is resolvable,[4] given a sufficiently high signal-to-noise ratio (SNR), from which source range ρ_o and bearing θ_o can be instantaneously estimated. Even for an extreme case when there is only one propagating mode surviving after long-range propagation, the array invariant approach is still applicable so long as the bandwidth of source signal is sufficiently large and the sufficient frequency contents of the source signal is above the modal cutoff frequency. This leads to a significant change in the group velocity and polar angle for that mode over the frequency band of the source, which makes the migration curve in the beam-time domain resolvable. Discrete migration line $\tilde{s}_n(t)$, which corresponds to the bearing of peak beamformed pressure for the nth mode at time t is specified by[4]

$$\tilde{s}_n(t) = \sin\tilde{\phi}_n \sin\theta_o. \tag{1}$$

An example, shown in Fig. 5 of Ref. 4, shows that the $\tilde{s}_n(t)$ segment for a given mode n migrates to a different part of a single migration curve $\tilde{s}(t)$ when the frequency band of the source signals changes. Once the bandwidth of the source signal is sufficiently large, the gaps between the discrete line segments are filled, and all $\tilde{s}_n(t)$ merge to a single migration curve $\tilde{s}(t)$,[4]

$$\tilde{s}_n(t) \equiv \tilde{s}(t) = \frac{\rho_o}{ct}\sin\theta_o, \tag{2}$$

as shown in Fig. 4 of Ref. 4.

However, this beam-time migration is not necessarily resolvable for a general broadband non-impulsive source with arbitrary time-dependence in an ideal waveguide, because modal arrivals at the receiver array may overlap in time. Besides temporal overlap in modal arrivals, the beam-time migration of modal arrivals for a non-impulsive source is similar to that of an impulsive source described above. To solve this time-overlapping issue, a generalized array invariant method is derived[5] based on the array invariant method for impulsive source localization[4] by introducing a temporal pulse compression, where a detailed derivation appears in Appendix A following Ref. 5. This is done by matched filtering received beam-time pressure data with the source signature either known as *a priori*, measured on a single hydrophone of the array, or measured on a single beam of the array after propagating through the ocean waveguide. After matched filtering, received beam-time pressure data are compressed and the resulting matched filter output forms multiple migration curves in the beam-time domain due to different time lag and polar angle between the mth mode in the received beam-time pressure data and the nth mode in the source signature. These migration curves, however, are bounded within a fan-shaped region, where the bounding curve in the $\tau > 0$ domain is the same as the single migration curve for an impulsive source, from which the source range can be obtained, and the bounding curve in the $\tau < 0$ domain corresponds to a constant array scan angle $\theta = \theta_o$, where the source bearing θ_o can be obtained. The temporal migration of the resulting matched filter peaks in the beam-time domain also falls onto these two bounding curves. Later in this section, an image transform technique similar to the Radon transform will be presented for automatic detection of these two bounding curves in the matched filter output, through which source range and bearing estimates of multiple sources can be simultaneously obtained without ambiguity even for sources located at different ranges but within the same beam.

For a horizontally stratified ocean waveguide, where the water column sound speed $c(z)$ varies with depth z, both modal group velocity $\tilde{v}_{gn}$ and modal polar angle $\tilde{\phi}_n$ for the nth mode vary as a function of $c(z)$ and are also dependent on the frequency derivative of the vertical wavenumber dk_{zn}/df arising from the depth-dependent sound speed structure.[4] If the condition that the modal cutoff frequency for all propagating modes is much lower than that of the source is satisfied,[4,5] change of the vertical wavenumber versus frequency is then typically negligible over the source frequency band. This makes both $\tilde{v}_{gn}$ and $\tilde{\phi}_n$ to be only dependent on the water column sound speed $c(z)$ and become approximately equal to those in an ideal waveguide. Modes that do not satisfy this condition are referred to as waterborne modes, similar to that used by Ref. 9, while modes that satisfy this condition are referred to as non-waterborne modes. Most of the propagating modes in a shallow water waveguide are non-waterborne modes if the dominant frequency contents is above the modal cutoff.[4] As a result, the array invariant method for impulsive broadband source localization[4] and the generalized array invariant method for non-impulsive broadband source localization are then both approximately valid in a horizontally stratified waveguide. A detailed derivation that shows the validity of this approximation in a horizontally stratified waveguide appears in Appendix A following the derivation of Ref. 5.

To demonstrate how source range ρ_o and bearing θ_o of a general broadband non-impulsive source can be directly obtained from the distribution of migration lines in the beam-time domain, an example is given in Fig. 2 for a

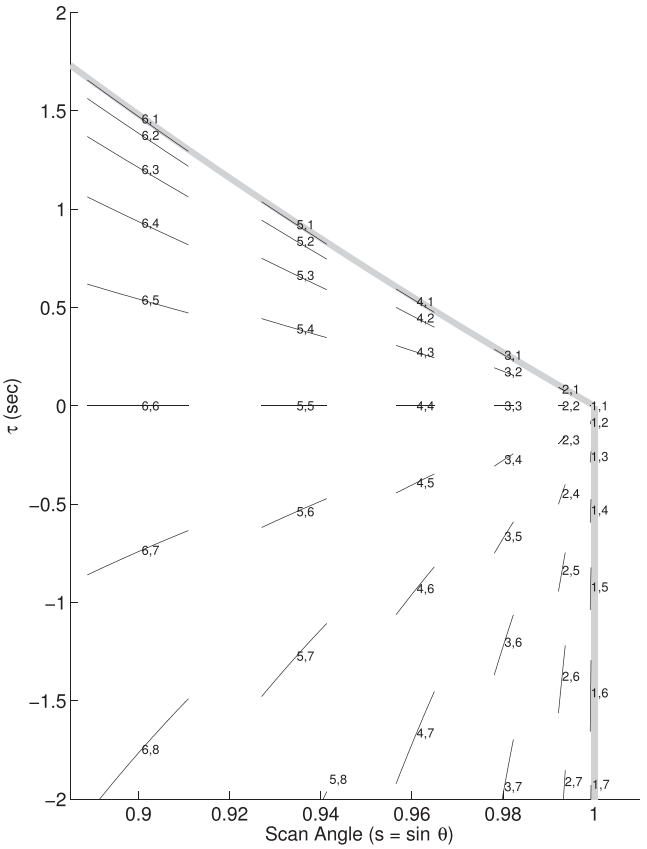

FIG. 2. Following Ref. 5, the migration lines $\tilde{s}_{mn}$ of all the cross-correlated modes m and n for a source located at $\rho_o = 20\,\text{km}$ and $\theta_o = \pi/2$, is shown as a function of the sine of array scan angle $s = \sin\theta$ and reduced travel time τ. The source is assumed to emit a broadband signal in the 90 to 100 Hz frequency range. The black solid lines are the migration lines $\tilde{s}_{mn}$ with corresponding mode numbers marked next to the lines. The slanted thick gray line in the $\tau > 0$ domain is the migration line for a broadband impulsive source defined in Eq. (A16), as a function of the reduced travel time $\tau = t - \rho_o/c$. The thick gray line in the $\tau < 0$ domain corresponds to $\sin\theta_o = 1$. It can be seen that the migration lines $\tilde{s}_{mn}$ are bounded by these two thick gray lines. This figure is reproduced from Fig. 3-2 of Ref. 5, with permission.

source located at $\rho_o = 20\,\text{km}$ and $\theta_o = \pi/2$, which transmits signals in the 90 to 100 Hz frequency band in an ideal waveguide.[5] Each migration line in Fig. 2 corresponds to beam-time migration $\tilde{s}_{mn}$ for a specific pair of mode numbers m and n, which is specified by[5]

$$\tilde{s}_{mn} = \sin\tilde{\phi}_m \sin\theta_o \equiv \sin\phi_m(\tilde{\tau}_{mn}) \sin\theta_o. \tag{3}$$

It is shown that $\tilde{s}_{mn}$ only depends on the polar angle $\tilde{\phi}_m$ of the mth mode in the received beam-time pressure data and the source bearing θ_o. However, this polar angle also depends on the travel time difference $\tilde{\tau}_{mn}$ between the mth mode in the received beam-time pressure data and the nth mode in the source signature. The travel time difference $\tilde{\tau}_{mn}$ is given by[5]

$$\tilde{\tau}_{mn} = \rho_o\left(\frac{1}{\tilde{v}_{gm}} - \frac{1}{\tilde{v}_{gn}}\right) = \frac{\rho_o}{c}\frac{\sin\tilde{\phi}_n - \sin\tilde{\phi}_m}{\sin\tilde{\phi}_m \sin\tilde{\phi}_n}, \tag{4}$$

where c is the water column sound speed, $\tilde{v}_{gm} = c\sin\tilde{\phi}_m$ is the group velocity of the mth mode, and $\tilde{v}_{gn} = c\sin\tilde{\phi}_n$ is the group velocity of the nth mode.

First, in the $\tau > 0$ domain, the thick gray line is the migration curve for an impulsive broadband source in an ideal waveguide as defined in Eq. (2), from which the source range can be directly obtained.[4] It is approximately a straight line for the sine of array scan angle $s = \sin\theta$ near the sine of source bearing $\sin\theta_o$ and for sufficiently small reduced travel time or time lag $\tau = t - \rho_o/c$. It can be seen from Fig. 2 that $\tilde{s}_{mn}$ aligns with this thick gray line when $n = 1$. As mode number m increases, $\tilde{s}_{m1}$ migrate along this slanted thick gray line because of (1) an increasing time lag between the mth mode in the received beam-time pressure data and the first order mode in the source signature and (2) an decreased polar angle $\tilde{\phi}_m$ associated with higher order modes m in the received beam-time pressure data. For fixed mode number m and increasing mode number n, e.g., $n = 1, 2, 3, 4$ and $m = 5$, $\tilde{s}_{mn}$ moves closer to the $\tau = 0$ axis, because the time lag between the mth mode and the nth mode decreases. So, the migration line $\tilde{s}_{m1}$ is the bounding line for all $\tilde{s}_{mn}$ in the $\tau > 0$ domain. Next, in the $\tau < 0$ domain, it can be seen from Fig. 2 that the migration line $\tilde{s}_{mn}$ for $m = 1$ aligns with $s = \sin\theta_o$ indicated by a vertical thick gray line. As mode number n increases, $\tilde{s}_{1n}$ migrates along this thick gray line because of (i) a decreasing time lag between the first order mode in the received beam-time pressure data and the nth mode in the source signature and (ii) a constant and near $\pi/2$ polar angle $\tilde{\phi}_m$ associated with the first order mode in the received beam-time pressure data leading to a constant array scan angle $s \simeq \sin\theta_o$. For fixed mode number n and increasing mode number m, e.g., $m = 1, 2, 3, 4$ and $n = 5$, migration lines $\tilde{s}_{mn}$ move toward the region where $|s| < |\sin\theta_o|$ in the s-domain because of a decreasing polar angle $\tilde{\phi}_m$ for higher order modes m and move toward the $\tau = 0$ axis in the τ-domain because of an increasing travel time difference $\tilde{\tau}_{mn}$ between the mth mode and the nth mode. So, the migration line $\tilde{s}_{1n} \simeq \sin\theta_o$ is the bounding line for all migration lines in the $\tau < 0$ domain. By detecting these two bounding lines, source range and bearing can be directly obtained.

B. Simulation results in a Pekeris waveguide

Here, an example is used to demonstrate that the generalized array invariant method can localize multiple uncorrelated broadband non-impulsive noise sources simultaneously in a typical continental shelf environment and horizontal array configurations via an image transform technique that is similar to the Radon transform.[5]

A case of simultaneously localizing three uncorrelated noise sources S_1, S_2, and S_3 in a 100-m deep Pekeris waveguide with sand bottom is considered.[5] The geometry of three uncorrelated noise sources and a horizontal receiver array is shown in Fig. 3(a).[5] The environment parameters are shown in Fig. 3(b).[5] The source and receiver array depths are $z_o = 50\,\text{m}$ and $z = 30$ m, respectively. The aperture length of the receiver array L is 150 m. The source S_1 is at range $\rho_1 = 8.2\,\text{km}$ and bearing $\theta_1 = 45°$. The other two sources S_2 and S_3 are at the same bearing $\theta_2 = \theta_3 = 60°$, and their ranges are $\rho_2 = 3\,\text{km}$ and $\rho_3 = 10$ km, respectively. An assumption is made that these three sources have the same source power spectral density of 0 dB re 1 $\mu\text{Pa}^2/\text{Hz}$

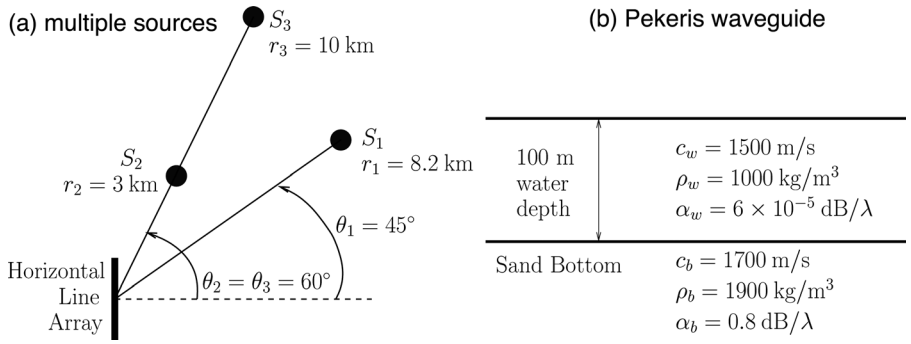

FIG. 3. (a) Top view of the geometry (not to scale) for the multiple sources localization example. Following Ref. 5, the source S_1 is located at 45° from the horizontal receiver array, and at 8.2-km range. The sources S_2 and S_3 are both located at 60° from the horizontal receiver array, and at 3- and 10-km range, respectively. All sources are at 30 m depth, and the receiver array depth is 50 m. The aperture length of the receiver array is 150 m. (b) The Pekeris waveguide with sand bottom, where c_w, ρ_w, and α_w are the sound speed, density, and attenuation of the water column, and c_b, ρ_b, and α_b are those of the sea-bottom. This figure is reproduced from Figs. 3-3 and 3-9 of Ref. 5, with permission.

in the 390 to 440 Hz frequency band, but they are uncorrelated with each other so that $\langle Q_i(f)Q_j^*(f)\rangle = \langle |Q_i|^2(f)\rangle \delta_{ij}$, where $Q_i(f)$ is the source spectrum of S_i, and δ_{ij} is the Kronecker delta function.

The two-dimensional (2D) beam-time intensity data $I_{Bo}(s,\tau)$ for this multiple source scenario is shown in Fig. 4,[5] where $I_{Bo}(s,\tau)$ exhibits two distinct intensity fields when $\tau < 0$, one bounded by a black solid line at $s_1 = \sin\theta_1$, which corresponds to the sine of the source bearing of S_1, and the other bounded by a black solid line at $s_2 = s_3 = \sin\theta_2 = \sin\theta_3$, which corresponds to the sine of the source bearing of S_2 and S_3. When $\tau > 0$, the field for S_2 and S_3 splits into two distinct fields and is bounded by two different migration lines $s_2(\tau)$ and $s_3(\tau)$, since source ranges from the receiver array are different. The field for S_1 in the $\tau > 0$ domain is bounded by another migration line $s_1(\tau)$.

In order to identify these bounding lines from Fig. 4,[5] an image transform technique, similar to the Radon transform,[21] is applied to the 2D beam-time intensity image.[5] The transformed intensity image $\mathcal{I}(s,\phi_r)$ is obtained by integrating

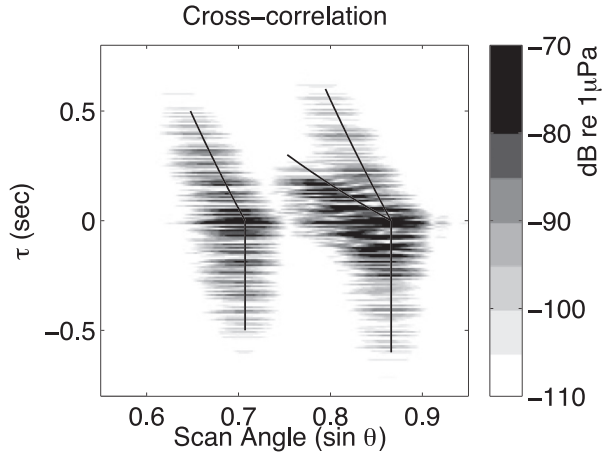

FIG. 4. Following Ref. 5, the 2D beam-time intensity data $I_{Bo}(s,\tau)$ for the source-receiver geometry shown in Fig. 3(a). The two solid black lines overlain in the $\tau < 0$ domain indicate $\sin\theta_1$ and $\sin\theta_2 = \sin\theta_3$, respectively. Three solid black lines in the $\tau > 0$ domain are the migration lines defined in Eq. (A16) for sources S_1, S_2, and S_3, respectively, as a function of the reduced travel time $\tau = t - \rho/c$. This figure is reproduced from Fig. 3-10 of Ref. 5, with permission.

$I_{Bo}(s,\tau)$ of Fig. 4 (Ref. 5) along all semi-infinite straight lines $\tau = (1/\tan\phi_r)(s - s')$ that start from $(s = s', t = 0)$ in the τs-plane via[5]

$$\mathcal{I}(s,\phi_r) = \int_0^\infty |I_{Bo}(s + l\sin\phi_r, l\cos\phi_r)|\, dl, \tag{5}$$

where $0 \le \phi_r < 2\pi$ is an angle measured with respect to the positive τ-axis in the clockwise direction. The normalized transformed image[5]

$$\mathcal{I}_n(s,\phi_r) = \frac{\mathcal{I}(s,\phi_r)}{\max_{s,\phi_r}\mathcal{I}(s,\phi_r)}, \tag{6}$$

for $150° \le \phi_r \le 210°$ is shown in Fig. 5,[5] from which the source bearing estimate $\hat{\theta}_o$ is taken as the array scan angle that corresponds to the maximum of $\mathcal{I}_n(s,\phi_r = 180°)$. In Fig. 5(a), it can be seen that $\mathcal{I}_n(s,\phi_r)$ exhibits two peaks along the horizontal dashed black line $\phi_r = 180°$ due to two distinct source bearings. The transect of $\mathcal{I}_n(s,\phi_r)$ along the line $\phi_r = 180°$ is plotted in Fig. 5(b), from which the three source bearing estimates[5]

$$\hat{\theta}_i = \sin^{-1}\hat{s}_i \equiv \sin^{-1}\left[\arg\max_s \mathcal{I}_n(s,\phi_r = 180°)\right] \quad \text{for } i = 1,\ 2,\ 3 \tag{7}$$

are determined to be $\hat{\theta}_1 = 44.5°$ or $\hat{s}_1 = 0.701$ and $\hat{\theta}_2 = \hat{\theta}_3 = 58.8°$ or $\hat{s}_2 = \hat{s}_3 = 0.855$, which are within 2% of the true source bearings $\theta_1 = 45°$ and $\theta_2 = \theta_3 = 60°$. The two sources S_2 and S_3 cannot be distinguished in the transect along $\phi_r = 180°$, since their bearings are identical.

Next, given the source bearing estimate $\hat{s}_i$ for each random noise source, the corresponding source range estimate $\hat{\rho}_i$ can be obtained from $\mathcal{I}_n(\hat{s}_i,\phi_r)$ using[5]

$$\hat{\rho}_i = -\frac{c(z)\sin\hat{\theta}_i}{\hat{\chi}_{l_i}} = -\frac{c(z)\sin\hat{\theta}_i}{\tan\hat{\phi}_r}, \tag{8}$$

where $\hat{\phi}_r = \arg\max_{\phi r}\mathcal{I}_n(s = \hat{s}_i, \phi_{r1} < \phi_r < \phi_{r2})$. Given the fact that the 2D beam-time intensity data measured by a horizontal line array in the $\tau > 0$ domain migrate from

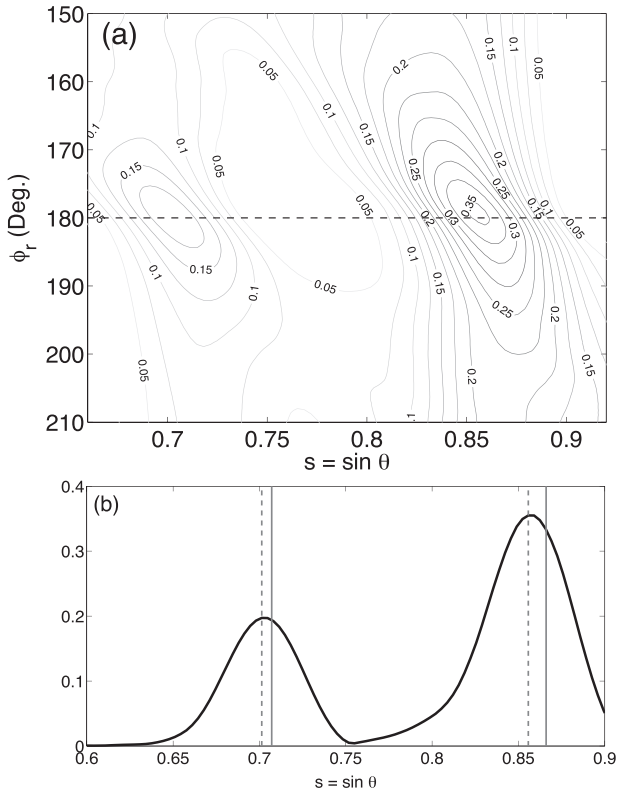

FIG. 5. (a) Following Ref. 5, the normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,t)$ shown in Fig. 4 near $\phi_r = 180^\circ$ is presented for $150^\circ \le \phi_r \le 210^\circ$. Two peaks exist along the $\phi_r = 180^\circ$ radial, indicated by the dashed black line, due to two different source bearings. (b) Following Ref. 5, transect of $\mathcal{I}_n(s,\phi_r)$ shown in (a) along the dashed black line at $\phi_r = 180^\circ$ is plotted. The two peaks in this transect, indicated by the dashed gray lines, correspond to the source bearing estimates $\hat{s}_1 = 0.701$ or $\hat{\theta}_1 = 44.5^\circ$ and $\hat{s}_2 = \hat{s}_3 = 0.855$ or $\hat{\theta}_2 = \hat{\theta}_3 = 58.8^\circ$. The dashed gray lines indicate the true source bearings $s_1 = \sin 45^\circ$ and $s_2 = s_3 = \sin 60^\circ$. This figure is reproduced from Figs. 3-11 and 3-12 of Ref. 5, with permission.

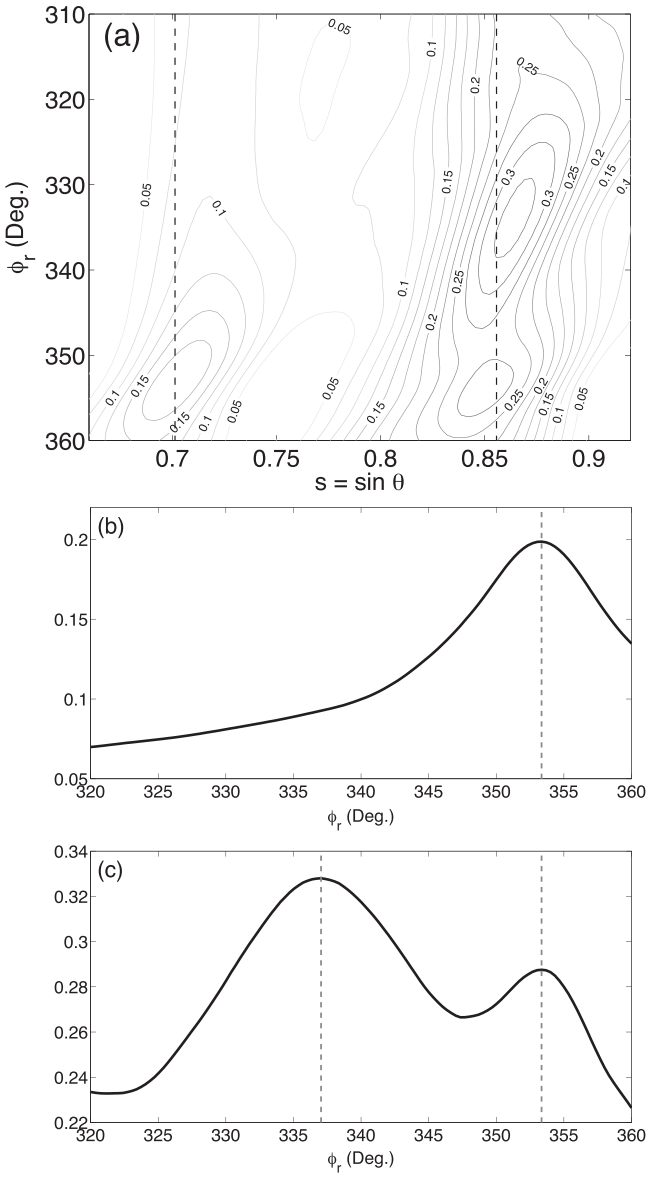

FIG. 6. (a) Following Ref. 5, the normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,\tau)$ shown in Fig. 4 for $310^\circ \le \phi_r \le 360^\circ$ is shown. Two dashed black lines are at $\hat{s}_1$ and $\hat{s}_2 = \hat{s}_3$, respectively. A peak along $\hat{s}_1 = 0.701$ and two peaks along $\hat{s}_2 = \hat{s}_3 = 0.855$ correspond to S_1, S_2, and S_3, respectively. (b) Following Ref. 5, transect of $\mathcal{I}_n(s,\phi_r)$ along $\hat{s}_1 = 0.701$ indicated by the dashed gray line in (a) is plotted as a function of ϕ_r for $320^\circ \le \phi_r \le 360^\circ$. The peak of this transect is at $\phi_r = 353.3^\circ$, which corresponds to the range estimate of $\hat{\rho}_1 = 9$ km. (c) Following Ref. 5, transect of $\mathcal{I}_n(s,\phi_r)$ along $\hat{s}_2 = \hat{s}_3 = 0.855$ indicated by the dashed gray line in (a) is plotted as a function of ϕ_r for $320^\circ \le \phi_r \le 360^\circ$. The first peak at $\phi_r = 337^\circ$ corresponds to S_2, which has a range estimate of $\hat{\rho}_2 = 3$ km. The second peak at $\phi_r = 353.3^\circ$ corresponds to S_3, which has a range estimate of $\hat{\rho}_3 = 11$ km. This figure is reproduced from Figs. 3-13 to 3-15 of Ref. 5, with permission.

$s = \sin\theta_o$ toward near the array broadside direction ($s = 0$), for $\hat{s}_i > 0$, $\phi_{r1} = 270^\circ$, and $\phi_{r2} = 360^\circ$ are used for image transform, while for $\hat{s}_i < 0$, $\phi_{r1} = 0^\circ$, and $\phi_{r2} = 90^\circ$ are used for image transform. For this example with $\hat{s}_1 = 0.701 > 0$ and $\hat{s}_2 = \hat{s}_3 = 0.855 > 0$, the transformed image $\mathcal{I}_n(s,\phi_r)$ for $310^\circ \le \phi_r \le 360^\circ$ is shown in Fig. 6(a),[5] where three distinct peaks in $\mathcal{I}_n(s,\phi_r)$ correspond to three sources in the waveguide.

A transect of Fig. 6(a) along $\hat{s}_1 = \sin\hat{\theta}_1$ is shown in Fig. 6(b).[5] Given the angle ϕ_r corresponding to the peak of $\mathcal{I}_n(\hat{s}_1,\phi_r)$, the source range estimate is obtained to be $\rho_1 = 9$ km using Eq. (8), which is within 10% of the true source range $\rho_1 = 8.2$ km. The ranges of the other two sources, S_2 and S_3, can be estimated from $\mathcal{I}_n(\hat{s} = \hat{s}_2,\phi_r)$ shown in Fig. 6(c), where it can be seen that there are two distinct peaks due to S_2 and S_3 located at different ϕ_r. The estimated source range of S_2 is $\hat{\rho}_2 = 3$ km, which is the same as the true source range of S_2. The estimated source range of S_3 is $\hat{\rho}_3 = 11$ km, which is also within 10% of the true source range $\rho_3 = 10$ km. Sound speed $c(z) = 1500$ m/s is used here to calculate all source ranges.

III. EXPERIMENTAL DEMONSTRATION OF THE GENERALIZED ARRAY INVARIANT METHOD FOR LOCALIZING A VERTICAL SOURCE ARRAY AND SIMULTANEOUSLY LOCALIZING MULTIPLE VOCALIZING HUMPBACK WHALES

Here, we experimentally demonstrate the performance of the generalized array invariant method by (1) localizing a vertical source array with known GPS measured ground truth positions deployed north of Georges Bank at ranges between

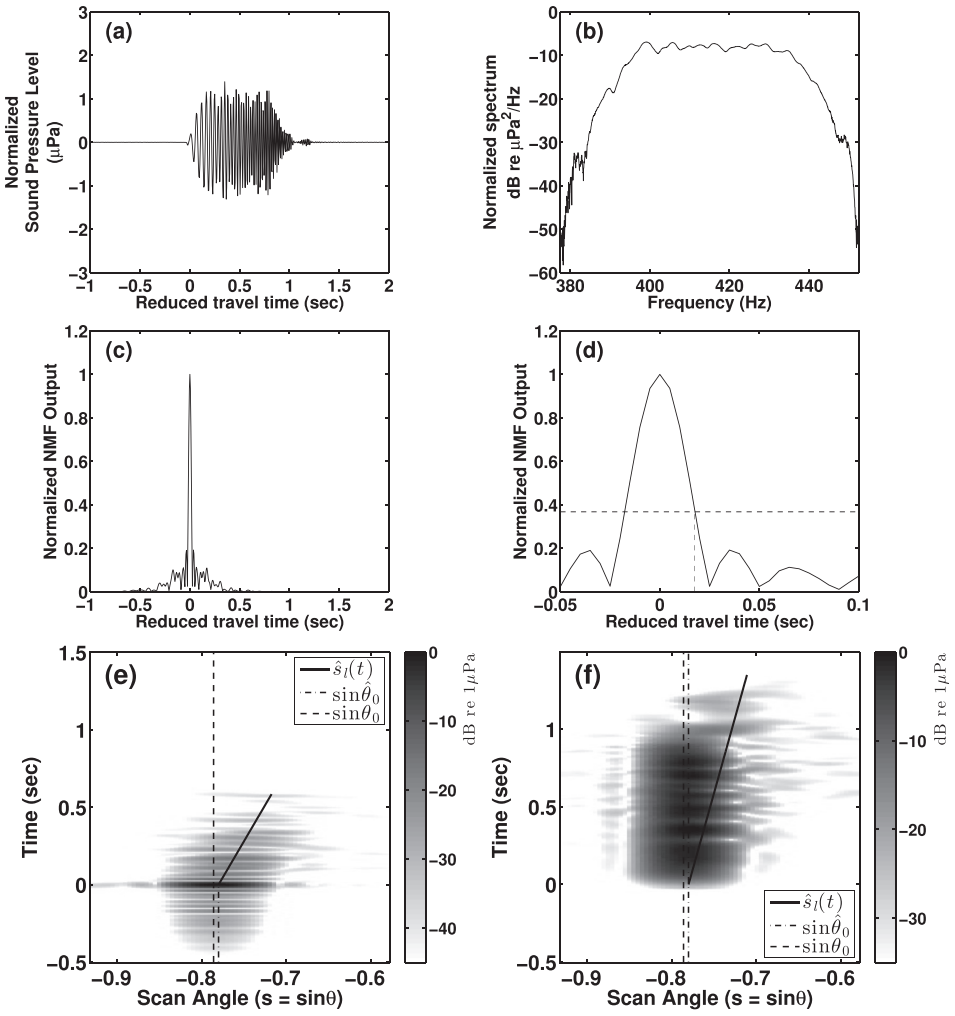

FIG. 7. Source localization with nonlinear matched filtering and without any matched filter for 1-s broadband LFM (50 Hz bandwidth) non-impulsive signal transmission centered at 415 Hz from $\rho_0 = 11.1$ km and bearing $\theta_0 = -51.8°$ from the receiver array broadside direction. (a) Normalized received SPL of the 1-s broadband LFM non-impulsive signal from a single beam of the horizontal receiver array measurements. (b) Normalized spectrum of the received signal shown in (a). (c) NMF output by cross-correlating a NMF replica, generated using the NMF kernel described in Appendix B of Ref. 6, with the measured signal shown in (a). (d) Same as (c) but zoomed around the correlation peak of the NMF output. Dashed lines indicate the e-folding time scale of the NMF output, which is roughly 0.017 s for this example and is over an order of magnitude smaller than the signal duration $T = 1$ s. (e) Beam-time intensity data obtained after nonlinear matched filtering. The dotted-dashed vertical line in $\tau < 0$ domain, determined by the image transform technique, is at $s = \sin\hat{\theta}_0$, and the dashed vertical line in $\tau < 0$ domain, obtained from GPS, is at $s = \sin\theta_0$, where $\hat{\theta}_0 = -51.3°$ and $\theta_0 = -51.8°$, respectively. The solid black line in the $\tau > 0$ domain is the migration line defined in Eq. (A16) and is also determined by the image transform technique. The range estimate is $\hat{\rho}_0 = 10.8$ km, which has a 2.8% error from the GPS measured source range. (f) Beam-time intensity data without any matched filtering. The range estimate is $\hat{\rho}_0 = 23.2$ km, which has an error 108.8% of the GPS measured source range. This highlights the need for pulse compression of the kind enabled by nonlinear matched filtering for non-impulsive signals whose modal arrivals may not be otherwise resolvable.

1 and 20 km from a towed horizontal receiver line array during GOME'06, and (2) simultaneously localizing multiple vocalizing humpback whales from passive array measurements by the same horizontal receiver array, at the same experimental site, and during the same time period.[22–24]

A. Experimental setup

During GOME'06, a vertical source array and a towed horizontal receiver array were separately deployed from two research vessels along the northern flank of Georges Bank over a 2 week period.[22–24] The vertical source array, centered at 60 to 70 m depth, was either moored or drifting with the current and located in the *far-field* of the horizontal receiver array with source-receiver separations ranging between 1 and 20 km. The source array transmitted over 2500, 0.6-s or 1-s broadband (50 Hz bandwidth), LFM non-impulsive signals centered at various frequencies in the 300 and 1200 Hz range. An example of a received source transmission of 1-s broadband LFM non-impulsive signal centered at 415 Hz on the horizontal receiver array is shown in Fig. 7. The same source array also transmitted a complex 2.4-s broadband non-impulsive LFM signal every 100-s, in the 275 to 760 Hz frequency band, as shown in Fig. 10, 32 times. Humpback whale (*Megaptera novaeangliae*) songs and non-song vocalizations from multiple humpback individuals, distributed widely over Georges Bank, were also continuously recorded by the same horizontal receiver array. These vocally active humpback whales were typically located tens of kilometers from the receiver array. At many instances, multiple humpback whale vocalizations were received at the receiver array roughly at the same time.

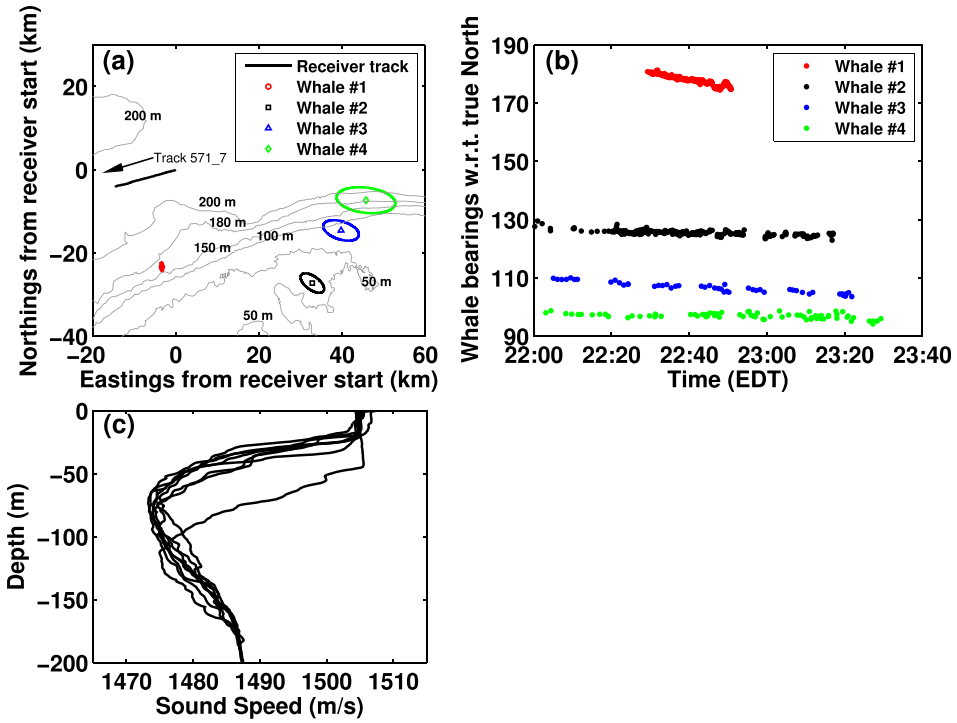

FIG. 8. (Color online) (a) The mean positions and corresponding position estimation uncertainties (ellipses) of four vocalizing humpback whales estimated using the MAT method are overlain with the receiver ship track line on October 3, 2006 between 22:00:00 EDT and 23:30:00 EDT. The depth contours of the sea-bottom in meters are also overlain. The arrow shows the heading of the receiver ship along track 571_7. The origin of the coordinates in this figure is at 42.274°N and 67.662°W. (b) Four sequences of vocalizing humpback whale bearings measured on October 3, 2006 between 22:00:00 EDT and 23:30:00 EDT. All vocalizing humpback whale bearings are measured from the true North in a clockwise direction with respect to the instantaneous spatial locations of towed horizontal receiver array center. (c) Profiles of water column sound speed from XBT measurements deployed on October 3, 2006 during GOME'06.

The horizontal receiver array was towed at a depth of 105 m and at a speed of roughly 4 knots along designated tracks just north of Georges Bank,[22–24] where the positions of the receiver array center were accurately measured by GPS. The multiple nested sub-apertures of the array contain a total of 160 hydrophones spanning a frequency range from below 50 to 3750 Hz for spatially unaliased sensing. A fixed sampling frequency of 8000 Hz (Ref. 22) was used so that acoustic signals with frequency contents up to 4000 Hz were recorded without temporal aliasing. Three linear apertures of the array, a low-frequency (LF) aperture, a mid-frequency (MF) aperture, and a high-frequency (HF) aperture, all of which consist of 64 equally spaced hydrophones with respective inter-element spacing of 1.5, 0.75, and 0.375 m, were used to analyze controlled broadband LFM source transmissions in the 300 to 1200 Hz frequency range and humpback whale vocalizations with fundamental frequency content below 1000 Hz. For controlled broadband source transmissions centered below 500 Hz and humpback whale vocalizations with frequency content below 500 Hz, the LF aperture was used, while for controlled broadband source transmissions centered at frequencies in the 500 Hz to 1 kHz range and humpback whale vocalizations with frequency content in the 500 Hz and 1 kHz range, the MF aperture was used, and for controlled broadband source transmissions centered at above 1 kHz, the HF aperture was used. For the complex 2.4-s broadband non-impulsive LFM signal with the 275 to 760 Hz frequency band, the LF and MF apertures were used. The angular resolution $\beta(\phi, f_c)$ of the horizontal receiver line array is $\beta(\phi, f_c) \approx 1.44(\lambda/L\cos\phi)$ for broadside ($\phi = 0$) through angles near endfire ($\phi = \pi/2$), where $\lambda = c/f_c$ is the acoustic wavelength, c is the sound speed, f_c is the center frequency, and L is the array aperture length. At endfire, the angular resolution is $\beta(\phi = \pi/2, f_c) \approx 2.8\sqrt{\lambda/L}$. The water column depth varied significantly between 50 and 200 m at the experimental site including areas on Georges Bank where many vocalizing humpback whales were found, as shown in Fig. 8(a). The seafloor at the GOME'06 experimental site is mostly sandy and the measured water-column sound speed profiles by XBTs and CTD sensors are shown in Fig. 3 of Ref. 22. The mean sound speed of 1477 m/s at receiver array depth was used for the range estimation of vertical source array and vocalizing humpback whales.

B. Experimental demonstration of the generalized array invariant method for localizing a vertical source array during GOME'06

Here, we experimentally explore the capability of nonlinear matched filtering to improve range estimation. This is done by localizing a vertical source array, with known ground truth positions measured by GPS, which transmitted a 0.6- or 1-s broadband LFM non-impulsive signal with 50 Hz bandwidth centered at various frequencies in the range between 300 and 1200 Hz every 50- or 75-s. The corresponding signal and its spectrum received by the horizontal towed receiver array after beamforming and propagating

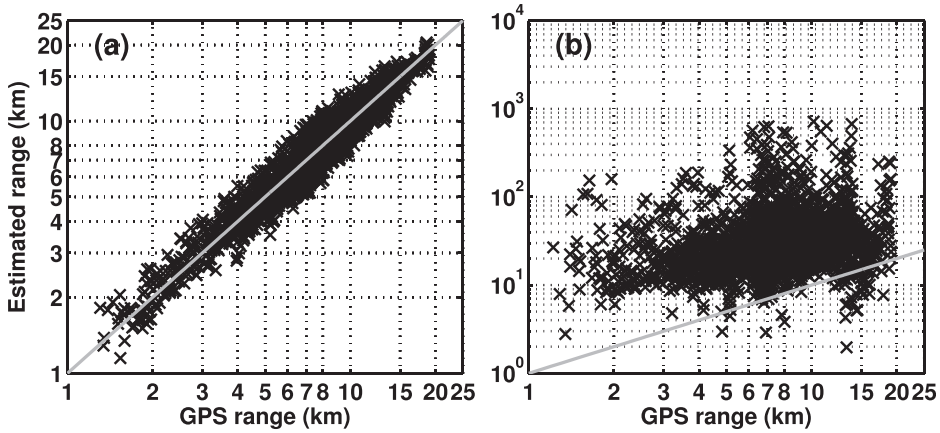

FIG. 9. A total of 2547 controlled source transmissions of 0.6- or 1-s broadband LFM non-impulsive signals with 50 Hz bandwidth are used to quantify the rms error of source range estimation by the generalized array invariant method with nonlinear matched filtering and without any matched filtering. (a) With nonlinear matched filtering, the source range estimates $\hat{\rho}_0$ of the vertical source array are plotted in logarithmic scale against GPS measured source range ρ_0 for the source range between 1 and 20 km and the source bearing between 0° and 80° from the receiver array endfire direction. The rms error of all range estimates is roughly 12% of the GPS measured source range. The solid gray line indicates the true source range measured by GPS. (b) Without any matched filtering, the error on average is roughly 800% of the GPS measured source range. Here, the solid black line also indicates the true source range measured by GPS. By comparing the range estimation results shown in (a) and (b), it highlights the need for pulse compression of the kind enabled by nonlinear matched filtering to improve the source localization accuracy for non-impulsive signals whose modal arrivals may not be otherwise resolvable.

11.1 km in a range-dependent ocean waveguide are shown in Figs. 7(a) and 7(b), respectively. This measured signal is then used to generate a replica for temporal pulse compression using a NMF kernel, as described in Ref. 6. The NMF output for the beam in the direction of the source bearing is shown in Fig. 7(c). The dashed lines in Fig. 7(d) indicate the e-folding time scale τ_e of the NMF output, which in this case is roughly 0.017 s. This τ_e is over an order of magnitude smaller than the signal duration $T = 1$ s and so is consistent with the definition of non-impulsive. For each source transmission, a NMF replica is reconstructed from the measured signal on the horizontal receiver array after beamforming. This replica is then matched filtered with each beam of the beam-time data, and the resulting output is shown in Fig. 7(e). Source range estimates obtained from the generalized array invariant method, first with nonlinear matched filtering [Fig. 7(e)] and then without any matched filtering [Fig. 7(f)], show that nonlinear matched filtering is necessary to make source localization possible in cases where temporal overlap between modal arrivals cannot be otherwise resolved. For all 2547 transmissions (0° to 80° from the receiver array endfire) of this 0.6- or 1-s broadband LFM non-impulsive signal, at ranges between 1 and 20 km, the generalized array invariant method, which includes nonlinear matched filtering, is empirically shown to enable source localization with a 12.0% rms error of the true source range [Fig. 9(a)] in a range-dependent ocean environment with variation of approximately 50% in water depth. Without any matched filtering, in contrast, the estimation error is on average roughly 800% of the true source range [Fig. 9(b)], highlighting the need for pulse compression of the kind enabled by nonlinear matched filtering. In theory, as the source range increases, the modal arrivals of non-impulsive source become more separated. This will lead to improved source range estimation by the array invariant method without any matched filtering, which is also experimentally demonstrated in Fig. 9(b). Source range estimates by the generalized array invariant method with nonlinear matched filtering have been shown to have an approximately 50% lower rms error than those with standard linear matched filtering.[6]

The accuracy of the generalized array invariant method with nonlinear matched filtering and without any matched filtering is also quantified by localizing the same vertical source array with 32 controlled broadband nonlinear LFM source transmissions (275 to 760 Hz) of 2.4-s signal duration. The corresponding signal and its spectrum received by the horizontal towed receiver array after beamforming and propagating 19.1 km in a range-dependent ocean waveguide are shown in Figs. 10(a) and 10(b), respectively. The NMF output for the beam in the direction of the source bearing is shown in Fig. 10(c), and the e-folding time scale $\tau_e \approx 0.14$ s of the NMF output is shown in Fig. 10(d), which is also orders of magnitude smaller than the signal duration $T = 2.4$ s and so is consistent with the definition of non-impulsive. Source range estimates obtained from the generalized array invariant method, first with nonlinear matched filtering [Fig. 10(e)] and then without any matched filtering [Fig. 10(f)], show that nonlinear matched filtering is necessary to make source localization possible. For all 32 transmissions (30° to 80° from the receiver array endfire) of this 2.4-s broadband non-impulsive LFM signal, at ranges between 6 and 20 km, the generalized array invariant method, which includes nonlinear matched filtering is empirically shown to enable source localization with a 9.0% rms error of the true source range [Fig. 11(a)] in a range-dependent ocean environment with variation of approximately 50% in water depth. Without any matched filtering, in contrast, the estimation error on average is over 10 000% of the true source range [Fig. 11(b)], again highlighting the need for pulse compression of the kind enabled by nonlinear matched filtering.

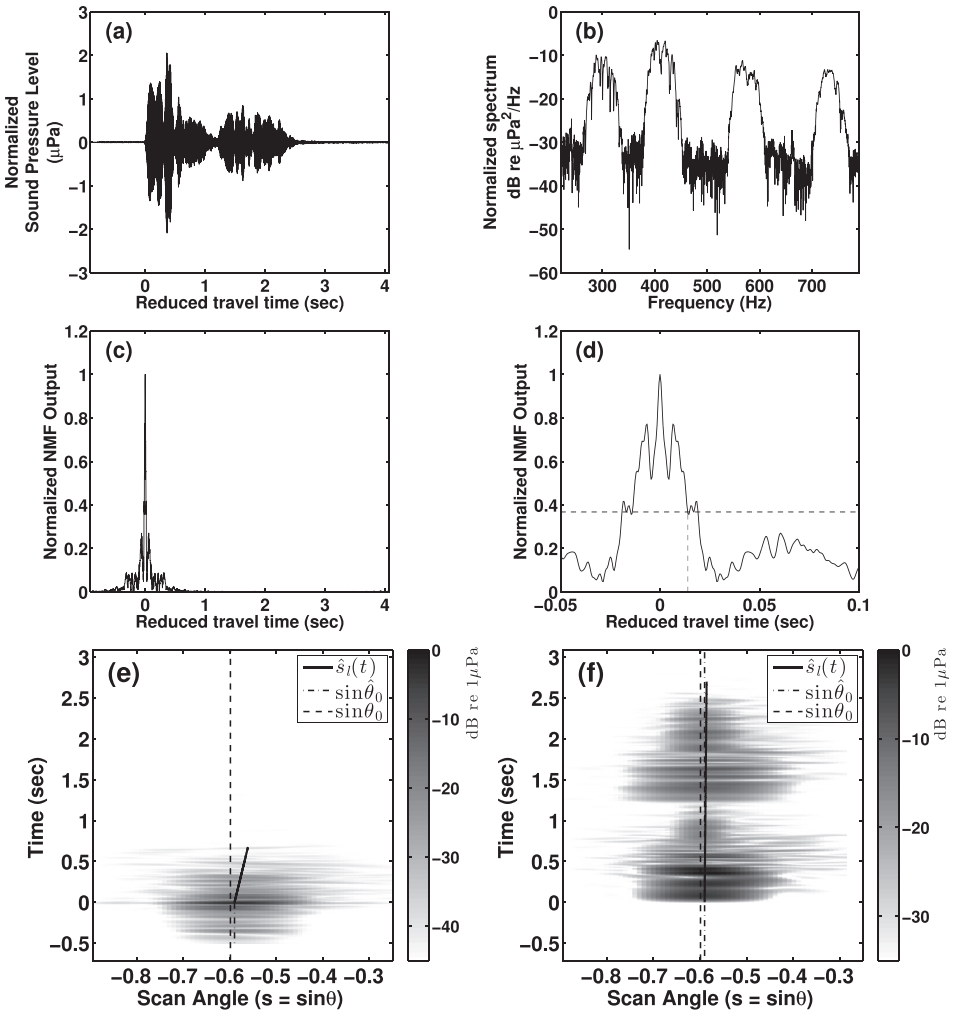

FIG. 10. Source localization with nonlinear matched filtering and without any matched filter for 2.4-s broadband non-impulsive LFM (275 to 760 Hz) signal transmission from $\rho_0 = 19.1$ km and bearing $\theta_0 = -36.9^\circ$ from the receiver array broadside direction. (a) Normalized received SPL from a single beam of the horizontal receiver array measurements. (b) Normalized spectrum of the received signal shown in (a). (c) NMF output. (d) Same as (c) but zoomed around the correlation peak of the NMF output. Dashed lines indicate the *e*-folding time scale of the NMF output, which is roughly 0.014 s for this example and is orders of magnitude smaller than the signal duration $T = 2.4$ s. (e) Beam-time intensity data obtained after nonlinear matched filtering. This is done by cross-correlating the NMF replica with each beam of the beam-time data. The dotted-dashed vertical line in the $\tau < 0$ domain, determined by the image transform technique, is at $s = \sin\hat{\theta}_0$, and the dashed vertical line in the $\tau < 0$ domain, obtained from GPS, is at $s = \sin\theta_0$, where $\hat{\theta}_0 = -36.2^\circ$ and $\theta_0 = -36.9^\circ$, respectively. The solid black line in the $\tau > 0$ domain is the migration line defined in Eq. (A16) and is also determined by the image transform technique. The range estimate is $\hat{\rho}_0 = 20.4$ km, which has a 6.6% error from the GPS measured source range. (f) Beam-time intensity data without any matched filtering. The migration line in the $\tau > 0$ domain is almost vertical, which corresponds to a range estimate of $\hat{\rho}_0 = 554.1$ km, and an over 1000% error of the GPS measured source range. By comparing the source range estimation results shown in (a) and (b), it highlights the need for pulse compression of the kind enabled by nonlinear matched filtering to improve range estimation accuracy for non-impulsive signals whose modal arrivals may not be otherwise resolvable.

C. Estimated positions of multiple vocalizing humpback whales

The receiver array track used for this analysis and the estimated mean positions of the vocalizing humpback whales are shown in Fig. 8(a). The humpback whale mean positions, averaged over roughly 1 h, are estimated by the MAT method[6] and they are used here for independent verification of whale position estimations obtained via the array invariant method. The MAT method uses measured bearings of humpback vocalizations, their arrival time at the receiver array, and GPS-measured receiver array positions as inputs to estimate whale mean positions. The humpback whales were observed to be in localized feeding activities near large herring concentrations[24] well represented by these mean positions. The resulting estimated whale mean positions and corresponding uncertainties in mean position estimation by using the MAT method as well as the bearings of humpback vocalizations measured clockwise from true north are shown in Figs. 8(a) and 8(b), respectively. Note that the bearings shown in Fig. 8(b) are converted from the bearing measurements directly obtained from conventional time-domain beamforming, which is with respect to the array broadside direction. Position estimation error, or the rms distance between the actual and estimated location, for the MAT method is a combination of range and bearing estimation errors. Range estimation error, expressed as the percentage of the range from the true source localization to the horizontal receiver array center, for the MAT method is determined by averaging over the sample variance. The range estimation

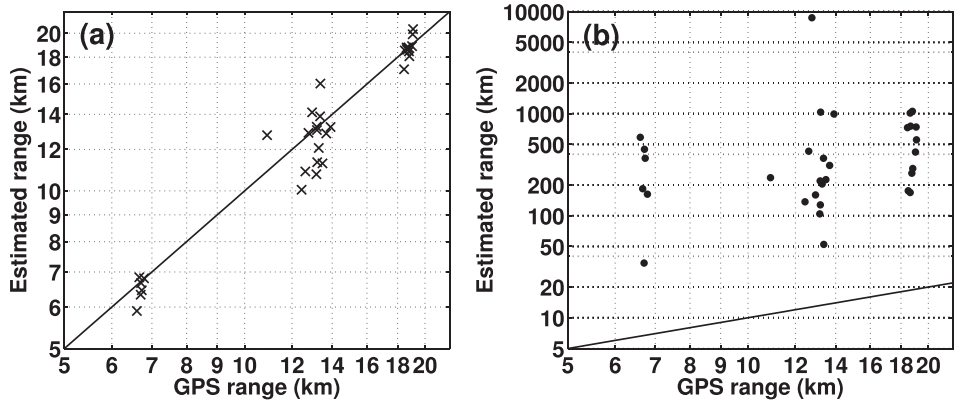

FIG. 11. A total of 32 controlled source transmissions of 2.4-s broadband LFM (275 to 760 Hz) non-impulsive signals are used to quantify the rms error of source range estimation by the generalized array invariant method with nonlinear matched filtering and without any matched filtering. (a) With nonlinear matched filtering, the source range estimate $\hat{\rho}_0$ is plotted in logarithmic scale against GPS measured range ρ_0 for the source range between 6 and 20 km and the source bearing between 10° and 60° from the receiver array endfire direction. The rms error of all range estimates is roughly 9.0% of the GPS measured source range. The solid black line indicates the true source range measured by GPS. (b) Without any matched filtering, the error on average is over 10 000% of the GPS measured source range. Here, the solid line also indicates the true source range measured by GPS. By comparing the source range estimation results shown in (a) and (b), it highlights the need for pulse compression of the kind enabled by nonlinear matched filtering to improve range estimation accuracy for non-impulsive signals whose modal arrivals may not be otherwise resolvable.

error is roughly 2% at array broadside and gradually increases to 10% at 65° from array broadside, and 25% at 90°, i.e., near or at endfire.[6] Bearing estimation error of the conventional time-domain beamformer is roughly 0.5° at array broadside and gradually increases to 6° at endfire.[6] These errors are determined at the same experimental site and time period as the whale position estimates presented in this paper, from thousands of controlled source transmissions with GPS measured ground-truth of source's position,[6] which are accurate to within 3–10 m.[25] As shown in Fig. 8(a), each vocalizing humpback whale position is characterized by a 2D Gaussian probability density function (ellipse) with mean equal to the estimated mean position by the MAT method and standard deviations in range and bearing determined by the measured range and bearing standard deviations. The uncertainty in whale position estimation is shown to be small for Whale #1, because it is located relatively close to the receiver array and in a direction near the array broadside, but this uncertainty is becoming larger as the whale positions are moving further away in range from the receiver array position or they are moving away from the array broadside direction or both, e.g., Whale #4, as can be seen from Fig. 8(a).

D. Data processing of simultaneously recorded humpback whale vocalizations

To obtain the 2D beam-time intensity data $I_{Bo}(s,\tau)$, acoustic recordings of humpback whale vocalizations measured by sensors across the horizontal receiver line array were first converted to 2D beamformed pressure data P_B by conventional time-domain beamforming, and then matched filtered with replica signals generated by a nonlinear matched-filter kernel proposed by Gong *et al.*,[6] which replicates the measured humpback whale vocalizations on the receiver line array. Instead of cross-correlating P_B with a single sensor measurement P_o of the received humpback whale vocalization, which has a low SNR due to a long range propagation through the ocean waveguide from the vocalizing whale to the receiver array, the replica signal P_b reconstructed from a single beam of the array measurements, which has a high SNR due to an array gain of 18 dB, is employed to do the matched filtering. The replica signal P_b can be expressed as

$$P_b(t|t_i) = \frac{1}{\sqrt{E_0}} A(t) e^{-j\nu(t)}, \tag{9}$$

where $A(t)$ and $\nu(t)$ are the amplitude and phase of the beamformed received signal at time t, respectively, and $E_0 = \int_{t_i}^{t_i+T_0} |A(t)|^2 dt$ is the total energy of the beamformed received signal of duration T_0 arriving at time t_i. The amplitude $A(t)$ and phase $\nu(t)$ can be obtained following Eqs. (B3)–(B6) of Ref. 6.

The image transform technique, as described in Sec. II, will be applied to the resulting 2D beam-time intensity image to estimate the source range and bearing. When applying the image transform technique to $I_{Bo}(s,\tau)$ that contains multiple whale vocalizations, a few modifications to Eqs. (5) and (6) are necessary in order to avoid ambiguity in range and bearing estimation. First, instead of integrating $I_{Bo}(s,\tau)$ along semi-infinite straight lines, we only integrate $I_{Bo}(s,\tau)$ of each vocalization within a predefined bounding box in the beam-time domain. Second, due to the fact that the reception times of humpback whale vocalizations are typically not the same, we determine $\tau_i = 0$ for each whale vocalization and form one independent transformed image $\mathcal{I}(s,\phi_r,\tau_i)$ for each whale vocalization via

$$\mathcal{I}(s,\phi_r,\tau_i) = \int_0^{l_i(\phi_r)} |I_{Bo}(s + l\sin\phi_r, \tau_i + l\cos\phi_r)|\, dl, \tag{10}$$

where $l_i(\phi_r)$ is the length of integral path at angle ϕ_r within the predefined bounding box for the i th vocalization. The normalized transformed image $\mathcal{I}_n(s,\phi_r,\tau_i)$ for ith whale vocalization is first obtained via

$$\mathcal{I}_n(s,\phi_r,\tau_i) = \frac{\mathcal{I}(s,\phi_r,\tau_i)}{\max_{s,\phi_r}\mathcal{I}(s,\phi_r,\tau_i)}, \tag{11}$$

before we linearly combine them to form a final normalized transformed image $\mathcal{I}_n(s,\phi_r)$ via

$$\mathcal{I}_n(s,\phi_r) = \frac{\sum_{i=1}^{N}\mathcal{I}_n(s,\phi_r,\tau_i)}{\mathcal{I}_{\max}}, \tag{12}$$

where $\mathcal{I}_{\max} = \max_{s,\phi r}\sum_{i=1}^{N}\mathcal{I}_n(s,\phi_r,\tau_i)$ and N is the number of vocalizations. This normalization procedure is necessary for eliminating the ambiguity in source range and bearing estimation arising from large differences in received sound pressure levels (SPLs) among vocalizations, especially when two vocalizing whales are located close to each other in the azimuthal direction.

E. Experimental results

Here, we show that multiple vocalizing humpback whales can be instantaneously, simultaneously, and unambiguously localized with the generalized array invariant method via the modified image transform technique.

During the receiver array track 571_7 conducted on October 3, 2006, four vocalizing humpback whale individuals were simultaneously present across wide azimuthal directions (Fig. 8) and a series of humpback whale songs and non-song vocalizations from these humpback whale individuals were simultaneously recorded by the horizontal receiver line array. Examples of spectrograms of humpback whale songs and non-song vocalizations from these four humpback individuals over two 30-s time periods are shown in Figs. 12 and 13, respectively. In Fig. 12, it is seen that a song theme from Whale #2 and a non-song vocalization from Whale #3 were recorded during the same period of time. In Fig. 13, another example is shown that a song theme from Whale #1 was simultaneously recorded with multiple non-song vocalizations from Whale #2 to Whale #4 during the same period of time.

We consider two representative cases to demonstrate the performance of the generalized array invariant method. First, two vocalizations from Whale #2 and Whale #3 (vocalizations within the gray boxes in Fig. 12), were received by the horizontal line array roughly at the same time. Second, three

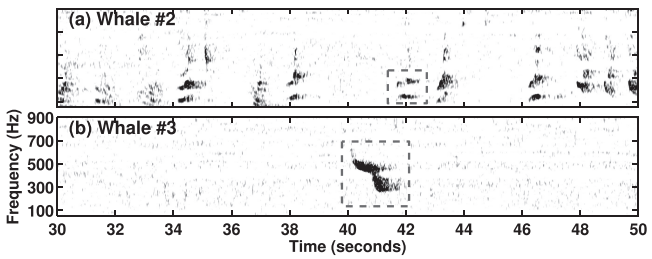

FIG. 12. Spectrograms of (a) a typical humpback whale song theme from Whale #2, and (b) one non-song vocalization from Whale #3 were simultaneously recorded by the horizontal receiver array over a 30-s interval starting at 22:23:00 EDT on October 3, 2006 during GOME'06. Spectrograms were generated after time-domain beamforming. The generalized array invariant method and modified image transform technique, as defined in Eqs. (10) and (12), are applied to the humpback whale vocalizations within the gray boxes to estimate the range and bearing of these vocalizing whales simultaneously.

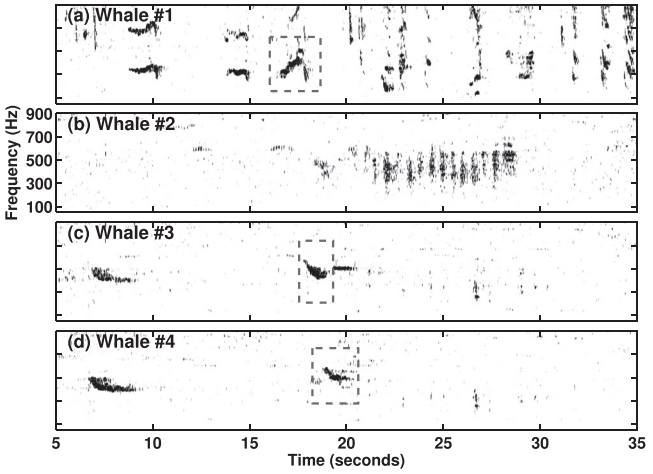

FIG. 13. Similar to Fig. 12, but for data recorded roughly 23 min later, where vocalizations from all four humpback whale individuals were simultaneously recorded over a 30-s interval. Spectrograms were generated after time-domain beamforming. The generalized array invariant method and modified image transform technique are applied to the humpback whale vocalizations within the gray boxes to estimate the range and bearing of these vocalizing whales simultaneously.

vocalizations from Whale #1, Whale #3, and Whale #4 (vocalizations within the gray boxes in Fig. 13), were received by the horizontal line array roughly at the same time.

1. Case I: Simultaneously localizing two vocalizing humpback whales

In this case, a vocalization from Whale #3 was received by the horizontal line array roughly at 22:23:10 EDT on October 3, 2006, approximately 1.5-s before the vocalization

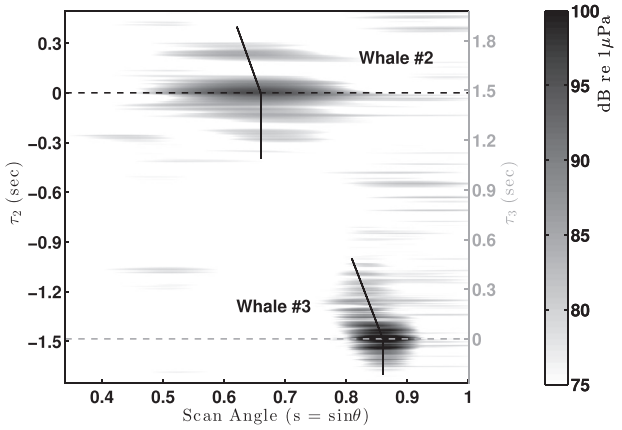

FIG. 14. The 2D beam-time intensity data $I_{Bo}(s,\tau)$ for the geometry of receiver array and two vocalizing humpback whales, Whale #2 and Whale #3, shown in Fig. 8(a). The two solid black lines overlain in the $\tau_2 < 0$ and $\tau_3 < 0$ domain indicate sin θ_2 and sin θ_3, respectively. Two solid black lines in the $\tau_2 > 0$ and $\tau_3 > 0$ domain are the migration lines defined in Eq. (A16) for Whale #2 and Whale #3, respectively, as a function of the reduced travel time $\tau_i = t - \rho_i/c$, where $i = 2, 3$. The dashed black line and dashed light gray line correspond to the reduced travel time $\tau_2 = 0$ and $\tau_3 = 0$ for Whale #2 and Whale #3, respectively. The slight difference in beam-time migration slope appears between $\tau_3 = 0$ and $\tau_3 = 0.25$ and that between $\tau_3 = 0.25$ and $\tau_3 = 0.5$ is due to the additive noise from adjacent beams and scintillation effect in the ocean waveguide. It is not due to the "waterborne modes," since they have insignificant contribution for the ranges examined, as has been noted in Ref. 4.

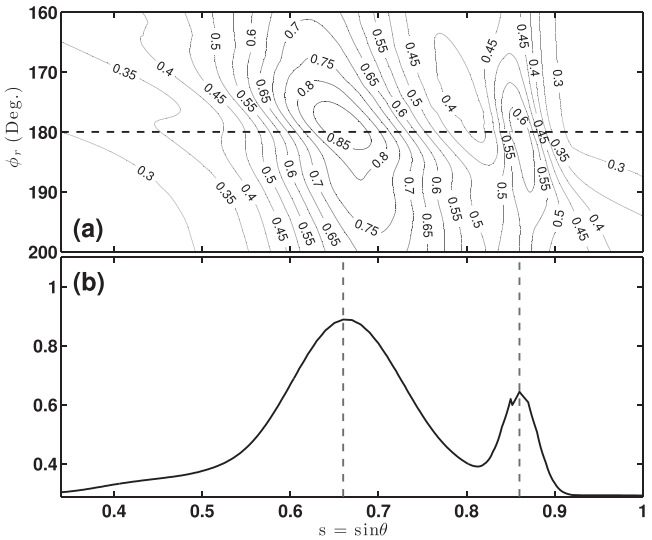

FIG. 15. (a) The normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,\tau)$ shown in Fig. 14 near $\phi_r = 180°$. Two peaks exist along the $\phi_r = 180°$ radial, indicated by the dashed black line, due to two different vocalizing humpback whale bearings. (b) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed black line at $\phi_r = 180°$ in (a). The two peaks in this transect, indicated by the dashed gray lines, correspond to the vocalizing humpback whale bearing estimates $\hat{s}_2 = 0.66$ or $\hat{\theta}_2^a = 41.3°$ and $\hat{s}_3 = 0.86$ or $\hat{\theta}_3^a = 59.3°$ for Whale #2 and Whale #3, respectively.

from Whale #2 was received, as shown in Fig. 14. Applying the MAT method[6] to the bearing estimates of whale vocalization $\hat{\theta}_i^m$, as shown in Fig. 8(b), Whale #3 was estimated to be at range $\hat{\rho}_3^m = 46.4 \pm 4.4$ km and bearing $\hat{\theta}_3^m = 107.5 \pm 2.8°$, and Whale #2 was estimated to be at range $\hat{\rho}_3^m = 44.5 \pm 3.2$ km and bearing $\hat{\theta}_2^m = 125.9 \pm 2.1°$ at these two time instances. Note that $\hat{\theta}_i^m$ is measured clockwise from true north and corresponding receiver array center coordinates. By a simple conversion that accounts for the corresponding receiver array heading $\alpha = 256.7$, which is also measured clockwise from true north, the bearing estimates of Whale #3 and Whale #2 measured from the array broadside direction are then $\hat{\theta}_3^b = 59.2 \pm 2.8°$ and $\hat{\theta}_2^b = 40.8 \pm 2.1°$, respectively.

We now apply the image transform technique, defined in Eqs. (10) and (12), to the 2D beam-time intensity image that contains two distinct humpback whale vocalizations, as shown in Fig. 14. In Fig. 15(a), the transformed image $\mathcal{I}_n(s,\phi_r)$ exhibits two peaks along the dashed black line $\phi_r = 180°$ due to two distinct whale vocalization bearings. The transect of $\mathcal{I}_n(s,\phi_r)$ along $\phi_r = 180°$ is plotted in Fig. 15(b), from which the two bearing estimates are determined to be $\hat{\theta}_3^a = 59.3°$ and $\hat{\theta}_2^a = 41.3°$ for Whale #3 and Whale #2, respectively, which are nearly identical to those estimated by the MAT method $\hat{\theta}_i^b$. Given that the bearing estimation error of the MAT method is typically less than 3° when the source is located outside of the endfire beam of the receiver array,[6] the estimated vocalizing humpback whale bearings $\hat{\theta}_i^a$ are expected to be within 5% of the true whale bearings.

Given the bearing estimate $\hat{s}_i = \sin\hat{\theta}_i^a > 0$ for each vocalizing humpback, the corresponding whale range ρ_i can then be estimated from the normalized transformed image

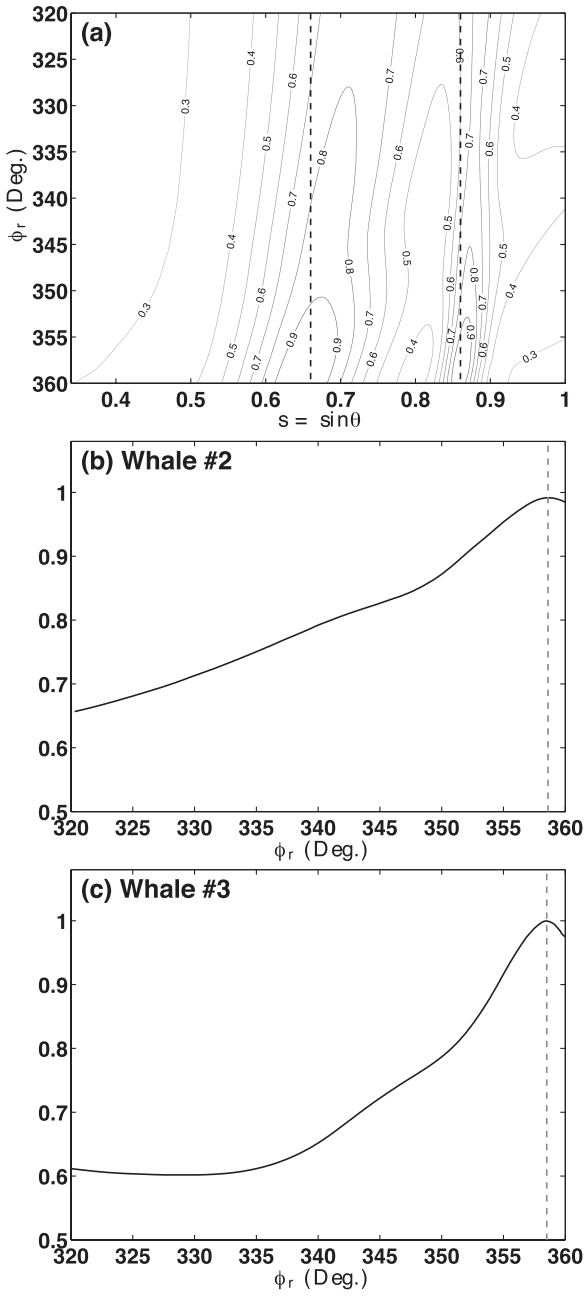

FIG. 16. (a) The normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,\tau)$ shown in Fig. 14 for $320° \le \phi_r \le 360°$. Two dashed black lines are at $\hat{s}_2$ and $\hat{s}_3$, respectively. Two peaks along $\hat{s}_2 = 0.66$ and $\hat{s}_3 = 0.86$ correspond to Whale #2 and Whale #3, respectively. (b) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed gray line at $\hat{s}_2 = 0.66$ in (a). The peak of this transect is at $\phi_r = 358.7°$, which corresponds to the range estimate of $\hat{\rho}_2 = 42.9$ km for Whale #2. (c) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed gray line at $\hat{s}_3 = 0.86$ in (a). The peak of this transect at $\phi_r = 358.5°$ corresponds to Whale #3, which has a range estimate of $\hat{\rho}_3 = 48.4$ km.

$\mathcal{I}_n(\hat{s}_i,\phi_r)$ along the transect $s = \hat{s}_i$. The normalized transformed image $\mathcal{I}_n(s,\phi_r)$ for $320° \le \phi_r \le 360°$ is shown in Fig. 16(a), where two distinct peaks in $\mathcal{I}_n(s,\phi_r)$ correspond to two received humpback whale vocalizations. Transects of Fig. 16(a) along $\hat{s}_3 = \sin\hat{\theta}_3^a = 0.86$ and $\hat{s}_2 = \sin\hat{\theta}_2^a = 0.66$ are shown in Figs. 16(b) and 16(c), where the peak of $\mathcal{I}_n(s_3,\phi_r)$ is seen to occur at $\hat{\phi}_r = 358.5°$ for Whale #3, and the peak of $\mathcal{I}_n(\hat{s}_2,\phi_r)$ is seen to occur at $\hat{\phi}_r = 358.7°$ for

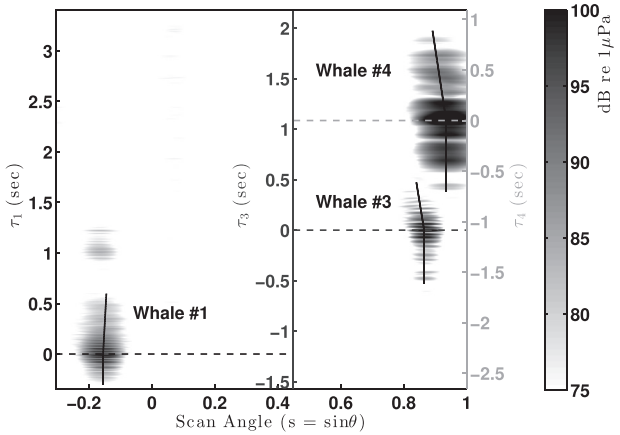

FIG. 17. The 2D beam-time intensity data $I_{Bo}(s,\tau)$ for the geometry of the receiver array and three vocalizing humpback whales, Whale #1, Whale #3, and Whale #4 shown in Fig. 8(a). The three solid black lines overlain in the $\tau_1<0$, $\tau_3<0$, and $\tau_4<0$ domain indicate sin θ_1, sin θ_3, and sin θ_4, respectively. Three solid black lines in the $\tau_1>0$, $\tau_3>0$, and $\tau_4>0$ domain are the migration lines defined in Eq. (A16) for Whale #1, Whale #3, and Whale #4, respectively, as a function of the reduced travel time $\tau_i = t-\rho_i/c$, where $i=1,3,4$. The dashed black line, dashed dark gray line, and dashed light gray line correspond to the reduced travel time $\tau_1=0$, $\tau_3=0$, and $\tau_4=0$ for Whale #1, Whale #3, and Whale #4, respectively.

Whale #2, respectively. The ranges of Whale #3 and Whale #2 are then estimated to be $\hat{\rho}_3^a = 48.4$ km and $\hat{\rho}_2^m = 42.9$ km using Eq. (8), which are both within the uncertainty of $\hat{\rho}_3^m = 46.4\pm4.4$ km and $\hat{\rho}_2^m = 44.5\pm3.2$ km estimated by the MAT method. Given that the position estimation error by the MAT method is roughly 2%–10% of the true range from the vocalizing humpback whale to the receiver array center for acoustic sources located outside of the receiver array endfire beam, the accuracy of range and bearing estimation obtained by the generalized array invariant method is shown to be of

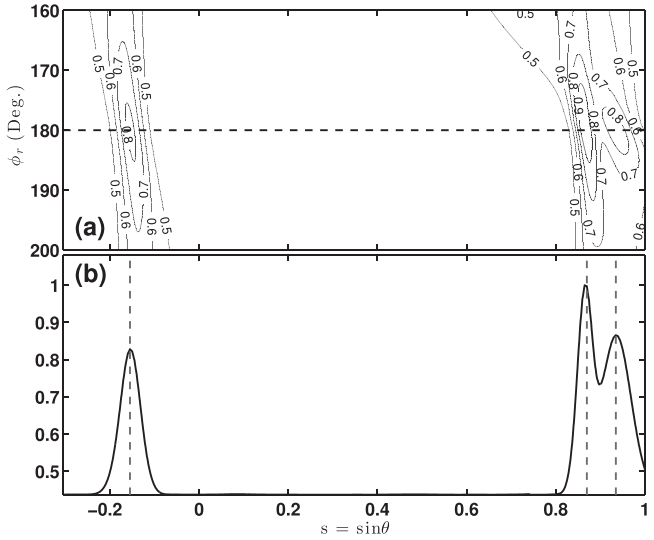

FIG. 18. (a) The normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,\tau)$ shown in Fig. 17 near $\phi_r = 180°$. Three peaks exist along the $\phi_r = 180°$ radial, indicated by the dashed black line, due to three different vocalizing humpback whale bearings. (b) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed black line at $\phi_r = 180°$ in (a). The three peaks in this transect, indicated by the dashed gray lines, correspond to the vocalizing humpback whale bearing estimates $\hat{s}_1 = -0.155$ or $\hat{\theta}_1^a = -8.9°$, $\hat{s}_3 = 0.865$ or $\hat{\theta}_3^a = 59.9°$, and $\hat{s}_4 = 0.935$ or $\hat{\theta}_4^a = 69.2°$ for Whale #1, Whale #3, and Whale #4, respectively.

practical value for instantaneously and simultaneously localizing multiple broadband uncorrelated non-impulsive acoustic sources in a range-dependent ocean waveguide.

2. Case II: Simultaneously localizing three vocalizing humpback whales

In this case, three vocalizations from Whale #1, Whale #3, and Whale #4 were received by the horizontal line array roughly at the same time around 22:46:33 EDT on October 3, 2006, as shown in Fig. 17. The estimated whale ranges and bearings by the MAT method are $\hat{\rho}_1^m = 22.0\pm1.1$ km and $\hat{\theta}_1^b = -8.8\pm0.8°$ for Whale #1, $\hat{\rho}_3^m = 44.4\pm4.2$ km and $\hat{\theta}_3^b = 59.9\pm2.8°$ for Whale #3, and $\hat{\rho}_4^m = 51.5\pm7.1$ km and $\hat{\theta}_4^b = 69.3\pm3.4°$ for Whale #4.

Applying the image transform technique to Fig. 17, the normalized transformed image $\mathcal{I}_n(s,\phi_r)$ shown in Fig. 18(a) exhibits three peaks along the dashed black line $\phi_r = 180°$. The transect of $\mathcal{I}_n(s,\phi_r)$ along $\phi_r = 180°$ is plotted in Fig. 18(b), from which the three bearing estimates are determined to be $\hat{\theta}_1^a = -8.9°$, $\hat{\theta}_3^a = 59.9°$, and $\hat{\theta}_4^a = 69.2°$ for Whale #1, Whale #3, and Whale #4, respectively, which are again nearly identical to those estimated by the MAT method.

Given the bearing estimate $\hat{s}_i = \sin\hat{\theta}_i^a$ for each vocalizing humpback whale, the corresponding whale range ρ_i can be estimated from $\mathcal{I}_n(\hat{s}_i,\phi_r)$. For this example with $\hat{s}_1 = -0.155 < 0$, $\hat{s}_3 = 0.865 > 0$, and $\hat{s}_4 = 0.935 > 0$, the normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ for $0° \le \phi_r \le 40°$ and $320° \le \phi_r \le 360°$ is shown in Fig. 19(a), where three distinct peaks in $\mathcal{I}_n(s,\phi_r)$ correspond to three received whale vocalizations. Transects of Fig. 19(a) along $\hat{s}_1$, $\hat{s}_3$, and $\hat{s}_4$ are, respectively, shown in Figs. 19(b)–19(d), where the peak of $\mathcal{I}_n(\hat{s}_i,\phi_r)$, where $i=1,3,4$, is seen to occur at $\hat{\phi}_r = 0.6°$ for Whale #1, at $\hat{\phi}_r = 358.4°$ for Whale #3, and at $\hat{\phi}_r = 358.5°$ for Whale #4, respectively. The ranges of Whale #1, Whale #3, and Whale #4 are then estimated to be $\hat{\rho}_1^a = 21.8$ km, $\hat{\rho}_3^a = 45.7$ km, and $\hat{\rho}_4^a = 52.7$ km using Eq. (8), which are all within the uncertainty of $\hat{\rho}_1^m = 22.0\pm1.1$ km, $\hat{\rho}_3^m = 44.4\pm4.2$ km and $\hat{\rho}_4^m = 51.5\pm7.1$ km obtained independently via the MAT method. This shows that the array invariant method can consistently provide robust source range estimation.

IV. CONCLUSION

The array invariant method, previously derived for instantaneous range and bearing estimation of a single broadband impulsive source in a horizontally stratified ocean waveguide, can be generalized to instantaneously and simultaneously localize multiple uncorrelated broadband noise sources that are not necessarily impulsive in the time domain. The theory is exact in an ideal waveguide, where the received beam-time intensity data along the array is bounded by the migration of matched filter peak in the beam-time domain arising from the differences in the modal group velocity and modal polar angle due to modal dispersion. These bounding lines follow a unique dependence on the source range and bearing and are invariant to the ocean environments. This theory is approximately valid in a horizontally

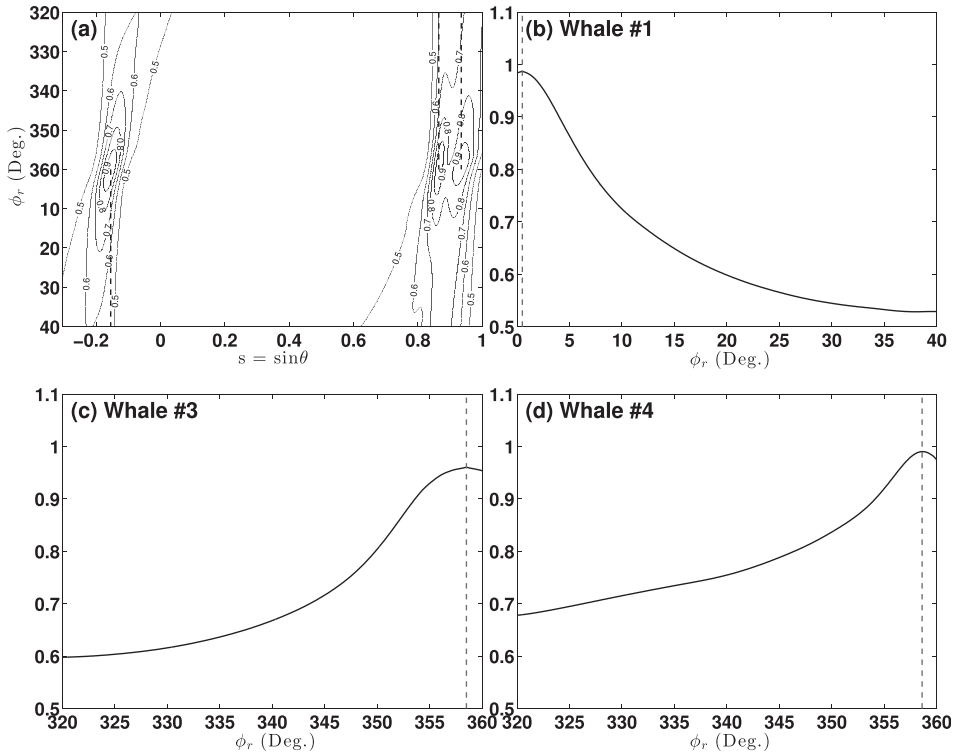

FIG. 19. (a) The normalized transformed intensity image $\mathcal{I}_n(s,\phi_r)$ of the 2D beam-time intensity data $I_{Bo}(s,\tau)$ shown in Fig. 17 for $320^\circ \le \phi_r \le 360^\circ$ and $0^\circ \le \phi_r \le 40^\circ$. Three dashed black lines are at $\hat{s}_1$, $\hat{s}_3$, and $\hat{s}_4$, respectively. Three peaks along $\hat{s}_1 = -0.155$, $\hat{s}_2 = 0.865$, and $\hat{s}_2 = 0.935$ correspond to Whale #1, Whale #3, and Whale #4, respectively. (b) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed gray line at $\hat{s}_1 = -0.155$ in (a). The peak of this transect is at $\phi_r = 0.6^\circ$, which corresponds to the range estimate of $\hat{\rho}_1^a = 21.8$ km for Whale #1. (c) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed gray line at $\hat{s}_3 = 0.865$ in (a). The peak at $\phi_r = 358.4^\circ$ corresponds to Whale #3, which has a range estimate of $\hat{\rho}_3^a = 45.7$ km. (d) Transect of $\mathcal{I}_n(s,\phi_r)$ along the dashed gray line at $\hat{s}_4 = 0.935$ in (a). The peak at $\phi_r = 358.5^\circ$ corresponds to Whale #4, which has a range estimate of $\hat{\rho}_4^a = 52.7$ km.

stratified ocean waveguide if the condition that there is at least one propagation mode whose cutoff frequency is sufficiently lower than that of the source is satisfied. A theoretical example for simultaneously localizing sources at different ranges but within the same beam in a Pekeris waveguide shows that unambiguous source range and bearing estimation can be achieved via a image processing technique similar to the Radon transform.

The generalized array invariant method is applied to localize a vertical source array with known GPS measured ground truth positions deployed north of Georges Bank during the GOME'06. This vertical source array transmitted a 0.6- or 1-s broadband (50 Hz bandwidth) LFM non-impulsive signal every 50- or 75-s at ranges between 1 and 20 km from a towed horizontal receiver array in a range-dependent ocean waveguide with approximately 50% variation in water depth. The rms error for source range estimates by the generalized array invariant method including the NMF of all 2547 controlled source transmissions of this waveform within 0° to 80° of the receiver array endfire is roughly 12% of the true source range and is on average roughly 800% without any matched filtering, indicating that pulse compression of the kind enabled by nonlinear matched filtering is necessary to make source localization possible in cases where temporal overlap between modal arrivals cannot be otherwise resolved. Applying the same approach to over 32 controlled source transmissions of 2.4-s broadband LFM (275 to 750 Hz) non-impulsive signals, the rms errors are found to be roughly 9% with nonlinear matched filtering and on average more than 10 000% without any matched filtering of the true source range for source ranges between 6 and 20 km and source bearings between 10° and 60°. This again highlights the need for pulse compression of the kind enabled by nonlinear matched filtering when temporal overlap between modal arrivals cannot be otherwise resolved.

The same approach is then applied to simultaneously localize multiple vocalizing humpback whales from passive acoustic recordings acquired by the same horizontal line array, at the same experimental site, and during the same time period. The vocalizing humpback whales used in this study are located tens of kilometers away from the horizontal line array and distributed over a vast swath of northern flank of Georges Bank. The ability to make simple, unambiguous, and instantaneous whale range and bearing estimates by the generalized array invariant method has been demonstrated with two representative examples. The resulting whale range and bearing estimates made by the generalized array invariant method are found to be highly consistent with those obtained independently via the MAT method with typical range estimation error below 10%.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research. We would like to thank Dr. Sunwoong Lee for giving us permission to reproduce some of the figures appearing in his doctoral thesis and many insightful comments he had provided for the progress of this work.

APPENDIX A: DERIVATION OF THE GENERALIZED ARRAY INVARIANT

The derivation of the generalized array invariant method shown in this section follows the derivation appearing in Chap. 3 of Ref. 5.

The same geometry of spatial and wavenumber coordinates, as shown in Figs. 1(a) and 2 of Ref. 4 for a horizontal line array, are employed in this paper, where the origin of the spatial coordinate system is placed at the air-water interface with the positive z axis pointing downward. The horizontal line array is parallel to the y axis with the array center located at the horizontal origin $\mathbf{r} = (x = 0, y = 0, z)$, and the acoustic source is located at $\mathbf{r_o} = (x_o, y_o, z_o)$. The locations of the array center and acoustic source are also defined as $\mathbf{r} = (\boldsymbol{\rho}, z)$, where $\rho = (0, 0)$ and $\mathbf{r_o} = (\boldsymbol{\rho_o}, z_o)$, where $\boldsymbol{\rho_o} = (x_o, y_o)$ and $\rho_o = |\boldsymbol{\rho_o}| = \sqrt{x_o^2 + y_o^2}$, respectively. The horizontal and vertical wavenumber components for the nth mode in an isovelocity layer[20,26] are $k_{rn} = -k \sin \phi_n$ and $k_{zn} = -k \cos \phi_n$, where ϕ_n is the polar angle of the mode measured from the z axis, and $k = \sqrt{k_{rm}^2 + k_{zn}^2}$ is the wavenumber.

In an ocean waveguide, the pressure field P_o at frequency f measured at the center of the array $\mathbf{r}$ due to a source at $\mathbf{r_o}$ can be expressed using normal mode theory as[5]

$$P_o(\boldsymbol{\rho} = \mathbf{0}, z, f) = 4\pi Q(f) \frac{i}{\sqrt{8\pi} d(z_o)} e^{-i\pi/4} \times \sum_n u_n(z_o)\, u_n(z) \frac{e^{ik_{rn}|\boldsymbol{\rho} - \boldsymbol{\rho_o}|}}{\sqrt{k_{rn}|\boldsymbol{\rho} - \boldsymbol{\rho_o}|}}, \tag{A1}$$

where $Q(f)$ is the source spectrum of broadband random noise which is not necessarily impulsive in the time domain, $d(z_o)$ is the density at source depth z_o, $u_n(z_o)$ and $u_n(z)$ are the amplitude of mode shape for the nth mode at source depth z_o and receiver depth z, respectively. The mode shape u_n satisfies $\int_0^\infty u_m(z) u_n^*(z)/d(z) dz = \delta_{mn}$.

When the source is located in the far-field of the receiver array, using the far-field approximation $|\boldsymbol{\rho} - \boldsymbol{\rho_o}| \simeq \rho_o - y \sin \theta_o$, where θ_o is the source bearing measured from the x axis, the beamformed pressure field P_B at frequency f can be expressed as a function of $s = \sin \theta$,[5]

$$\begin{aligned} P_B(s,f) &= \int_{-\infty}^{\infty} T(v) P(v, z, f) e^{i2\pi v \sin\theta} dv \\ &= 4\pi Q(f) \frac{i}{\sqrt{8\pi} d(z_o)} e^{-i\pi/4} \\ &\quad \times \sum_n u_n(z_o) u_n(z) \frac{e^{ik_{rn}\rho_o}}{\sqrt{k_{rn}\rho_o}} B(s - s_n), \end{aligned} \tag{A2}$$

where $v = ky/2\pi$, $s_n = \sin \phi_n \sin \theta_o$, $\sin \phi_n = k_{rn}/k$, and $B(s)$ is the array beam pattern, which is the spatial Fourier transform of the array taper function $T(v)$. The far field approximation is valid when $|\boldsymbol{\rho} - \boldsymbol{\rho_o}| > L^2/\lambda$, where L is the receiver array aperture length and $\lambda = 2\pi/k$ is the acoustic signal wavelength.

In Ref. 4, it has been shown that the array invariant method derived specifically for an impulsive source in the time domain can be applied to any continuous broadband noise signatures if the received pressure field is matched filtered with the source spectrum. This can be achieved by cross-correlating the beamformed received pressure field P_B with the single sensor measured pressure field P_o. The resulting matched filtered output I_{Bo} as a function of $s = \sin \theta$ and reduced travel time τ can be expressed as[5]

$$\begin{aligned} I_{Bo}(s, \tau) &= \lim_{T \to \infty} \frac{1}{T} \int_0^T P_o^*(t) P_B(s, t + \tau) dt \\ &= \mathcal{F}^{-1} \{ P_o^*(\boldsymbol{\rho} = \mathbf{0}, z, f) P_B(s, f) \} \\ &= 2\mathrm{Re} \left\{ \int_0^\infty P_o^*(\boldsymbol{\rho} = \mathbf{0}, z, f) P_B(s, f) e^{-i2\pi f t} df \right\} \\ &= 2\mathrm{Re} \left\{ \sum_m \sum_n E_{mn}(s, \tau) \right\}, \end{aligned} \tag{A3}$$

where $\mathrm{Re}\{\cdot\}$ represents the real part, and $E_{mn}(s, \tau)$ is the complex envelope of $I_{Bo}(s, \tau)$ between the mth mode in P_B and the nth mode in P_o given by[5]

$$E_{mn}(s, \tau) = \frac{2\pi}{d^2(z_o)} \int_0^\infty |Q(f)|^2 u_m(z_o) u_m(z) u_n^*(z_o) u_n^*(z) \times \frac{B(s - s_{mn})}{\sqrt{k_{rm} k_{rn}} \rho_o} e^{i\rho_o \psi_{mn}} df, \tag{A4}$$

and $\psi_{mn} = (k_{rm} - k_{rn}) - 2\pi f \tau/\rho_o$. Note that $I_{Bo}(s, \tau)$ defined in Eq. (A3) only depends on the source power spectrum $|Q(f)|^2$ but not on the detailed shape of the source signal in the time domain. The correlation between modes, as indicated by the double summation in Eq. (A3), reveals the underlying physics and modal arrival structure that make this method unique for localizing any signals with arbitrary time dependence in an ocean waveguide.

Now, the structure of $E_{mn}(s, \tau)$ is analyzed in the beam-time domain. First, in case of $m = n$, Eq. (A4) is simplified to[5]

$$E_{mm}(s, \tau) = \frac{2\pi}{d^2(z_o)} \int_0^\infty |Q(f)|^2 |u_m(z_o)|^2 |u_m(z)|^2 \times \frac{B(s - s_m)}{k_{rm} \rho_o} e^{-i2\pi f \tau} df, \tag{A5}$$

which is identical to the auto-correlation of the mth mode, except the beam pattern $B(s - s_m)$ is inside the integral. If the change of $B(s - s_m)$ is relatively slow as a function of frequency, E_{mm} is confined near $\tau \simeq 0$, and the width of E_{mm} in the τ domain is inversely proportional to the bandwidth of the source power spectrum $|Q(f)|^2$. Second, in case of $m \neq n$, the complex $E_{mn}(s, \tau)$ can then be approximated as

$$E_{mn}(s, \tau) \simeq \frac{2\pi}{d^2(z_o)} |Q(\tilde{f})|^2 \tilde{u}_m(z_0) \tilde{u}_m(z) \tilde{u}_n^*(z_0) \tilde{u}_n^*(z) \times \frac{\tilde{B}(s - \tilde{s}_{mn})}{\sqrt{\tilde{k}_{rm} \tilde{k}_{rn}} \rho_o} F_n(\tilde{f}), \tag{A6}$$

using the method of stationary phase[27–30] when $\psi_{mn} \rho_o \gg 1$, where $\tilde{f}$ is the dominant frequency component within the source band that satisfies[5]

$$\left[\frac{\partial \psi_{mn}}{\partial f} \right]_{f = \tilde{f}} = \left[\frac{\partial k_{rm}}{\partial f} - \frac{\partial k_{rn}}{\partial f} - \frac{2\pi\tau}{\rho_o} \right]_{f = \tilde{f}} = 0, \tag{A7}$$

and $\tilde{u}_n$, $\tilde{u}_m$, $\tilde{B}$, $\tilde{s}_{mn}$, $\tilde{k}_{rn}$, $\tilde{k}_{rm}$ are the corresponding values of u_n, u_m, B, s_{mn}, k_{rn}, k_{rm} at $f = \tilde{f}$. The function $F_n(\tilde{f})$ in Eq. (A6) is given in Eq. (A9) of Appendix A of Ref. 4.

The delay time $\tilde{\tau}_{mn}$ is defined to be the time that satisfies Eq. (A7),[5]

$$\tilde{\tau}_{mn} \equiv \tau_{mn}(\tilde{f}) = \frac{\rho_o}{2\pi}\left[\frac{\partial k_{rm}}{\partial f} - \frac{\partial k_{rn}}{\partial f}\right]_{f=\tilde{f}} = \rho_o\left(\frac{1}{v_{gm}(\tilde{f})} - \frac{1}{v_{gn}(\tilde{f})}\right) = \rho_o\left(\frac{1}{\tilde{v}_{gm}} - \frac{1}{\tilde{v}_{gn}}\right), \quad \text{(A8)}$$

which is also the travel time difference between the mth mode and the nth mode and at $f = \tilde{f}$. The peak of $E_{mn}(s,\tau)$ at this stationary point $\tau = \tilde{\tau}_{mn}$ in the s domain occurs at[5]

$$\tilde{s}_{mn} \equiv s(\tilde{\tau}_{mn}) = \sin\phi_m(\tilde{\tau}_{mn})\sin\theta_o, \quad \text{(A9)}$$

which is the zero argument of the beam pattern $\tilde{B}(s - \tilde{s}_{mn})$ in Eq. (A6). Equation (A9) shows that $\tilde{s}_{mn}$ only depends on the polar angle ϕ_m of the mth mode and the source bearing θ_o. However, this polar angle depends on the travel time difference $\tilde{\tau}_{mn}$ between the mth and nth mode.

In Appendix A 1 and A 2, the migration of $\tilde{s}_{mn}$ in Eq. (A9) as a function of mode numbers m and n, and the delay time $\tilde{\tau}_{mn}$ in the beam-time domain is analyzed to identify its unique relationship with the source range and bearing. It is shown that the migration of $\tilde{s}_{mn}$ reveals the source range ρ_o when $\tilde{\tau}_{mn} > 0$, and the migration of $\tilde{s}_{mn}$ reveals the source bearing θ_o when $\tilde{\tau}_{mn} < 0$. The structure of E_{mn} in an ideal waveguide with rigid or pressure-release boundaries is first analyzed. Then, the analysis is shown to be approximately valid in a horizontally stratified ocean waveguide, where the source range can be instantaneously obtained without *a priori* knowledge of the environment, except the sound speed $c(z)$ at the receiver depth z.

A. Case I: $\tilde{\tau}_{mn} > 0$

This occurs when the group velocity of the nth mode in P_o is faster than that of the mth mode in P_B, as can be seen from Eq. (A8). It is shown that the spread of E_{mn} in the s and $\tau > 0$ domain is bounded by the migration line for an impulsive source, defined in Eq. (8) of Ref. 4, from which estimates of the array invariant $\hat{\chi}_h$ and the source range $\hat{\rho}_o$ can be obtained.

1. Ideal waveguide

In an ideal waveguide, the group velocity of the nth mode v_{gn} is related to the polar angle ϕ_n via[5]

$$[v_{gn}]^{-1} = \frac{1}{2\pi}\frac{d}{df}\sqrt{k^2 - k_{zn}^2} = \frac{1}{2\pi}\frac{dk}{df}\frac{k}{k_{rn}} = [c\sin\phi_n]^{-1}, \quad \text{(A10)}$$

for all frequencies f. Then the travel time difference between two different modes in Eq. (A8) can be expressed as[5]

$$\tilde{\tau}_{mn} = \rho_o\left(\frac{1}{\tilde{v}_{gm}} - \frac{1}{\tilde{v}_{gn}}\right) = \frac{\rho_o}{c}\frac{\sin\tilde{\phi}_n - \sin\tilde{\phi}_m}{\sin\tilde{\phi}_m\sin\tilde{\phi}_n}, \quad \text{(A11)}$$

which leads to[5]

$$\sin\tilde{\phi}_m = \frac{\rho_o\sin\tilde{\phi}_n}{c\tilde{\tau}_{mn}\sin\tilde{\phi}_n + \rho_o}, \quad \text{(A12)}$$

where $\sin\tilde{\phi}_m = \sin\phi_m(\tilde{f})$, and $\sin\tilde{\phi}_n = \sin\phi_n(\tilde{f})$. Substituting Eq. (A12) into Eq. (A9), $\tilde{s}_{mn}$ can be rewritten as[5]

$$\tilde{s}_{mn} = \frac{\rho_o\sin\tilde{\phi}_n}{c\tilde{\tau}_{mn}\sin\tilde{\phi}_n + \rho_o}\sin\theta_o. \quad \text{(A13)}$$

If $\sin\tilde{\phi}_n$ is approximately constant over the source frequency band, the derivative of $\tilde{s}_{mn}^{-1}$ with respect to $\tilde{\tau}_{mn}$ yields[5]

$$\frac{\partial\tilde{s}_{mn}^{-1}}{\partial\tilde{\tau}_{mn}} \simeq \frac{c\sin\tilde{\phi}_n}{\rho_o\sin\tilde{\phi}_n\sin\theta_o} = \frac{c}{\rho_o\sin\theta_o} = \chi_h, \quad \text{(A14)}$$

where χ_h is the array invariant for a horizontal line array in an ideal waveguide, as defined in Eq. (9) of Ref. 4. This is satisfied when (1) the source frequencies are much higher than the cutoff frequency of the nth mode so that $\sin\tilde{\phi}_n$ is nearly constant and is close to 1, as has been shown in Fig. 3 of Ref. 4, or (2) the frequency band of the source is sufficiently narrow so that the change of $\sin\tilde{\phi}_n$ over the frequency band is essentially negligible. In the context of this paper, the first case is of primary interest, where $\sin\tilde{\phi}_n \simeq 1$. Then group velocity $\tilde{v}_{gn}$ of the nth mode is approximately equal to the water column sound speed c, as can be seen from Eq. (A10), and Eq. (A13) reduces to[5]

$$\tilde{s}_{mn} \simeq \frac{\rho_o}{c\tilde{\tau}_{mn} + \rho_o}\sin\theta_o = \frac{\rho_o\sin\theta_o}{c(\tilde{\tau}_{mn} + \rho_o/c)} = \frac{\rho_o}{ct}\sin\theta_o, \quad \text{(A15)}$$

where $t = \tilde{\tau}_{mn} + \rho_o/c$ is approximately equal to the total travel time of the mth mode, since $\rho_o/c \simeq \rho_o/\tilde{v}_{gn}$ is the travel time of the nth mode, and $\tilde{\tau}_{mn}$ is the travel time difference between the mth and nth modes. Equation (A15) is then equivalent to Eq. (8) of Ref. 4, where[5]

$$\tilde{s}(t) = \frac{\rho_o}{ct}\sin\theta_o \quad \text{(A16)}$$

is the migration curve for an impulsive source in an ideal waveguide. This migration curve is approximately a straight line when s is close to the sine of the source bearing $\sin\theta_o$ and the reduced travel time τ is sufficiently small, as shown in Fig. 4 of Ref. 4. Additionally, when $\tilde{\tau}_{mn}$ is sufficiently close to zero, e.g., $\tilde{s}_{mn}(\tilde{\tau}_{mn} \simeq 0) \simeq \sin\theta_o$, it is shown that $\tilde{s}_{mn}$ emerges from the true source bearing $\sin\theta_o$.

2. Horizontally stratified waveguide

For general horizontally stratified waveguides, the relation between group velocity v_{gn} and modal polar angle ϕ_n for the nth mode is[5]

$$v_{gn} = \frac{c(z)\sin\phi_n(z)}{1 + D_n(z)}, \quad \text{(A17)}$$

where $D_n(z) = [c(z)/2\pi]\cos\phi_n(z)[dk_{zn}(z)/df]$ is the correction term when there is a variation in sound speed structure versus depth, as defined in Eq. (11) of Ref. 4. Substituting Eq. (A17) into Eq. (A8), the delay time $\tilde{\tau}_{mn}$ can be expressed as[5]

$$\tilde{\tau}_{mn} = \rho_o\left[\frac{1}{\tilde{v}_{gm}} - \frac{1}{\tilde{v}_{gn}}\right] = \frac{\rho_o}{c(z)}\frac{\sin\tilde{\phi}_n(z)\left[1+\tilde{D}_m(z)\right] - \sin\tilde{\phi}_m(z)\left[1+\tilde{D}_n(z)\right]}{\sin\tilde{\phi}_m \sin\tilde{\phi}_n(z)}. \tag{A18}$$

Now, an assumption is made that there is at least one propagating mode whose cutoff frequency is much lower than that of the source, which leads to $\sin\tilde{\phi}_n \simeq 1$, $\cos\tilde{\phi}_n \simeq 0$, and $\tilde{D}_n(z) \simeq 0$ within the source frequency band. In addition, for most of the propagating modes in shallow-water waveguides, Eq. (13) of Ref. 4 is satisfied, since change of vertical wavenumber versus frequency is typically negligible for frequencies not near modal cutoff, as shown in Fig. 8 of Ref. 4. Therefore, the correction term $\tilde{D}_m(z)$ is also negligible. Then, $\tilde{\tau}_{mn}$ in Eq. (A18) can be simplified to[5]

$$\tilde{\tau}_{mn} \simeq \frac{\rho_o}{c(z)}\frac{1-\sin\tilde{\phi}_m(z)}{\sin\tilde{\phi}_m(z)}, \tag{A19}$$

which leads to[5]

$$\tilde{s}_{mn} = \sin\tilde{\phi}_m(z)\sin\theta_o \simeq \frac{\rho_o \sin\theta_o}{c(z)\tilde{\tau}_{mn} + \rho_o}. \tag{A20}$$

Now, the array invariant χ_h for a general horizontally stratified waveguide can be obtained from Eq. (A20) as[5]

$$\chi_h \equiv \frac{\partial\tilde{s}_{mn}^{-1}}{\partial\tilde{\tau}_{mn}} \simeq \frac{c(z)}{\rho_o \sin\theta_o}, \tag{A21}$$

which is similar to Eq. (A14) but with $c(z)$ substituted for c.

In a horizontally stratified ocean waveguide, the maximum extent of $\tilde{s}_{mn}$ in the τ-domain is finite as is the beam-time migration line for an impulsive source,[4] since the group velocities of the non-waterborne modes are bounded by the minimum group velocity at the Airy phase.[31] Equation (A20) can then be expanded using the Taylor series at $\tilde{\tau}_{mn} = 0$, and the linearized array invariant χ_l can be defined as[5]

$$\chi_l \equiv \left.\frac{\partial\tilde{s}_{mn}}{\partial\tilde{\tau}_{mn}}\right|_{\tilde{\tau}_{mn}=0} \simeq -\frac{c(z)\sin\theta_o}{\rho_o}, \tag{A22}$$

which is convenient for practical use.

B. Case II: $\tau_{mn} < 0$

This occurs when the group velocity of the mth mode in P_B is faster than the group velocity of the nth mode in P_o. It is shown that the migration of $\tilde{s}_{mn}$ in this case is bounded by the bearing of the source. The bearing of the source can then be estimated from this bounding line.

1. Ideal waveguide

An assumption is made that there is at least one propagating mode whose cutoff frequency is much lower than that of the source, which lead to $\sin\tilde{\phi}_m \simeq 1$. As a result, the group velocity $\tilde{v}_{gm}$ of the mth mode is approximately equal to the water column sound speed c. The travel time difference $\tilde{\tau}_{mn}$ and the migration line $\tilde{s}_{mn}$ can then be approximated as[5]

$$\tilde{\tau}_{mn} = \rho_o\left(\frac{1}{\tilde{v}_{gm}} - \frac{1}{\tilde{v}_{gn}}\right) \simeq \frac{\rho_o}{c}\left(1 - \frac{1}{\sin\tilde{\phi}_n}\right) \tag{A23}$$

and

$$\tilde{s}_{mn} = \sin\phi_m(\tilde{\tau}_{mn})\sin\theta_o \simeq \sin\theta_o, \tag{A24}$$

respectively.

Equations (A23) and (A24) show that the location of $\tilde{s}_{mn}$ in the τ-domain is determined by the polar angle $\tilde{\phi}_n$ of the nth mode; however, the location of $\tilde{s}_{mn}$ in the s-domain is independent of $\tilde{\phi}_n$, and is always at the sine of the source bearing $\sin\theta_o$.

2. Horizontally stratified waveguide

For general horizontally stratified waveguides, the travel time difference $\tilde{\tau}_{mn}$ defined in Eq. (A23) for an ideal waveguide now changes to[5]

$$\tilde{\tau}_{mn} = \frac{r_o}{c(z)}\left[1 - \frac{1+\tilde{D}_n(z)}{\sin\tilde{\phi}_n(z)}\right], \tag{A25}$$

and the migration of $\tilde{s}_{mn}$, however, is still given by

$$\tilde{s}_{mn} = \sin\phi_m(\tilde{\tau}_{mn})\sin\theta_o \simeq \sin\theta_o, \tag{A26}$$

if the assumption that there is at least one propagating mode that satisfies the condition $\sin\tilde{\phi}_m \simeq 1$, $\cos\tilde{\phi}_m \simeq 0$, and $\tilde{D}_m(z) \simeq 0$ is valid. Equations (A25) and (A26) show that the correction term $\tilde{D}_n(z)$ can change the travel time difference $\tilde{\tau}_{mn}$ and thus the location of $\tilde{s}_{mn}$ in the τ-domain, but not the location of $\tilde{s}_{mn}$ in the s-domain. Then, the migration line $\tilde{s}_{mn}$ in the $\tau < 0$ domain is still bounded by the sine of the source bearing $\sin\theta_o$, as was the case in an ideal waveguide.

[1]D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Prentice Hall, Upper Saddle River, NJ, 1993), Chap. 1.
[2]G. C. Carter, "Time delay estimation for passive sonar signal processing," IEEE Trans. Acoust., Speech, Signal Process. **29**, 463–470 (1981).
[3]J. C. Hassab, *Underwater Signal and Data Processing* (CRC Press, Boca Raton, FL, 1989), pp. 205–308.
[4]S. Lee and N. C. Makris, "The array invariant," J. Acoust. Soc. Am. **119**, 336–351 (2006).
[5]S. Lee, "Efficient localization in a dispersive waveguide: Applications in terrestrial continental shelves and on Europa," Ph.D. thesis, Massachusetts Institute of Technology (2006).
[6]Z. Gong, D. D. Tran, and P. Ratilal, "Comparing passive source localization and tracking approaches with a towed horizontal receiver array in an ocean waveguide," J. Acoust. Soc. Am. **134**, 3705–3720 (2013).

[7]S. Lee and N. C. Makris, "A new invariant method for instantaneous source range estimation in an ocean waveguide from passive beam-time intensity data," J. Acoust. Soc. Am. **116**, 2646 (2004).
[8]K. Søstrand, "Range localization of 10-100 km explosions by means of an endfire array and a waveguide invariant," IEEE J. Ocean. Eng. **30**, 207–212 (2005).
[9]T. C. Yang, "Beam intensity striations and applications," J. Acoust. Soc. Am. **113**, 1342–1352 (2003).
[10]References 7 and 8 have temporally interwoven publication and submission dates, while Ref. 9 was published before any of these, with similarities and differences noted in Sec. I.
[11]M. J. Noad, D. H. Cato, and M. D. Stokes, "Acoustic tracking of humpback whales: Measuring interactions with the acoustic environment," in *Proceedings of Acoustics*, Gold Coast, Australia (2004), pp. 353–358.
[12]D. Cato, R. McCauley, T. Rogers, and M. Noad, "Passive acoustics for monitoring marine animals—progress and challenges," in *Proceedings of Acoustics*, Christchurch, New Zealand (2006), pp. 453–460.
[13]W. W. L. Au, J. Mobley, W. C. Burgess, M. O. Lammers, and P. E. Nachtigall, "Seasonal and diurnal trends of chorusing humpback whales wintering in waters off Western Maui," Mar. Mammal Sci. **16**, 530–544 (2000).
[14]P. O. Thompson, W. C. Cummings, and S. J. Ha, "Sounds, source levels and associated behavior of humpback whales, Southeast Alaska," J. Acoust. Soc. Am. **80**, 735–740 (1986).
[15]M. Noad and D. Cato, "A combined acoustic and visual survey of humpback whales off southeast Queensland," Mem. Queensl. Mus. **47**, 507–523 (2001).
[16]M. A. McDonald and C. G. Fox, "Passive acoustic methods applied to fin whale population density estimation," J. Acoust. Soc. Am. **105**, 2643–2651 (1999).
[17]W. Watkins, M. Daher, G. Reppucci, J. George, D. Martin, N. DiMarzio, and D. Gannon, "Seasonality and distribution of whale calls in the North Pacific," Oceanography **13**, 62–67 (2000).
[18]J. Barlow and B. Taylor, "Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey," Mar. Mammal Sci. **21**, 429–445 (2005).
[19]A. Thode, "Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array," J. Acoust. Soc. Am. **116**, 245–253 (2004).
[20]F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. (Springer-Verlag, New York, 2011), Chap. 5.
[21]S. R. Deans, "Radon and Abel transform," in *The Transforms and Applications Handbook*, 2nd ed., edited by A. D. Poularikas (CRC Press, Boca Raton, FL, 2000), Chap. 8.
[22]Z. Gong, M. Andrews, S. Jagannathan, R. Patel, J. M. Jech, N. C. Makris, and P. Ratilal, "Low-frequency target strength and abundance of shoaling Atlantic herring (*Clupea harengus*) in the Gulf of Maine during the Ocean Acoustic Waveguide Remote Sensing 2006 Experiment," J. Acoust. Soc. Am. **127**, 104–123 (2010).
[23]N. C. Makris, P. Ratilal, S. Jagannathan, Z. Gong, M. Andrews, I. Bertsatos, O. R. Gødo, R. W. Nero, and J. M. Jech, "Critical population density triggers rapid formation of vast oceanic fish shoals," Science **323**, 1734–1737 (2009).
[24]Z. Gong, A. D. Jain, D. D. Tran, D. H. Yi, W. Fan, A. Zorn, P. Ratilal, and N. C. Makris, "Ecosystem scale acoustic sensing reveals humpback whale behavior synchronous with herring spawning processes and re-evaluation finds no effect of sonar on humpback song occurrence in the Gulf of Maine in fall 2006," PLoS One **9**, e104733 (2014).
[25]FURUNO Marine GPS Navigator Model GP-90 brochure, FURUNO Electric Co., Ltd., URL http://techserv.gso.uri.edu/DownLoad/FurunoGP90.pdf (Last viewed April 18, 2013).
[26]F. Ingenito, "Scattering from an object in a stratified medium," J. Acoust. Soc. Am. **82**, 2051–2059 (1987).
[27]W. M. Ewing, W. S. Jardetzky, and F. Press, *Elastic Waves in Layered Media* (McGraw-Hill, New York, 1957), pp. 341–367.
[28]J. Miklowitz, *The Theory of Elastic Waves and Waveguides* (North-Holland Publishing Company, Amsterdam, 1978), pp. 273–277.
[29]C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978), Chap. 6.
[30]A. Walther, *The Ray and Wave Theory of Lenses*, Vol. 15 (Cambridge University Press, Cambridge, United Kingdom, 1995), pp. 154–158.
[31]C. S. Clay and H. Medwin, *Acoustical Oceanography* (John Wiley & Sons, Inc., New York, 1977), pp. 312–313.