High-frequency (8–16 kHz) model-based source localization[a)]

Paul Hursky,[b)] Michael B. Porter, and Martin Siderius
Science Applications International Corporation, 10260 Campus Point Drive, San Diego, California 92121

Vincent K. McDonald
Space and Naval Warfare Systems Center, San Diego, 53560 Hull Street, San Diego, California 92152-5001

(Received 1 October 2003; revised 3 February 2004; accepted 3 February 2004)

Matched-field or model-based processing has now been widely demonstrated for improving source localization and detection in ocean waveguides. Most of the processing approaches become increasingly sensitive to fluctuations or uncertainties as the frequency increases. As a result, there has been very limited work above 1 kHz and there is a perception that above several kilohertz the technique cannot be applied. We have conducted acoustic communications experiments in a variety of shallow water sites around coastal areas of the United States. These experiments show that a clear multipath structure is readily observed even in the 8–16 kHz band. Furthermore, it is shown that model-based processing can then be exploited to localize sources at these high frequencies out to ranges of several kilometers. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1690078]

PACS numbers: 43.60.Kx, 43.30.Wi, 43.60.Uv, 43.30.Re [AIT] Pages: 3021–3032

I. INTRODUCTION

Over the last 20 years there has been a great deal of research on using or embedding acoustic models in signal processing algorithms. An example application is for vertical or horizontal line arrays in which the spatial and temporal multipath structure is used to determine the location of a source in the ocean waveguide.

The terminology for this work is not standardized. We use the term ''model-based'' processing to refer to any technique that uses a computer model of acoustic propagation in the ocean waveguide. This term encompasses:

(1) matched-field processing, which exploits the phase-amplitude structure of some small set of narrowband signals,[1–3]
(2) backpropagation or time-reversal techniques which use a computer model to propagate the field observed on the receive array and (under certain conditions) refocus it at the source location,[4–7]
(3) correlation processing, which exploits the temporal multipath structure.[8–21]

These techniques are all closely related and in some cases actually identical. However, they suggest different ways of organizing the processing and sometimes lead to different insights about how to exploit the space–time structure of the acoustic field. We will present source localization results obtained using correlation techniques. As indicated in the above mentioned citations, these techniques have been addressed in the literature at least as far back as 1971. The main thrust of this work is to extend the correlation-based techniques to significantly higher frequencies.

The motivation for working at higher frequencies includes:

(1) high spatial resolution can be achieved using small apertures,
(2) sources such as dolphins or AUVs have a signature in this band,
(3) the ambient noise background is significantly lower at high frequencies than in lower bands (where the clutter is dominated by surface shipping),
(4) understanding the propagation physics in this band will lead to insights into the performance of acoustic communications systems, and suggest better wave form and receiver designs.

Previously, source localization has been demonstrated in a midfrequency band,[22] using a 22-element vertical line array in very shallow water (<10 m) to process four tones at 2.4, 3.5, 4.6, and 5.7 kHz out to 200 m in range. In other notable work at midfrequency,[23] 15 tones in the 3–4 kHz band from an 8-element vertical line array at a range of 1.5 km were processed using an approach that also optimizes the parameters characterizing the environment, such as ocean depth, bottom properties, and sound speed profile.

The first issue that arises is that of understanding qualitatively the propagation physics in this band. Should we expect distinct echoes from the surface and bottom? A variety of phenomena might conspire to produce a diffuse smear of acoustic energy, providing little structure to be exploited by a source location estimator:

(1) surface and bottom roughness,
(2) small-scale ocean variability,
(3) source/receiver motion,
(4) near-surface bubbles.

These phenomena have been addressed in the literature,[24] although most of the studies have been devoted to single boundary interactions and/or the backscattered field, and thus

[a)]Portions of this work were presented in ''High-frequency broadband matched field processing in the 8–16 kHz band,'' Proceedings of the IEEE Oceans 2003 Marine Technology and Ocean Science Conference, San Diego, CA, 22–26 September 2003.
[b)]Electronic mail: paul.hursky@saic.com

provide little guidance as to what sort of multipath signature might be observed, or whether it exhibits a reliable enough structure to be the basis for multipath ranging. As we will show, our experiments at a variety of typical shallow water sites reveal *a clear set of surface and bottom echoes rising well above the reverberant haze*.

The second issue is whether we can predict the field accurately enough to localize a source using the echo pattern as a fingerprint of target location. Model-based source localization at high frequency either requires very accurate modeling or must be made inherently robust against model mismatch. Our results demonstrate that *source localization at high frequencies is possible out to ranges of at least several kilometers* with even a minimal receiver configuration (a single phone in Sec. II, and a pair of phones in Sec. III). However, special techniques must be applied to exploit the reliable features of the propagation.

A series of experiments, part of ONR's SignalEx program,[25] have been conducted in a variety of shallow water coastal environments to relate the performance of acoustic communications systems to different propagation and oceanographic phenomena. These experiments consisted of using a fixed receiver to record wave forms from a transmitter allowed to drift out to ranges beyond which the signal was no longer detectable (on a single hydrophone). Typically, channel probe wave forms were alternated with wave forms to test specific communication modulation and coding schemes. The probes consisted of 50 ms LFM chirps, sweeping from 8 to 16 kHz, repeated four times per second. The probe pulses have enabled the channel impulse response to be measured in a variety of acoustic waveguides. We will present the results of applying our model-based source localization (correlation-based multipath ranging) to the probe signals.

Applying a matched filter to the probe pulses (i.e., using the known LFM chirp wave form as the correlation kernel) produces measurements of the channel impulse response at the pulse repetition rate. These measurements reveal the fine-scale time variations of the individual multipath components of the channel impulse response. Using a ray-based propagation model (which tells us which eigenray interacted with the surface, for example), enables us to isolate the physical mechanisms causing these multipath fluctuations that have such a major-impact on acoustic communications, environmental inversion, and source localization. As we will show, even after stabilizing the first couple of multipath arrivals from ping to ping, the later arrivals exhibit fluctuating amplitudes and times of arrival, due to the motion of the ocean surface, water column variability, and the varying bathymetry as the transmitter drifts in range. Besides measuring the fine scale temporal structure over 10 and 30 s intervals (at the two sites which we will discuss), having the source drift out to ranges until the probe signals were no longer detectable at the receiver (typically at 6–8 km) revealed how the channel impulse response varied as a function of range.

This paper presents results using the 8–16 kHz channel probes from sites at the New England Front and off the coast of La Jolla in San Diego, CA. Section II will show data from a SignalEx test on the New England Shelf in 2000. At this site, we will present source localization results based upon matching measured and modeled impulse response functions. The impulse response is measured by applying a matched filter, using the source wave form as the correlation kernel. Obviously, in many applications, the source would not be so cooperative as to let its wave form be known. We have previously[21] demonstrated how to extend the impulse response method to auto- and cross-correlation wave forms. Section III will show data from a SignalEx test off the coast of La Jolla in San Diego in 2002. At this site, we will present results using impulse response functions, and also cross-correlation wave forms. Matching cross-correlation wave forms does not require the source signature to be known, only that it be sufficiently broadband that its cross-correlation in the time-domain produces a narrow enough pulse to resolve multipath arrivals. Comparisons will be made between measured and modeled channel impulse response functions at both sites.

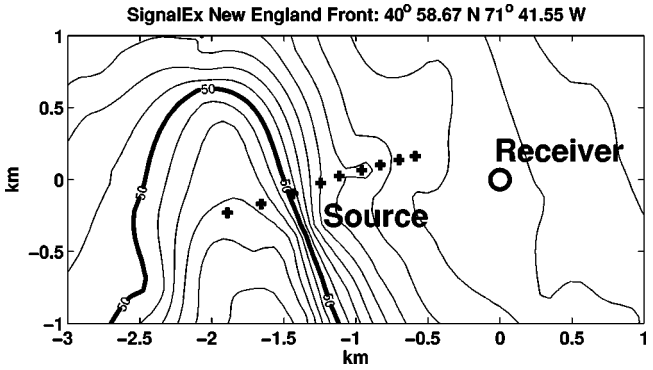

FIG. 1. SignalEx Front 2000 experiment configuration. Contour lines are spaced at 1 m intervals, with depth increasing moving west from the receiver to the heavy contour line at 50 m.

II. NEW ENGLAND SHELF RESULTS

Figure 1 shows the bathymetry on the New England Shelf where the SignalEx Front 2000 experiment was performed. The bathymetry between the receiver (indicated by the circle) and the drifting source (whose track is marked by plus marks) is mildly sloped, at least for the first several kilometers. A single radial from the receiver was used to set the bathymetry that was used to model the propagation at this site.

Figure 2 shows the measured sound speed profile and the depths of the source (29 m) and receiver (40 m). The sound speed profile is upward refracting and contains a surface duct, so there are fewer interactions with the bottom than in the La Jolla environment we will look at in Sec. III. The bottom sound speed is greater and the ocean is shallower at the New England site than in La Jolla. These effects combine to produce more multipath arrivals at the New England site than at the La Jolla site. However, the surface was rougher at the New England site, which caused the surface-interacting paths to exhibit arrival times that were not as well defined as at the La Jolla site.

All of the model-based techniques rely upon boundary interactions in the waveguide to produce multiple virtual images of the source (or receiver), which form an effective

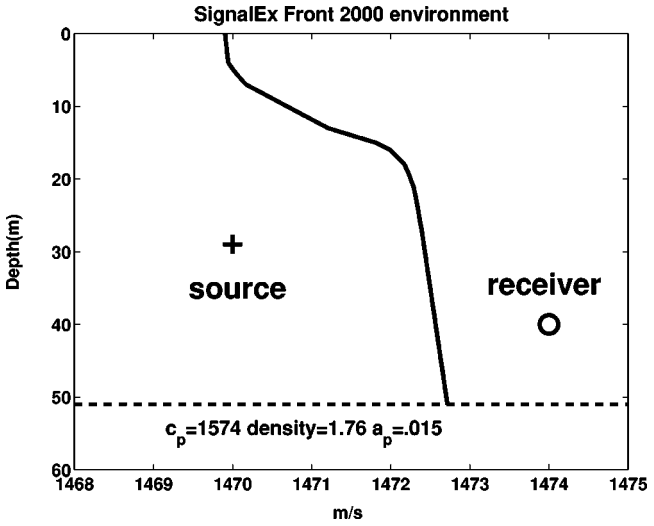

FIG. 2. SignalEx Front 2000 measured sound speed profile and presumed bottom properties, showing depths of source and receiver relative to the profile.

aperture much larger than the actual physical aperture, making it possible to estimate ranges and depths with arrays much smaller than would be needed in free space conditions.[26] As we will see, this makes source localization easier at the New England site.

Figure 3 shows two views of the wave form sequence during the SignalEx test on the New England Shelf in April of 2000. The upper plot in Fig. 3 covers 30 min, and shows different colored rectangles. The gray rectangles labeled ''P'' represent probe sequences. The rectangles labeled ''A''–''F'' represent acoustic communications test sequences. The six different test sequences are preceded by identical probe sequences. Each probe/test sequence pair occupies 5 min. The lower plot in Fig. 3 is a 10 s excerpt from the probes and shows repeating probe wave forms. Each probe interval (the ''P'' blocks) contains 40 LFM chirps. Each chirp sweeps from 8 to 16 kHz in 50 ms. The chirps are repeated every 250 ms (40 chirps in 10 s).

A. Channel impulse response measurements

The upper part of Fig. 4 shows 40 processed chirps (each row of the image is a matched filter output), from the

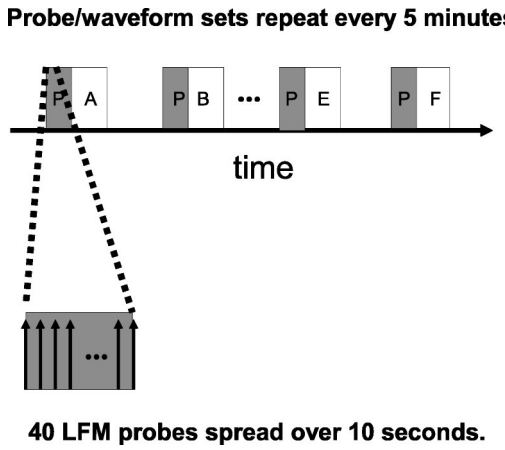

FIG. 3. Timeline of LFM probe signals (labeled P) and acoustic communications wave forms (labeled A–F) during SignalEx 2002.

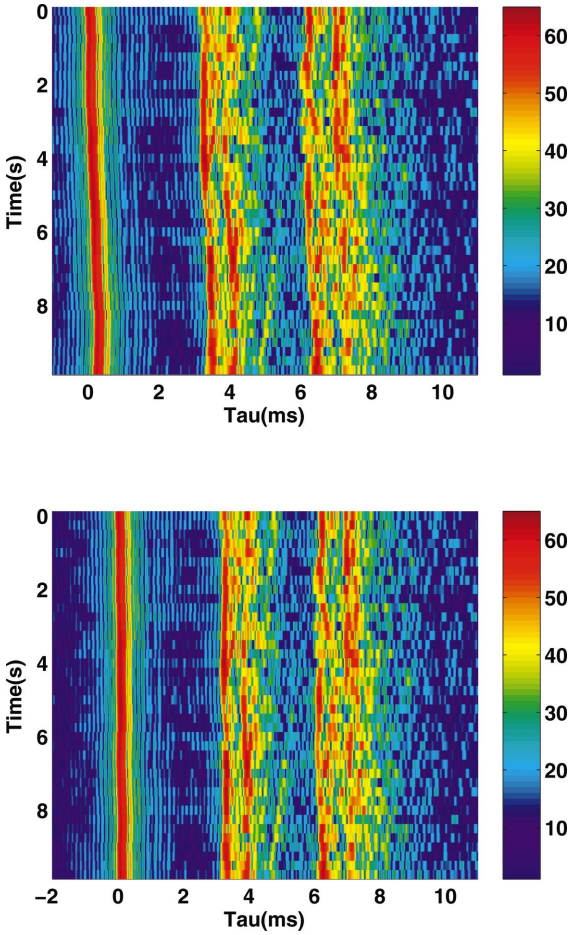

FIG. 4. The upper plot shows stacked impulse responses, aligned according to a constant Doppler correction. The lower plot shows stacked impulse responses, aligned by cross-correlating consecutive pairs of responses. Each row of these images contains the log-envelope (in decibels) of the matched filter output corresponding to a single LFM chirp. There were 40 8–16 kHz LFM chirps, each 50 ms long, transmitted every 250 ms. Only the first 10 ms of the measured channel impulse response is shown to isolate the details of the earliest arrivals (at this 500 m range, the entire impulse response lasted 50 ms).

SignalEx test on the New England Shelf, stacked one on top of the other. The rows of this image have been aligned by spacing them according to the known pulse repetition interval of 250 ms, corrected for a constant Doppler. Each row of this image contains the envelope of the matched filter output on a decibel scale, calculated using the known probe waveform as the matched filter replica.

The lower part of Fig. 4 shows the same 40 chirps, aligned by correlating each row with its predecessor. That is, each row is offset relative to the previous row so that the peak of their cross-correlation is at the zeroth lag. This method of aligning one matched filter output with respect to its predecessor is only one of many techniques we attempted, including peak picking. Using cross-correlation to align these wave forms turned out to be the most robust for this and other data sets. Because the cross-correlation is driven by the higher amplitude earlier arrivals (at least in this data set), the fluctuations are all but removed from these earlier arrivals by this process. At the same time, an alignment based on cross-correlation enables all arrivals to contribute to the alignment (and so is more robust than simply aligning on

the basis of the first peak location). However, the higher amplitude earliest arrival will clearly have the most influence. In other configurations (i.e., a different geometry, or a different propagation environment), later arrivals could very well dominate. Aligning the probes enables the structure in the later arrivals to be clearly seen, independent of the fluctuations in the early arrivals.

Although the entire impulse response has a duration of 50 ms at the range shown (500 m), we have displayed only 12 ms to isolate the first few arrivals, so that the fine structure in these arrivals can be observed. After the first arrival, two pairs of arrivals can be made out, but these later arrivals are not as stable as the first arrival, exhibiting some spread in the time of arrival and quite a bit of amplitude fading. Note that it is difficult to pick out a distinct track in any of these later arrivals. Later arrivals (not shown here) exhibited even worse fading. We speculate that this is due to the rougher surface at this site.

B. Modeling

In the previous section, Fig. 4 shows a single probes interval (10 s long, with 40 probes, each probe an LFM chirp). Rather than attempting to estimate source location from a single impulse response measurement (rarely are all the arrivals present in a single measurement, because of the amplitude fading), we averaged all 40 probe response envelopes to form a single composite impulse response estimate. Clearly, the alignment described in the previous section (based on cross-correlating consecutive impulse response measurements) is critical for this averaging to be effective. The upper part of Fig. 5 shows the result of stacking 26 such averages (roughly 2 h of the drift, out to 3 km range). Each row of the image shown in the upper part of Fig. 5 contains the average of one 10 s, 40-chirp time interval. Each of the 10 s intervals was Doppler corrected (as in the upper part of Fig. 4), aligned (chirp-to-chirp, as in the lower part of Fig. 4), and summed to form a single average impulse response function estimate. The signal-to-noise ratio (at the output of the matched filter and after averaging 40 chirps having a duty cycle of 20 percent) can be read from the upper part of Fig. 5 (the color scale is in dB). The earlier arrivals have a signal-to-noise ratio (SNR) of roughly 15 dB and the later arrivals (somewhat smeared by the averaging) have a SNR of roughly 5 dB. The averaging serves to stabilize the channel impulse response spreading and amplitude fading, but also smears out the later arrivals, whose time of arrival is not as consistent as the earliest arrival. These are the measurements that we must duplicate with a propagation model to form the source location estimate.

To reproduce the range-dependent impulse response function shown in the upper part of Fig. 5, the broadband channel impulse response function was modeled using the Bell-hop ray/beam tracing program.[27–29] This model calculates magnitudes, phases (note that envelopes are shown in the plots of modeling calculations), and times of travel of all multipath components for a particular source and receiver geometry, given a sound speed profile, geo-acoustic properties of the surface and bottom, and a potentially range-dependent bathymetry. No roughness was incorporated in the

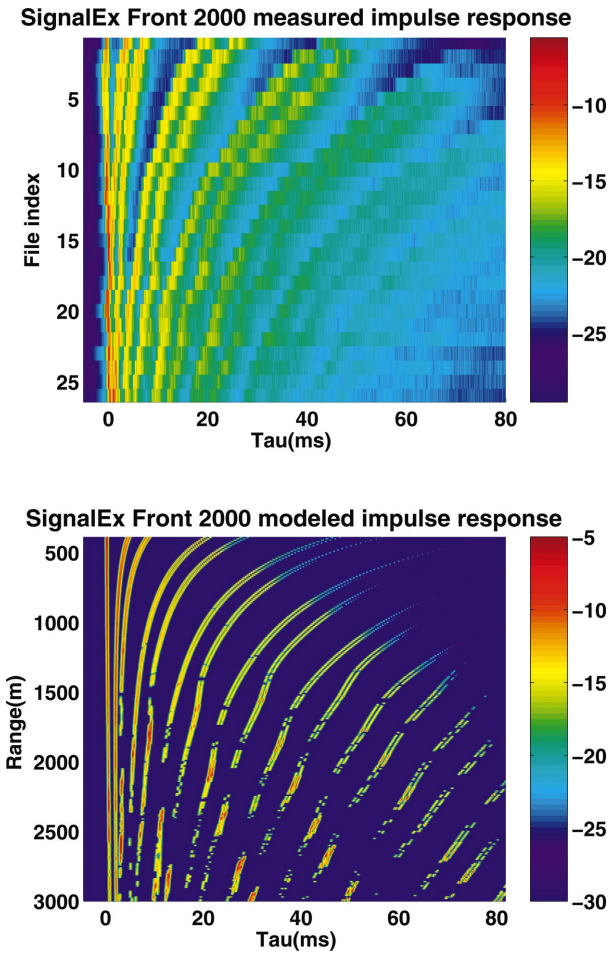

FIG. 5. Measured (upper) and modeled (lower) channel impulse response functions, as a function of range (receiver depth of 40 m and source depth of 29 m). Each scan line in the measured data is an average over 40 chirps spanning 10 s.

modeling, except for the coarse features specified by a bathymetry map. A band-limited impulse response function is synthesized from these multipath arrival parameters.

The lower part of Fig. 5 shows the multipath structure calculated by Bellhop for the experiment configuration during the New England Front SignalEx test (relative time of arrival is shown along the horizontal axis, and the ranges shown along the vertical axis are the ranges at which the data shown in Fig. 5 were measured, as calculated from GPS measurements). The agreement between the coarse features of the measured and modeled data shown in Fig. 5 is excellent, which bodes well for our model-based source localization. However, the later arrivals in the measured data, whose arrival times exhibit the fluctuations seen in the upper part of Fig. 4, have been smeared out by the averaging process, grossly underestimating the amplitudes of the later arrivals, compared to the amplitudes predicted by the model for the later arrivals in the lower part of Fig. 5. We will show how to reduce the impact of this mismatch so that source localization is possible even with significant fading.

The measured impulse response functions shown in Fig. 5 were aligned by cross-correlating consecutive rows, just as in Fig. 4 in the previous section. That is why the first arrivals do not all line up on a vertical line (the other arrivals con-

tribute and cause the alignment to move around a bit). The measured response functions are displayed using a detected early arrival to set the left edge of the first image row, and the peak cross-correlation (row to row) to set subsequent rows (as discussed earlier). The modeled results are displayed using a reduced time (range/sound speed) to set the left edge of each image row. As we will discuss in the next section, these time offsets (of one row with respect to the next, and between measured and modeled wave forms) do not impact the source location estimate.

We have used normal mode and parabolic equation models at lower frequencies to synthesize the time-domain impulse response functions (by transforming the spectrum calculated by runs of these full-wave models at each frequency), but a ray-based model provided adequate fidelity in this high frequency band, and was more convenient because it enabled direct manipulation of the time domain features of the impulse response function. Normal mode and parabolic equation models become computationally cumbersome at higher frequency.

C. Localization

As mentioned earlier, we are not introducing a novel source location estimator, but are extending correlation-based multipath ranging algorithms[8–21] to much higher frequencies, with some minor tricks to overcome difficulties peculiar to this band.

The source location metric $b(r,z)$ at source range r and depth z,

$$b(r,z)=\max_{n}\sum_{i=0}^{N-1} d_{i+n}m_i(r,z), \tag{1}$$

is calculated by cross-correlating d_{i+n}, the measured impulse response envelope, and $m_i(r,z)$, the modeled impulse response envelope, and then taking the maximum cross-correlation peak. Both d_{i+n} and m_i are wave form values, sampled at times i (or $i+n$, with n a correlation lag). The uper part of Fig. 5 shows the d_i values we will use (i corresponds to multipath time of arrival along the horizontal axis), with each row containing the impulse response function measured from a single 40-chirp interval (at a particular source range, to be estimated). The lower part of Fig. 5 shows the analogous m_i values as a function of time of arrival (along the horizontal) and source range (along the vertical). Because there is no time reference for the probe arrivals, we must match measured and modeled wave forms at all possible offsets of one with respect to the other, by cross-correlating them, as opposed to simply forming an inner product (which would have been possible, if we had a time reference). The metric used to match measured and modeled wave forms, $b(r,z)$ in Eq. (1), is taken to be the maximum value of their correlation (i.e., at whatever lag it occurs). The lag n at which this maximum value occurs does not enter into our source location estimate (because we can only measure relative times of arrival, and not the actual travel times).

Combining the multipath arrivals predicted by the ray-based model coherently, using the true spectrum, produces very short duration arrival pulses (1/bandwidth, where our chirps had a bandwidth of 8 kHz), and with complicated interference between overlapping arrivals. It is significantly more difficult to model the phases of the multipath components than the envelopes and times of arrival, even at lower frequencies.[21] Therefore, for this high frequency band, we chose to operate on the envelopes of the measured and modeled matched filter outputs, rather than the raw wave forms. Figure 5 shows envelopes (displayed using a decibel scale). Furthermore, it was very difficult to model the arrival times accurately to within the time resolution (1/8000 s) provided by our signal bandwidth of 8 kHz. In order to desensitize our modeling, we artificially reduced the bandwidth of the modeled signal (to broaden the multipath arrivals in the time domain) and combined the multiple arrivals incoherently (to avoid the increased opportunities for interference between arrivals that was a result of the reduced bandwidth).

Using the above-outlined process, a source location metric was calculated for every candidate source range r and depth z, at every time epoch for which we have measured the impulse response (as the source drifts in range). A two-dimensional ambiguity surface was produced for each time epoch, so that the overall output for the entire source drift was a three-dimensional ambiguity volume, indexed on source range, source depth, and time epoch.

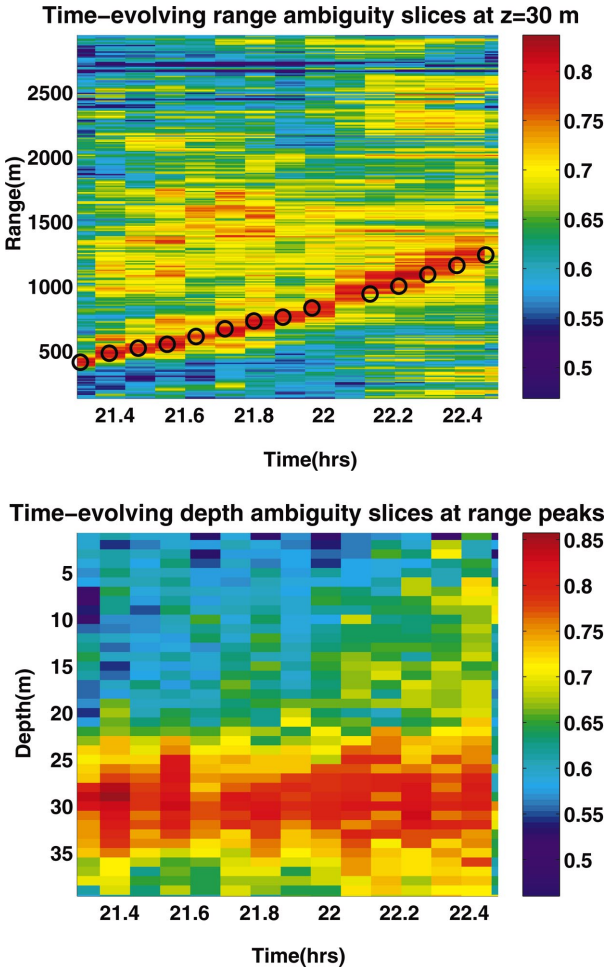

FIG. 6. Upper plot shows range track at source depth of 29 m (and receiver depth of 40 m). Black circles indicate ranges calculated from GPS measurements. Lower plot shows depth track along estimated range track.

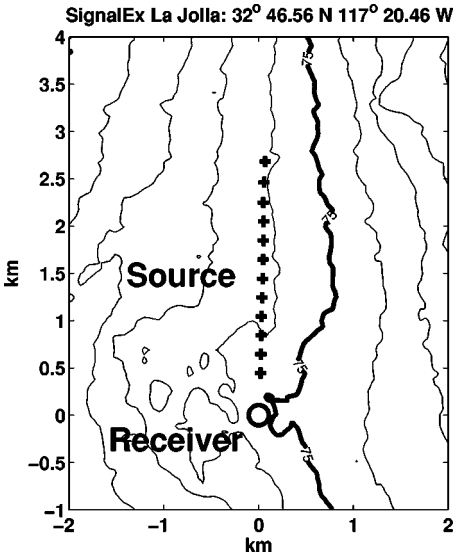

FIG. 7. SignalEx La Jolla 2002 experiment configuration. Contour lines are spaced at 5 m intervals, with depth increasing moving west. The heavy contour line indicates a depth of 75 m.

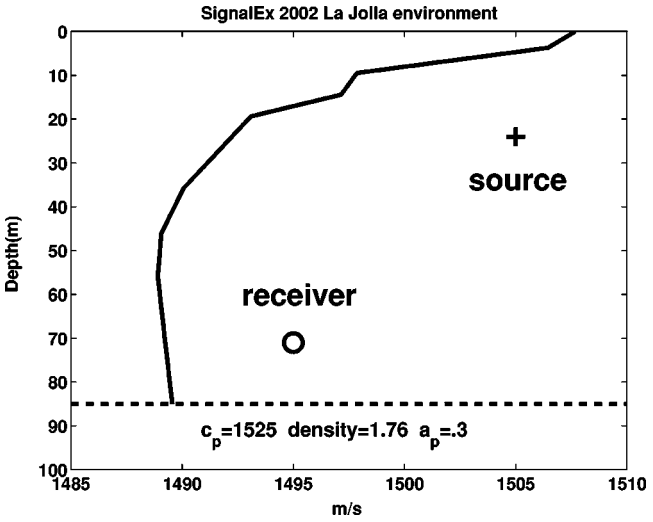

FIG. 8. SignalEx La Jolla 2002 measured sound speed profile and presumed bottom properties, showing depths of source and receiver relative to the profile.

The upper part of Fig. 6 shows the 2D slice versus range and time for the known source depth of 29 m. The circles indicate the known source range, calculated from GPS measurements at each time epoch. The range track is consistent with the GPS measurements. The lower part of Fig. 6 shows the slices versus depth that follow the source track in range. A very strong track is apparent at the known source depth of 29 m. A persistent track is apparent in both range and depth.

Note that the slice through the 3D ambiguity volume shown in Fig. 6 was selected (from among all the possible source depths) knowing the source depth. A way to independently determine the source depth from the 3D ambiguity volume would be to search for a continuous track among all depths, presuming the source depth was constant. We have scanned range-time slices (such as the one shown in Fig. 6) at all depths, and found that the most persistent track is found at the correct depth. Unfortunately, the individual 2D ambiguity surfaces in range and depth at each time epoch have so many spurious peaks that a 2D peak-picking process repeated at each time epoch does not produce a consistent track in range and depth. It is only when we seek a persistent track over many epochs that the source location reveals itself.

Despite these shortcomings, the results shown in Fig. 6 are quite surprising, given the high frequency band. Note that to get a track beyond the 500 m starting range required us to:

(1) average multiple measurements of the impulse response,
(2) operate on the envelopes of the data and the modeled wave forms (disregarding the phase),
(3) artificially increase the bandwidth of our modeled multipath arrivals to broaden them in the time domain, to reduce the sensitivity to model mismatch at these high frequencies.

III. LA JOLLA 2002 RESULTS

Figure 7 shows the bathymetry and locations of the receiver (circle) and transmitter (plus marks) at the La Jolla site, 5 km off the coast of San Diego in California. The receiver was moored to the bottom and suspended at a depth of 71 m. The transmitter was suspended at a depth of 24 m from a boat that drifted away from the receiver out to a range of roughly 7 km (only the first part of the track is shown in Fig. 7). The transmitter track follows an isobath at roughly 80 m (the heavy contour line to the right of the transmitter track is at a depth of 75 m).

Figure 8 shows the sound speed profile measured at this site and the receiver and transmitter depths relative to this profile. Note that compared with the profile at the New England site (see Fig. 2), where there is a surface duct keeping sound away from the bottom and favoring the surface, the La Jolla site has a downward refracting profile that favors interactions with the bottom. The bottom properties here, the slower compressional wave speed and greater attenuation, result in a less reflective bottom. As a result, we will see that there are fewer arrivals at the La Jolla site than at the New

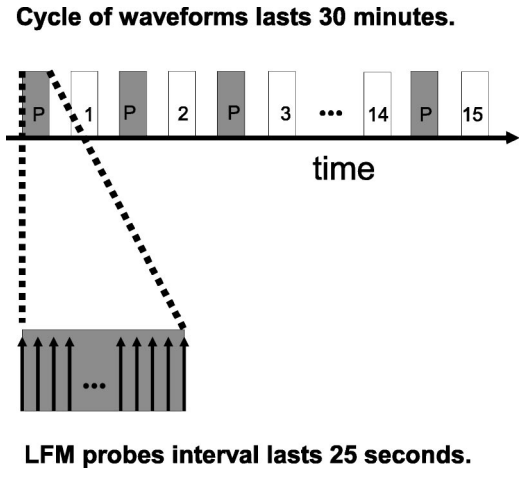

FIG. 9. Timeline of LFM probe signals (labeled P) and acoustic communications wave forms (labeled 1–15) during SignalEx 2002 in La Jolla.

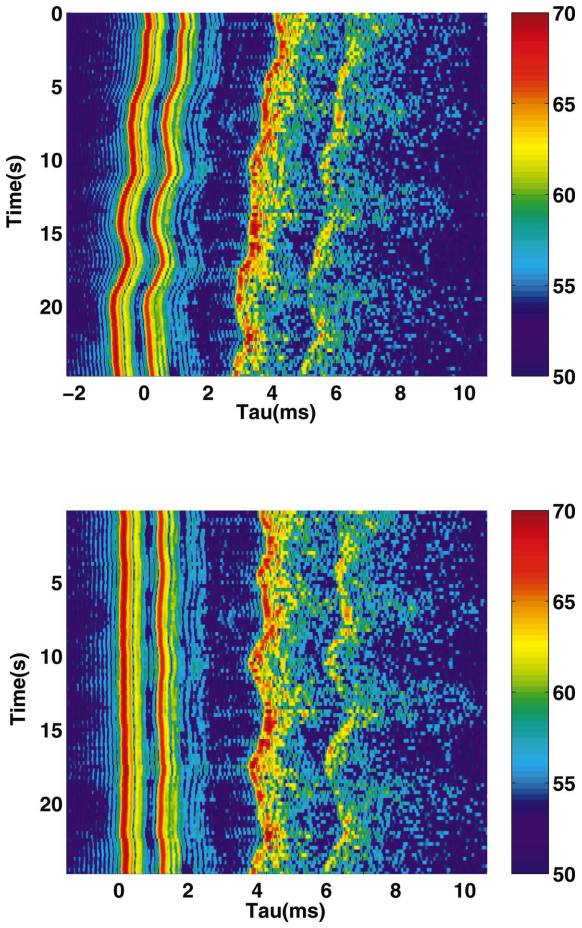

FIG. 10. The upper plot shows stacked impulse responses, aligned according to a constant Doppler correction. The lower plot shows stacked impulse responses, aligned by cross-correlating consecutive pairs of responses. Each row of these images contains the log-envelope (in decibels) of the matched filter output corresponding to a single LFM chirp. There were 100 8–16 kHz LFM chirps, each 50 ms long, transmitted every 250 ms. Only the first 10 ms of the measured channel impulse response is shown to isolate the details of the earliest arrivals (at this range, the entire impulse response lasted 60 ms).

England site. The surface was very calm, less rough than at the New England site.

Figure 9 is similar in format to Fig. 3 in Sec. II. In the 2002 SignalEx test at the La Jolla site, there were 15 acoustic communications wave forms being tested, each allotted 1 min of transmit time (the transparent boxes, labeled from 1 to 15, in the upper timeline). Each such test wave form was preceded by a probes interval (the gray boxes, labeled "P"), 1 min long, which contained 25 s of LFM chirps. Each chirp was 50 ms long, sweeping up from 8 to 16 kHz. These chirps were transmitted at a rate of 4 per s (i.e., every 250 ms). Thus, each 25 s probes interval produced 100 transmitted LFM chirps.

A. Impulse response measurements

Figure 10 is analogous to Fig. 4 in Sec. II A. The upper part of Fig. 10 shows 100 processed chirps (matched filter outputs), stacked one on top of the other, from the La Jolla 2002 site at a range of 450 m. There are significant fluctuations in all arrivals. The lower part of Fig. 10 shows the same 100 chirps, aligned by cross-correlating each row with its predecessor as described in Sec. II A. Here too, the later arrivals seem to be driven by a process that is independent of the process governing the earliest arrivals.

Note that, compared to the New England site, each arrival at the La Jolla site (in Fig. 10) is concentrated along a single well-defined track (in correlation lag time). We also see some reverberation following the third and fourth arrivals. In impulse response measurements made at the New England site (see Fig. 4), the multipath arrivals following the earliest arrival were broken up to such an extent that it was difficult to identify distinct tracks (in correlation lag time). Nevertheless, the arrivals following the first two arrivals at the La Jolla site have perhaps as much amplitude fading as those at the New England site.

Given the La Jolla configuration with the receiver close to the bottom and the source in the water column (see Fig. 8), and having modeled this configuration with a ray trace model, we can identify the first and second arrivals as direct and bottom-reflected paths, and the third and fourth arrivals as surface interacting paths. Although the fluctuations seen in the third and fourth arrivals (which appear to be strongly correlated) could be due to water column phenomena, they are probably due to the motion of the surface. The duration of the entire impulse response at this range was roughly 60 ms, but we are only showing the first 12 ms of this, so that the fine structure in the earlier arrivals can be observed. Later arrivals (not shown in this 12 ms excerpt) had similar fluctuations in (relative) arrival times and amplitudes.

B. Modeling

Figure 11 is analogous to Fig. 5 in Sec. II B. Each row in the upper part of Fig. 11 is the result of averaging the aligned matched filter outputs corresponding to 100 chirps (from a 25 s probes interval), as described in Sec. II B. The signal-to-noise ratio (at the output of the matched filter and after averaging 100 chirps having a duty cycle of 20%) can be read from Fig. 11 (the color scale is in decibels). The earlier arrivals have a SNR of roughly 15 dB and the later arrivals (somewhat smeared by the averaging) have a SNR of roughly 5 dB. These SNRs are at the output of the matched filter. The lower part of Fig. 11 shows the modeled impulse response envelopes as a function of range (the color scale is in decibels). The ranges along the vertical axis are the ranges at which the data shown in the upper part of Fig. 11 were measured, according to GPS measurements. The modeling has already been described in Sec. II B: the Bellhop ray/beam model was used along a radial (from the receiver) with a range-dependent bathymetry.

The dropouts seen along some of the later arrivals in the modeled data are due to the range-dependent bathymetry (they disappeared when a flat bottom was modeled). These were duplicated by a broadband parabolic equation calculation, run as a check on the ray tracing results. Because the measured data are the result of averaging over 25 s of drift, these dropouts are not observed in the measured data.

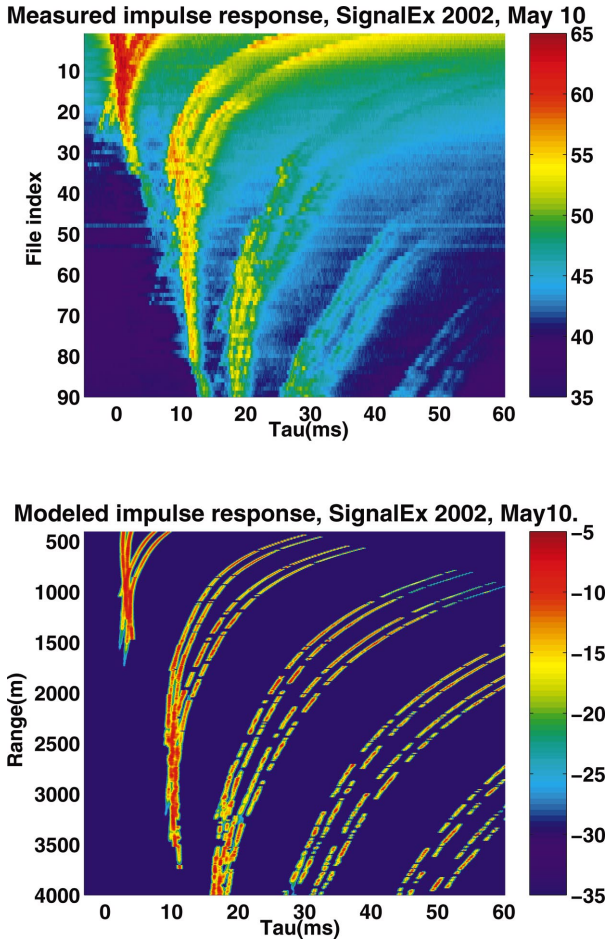

FIG. 11. Measured (upper) and modeled (lower) channel impulse response functions, as a function of range (receiver depth of 71 m and source depth of 24 m). Each scan line in the measured data is an average over 100 chirps spanning 25 s.

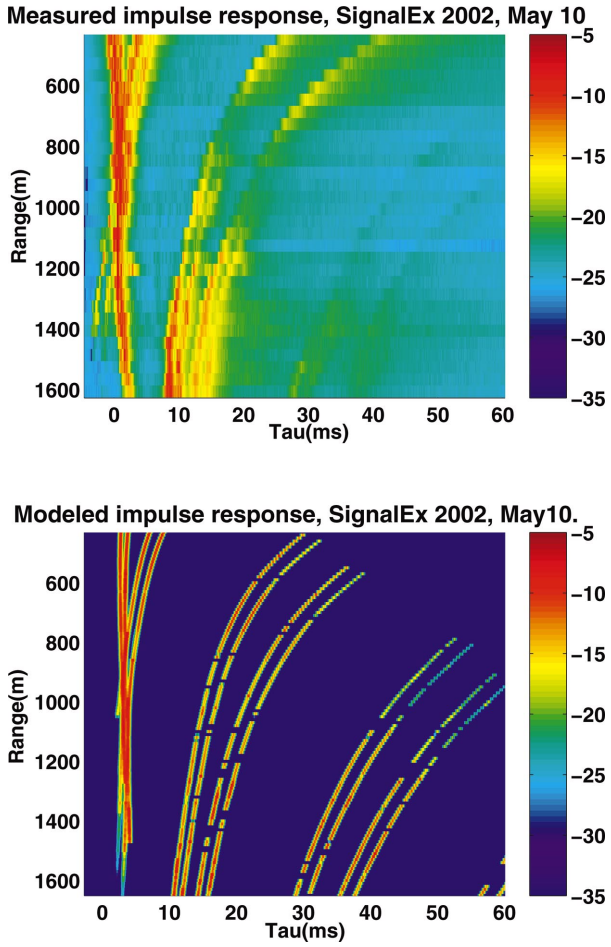

FIG. 12. Excerpts of measured (upper plot) and modeled (lower plot) impulse responses over range interval being processed.

C. Localization (source wave form known—matching matched filter outputs)

This section is analogous to Sec. II C: we present the results of matching measured and modeled impulse response functions (shown in the previous two sections) to estimate source location. However, using data from the La Jolla site, it was only possible to locate the source out to a range of 900 m using the techniques described in Sec. II C. We were able to obtain source location estimates beyond this range only after applying several transforms to the wave forms being matched. These transforms served to appropriately weight the features that served to distinguish different source ranges and depths.

When the technique presented in Sec. II C for the New England data was applied to the La Jolla data, plausible source location peaks were produced for only the first few (short) ranges. There were several reasons for this. Looking at the measured and modeled data, the mismatch in the higher amplitude earlier arrivals was dominating the information provided by the later arrivals. This was further compounded by the later arrivals being smeared out by our averaging process, due to the fluctuations in their time of travel (see Fig. 10), causing their amplitudes to be grossly overestimated by the modeling. There were also far fewer arrivals at this site (because of the downward refracting profile, a deeper ocean, and a softer bottom).

Figure 12 shows blowups of the measured and modeled data (seen in Fig. 11), showing what happens to the impulse response over ranges from 400 to 1600 m. Our initial attempt at source tracking failed at ranges around 1000 m. From 1000 to 1400 m, the pair of earliest arrivals is not predicted by the ray model. From 600 to 1200 m, the later set of arrivals (at ranges from 600 to 1100 m, between 10 and 20 ms in Fig. 12) show significant fading that is not predicted by the ray model. These differences between the measured and modeled data can be expected to cause problems for any source localization based on matching this measured and modeled data.

In previous work,[30] the log-envelope of impulse response data served to emphasize the contribution of later arrivals, which otherwise had much lower amplitudes than the earlier arrivals. Following this example, we tested several transformations of the measured and modeled impulse response wave forms to emphasize later arrivals, and as it turned out, to reduce the impact of fluctuations in the earlier arrivals.

The measured wave form was whitened using a three-pass, split-window moving average process to estimate both the mean and the standard deviation at each sample (many similar whiteners have been previously described[31]). The ini-

tial passes are used to form preliminary estimates of the mean and standard deviation (at each point in the data). These estimates are potentially biased by the presence of strong peaks. A "shearing" threshold is set based on these preliminary estimates (at a selected number of standard deviations above the mean), and any peaks exceeding this threshold are replaced by the current mean estimate (at that point). Once the peaks have been truncated, the moving average process is repeated on data that should no longer be corrupted by the presence of strong signals (an ideal noise measurement would be based on noise alone). After three passes of this process, estimated means and standard deviations have typically become reliable, and the data are whitened using estimates of its (noise) mean $\overline{x_i}$ and standard deviation σ_{xi}:

$$w_i = \frac{x_i - \overline{x_i}}{\sigma_{xi}}. \tag{2}$$

Low amplitude peaks surrounded by low power noise are transformed to values comparable to high amplitude peaks surrounded by high power noise. Since the noise typically is higher around the earlier peaks, this serves to emphasize the later arrivals and de-emphasize the earlier arrivals.

The above-described whitening process (applied to the measured data) was not appropriate for the modeled wave forms, since they are noiseless. Instead, the modeled wave form was raised to a fractional power (0.1) in order to reduce the disparity between the early and late arrival amplitudes (a transform that produces similar results to the log-envelope). In addition, the transformed modeled wave forms were forced to be zero-mean by subtracting their average value.

These somewhat *ad hoc* transforms, in addition to the steps described in Sec. II C for the New England data, resulted in the much-improved results shown in Fig. 13. The black circles in the upper plot indicate the known source ranges (as measured using GPS instruments). The lower part of Fig. 13 shows the slices versus depth intersecting the estimated ranges in the upper part of Fig. 13. In both the range and depth track plots, one-dimensional (1D) slices from a sequence of 3D volumes were stacked to form images. Because the dynamic range was not consistent across these 1D slices, each slice was rescaled to have a unit norm.

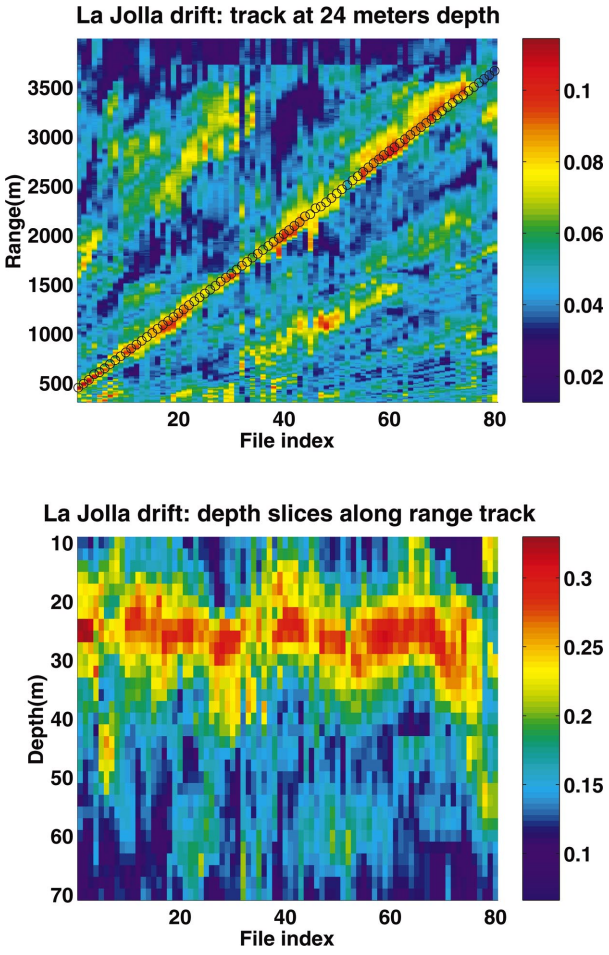

FIG. 13. Upper plot shows range track at source depth of 24 m (and receiver depth of 71 m). Black circles indicate ranges calculated from GPS measurements. Lower plot shows depth track along estimated range track.

D. Localization (source wave form unknown—matching cross-correlation wave forms)

In the previous sections, the source location statistic or metric upon which the location estimate was based was calculated using the channel impulse response measured by a matched filter. These estimates presume that the source wave form is known. The more practical case occurs when we do not know the source wave form. Instead of operating on the measured impulse response, which can only be measured directly by knowing the source wave form, we operate on the auto- or cross-correlation of the impulse response function. Let $s(t)$ be the source wave form, and $S(\omega)$ its spectrum. Similarly, let $h(t)$ be the channel impulse response function, and $H(\omega)$ its spectrum. The received signal, $r(t)$, is the convolution of $s(t)$ and $h(t)$,

$$r(t) = h(t) \otimes s(t),$$

where $\otimes$ is the convolution operator. Equivalently, in the frequency domain

$$R(\omega) = H(\omega) S(\omega).$$

Applying a matched filter is the same as correlating $r(t)$ with $s(t)$, which can be written in the frequency domain as

$$X(\omega) = S^*(\omega) R(\omega) = H(\omega) |S(\omega)|^2. \tag{3}$$

The superscript asterisk indicates a complex conjugate. Note that the phase of $s(t)$ no longer appears in the expression. Similarly, when we have two receiving elements, with impulse response functions $h_1(t)$ and $h_2(t)$, so that the spectra at the two receiver elements are $H_1(\omega)S(\omega)$ and $H_2(\omega)S(\omega)$, cross-correlating the two wave forms results in

$$\begin{aligned} C_{12}(\omega) &= [H_1(\omega) S(\omega)]^* [H_2(\omega) S(\omega)] \\ &= H_1^*(\omega) H_2(\omega) |S(\omega)|^2, \end{aligned} \tag{4}$$

where again the phase of the source wave form is not a factor. In Eq. (3), we recreate $H(\omega)$ using an acoustic propagation model. In Eq. (4), we use the same acoustic propagation model to recreate $H_1^*(\omega)H_2(\omega)$.

The fact that source wave form phase does not appear in the correlation wave forms means our process will work re-

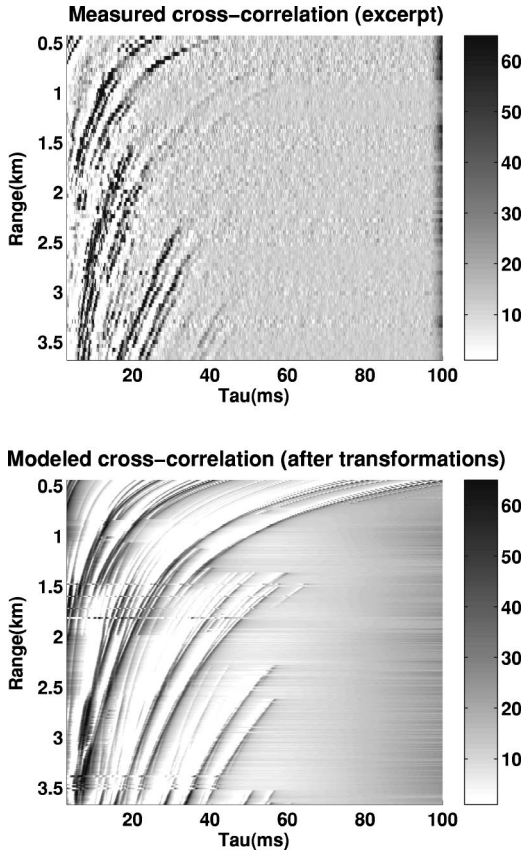

FIG. 14. Upper plot shows measured cross-correlation after whitening and excerpting. Lower plot shows the modeled cross-correlation after transformations and excerpting. An interval of 3–100 ms is shown.

gardless of whether the source wave form is a well-behaved wave form like a chirp, or whether it is closer to a completely random process. The source must have a wide enough spectrum so that its auto-correlation (the term $|S(\omega)|^2$) produces a narrow enough pulse in the time domain to resolve the multipath arrivals.

Note that using the cross-correlation envelope essentially removes the sensitivity to the center frequency—this means requirements for array element localization and source location search grid step size are set not by the high center frequency, but by the source bandwidth (which sets the pulse width in the time domain of the multipath arrivals). However, these benefits come at the cost of losing the gain otherwise available from coherently summing over frequency (if the source wave form phase is known, which for an uncooperative source is not likely anyway).

The impulse measurements do not have a time reference, so it is necessary to calculate the match at all possible relative time offsets between the measured and modeled wave forms (i.e. using a cross-correlation operation). The cross-correlation waveforms are functions of the time-difference-of-arrival only, so the measured and modeled wave forms are implicitly aligned and the match for each candidate source location is calculated using an inner product, which (in the time-domain at least) is more efficient than a cross-correlation.

The upper part of Fig. 14 shows the measured cross-correlation of two adjacent receive elements (arranged vertically and spaced 14 in. apart). The lower part of Fig. 14 shows their modeled cross-correlation. Both of these figures show wave forms after the transforms described in Sec. III C have been applied (prewhitening of the data wave form using a split-window three pass moving average process; raising the modeled wave form to the power 0.1, subtracting its mean, and re-scaling it to have unit norm).

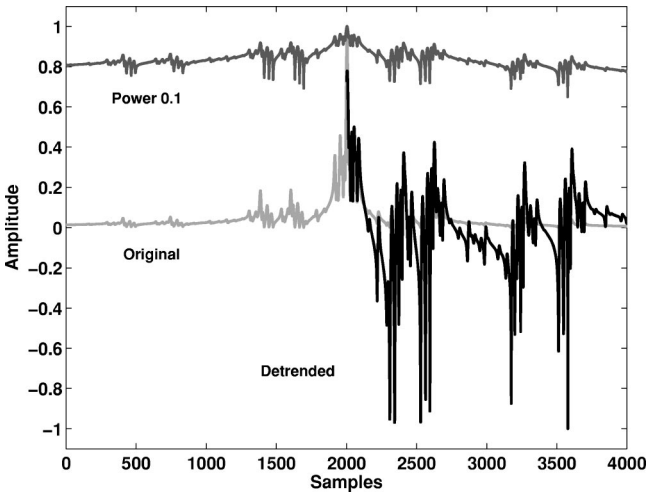

FIG. 15. This plot shows the transformations applied to the modeled cross-correlation wave forms. Shown are the original modeled wave form (light gray), this wave form raised to the power 0.1 (dark gray), and the result of subtracting the mean and rescaling so that the wave form has unit norm (black).

Figure 15 shows the effect of the transformations applied to the modeled wave form. The light gray curve is the original modeled cross-correlation wave form. The dark gray curve is the result of raising it to the power 0.1. The black curve shows the result of subtracting the overall mean and scaling the result so that it has unit norm. Only the right half of the black curve is shown—we only matched on the right half of the curve (other geometries would require matching over the entire wave form). In addition to these transformations, we omitted the first 3 ms of each correlation wave form from the source location metric calculation, because they were so unpredictable.

Figure 16 shows the range and depth tracks resulting from matching measured and modeled cross-correlations. The dynamic range is different in these two plots because the 1D slices being stacked were scaled to have unit norm.

IV. CONCLUSIONS

The most striking finding is that there seems to be a stable, exploitable impulse response of distinct and predictable multipath arrivals at these high frequencies. Although we only show results for two sites, we have seen qualitatively similar results at a number of sites where SignalEx experiments were performed.

The measured impulse response can be reproduced by standard acoustic propagation models well enough to support source localization using even a minimal receiver (we used either a single hydrophone, or a pair of hydrophones), although this was much more difficult than we have found at low frequencies.[21] At the New England site, we were able to

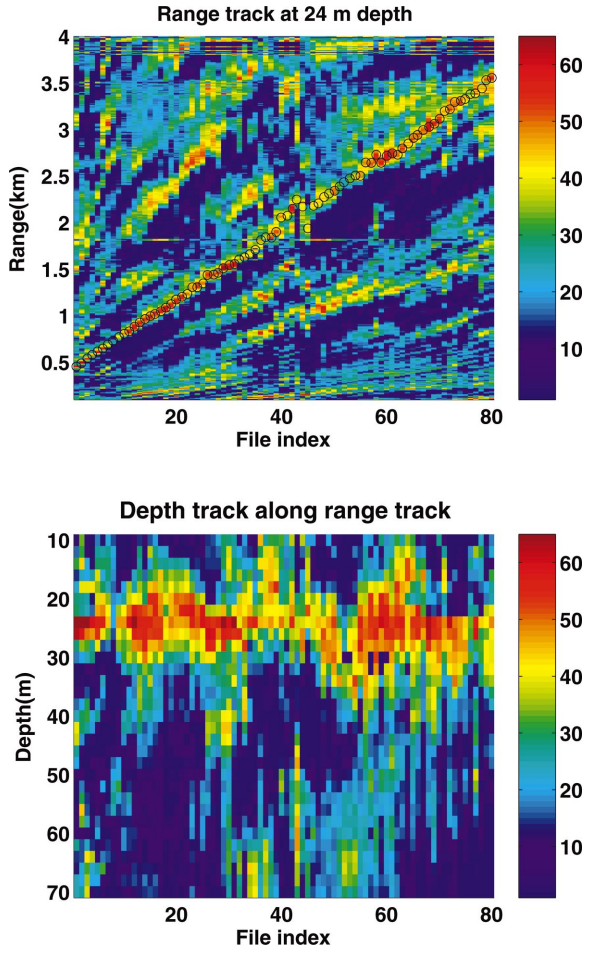

FIG. 16. Upper plot shows range track at 24 m depth. Ranges at which a depth slice is taken are indicated by black circles. The lower plot shows the depth track along the range track. The black circles over the range track show the peaks picked along the range track—these were used to set the ranges at which the slices versus depth were taken to form the depth track.

produce a persistent range and depth track out to 1.4 km, using relatively simple modeling. At the La Jolla site, where we had less multipath due to several environmental factors (downward refracting profile, softer bottom, deeper water), we had to work harder to pre-emphasize features of the impulse response function. Nevertheless, after these modifications, we were able to produce a persistent range and depth track out to 3.5 km. The additional transforms applied to the La Jolla site were not tested on the New England data.

Admittedly, using a known source wave form to measure the impulse response would not be possible with an uncooperative source, so the matched filter examples of source localization in Secs. II C and III C are perhaps academic demonstrations of feasibility only. To show that similar techniques could be applied to the more realistic problem of tracking an uncooperative source, whose wave form would not be known, in Sec. III D, we repeated the tracking at the La Jolla site using cross-correlation wave forms as data inputs (similarly to how we have done this in previous work at lower frequencies[21]). Note that only a time-domain representation could have enabled us to manipulate the relative importance of the different multipath arrivals, which produced dramatically improved localization ranges in Secs. III C and III D.

The averaging over multiple impulse response measurements proved to be a valuable preprocessing step that produced a very structured impulse response, suitable for comparison with model predictions. To reduce the sensitivity to mismatch in the times of arrival, the signal bandwidth was artifically reduced in the modeled waveforms. To obtain source location estimates beyond a kilometer in the La Jolla data, somewhat *ad hoc* transformations of the impulse response (data prewhitening and raising the modeled wave forms to a small power) were used to reduce the impact of unpredictable early arrivals and to emphasize the later arrivals, as described in Secs. III C and III D. These additional transforms were not applied to the New England data. These machinations were needed because our modeling did not incorporate phenomena that cause fluctuations in the impulse response. We speculate that incorporating better models of such phenomena would improve the source location estimates (different arrivals will behave differently, for example, depending on whether they interact with a dynamic surface, a rough bottom, or a part of the water column where the sound speed is changing). Extending our propagation models to incorporate knowledge of such ocean dynamics or its statistics is a topic for future work.

ACKNOWLEDGMENTS

This work was funded by ONR under Contract No. N00014-00-D-0115. The SignalEx experiments were funded by ONR 322OM. The tracking algorithms were developed as part of the Hydra program, supported by ONR 321SS. The propagation modeling was supported by ONR 321OA.

[1] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," J. Acoust. Soc. Am. **59**, 368–373 (1976).
[2] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," IEEE J. Ocean. Eng. **18**, 401–424 (1993).
[3] N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, "Source localization with broad-band matched-field processing in shallow water," IEEE J. Ocean. Eng. **21**, 402–412 (1996).
[4] A. Parvulescu, "Signal detection in a multipath medium by M.E.S.S. processing," J. Acoust. Soc. Am. **33**, 1674 (1961).
[5] A. Parvulescu and C. S. Clay, "Reproducibility of signal transmissions in the ocean," Radio Electron. Eng. **29**, 223–228 (1965).
[6] L. Nghiem-Phu, F. D. Tappert, and S. C. Daubin, "Source localization by cw acoustic retrogration," 1985.
[7] D. R. Jackson and D. R. Dowling, "Phase conjugation in underwater acoustics," J. Acoust. Soc. Am. **89**, 171–181 (1991).
[8] R. E. Williams and H. F. Battestin, "Coherent recombination of acoustic multipathsignals propagated in the deep ocean," J. Acoust. Soc. Am. **50**, 1433–1442 (1971).
[9] W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation," J. Acoust. Soc. Am. **58**, 201–207 (1975).
[10] N. L. Owsley and G. R. Swope, "Time delay estimation in a sensor array," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-29**, 519–523 (1981).
[11] E. K. Westwood, "Broadband matched-field source localization," J. Acoust. Soc. Am. **91**, 2777–2789 (1992).
[12] M. Hamilton and P. M. Schultheiss, "Passive ranging in multipath dominant environments. I. Known multipath parameters," IEEE Trans. Signal Process. **40**, 1–12 (1992).
[13] S. M. Jesus, "Broadband matched-field processing of transient signals in

shallow water,'' J. Acoust. Soc. Am. **93**, 1841–1850 (1993).
[14] R. K. Brienzo and W. Hodgkiss, ''Broadband matched-field processing,'' J. Acoust. Soc. Am. **94**, 2821–2831 (1993).
[15] M. Hamilton and P. M. Schultheiss, ''Passive ranging in multipath dominant environments. II. Unknown multipath parameters,'' IEEE Trans. Signal Process. **41**, 1–12 (1993).
[16] Z.-H. Michalopoulou, M. B. Porter, and J. P. Ianniello, ''Broadband source localization in the Gulf of Mexico,'' J. Comput. Acoust. **4**, 361–370 (1996).
[17] Z.-H. Michalopoulou and M. B. Porter, ''Source tracking in the Hudson Canyon Experiment,'' J. Comput. Acoust. **4**, 371–383 (1996).
[18] E. K. Westwood and D. P. Knobles, ''Source track localization via multipath correlation matching,'' J. Acoust. Soc. Am. **102**, 2645–2654 (1997).
[19] J. P. Ianniello, ''Recent developments in sonar signal processing,'' IEEE Signal Process. Mag. **15**, 27–40 (1998).
[20] Z.-H. Michalopoulou, ''Matched impulse-response processing for shallow-water localization and geoacoustic inversion,'' J. Acoust. Soc. Am. **108**, 2082–2090 (2000).
[21] M. B. Porter, P. Hursky, C. O. Tiemann, and M. Stevenson, ''Model-based tracking for autonomous arrays,'' in *MTS/IEEE Oceans 2001—An Ocean Odyssey, Conference Proceedings* (IEEE Press, Honolulu, HI, 2001), pp. 786–792.
[22] W. S. Hodgkiss, W. A. Kuperman, J. J. Murray, G. L. D'Spain, and L. P. Berger, ''High frequency matched field processing,'' in *High Frequency Acoustics in Shallow Water*, edited by N. G. Pace, E. Pouliquen, O. Bergem, and A. P. Lyons (NATO SCALANT Undersea Research Centre, La Spezia, Italy, 1997), pp. 229–234.
[23] S. B. Suppappola and B. F. Harrison, ''Experimental results for matched-field processing with a small-aperture mid-frequency array,'' in *Oceans 2000 MTS/IEEE Conference and Exhibition, Conference Proceedings* (IEEE Press, 2000), Vol. 1, pp. 439–446.
[24] E. I. Thorsos, ''Report on the Office of Naval Research High-Frequency Acoustics Workshop,'' Technical Report No. APL-UW TR 9702, Applied Physics Laboratory, University of Washington, 16–18 April, 1996 (unpublished).
[25] M. B. Porter, V. K. McDonald, P. A. Baxley, and J. A. Rice, ''SignalEx: Linking environmental acoustics with the signaling schemes,'' in *MTS/IEEE OCEANS'00 Conference Proceedings of* (IEEE Press, 2000), pp. 595–600.
[26] S. Kim, G. F. Edelmann, W. A. Kuperman, W. S. Hodgkiss, H. C. Song, and T. Akal, ''Spatial resolution of time-reversal arrays in shallow water,'' J. Acoust. Soc. Am. **110**, 820–829 (2001).
[27] M. B. Porter, Acoustics Toolbox, http://oalib.saic.com/Modes/AcousticsToolbox.
[28] M. B. Porter, ''Acoustic models and sonar systems,'' IEEE J. Ocean. Eng. **OE-18**, 425–437 (1994).
[29] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP Press, Woodbury, NY, 1994).
[30] M. B. Porter, S. M. Jesus, Y. Stéphan, X. Démoulin, and E. Coelho, ''Exploiting reliable features of the ocean channel response,'' in *Shallow Water Acoustics*, edited by R. Zhang and J. Zhou (China Ocean Press, Beijing, China, 1998), pp. 77–82.
[31] W. A. Struzinski and E. D. Lowe, ''A performance comparison of four noise-background normalization schemes proposed for signal detection systems,'' J. Acoust. Soc. Am. **76**, 1738–1742 (1984).