The Journal of the Acoustical Society of America

JULY 13 2021

Range-coherent matched field processing for low signal-to-noise ratio localization

F. Hunter Akins; W. A. Kuperman

J. Acoust. Soc. Am. 150, 270–280 (2021)
https://doi.org/10.1121/10.0005586

JASA ARTICLE

Range-coherent matched field processing for low signal-to-noise ratio localization

F. Hunter Akins[a] and W. A. Kuperman

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0701, USA

ABSTRACT:

Range-coherent matched field processing (MFP) coherently combines snapshots to localize a moving, narrowband source. This approach differs from existing MFP approaches that treat each snapshot as having a random phase due to both unknown motion through the medium and imprecise knowledge of the source frequency. Range-coherent MFP requires determination of the source phase acquired between snapshots. With that information, MFP can be applied to the cross-spectrum of snapshots acquired at different times, since relative phase between snapshots is determined by the medium properties, source location, and source velocity. Viewed another way, range-coherent MFP is simply MFP applied to a passive synthetic aperture formed from a moving source. The synthetic aperture geometry depends on source velocity, which is included in the MFP search space. Range-coherent MFP produces robust velocity estimates at low signal-to-noise ratio (SNR), which permits the use of a longer fast Fourier transform in pre-processing. The synthetic aperture array gain plus the increased input SNR afforded by the enhanced pre-processing significantly lowers the required signal level for successful localization. In data from the SWellEx-96 experiment, range-coherent MFP successfully localizes a source that is too quiet for conventional methods to localize.

 https://doi.org/10.1121/10.0005586

(Received 13 April 2021; revised 17 June 2021; accepted 17 June 2021; published online 13 July 2021)

[Editor: Kainam Thomas Wong] Pages: 270–280

I. INTRODUCTION

Matched field processing (MFP), a generalized beamforming method, correlates data acquired on acoustic arrays with model outputs to localize sources.[1] For each candidate source position, a physical model provides a prediction, which is then compared with the data in either a linear (e.g., the Bartlett beamformer) or nonlinear/adaptive (e.g., maximum-likelihood estimator) fashion to estimate the likelihood of the candidate being the true position. MFP remains a topic of largely theoretical interest because it requires (a) high input signal-to-noise ratio (SNR) and (b) accurate environmental knowledge, both of which are rarely available in situations of practical interest. In this paper, we confront the limitation due to SNR by applying MFP to a passive synthetic aperture formed from a moving source.

Passive synthetic apertures are a method of increasing system gain and resolution.[2,3] Several works demonstrate the formation of towed-array synthetic apertures for bearing estimation,[3–5] as well as for improving geoacoustic inversion results in shallow water.[6,7] The goal of this paper is to incorporate synthetic aperture theory directly into an MFP algorithm in a simple and robust way, using the increased gain to lower the SNR regime in which localization efforts can succeed.

A few conclusions may be drawn from a review of passive synthetic aperture work in underwater acoustics. First, coherence times for stable, low-frequency continuous wave (CW) transmissions in the ocean may be quite long.[8] Second, the main limitation in synthetic aperture implementations comes from tow-path irregularities.[8–11] Third, synthetic aperture source localization efforts have been largely limited to processing data from towed horizontal line arrays (HLAs) to obtain bearing estimates.[3,4,12] Range-coherent MFP takes advantage of signal coherence at low frequencies to extend synthetic aperture methods to source localization in multipath shallow water environments.

Range-coherent MFP also has much in common with coherent multi-frequency processors.[13–16] Those processors form a synthetic array (or "supervector") by stacking measurements made on an array at various *frequencies*. The proposed processor forms a synthetic array by stacking measurements made on an array at various *ranges*. Those processors require knowledge of the source spectrum at each frequency, either *a priori* or via an estimator. The proposed narrowband processor uses the source phase accumulated between snapshots. In both cases, additional gain comes at the price of additional knowledge or algorithm complexity.

It is worth distinguishing the proposed method from a third category of research: techniques that use source motion to incoherently combine snapshots from different ranges. Explicit target motion compensation (ETMC)[17,18] addresses the "blurring" effect that source motion has on sample covariance matrix (SCM) formation by using a velocity hypothesis to correct for source motion. A recently developed method reformulates ETMC using waveguide invariant theory,[19] although source velocity still appears implicitly.

[a]Electronic mail: fakins@ucsd.edu, ORCID: 0000-0003-1191-0790.

 0001-4966/2021/150(1)/270/11/$30.00

Matched field tracking (MFT),[20,21] another incoherent approach, sums ambiguity surfaces from separate covariance integration periods according to a velocity hypothesis. Whereas ETMC improves estimation of the usual SCM, range-coherent MFP estimates a larger SCM whose elements include the cross-spectrum of the synthetic aperture elements. In this sense, it has more in common with coherent passive synthetic aperture methods and frequency-coherent processors than with methods such as ETMC and incoherent tracking algorithms.

During data pre-processing, range-coherent MFP demonstrates an additional advantage over traditional MFP. At low SNR, traditional MFP is constrained to use a fast Fourier transform (FFT) bin width that encompasses the unknown Doppler shift. Range-coherent MFP's capability to produce robust velocity estimates at low SNR removes this constraint, permitting a longer FFT in pre-processing and accordingly improved input SNR.

This paper formulates range-coherent MFP on a vertical line array (VLA) and compares its performance to existing MFP approaches. Section II presents background theory for a moving, narrowband source in shallow water and shows how to form a passive synthetic aperture. Section III reviews relevant aspects of linear (Bartlett) and white noise constraint (WNC) matched field processors and SCM estimation. Section IV presents simulations and Cramér–Rao lower bound (CRLB) calculations that compare range-coherent MFP and traditional MFP performance. Section V discusses the SWellEx-96 experiment, along with pre-processing and modeling considerations. Section VI presents the results of applying range-coherent MFP to the SWellEx-96 dataset. The range-coherent WNC processor localizes the source using a received multi-tone signal 36 dB quieter than the pilot tones used for localization in previous studies. To the best of the authors' knowledge, this is the first study demonstrating successful localization on any of the low-level tonal sets transmitted during SWellEx-96.

II. CONSTRUCTING A PASSIVE SYNTHETIC APERTURE

This section shows how to form a synthetic aperture from a narrowband source moving at constant range-rate and depth through shallow water. The material presented here will be applied to data in Sec. VI.

A. Uniform range-rate source in shallow water

Modal doppler theory provides a description of the distant field radiated by a narrowband source moving at constant depth z_s and uniform range-rate (radial velocity) v (no acceleration) through a range-independent shallow water waveguide.[22–24] Even though the source is narrowband, the received signal has a finite bandwidth that is determined by v, along with the waveguide properties. Wavenumbers computed for a stationary source will differ from those for a moving source by a factor on the order of v/c, where c is the sound speed. Here, we assume the source moves slowly enough that use of the stationary source wavenumbers introduces negligible errors. The position track of the source is given in cylindrical coordinates as

$$(r(t), z(t)) = (r_0 + vt, z_s), \tag{1}$$

and it radiates at angular frequency ω_0.

For a stationary sensor at a depth z on the z axis of a cylindrical coordinate system, the received field from the moving source is

$$p(r,z,t) = e^{j\omega_0 t} A(t) \sum_i a_i(z,z_s) e^{-jk_i r_0} e^{-jk_i vt}. \tag{2}$$

The a_i are the modal amplitudes

$$a_i = \frac{\Phi_i(z)\Phi_i(z_s)}{\sqrt{k_i}}, \tag{3}$$

where Φ is the ith modal eigenfunction and k_i is its associated eigenvalue.

$$A(t) = \frac{e^{-j\pi/4 + j\phi_0}}{\sqrt{8\pi(r_0 + vt)}} \tag{4}$$

is a slowly varying envelope that contains the source phase term, $e^{j\phi_0}$, and the range spreading term.

It is convenient to introduce the energy-averaged wavenumber $\bar{k} \equiv \sum_i a_i^2 k_i / \sum_i a_i^2$, and the received "carrier frequency" $\omega_d \equiv \omega_0 - \bar{k}v$. Then we can rewrite Eq. (2) as

$$p(r,z,t) = e^{j\omega_d t} A(t) \sum_i a_i(z,z_s) e^{-jk_i r_0} e^{-j(k_i - \bar{k})vt}. \tag{5}$$

The bandwidth of the signal is largely determined by the terms of the sum and is well-approximated by the Doppler spread between the highest and lowest order modes

$$B \approx \frac{1}{2}(\max\{k_i\} - \min\{k_i\})v/2\pi. \tag{6}$$

The coherence time, τ, is defined as the inverse of B,

$$\tau = \frac{4\pi}{v \times (\max\{k_i\} - \min\{k_i\})}. \tag{7}$$

As a reference example, for a 100 Hz signal moving through the far-field at $v = 2.5$ m/s in a 100 m deep Pekeris waveguide with bottom speed $c_b = 1600$ m/s and density $\rho_b = 1800\,\mathrm{kg/m^3}$, τ is roughly 200 s.

If we denote the time at the beginning of a short snapshot window (duration $T \ll \tau$) by t_0, the complex baseband pressure at the Doppler-shifted frequency ω_d is well-approximated by

$$\tilde{p}(z,t_0) = W e^{j\omega_d t_0} A(t_0) \sum_i a_i(z,z_s) e^{-jk_i r(t_0)}, \tag{8}$$

which is the field due to a stationary source at $r(t_0)$ (W, a real number, accounts for the scaling effect of the window on the signal amplitude).

https://doi.org/10.1121/10.0005586

Note the source phase term in front of the modal sum (8): $e^{j\omega_d t_0} = e^{j\omega_0 t_0 + j\bar{k}vt_0}$. As pointed out by Bucker in one of the earliest MFP papers,[25] this phase term (a) naturally rotates with time if the reference frequency is not precisely matched to ω_0 and (b) rotates with (small) changes in source location. This motivates the use of the cross-spectral density matrix in array processing, which only emphasizes the relative phase between array elements. However, with (a) precise knowledge of the source frequency ω_0 and (b) an estimate of the range-rate v, the relative phase between snapshots can be modeled, and a synthetic aperture can be formed.

B. Forming a passive synthetic aperture

To form a synthetic array, take a set of snapshots at a single array element at depth z for a sequence of initial times $\{t_n\} = \{t_0 + nT_{syn}\}$ to obtain a corresponding sequence of complex values $\{\tilde{d}_n(z)\} = \{\tilde{p}(z, t_n) + \tilde{n}(z, t_n)\}$ that describes the evolution of the pressure field at that sensor. $\tilde{p}$ represents the signal portion of the data, and $\tilde{n}$ is additive noise. From (8), note that the signal portion of each snapshot has a source phase component $e^{j\omega_0 t_n}$ related to the source frequency and the start of the snapshot time. Remove the source phase component from each snapshot by multiplying by its conjugate

$$\tilde{d}'_n(z) \equiv e^{-j\omega_0 t_n}\tilde{d}_n(z). \tag{9}$$

From Eq. (5), the modified snapshots will then be described by

$$\tilde{d}'_n(z) = WA(t_n)\sum_i a_i(z, z_s)e^{-jk_i r(t_n)} + \tilde{n}'_n. \tag{10}$$

The noise has been modified to $\tilde{n}'_n = \tilde{n}_n e^{-j\omega_0 t_n}$.

Now the difference between $\tilde{d}'_m$ and a later measurement $\tilde{d}'_n$ is only dependent on the range traversed by the source from time t_m to time t_n. In other words, the modified $\{\tilde{d}'_n\}$ are identical to the signal received from a stationary source on a stationary HLA with element spacing vT_{syn}.

If the original measurement system is an M-element VLA with sensors at depths $z_0 + m\Delta z$, $m = 0, 1, \ldots, M-1$, then the moving source generates a planar synthetic array oriented endfire to a source. The array "elements" are indexed by l, with a column-major indexing. With this indexing, the lth array element corresponds to a sensor at the $(l \bmod M)$ th sensor depth and the time sample corresponding to the integer part of l/M. Put more simply, the N_{syn} column vectors of modified VLA measurements are stacked into a "supervector" $\tilde{\mathbf{d}}$ of length $M \cdot N_{syn}$, which is then fed into an MFP algorithm.

III. MFP

As shown in Sec. II, data from sequential times may be formed into a single array of measurements, $\tilde{\mathbf{d}}$. A SCM, $\hat{\mathbf{K}} = (1/L)\sum_i \tilde{\mathbf{d}}_i\tilde{\mathbf{d}}_i^\dagger$, can be formed from L samples $\{\tilde{\mathbf{d}}_i\}$ of $\tilde{\mathbf{d}}$. Note that for range-coherent MFP, the SCM is formed from the "supervectors" described in Sec. II.

MFP algorithms compute a weighting vector $\mathbf{w}(\boldsymbol{\theta})$ for each potential vector of source parameters $\boldsymbol{\theta}$ and combine them with $\hat{\mathbf{K}}$ to produce an ambiguity surface

$$P(\hat{\mathbf{K}}, \boldsymbol{\theta}) = \mathbf{w}^\dagger(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{w}(\boldsymbol{\theta}). \tag{11}$$

In general, the weighting vector $\mathbf{w}(\boldsymbol{\theta})$ is some function of the steering vector $\mathbf{s}(\boldsymbol{\theta})$, computed from an acoustic model for each search $\boldsymbol{\theta}$. In the absence of any mismatch, the steering vector $\mathbf{s}(\boldsymbol{\theta})$ will match the source signal $\tilde{\mathbf{p}}(\boldsymbol{\theta}_T)$ received on the array when $\boldsymbol{\theta}$ matches $\boldsymbol{\theta}_T$, the true source parameter value. Steering vectors $\mathbf{s}$ are normalized to the number of array elements M: $\sqrt{\mathbf{s}^\dagger\mathbf{s}} = M$.

In a range-independent environment, MFP uses the source position in cylindrical coordinates as the source parameter vector: $\boldsymbol{\theta} = (r\ z)^T$. Range-coherent MFP for a constant depth, uniform range-rate source includes source range-rate v in the parameter vector: $\boldsymbol{\theta} = (r\ z\ v)^T$.

A. Bartlett processor

The Bartlett (also called "linear" or "conventional") processor[25] uses a scaled steering vector as the weighting vector, $\mathbf{w}(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta})/M$, so the Bartlett ambiguity surface is simply

$$P_{BART}(\hat{\mathbf{K}}, \boldsymbol{\theta}) = \frac{1}{M^2}\mathbf{s}^\dagger(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{s}(\boldsymbol{\theta}). \tag{12}$$

The Bartlett processor is cheap to compute and robust to model mismatch but performs poorly in the presence of interferers and significant sidelobe ambiguity structure.

B. WNC processor

The WNC processor[26,27] strives to balance the interference rejection and sidelobe suppression of a minimum-variance (MV) beamformer (also called Capon's method and the maximum-likelihood method) with the robustness of the Bartlett processor. To accomplish this, it seeks an adaptive weighting vector $\mathbf{w}_a(\theta)$ that minimizes total output power, maintains a borehole constraint, and satisfies an inequality constraint on the white noise gain (defined below)

$$\begin{aligned} &\min \mathbf{w}_a^\dagger(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{w}_a(\boldsymbol{\theta}) \\ \text{subject to}\quad &\mathbf{w}_a(\boldsymbol{\theta})^\dagger\mathbf{s}(\boldsymbol{\theta}) = 1 \\ &\left(\mathbf{w}_a(\boldsymbol{\theta})^\dagger\mathbf{w}_a(\boldsymbol{\theta})\right)^{-1} \geq \delta^2. \end{aligned} \tag{13}$$

The white noise gain is defined as

$$G_W = \frac{|\mathbf{w}_a^\dagger\mathbf{s}|^2}{\mathbf{w}_a^\dagger\mathbf{w}_a}. \tag{14}$$

Its inverse $S_W = 1/G_W$ is a measure of sensitivity to errors between the actual signal $\tilde{\mathbf{p}}$ and the steering vector $\mathbf{s}$.[26] From the Cauchy–Schwarz inequality and the normalization of $\mathbf{s}$,

$$G_W \leq M \tag{15}$$

with equality when $\mathbf{w} = \alpha\mathbf{s}$ for any $\alpha \in \mathbb{R}$. Hence, the Bartlett weight vector $\mathbf{w} = \mathbf{s}/M$ maximizes white noise gain and minimizes sensitivity. Removing the WNC entirely (or, equivalently, setting $\delta^2 = 0$) results in the MV processor, which maximizes sensitivity. The user's choice of δ^2 controls the trade-off between robustness and sidelobe suppression/interference rejection.

Solving the constrained optimization problem posed in Eq. (13) yields the corresponding WNC weighting vector

$$\mathbf{w}_a(\hat{\mathbf{K}},\boldsymbol{\theta},\delta^2) = \frac{(\hat{\mathbf{K}}+\epsilon\mathbf{I})^{-1}\mathbf{s}(\boldsymbol{\theta})}{\mathbf{s}(\boldsymbol{\theta})(\hat{\mathbf{K}}+\epsilon\mathbf{I})^{-1}\mathbf{s}(\boldsymbol{\theta})}, \tag{16}$$

where $\mathbf{I}$ is the identity matrix and $\epsilon(\hat{\mathbf{K}},\boldsymbol{\theta},\delta^2)$ is a regularization parameter. An eigenvector decomposition of $\hat{\mathbf{K}}$ permits efficient calculation of (16) (see 10.3.4 of Jensen *et al.*[23]) to find, for each search location, an ϵ that satisfies the white noise gain constraint.

Often δ^2 is measured as some number of dB down from the maximum white noise gain, so the expression "the white noise gain is constrained to X dB" translates to $10\log_{10}(\delta^2/M) = X$. Thus, the 0 dB WNC processor is a Bartlett processor, and the $-\infty$ dB WNC processor is the MV processor.

The output of the WNC processor is, from Eq. (11),

$$P_{WNC}(\hat{\mathbf{K}},\boldsymbol{\theta},\delta^2) = \mathbf{w}_a(\hat{\mathbf{K}},\boldsymbol{\theta},\delta^2)\hat{\mathbf{K}}\mathbf{w}_a(\hat{\mathbf{K}},\boldsymbol{\theta},\delta^2). \tag{17}$$

C. Snapshot deficiency

Adaptive methods presuppose an accurate estimate of the SCM $\hat{\mathbf{K}}$. Snapshot deficiency occurs when a limited number of snapshots available to form $\hat{\mathbf{K}}$ results in a poor, possibly singular estimate of the SCM. For a moving source, the number of available snapshots is limited by the amount of time spent by the source in a range cell (the "beam width" of a Bartlett range estimator[18]). The synthetic array SCM dimension is a factor of N_{syn} larger than the dimension of the simple VLA SCM; range-coherent MFP requires N_{syn} times more snapshots than "traditional" MFP. As a result, snapshot deficiency is likely to occur for all but the slowest moving sources.

There are a number of methods for mitigating snapshot deficiency. The simplest method is diagonal loading.[28] This strategy will fail to null a loud interferer creating colored noise along the array but solves the singularity issue of a severely deficient $\hat{\mathbf{K}}$. To deal with a loud interferer or colored environmental noise, a full-rank SCM can be formed by averaging over nearby frequency bins[29] or by averaging sequences of shorter snapshots.[30] Alternatively, subspace methods can be applied to the snapshot-deficient SCM.[31] The SWellEx-96 event analyzed in this paper has no loud interferers, so the simplest approach is taken: the SCMs used are diagonally loaded

$$\hat{\mathbf{K}}' \equiv \hat{\mathbf{K}} + \epsilon\mathbf{I}, \tag{18}$$

with $\epsilon = 10^{-20}$ (around 8 orders of magnitude lower than the noise power on a single phone). However, in a more complicated noise environment, one of the above strategies would be necessary.

D. Multi-tone MFP

When a source transmits at multiple frequencies, a single ambiguity surface may be obtained by performing a geometric mean of the ambiguity surfaces from processing each transmitted frequency.[1] The result is a single ambiguity surface

$$P = \frac{1}{N_f}\sum_{k=1}^{N_f} 10\log_{10}P(f_k). \tag{19}$$

More optimal approaches that compensate for the frequency-dependence of background noise exist. However, the simplicity and effectiveness of the naive summation are deemed sufficient for the present work. Whenever MFP results are presented in this paper, they will be formed from the incoherent frequency sum in Eq. (19).

IV. PERFORMANCE OF RANGE-COHERENT MFP

A synthetic array demonstrates two advantages over the simple array from which it is formed: array gain and improved resolution. Both of these advantages follow from the near equivalence between the passive synthetic array formed from a moving source and a VLA and a planar endfire array recording a stationary source. The only difference between these two measurement scenarios is that a planar endfire array recording a stationary source does not suffer from Doppler broadening and limited covariance integration time.

The CRLB provides a tool to quantify both the resolution and threshold effect for both simple VLAs and synthetic planar arrays. The CRLB provides a lower bound on the variance of any statistical estimator of a given parameter.[1,32] The CRLB shows that although there is an increase in resolution from adding synthetic array elements, the main benefit of the synthetic aperture method is the lowered required input SNR.

We computed the CRLB for the case of a source transmitting at 100 Hz in an ideal waveguide (pressure release surface and rigid bottom) with sound speed $c = 1500$ m/s and depth $D = 216$ m. The parameter vector to be estimated is $\mathbf{a}_T = (r_s, z_s)$, the source range and depth. The data model consists of the signal from the source with uniformly distributed random source phase $e^{j\phi_0}\mathbf{p}(r_s, z_s)$ plus additive, white complex Gaussian noise received on the VLA shown in Fig. 1. The simulated data follow a zero-mean, complex Gaussian distribution with covariance matrix $\mathbf{K}(r_s, z_s)$ determined by the waveguide properties, source location, and noise power σ_W^2,

$$\mathbf{K}(r_s, z_s) = \sigma_W^2\mathbf{I} + \mathbf{p}(r_s, z_s)\mathbf{p}^\dagger(r_s, z_s). \tag{20}$$

Input power is defined to be $\sigma_S^2 = (1/N)\mathrm{trace}(\mathbf{p}\mathbf{p}^\dagger)$, and input SNR in dB is $10\log_{10}(\sigma_S^2/\sigma_W^2)$. From Eq. (15.52) of Kay,[32] the Fisher information matrix (FIM) is

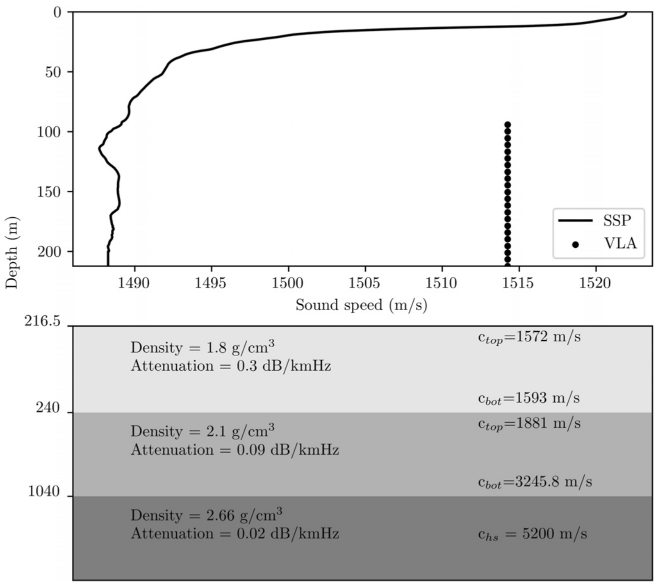

FIG. 1. Environment used for modeling the SWellEx-96 S5 event. Consists of water column with sound speed profile (SSP) from a conductivity–temperature–depth (CTD) cast during the experiment overlying the geoacoustic model derived in Baxley *et al.* (Ref. 33). The recording system consists of a VLA with 21 elements spaced at 5.625 m.

$$\mathbf{J}_{ij} = \mathrm{trace}\left(\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \mathbf{a}_{\mathrm{T}i}}\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \mathbf{a}_{\mathrm{T}j}}\right). \tag{21}$$

The resulting bounds on the variance of any estimator, $\hat{\mathbf{a}}_i$, of $\mathbf{a}_{Ti}$ follow from the diagonal elements of the inverse of the FIM,

$$E((\hat{\mathbf{a}}_i - \mathbf{a}_{Ti})^2) \geq [\mathbf{J}^{-1}]_{ii}. \tag{22}$$

The results from applying the CRLB to three array geometries are shown in Fig. 2(a). As expected, the array gain of additional VLA staves lowers the input SNR threshold required for accurate localization.

CRLB calculations for SNR thresholds are also consistent with Monte Carlo simulations of the (inefficient) Bartlett processor applied to a snapshot-deficient SCM. For each of the three array geometries and input SNR values, we generated 1000 realizations of the SCM. The SCM for each realization was formed from four independent snapshots (signal plus complex white Gaussian noise). Applying the Bartlett range estimator to each realization generates an ensemble of 1000 range estimates. From these we form an estimate of the root-mean-square error (RMSE). The RMSE estimates are shown in Fig. 2, and the SNR thresholds at which they degrade match the CRLB-derived thresholds.

The CRLB and simulations show that range-coherent MFP can localize a source in SNR regimes where traditional MFP cannot.

V. SWELLEX-96 EXPERIMENT

Here, we provide a brief orientation to the SWellEx-96 data set[34] (and more specifically the S5 event) as well as the pre-processing and modeling required to perform MFP and range-coherent MFP.

A. S5 event

SWellEx-96 refers to a well-known shallow water MFP experiment that was conducted off the coast of San Diego in 1996. The S5 event consists of a 75 min source tow at roughly 2.5 m/s under a smooth sea state and over variable bathymetry. The S5 event has a number of features that make it a good candidate for demonstrating our proposed method. It contains a slowly moving source transmitting CW signals, so the model in Sec. II matches the scenario. A geoacoustic model that optimizes the match between replicas and data was developed by Bachman *et al.*[35] and later refined by Baxley *et al.*,[33] so accurate replica calculation is not an issue. It also contains low SNR data that have never been used successfully for localization.

The environment is shown in Fig. 1, and the range of the ship, which towed the two acoustic sources, for the S5 event is shown in Fig. 3. The water column depth at the

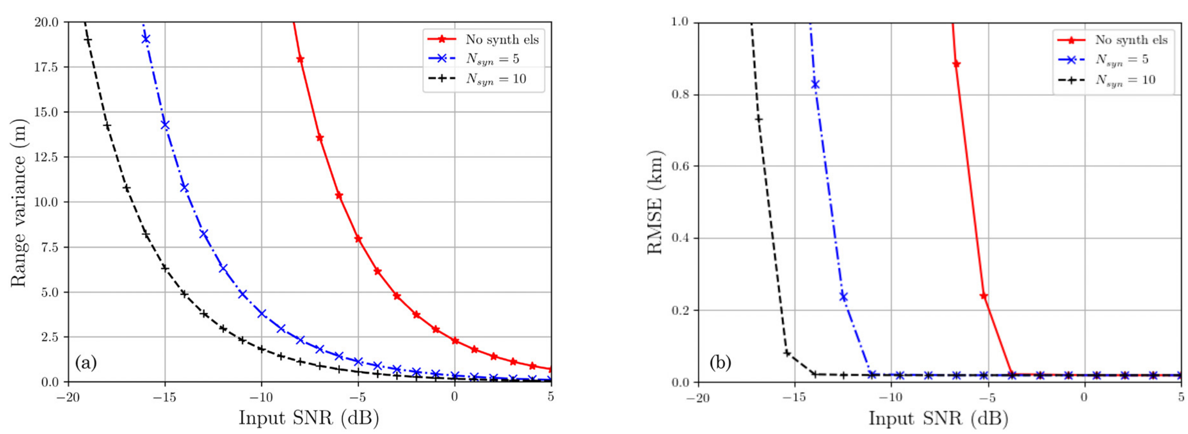

FIG. 2. (Color online) Performance comparison of a simple VLA, a five-element planar array, and a ten element planar array. (a) shows the CRLB for range estimators as a function of input SNR for the three arrays. In (b), we show the RMSE for Bartlett range estimates from 1000 simulation runs for a range of input SNR. Note the different units between (a) and (b). For each simulation, four snapshots are used to form the SCM (four realizations of noise + signal).

https://doi.org/10.1121/10.0005586

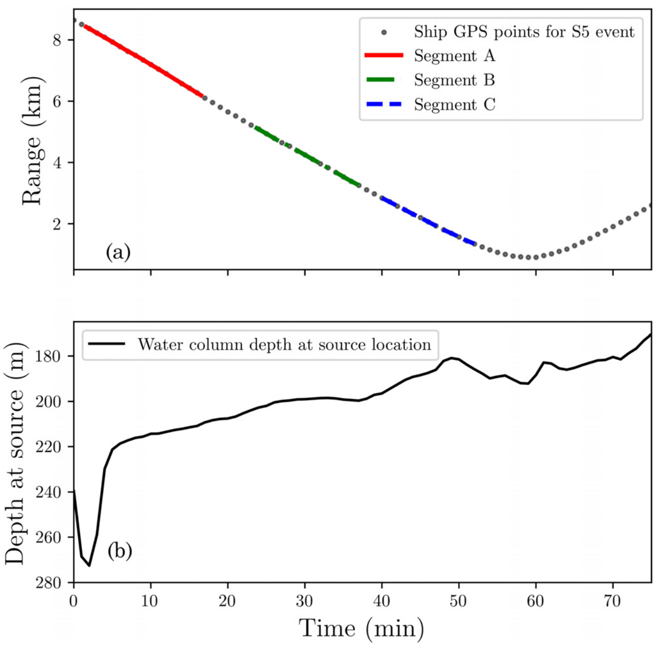

FIG. 3. (Color online) S5 event. The ship's range from the VLA is shown in (a). The three segments of the track that are analyzed, labeled A, B, and C, are superimposed over the GPS track. (b) shows the water column depth at the source location over the course of the event.

ship's location as a function of experiment time is also shown. The two sources consist of a shallow source at 9 m depth and a deep source at around 55 m depth.

The deep source simultaneously transmitted five sets of tonals, which are labeled here as TSL ("L" for "loud"), TSQ1, TSQ2, TSQ3, and TSQ4 ("Q" for "quiet"). As shown in Table I, TSL is 26 dB above TSQ1, and each subsequent quiet tonal set is 4 dB below the previous set. Each tonal set contains 13 tones spaced between 49 and 400 Hz, and the tones of each set differ from those of the adjacent sets by 3 Hz.

The data are received on a VLA with 64 elements spanning the water column from 100 to 200 m with a sampling rate of 1500 Hz. Only every third element from the array was included in the processing, resulting in a 21 element VLA with 5.625 m spacing.

TABLE I. Signal level and maximum input SNR, averaged over tones in each set, for processed segments of the S5 event for two different FFT lengths. The maximum SNR occurs at the closest range analyzed, around 1.3 km. As the source moves away, SNR drops approximately by the cylindrical spreading law $10\log_{10}(r/r_{min})$. For example, for the 8 km range, the SNR will be 8 dB lower than the maximum value shown here.

Tonal set	Source level (dB re 1 μPa)	Maximum SNR, $N_{FFT} = 2048$ (dB)	Maximum SNR, $N_{FFT} = 16\,384$ (dB)
TSL	158	11	21.5
TSQ1	132	−15	−4.5
TSQ2	128	−19	−8.5
TSQ3	124	−23	−12.5
TSQ4	120	−27	−16.5

B. FFT length, SNR and SCM formation

FFT length T simultaneously sets the input SNR to the matched field processor and limits the number of snapshots available for forming the SCM. When processing data from a narrowband source moving at unknown velocity, a wide bin width is required to account for the unknown Doppler shift. For a source moving at 2.5 m/s with frequency 400 Hz, the Doppler shift is 0.66 Hz (as calculated from $\bar{k}v/2\pi$). Calculations (not shown for brevity) show that for a Hamming window longer than 1 s, the leakage loss for the 400 Hz band will exceed 3 dB. The lack of source velocity knowledge limits FFT length and, hence, input SNR. Since range-coherent MFP includes velocity as a search parameter, the leakage constraint derived above no longer applies. For each search range-rate, the Doppler shift can be computed, and the appropriate bin can be selected for input to the matched field processor. Of course, one could add search range-rate and implement this bin-selection algorithm in traditional MFP. However, we found that simply using search range-rate to estimate the proper bins for traditional MFP completely fails at low SNR. Therefore, this strategy is only available to the range-coherent method, which can accurately estimate range-rate at low SNR.

SNR estimations performed for TSL, shown in Fig. 4, demonstrate the received signal coherence properties and inform FFT length selection when the Doppler shift is known. For a coherent signal in white noise, doubling FFT length T should increase SNR by around 3 dB until T approaches the coherence time τ (the time at which the source leaves its range cell). Once T reaches τ, modal leakage out of the FFT bin cancels out the coherent gain, and the SNR reaches a plateau.

With these considerations in mind, the data were processed in two ways: one with a 2048 point FFT ($T = 1.4$ s) with 50% overlap and another with a 16 384 point FFT ($T = 10.9$ s) with 50% overlap. Table I shows the resulting input SNR for these two settings for each tonal set. Bins corresponding to the nominal source frequency are selected from the 2048-point FFT to feed into traditional, "velocity-

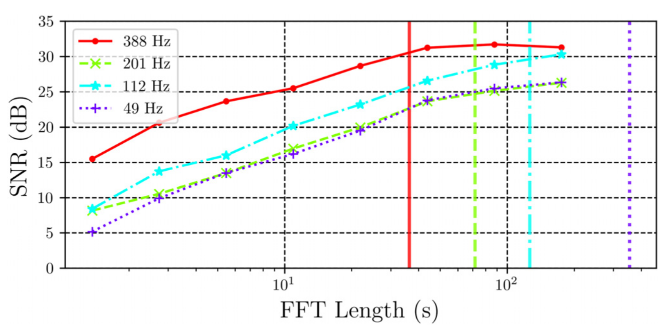

FIG. 4. (Color online) SNR estimates as a function of FFT length shown for four tones from the loud tonal set. The dashed-dotted lines show the theoretical coherence times τ calculated from Eq. (7) for the SWellEx environment with $v = -2.5$ m/s. For a coherent source, doubling the FFT length should increase SNR to $10\log_{10}2$ dB. Both 49 and 112 Hz SNR estimates follow this trend quite accurately. The SNRs for the higher frequencies follow a linear trend until they exceed their coherence time, at which point they taper off.

ignorant" matched field processors. For range-coherent processing, the bin is selected by estimating the Doppler shift from the known source frequency f_0 and the given search range-rate v. For example, the signal from source frequency f_0 will be associated with the bin closest to $f_0(1-v/c)$ for each search range-rate v (c is average modal group velocity). Range-coherent MFP's capacity to estimate source range-rate therefore adds gain at the outset by simply increasing the input SNR.

The covariance integration time is set to the time spent in a range cell of a 2.5 m/s source transmitting at the maximum frequency of 400 Hz: around 25 s.[36] For the 2048 point FFT sequence, 36 snapshots are used to form the SCM. For the 16 384 point FFT sequence, four snapshots are used to form the SCM. The synthetic element spacing T_{syn} is chosen to equal the covariance integration time. The result is two sequences of covariance matrices for each frequency, $\{\hat{\mathbf{K}}_i(f)\}_{2048}$ formed from simple snapshots and $\{\hat{\mathbf{K}}_i(f, N_{syn})\}_{16384}$ formed from "supervectors" of N_{syn} stacked snapshots.

An explicit example for computing $\hat{\mathbf{K}}_0(f,3)$ follows. Let $\{\tilde{\mathbf{d}}_i(f)\}$, denote the sequence of snapshots that results from applying the 16 384 point FFT to the VLA and selecting the bin corresponding to frequency f. First, form a sequence of supervectors $\{\tilde{\mathbf{d}}_i^{(3)}(f)\}$,

$$\tilde{\mathbf{d}}_i^{(3)}(f) = \left[\tilde{\mathbf{d}}_i^T(f)\ \tilde{\mathbf{d}}_{i+4}^T(f)\ \tilde{\mathbf{d}}_{i+8}^T(f)\right]^T. \quad (23)$$

Note that the supervector is formed by stacking every fourth snapshot. Then

$$\hat{\mathbf{K}}_0(f,3) = \frac{1}{4}\sum_{i=0}^{3}\tilde{\mathbf{d}}_i^{(3)}(f)\tilde{\mathbf{d}}_i^{(3)}(f)^\dagger. \quad (24)$$

C. Modeling

The KRAKENC normal mode program[37] is used to compute the steering vectors $\mathbf{s}(r,z)$ for the simple VLA. The steering vectors account for the known array tilt of 1°.[34] For a synthetic array with N_{syn} synthetic elements, the steering "supervector" $\mathbf{s}_{syn}$ is formed by stacking a collection of VLA replicas that depend on the search parameters r, z, v, and the time T_{syn} between snapshots stacked into the data supervector

$$\mathbf{s}_{syn}(r,z,v) = \begin{bmatrix} \mathbf{s}(r,z) \\ \mathbf{s}(r+vT_{syn},z) \\ \vdots \\ \mathbf{s}(r+(N_{syn}-1)vT_{syn},z) \end{bmatrix}. \quad (25)$$

As shown in Fig. 1, the bathymetry varies by almost 100 m over S5. Range-dependent modeling can account for this. However, the results of D'Spain *et al.* on mirages in shallow water[38] show that a range-independent model will still produce a significant correlation maximum, but with the location (range and depth) of maximum correlation shifted by an amount related to the error in assumed source depth. By modeling the bottom as a simple linear segment between the receiver and source, the range and depth of the correlation maximum will be distorted from the true source range and depth by a factor of $\gamma \equiv D/d_s$, where d_s is the true depth at the source location and D is the depth at the receiver. Therefore, the receiver depth is used as the depth of a range-independent normal mode model, and the factor γ is computed from the known bathymetry and used to correct the ambiguity surfaces (this same method is used by Booth *et al.*[39]).

VI. MFP RESULTS

Traditional MFP and range-coherent MFP were applied to three segments, labeled A, B, and C, of the approaching track (see Fig. 3). We restrict our analysis to these segments to avoid portions of the track during which the deep source was turned off. The ship approaches the VLA over the three segments, so segment A is the most distant segment and segment C is the closest segment.

For each tonal set, the ambiguity surfaces from each tone are summed incoherently as in Eq. (19). None of the implementations managed to consistently localize TSQ4, so we focus on TSL, TSQ1, TSQ2, and TSQ3 (see Table I for a review of the source level and SNR of the tonal sets).

To provide an example of "traditional" MFP, Bartlett and WNC processors were applied to the SCM sequences $\{K_i(f)\}_{2048}$. The result for each tonal set is a three-dimensional ambiguity surface over range, depth, and time for each of the two processors. After varying δ^2 for the WNC, we settled on using white noise gain of –2 dB for the "traditional" method.

Then Bartlett and WNC processors were applied to the range-coherent SCM sequences $\{K_i(f,N_{syn})\}_{16384}$ for various values of N_{syn} to investigate range-coherent MFP performance. For the range-coherent results, a four-dimensional Bartlett ambiguity surface over source parameter vector $\boldsymbol{\theta} = (r\ z\ v)^T$ and time t_i was formed for each tonal set

$$P_B(\boldsymbol{\theta},t_i) = \frac{1}{13}\sum_{k=1}^{13}10\log_{10}P_{BART}\left(\hat{\mathbf{K}}_i(f_k',N_{syn}),\boldsymbol{\theta}\right), \quad (26)$$

where $f_k' = f_k(1-v/c)$ is the received frequency of the kth source tone f_k in the tonal set. The range-coherent Bartlett estimate is $\hat{\boldsymbol{\theta}}(t) = \mathrm{argmax}_{\boldsymbol{\theta}}P_B(\boldsymbol{\theta},t)$. The range-rate/time ambiguity surfaces for the different tonal sets are shown in Fig. 5.

Then, for various choices of the white noise gain δ^2, a three-dimensional WNC surface is formed over $\boldsymbol{\theta} = (r\ z\ \hat{v}(t_i))^T$ and time t_i,

$$P_W(\boldsymbol{\theta},t_i) = \frac{1}{13}\sum_{k=1}^{13}10\log_{10}P_{WNC}\left(\hat{\mathbf{K}}_i(f_k',N_{syn}),\boldsymbol{\theta}\right). \quad (27)$$

The search range-rate is constrained to the Bartlett estimate $\hat{v}$. All presented results for range-coherent processing use a white noise gain of –0.5 dB, which seemed to work best.

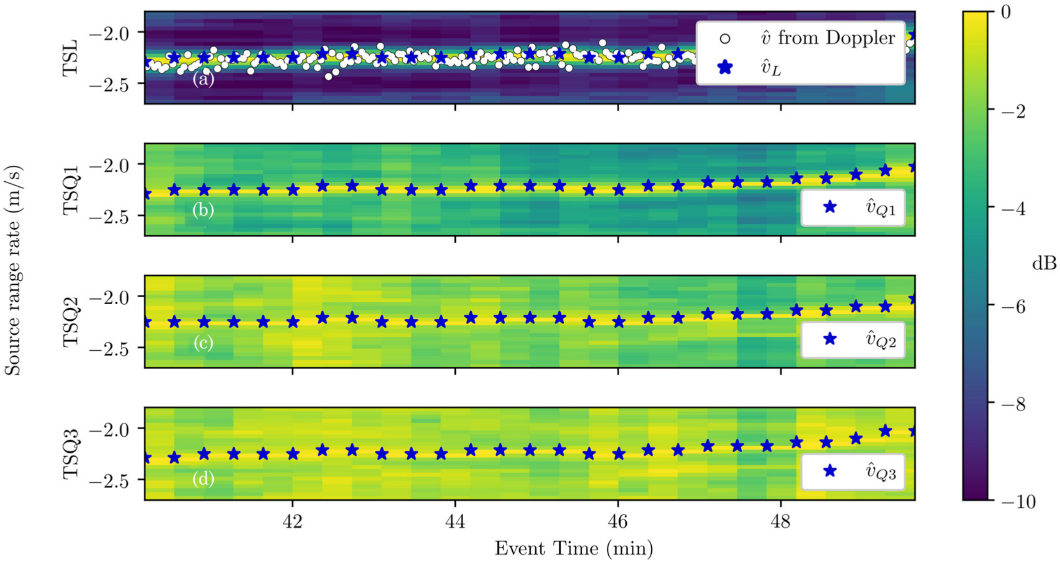

FIG. 5. (Color online) Multi-tone Bartlett range-rate/time ambiguity surfaces for segment C for a five synthetic element array ($N_{syn} = 5$). The results using TSL are shown in (a), TSQ1 in (b), TSQ2 in (c), and TSQ3 in (d). In each figure, a blue star marks the location of the maximum of the ambiguity surface for each time step. In (a), range-rate estimates from Doppler shifts of TSL are superimposed. In (b)–(d), the SNR is too low to estimate range-rate from Doppler, so only v_{Qi}, estimates from range-coherent MFP applied to TSQi, are shown.

The number of synthetic elements N_{syn} that produced the best results depended on the experimental segment. In segment C, the proximity to the CPA created a non-uniform range-rate that limited the synthetic aperture length to around 300 m ($N_{syn} = 5$). In segment A and segment B, the range-rate is more uniform. As such, range-coherent MFP worked on synthetic aperture lengths up to 600 m ($N_{syn} = 10$).

A representative comparison of ambiguity surfaces for the four implementations is shown in Fig. 6 for TSQ3. The signal level is too low for the simple Bartlett or WNC to localize the source. In contrast, range-coherent WNC accurately estimates the source range and depth while also providing an accurate range-rate estimate.

To present comparisons over time graphically, we form two-dimensional ambiguity surfaces by holding time and a single parameter fixed and take the maximum over the remaining parameters. For example, to form the range-rate/time ambiguity surface shown in Fig. 5, take the maximum over range and depth for each range-rate and time value. Doppler shift range-rate estimates from TSL agree with the range-coherent Bartlett range-rate estimates for each tonal set. The quiet tonal sets are too quiet for Doppler shift

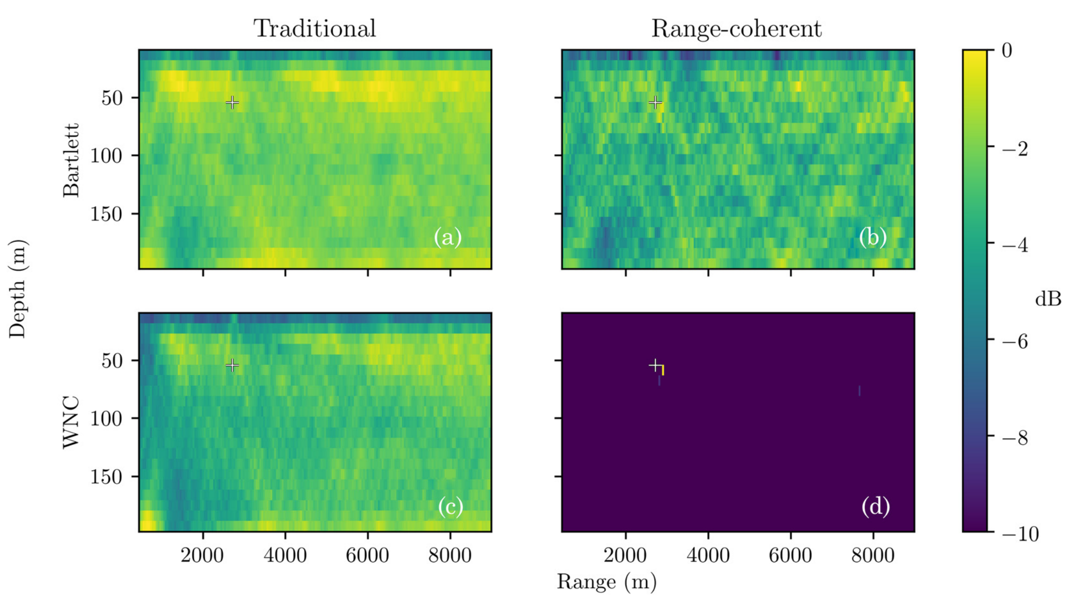

FIG. 6. (Color online) Comparison of benchmark and range-coherent ambiguity surfaces for a sample interval for TSQ3. (a) and (c) show the Bartlett and WNC results on a simple VLA. (b) and (d) show the Bartlett and WNC range-coherent MFP for a five synthetic element array ($N_{syn} = 5$). The cross-shows the ship range from GPS and nominal source depth of 55 meters.

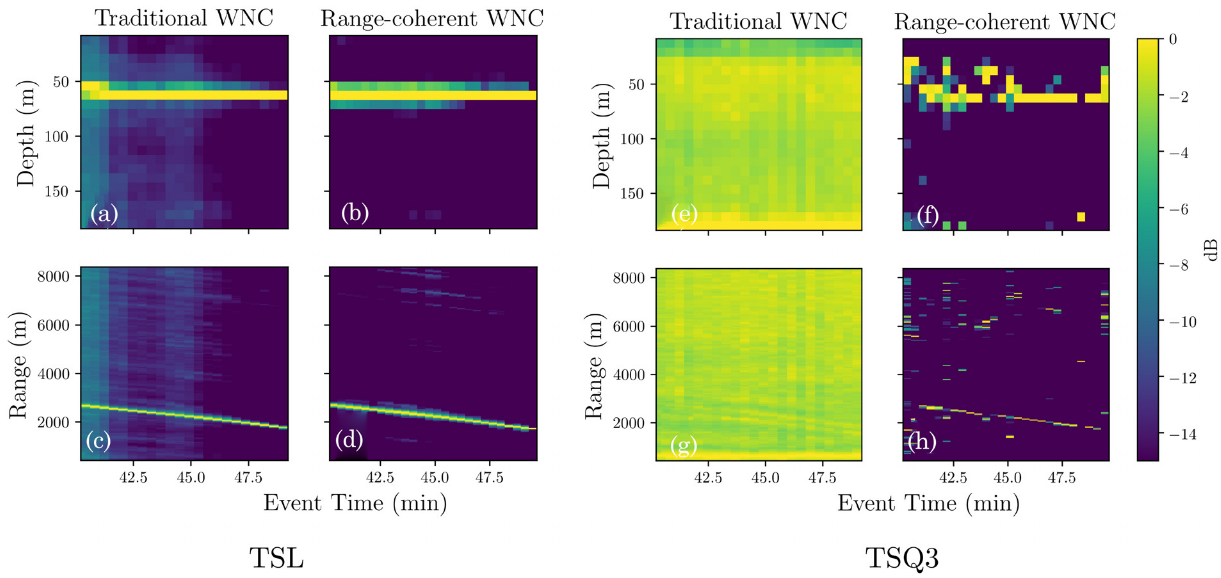

FIG. 7. (Color online) Comparing "traditional" WNC to range-coherent WNC. Tracks are shown for segment C for TSL (left panel) and TSQ3 (right panel) tonal sets. In the top row, (a), (b), (e), and (f) show depth/time ambiguity surfaces. In the bottom row, (c), (d), (g), and (h) show range/time ambiguity surfaces. (a), (c), (e), and (g) show "traditional" WNC. (b), (d), (f), and (h) show range-coherent WNC with five synthetic elements. Both methods localize the source over the whole segment on TSL. Only range-coherent WNC localizes the source using TSQ3.

estimation, so the only estimates of range-rate come from range-coherent MFP.

We also compare range/time and depth/time ambiguity surfaces to assess the relative performance of the four implementations over the segments of the experiment. A comparison between range/time and depth/time ambiguity surfaces over segment C for traditional WNC and the range-coherent WNC with five synthetic elements is shown for TSL and TSQ3 in Fig. 7. Both processors easily localize the source over the full segment using TSL. For TSQ3, traditional WNC completely breaks down, but range-coherent WNC produces a nice track. In fact, traditional WNC fails on all three segments using any of the quiet tonal sets. We show the range-coherent result for TSQ3 because it was the lowest set that produced a clear source track.

Range-coherent WNC range/time and depth/time ambiguity surfaces for all three segments for TSQ1 are shown in Fig. 8. For segments A and B, the synthetic array consists of

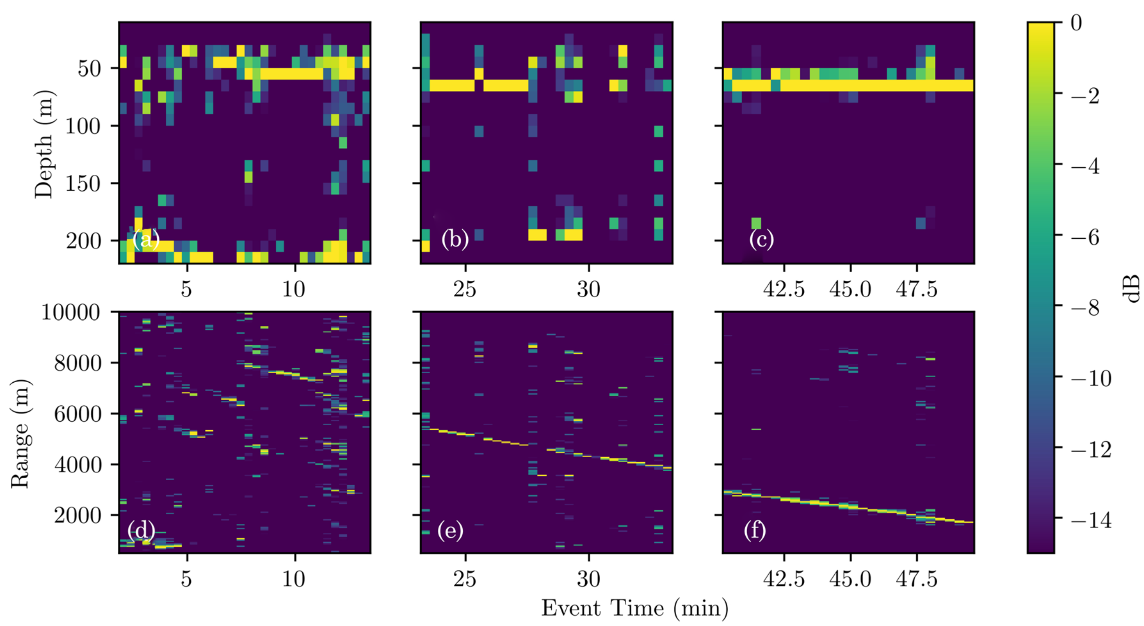

FIG. 8. (Color online) Range-coherent WNC depth/time and range/time ambiguity surfaces for TSQ1 for all three data segments. Segment C, shown in (c) and (f), uses a five-element synthetic array. Segments A and B, shown in (a) and (d) and in (b) and (e) use an eight-element synthetic array. Further from the closest-point of approach (CPA), constant velocity translates to a nearly uniform range-rate. This permits formation of a longer synthetic aperture in segments A and B than in segment C.

eight elements, whereas for segment C, the synthetic array had only five elements. The ship track in Fig. 3 shows the slight curvature of segment C. The curvature manifests as a non-uniform range-rate. The assumption of uniform range-rate limits the synthetic array length to intervals over which the acceleration creates minimal mismatch. The uniform range-rate at greater range permitted the longer synthetic aperture used in segments A and B.

Range-coherent MFP demonstrates superior performance to the benchmark MFP implementations on the quiet tonal sets. Using a 500 m synthetic aperture, range-coherent WNC localizes TSQ1 out to a range of 8 km and TSQ2 out to a range of 5 km.

VII. CONCLUSION

Range-coherent MFP can localize quieter sources than traditional methods. It does this by combining MFP with a passive synthetic aperture that provides both spatial array gain and gain from increased integration time. In the presented application to the SWellEx-96 data, these two features result in a total gain of roughly 16 dB over previous MFP results.

The assumption of uniform source motion limits the maximum synthetic aperture length. Future work could incorporate a higher order motion model, for example, by applying the array shape matched field inversion of Hodgkiss *et al.*[40] to the synthetic array, or could minimize the penalty for mismatch by incorporating an array geometry robustness constraint like that used in Byun *et al.*[41] Fusing velocity, range, and depth measurements produced by range-coherent MFP in a tracking algorithm could further improve results.

ACKNOWLEDGMENTS

This work was funded by the Task Force Ocean (TFO) program of the United States Office of Naval Research. Ryan Saenger, Ludovic Tenorio, and Paul Akins provided helpful comments on an early draft.

[1]A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," J. Acoust. Soc. Am. **83**(2), 571–587 (1988).
[2]S. W. Autrey, "Passive synthetic arrays," J. Acoust. Soc. Am. **84**(2), 592–598 (1988).
[3]N. Yen and W. Carey, "Application of synthetic-aperture processing to towed-array data," J. Acoust. Soc. Am. **86**(2), 754–765 (1989).
[4]S. Stergiopoulos and E. J. Sullivan, "Extended towed array processing by an overlap correlator," J. Acoust. Soc. Am. **86**(1), 158–171 (1989).
[5]E. J. Sullivan, J. D. Holmes, W. M. Carey, and J. F. Lynch, "Broadband passive synthetic aperture: Experimental results," J. Acoust. Soc. Am. **120**(4), EL49–EL54 (2006).
[6]B. A. Tan, P. Gerstoft, C. Yardim, and W. S. Hodgkiss, "Broadband synthetic aperture geoacoustic inversion," J. Acoust. Soc. Am. **134**(1), 312–322 (2013).
[7]J. Unpingco, W. A. Kuperman, W. S. Hodgkiss, and R. Hecht-Nielsen, "Single sensor source tracking and environmental inversion," J. Acoust. Soc. Am. **106**(3), 1316–1329 (1999).
[8]R. E. Williams, "Creating an acoustic synthetic aperture in the ocean," J. Acoust Soc. Am. **60**(1), 60–73 (1976).
[9]T. C. Yang, "Source depth estimation based on synthetic aperture beamfoming for a moving source," J. Acoust. Soc. Am. **138**(3), 1678–1686 (2015).
[10]T. C. Yang, "Data-based matched-mode source localization for a moving source," J. Acoust. Soc. Am. **135**(3), 1218–1230 (2014).
[11]G. V. Frisk, K. M. Becker, and, and J. A. Doutt, "Modal mapping in shallow water using synthetic aperture horizontal arrays," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158)*, Providence, RI (September 11–14, 2000), Vol. 1, pp. 185–188.
[12]S. Stergiopoulos and H. Urban, "An experimental study in forming a long synthetic aperture at sea," IEEE J. Oceanic Eng. **17**(1), 62–72 (1992).
[13]Z. Michalopoulou and M. B. Porter, "Matched-field processing for broadband source localization," IEEE J. Ocean. Eng. **21**(4), 384–392 (1996).
[14]G. J. Orris, M. Nicholas, and J. S. Perkins, "The matched-phase coherent multi-frequency matched-field processor," J. Acoust. Soc. Am. **107**(5), 2563–2575 (2000).
[15]Z. Michalopoulou, "Robust multi-tonal matched-field inversion: A coherent approach," J. Acoust. Soc. Am. **104**(1), 163–170 (1998).
[16]C. Debever and W. A. Kuperman, "Robust matched-field processing using a coherent broadband white noise constraint processor," J. Acoust. Soc. Am. **122**(4), 1979–1986 (2007).
[17]L. M. Zurk, N. Lee, and, and J. Ward, "3D adaptive matched field processing for a moving source in a shallow water channel," in *Oceans '99. MTS/IEEE. Riding the Crest into the 21st Century. Conference and Exhibition. Conference Proceedings (IEEE Cat. No.99CH37008)*, Seattle, WA (September 13–16, 1999), Vol. 2, pp. 728–731.
[18]L. M. Zurk, N. Lee, and J. Ward, "Source motion mitigation for adaptive matched field processing," J. Acoust. Soc. Am. **113**(5), 2719–2731 (2003).
[19]F. Li, F. Zhu, Y. Zhang, B. Zhang, W. Li, and W. Luo, "Synthetic adaptive matched field processing for moving source with a horizontal line array," J. Acoust. Soc. Am. **149**(2), 1138–1146 (2021).
[20]H. Bucker, "Matched-field tracking in shallow water," J. Acoust. Soc. Am. **96**(6), 3809–3811 (1994).
[21]L. T. Fialkowski, J. S. Perkins, M. D. Collins, M. Nicholas, J. A. Fawcett, and W. A. Kuperman, "Matched-field source tracking by ambiguity surface averaging," J. Acoust. Soc. Am. **110**(2), 739–746 (2001).
[22]A. N. Guthrie, R. M. Fitzgerald, D. A. Nutile, and J. D. Shaffer, "Long-range low-frequency CW propagation in the deep ocean Antigua-Newfoundland," J. Acoust. Soc. Am. **56**(1), 58–69 (1974).
[23]F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (Springer, New York, 2011).
[24]K. E. Hawker, "A normal mode theory of acoustic doppler effects in the oceanic waveguide," J. Acoust. Soc. Am. **65**(3), 675–681 (1979).
[25]H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," J. Acoust. Soc. Am. **59**(2), 368–373 (1976).
[26]H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," IEEE Trans. Acoust. Speech Signal Process. **35**(10), 1365–1376 (1987).
[27]J. N. Maksym, "A robust formulation of an optimum cross-spectral beamformer for line arrays," J. Acoust. Soc. Am. **65**(4), 971–975 (1979).
[28]B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," IEEE Trans. Aerosp. Electron. Syst. **24**(4), 397–401 (1988).
[29]K. Hsu and A. B. Baggeroer, "Application of the maximum-likelihood method (MLM) for sonic velocity logging," Geophysics **51**(3), 780–787 (1986).
[30]H. Cox, "Multi-rate adaptive beamforming (MRABF)," in *Proceedings of the 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop. SAM 2000 (Cat. No.00EX410)*, Cambridge, MA (March 17, 2000), pp. 306–309.
[31]I. Kirsteins and D. Tufts, "On the probability density of signal-to-noise ratio in an improved adaptive detector," in *ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tampa, FL (April 26–29, 1985), Vol. 10, pp. 572–575.
[32]S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory, Fundamentals of Statistical Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[33]P. A. Baxley, N. O. Booth, and W. S. Hodgkiss, "Matched-field replica model optimization and bottom property inversion in shallow water," J. Acoust. Soc. Am. **107**(3), 1301–1323 (2000).
[34]J. Murray and D. Ensberg, "The SWellEx-96 experiment," http://swellex96.ucsd.edu/ (Last viewed June 6, 2021).
[35]R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water off San Diego, California," J. Acoust. Soc. Am. **99**(4), 2077–2085 (1996).
[36]A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No. CH37020)*, Pacific Grove, CA (October 24–27, 1999). Vol. 1, pp. 103–108.
[37]M. B. Porter, "The KRAKEN normal mode program," Report SM-245, SACLANT Undersea Research Centre, La Spezia, Italy, 1991.
[38]G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, N. O. Booth, and P. W. Schey, "Mirages in shallow water matched field processing," J. Acoust. Soc. Am. **105**(6), 3245–3265 (1999).
[39]N. O. Booth, A. T. Abawi, P. W. Schey, and W. S. Hodgkiss, "Detectability of low-level broad-band signals using adaptive matched-field processing with vertical aperture arrays," IEEE J. Ocean. Eng. **25**(3), 296–313 (2000).
[40]W. S. Hodgkiss, D. E. Ensberg, J. J. Murray, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Direct measurement and matched-field inversion approaches to array shape estimation," IEEE J. Ocean. Eng. **21**(4), 393–401 (1996).
[41]G. Byun, F. Hunter Akins, K. L. Gemba, H. C. Song, and W. A. Kuperman, "Multiple constraint matched field processing tolerant to array tilt mismatch," J. Acoust. Soc. Am. **147**(2), 1231–1238 (2020).
[42]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0005586 for an analysis of the relationship between range-coherent matched field processing and matched field tracking.