Robust Ocean Acoustic Localization With Sparse Bayesian Learning

Kay L. Gemba, Santosh Nannuru, and Peter Gerstoft, *Senior Member, IEEE*

***Abstract*—Matched field processing (MFP) compares the measures to the modeled pressure fields received at an array of sensors to localize a source in an ocean waveguide. Typically, there are only a few sources when compared to the number of candidate source locations or range-depth cells. We use sparse Bayesian learning (SBL) to learn a common sparsity profile corresponding to the location of present sources. SBL performance is compared to traditional processing in simulations and using experimental ocean acoustic data. Specifically, we localize a quiet source in the presence of a surface interferer in a shallow water environment. This multi-frequency scenario requires adaptive processing and includes modest environmental and sensor position mismatch in the MFP model. The noise process changes likely with time and is modeled as a non-stationary Gaussian process, meaning that the noise variance changes across snapshots. The adaptive SBL algorithm models the complex source amplitudes as random quantities, providing robustness to amplitude and phase errors in the model. This is demonstrated with experimental data, where SBL exhibits improved source localization performance when compared to the white noise gain constraint (–3 dB) and Bartlett processors.**

***Index Terms*—Robust beamforming, sparse Bayesian learning, matched field processing, non-stationary noise, sparse reconstruction, array processing.**

I. Introduction

WITH long observation times weak signals can be extracted in a noisy environment. Most analytic treatments analyze these cases assuming Gaussian noise with constant variance. However, for long observation times, the noise process though is likely to change with time. This means that the noise variance is non-stationary in time (taken across snapshots), and/or even across sensors. Modeling of these nuisance parameters in the processing is useful to improve the signal estimate.

Accounting for the noise variation is certainly useful in statistical signal processing. For example, the noise has been assumed to vary spatially [1] and spatiotemporally [2]. This has lead to so-called lucky imaging in astronomy [3] or lucky ranging in ocean acoustics [4], where only the measurements giving good imaging results are used. In contrast, we use all measurements either by estimating the noise per snapshot or by normalizing the data on a per snapshot basis, demonstrated here for the matched field processing application.

Matched field processing (MFP) is a generalized beamforming method which matches received array data to a dictionary of replica vectors to localize one or several sources [5]–[10]. Source localization parameters of interest include both range and depth in an ocean acoustic waveguide. Modeling the transfer function (e.g., the propagating modes) between a candidate source position and elements deployed as a vertical line array (VLA) requires regional knowledge of environmental parameters such as water column sound speed, seabed depth, and geoacoustic parameters.

On the one hand, field complexity is desirable because it increases localization ability [11, p.86], [12, p.572]. On the other hand, field diversity increases model complexity, often beyond our knowledge for dynamic ocean environments. In such situations, one may choose to model a range-dependent waveguide as range-independent. For example, bottom depth or sound speed change in range and are approximated with a flat bottom and an average sound speed. Further, localization ambiguity decreases if multiple frequencies (below 1 kHz) are processed. Since the source function (i.e., phase between frequencies) is not know [13], localization results for each center frequency usually are combined incoherently. In contrast, it is still possible to estimate range [14], depth [15] or source bearing [16] independently without environmental information.

Traditionally, the Bartlett processor [17] is used as a point of departure to estimate MFP source location. High resolution localization using the minimum variance distortionless response (MVDR) processor is not practical due to encountered (e.g., environmental) mismatch (through-the-sensor environmental characterization are an active area of research [18]–[20]). The adaptive white noise gain constraint [21] (WNGC) processor is more versatile because it can adjust its behavior (thus resolution and sidelobe suppression) from Bartlett to MVDR [22].

The MFP localization problem can be reformulated as an underdetermined system of linear equations and source parameters are estimated using compressive sensing (CS) [23]–[26]. When implemented using basis pursuit (BP), CS possesses properties similar to an adaptive processor and offers localization improvement compared to the WNGC processor for single- and

Manuscript received June 29, 2018; revised November 11, 2018 and February 12, 2019; accepted February 16, 2019. Date of publication February 22, 2019; date of current version April 11, 2019. This work was supported by the Office of Naval Research under Grant N00014-18-1-2118. The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Peter. K. Willet. *(Corresponding author: Kay L. Gemba.)*

K. L. Gemba is with the Marine Physical Laboratory of the Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0238 USA (e-mail: gemba@ucsd.edu).

S. Nannuru is with the Signal Processing and Communications Research Center, International Institute of Information Technology Hyderabad, Hyderabad 500032, India (e-mail: santosh.nannuru@iiit.ac.in).

P. Gerstoft is with the Noise Laboratory of the Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0238 USA (e-mail: gerstoft@ucsd.edu).

Digital Object Identifier 10.1109/JSTSP.2019.2900912

multiple-sources [27]. CS has demonstrated improved performance in the presence of perturbations or mismatch [27]–[29].

However, CS implemented using the non-greedy BP method has a significant shortcoming in data applications: for K unknown sources, the number of sparse solutions K_s required to localize all sources is $K_s \geq K$ [27]. Additional ambiguous solutions are due to mismatch between the observed field at the VLA and the modeled replica vector.

An alternative CS implementation, sparse Bayesian learning (SBL) [30]–[33], offers relief to this shortcoming for multi-frequency MFP [34]. SBL has the advantage that it determines sparsity automatically. Here, SBL learns a sparsity profile (common across snapshots and frequencies) corresponding to the location of present sources.

Following the CS approach, SBL also reformulates the parameter estimation problem as an underdetermined linear problem. The variables are treated as Gaussian random vectors and Bayesian evidence maximization is performed to obtain a sparse solution [32]. SBL can be interpreted as iterative, re-weighted BP [35] and exhibits similar sparse signal recovery compared to the BP method [36]. BP's execution time increases quadratically with number of snapshots while SBL is nearly snapshot independent [33], [34]. SBL can be viewed as a stochastic maximum likelihood approach, it has the same objective function as LIKES [37].

This paper compares processor localization performance using simulated and SWellEx-96 data subject to environmental and sensor location mismatch. Further, we implement SBL with two noise models for scenarios with stationary and non-stationary temporal noise statistics. Previously [34], we compared SBL to the non-adaptive Bartlett and adaptive WNGC processors in a two-source localization scenario.

The two-source scenario required adaptive processing to identify the location of a weaker source in the presence of a stronger source. Further, robust adaptive processing is desirable because environmental parameters, sensor locations, or the noise process are generally unknown. Robust to mismatch means the beamformer is less sensitive to small amplitude or phase errors [21].

In this manuscript, using a similar two-source scenario, we process another event than [34] and localize the quiet submerged source in the presence of an uncontrolled, broadband surface interferer (a passing ship).

A. Matched Field Problem Formulation

For the lth observation snapshot, we assume the linear model

$$\mathbf{y} = \sum_K \mathbf{a}(\boldsymbol{\theta}_K) x_K + \mathbf{n}. \tag{1}$$

The snapshot $\mathbf{y}_l \in \mathbb{C}^N$ consists of a vector of Fourier coefficients at a single frequency f obtained via a fast Fourier transform (FFT) of the lth data segment from N receivers. The replica vector $\mathbf{a}(\boldsymbol{\theta}) \in \mathbb{C}^N$ is the Green's function (computed with a normal mode code) for a candidate source position $\boldsymbol{\theta}$ (i.e., range and depth) to each receiver. The complex source amplitude for the Kth source is x. Further, $\mathbf{n}_l \in \mathbb{C}^N$ is additive zero-mean circularly symmetric complex Gaussian noise, which is generated from a Gaussian process $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{n}_l; \mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{n}_l})$. Due to the circular symmetry of the noise the phase is uniformly distributed.

To localize a source, the waveguide is discretitized to M_d depths and M_r ranges, for $M = M_d \times M_r$ candidate source positions. The dictionary $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_M] \in \mathbb{C}^{N\times M}$ contains all M replica vectors as columns.

We observe narrowband waves on N sensors for L snapshots $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L] \in \mathbb{C}^{N\times L}$. Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_L] \in \mathbb{C}^{M\times L}$ be the unknown complex source amplitudes for a total of M range-depth positions. Only a few sources K are typically present in practical applications (hence $K \ll M$). The source amplitudes x_K in (1) form the non-zero entries of the sparse vector $\mathbf{x}_l$. These sources are assumed to be spatially stationary across snapshots. A linear regression model relates the array data $\mathbf{Y}$ to the source amplitudes $\mathbf{X}$ as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}. \tag{2}$$

This system of equations typically is underdetermined ($N \ll M$). The noise $\mathbf{N} = [\mathbf{n}_1, \ldots, \mathbf{n}_L] \in \mathbb{C}^{N\times L}$ is in general non-stationary across snapshots.

Under these assumptions, (2) is solved using ℓ_2 regularization leading to Bartlett and WNGC processors (Sec. II). Using ℓ_1 regularization leads to the sparsity promoting LASSO solution, briefly discussed in Sec. III-C. The SBL approach to solving (2) in a Bayesian framework is the focus of this paper and discussed in Sec. III.

B. Noise Model

The noise is modeled as a diagonal covariance matrix, parameterized as

$$\boldsymbol{\Sigma}_{\mathbf{n}_l} = \sigma_l^2 \mathbf{I}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix. Note that the covariance matrices $\boldsymbol{\Sigma}_{\mathbf{n}_l}$ are varying over the snapshot index $l = 1, \ldots, L$ and $\sigma_l \in \mathbb{R}_{0+}$ where $\mathbb{R}_{0+}$ denotes non-negative real numbers.

We consider two special cases for the a priori knowledge on the noise covariance model (3) as follows:

Case I: We assume wide-sense stationarity of the noise in space and time: $\sigma_l^2 = \sigma^2, \forall l$. This is a commonly used noise model, i.e. stationary white noise.

Case II: We assume wide-sense stationarity of the noise in space only. The noise variance for all sensor elements is equal across the array and it varies over snapshots, i.e., non-stationary white noise.

A Case III [2], in which the noise statistics vary across both time and space, was considered. Results did not indicate a significant improvement over Case II for this data set and are therefore not presented. However, Case III may be of interest when sensors are deployed close to the sea surface [38], [39]. While focusing on narrowband (single frequency) observations, we emphasize that the noise varies across frequencies.

C. Data Normalization

An alternate approach to model non-stationary noise explicitly as in Sec. I-B is to perform snapshot normalization. Let us introduce the factorization $\boldsymbol{\Sigma}_{\mathbf{n}_l}^{-1} = \mathbf{W}_l^H \mathbf{W}_l$, where $\mathbf{W}_l$ is a square and non-singular matrix. For our particular setup, we have $\mathbf{W}_l = \sigma_l^{-1}\mathbf{I}$. The matrix $\mathbf{W}_l$ is useful for normalizing the sensor data. The corresponding normalized sensor data and noise are

$$\widetilde{\mathbf{y}}_l = \mathbf{W}_l \mathbf{y}_l, \quad \widetilde{\mathbf{n}}_l = \mathbf{W}_l \mathbf{n}_l. \tag{4}$$

For known diagonal noise covariance $\mathbf{\Sigma}_{\mathbf{n}_l}$ the above means we have to normalize (1) with σ_l as then the resulting noise satisfies $\widetilde{\mathbf{n}}_l \sim \mathcal{C}N(\widetilde{\mathbf{n}}_l; \mathbf{0}, \mathbf{I})$, and thus all entries are identically distributed.

In practice, we do not have access to the true noise variances σ_l. For low SNR scenario, i.e. when the noise $\mathbf{n}_l$ is significantly larger than the signal $\mathbf{a}(\theta)x_l$, we can approximate the data to be equal to the noise leading to the approximation $\sigma_l \approx \|\mathbf{y}_l\|_2$. We consider two cases:

$$\mathbf{W}_l \approx \begin{cases} \frac{\sqrt{NL}}{\|\mathbf{Y}\|_F}\mathbf{I} & \text{for Case I} \\ \frac{\sqrt{N}}{\|\mathbf{y}_l\|_2}\mathbf{I} & \text{for Case II,} \end{cases} \tag{5}$$

leading to

$$\widetilde{\mathbf{y}}_l \approx \begin{cases} \frac{\sqrt{NL}}{\|\mathbf{Y}\|_F}\mathbf{y}_l & \text{for Case I} \\ \frac{\sqrt{N}}{\|\mathbf{y}_l\|_2}\mathbf{y}_l & \text{for Case II,} \end{cases} \tag{6}$$

where $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L]$ is the matrix of all data snapshots and $\|\mathbf{Y}\|_F$ denotes its Frobenius norm. We refer to Case I as 'not normalized' and Case II as 'normalized' since $\widetilde{\mathbf{y}}_l$ has unit norm.

II. Bartlett and WNGC Processors

Bartlett is a spatial matched-filter processor which matches replica vectors $\mathbf{a}(\boldsymbol{\theta})$ to the data $\mathbf{y}$ (1):

$$P_B(\boldsymbol{\theta}) = \mathbf{a}^H(\boldsymbol{\theta})\mathbf{S}_{\mathbf{y}}\mathbf{a}(\boldsymbol{\theta}), \tag{7}$$

where the superscript H denotes the Hermitian operator and the sample covariance matrix (SCM) $\mathbf{S}_{\mathbf{y}} \in \mathbb{C}^{N\times N}$ is obtained using L snapshots:

$$\mathbf{S}_{\mathbf{y}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}_l\mathbf{y}_l^H. \tag{8}$$

$P_B(\boldsymbol{\theta})$ denotes the Bartlett power at position $\boldsymbol{\theta}$ using normalized replicas vectors (i.e., $||\mathbf{a}(\boldsymbol{\theta})||_2 = 1$). Processor output power is plotted for for a total of M candidate positions $\boldsymbol{\theta}$, conventionally arranged in range and depth and referred to as an ambiguity surface. While Bartlett does not invert $\mathbf{S}_{\mathbf{y}}$ and thus does not have a minimum number of required snapshots, it suffers from high sidelobes. Sidelobe suppression is important if a combination of sources (or a combination of sources and interferers) are present.

The WNGC processor P_{wngc} discriminates against other sources/interferers while offering a degree of robustness in frequently encountered mismatch scenarios. The WNGC is considered versatile because of its ability to adjust its behavior (thus resolution and sidelobe suppression) from Bartlett to MVDR [22] at the expense of inverting $\mathbf{S}_{\mathbf{y}}$ (8). To have $\mathbf{S}_{\mathbf{y}}$ invertible, we require $L \geq N$ (diagonal loading of $\mathbf{S}_{\mathbf{y}}$ can mitigate this requirement). The WNGC processor is:

$$P_{wngc}(\boldsymbol{\theta}) = \mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{S}_{\mathbf{y}}\mathbf{a}_w(\boldsymbol{\theta}),$$
$$\text{where } \mathbf{a}_w(\boldsymbol{\theta}) = \frac{(\mathbf{S}_{\mathbf{y}} + \epsilon\mathbf{I})^{-1}\mathbf{a}(\boldsymbol{\theta})}{\mathbf{a}^H(\boldsymbol{\theta})(\mathbf{S}_{\mathbf{y}} + \epsilon\mathbf{I})^{-1}\mathbf{a}(\boldsymbol{\theta})}. \tag{9}$$

The adaptive weights $\mathbf{a}_w(\boldsymbol{\theta})$ correspond to diagonally loaded MVDR weights and are obtained by solving:

$$\begin{aligned} &\min_{\mathbf{a}_w} \mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{S}_{\mathbf{y}}\,\mathbf{a}_w(\boldsymbol{\theta}) \text{ subject to} \\ &\mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{a}(\boldsymbol{\theta}) = 1, \\ &\left|\mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{a}_w(\boldsymbol{\theta})\right|^{-1} \geq \delta^2, \end{aligned} \tag{10}$$

for each replica vector at position $\boldsymbol{\theta}$. The constraining value δ^2 imposes a gain constraint on the adaptive weights and the white noise gain constraint G_{wngc} such that

$$\delta^2 \leq G_{wng} = \left|\mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{a}_w(\boldsymbol{\theta})\right|^{-1} < N, \tag{11}$$

which in practice is normalized and expressed as $10\log_{10}(\delta^2/N) \leq 0$ dB. The selected constraining value δ^2 determines ϵ in (9) and hence WNGC output power.

An initialization value ϵ_0 for this optimization is subject to the inequality constraint and is selected conservatively to ensure that $\left|\mathbf{a}_w^H(\boldsymbol{\theta})\mathbf{a}_w(\boldsymbol{\theta})\right|_{\epsilon_0}^{-1} < \delta^2$. Since $\epsilon(\boldsymbol{\theta})$ can span many orders of magnitude, ϵ is parameterized on a dB scale. The search uses a singular value decomposition of $\mathbf{S}_{\mathbf{y}}$ with $\epsilon_0 = 10\log_{10}(\mathrm{tr}[\mathbf{S}_{\mathbf{y}}]/N) - 30$, where $\mathrm{tr}[\cdot]$ denotes the trace of a matrix. The iterative algorithm converges when a selected constraint is satisfied within ± 0.1 dB. Thus, $P_{wngc}(\boldsymbol{\theta})$ denotes the WNGC power at position $\boldsymbol{\theta}$ for a selected (white noise gain) constraint, which frequently falls within [–6 –2] dB.

To localize a source, Eqs. (7) and (9) are evaluated at M range-depth candidate positions $\boldsymbol{\theta}$. Computed ambiguity surfaces for F processed frequencies are averaged:

$$P^F(\boldsymbol{\theta}) = \sum_{f=1}^{F} P(\boldsymbol{\theta}, f). \tag{12}$$

Processing additional frequencies improves source localization performance for a weaker source in the presence of a stronger source and environmental or model mismatch. The mainlobes are in the same locations while the sidelobes are not. It is desirable that these frequencies span multiple octaves which may increase the sidelobe diversity in (12).

III. Sparse Bayesian Learning

In this section we discuss our SBL approach and develop a multi-frequency solution to (2) from a Bayesian point of view. Starting from (2), the complex source amplitudes $\mathbf{x}_l$ and the noise $\mathbf{n}_l$ are assumed independent with each other and across snapshots. The data likelihood is then given by

$$p(\mathbf{Y}|\mathbf{X}) = \prod_{l=1}^{L} p(\mathbf{y}_l|\mathbf{x}_l) = \prod_{l=1}^{L} \mathcal{C}N(\mathbf{y}_l; \mathbf{A}\mathbf{x}_l, \mathbf{\Sigma}_{\mathbf{n}_l}). \tag{13}$$

A. Prior on the Sources

We assume that the complex source amplitudes x_{ml} are independent both across snapshots (i.e. index l) and across space (i.e. index m) and follow a zero-mean circularly symmetric complex Gaussian distribution with space-dependent variance γ_m, $m = 1, \ldots, M$,

$$p(x_{ml}; \gamma_m) = \begin{cases} \delta(x_{ml}), & \text{for } \gamma_m = 0 \\ \frac{1}{\pi \gamma_m} \mathrm{e}^{-|x_{ml}|^2/\gamma_m}, & \text{for } \gamma_m > 0 \end{cases}, \tag{14}$$

$$p(\mathbf{X}; \boldsymbol{\gamma}) = \prod_{l=1}^{L} \prod_{m=1}^{M} p(x_{ml}; \gamma_m) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{x}_l; \mathbf{0}, \boldsymbol{\Gamma}), \tag{15}$$

i.e., the source vector $\mathbf{x}_l$ at each snapshot $l \in \{1, \ldots, L\}$ is multivariate Gaussian with covariance matrix,

$$\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma}) = \mathsf{E}[\mathbf{x}_l \mathbf{x}_l^H; \boldsymbol{\gamma}], \ \boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_M]. \tag{16}$$

When the variance $\gamma_m = 0$, then $x_{ml} = 0$ with probability 1 implying that there is no source at the range-depth position corresponding to index m. Assuming the sources to be stationary across snapshots, the sparsity profile (i.e. the location of non-zero entries in $\mathbf{x}_l$) is the same across snapshots. The sparsity of the model is thus controlled with the parameter $\boldsymbol{\gamma}$, and the working set $\mathcal{M}$ is equivalently

$$\mathcal{M} = \{m \in \mathbb{N} | \gamma_m > 0\}. \tag{17}$$

The covariance matrix $\boldsymbol{\Gamma}$ can potentially be singular as $\mathrm{rank}(\boldsymbol{\Gamma}) = \mathrm{card}(\mathcal{M}) = K \leq M$. Since $\boldsymbol{\Gamma}$ is a diagonal matrix, its rank is equal to the number of non-zero diagonal entries (and there are K such entries each of which correspond to a true source). The SBL algorithm ultimately estimates $\boldsymbol{\gamma}$ rather than the complex source amplitudes $\mathbf{X}$. This amounts to a significant reduction of the degrees of freedom in the problem and potentially causes less variance in the estimate.

B. Stochastic Maximum Likelihood

Here, we derive the well-known stochastic maximum likelihood function [40]–[42]. Given the linear model (2) with Gaussian likelihood (13) and prior (15) the array data $\mathbf{Y}$ is Gaussian for each snapshot l with the covariance $\boldsymbol{\Sigma}_{\mathbf{y}_l}$ given by

$$\boldsymbol{\Sigma}_{\mathbf{y}_l} = \mathsf{E}[\mathbf{y}_l \mathbf{y}_l^H] = \boldsymbol{\Sigma}_{\mathbf{n}_l} + \mathbf{A} \boldsymbol{\gamma} \mathbf{A}^H. \tag{18}$$

The probability density function of $\mathbf{Y}$ is thus given by

$$p(\mathbf{Y}) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_l; \mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{y}_l}) = \prod_{l=1}^{L} \frac{\mathrm{e}^{-\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{y}_l}}{\pi^N \det \boldsymbol{\Sigma}_{\mathbf{y}_l}}. \tag{19}$$

The L-snapshot log-likelihood for estimating $\boldsymbol{\gamma}$ and $\sigma_{1:L} = [\sigma_1, \ldots, \sigma_L]$ is

$$\log p(\mathbf{Y}; \boldsymbol{\gamma}, \sigma_{1:L}) \propto - \sum_{l=1}^{L} \left(\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{y}_l + \log \det \boldsymbol{\Sigma}_{\mathbf{y}_l} \right). \tag{20}$$

This likelihood function is identical to the Type II likelihood function (evidence) in standard SBL [31] which is obtained by treating $\boldsymbol{\gamma}$ as a hyperparameter similar to LIKES [37]. The parameter estimates $\hat{\boldsymbol{\gamma}}$ and $\widehat{\sigma}_{1:L}$ are obtained by maximizing the log-likelihood, leading to

$$(\widehat{\boldsymbol{\gamma}}, \widehat{\sigma}_{1:L}) = \underset{\boldsymbol{\gamma} \geq 0, \ \sigma_{1:L} \in \mathbb{R}_{0+}^L}{\arg\max} \ \log p(\mathbf{Y}; \boldsymbol{\gamma}, \sigma_{1:L}). \tag{21}$$

C. LASSO Versus SBL

Within a Bayesian framework, both LASSO and SBL use the linear model (2) with complex zero-mean Gaussian random noise but they differ in the modeling of the source matrix $\mathbf{X}$. The LASSO approach assumes a priori that $\mathbf{X}$ is random with uniformly i.i.d. distributed phase and Laplace-like prior amplitudes,

$$p(\mathbf{X}) = p(\mathbf{x}^{\ell_2}) \propto \exp(-\|\mathbf{x}^{\ell_2}\|_1 / \nu), \tag{22}$$

$$[\mathbf{x}^{\ell_2}]_n = \left(\sum_{l=1}^{L} |[\mathbf{x}_l]_n|^2 \right)^{1/2}. \tag{23}$$

Thus only the summed amplitudes (23) are Laplacian. The elements in $\mathbf{X}$ are unknown and must be estimated. The LASSO approach uses the conditional likelihood (Type I) for $p(\mathbf{Y}|\mathbf{X})$ and applies Bayes rule with the prior $p(\mathbf{X})$ giving the MAP estimate

$$\begin{aligned} \widehat{\mathbf{X}} &= \arg\max p(\mathbf{Y}|\mathbf{X}) p(\mathbf{X}) \\ &= \underset{\mathbf{X} \in \mathbb{C}^{N \times L}}{\arg\min} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_{\mathcal{F}}^2 + \mu \|\mathbf{x}^{\ell_2}\|_1. \end{aligned} \tag{24}$$

The approach in (24) estimates the realization of $\mathbf{x}_l$ for each snapshot l. The number of parameters to be estimated for the LASSO approach grow linearly with the number of snapshots. This differs from our SBL approach, which estimates the source power/variance $\boldsymbol{\gamma}$ across all snapshots.

D. Source Power Estimation Using SBL

SBL follows the stochastic maximum likelihood approach (21) to estimate source powers (i.e. γ_m). Our derivation follows [33], [43]. We impose the diagonal structure $\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$, in agreement with (15), to form derivatives of (20) with respect to the diagonal elements γ_m. Using

$$\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1}}{\partial \gamma_m} = -\boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_l}}{\partial \gamma_m} \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} = -\boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1}, \tag{25}$$

$$\frac{\partial \log \det(\boldsymbol{\Sigma}_{\mathbf{y}_l})}{\partial \gamma_m} = \mathrm{tr}\left[\boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_l}}{\partial \gamma_m} \right] = \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m, \tag{26}$$

the derivative of (20) is formulated as

$$\begin{aligned} \frac{\partial \log p(\mathbf{Y}; \boldsymbol{\gamma}, \sigma_{1:L})}{\partial \gamma_m} &= \sum_{l=1}^{L} \left(\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{y}_l - \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m \right) \\ &= \sum_{l=1}^{L} |\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m|^2 - \sum_{l=1}^{L} \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m. \end{aligned} \tag{27}$$

In order to solve the stochastic maximum likelihood equation (21), we impose the necessary condition $\frac{\partial \log p(\mathbf{Y}; \boldsymbol{\gamma}, \sigma_{1:L})}{\partial \gamma_m} = 0$

in (27), giving

$$\sum_{l=1}^{L} |\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m|^2 = \sum_{l=1}^{L} \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m . \tag{28}$$

We are not aware of any closed form solution when solving (28) for γ_m and formulate a fixed point update rule:

$$\gamma_m \sum_{l=1}^{L} |\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m|^2 = \gamma_m \sum_{l=1}^{L} \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m . \tag{29}$$

Assuming γ_m and $\boldsymbol{\Sigma}_{\mathbf{y}_l}$ are known from previous iterations or initialization, we rewrite (29) to obtain a fixed point iteration for γ_m (with similarities to [30]–[33], [43]):

$$\gamma_m^{\text{new}} = \gamma_m^{\text{old}} \left(\frac{\sum_{l=1}^{L} |\mathbf{y}_l^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m|^2}{\sum_{l=1}^{L} \mathbf{a}_m^H \boldsymbol{\Sigma}_{\mathbf{y}_l}^{-1} \mathbf{a}_m} \right). \tag{30}$$

This fixed point iteration is not guaranteed to converge [31] and the number of iterations required for a reasonable convergence parameter (see SBL pseudo-code in Sec. III-G) is typically large (on the order of 1000).

For a single snapshot, the numerator in (30) belongs to the class of generalized matched filters $|\mathbf{y}^H \boldsymbol{\Sigma}^{-1} \mathbf{a}|^2$ [44, p.137], [45, p.351] because the data is whitened and projected onto the replica. SBL can be considered an adaptive version of this class since its whitener [i.e., the model covariance (18)] is estimated iteratively. For example, selecting an appropriate noise model (Sec. I-B) improves whitening on a per-snapshot basis if the noise process is non-stationary over snapshots. The inverse of this model covariance exits (if the noise variance is non-zero), whereas the inverse of a data-based estimated covariance may be singular (e.g., when $L < N$).

E. Noise Variance Estimation

While there has been more focus on estimating the source location parameters or $\boldsymbol{\gamma}$, noise is also part of the model and physical system. In SBL, the noise variance controls the sharpness of the peaks in the $\boldsymbol{\gamma}$ spectrum, with higher noise levels giving broader peaks. Thus, as we optimize $\boldsymbol{\gamma}$, we expect that a better noise model will improve convergence properties and processor localization performance.

In this section we give equations to estimate the noise variance for Cases I and II in Sec. I-B. For derivations of these noise estimates in detail see [2]. We will assume that the sparsity of $\boldsymbol{\gamma}$ is known.

1) Noise estimate, Case I: Under Case I, where $\boldsymbol{\Sigma}_{\mathbf{n}_l} = \sigma^2 \mathbf{I}_N$, the stochastic maximum likelihood [46] can provide an asymptotically efficient estimate of σ^2 if the number of sources $\mathcal{M}$ is known. Let $\boldsymbol{\Gamma}_{\mathcal{M}} = \text{diag}(\boldsymbol{\gamma}_{\mathcal{M}}^{\text{new}})$ be the covariance matrix of the K active sources obtained above with corresponding active replica vector matrix $\mathbf{A}_{\mathcal{M}}$ which maximizes (20). The noise variance estimate is given by

$$\hat{\sigma}^2 = \frac{\text{tr}[(\mathbf{I}_N - \mathbf{P})\mathbf{S}_{\mathbf{y}}]}{N-K}, \tag{31}$$

where $\mathbf{P} = \mathbf{A}_{\mathcal{M}}(\mathbf{A}_{\mathcal{M}}^H \mathbf{A}_{\mathcal{M}})^{-1} \mathbf{A}_{\mathcal{M}}^H$ is the projection matrix onto the subspace spanned by the active replica vectors. Note that the Case I estimate (31) is valid for any number of snapshots, even for just one snapshot. The algorithm using (30) for source power estimates and (31) for noise variance estimate is referred to as SBL1.

2) Noise estimate, Case II: For Case II, the noise variance changes with snapshot l, i.e. $\boldsymbol{\Sigma}_{\mathbf{n}_l} = \sigma_l^2 \mathbf{I}_N$. We apply (31) for each snapshot l individually leading to

$$\hat{\sigma}_l^2 = \frac{\text{tr}[(\mathbf{I}_N - \mathbf{P})\mathbf{y}_l \mathbf{y}_l^H]}{N-K} = \frac{\|(\mathbf{I}_N - \mathbf{P})\mathbf{y}_l\|_2^2}{N-K}. \tag{32}$$

The algorithm which uses (30) for source power estimates and (32) for noise variance estimates is referred to as SBL2.

F. Multi-Frequency SBL

Let observations be available at multiple frequencies and the fth frequency array data be denoted by $\mathbf{Y}_f$, where $f = 1, \ldots, F$. The observation model is given by

$$\mathbf{Y}_f = \mathbf{A}_f \mathbf{X}_f + \mathbf{N}_f, \tag{33}$$

where $\mathbf{A}_f$ is the replica dictionary, $\mathbf{X}_f$ is the unknown source amplitude matrix, and $\mathbf{N}_f$ is the noise at the fth frequency. Since the same set of sources are generating these observations, the sparsity profile of $\mathbf{X}_f$ is the same for all frequencies.

For computational tractability of the update equations, we assume that the source amplitudes $\mathbf{X}_f$ and the noise $\mathbf{N}_f$ are independent across frequencies. The priors for $\mathbf{X}_f$ also share the same covariance $\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma})$ for all frequencies, which further reduces the number of parameters and in turn increases identifiability. In other words, source amplitudes $\hat{\mathbf{X}}_f$ can vary across frequency while the source's power spectrum is represented by a single parameter (this can be viewed as an ideal bandpass approximation to the source's power spectrum). This tractable multi-frequency SBL model thus allows for estimating source location parameters rather than source power (note also the normalized replica vectors).

Proceeding as in Sec. III-B, the multi-frequency log-likelihood is then given by:

$$\log p(\mathbf{Y}_{1:F}; \boldsymbol{\gamma}, (\sigma_{1:L})_{1:F}) \propto -\sum_{f=1}^{F} \sum_{l=1}^{L} \left(\mathbf{y}_{fl}^H \boldsymbol{\Sigma}_{\mathbf{y}_{fl}}^{-1} \mathbf{y}_{fl} + \log \det \boldsymbol{\Sigma}_{\mathbf{y}_{fl}} \right), \tag{34}$$

where $$\boldsymbol{\Sigma}_{\mathbf{y}_{fl}} = \boldsymbol{\Sigma}_{\mathbf{n}_{fl}} + \mathbf{A}_f \boldsymbol{\Gamma} \mathbf{A}_f^H, \quad \boldsymbol{\Sigma}_{\mathbf{n}_{fl}} = \sigma_{fl}^2 \mathbf{I}. \tag{35}$$

σ_{fl}^2 is the noise variance of the fth frequency at the lth snapshot. As before, computing the derivative of (34) and equating it to zero gives the multi-frequency update rule: [34], [43], [47]

$$\gamma_m^{\text{new}} = \gamma_m^{\text{old}} \left(\frac{\sum_{f=1}^{F} \sum_{l=1}^{L} |\mathbf{y}_{fl}^H \boldsymbol{\Sigma}_{\mathbf{y}_{fl}}^{-1} \mathbf{a}_{fm}|^2}{\sum_{f=1}^{F} \sum_{l=1}^{L} \mathbf{a}_{fm}^H \boldsymbol{\Sigma}_{\mathbf{y}_{fl}}^{-1} \mathbf{a}_{fm}} \right). \tag{36}$$

A plot of Eq. (36) can be interpreted as a broadband ambiguity surface. The noise variance estimates for each frequency can be computed using either (31) or (32) depending on the noise type. Ideally, this estimate accounts for noise effects across snapshots and frequencies.

TABLE I
SBL ALGORITHM

0	Initialize: $\boldsymbol{\gamma}^{\text{new}} = \mathbf{1}$, $\boldsymbol{\Sigma}_{\mathbf{n}_{fl}}^{\text{new}}$ = Eq. (31) or (32), with $\mathbf{P} = \mathbf{0}, K = 2$ $\epsilon_{\min} = 0.001, \epsilon = 5 \times 10^{-6}, j = 0, j_{\max} = 3000$
1	while $(\epsilon > \epsilon_{\min})$ and $(j < j_{\max})$
2	$\boldsymbol{\gamma}^{\text{old}} = \boldsymbol{\gamma}^{\text{new}},\ \boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma}^{\text{old}}),\ \boldsymbol{\Sigma}_{\mathbf{n}_{fl}}^{\text{old}} = \boldsymbol{\Sigma}_{\mathbf{n}_{fl}}^{\text{new}}$
3	$\boldsymbol{\Sigma}_{\mathbf{y}_{fl}} = \boldsymbol{\Sigma}_{\mathbf{n}_{fl}}^{\text{old}} + \mathbf{A}_f \boldsymbol{\Gamma} \mathbf{A}_f^H$ use (35)
4	γ_m^{new} use (36) $\forall m$
5	$\mathcal{M} = \{m \in \mathbb{N} \mid$ K largest peaks in $\boldsymbol{\gamma}\} = \{m_1 \ldots m_K\}$
6	$\mathbf{A}_{f\mathcal{M}} = (\mathbf{a}_{fm_1}, \ldots, \mathbf{a}_{fm_K})$
7	$\boldsymbol{\Sigma}_{\mathbf{n}_{fl}}^{\text{new}}$ = choose from (31) or (32)
8	$\epsilon = \Vert\boldsymbol{\gamma}^{\text{new}} - \boldsymbol{\gamma}^{\text{old}}\Vert_1 / \Vert\boldsymbol{\gamma}^{\text{old}}\Vert_1, \quad j = j+1$
9	Output: $\mathcal{M}, \boldsymbol{\gamma}^{\text{new}}$

G. SBL Pseudo-Code

The complete multi-frequency SBL algorithm is summarized in Table I. The same algorithm is valid for Cases I and II. Given the observed $\mathbf{Y}_f$, we iteratively update $\boldsymbol{\Sigma}_{\mathbf{y}_{fl}}$ (18) by using the current $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\mathbf{n}_{fl}}$. The $\boldsymbol{\Sigma}_{\mathbf{y}_{fl}}^{-1}$ is computed directly as the numerical inverse of $\boldsymbol{\Sigma}_{\mathbf{y}_{fl}}$. For updating $\gamma_m, m = 1, \ldots, M$ we use (36). For the initialization of $\boldsymbol{\gamma}$ we set it to the constant $\mathbf{1}$.

Based on the corresponding Noise Case I or II, (31) or (32) is used to estimate $\boldsymbol{\Sigma}_{\mathbf{n}_{fl}}$. The noise is initialized using the same equations with $\mathbf{P} = \mathbf{0}, K = 0$, which provides an over estimate of the noise variance. The convergence parameter ϵ measures the relative change in estimated total source power,

$$\epsilon = \Vert\boldsymbol{\gamma}^{\text{new}} - \boldsymbol{\gamma}^{\text{old}}\Vert_1 \;/\; \Vert\boldsymbol{\gamma}^{\text{old}}\Vert_1. \tag{37}$$

The algorithm stops when $\epsilon \leq \epsilon_{\min}$ and the output is the working set $\mathcal{M}$ (17) from which all source parameters are computed.

IV. Data Selection and Processing

A. SWellEx-96 Data

We use the relatively range-independent SWellEx-96 Event S59 data set [27], [34], [48] for processing, recorded on Julian Day 134, 11:45–12:50 GMT. The surface ship R/V Sproul traveled with a speed of 2.5 m/s towards the VLA with closest point of approach (CPA) at 1 km (Fig. 1). The ship towed a submerged source at 60 m depth emitting multiple tonals.

For the 65 min Event S59, data are sampled at $f_s = 1500$ Hz on a 64 element vertical line array (VLA) but only $N = 21$ elements are used for processing (see Fig. 2). To localize the submerged source, we processes 14 frequencies: 52, 67, 82, 97, 115, 133, 151, 166, 169, 204, 238, 286, 341, and 391 Hz. Observed received levels for an unknown Source 1 location were 150 dB re 1 μPa for 166 Hz and 122–132 dB re 1 μPa for the other frequencies.

The data are split into 135 segments (segment length of 29 s). A FFT length of 4096 samples (2.7 s, 0.37 Hz bins) with 50% overlap results in $L = 21$ snapshots per segment. These snapshots are used to construct the SCM (8) for Noise Case I. Snapshots are normalized prior to constructing the SCM for Noise Case II (6). Our algorithm also processes the adjacent FFT bins

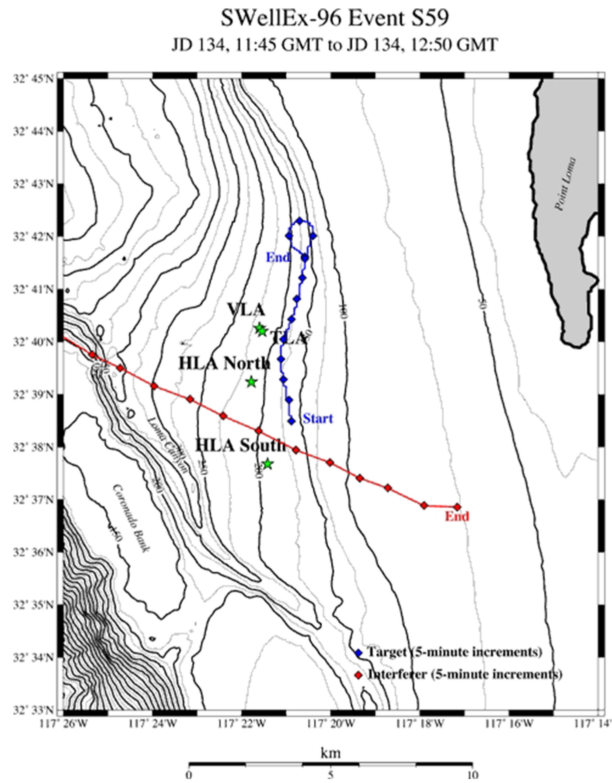

Fig. 1. (Color online) SWellEx-96 Event S59 showing the path of the surface ship R/V Sproul in blue and an opportunistically recorded surface ship (an interferer) in red. The R/V towed a deep source at approximately 60 m depth along roughly a 175 m isobath during the first part of the 65 min VLA recording, ending with a "U-turn".

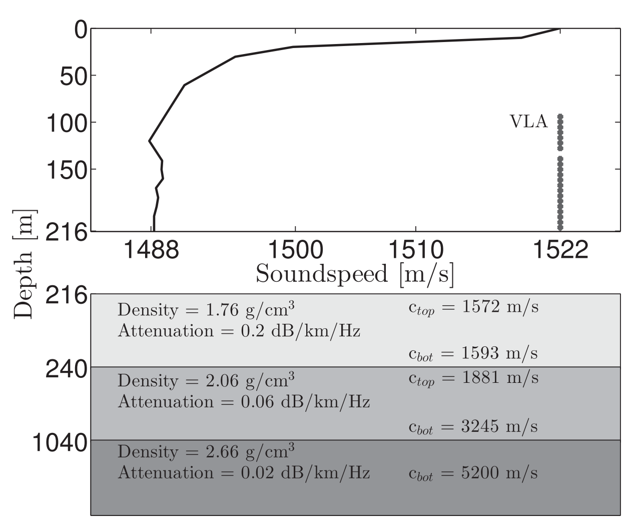

Fig. 2. Waveguide with sound speed profile, VLA, and geoacoustic parameters for range-independent SWellEx-96 Event S59. A single element out of the 22 element subset is excluded from processing.

at each frequency to accommodate an unknown Doppler shift. Thus, we process $F = 14 * 3 = 42$ frequencies using Bartlett (7), WNGC (9), and SBL (36).

For the range-independent waveguide geoacoustic model (Fig. 2) [48], the water depth is 216 m, 4 m below the deepest array element. The 22 element VLA spans the lower half of the water column and the inter-element spacing is 5.6 m (design frequency of 133 Hz at 1488 m/s). The seafloor is composed of a 24 m thick sediment layer, overlaying a 800 m thick mudstone layer. $M = 4000$ replica vectors $\mathbf{a}(\boldsymbol{\theta})$ per frequency are computed using the Kraken normal mode code [49] with a range and depth discretization step-size of 50 m and 10 m on a 10 km x 200 m grid, respectively. This step-size choice ideally is selected according to the (physically) interfering modes excited at the source location [50].

B. Simulations

To explore processor performance in a controlled environment, we use the SWellEx-96 replica vectors to simulate a multi-frequency scenario. Source 1 is the submerged source and Source 2 is a surface source (interferer). To simulate $L = 21$ snapshots while ensuring that both sources are incoherent, each source phase is selected independently from a uniform distribution $[0, 2\pi)$ for each snapshot. This requirement is necessary for eigenanalysis methods (e.g., MVDR, WNGC, and MUSIC) because signal coherence affects processor performance when inverting the SCM [45, p.385], [41]. SBL does not invert the SCM and can accommodate coherent sources [27].

The simulated received signal and noise vectors are added:

$$\mathbf{Y} = \mathbf{a}(\boldsymbol{\theta}_1)\mathbf{x}_1^T + \mathbf{a}(\boldsymbol{\theta}_2)\mathbf{x}_2^T + \mathbf{N}, \tag{38}$$

where each $\mathbf{x}_K \in \mathbb{C}^L$ contains L complex amplitudes for the Kth source. These observations $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_L]$ are used to construct the SCM (8) over L snapshots for Noise Case I or normalized snapshots for Noise Case II (6). At each frequency, source amplitudes are drawn from a uniform distribution with zero mean and specified variance. Similar to data observed during the experiment, Source 2 (interferer) power is 12, 12, 10, and 5 dB above Source 1 for 52, 82, 204, 286 Hz. Source 1 is dominant for 166 Hz (15 dB above Source 2).

Processor performance at a particular frequency is evaluated for additive noise (SNR discretization is 0.5 dB). Here, array (or average single element) SNR is defined as the ratio of the power of the weaker source to independent and identically distributed (i.i.d.) complex Gaussian noise $\mathbf{n}$:

$$\text{SNR} = 10 \log_{10} \frac{\mathbf{E}\{||\mathbf{a}x||_2^2\}}{\mathbf{E}\{||\mathbf{n}||_2^2\}} [\text{dB}]. \tag{39}$$

Equation (39) corresponds to the single snapshot SNR, where $\mathbf{a}$ is the source replica and x its complex amplitude. When simulating $L = 21$ snapshots, the signals are added to i.i.d. complex Gaussian noise in (38). When processing multiple frequencies, each frequency is generated with a different noise seed.

The source wavefield does not correspond exactly to a dictionary entry due to environmental uncertainty. We simulate the stationary sources (38) on a more finely spaced grid of replica vectors (2 m in depth and 10 m in range). The finely spaced replica set allows each snapshot to be drawn randomly from 5*5-1=24 additional positions while remaining within ± 1/2 cell to the grid point (10 m depth and 50 m range discretization).

Processor performance is measured by comparing the location of five dominant peaks found on the ambiguity surface to the location of the grid point of the submerged Source 1. The localization statistic P_L for this source is computed by:

$$P_L = \frac{C}{Q}, \tag{40}$$

where C is the number of correctly found peaks for Q simulations. For the single (multi) frequency case, we set $Q = 500$ ($Q = 200$). In order to effectively assess and compare processor performance, 95% credibility intervals (CI) are computed in addition to localization statistics P_L (which are representative of the mean).

CIs are calculated in three steps: 1) a single peak belonging to the set of five dominant (candidate) peaks having minimum distance (measured separately in range and depth) to the location of the submerged Source 1 is retained. 2) 5% of data points with the largest localization error are removed from the simulation set Q. 3) this process is repeated for all SNRs and each processor. In addition to CIs, selected histograms further illustrate the distributions of localization error for this joint localization-detection problem.

V. Results

A. Simulations

Processor localization performance is investigated first in a controlled simulation environment. The weaker Source 1 is placed in the SWellEx-96 environment at 2.5 km range and 60 m depth, Source 2 (i.e., the surface Interferer) at 4.5 km range and 15 m depth. Panels in Fig. 3(a)–(c) show ambiguity surfaces for Bartlett, WNGC –3 dB, and SBL1, respectively. SBL1 and SBL2 use Noise Case I (31) and Noise Case II (32), respectively.

To facilitate a comparison, each ambiguity surface is normalized by its maximum value and processor output power P is decibel [dB, i.e., $10 \log_{10}(P)$]. Array SNR is 10 dB and all processors use $L = 21$ snapshots at 204 Hz. The Interferer is 10 dB above Source 1.

Bartlett exhibits a strong sidelobe behavior [Fig. 3(a)]. WNGC and SBL1 suppress ambiguous source positions but the location output power at Source 1 remains comparable to ambiguous localizations [Fig. 3(b), (c)]. Hence processor performance is measured by comparing the location of five dominant peaks found on the ambiguity surface to the location of the grid point of the submerged Source 1.

Figure 3(d) compares processor probability of localization (P_L) [same scenario as Fig. 3(a)–(c)] versus SNR using $Q = 500$ simulations at each SNR. Bartlett exhibits poor performance because the 5 highest peaks are not in the vicinity of Source 1 ($P_L = 0$). To suppress sidelobes, adaptive processing is required. The localization performance for SBL1 and SBL2 is similar.

The benefit of SBL2 becomes apparent in a controlled simulation environment subject to non-stationary noise across snapshots. Figure 4(a)–(d) show simulation 10 out of $Q = 500$, the array SNR is 0 dB. SBL2 (d) outperforms SBL1 (c) because its noise model can account for non-stationary noise. The localization performance gain of SBL2 over SBL1 is about 20 dB in SNR as observed in Panel (e).

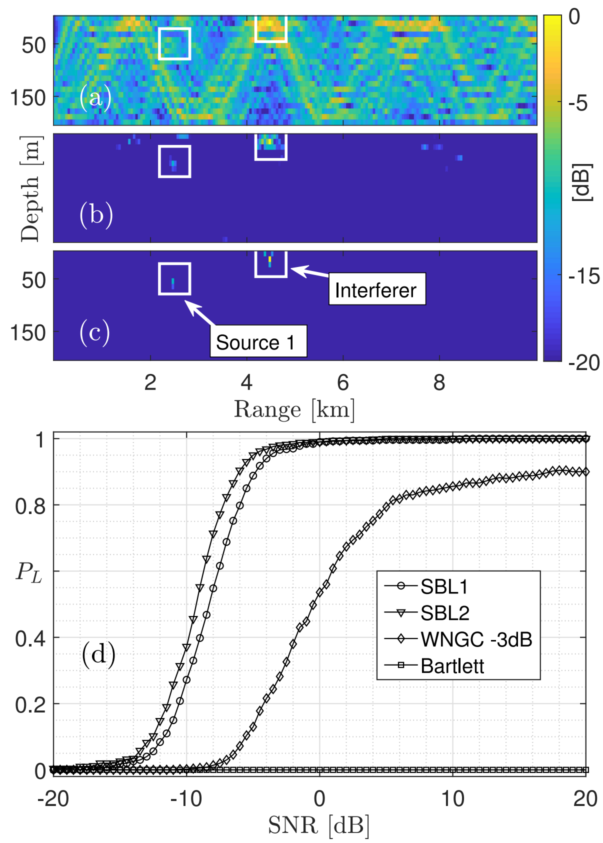

Fig. 3. Localization for Noise Case I with two simulated sources at 204 Hz and SNR 10 dB. True positions are indicated by white squares and each panel is normalized by its respective peak value. 21 data snapshots are drawn randomly from a finer replica vector mesh for stationary sources. The weaker Source 1 is located at 2.5 km with power 10 dB below the surface Source 2 located at 4.5 km: (a) Bartlett; (b) WNGC –3 dB; and (c) SBL1. Panel (d) shows the probability of localizing (P_L) the weaker Source 1.

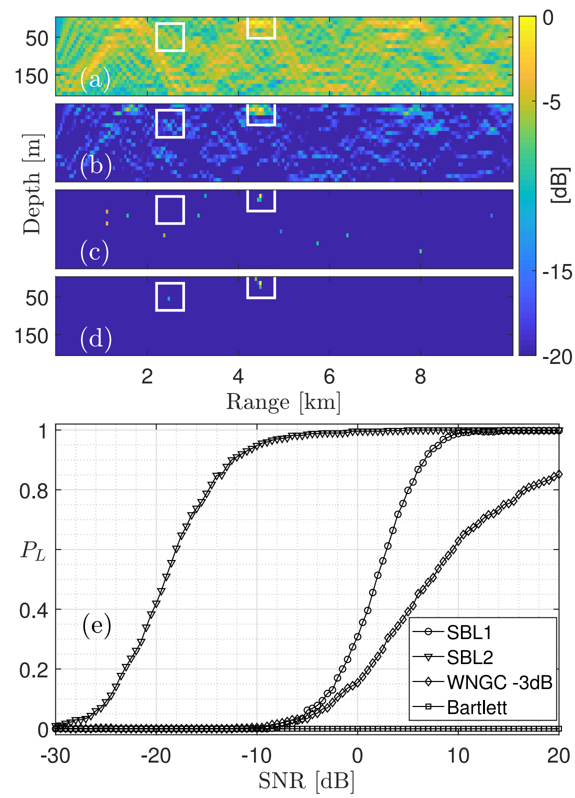

Fig. 4. Probability of localizing the weaker Source 1 for Noise Case II (see Sec. I-B). The single-frequency scenario is the same as in Fig. 3 but with panel (d) SBL2 and SNR of 0 dB in Panels (a)–(d).

Processor performance is evaluated by comparing 95% Credibility Intervals (CI) shown in Fig. 5 (same SNR as in Fig. 4 but for simulation 200 out of $Q = 500$, to get a sense of processor output variability). Panels (a)–(d) show Source 1 CIs by the size of the white rectangles, the CI is zero for SBL2 at SNR $= 0$ dB. The 95% range error CI versus SNR in Panel (e) indicates a similar behavior for Bartlett and WNGC –3 dB, and for SBL1 and SBL2.

Ambiguity surfaces for Bartlett and WNGC illustrate that CIs strongly depend on the interference pattern (source coordinates) for this joint localization-detection problem [51]. For example, a $CI_R = 900$ m in Fig. 5(e) from 0–20 dB SNR likely corresponds to the sidelobe structure observed at 1.5 km range and 15 m depth in Panels (a) and (b). All processors have a maximum CI_R of 2 km, which is the range separation between Source 1 and the dominant surface Interferer. The CI_R curves for SBL1 and SBL2 are of similar shape but shifted by about 20 dB [see also Fig. 4(e)].

The SBL CI is binary in Fig. 5(e) (0 or 2 km), lacking the $CI_R = 900$ m for Bartlett and WGNC. This suggests that the underlying PDF is of different shape. The histogram in Fig. 6(a) is used to compute 95% CIs (range only) in Fig. 5(a)–(e) at 0 dB SNR. Range error in these histograms is not entirely indicative of

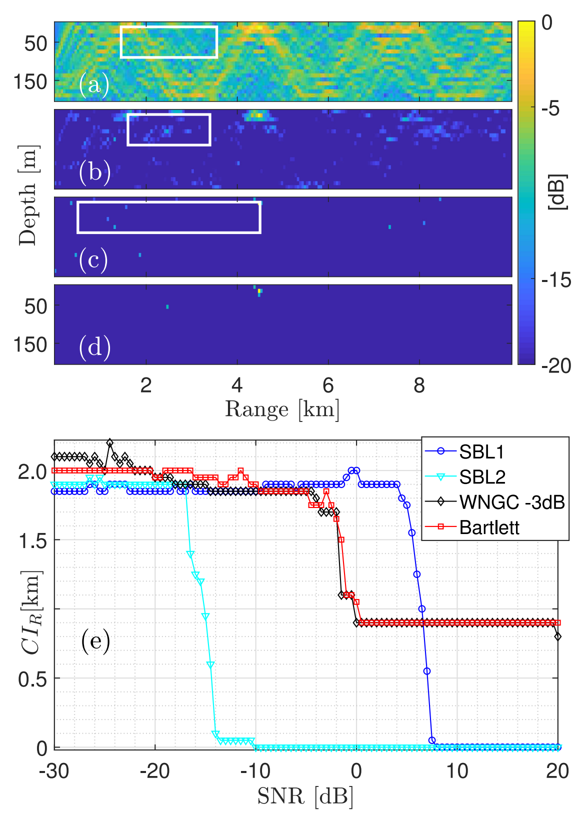

Fig. 5. 95% Credibility interval (CI) for localizing Source 1 for Noise Case II. Panels (a)–(d) show 95% CI for range and depth at SNR 0 dB as a white rectangle, Panel (e) shows 95% CI for range (CI_R) only.

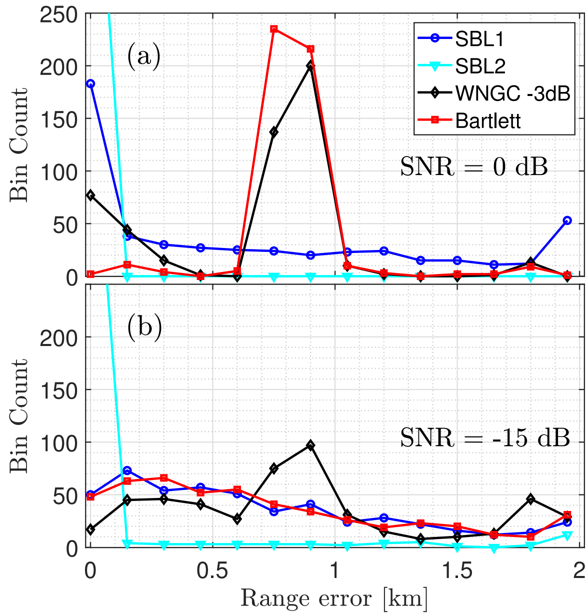

Fig. 6. Range error histograms for localizing Source 1 for Noise Case II. The 95% CIs in Fig. 5(d) at 0 and -15 dB SNR are computed from Panels (a) and (b), respectively. The plotted range of the y-axis is reduced from $Q = 500$ to 250 bins. The bin width is 150 m and range errors occurring at ranges greater than 2 km are shifted to the bin closest to 2 km.

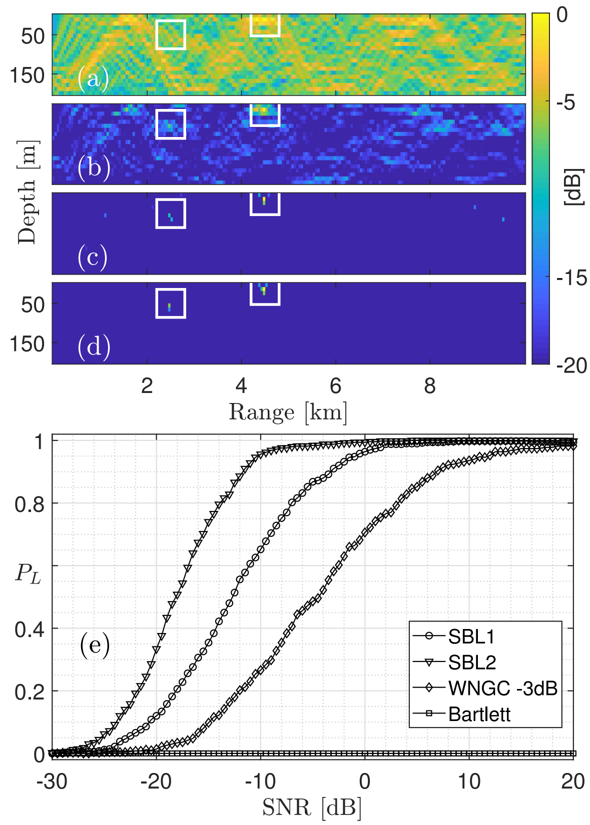

Fig. 7. As Fig. 4 but snapshots are normalized prior to processing.

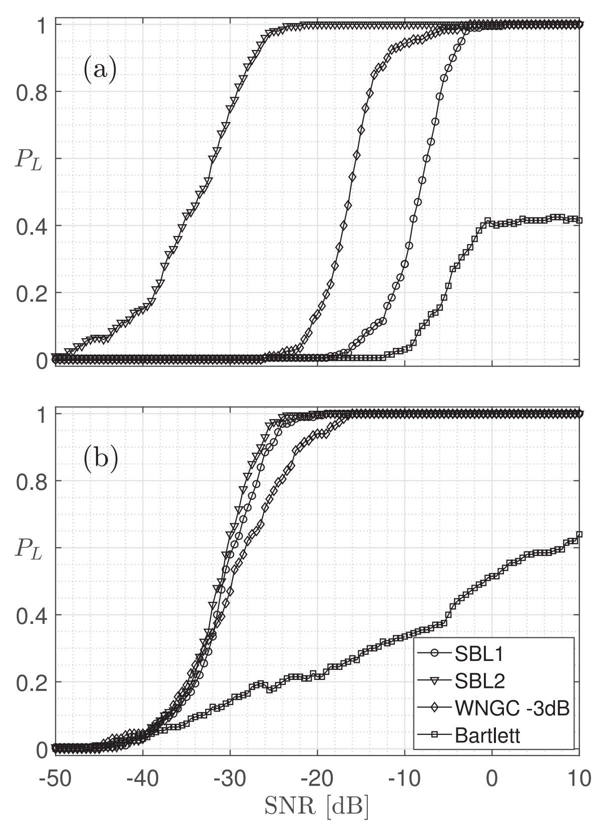

Fig. 8. Probability of localizing the weaker Source 1 for Noise Case II in a multi-frequency scenario (52, 82, 166, 204, and 286 Hz). Snapshots are (a) not normalized and (b) normalized.

performance because the peak label (i.e., sidelobe or mainlobe of the Kth source) at 900 m must be considered. SBL1 suppresses this sidelobe while Bartlett and WNGC can not (Fig. 5(e), 0 to 6 dB SNR). Note that the WNCG can suppress the sidelobe structures at 900 m with a –6 dB constraint (not shown), yielding a similar binary CI as SBL (see also Fig. 4 in [27]).

For non-stationary noise, snapshots can be normalized prior to processing (see Sec. I-C). In this case, performance increases for SBL1 and WNGC when comparing results in Fig. 7 to Fig. 4. SBL2 performance is similar in both cases, indicating that it can accommodate both stationary or non-stationary noise.

Probability of localizing the weaker Source 1 increases at lower SNRs when processing multiple frequencies (Fig. 8) for Noise Case II with $Q = 200$ simulations per frequency. For Bartlett and WNGC, a multi-frequency ambiguity surface is the average of individually computed ambiguity surfaces (12), SBL uses (36). The selected frequencies and the Source 1 to Source 2 power ratio are selected to be representative of the data, see below (38).

Bartlett can localize the quiet Source 1 at higher SNR in Fig. 8 than in Fig. 4. Processing additional frequencies improves the location of the weaker source because the mainlobes are in the same location while the sidelobes are not. Notably, the WNGC –3 dB exhibits improved performance when compared to SBL1 [Panel (a)]. All adaptive processors exhibit similar performance in Panel (b) when snapshots are normalized.

B. SWellEx-96 S59

We investigate processor performance to localize the weaker Source 1 in a multi-frequency scenario 35 min into the Event S59. Figure 9 shows the weaker, 60 m deep SWellEx-96

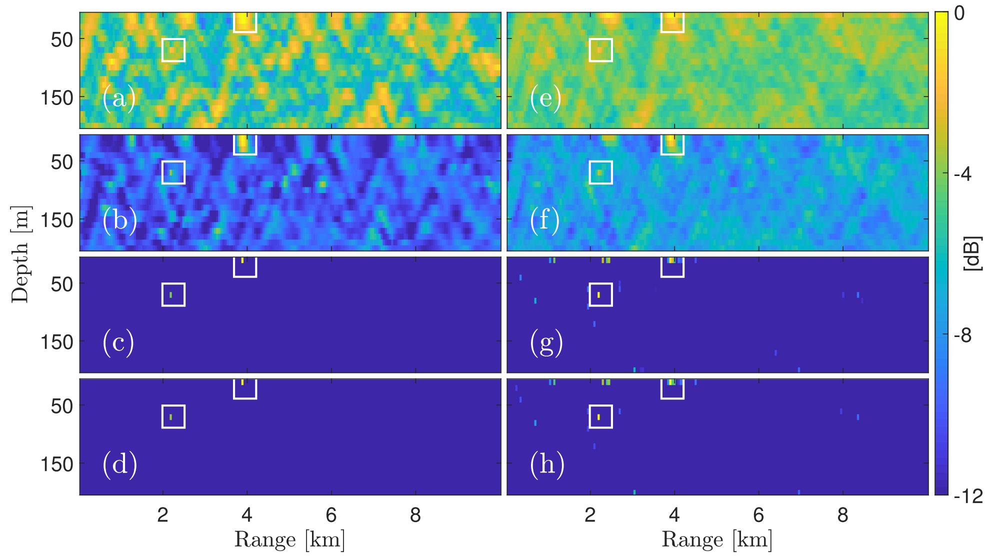

Fig. 9. Multi-frequency localization for the SWellEx-96 deep Source 1 and surface Interferer for data segment 71 (35 min into the event). True positions are indicated by white squares. Snapshots are not normalized (left, Noise Case I) and normalized (right, Noise Case II): (a), (e) Bartlett; (b), (f) WNGC –3 dB; (c), (g) SBL1; and (d), (h) SBL2. The 14 processed frequencies range from 52–391 Hz and include the adjacent bins, in total $F = 42$ frequencies.

Source 1 and the surface Interferer at range 2.5 and 4.5 km, respectively. Results in the left panels use non-normalized snapshots (Noise Case I). These multi-frequency ambiguity surfaces show a similar scenario as in Fig. 3.

SBL exhibits improved localization performance for the submerged source in Fig. 9(c) and (d) when compared to Bartlett (a) and WNGC (b). Both Bartlett and WNGC display power at the Interferer location and in adjacent cells. The WNGC can discriminate against many of the ambiguous solutions compared to Bartlett.

Modest mismatch is present in the SWellEx96 data set. SBL robustness to VLA tilt was rigorously investigated in a similar event of the same experiment [34], where the projected tilt on the source-VLA plane was on the order of 1–2°. Environmental mismatch exists because the acoustic waveguide parameters are unknown (approximated with a range-independent model). SBL displays no ambiguity in Fig. 9(c–d), which demonstrates that SBL has properties similar to an adaptive processor while exhibiting robustness to modest mismatch in the form of environmental and sensor location errors. Since SBL models the complex source amplitudes as random quantities, it provides robustness to amplitude and phase errors in the model.

Right panels in Fig. 9 use normalized snapshots. SBL1 and SBL2 exhibit the best performance while WNGC –3 dB exhibit some and Bartlett the most ambiguity for source localization. Normalizing snapshots reduces ambiguity for localizing the weaker source. Note that all processed frequency bins contain energy from the Interferer while this is not necessarily the case for the narrow band tonals emitted by the quiet Source 1.

Results in Fig. 10 show a similar analysis as Fig. 9 but for segment 93 of 135 (45 min into the event). The Interferer is past 10 km range yet interference patterns result in ambiguous source localizations near the surface. Normalized snapshots increase processor output power [Fig. 10(e–h)] at the location of the weaker source [e.g., compare Panel (b) to Panel (f)], which may help source localization. Localization ambiguity is increased in Fig. 10 compared to Fig. 9, which indicates that modeled replicas are increasingly mismatched to the data.

To further demonstrate SBL performance in localizing the weaker Source 1, we extend our processing to the entire Event S59 in Fig. 11. We extract the five highest power levels for each of the 135 segments (e.g., from results similar to Figs. 9–10) and plot corresponding range information for each ambiguity surface. This data then is displayed as a vertical stripe, containing only these five power levels in their respective range cells. The vertical stripes are assembled in temporal order and individually normalized range-time panels are shown in Fig. 11 for each processor.

Results shown in Fig. 11 allow for a qualitative assessment of processor performance in localizing both sources over the entire event, split into 135 individual localization realizations. Panels (c), (d) demonstrate the processing advantage of SBL, where SBL2 performs equal or better than SBL1. The WNGC -3 dB in Panel (b) can suppress some ambiguous tracks due to sidelobes relative to Bartlett in Panel (a). All Processors exhibit a gain in localizing both sources when snapshots are normalized in Panels (e)–(h) and the localization performance of Bartlett, SBL1 and SBL2 is similar. Comparing these results to the approximately true range cell ($\pm$ 2 cells) of the source, processors in Panels (a)–(h) yield 41, 26, 78, 83, 80, 62, 86, and 89 successful localizations, respectively.

The encountered modest mismatch is not a constant over the 65 min Event S59. The acoustic waveguide parameters used to calculate replicas in $\mathbf{A}$ (2) are assumed range-independent and time invariant over the experiment duration. Clearly, this is an attractive simplification but replica fitness degrades over time in dynamic ocean environments. All processors are subjected to environmental variability and array element position deviations. The SBL algorithm displays robustness to these kinds of changing mismatch.

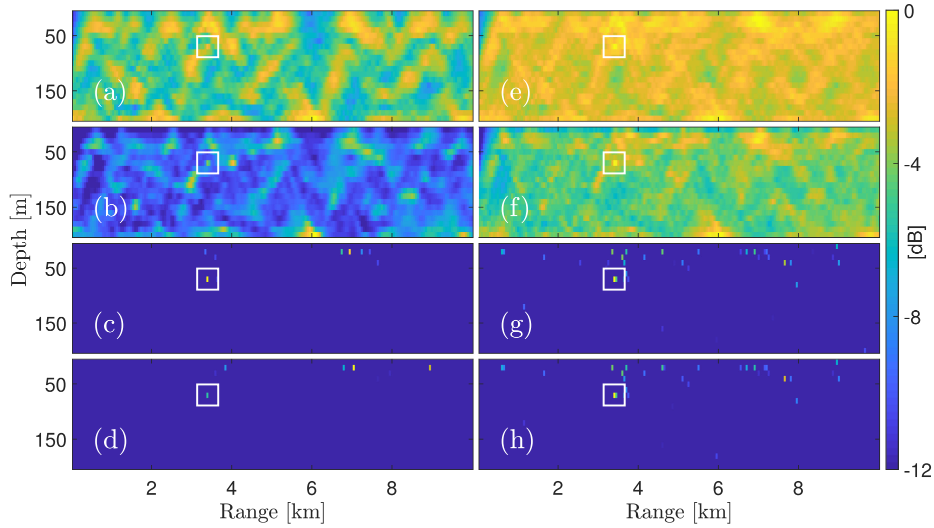

Fig. 10. As Fig. 9 but for segment 93 (45 min into the event). Normalizing snapshots improves localization of the weaker Source 1. The Interferer location is past 10 km, outside the processed range.

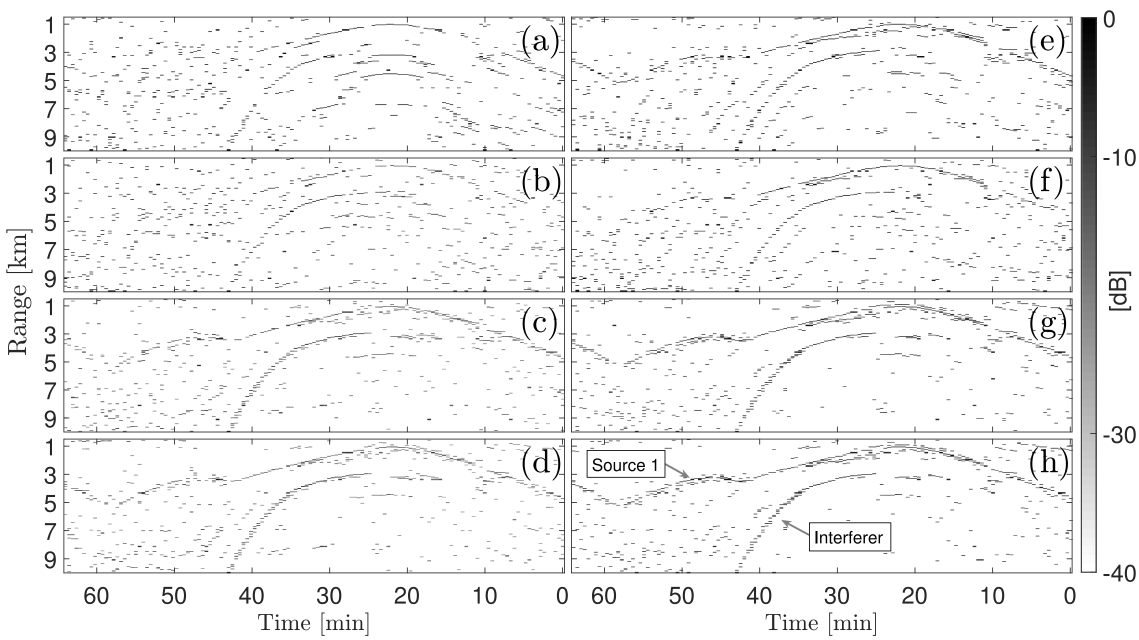

Fig. 11. Multi-frequency range localization of Source 1 and the Interferer for the SWellEx-96 Event S59. For each of the 135 processed segments/ambiguity surfaces (Fig. 9 shows No. 71 and Fig. 10 No. 93), five peaks corresponding to the highest power levels are plotted. Snapshots are not normalized (left, Noise Case I) and normalized (right, Noise Case II): (a), (e) Bartlett; (b), (f) WNGC −3 dB; (c), (g) SBL1; and (d), (h) SBL2.

It appears a *constant* WNGC constraint of –3 dB is inappropriate at times, because the non-adaptive Bartlett processor in Fig. 11(e) outperforms the WNGC in Fig. 11(f) from 40–65 min. This may be due to the joint processing of adjacent bins and changing mismatch. The broadband Interferer is present in all 3 adjacent bins but the narrowband Source 1 is just in a single bin. Model mismatch increases with interferer range, in which case ambiguous solutions become part of the 5 highest peaks. Likely WNGC performance improves when allowing the constraint to change. In this sense, the single SBL solution [computed across frequency (36)] can better accommodate scenarios with operational uncertainty.

VI. Conclusion

We demonstrated that sparse Bayesian learning (SBL) behaves similar to an adaptive processor and that it is robust to modest mismatch in simulations and with the SWellEx-96 data set. SBL outperforms WNGC −3 dB when localizing a weaker source in the presence of an interferer in a multi-frequency data scenario. Unlike other high-resolution sparse methods, SBL automatically determines sparsity. We further demonstrated the usefulness of estimating non-stationary noise for SBL. Normalizing snapshots prior to computing the sample covariance increases the localization performance for Bartlett, WNGC −3 dB, and SBL.

References

[1] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3038–3044, Jul. 2008.

[2] P. Gerstoft, S. Nannuru, C. F. Mecklenbräuker, and G. Leus, "Sparse bayesian learning for DOA estimation in heteroscedastic noise," 2017, arXiv:1711.03847v1.

[3] N. M. Law, C. D. Mackay, and J. E. Baldwin, "Astrophysics Lucky imaging: High angular resolution imaging in the visible from the ground," *Astron. Astrophys.*, vol. 446, no. 2, pp. 739–745, 2006.

[4] H. Ge and I. P. Kirsteins, "Lucky ranging with towed arrays in underwater environments subject to non-stationary spatial coherence loss," in *Proc. IEEE Intl. Conf. Acoust., Speech Signal Process.,* 2016, pp. 3156–3160.

[5] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 18, no. 4, pp. 401–424, Oct. 1993.

[6] T. Meng and J. R. Buck, "Rate distortion bounds on passive sonar performance," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 326–336, Jan. 2010.

[7] Y. Le Gall, F.-X. Socheleau, and J. Bonnel, "Matched-field processing performance under the stochastic and deterministic signal models," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5825–5838, Nov. 2014.

[8] S. Finette and P. C. Mignerey, "Stochastic matched-field localization of an acoustic source based on principles of Riemannian geometry," *J. Acoust. Soc. Amer.*, vol. 143, no. 6, pp. 3628–3638, 2018.

[9] K. L. Gemba and E.-M. Nosal, "Source characterization using recordings made in a reverberant underwater channel," *Appl. Acoust.*, vol. 105, pp. 24–34, 2016.

[10] M. D. Collins, N. C. Makris, and L. T. Fialkowski, "Noise cancellation and source localization," *J. Acoust. Soc. Am.*, vol. 96, no. 3, pp. 1773–1776, 2005.

[11] H. Kuttruff, *Room Acoustics*, 5th ed. New York, NY, USA: Taylor & Francis, 2009.

[12] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. Berlin, Germany: Springer, 2011.

[13] G. J. Orris, M. Nicholas, and J. S. Perkins, "The matched-phase coherent multi-frequency matched-field processor," *J. Acoust. Soc. Amer.*, vol. 107, no. 5, pp. 2563–2575, 2000.

[14] B. G. Ferguson, K. W. Lo, and R. A. Thuraisingham, "Sensor position estimation and source ranging in a shallow water environment," *IEEE J. Ocean. Eng.*, vol. 30, no. 2, pp. 327–337, Apr. 2005.

[15] V. E. Premus and M. N. Helfrick, "Use of mode subspace projections for depth discrimination with a horizontal line array: Theory and experimental results," *J. Acoust. Soc. Amer.*, vol. 133, no. 6, pp. 4019–4031, 2013.

[16] J. S. Rogers and J. L. Krolik, "Time-varying spatial spectrum estimation with a maneuverable towed array," *J. Acoust. Soc. Amer.*, vol. 128, no. 6, pp. 3543–3553, 2010.

[17] M. J. Hinich, "Maximum-likelihood signal processing for a vertical array," *J. Acoust. Soc. Amer.*, vol. 54, no. 2, pp. 499–503, 1973.

[18] M. J. Bianco and P. Gerstoft, "Dictionary learning of sound speed profiles," *J. Acoust. Soc. Amer.*, vol. 141, no. 3, pp. 1749–1758, 2017.

[19] C. M. A. Verlinden, J. Sarkar, W. S. Hodgkiss, W. A. Kuperman, and K. G. Sabra, "Passive acoustic tracking using a library of nearby sources of opportunity," *J. Acoust. Soc. Amer.*, vol. 143, no. 2, pp. 878–890, 2018.

[20] K. L. Gemba, J. Sarkar, B. Cornuelle, W. S. Hodgkiss, and W. A. Kuperman, "Estimating relative channel impulse responses from ships of opportunity in a shallow water environment," *J. Acoust. Soc. Amer.*, vol. 144, no. 3, pp. 1231–1244, Sep. 2018.

[21] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.

[22] J. N. Maksym, "A robust formulation of an optimum cross-spectral beamformer for line arrays," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 971–975, 1979.

[23] R. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[24] P. A. Forero and P. A. Baxley, "Shallow-water sparsity-cognizant source-location mapping," *J. Acoust. Soc. Amer.*, vol. 135, no. 6, pp. 3483–3501, 2014.

[25] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Amer.*, vol. 136, no. 1, pp. 260–271, 2014.

[26] P. Gerstoft, C. F. Mecklenbräuker, W. Seong, and M. Bianco, "Introduction to special issue on compressive sensing in acoustics," *J. Acoust. Soc. Amer.*, vol. 143, no. 6, pp. 3731–3736, 2018.

[27] K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," *J. Acoust. Soc. Amer.*, vol. 141, no. 1, pp. 92–103, 2017.

[28] M. A. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 342–349, Apr. 2010.

[29] H. Xu, C. Caramanis, and S. Mannor, "Robust regression and lasso," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3561–3574, Jul. 2010.

[30] M. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[31] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[32] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, Jul. 2007.

[33] P. Gerstoft, C. F. Mecklenbrauker, A. Xenaki, and S. Nannuru, "Multi-snapshot sparse Bayesian learning for DOA," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1469–1473, Oct. 2016.

[34] K. L. Gemba, S. Nannuru, P. Gerstoft, and W. S. Hodgkiss, "Multi-frequency sparse Bayesian learning for robust matched field processing," *J. Acoust. Soc. Amer.*, vol. 141, no. 5, pp. 3411–3420, 2017.

[35] D. Wipf and S. Nagarajan, "Iterative reweighted L1 and L2 methods for finding sparse solutions," *IEEE J. Sel. Topics Signal Process.*, vol. 4, pp. 317–329, 2010.

[36] R. Giri and B. D. Rao, "Type I and type II Bayesian methods for sparse signal recovery using scale mixtures," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3418–3428, Jul. 2016.

[37] P. Stoica and P. Babu, "SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation," *Signal Process.*, vol. 92, no. 7, pp. 1580–1590, 2012.

[38] H. Cox, "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *J. Acoust. Soc. Amer.*, vol. 54, no. 5, pp. 1289–1301, 1973.

[39] W. A. Kuperman and F. Ingenito, "Spatial correlation of surface generated noise in a stratified ocean," *J. Acoust. Soc. Amer.*, vol. 67, no. 6, pp. 1988–1996, 1980.

[40] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Process.*, vol. 11, no. 4, pp. 329–337, 1986.

[41] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[42] P. Stoica and A. Nehorai, "On the concentrated stochastic likelihood function in array signal processing," *Circuits Syst. Signal Process.*, vol. 14, no. 5, pp. 669–674, 1995.

[43] S. Nannuru, K. L. Gemba, P. Gerstoft, W. S. Hodgkiss, and C. F. Mecklenbräuker, "Sparse Bayesian learning with multiple dictionaries," *Signal Process.*, vol. 159, pp. 159–170, Jun. 2019.

[44] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[45] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[46] J. F. Böhme, "Source-parameter estimation by approximate maximum likelihood and nonlinear regression," *IEEE J. Ocean. Eng.*, vol. OE-10, no. 3, pp. 206–212, Jul. 1985.

[47] S. Nannuru, K. L. Gemba, and P. Gerstoft, "Sparse Bayesian learning with multiple dictionaries," in *Proc. IEEE Global Conf. Signal Inf. Process.*, 2017, pp. 1190–1194.

[48] 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.*, vol. 25, no. 3, pp. 296–313, Jul. 2000.

[49] M. B. Porter, "The KRAKEN normal mode program," SACLANT Undersea Research Centre Memorandum (SM-245) and Naval Research Laboratory, Norfolk, VA, USA, Mem. Rep. 6920, 1992.

[50] B. H. Tracey, "Matched field range and depth resolution in range-dependent waveguides," *Acoust. Res. Lett. Online*, vol. 6, no. 4, 2005, Art. no. 274.

[51] A. N. Mirkin and L. H. Sibul, "Maximum likelihood estimation of the locations of multiple sources in an acoustic waveguide," *J. Acoust. Soc. Am.*, vol. 95, no. 2, pp. 877–888, 1994.

Authors' photographs and biographies not available at the time of publication.