

Adaptive and compressive matched field processing

Kay L. Gemba,^{a)} William S. Hodgkiss, and Peter Gerstoft

*Marine Physical Laboratory of the Scripps Institution of Oceanography,
University of California at San Diego, La Jolla, California 92093-0238, USA*

(Received 2 May 2016; revised 5 October 2016; accepted 14 December 2016; published online 10 January 2017)

Matched field processing is a generalized beamforming method that matches received array data to a dictionary of replica vectors in order to locate one or more sources. Its solution set is sparse since there are considerably fewer sources than replicas. Using compressive sensing (CS) implemented using basis pursuit, the matched field problem is reformulated as an underdetermined, convex optimization problem. CS estimates the unknown source amplitudes using the replica dictionary to best explain the data, subject to a row-sparsity constraint. This constraint selects the best matching replicas within the dictionary when using multiple observations and/or frequencies. For a single source, theory and simulations show that the performance of CS and the Bartlett processor are equivalent for any number of snapshots. Contrary to most adaptive processors, CS also can accommodate coherent sources. For a single and multiple incoherent sources, simulations indicate that CS offers modest localization performance improvement over the adaptive white noise constraint processor. SWellEx-96 experiment data results show comparable performance for both processors when localizing a weaker source in the presence of a stronger source. Moreover, CS often displays less ambiguity, demonstrating it is robust to data-replica mismatch.

© 2017 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4973528>]

[JFL]

Pages: 92–103

I. INTRODUCTION

Matched field processing (MFP) is a generalized beamforming method that matches received array data to a dictionary of replicas to localize a source (see [Baggeroer et al., 1993](#), for a historical overview). The solution set inherently is sparse, having substantially fewer sources than possible positions and is a candidate application for compressible (approximately sparse) processors.

Compressive sensing (CS; e.g., see [Candès and Wakin, 2008](#)) provides the framework to obtain sparse solutions to underdetermined problems with convex optimization (ℓ_1 -norm minimization) rather than an exhaustive search (ℓ_0 -norm minimization). This is possible if the underlying signal is sparse and if the replica dictionary that maps the underlying signal to the observations is sufficiently incoherent. However, dictionary entries may be coherent or even redundant in practice because the entries represent a physical process and therefore cannot be chosen arbitrarily ([Candès et al., 2011](#)). This is the case for plane wave beamforming, and [Xenaki et al. \(2014\)](#) discuss the relationship between replica coherence and signal-to-noise ratio (SNR). We extend the CS approach from the beamforming ([Edelmann and Gaumond, 2011](#)) to the MFP application, and estimate location and amplitude parameters in single- and multi-source scenarios.

CS, potentially, is attractive because it can achieve super-resolution beyond the Rayleigh limit even for the single snapshot case ([Fortunati et al., 2014](#)). CS has been investigated in the time domain for ultrasound signals ([David et al., 2015](#)) and wavenumber tracking ([Le Courtois and Bonnel, 2015](#)).

Singular value decomposition can help to improve the rank of the dictionary ([Malioutov et al., 2005](#); [Edelmann and Gaumond, 2011](#)). However, CS performance degrades in the presence of mismatch due to a misalignment between the actual source field observed at the array and the modeled replica vector (also known as sources off the grid). In the presence of mismatch, grid refinement or a continuous grid might improve performance ([Xenaki and Gerstoft, 2015](#)). For single snapshot data using the basis pursuit method ([Chen et al., 1998](#)), CS can outperform traditional beamforming methods under challenging scenarios ([Xenaki et al., 2014](#)) at moderate to high SNRs. Beamforming results using multiple snapshots ([Gerstoft et al., 2015](#)) indicate that CS outperforms the minimum variance (MV; [Capon, 1969](#)) and MUSIC ([Schmidt, 1986](#)) processors in multi-source scenarios.

CS algorithms also have been investigated in underwater source localization applications. In addition to basis pursuit de-noising ([Liu et al., 2012](#)) and regularization via the Lasso path ([Tibshirani, 1996](#)), [Forero and Baxley \(2014\)](#) showed that an elastic-net ([Zou and Hastie, 2005](#)) might be a useful additional constraint for source localization. Localization performance for closely spaced sources ([Forero and Baxley, 2014](#)) and broadband underwater source localization ([Forero, 2014](#)) has been investigated. It was shown in [Xu et al. \(2010\)](#) that the solution to the Lasso has desirable robustness properties and offers protection from noise. Considering the wide applicability of CS, we investigate the performance of CS (implemented using the basis pursuit method and Lasso) and compare it to non-adaptive and adaptive matched field processors.

The non-adaptive Bartlett processor ([Bucker, 1976](#)) has been used in most papers on MFP, and thus it is an

^{a)}Electronic mail: gemba@ucsd.edu

appropriate candidate for a baseline comparison. We consider this processor using single and multiple (incoherently combined) frequencies. The adaptive white noise constraint (WNC; Cox *et al.*, 1987) processor offers potentially high-resolution source localization performance (Krolik, 1992; Debever and Kuperman, 2007). The processor is a good benchmark candidate because, unlike MV, the WNC is more robust to data-replica mismatch in realistic scenarios [see Eq. (3)]. The WNC will be considered using single and multiple frequencies as well.

This paper compares processor localization performance using single and multiple frequencies, as well as single and multiple snapshots to track single and multiple sources. For the matched field processing application, we demonstrate the following:

- For a single source, results show that the performances of CS and the Bartlett processor are equivalent for any number of snapshots (i.e., the one-sparse solution is equivalent to the least-squares one, see Sec. IV A and the Appendix);
- Simulations using multiple incoherent sources indicate that CS offers modest localization performance improvement over the adaptive WNC processor in localizing a weaker source in the presence of a stronger source (Sec. IV A, third paragraph);
- Unlike adaptive processors such as WNC, CS can accommodate coherent sources in multiple-source scenarios (Sec. IV A, fourth paragraph);
- CS often displays less ambiguity when compared to the WNC processor in data scenarios, demonstrating it is robust to data-replica mismatch (Sec. IV B).

First, we introduce processors in Sec. II followed by an overview of data selection and processing in Sec. III. Section IV presents simulations analyzing processor performance for both single- and two-source scenarios. We also investigate processor performance for equivalent scenario configurations using the SWellEx-96 data set (Booth *et al.*, 2000; Hursky *et al.*, 2001). This paper concludes with a discussion and summary in Secs. V and VI, respectively.

II. PROCESSORS

A. Bartlett and WNC processors

CS performance is benchmarked using Bartlett (P_B) and WNC (P_{wnc}) processors. Bartlett is a spatial matched-filter processor that matches normalized replica vectors \mathbf{a} (corresponding to the complex wavefield of a source at frequency f and position θ received at an array of N elements) to the data \mathbf{y} ,

$$P_B(\theta) = \mathbf{a}^H(\theta)\mathbf{K}\mathbf{a}(\theta), \quad (1)$$

where H denotes the Hermitian operator and $P_B(\theta)$ denotes the Bartlett power at position θ . The sample covariance matrix (SCM) $\mathbf{K} \in \mathbb{C}^{N \times N}$ is obtained using L snapshots

$$\mathbf{K} = \frac{1}{L} \sum_{l=1}^L \mathbf{y}_l \mathbf{y}_l^H. \quad (2)$$

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 l th data segment from each of the array elements. Averaging over multiple snapshots, generally, is desirable and sometimes required to improve SNR and, hence, source localization. However, the number of available snapshots to obtain \mathbf{K} depends on the time a source remains in a resolution cell (Baggeroer and Cox, 1999; Cox, 2002). While Bartlett does not invert \mathbf{K} and thus does not have a minimum number of required snapshots, it suffers from high sidelobes. Sidelobe suppression is important if several sources (or a combination of sources and interferers) are present.

The WNC processor P_{wnc} discriminates against sidelobes (and other sources/interferers) while offering a degree of robustness in frequently encountered replica-data mismatch scenarios. The WNC is versatile because of its ability to adjust its behavior (thus resolution) from Bartlett to MV (Maksym, 1979; Song *et al.*, 2003b) at the expense of inverting \mathbf{K} . To have \mathbf{K} invertible, we require $L > N$ (diagonal loading of \mathbf{K} can be used to mitigate this requirement). The WNC processor is given by

$$P_{\text{wnc}}(\theta) = \mathbf{a}_w^H(\theta)\mathbf{K}\mathbf{a}_w(\theta), \quad \text{where } \mathbf{a}_w = \frac{(\mathbf{K} + \epsilon\mathbf{I})^{-1}\mathbf{a}}{\mathbf{a}^H(\mathbf{K} + \epsilon\mathbf{I})^{-1}\mathbf{a}}. \quad (3)$$

The adaptive weights \mathbf{a}_w correspond to diagonally loaded MV weights and are obtained by solving

$$\begin{aligned} & \min_{\mathbf{a}_w} \mathbf{a}_w^H(\theta)\mathbf{K}\mathbf{a}_w(\theta) \text{ subject to} \\ & \mathbf{a}_w^H\mathbf{a} = 1, \\ & |\mathbf{a}_w^H\mathbf{a}_w|^{-1} \geq \delta^2. \end{aligned} \quad (4)$$

The iterative algorithm finds the diagonal loading ϵ for each replica vector or position θ to satisfy a white noise gain constraint G_{wng} such that

$$\delta^2 < G_{\text{wng}} = |\mathbf{a}_w^H\mathbf{a}_w|^{-1} < N, \quad (5)$$

which, in practice, is normalized and expressed as $10 \log_{10}(\delta^2/N) \leq 0$ dB. Thus, $P_{\text{wnc}}(\theta)$ denotes the WNC power at position θ for a selected (white noise gain) constraint.

Similar to the MV, the WNC maintains a distortionless response (or gain of 1) for the field component corresponding to the replica vector \mathbf{a} and minimizes the projection of the weights on directional noise or interferers possibly present in \mathbf{K} . However, in the presence of mismatch (i.e., a misalignment between the actual source field observed at the array and the modeled replica vector), a source might appear as an interference to the high-resolution MV and be suppressed by means of adaptive nulling. The WNC addresses this shortcoming at the expense of reduced resolution by imposing a gain constraint on the adaptive weights. Depending on the amount of mismatch encountered, an appropriate constraint frequently falls within $10 \log_{10}(\delta^2/N) \in [-6 - 2]$ dB (Booth

et al., 2000; Song *et al.*, 2003b). The WNC performs closer to Bartlett when using a constraint of -2 dB and more like MV when using a constraint of -6 dB.

To localize a source, Eqs. (1) and (3) are evaluated at M range-depth positions or cells θ . Processor output is arranged corresponding to each replica position on a range-depth plot, conventionally called an ambiguity surface. If the modeled replica vector is a close match to the acoustic field observed at the array, then the surface maximum (mainlobe) corresponds to the source position. A robust method to increase localization performance in the presence of environmental uncertainty is to sum over multiple ambiguity surfaces corresponding to Ω processed frequencies

$$\begin{aligned} P_B^\Omega(\theta) &= \sum_{i=1}^{\Omega} P_B(\theta, f_i), \\ P_{\text{wnc}}^\Omega(\theta) &= \sum_{i=1}^{\Omega} P_{\text{wnc}}(\theta, f_i). \end{aligned} \quad (6)$$

Localization ambiguity decreases if the mainlobe locations are consistent across frequency. Ideally, the processed frequencies span at least an octave, which increases the sidelobe diversity on the surface. For multiple sources, localization performance improves using multiple frequencies when localizing a weaker source in the presence of a stronger source. The power of the weaker source mainlobe might be below the power of a sidelobe of the stronger source. Consequently, using multiple frequencies is beneficial since Ω corresponds to a gain factor for the mainlobe of the weaker source. In most situations, we primarily are interested in identifying the mainlobe positions on this surface, which is an application for CS.

B. CS

1. Single snapshot CS

The CS approach reformulates the traditional spatial matched-filter problem as a convex optimization problem. A linear model relates the array data \mathbf{y} to the complex source amplitudes \mathbf{x} with additive noise \mathbf{n} ,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (7)$$

where $\mathbf{A} = [\mathbf{a}(1), \dots, \mathbf{a}(M)] \in \mathbb{C}^{N \times M}$ is the dictionary of replicas using an array of N hydrophones, M range-depth replicas $\mathbf{a} = [a_1, \dots, a_N]^T$, a single measurement vector $\mathbf{y}(1) = [y_1, \dots, y_N]^T$ and noise $\mathbf{n} = [n_1, \dots, n_N]^T$. T denotes the transpose operator. Localization estimation can be expressed as a linear underdetermined problem ($N \ll M$) with a sparsity constraint enforced on its solution

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathbb{C}^M} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1. \quad (8)$$

The ℓ_1 -norm $\|\cdot\|_1$ promotes sparsity and the regularization parameter λ controls the number of sparse solutions, which correspond to non-zero entries in \mathbf{x} . An approximate value for λ for any number of sparse solutions can be found via the

Lasso path (Tibshirani, 1996; Gerstoft *et al.*, 2015). The initialization value is set to $\lambda_0 = \|2\mathbf{A}^H \mathbf{y}\|_\infty$ and marks an upper bound corresponding to a set of values denoted by $\lambda^{(1)}$ at which a first non-zero entry in \mathbf{x} appears. The transition from the first to the second solution is observed for $\lambda^{(2)} < \lambda_1 = \|2\mathbf{A}^H(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_1)\|_\infty$. Successive upper bounds for $\lambda^{(i)}$ are found using $\lambda_{i-1} = \|\mathbf{r}_{i-1}\|_\infty$ with the residual $\mathbf{r}_{i-1} = 2\mathbf{A}^H(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_{i-1})$, where $\hat{\mathbf{x}}_{i-1}$ contains $i-1$ non-zero entries. The condition $\lambda^{(i)} < \lambda_{i-1}$ is the Lasso path indicator that the i th solution set is found. The optimization problem in Eq. (8) is solved for a given λ using the CVX toolbox (Grant and Boyd, 2016) for disciplined convex optimization. It uses interior point solvers to obtain the global solution of a well-defined optimization problem.

It is useful to distinguish between two main classes of methods to implement CS: greedy algorithms (not used here) such as matching pursuit (Mallat and Zhang, 1993) and non-greedy algorithms such as basis pursuit [corresponding to the ℓ_1 -norm minimization problem in Eq. (8)]. Both methods find the complex source amplitudes and related replica vectors by subtracting the modeled pressure field for the determined sparse sources from the observations. In principle, the algorithms have the ability to discriminate against sidelobes if they correctly separate each received source pressure from the data. However, they differ in how multiple sources are localized.

The term ‘‘greedy’’ implies that the algorithm makes a ‘‘hard’’ decision based upon some optimal minimization criterion (Eldar and Kutyniok, 2013). These algorithms make locally optimal choices and here are defined as finding the active indices one-by-one without revisiting previous solutions. Note that the definition of greedy can be relaxed to include more sophisticated (also termed greedy) algorithms, which revisit previous solutions to perform similarly to the ℓ_1 -norm minimization. Greedy implementations clean the data *sequentially* (e.g., in underwater acoustics the CLEAN algorithm; Song *et al.*, 2003a). A greedy algorithm first cleans the data with the best estimate of a stronger source 1 pressure, reducing ambiguity for identifying a weaker source 2 at the next iteration. This approach is problematic for sources of similar power, in scenarios with environmental uncertainty, or with non-orthogonal dictionary entries because an early false selection cannot be altered. See Chen *et al.* (1998) for additional discussion (and references), including an example in which matching pursuit does not converge.

In contrast to greedy implementations, the basis pursuit method has no location-memory of previously selected solutions because it performs a global optimization for a given regularizer λ . The algorithm can either rediscover or alter previously selected solutions as it descends on the Lasso path, giving CS (as implemented here) the ability to *adapt* in order to best explain the observations.

2. Multiple snapshots CS

The model in Eq. (7) can be extended to include L snapshots for an approximately stationary source

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}, \quad (9)$$

where $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(L)] \in \mathbb{C}^{N \times L}$ and $\mathbf{N} \in \mathbb{C}^{N \times L}$. Similarly, Eq. (8) can be extended to using multiple snapshots

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathbb{C}^{M \times L}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 + \lambda \sum_{j=1}^M \|\mathbf{X}_j\|_2, \quad (10)$$

where the Frobenious norm $\|\cdot\|_F$ for a matrix $\mathbf{B} \in \mathbb{C}^{N \times L}$ is defined as $\|\mathbf{B}\|_F = (\sum_{i,j} |b_{i,j}|^2)^{1/2}$ (the norm measures the misfit for a matrix). In the last term of Eq. (10), the ℓ_2 -norm acts on the j th row of the matrix \mathbf{X} and the ℓ_1 -norm (replaced by the summation operator) acts on the resulting column vector. The combination of these two norms ($\|\mathbf{X}\|_{1,2}$) enforces row-sparsity in the solution, but does not force the row-entries to be sparse (Malioutov *et al.*, 2005). Since \mathbf{X} is row sparse (i.e., for each sparse solution, most of its row is non-zero), multiple snapshots should be taken with the source in the same resolution cell.

Both Bartlett and WNC yield a power for every replica vector, while for CS every solution in $\hat{\mathbf{X}}$ corresponds to a biased estimate of the complex amplitude of the source. This bias is intrinsic to CS due to the additive sparsity constraint. Additional uncertainty arises since the number of solutions are constant over the interval $\lambda_i < \lambda^{(i)} < \lambda_{i-1}$. In this band, the complex amplitude alone varies.

The unbiased complex source amplitudes are estimated as

$$\hat{\mathbf{X}}_{\text{CS}} = \mathbf{A}_S^+ \mathbf{Y}, \quad (11)$$

where $\hat{\mathbf{X}}_{\text{CS}} \in \mathbb{C}^{S \times L}$ and $\mathbf{A}_S \in \mathbb{C}^{N \times S}$ contain only the active replicas associated with S non-zero solutions and $\mathbf{A}_S^+ = (\mathbf{A}_S^H \mathbf{A}_S)^{-1} \mathbf{A}_S^H$ is its Moore-Penrose pseudoinverse. The adjoint \mathbf{A}_S^H reverses the medium's phase shift and the "denominator" corrects for the numerator's array response. For generating a range-depth surface for comparing the CS results with Bartlett and WNC ambiguity surfaces, we computed the ℓ_2 -norm over the rows of $\hat{\mathbf{X}}_{\text{CS}}$. Since the range-depth surface generated by CS might include erroneous source locations, we will refer to this surface as an ambiguity surface.

Incorporating additional snapshots for sparse processing is beneficial to aid in source localization since data corresponding to the same resolution cell share the same replica vector in the dictionary. Imposing this dependency with $\|\mathbf{X}\|_{1,2}$ [the last term in Eq. (10)] is called a collaborative or a multitask sparse coding problem (Eldar and Kutyniok, 2013) or group Lasso. Adding noisy snapshots to high SNR snapshots does not result in incorrect support because the worst case performance of using multiple snapshots with $\|\mathbf{X}\|_{1,2}$ is equivalent to the worst case bounds of the single snapshot case (Davies and Eldar, 2012).

3. Extension to multiple frequencies: Incoherent processor

One MFP approach to improving source localization performance is to average incoherently over multiple ambiguity surfaces, whereas each surface corresponds to a different source frequency [see Eq. (6) in Sec. II A]. Localization performance improves when mainlobe locations are

approximately in the same range-depth cell and sidelobes are in different cells. For CS, averaging over independently computed surfaces generally is not desirable since there are no mainlobes or sidelobes—only sparse solutions. Therefore, we must minimize one multi-frequency cost function and enforce a sparsity constraint for all frequencies *simultaneously*

$$\hat{\mathbf{X}}^\Omega = \arg \min_{\mathbf{X}^\Omega \in \mathbb{C}^{M \times (L\Omega)}} \sum_{i=1}^{\Omega} \|\mathbf{Y}_{f_i} - \mathbf{A}_{f_i} \mathbf{X}_{f_i}\|_F^2 + \lambda \sum_{j=1}^M \|\mathbf{X}_j^\Omega\|_2. \quad (12)$$

In Eq. (12), Ω is the number of frequencies and $\mathbf{X}^\Omega = [\mathbf{X}_{f_1}, \dots, \mathbf{X}_{f_\Omega}] \in \mathbb{C}^{M \times L\Omega}$ is the multi-snapshot multi-frequency solution. The i th term in the cost function's sum only uses a subset of \mathbf{X}^Ω corresponding to one frequency. The second term in Eq. (12) enforces the row-sparsity constraint for all frequencies simultaneously in \mathbf{X}^Ω . The dimension of \mathbf{X}^Ω is such that every frequency uses the same number of snapshots. For a single sparse solution, the mostly non-zero row in $\hat{\mathbf{X}}^\Omega$ maps to a unique range-depth cell and its entries correspond to the biased complex amplitude estimate of the source at each frequency. Unbiased results are estimated for each frequency separately ($\hat{\mathbf{X}}_{\text{CS}}^\Omega \in \mathbb{C}^{S \times L\Omega}$) as in Eq. (11) and plotted using the ℓ_2 -norm over the rows of $\hat{\mathbf{X}}_{\text{CS}}^\Omega$. Note that if $\Omega = 1$, Eq. (12) reduces to Eq. (10).

III. DATA SELECTION

A. SWellEx-96 data

To compare processors, we use the relatively range-independent SWellEx-96 Event S5 data set (Booth *et al.*, 2000; Orris *et al.*, 2000; Hursky *et al.*, 2001) recorded on a 64 element vertical line array (VLA) with $N=21$ of those elements used for processing. The surface ship R/V Sproul traveled with a radial velocity of 2.5 m/s toward the VLA (Fig. 1) and towed a deep and a shallow source, both projecting different multi-tonal sets. Data are sampled at 1500 Hz and we use the deep source at frequencies 94, 112, 130, 148, 166, 235 Hz for processing. The first 53 min of the Event S5 are split into 160 segments with 50% overlap, resulting in a single segment length of 40 s. An FFT length of 4096 samples (~ 2.7 s) with 50% overlap results in 28 snapshots for each segment with a FFT bin width of 0.37 Hz. Our algorithm searches the adjacent ± 1 FFT bins and extracts the FFT value corresponding to the maximum bin power to accommodate Doppler shift. Each snapshot is windowed with a normalized Kaiser window (Kaiser, 1974) with $\beta=2.5$.

For the range-independent waveguide geoacoustic model (Fig. 2), the water depth is assumed to be the water depth at the array (216 m). The VLA spans the lower half of the water column and the inter-element spacing is 5.6 m, which corresponds to a $\lambda/2$ spacing at 133 Hz using a sound speed of 1488 m/s. The seafloor is composed of a 23.5 m thick sediment layer, overlaying an 800 m thick mudstone layer (Booth *et al.*, 2000). Replica vectors are computed using the Range-dependent Acoustic Model (RAM) parabolic equation code (Collins, 1993, 1994) with a range and

SWellEx-96 Event S5
 JD 131, 23:15 GMT to JD 132, 00:30 GMT

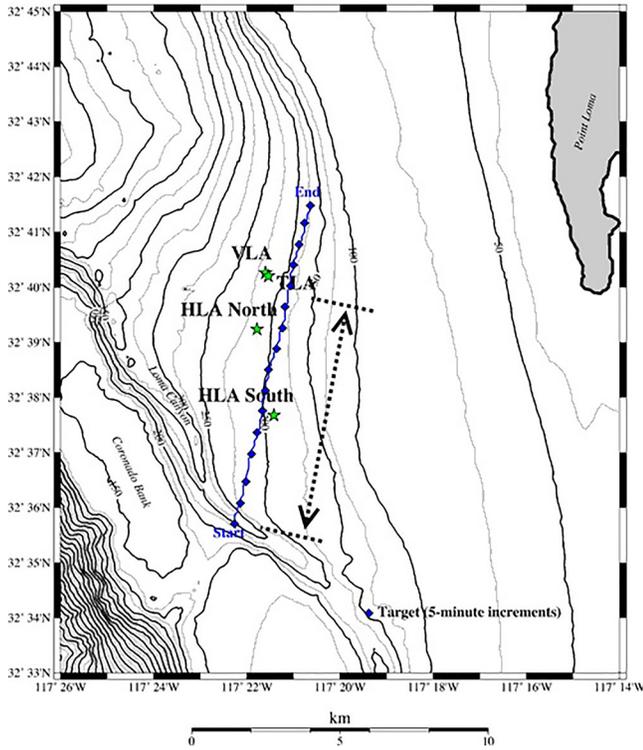


FIG. 1. (Color online) SWellEx-96 Event S5 showing the path of the surface ship R/V Sproul. The ship towed a deep source at ~ 60 m depth along roughly a 200 m isobath during the 75 min VLA recording. The source track used for processing is indicated by the black arrows.

depth discretization of 50 m and 10 m on a $10 \text{ km} \times 200 \text{ m}$ grid, respectively.

B. Simulations

To explore processor performance in a controlled environment, we use the SWellEx-96 replica vectors to simulate

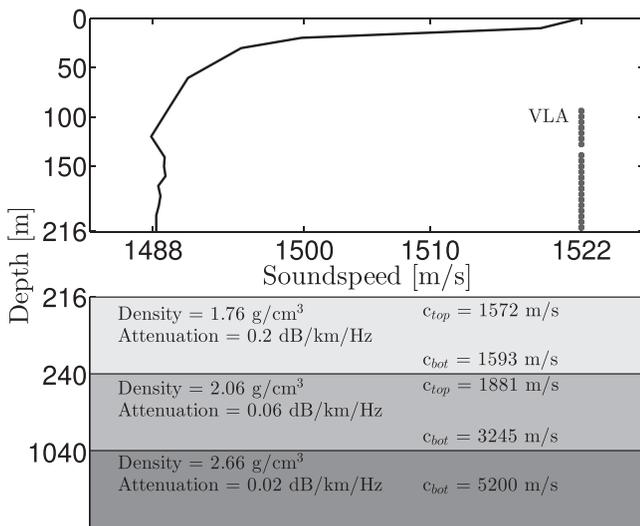


FIG. 2. Waveguide with sound speed profile, VLA, and geo-acoustic parameters for range-independent SWellEx-96 Event S5. A single element out of the 22 element subset is excluded from processing.

both single- and two-source scenarios. In the two-source scenario, source magnitudes are selected such that the source 1 to source 2 ratio (SSR) is 3 dB. Source 2 is the weaker source and is farther in range relative to source 1 by 1 km. To simulate L snapshots while ensuring that both sources are incoherent, each source phase is selected independently from a uniform distribution $[0, 2\pi)$ for each snapshot. The simulated received data vectors are added

$$\mathbf{Y} = \mathbf{a}(\theta_1)\mathbf{x}_1^T + \mathbf{a}(\theta_2)\mathbf{x}_2^T + \mathbf{N}, \quad (13)$$

where each $\mathbf{x}_i \in \mathbb{C}^L$ contains L complex amplitudes for the i th source. To construct coherent sources, the phase for sources 1 and 2 is the same for the l th snapshot. For both coherent and incoherent scenarios, the source level is constant and sources are stationary across snapshots. CS operates on the observations $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(L)]$ directly, while Bartlett and WNC use the same SCM averaged over L snapshots [Eq. (2)].

Processor performance for 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 weakest signal present to independent and identically distributed (i.i.d.) complex Gaussian noise \mathbf{n} ,

$$\text{SNR} = 10 \log_{10} \frac{\mathbf{E}\{\|\mathbf{ax}\|_2^2\}}{\mathbf{E}\{\|\mathbf{n}\|_2^2\}} \text{ (dB)}. \quad (14)$$

Equation (14) corresponds to the single snapshot SNR, where \mathbf{a} is the source replica and x its complex amplitude. When simulating L snapshots, the signals are added to i.i.d. complex Gaussian noise. When processing multiple frequencies, each frequency has the same SNR.

Processor performance is measured using root-mean-square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{Q} \sum_{q=1}^Q \Delta\theta_q^2}, \quad (15)$$

where the statistics are approximated using $Q = 1000$ realizations at each SNR and $\Delta\theta$ denotes the range error between true and estimated source location. For the single source scenario, the estimated location corresponds to the peak of the ambiguity surface. For the two source scenario, the estimated location corresponds to the location of the weaker source. In particular, we extract the two largest peaks on the ambiguity surface and assign the appropriate source number to those peaks by comparing the found locations to the true locations. If the stronger source is found correctly but the weaker source is not, the error can be computed directly. If both locations do not correspond to the true locations, we use the minimum error of all possible combinations. Note that the latter metric is not too important because source 2 dominates early RMSE behavior (see Table I for a RMSE comparison between sources 1 and 2). Since RMSE is sensitive to outliers, the computed RMSE curves are smoothed with a zero-phase forward and reverse

TABLE I. RMSE for CS using 28 snapshots in Fig. 4(b). Units: m.

SNR (dB)	-4	-5	-6	-7	-8	-9	-10
S1 range	0	0	0	0	0	42	280
S1 depth	0	0	0	0	0	1	3
S2 range	9	36	72	180	355	607	865
S2 depth	0	1	2	6	12	23	33

two-point moving average filter to avoid a potential biased representation of processor performance. For plotting, surfaces for all processors are normalized by their respective peak values.

IV. RESULTS

A. Simulation results

Processor localization performance is investigated first for a single source. Source 1 is placed in the SWellEx-96 environment at 2.5 km range and 60 m depth. Figure 3 shows the ambiguity surfaces for each of the three processors. Array SNR is 0 dB and all processors use $L = 28$ snapshots at 166 Hz. Bartlett exhibits a strong sidelobe behavior, the highest sidelobe is -2.5 dB below the mainlobe. The WNC is able to suppress ambiguous source positions, and the highest sidelobe is -9 dB below the mainlobe. The single sparse CS solution is clearly visible.

Figure 4(a) shows processor localization performance at different SNRs. Localization curves are only shown for non-zero RMSE. In agreement with theory (see the Appendix), both CS and Bartlett perform identically for a single and for 28 snapshots in localizing the source and additional snapshots help to localize the source at lower SNRs. The WNC performs worse than the Bartlett processor. As the constraint is decreased from -2 to -6 dB, localization performance decreases by roughly 2 dB in SNR. The scenario is extended to include additional frequencies in Fig. 4(c). Similar to Fig. 4(a), the performance for Bartlett and CS is identical using two and three incoherent frequency combinations with one and two snapshots each, respectively.

For the two-source scenario, we add a second source to the single source scenario at 3.5 km range and 60 m depth (see Fig. 5). The SSR is 3 dB and source 2 is the weaker

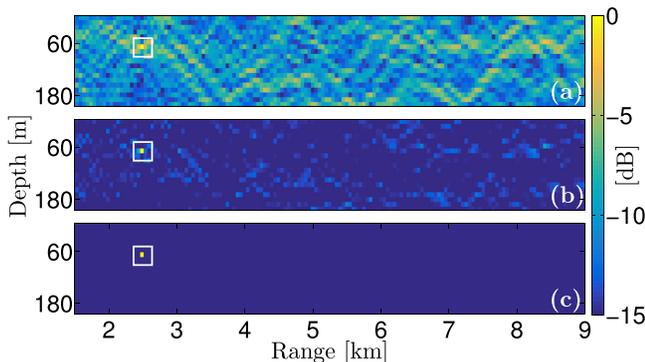


FIG. 3. (Color online) Single source localization simulation at SNR 0 dB and 166 Hz: (a) Bartlett, (b) WNC -2 dB, and (c) CS. True source locations are marked by white squares, and all processors localize the single source.

source. The array SNR for source 2 is 0 dB and all processors use 28 snapshots at 166 Hz. The Bartlett processor exhibits a strong sidelobe behavior. Several sidelobes at about -2 dB are higher than the level at the true position of source 2 at -2.5 dB. Both WNC and CS correctly locate each source. For WNC, source 2 is at -2.8 dB and the highest sidelobe is at -9 dB. For CS, source 2 is at -2.6 dB.

Processor ambiguity surfaces corresponding to two fully coherent sources are displayed in Fig. 6 (same configuration as in Fig. 5). Neither Bartlett nor the WNC exhibit a dominant peak corresponding to source 2. Only CS localizes both sources. While it is possible to decorrelate sources (e.g., Pillai and Kwon, 1989) for adaptive processing, we implement incoherent sources for all of the following scenarios because there is little reason to assume that source signatures will be coherent in practice.

Figure 4(b) compares processor localization performance for two-sources (same range-depth cells and SSR as shown in Fig. 5) versus SNR. Bartlett results are not plotted because the power of the second highest peak is not in the vicinity of source 2, resulting in a $RMSE > 1$ km. To suppress sidelobes and identify the position of source 2, adaptive processing is required. For the WNC, we selected two constraints: -2 and -6 dB. The WNC curves roughly are between the curves corresponding to CS solutions computed with 14 and 28 snapshots. A WNC constraint of -2 dB performs better than a constraint of -6 dB.

We investigate processor performance using two frequencies for the two-source scenario in Fig. 4(d). Bartlett now can localize the weaker source until a SNR of -4.5 dB. Similar to Figs. 4(a) and 4(b), a WNC constraint of -2 dB is preferred. WNC -2 dB has non-zero RMSE at SNR greater than -6.5 dB. CS has non-zero RMSE at SNR greater than -7.5 dB using 28 snapshots. The number of realizations is decreased from 1000 to 100 in order to reduce computation time for Fig. 4(d). Note that the SNR curves for a multi-source scenario depend on the relative position of the two sources. A different configuration changes the sidelobe structure and can shift the localization performance with respect to SNR, whereas relative processor performance largely is invariant to the actual source positions.

RMSE in Fig. 4 has been plotted using range error for the weakest source present. Table I shows a typical RMSE breakdown for the two source scenario split separately into range and depth. The processor has non-zero RMSE for the weaker source 2 (S2) at a higher SNR than the stronger source 1 (S1), and the source 2 range error is the dominant error.

Results in Figs. 3–5 compare processor performance for sources simulated with a dictionary replica (i.e., the simulated source is identical to a dictionary entry). To illustrate processor performance when the source field does not correspond exactly to a dictionary entry, we simulate the sources [Eq. (13)] on a more finely spaced grid of replica vectors (2 m in depth and 10 m in range). The denser replica set allows the source to be placed in 24 different positions in each of the 10 m depth and 50 m range discretization grid cells [i.e., the replica dictionary \mathbf{A} in Eq. (10)] used to localize sources. Figures 7(a)–7(c) shows processor output when

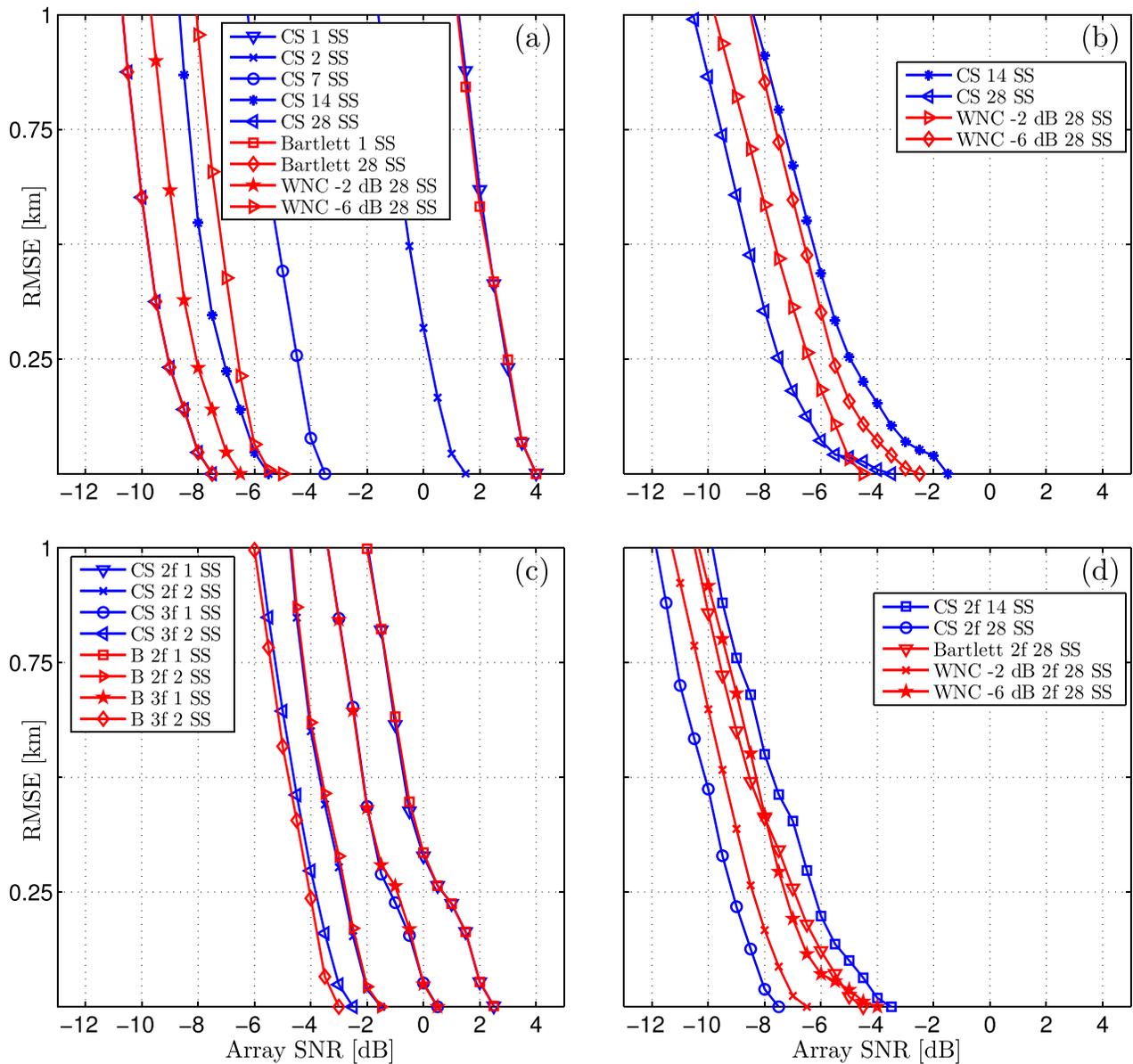


FIG. 4. (Color online) Range RMSE localization performance for single and multiple snapshots (SS). Single frequency 166 Hz top panels show (a) source 1 only and (b) sources 1 and 2. Multi-frequency bottom panels show (c) source 1 only and (d) sources 1 and 2. The two sets of frequencies are $2f = \{148, 166\}$ Hz and $3f = \{148, 166, 235\}$ Hz. In (b) and (d), SNR and RMSE correspond to the weaker source 2.

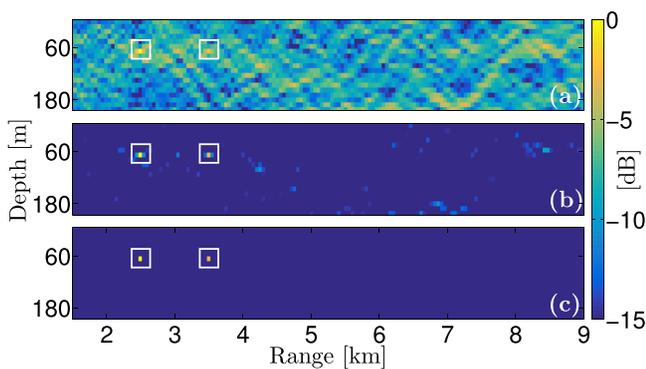


FIG. 5. (Color online) Two-source localization simulation at SNR 0 dB and 166 Hz: (a) Bartlett, (b) WNC -2 dB, and (c) CS. Bartlett has several competing sidelobes at higher levels than source 2. WNC and CS localize both sources.

28 snapshots are selected randomly from the finely spaced replicas. Both sources remain stationary in their respective discretization cells. To simulate non-stationary sources, 28 snapshots are split into 2 sets of 14 and drawn from locations between two adjacent dictionary entries in Figs. 7(d)–7(f).

Both mismatch simulations increase localization ambiguity for all processors when compared to results in Fig. 5. In particular, mismatch and source movement across discretization cells increases the processor output power in cells adjacent to the true source positions, which easily can be observed in the WNC and CS panels. Note that in Fig. 7(c) there are two solutions in the vicinity of source 1 (one above and one below the true position) at -17 dB. CS computes eight sparse solutions in Figs. 7(c) and 7(f), some with power below -15 dB. The highest WNC sidelobes in Figs. 7(b) and 7(e) are at -5 and -3.3 dB, respectively (both located at approximately 60 m depth and 8.5 km range).

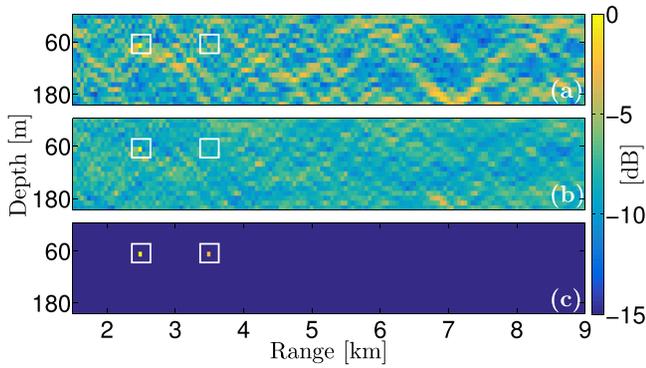


FIG. 6. (Color online) Same scenario as Fig. 5, but the two sources are coherent.

B. SWellEx-96 results

CS performance first is illustrated by comparing it to the Bartlett processor for the single SWellEx-96 source at 166 Hz traveling toward the VLA. Figures 8(a)–8(e) show P_B output with 1, 2, 7, 14, and 28 snapshots, respectively, to calculate \mathbf{K} . Replica vector depth discretization is changed from 10 m to 2 m to improve visualization. Only the maximum value of each Bartlett ambiguity surface is plotted. Using the Lasso path, the first sparse solution for CS is plotted in Figs. 8(f)–8(j) with the same number of snapshots in \mathbf{Y} to allow for a direct comparison. Increasing the number of snapshots results in fewer false localizations for both processors and tracks using the same number of snapshots are comparable in the presence of environmental uncertainty.

To illustrate the adaptive capabilities of CS, we construct several two-source scenarios. A copy of the SWellEx-96 data is added to the original SWellEx-96 data, resulting in a source separation of 1 km. Similar to the simulation, source 2 is the weaker source and located further in range than source 1 (with respect to the VLA). Each snapshot is multiplied by a random phase to ensure the sources are not coherent. The snapshots of source 2 subsequently are added to the snapshots of source 1: $\mathbf{Y} = \mathbf{Y}_1 + \xi \mathbf{Y}_2$. The SSR scale

factor $\xi = 10^{-3/20} \|\mathbf{Y}_1\|_F / \|\mathbf{Y}_2\|_F$, hence, source 2 is 3 dB below the power of source 1. CS operates on \mathbf{Y} directly. For Bartlett and WNC, \mathbf{Y} is used to construct the SCM. For all processors, 28 snapshots are used in the processing.

Ambiguity surfaces in Fig. 9 show scenario 1 as indicated by the arrows in Figs. 8(e) and 8(j). Single frequency (166 Hz) left panels show that the level of the weaker source is competitive to other positions. The Bartlett processor [Fig. 9(a)] exhibits the most ambiguity. The WNC -6 dB (not shown) is too sensitive and misses source 1, which is found using a constraint of -2 dB [Fig. 9(b)]. While WNC -2 dB is able to reduce localization ambiguity in comparison to Bartlett, the adaptive processor displays many ambiguous positions for the weaker source. The CS performance [Fig. 9(c)] is comparable to the WNC performance in the number and positions of ambiguous locations.

To improve localization performance for source 2, we process additional frequencies and sum the corresponding Bartlett and WNC ambiguity surfaces incoherently, and similarly combine frequencies via Eq. (12) with CS. Increasing the number of processed frequencies from one to six [Figs. 9(d)–9(f)] increases the processor output power of source 2. While many ambiguous positions with similar or higher power than source 2 remain for Bartlett, both WNC and CS offer a degree of robustness and are able to discriminate against some of the false peaks. The solutions offered by WNC and CS are similar for true and ambiguous source locations.

Figure 10 presents two additional multi-frequency results. Panels on the left show scenario 2 and panels on the right show scenario 3 [see arrows in Figs. 8(e) and 8(j)]. While Bartlett exhibits a large number of ambiguous positions with similar or greater power than source 2, both WNC and CS offer solutions with fewer false localizations. CS displays the least number of ambiguous peaks.

V. DISCUSSION

CS and Bartlett yield the same localization results for a single source as demonstrated analytically (see the

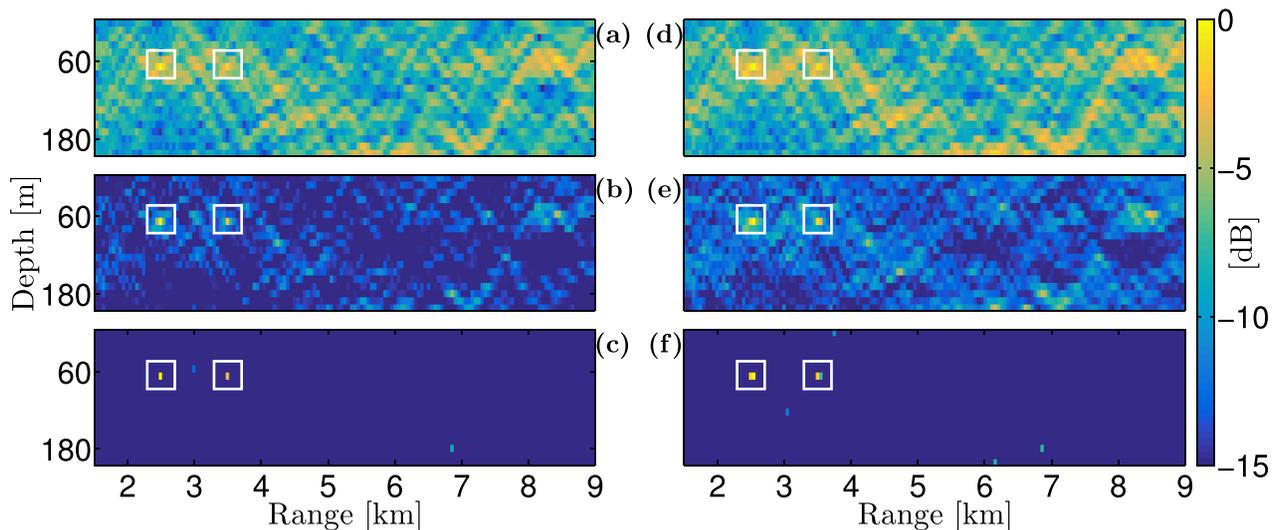


FIG. 7. (Color online) Localization simulation with replica mismatch at SNR 0 dB and 166 Hz showing (a),(d) Bartlett, (b),(e) WNC -2 dB, and (c),(f) CS. True positions are indicated by white squares. Left panels: sources are stationary. Right panels: source snapshots are drawn from two adjacent discretization cells.

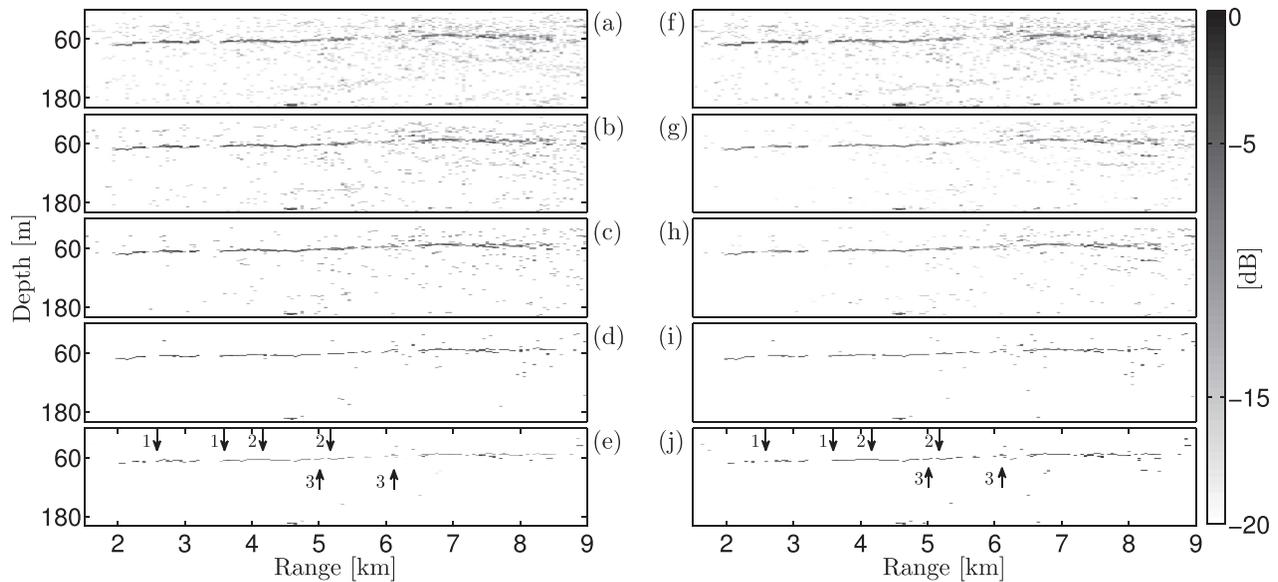


FIG. 8. Localization results for depth vs range of the ~ 60 m deep SWellEx-96 source at 166 Hz for (left) Bartlett and (right) CS using (a),(f) 1, (b),(g) 2, (c),(h) 7, (d),(i) 14, and (e),(j) 28 snapshots. Only the maxima of the Bartlett ambiguity surfaces are used to construct the left panels. The source briefly stopped transmitting tonals at ranges 2.4, 3.5, 5.8, and 6.3 km. Arrows indicate data combinations used for three two-source scenarios.

Appendix) with simulated and SWellEx-96 data. Figure 4(a) shows that Bartlett and CS performance is identical. Processor performance degrades at the same SNR threshold and at higher SNRs, the localization error is quantized within one discretization cell. These results extend to multiple frequencies in Fig. 4(c). Figure 8 shows that both processors exhibit comparable performance for any number of snapshots in scenarios with environmental uncertainty. CS yields the same localization result as the Bartlett processor at an increased computational cost.

The two-source scenarios demonstrate that CS and WNC performance is similar and CS often displays less ambiguous peaks than WNC. Forero and Baxley (2014) showed that a theoretical connection can be established between the elastic-net regularized CS framework and the

WNC processor. Simulation and data results here suggest that there exists a connection even without the additional elastic-net regularizer [Forero and Baxley (2014) also observed good results when setting the elastic net regularizer close to zero]. In our simulations [Fig. 4(b)], CS and WNC locate the weaker source while Bartlett has a RMSE greater than 1 km. Figures 4(b) and 4(d) further indicate that CS has a SNR advantage over the WNC. The WNC performance depends on the selected constraint: decreasing the constraint from -2 to -6 dB reduces localization performance by about 2 dB. CS possesses properties similar to an adaptive processor without being subjected to the constraint-dependent SNR penalty encountered by the WNC.

The ℓ_1 minimization has been extensively applied to approximate the ℓ_0 -norm for signal recovery and variable

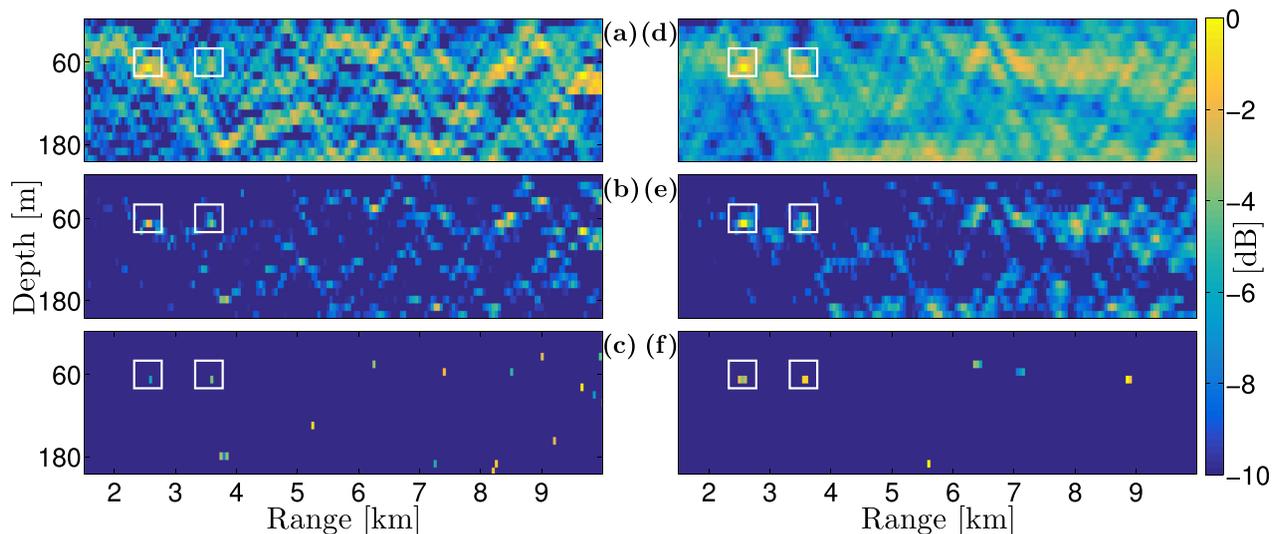


FIG. 9. (Color online) Localization for two SWellEx-96 sources (scenario 1). True positions are indicated by white squares. Panels on the left use a single frequency (166 Hz), panels on the right use six frequencies (94, 112, 130, 148, 166, 235 Hz): (a),(d) Bartlett, (b),(e) WNC -2 dB, and (c),(f) CS. Incorporating multiple frequencies reduces ambiguity and helps localize source 2.

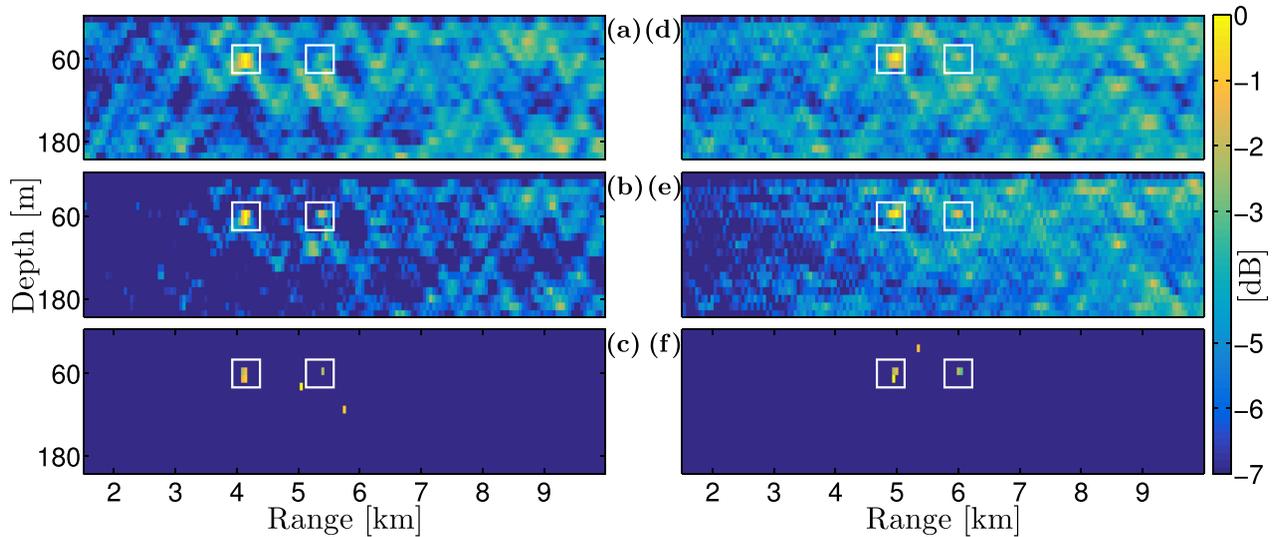


FIG. 10. (Color online) Localization for two SWellEx-96 sources using six frequencies (94, 112, 130, 148, 166, 235 Hz). True positions are indicated by white squares. Left panels show scenario 2 and right panels show scenario 3 for: (a),(d) Bartlett, (b),(e) WNC -2 dB, and (c),(f) CS. Bartlett displays the most ambiguity while both WNC and CS exhibit good performance with fewer false localizations.

selection in regression (Eldar and Kutyniok, 2013). Results in Gerstoft *et al.* (2015) indicate that CS performs closely to an ℓ_0 -norm (demonstrated by means of an exhaustive search) on similar SNR curves (for plane wave beamforming). Results in this paper complement their MV results (which the WNC -6 dB approximates, see Sec. II A) and further illustrate why these adaptive processors exhibit a threshold in localization error performance at higher SNR than CS. The selected WNC constraint essentially is a sidelobe suppression tuning parameter. The constraint should be selected small enough to suppress sidelobes and large enough to maintain localization performance at low SNR. For the presented simulation and data scenarios, a constraint of -2 dB yielded good results. For CS the regularization parameter λ controls the number of sparse solutions which is selected to localize source 2 for the two-source scenarios. In data cases where some degree of mismatch is present, the number of sparse solutions might be difficult to set ahead of time if some solutions are due to ambiguous localizations.

Figure 9(a)–9(c) displays single frequency data scenario 1, which include ambiguous localizations. Bartlett exhibits numerous sidelobes with greater power than the weaker source. Both WNC -2 dB and CS localize source 2 along with many other false peaks. Processing multiple frequencies [Figs. 9(d)–9(f)] helps to localize the weaker source, and both WNC and CS are able to reduce the number of ambiguous peaks. Data results and mismatch simulation Fig. 7 show that the solutions offered by WNC and CS are similar, demonstrating that CS is able to discriminate against sidelobes. Unlike WNC, CS yields a complex source amplitude for every snapshot vector that might be useful for further processing [see Eq. (11)].

Extending the row-sparsity constraint to include snapshots corresponding to multiple frequencies with a single regularizer [Eq. (12)] yields good results in realistic scenarios. Data results indicate that the CS solutions [Fig. 10(c)] still coincide with those of the WNC [Fig. 10(b)] even if the source

power is split between range-depth cells. Multi-frequency solutions in Fig. 9(f) are adjacent to each other, which might indicate that the source position is at the boundary between adjacent range cells. CS finds significant energy in the adjacent cell and treats it as a new sparse solution. This also might be due to a slight misalignment of the mainlobe of each frequency. The peaks of both sources extend over adjacent range cells for multi-frequency Figs. 9(d)–9(f) compared to the single frequency Figs. 9(a)–9(c).

This behavior also is evident in Fig. 10, displaying processor comparison for two additional scenarios when the sources are separated 1 km in range. Source 1 in Fig. 10 is localized in two depth cells on most panels, while both sources exclusively are in a deeper depth cell in Fig. 9. The transition to a deeper depth cell might be an artifact of using a range-independent environmental model (D’Spain *et al.*, 1999).

Incorporating multiple realizations and multiple frequencies using a row-sparsity constraint improves source localization as demonstrated with the simulations and processing the SWellEx-96 data. Lacking required spatial and temporal parameter knowledge for a detailed environmental model, it is likely that the source pressure field will not exactly match the corresponding range-depth cell replica. While replica mismatch may cause unexpected behavior, CS yields reasonable results for the processed data set.

VI. SUMMARY

For the matched field processing application, we demonstrate that CS has characteristics similar to an adaptive processor. CS performs slightly better than the WNC processor under ideal conditions and often displays less ambiguity using data for two incoherent sources. CS performance is equivalent to the Bartlett processor for a single source scenario for any number of snapshots and frequencies at an increased computational burden. The row-sparsity constraint combines multiple temporal snapshots in order to increase

localization performance at low SNR. This formalism has been extended to combine snapshots corresponding to multiple frequencies, which reduces ambiguous peaks and further increases localization performance of a weaker source. In addition, CS appears robust to modest data-replica mismatch and situations when multiple snapshots correspond to the source(s) occupying adjacent range-depth cells at the expense of possible additional solutions.

ACKNOWLEDGMENTS

The authors thank William A. Kuperman, Heechun Song, Santosh Nannuru, Jeff Tippmann, and Ned Richards for their helpful discussions related to MFP and CS. The authors thank David Ensberg for providing help with the SWellEx-96 data set. We also thank the three anonymous reviewers for their constructive comments, which helped to improve the manuscript. This work was supported by the Office of Naval Research under Grant No. N00014-13-1-0632.

APPENDIX: ONE SOURCE FOR CS AND BARTLETT

Consider an arbitrary data matrix \mathbf{Y} , consisting of sources and additive noise. For locating the strongest source, only one row in \mathbf{X} is active, row j , and this row vector is denoted \mathbf{X}_j . Then the data fit is minimized for CS,

$$\begin{aligned}\Phi &= \|\mathbf{Y} - \mathbf{a}_j \mathbf{X}_j\|_F^2 = \text{Tr}[(\mathbf{Y} - \mathbf{a}_j \mathbf{X}_j)^H (\mathbf{Y} - \mathbf{a}_j \mathbf{X}_j)] \\ &= \text{Tr}[\mathbf{Y}^H \mathbf{Y} + \mathbf{X}_j^H \mathbf{a}_j^H \mathbf{a}_j \mathbf{X}_j - \mathbf{X}_j^H \mathbf{a}_j^H \mathbf{Y} - \mathbf{Y}^H \mathbf{a}_j \mathbf{X}_j].\end{aligned}\quad (\text{A1})$$

Solving

$$\frac{\partial \Phi}{\partial \mathbf{X}_j^H} = \mathbf{a}_j^H \mathbf{a}_j \mathbf{X}_j - \mathbf{a}_j^H \mathbf{Y} = 0 \quad (\text{A2})$$

gives the sparse estimate

$$\rightarrow \mathbf{X}_j = \frac{\mathbf{a}_j^H \mathbf{Y}}{\mathbf{a}_j^H \mathbf{a}_j} = \mathbf{a}_j^H \mathbf{Y}, \quad (\text{A3})$$

since we have assumed \mathbf{a}_j is normalized $\|\mathbf{a}_j\| = 1$. Thus, the CS solution seeks the replica vector that is best aligned with the data \mathbf{Y} and gives the solution θ_j . The average power per snapshot is then

$$P_{CS} = \|\mathbf{X}_j\|_2^2 / L = \|\mathbf{a}_j^H \mathbf{Y}\|_2^2 / L. \quad (\text{A4})$$

For the Bartlett processor evaluated at θ_j ,

$$\begin{aligned}P_B(\theta_j) &= \frac{\mathbf{a}_j^H \mathbf{K} \mathbf{a}_j}{\left(\mathbf{a}_j^H \mathbf{a}_j\right)^2} = \frac{\mathbf{a}_j^H \mathbf{Y} \mathbf{Y}^H \mathbf{a}_j}{L} \\ &= \|\mathbf{a}_j^H \mathbf{Y}\|_2^2 / L.\end{aligned}\quad (\text{A5})$$

Comparing this to the CS solution, the Bartlett processor also is maximum at θ_j and has an identical power estimate for one source. However, for multiple sources their solution will differ.

- Baggeroer, A., and Cox, H. (1999). "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," in *33rd Asilomar Conference on Signals, Systems, and Computers* (IEEE, New York), pp. 103–108.
- Baggeroer, A. B., Kuperman, W. A., and Mikhalevsky, P. N. (1993). "An overview of matched field methods in ocean acoustics," *IEEE J. Oceanic Eng.* **18**, 401–424.
- Booth, N. O., Abawi, A. T., Schey, P. W., and Hodgkiss, W. S. (2000). "Detectability of low-level broad-band signals using adaptive matched-field processing with vertical aperture arrays," *IEEE J. Oceanic Eng.* **25**, 296–313.
- Bucker, H. P. (1976). "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *J. Acoust. Soc. Am.* **59**, 368–373.
- Candès, E. J., Eldar, Y. C., Needell, D., and Randall, P. (2011). "Compressed sensing with coherent and redundant dictionaries," *Appl. Comp. Harmonic Anal.* **31**, 59–73.
- Candès, E., and Wakin, M. (2008). "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, 21–30.
- Capon, J. (1969). "High-resolution frequency-wavenumber spectrum analysis," *IEEE Proc.* **57**, 1408–1418.
- Chen, S. S., Donoho, D. L., and Saunders, M. A. (1998). "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.* **20**, 33–61.
- Collins, M. D. (1993). "A split-step Padé solution for the parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742.
- Collins, M. D. (1994). "Generalization of the split-step Padé solution," *J. Acoust. Soc. Am.* **96**, 382–385.
- Cox, H. (2002). "Adaptive beamforming in non-stationary environments," in *36th Asilomar Conference on Signals, Systems and Computers* (IEEE, New York), pp. 431–438.
- Cox, H., Zeskind, R., and Owen, M. (1987). "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 1365–1376.
- David, G., Robert, J.-L., Zhang, B., and Laine, A. F. (2015). "Time domain compressive beam forming of ultrasound signals," *J. Acoust. Soc. Am.* **137**, 2773–2784.
- Davies, M. E., and Eldar, Y. C. (2012). "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory* **58**, 1135–1146.
- Debever, C., and Kuperman, W. A. (2007). "Robust matched-field processing using a coherent broadband white noise constraint processor," *J. Acoust. Soc. Am.* **122**, 1979–1986.
- D'Spain, G. L., Murray, J. J., Hodgkiss, W. S., Booth, N. O., and Schey, P. W. (1999). "Mirages in shallow water matched field processing," *J. Acoust. Soc. Am.* **105**, 3245–3265.
- Edelmann, G. F., and Gaumont, C. F. (2011). "Beamforming using compressive sensing," *J. Acoust. Soc. Am.* **130**, EL232–EL237.
- Eldar, Y., and Kutyniok, G., eds. (2013). *Compressed Sensing, Theory and Applications*, 3rd ed. (Cambridge University Press, Cambridge, UK), pp. 1–544.
- Forero, P. A. (2014). "Broadband underwater source localization via multi-task learning," in *48th Annual Conference on Information Sciences and Systems (CISS)* (IEEE, New York), pp. 1–6.
- Forero, P. A., and Baxley, P. A. (2014). "Shallow-water sparsity-cognizant source-location mapping," *J. Acoust. Soc. Am.* **135**, 3483–3501.
- Fortunati, S., Grasso, R., Gini, F., Greco, M. S., and LePage, K. (2014). "Single-snapshot DOA estimation by using compressed sensing," *EURASIP J. Adv. Signal Process.* **120**, 120.
- Gerstoft, P., Xenaki, A., and Mecklenbräuker, C. F. (2015). "Multiple and single snapshot compressive beamforming," *J. Acoust. Soc. Am.* **138**, 2003–2014.
- Grant, M., and Boyd, S. (2016). "CVX: MATLAB software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx/> (Last viewed June 15, 2016).
- Hursky, P., Hodgkiss, W. S., and Kuperman, W. A. (2001). "Matched field processing with data-derived modes," *J. Acoust. Soc. Am.* **109**, 1355–1366.
- Kaiser, J. F. (1974). "Nonrecursive digital filter design using the Io-sinh window function," in *IEEE Int. Symp. Circuits and Systems (ISCAS'74)*, San Francisco, pp. 20–23.
- Krolik, J. L. (1992). "Matched-field minimum variance beamforming in a random ocean channel," *J. Acoust. Soc. Am.* **92**, 1408–1419.
- Le Courtois, F., and Bonnel, J. (2015). "Compressed sensing for wideband wavenumber tracking in dispersive shallow water," *J. Acoust. Soc. Am.* **138**, 575–583.
- Liu, C., Zakharov, Y. V., and Chen, T. (2012). "Broadband underwater localization of multiple sources using basis pursuit de-noising," *IEEE Trans. Signal Process.* **60**, 1708–1717.

- Maksym, J. (1979). "A robust formulation of an optimum cross-spectral beamformer for line arrays," *J. Acoust. Soc. Am.* **65**, 971–975.
- Malioutov, D., Çetin, M., and Willsky, A. (2005). "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.* **53**, 3010–3022.
- Mallat, S., and Zhang, Z. (1993). "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.* **41**, 3397–3415.
- Orris, G. J., Nicholas, M., and Perkins, J. S. (2000). "The matched-phase coherent multi-frequency matched-field processor," *J. Acoust. Soc. Am.* **107**, 2563–2575.
- Pillai, S., and Kwon, B. (1989). "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Process.* **37**, 8–15.
- Schmidt, R. (1986). "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.* **34**, 276–280.
- Song, H. C., de Rosny, J., and Kuperman, W. A. (2003a). "Improvement in matched field processing using the CLEAN algorithm," *J. Acoust. Soc. Am.* **113**, 1379–1386.
- Song, H., Kuperman, W. A., Hodgkiss, W. S., Gerstoft, P., and Kim, J. S. (2003b). "Null broadening with snapshot-deficient covariance matrices," *IEEE J. Oceanic Eng.* **28**, 250–261.
- Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso," *J. R. Statist. Soc. B* **58**, 267–288.
- Xenaki, A., and Gerstoft, P. (2015). "Grid-free compressive beamforming," *J. Acoust. Soc. Am.* **137**, 1923–1935.
- Xenaki, A., Gerstoft, P., and Mosegaard, K. (2014). "Compressive beamforming," *J. Acoust. Soc. Am.* **136**, 260–271.
- Xu, H., Caramanis, C., and Mannor, S. (2010). "Robust regression and lasso," *IEEE Trans. Inf. Theory* **56**, 3561–3574.
- Zou, H., and Hastie, T. (2005). "Regularization and variable selection via the elastic net," *J. R. Statist. Soc. B* **67**, 301–320.