Sparse Bayesian learning for beamforming using sparse linear arrays

Santosh Nannuru, Ali Koochakzadeh, Kay L. Gemba, Piya Pal, and Peter Gerstoft

Citation: The Journal of the Acoustical Society of America **144**, 2719 (2018); doi: 10.1121/1.5066457
View online: https://doi.org/10.1121/1.5066457
View Table of Contents: http://asa.scitation.org/toc/jas/144/5
Published by the Acoustical Society of America

Sparse Bayesian learning for beamforming using sparse linear arrays

Santosh Nannuru,[1,a)] Ali Koochakzadeh,[2] Kay L. Gemba,[3] Piya Pal,[2] and Peter Gerstoft[3]
[1]*Signal Processing and Communications Research Center, IIIT Hyderabad, India*
[2]*ECE Department, University of California San Diego, La Jolla, California 92093, USA*
[3]*Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA*

(Received 1 June 2018; revised 20 September 2018; accepted 14 October 2018; published online 8 November 2018)

Sparse linear arrays such as co-prime and nested arrays can resolve more sources than the number of sensors. In contrast, uniform linear arrays (ULA) cannot resolve more sources than the number of sensors. This paper demonstrates this using Sparse Bayesian learning (SBL) and co-array MUSIC for single frequency beamforming. For approximately the same number of sensors, co-prime and nested arrays are shown to outperform ULA in root mean squared error. This paper shows that multi-frequency SBL can significantly reduce spatial aliasing. The effects of different sparse sub-arrays on SBL performance are compared qualitatively using the Noise Correlation 2009 experimental data set. © *2018 Acoustical Society of America*.
https://doi.org/10.1121/1.5066457

[JL] Pages: 2719–2729

I. INTRODUCTION

Compressive sensing is a processing paradigm to estimate high dimensional sparse vectors using as few measurements as possible. Significant research has been performed to develop algorithms for this task, some of which include basis pursuit,[1] matching pursuit,[2] and sparse Bayesian learning.[3] These algorithms have been used in various physical applications for parameter estimation,[4–7] including beamforming.[8–10] We use the term beamforming and direction-of-arrival (DOA) estimation interchangeably in this paper. In these applications, the traditional (usually) non-linear problem is reformulated as an underdetermined linear problem and its sparse solutions provide the required parameter estimates.

Sparse Bayesian learning (SBL)[3,5] is a compressive sensing technique that uses a Bayesian framework to find a sparse solution to an underdetermined linear problem. In the context of DOA estimation using uniform linear array (ULA) sensor data, it has been applied to resolve nearby sources.[8,11,12] A limitation of ULA is that the maximum number of resolvable sources is limited by the number of sensors. This is confirmed for SBL as well using a Cramér-Rao based analysis.[13–15]

SBL has many advantages over other array processing methods. SBL estimates the covariance of the weights rather than the weights themselves, significantly reducing the number of parameters to estimate. SBL uses iterative updates for parameter estimation that is computationally faster compared to convex optimization used by basis pursuit. Algorithms such as MUSIC require many snapshots to build the covariance matrix, whereas SBL can show better performance with fewer snapshots. It is also less sensitive to array geometry, unlike MUSIC, which typically requires ULA.

Recently, various sparse array geometries have been proposed such as nested arrays[16] and co-prime arrays,[17] which can resolve more sources than the number of sensors.[18–25] This has been demonstrated with MUSIC,[26–29] SBL,[30] and justified theoretically[31,32] using the co-array covariance matrix.

We use SBL to directly process observation vectors from nested and co-prime arrays without constructing a co-array based covariance matrix. We also apply MUSIC to the contiguous portion of difference co-array. For plane wave arrivals, the difference co-array covariance matrix is Hermitian Toeplitz. Thus, this covariance matrix can be formed by identifying the relevant entries in the physical sensor array covariance matrix.

We show using simulations that both SBL and co-array MUSIC can identify more sources than the number of sensors. A root mean squared error (RMSE) comparison of ULA, nested and co-prime arrays, all with the same number of sensors, is performed with respect to number of snapshots, number of sources, and signal-to-noise ratio. We demonstrate utility of sparse arrays in combination with multi-frequency SBL by processing Noise Correlation 2009 (NC09) experiment data.[33] Various sparse arrays including co-prime, nested, and ULA (with large inter-sensor spacing) are constructed by appropriately sampling the full ULA. The robustness of DOA estimation is increased by processing multi-frequency data.[34,35] Parts of this paper are published in a conference proceeding.[36]

This paper is organized as follows: Sec. II gives a brief overview of the signal model and co-prime and nested sparse linear arrays. It also discusses co-array processing and the MUSIC algorithm. The SBL algorithm is presented in Sec. III and a pseudocode is provided for implementation. The multi-frequency SBL is also discussed. Beamforming simulations are performed in Sec. IV to study SBL and MUSIC with respect to various parameters while processing observations from sparse arrays. Section V applies SBL to process

a)Electronic mail: santosh.nannuru@iiit.ac.in

experimental ship noise data for DOA estimation using various array geometries. Conclusions are discussed in Sec. VI.

II. SIGNAL MODEL AND SPARSE LINEAR ARRAYS

A. Signal model

The lth observation snapshot $\mathbf{y}_l$ recorded at a sensor array due to impinging plane waves is given by

$$\mathbf{y}_l = \mathbf{A}\mathbf{x}_l + \mathbf{n}_l, \tag{1}$$

where $\mathbf{y}_l \in \mathbb{C}^N$, N is the number of sensors in the array, $\mathbf{x}_l \in \mathbb{C}^M$, M is the number of grid points in which the angle space $[-90, 90]$ is divided, and $\mathbf{n}_l$ is the additive zero-mean circularly symmetric complex Gaussian noise. The vector $\mathbf{x}_l$ is sparse and has at most $K \ll M$ non-zero entries corresponding to the complex amplitudes of the impinging waves. Moreover, $\mathbf{x}_l$ is a zero-mean random vector such that $\mathbb{E}(\mathbf{x}_l\mathbf{x}_l^H) = \mathbf{\Gamma}$, $\mathbb{E}(\mathbf{n}_l\mathbf{n}_l^H) = \sigma^2\mathbf{I}$, $\mathbb{E}(\mathbf{x}_l\mathbf{n}_l^H) = 0$, where $\mathbf{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$ and $\boldsymbol{\gamma} \in \mathbb{R}^M$ is a sparse vector. The objective is to find the unknown vector $\mathbf{x}_l$ given the observations $\mathbf{y}_l$ and the dictionary $\mathbf{A}$. Typically multiple (L) snapshots are processed $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L]$.

The columns of dictionary $\mathbf{A}$ are composed of the steering vectors corresponding to the M discrete angles $\{\theta_1, \ldots, \theta_M\}$. For a narrow-band plane wave signal of wavelength λ and sensor locations given by $\{d_1, \ldots, d_N\}$, the mth column is

$$\mathbf{a}_m = \left[1, e^{j2\pi(d_1/\lambda)\sin(\theta_m)}, \ldots, e^{j2\pi(d_N/\lambda)\sin(\theta_m)}\right]^T, \tag{2}$$

for $m = 1 \ldots M$, where θ_m is the mth discretized angle. The signal wavelength and sensor locations are assumed to be known.

Let the sensor positions in an array be given by $d_n = z_n d$ where the integer z_n is the normalized sensor location of nth sensor and d is the minimum sensor spacing. Define the set $\mathbb{S} = \{z_n | n = 1, \ldots, N\}$. A ULA consists of uniformly spaced sensors with $z_n = n$ and

$$\mathbb{S}_{\mathrm{ULA}} = \{1, 2, \ldots, N\}. \tag{3}$$

Typically the minimum spacing d is chosen to be $d = \lambda/2$. We now briefly discuss the sparse linear arrays of co-prime[17] and nested arrays.[16]

B. Co-prime array

A co-prime array consists of two ULAs with spacing $N_1^c(\lambda/2)$ and $N_2^c(\lambda/2)$ such that N_1^c and N_2^c are co-prime (i.e., their greatest common divisor is 1). Also let $N_1^c > N_2^c$ without loss of generality. The sensors for a co-prime array are given by the following set

$$\begin{aligned}\mathbb{S}_{\mathrm{coprime}} = &\{1, N_2^c + 1, \ldots, (N_1^c - 1)N_2^c + 1\} \\ &\cup \{N_1^c + 1, 2N_1^c + 1, \ldots, (2N_2^c - 1)N_1^c + 1\}.\end{aligned} \tag{4}$$

A co-prime array has a total of $(N_1^c + 2N_2^c - 1)$ sensors. Figure 1 shows an example of co-prime array with $N_1^c = 5$, $N_2^c = 2$.

C. Nested array

A nested array also consists of two ULAs, the dense ULA portion with N_1^n sensors (spacing $\lambda/2$), and the second ULA portion with N_2^n sensors [spacing $(N_1^n + 1)\lambda/2$]. The sensors for a nested array are given by

$$\begin{aligned}\mathbb{S}_{\mathrm{nested}} = &\{1, 2, \ldots, N_1^n\} \\ &\cup \{(N_1^n + 1), 2(N_1^n + 1), \ldots, (N_1^n + 1)N_2^n\}.\end{aligned} \tag{5}$$

A nested array consists of a total of $(N_1^n + N_2^n)$ sensors. An example of a nested array is shown in Fig. 1 with $N_1^n = 4, N_2^n = 3$.

D. Co-array and holes

Let the set $\mathbb{D}$ denote the difference set (co-array) corresponding to the set $\mathbb{S}$ defined as

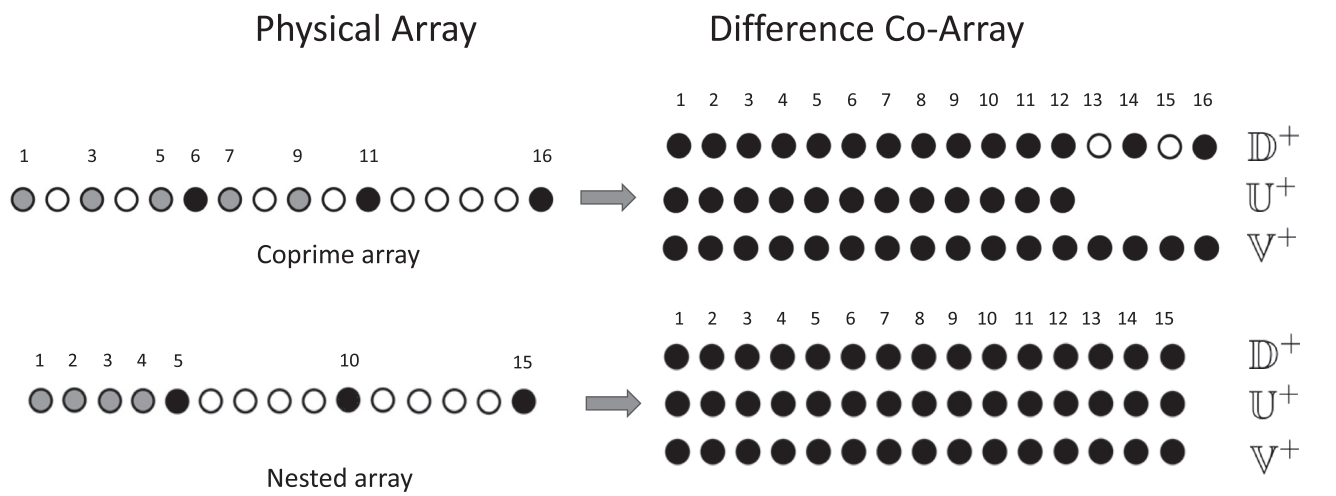

FIG. 1. Used co-prime and nested arrays and their corresponding difference co-arrays. In the physical array, the shaded and full circles correspond to the first and second ULA in each sparse array corresponding to the set S. In the difference co-array, we construct the three sets $\mathbb{D}^+$, $\mathbb{U}^+$, and $\mathbb{V}^+$. The empty circles in the difference co-array show the holes. For co-prime array $N_1^c = 5, N_2^c = 2$ and for nested array, $N_1^n = 4, N_2^n = 3$.

$$\mathbb{D} = \{z_i - z_j | 1 \le i, j \le N\}, \tag{6}$$

containing all the pairwise differences of the normalized sensor locations. In order to more carefully study the structure of $\mathbb{D}$, we adopt the following definition:[37]

Definition 1. *For a sensor array* $\mathbb{S}$*, with difference coarray* $\mathbb{D}$*, the maximum central contiguous ULA in* $\mathbb{D}$ *is defined as* $\mathbb{U} = \{r | \{-|r|, \ldots, 0, \ldots, |r|\} \subseteq \mathbb{D}\}$. *Moreover, the shortest ULA containing* $\mathbb{D}$ *is denoted as* $\mathbb{V} = \{r | \min(\mathbb{D}) \le r \le \max(\mathbb{D})\}$.

For any array we have $\mathbb{U} \subseteq \mathbb{V}$, where $\subseteq$ denotes the subset operator. For certain array geometries (such as ULA or nested arrays), the equality $\mathbb{U} = \mathbb{V} = \mathbb{D}$ holds, meaning that the difference co-array does not have any "holes."[16] However, in certain cases (such as coprime arrays), the difference co-array can contain holes, i.e., $\mathbb{U} \subset \mathbb{V}$, and $\mathbb{U} \neq \mathbb{V}$. An example of $\mathbb{U}, \mathbb{V}, \mathbb{D}$ for a co-prime array and a nested array is shown in Fig. 1. Let $\mathbb{U}^+, \mathbb{V}^+$, and $\mathbb{D}^+$ represent the nonnegative subsets of $\mathbb{U}$, $\mathbb{V}$, and $\mathbb{D}$.

E. Co-array processing

Traditional DOA estimation methods such as MUSIC are based on the second order moment of the data[38] and work for ULAs. The number of sources that can be resolved depends on the degrees of freedom in the covariance matrix (which is equal to the number of sensors for a ULA). In contrast, for certain sparse arrays such as co-prime array, the covariance matrix has more degrees of freedom than the number of sensors in the sparse array. Exploiting this property of the second order moment allows us to resolve more sources than the number of sensors for a sparse array.

From Eq. (1), the $N \times N$ dimensional covariance matrix of $\mathbf{y}_l$ is

$$\begin{aligned} \mathbf{R}_{\mathbb{S}} &= \mathbb{E}(\mathbf{y}_l \mathbf{y}_l^H) = \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^H + \sigma^2\mathbf{I} \\ &= \sum_{m \in \mathcal{M}} \gamma_m \mathbf{a}_m \mathbf{a}_m^H + \sigma^2\mathbf{I}, \end{aligned} \tag{7}$$

where $\mathcal{M}$ represent the active sources (non zero elements of $\boldsymbol{\gamma}$). The elements of the matrix $\mathbf{a}_m\mathbf{a}_m^H$ has entries of the form $e^{j(2\pi/\lambda)(d_{n_1} - d_{n_2})\sin(\theta_m)}$ for $z_{n_1}, z_{n_2} \in \mathbb{S}$. Thus, the entries only depend on the differences in sensor locations. We make the following observations:

(1) The upper triangular part (including the diagonal) of the matrix $\mathbf{R}_{\mathbb{S}}$ corresponds to the set $\mathbb{D}^+$, and the lower triangular part (including the diagonal) corresponds the set $\mathbb{D}^-$, where $\mathbb{D}^-$ denotes the nonpositive subset of $\mathbb{D}$.
(2) For a ULA, the covariance matrix is Hermitian Toeplitz and can be fully constructed given its first column $\mathbf{r}_{\mathbb{S}}$.
(3) For sparse arrays, $\mathbf{R}_{\mathbb{S}}$ is not Hermitian Toeplitz in general, but by appropriately choosing from its entries we can construct a column vector $\mathbf{r}_{\mathbb{U}^+}$ for the central contiguous ULA segment $\mathbb{U}^+$ and the corresponding $|\mathbb{U}^+| \times |\mathbb{U}^+|$ dimensional Hermitian Toeplitz matrix $\mathbf{R}_{\mathbb{U}^+}$

$$\mathbf{R}_{\mathbb{U}^+} = \begin{pmatrix} \mathbf{r}_{\mathbb{U}^+,0} & \mathbf{r}^H_{\mathbb{U}^+,1} & \cdots & \mathbf{r}^H_{\mathbb{U}^+,|\mathbb{U}^+|} \\ \mathbf{r}_{\mathbb{U}^+,1} & \mathbf{r}_{\mathbb{U}^+,0} & \cdots & \mathbf{r}^H_{\mathbb{U}^+,|\mathbb{U}^+|-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_{\mathbb{U}^+,|\mathbb{U}^+|} & \mathbf{r}_{\mathbb{U}^+,|\mathbb{U}^+|-1} & \cdots & \mathbf{r}_{\mathbb{U}^+,0} \end{pmatrix}, \tag{8}$$

where $\mathbf{r}_{\mathbb{U}^+,n}$ denotes the nth element of vector $\mathbf{r}_{\mathbb{U}^+}$.

In a practical setting, however, only a finite number of snapshots are available. We compute the sample covariance matrix (SCM) using the samples $\mathbf{y}_l$, for $l = 1, \ldots, L$

$$\hat{\mathbf{R}}_{\mathbb{S}} = \frac{1}{L}\sum_{l=1}^{L} \mathbf{y}_l \mathbf{y}_l^H. \tag{9}$$

From $\hat{\mathbf{R}}_{\mathbb{S}}$ we can form a sample auto correlation vector $\hat{\mathbf{r}}_{\mathbb{U}^+}$ (and hence $\hat{\mathbf{R}}_{\mathbb{U}^+}$) for the $\mathbb{U}^+$ part of the contiguous co-array by averaging of relevant elements from $\hat{\mathbf{R}}_{\mathbb{S}}$, see Refs. 28 and 37. The nth entry of $\hat{\mathbf{r}}_{\mathbb{U}^+}$ can be computed as

$$\hat{\mathbf{r}}_{\mathbb{U}^+,n} = \frac{1}{k_n} \sum_{n_1 - n_2 = n} \hat{\mathbf{R}}_{\mathbb{S}}(n_1, n_2), \tag{10}$$

$$k_n = |\{(n_1, n_2) | n_1 - n_2 = n, n_1, n_2 \in \mathbb{S}\}|. \tag{11}$$

For example, for the co-prime array shown in Fig. 1, $|\mathbb{U}^+| = 12$, and the values of k_n for $n = 1$ to 12 are 8, 2, 5, 2, 4, 3, 3, 1, 2, 1, 2, and 1. Note that the number of terms k_n contributing towards the average in Eq. (10) are significantly different, thus complicating the statistical properties of $\hat{\mathbf{R}}_{\mathbb{U}^+}$. The empirical SCM $\hat{\mathbf{R}}_{\mathbb{U}^+}$ can be used with MUSIC. The contiguous difference co-array $\mathbb{U}^+$ avoids aliasing as well as provides higher degrees of freedom to resolve more sources than sensors (N), while the large aperture ($|\mathbb{U}^+|$) provides resolution gain.

F. MUSIC

Direct application of MUSIC to sparse array data can resolve few sources as it is limited by the number of sensors. To resolve more sources than the number of sensors, MUSIC can be applied to the central ULA portion of the co-array[26,28] or to the covariance matrix obtained after co-array interpolation.[37]

Given an empirical covariance matrix $\mathbf{R}$ and the number of sources K, if $\mathbf{B_n}$ is the $N \times (N - K)$ dimensional matrix of its noise eigenvectors, then the MUSIC spectrum is given by

$$P(\theta) = \frac{1}{\mathbf{a}(\theta)^H \mathbf{B_n} \mathbf{B_n}^H \mathbf{a}(\theta)}, \tag{12}$$

where $\mathbf{a}(\theta)$ is the steering vector at angle θ. To apply MUSIC, the number of sources K is required to be known.

III. SPARSE BAYESIAN LEARNING

The multi-snapshot signal model is given by

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}, \tag{13}$$

where the noise $\mathbf{N} = [\mathbf{n}_1, \ldots, \mathbf{n}_L]$ is zero-mean complex Gaussian with variance σ^2, $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{n}_l; 0, \sigma^2\mathbf{I})$; $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_L]$ is the matrix of sparse weights with all the columns sharing the same sparsity profile. The observations are assumed to be independent across snapshots giving the multi-snapshot likelihood function

$$p(\mathbf{Y}|\mathbf{X}) = \prod_{l=1}^{L} p(\mathbf{y}_l|\mathbf{x}_l) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_l; \mathbf{A}\mathbf{x}_l, \sigma^2\mathbf{I}). \tag{14}$$

Prior: In SBL, the sparse weights are treated as zero-mean complex Gaussian random vectors with diagonal covariance $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1 \ldots \gamma_M) = \mathrm{diag}(\boldsymbol{\gamma})$. The prior model is given by

$$p(\mathbf{X}) = \prod_{l=1}^{L} p(\mathbf{x}_l) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{x}_l; \mathbf{0}, \mathbf{\Gamma}). \tag{15}$$

Evidence: From the Gaussian prior [Eq. (15)] and likelihood [Eq. (14)], the evidence $p(\mathbf{Y})$ is Gaussian and given by

$$p(\mathbf{Y}) = \int p(\mathbf{X})p(\mathbf{Y}|\mathbf{X})d\mathbf{X} = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_l; \mathbf{0}, \Sigma_{\mathbf{y}}), \tag{16}$$

where $\Sigma_{\mathbf{y}} = \sigma^2\mathbf{I} + \mathbf{A}\mathbf{\Gamma}\mathbf{A}^H$. The SBL approach is to estimate the diagonal entries of $\mathbf{\Gamma}$ by maximizing the (log) evidence

$$\begin{aligned}(\hat{\gamma}_1 \ldots \hat{\gamma}_M) &= \arg\max_{\boldsymbol{\gamma}} \log\ p(\mathbf{Y}) \\ &= \arg\max_{\boldsymbol{\gamma}} \left\{ -\sum_{l=1}^{L} (\mathbf{y}_l^H \Sigma_{\mathbf{y}}^{-1} \mathbf{y}_l + \log|\Sigma_{\mathbf{y}}|) \right\}. \end{aligned} \tag{17}$$

Differentiating the above objective function and equating the derivatives to zero gives the fixed point update rule[3,5,11]

$$\begin{aligned}\gamma_m^{\mathrm{new}} &= \gamma_m^{\mathrm{old}} \frac{1}{L} \frac{||\mathbf{Y}^H \Sigma_{\mathbf{y}}^{-1} \mathbf{a}_m||_2^2}{\mathbf{a}_m^H \Sigma_{\mathbf{y}}^{-1} \mathbf{a}_m} \\ &= \gamma_m^{\mathrm{old}} \frac{\mathrm{Tr}\left[\mathbf{S}_{\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \mathbf{a}_m \mathbf{a}_m^H \Sigma_{\mathbf{y}}^{-1}\right]}{\mathbf{a}_m^H \Sigma_{\mathbf{y}}^{-1} \mathbf{a}_m}, \end{aligned} \tag{18}$$

where $\mathbf{S}_{\mathbf{y}} = (1/L)\mathbf{Y}\mathbf{Y}^H$ is the SCM and Tr[·] denotes the trace operator for a matrix. The pseudocode of the SBL algorithm is given in Table I. The noise variance is also required to be estimated if unknown, see Sec. III A. The parameters ε and N_t are the convergence error threshold and number of iterations of the algorithm. The unknown vector $\boldsymbol{\gamma}$ is initialized to the unnormalized conventional beamformer (CBF) output.

To resolve more sources than the number of sensors, SBL algorithm in Table I can directly be applied to observations from co-prime and nested arrays. This is because the SBL update rule in Eq. (18) depends on the sample covariance matrix $\mathbf{S}_{\mathbf{y}}$, which has more degrees of freedom than the raw observations $\mathbf{Y}$. The dimension of the covariance matrix required by SBL is $O((N_1 + N_2)^2)$.

By comparison, MUSIC based on direct SCM can only find up to $O(N_1 + N_2)$ sources as there are at most $O(N_1 + N_2)$ eigenvalues. MUSIC with co-array based covariance can resolve more sources than number of sensors.[26] This requires construction of a higher dimensional covariance matrix of size $O((N_1N_2)^2)$ from the smaller $O((N_1 + N_2)^2)$ dimensional direct SCM. SBL does not rely on eigendecomposition and is thus able to extract a higher number of sources from the smaller, direct SCM itself. Alternately, a covariance based LASSO[39] could recover more sources but the computational costs would be higher than for SBL (see Ref. 11, Fig. 3).

TABLE I. SBL algorithm pseudocode: Input consists of data $\mathbf{Y}$, dictionary $\mathbf{A}$, and if available, noise variance σ^2. Convergence is controlled by the error threshold ϵ and maximum number of iterations N_t.

SBL Algorithm
1. Input: $\mathbf{Y}$, $\mathbf{A}$, σ^2 (optional)
2. Parameters: $\epsilon = 10^{-3}$, $N_t = 500$
3. Initialization: $\gamma_m^{\mathrm{old}} = \mathbf{a}_m^H \mathbf{S}_{\mathbf{y}} \mathbf{a}_m$, $\forall m$
4. **for** $i = 1$ to N_t
5. Compute: $\Sigma_{\mathbf{y}} = \sigma^2\mathbf{I}_N + \mathbf{A}\mathbf{\Gamma}^{\mathrm{old}}\mathbf{A}^H$
6. γ_m^{new} update $\forall m$ using Eq. (18)
7. σ^2 update using Eq. (19)
8. If $\|\boldsymbol{\gamma}^{\mathrm{new}} - \boldsymbol{\gamma}^{\mathrm{old}}\|_1 / \|\boldsymbol{\gamma}^{\mathrm{old}}\|_1 < \epsilon$, **break**
9. $\boldsymbol{\gamma}^{\mathrm{old}} = \boldsymbol{\gamma}^{\mathrm{new}}$, $\mathbf{\Gamma}^{\mathrm{old}} = \mathrm{diag}(\boldsymbol{\gamma}^{\mathrm{new}})$
10. **end**
11. Output: $\boldsymbol{\gamma}$, σ^2 (optional)

A. Noise estimate

We use a stochastic maximum likelihood based method to estimate the noise variance σ^2. Let $\mathbf{A}_{\mathcal{M}}$ be formed by K columns of $\mathbf{A}$ indexed by $\mathcal{M}$, where the set $\mathcal{M}$ indicates the location of non-zero entries of $\mathbf{x}$ with cardinality $|\mathcal{M}| = K$. The noise variance estimate is then[8,11,40]

$$\hat{\sigma}^2 = \frac{\mathrm{Tr}\left[(\mathbf{I}_N - \mathbf{A}_{\mathcal{M}}\mathbf{A}_{\mathcal{M}}^{+})\mathbf{S}_{\mathbf{y}}\right]}{N - K}, \tag{19}$$

where $\mathbf{A}_{\mathcal{M}}^{+}$ denotes the Moore-Penrose pseudo-inverse. Note that this noise estimate is not valid for more sources than sensors ($K \geq N$). To keep simulations simple, in Sec. IV it is assumed that the noise is known exactly and we focus on estimating $\boldsymbol{\gamma}$. The estimate in Eq. (19) requires knowledge of number of sources K, which could be obtained using model order selection methods[41] such as AIC and BIC.

B. Multi-frequency SBL

Let observations be available at multiple frequencies given by

$$\mathbf{Y}_f = \mathbf{A}_f\mathbf{X}_f + \mathbf{N}_f, \quad f = 1, 2, \ldots, F, \tag{20}$$

where $\mathbf{Y}_f$, $\mathbf{A}_f$, $\mathbf{X}_f$, and $\mathbf{N}_f$ are, respectively, the observations, dictionary, prior, and the noise at the fth frequency. The observations $\mathbf{Y}_{1:F} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_F]$ and the noise $\mathbf{N}_{1:F} = [\mathbf{N}_1, \ldots, \mathbf{N}_F]$ are assumed to be independent across frequencies. Following Eq. (14), the likelihood for fth frequency is given as

$$p(\mathbf{Y}_f|\mathbf{X}_f) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_{fl}; \mathbf{A}_f \mathbf{x}_{fl}, \sigma_f^2 \mathbf{I}), \tag{21}$$

where σ_f^2 is the noise variance at fth frequency.

Prior: Assuming the source variance to be same across frequencies (i.e., flat source power spectrum) the prior is

$$p(\mathbf{X}_f) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{x}_{fl}; \mathbf{0}, \Gamma), \quad f = 1, 2, \ldots, F. \tag{22}$$

The flat power spectrum assumption implies that the covariance parameter Γ is the same across frequencies. This reduces the number of unknown parameters to be estimated while enforcing a common sparsity profile[12,34,42] across all the observations.

Evidence: The multi-frequency evidence is

$$\begin{aligned} p(\mathbf{Y}_{1:F}) &= \prod_{f=1}^{F} \int p(\mathbf{X}_f) p(\mathbf{Y}_f|\mathbf{X}_f) d\mathbf{X}_f \\ &= \prod_{f=1}^{F} \prod_{l=1}^{L} \mathcal{CN}(\mathbf{y}_{fl}; \mathbf{0}, \Sigma_{\mathbf{y}_f}), \end{aligned} \tag{23}$$

where $\Sigma_{\mathbf{y}_f} = \sigma_f^2 \mathbf{I} + \mathbf{A}_f \Gamma \mathbf{A}_f^H$. Proceeding as before [see Eq. (17)], to estimate Γ we maximize the (log) evidence

$$\begin{aligned} (\hat{\gamma}_1 \ldots \hat{\gamma}_M) &= \arg\max_{\gamma} \log\, p(\mathbf{Y}_{1:F}) \\ &= \arg\max_{\gamma} \left\{ -\sum_{f=1}^{F} \sum_{l=1}^{L} (\mathbf{y}_{fl}^H \Sigma_{\mathbf{y}_f}^{-1} \mathbf{y}_{fl} + \log |\Sigma_{\mathbf{y}_f}|) \right\}. \end{aligned} \tag{24}$$

We get the following update rule by differentiation

$$\gamma_m^{\text{new}} = \gamma_m^{\text{old}} \frac{1}{L} \frac{\sum_{f=1}^{F} \sum_{l=1}^{L} |\mathbf{y}_{fl}^H \Sigma_{\mathbf{y}_f}^{-1} \mathbf{a}_{fm}|^2}{\sum_{f=1}^{F} \mathbf{a}_{fm}^H \Sigma_{\mathbf{y}_f}^{-1} \mathbf{a}_{fm}}. \tag{25}$$

We refer to this multi-frequency SBL as SBL MF1. Alternately, we could process observations from each frequency independently and average the estimated γ across frequencies. We will call this SBL MF2. A simulation example of these multi-frequency SBL methods is shown later in Sec. IV D.

Remark: A detailed description of multi-frequency SBL, including simulations and its impact on spatial aliasing, is discussed in another paper.[12] Simulations there clearly show the advantage of using multiple frequencies to get processing gains especially at low signal-to-noise ratio (SNR). Additionally, it is demonstrated that when processing multiple frequencies, since the location of aliased peaks is frequency dependent, jointly processing multi-frequency data using SBL can significantly suppress spatial aliasing.

IV. SIMULATIONS

A. Gram matrices

In simulations, we apply SBL to measurements obtained from three array geometries: uniform linear array (ULA), co-prime array, and nested array. These configurations are inspired from the array geometry constraints we have for the experiment data in Sec. V. Specifically we consider the co-prime array with $N_1^c = 5, N_2^c = 2$ for a total of eight sensors and the nested array is constructed with $N_1^n = 4, N_2^n = 3$ for a total of seven sensors. The full ULA has 16 sensors, whereas ULA has 8 sensors, both have inter-element spacing d. The sensor positions and the total array length (i.e., aperture) for the four cases are given in Table II. We use $d = \lambda/2$.

The corresponding Gram matrices $|\mathbf{A}^H\mathbf{A}|$ for each of the arrays is shown in Fig. 2. The angle space $[-90, 90]$ is discretized using a grid of size $M = 360$ giving a resolution of $\Delta\theta = 0.5°$. We note that due to the specific sensor arrangement, co-prime and nested arrays have larger aperture than ULA, see Table II. Though co-prime and nested arrays have much larger apertures than ULA, they never cause aliasing (since $d = \lambda/2$, the aperture of ULA cannot be increased without causing aliasing).

B. Resolving more sources than sensors

The SBL algorithm in Sec. III is applied for DOA estimation. We simulate observations assuming multiple sources with equal source power of 1, equally spaced between $-60°$ and $60°$. The sources are always on a grid point.

The mean of $\hat{\gamma}$ estimate from 1000 Monte Carlo (MC) runs of SBL is shown in Fig. 3 for the three arrays and three different sparsity values, i.e., $K = 4$, 8, and 10 sources. Data is processed using $L = 100$ snapshots with an SNR of 20 dB. When non-zero, the true value of γ is 1. The co-prime and nested arrays can resolve all the DOAs even when more sources are present than the number of sensors ($K = 10$, $N = 7$ or 8), while the ULA solution is not sparse. The mean $\hat{\gamma}$ could be less than 1 as the power spreads out in neighboring grid points due to high coherence of steering vectors. The last column shows the normalized co-array MUSIC spectrum discussed in Secs. II E and II F applied to the co-prime array in Table II. As can be seen, co-array MUSIC can also identify more sources than the number of sensors. The interpolated co-array has $|\mathbb{U}^+| = 12$ and hence co-array MUSIC cannot identify more than 11 sources.

C. DOA estimation

The performance of SBL algorithm processing measurements from different array types can be quantified using the RMSE of the DOA estimates. The DOAs are estimated as the location of the strongest K peaks in γ. The DOA RMSE is calculated as

TABLE II. The three arrays used in simulations along with the location of the sensors and the array aperture.

Array	Sensor positions	Length
Full ULA	$\{1, 2, \ldots, 16\}d$	$15\,d$
ULA	$\{1, 2, \ldots, 8\}d$	$7d$
Co-prime	$\{1, 3, 5, 6, 7, 9, 11, 16\}d$	$15d$
Nested	$\{1, 2, 3, 4, 5, 10, 15\}$d	$14d$

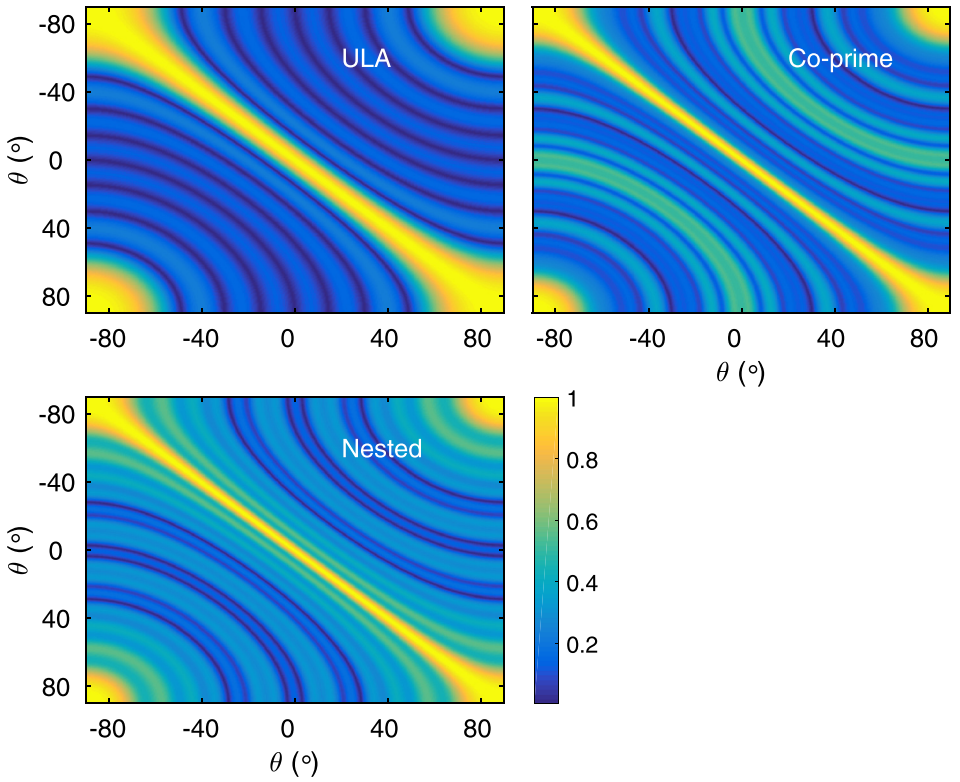

FIG. 2. (Color online) Gram matrices $|\mathbf{A}^H\mathbf{A}|$ for the three array configurations: ULA, Co-prime, and Nested. The arrays have 8, 8, and 7 sensors, respectively, and grid size $M = 360$.

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{sim}}}\frac{1}{K}\sum_{i=1}^{N_{\text{sim}}}\sum_{k=1}^{K}\left(\theta_{i,k} - \hat{\theta}_{i,k}\right)^2}, \tag{26}$$

where θ_i and $\hat{\theta}_i$ are the true and the estimated DOAs for the ith MC simulation run, N_{sim} is the number of MC runs, and K DOAs are present.

The DOA RMSE versus number of snapshots is in Fig. 4 for $K = 4$, 8, and 12 sources. SNR $= 0$ dB. For $K = 4$, ULA has the lowest RMSE. For $K = 8$ and $K = 12$, both co-prime and nested arrays can localize all the sources, whereas the ULA is unable to localize all the sources (see Fig. 3). As a result, the co-prime and nested arrays have lower error and the ULA has high error (not visible in the current plot).

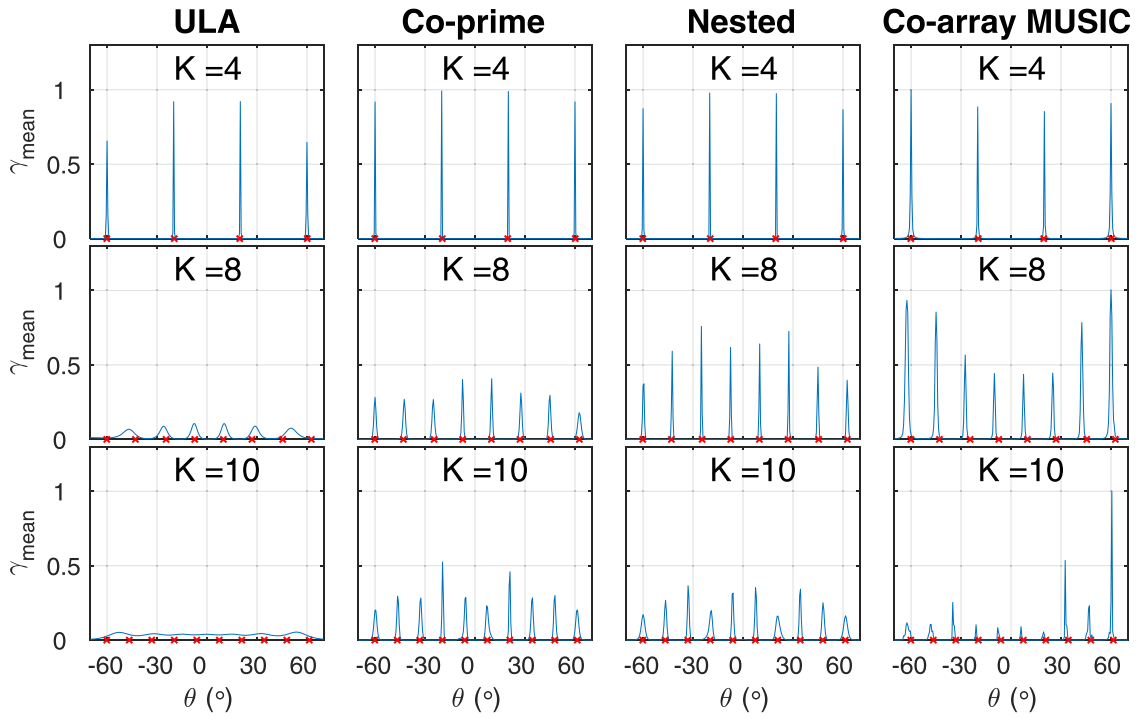

FIG. 3. (Color online) Mean γ over 1000 MC runs for array configurations of ULA, co-prime, and nested. Number of snapshots $L = 100$ and SNR is 20 dB. The red marks indicate true source locations.

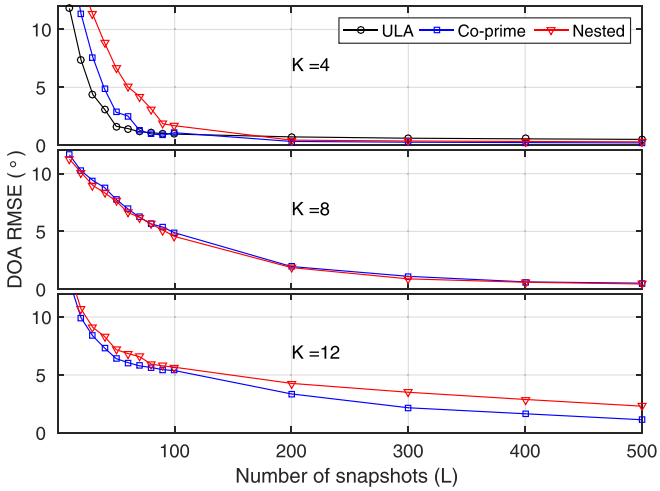

FIG. 4. (Color online) DOA RMSE versus L for ULA, co-prime, and nested arrays. SNR is set to 20 dB.

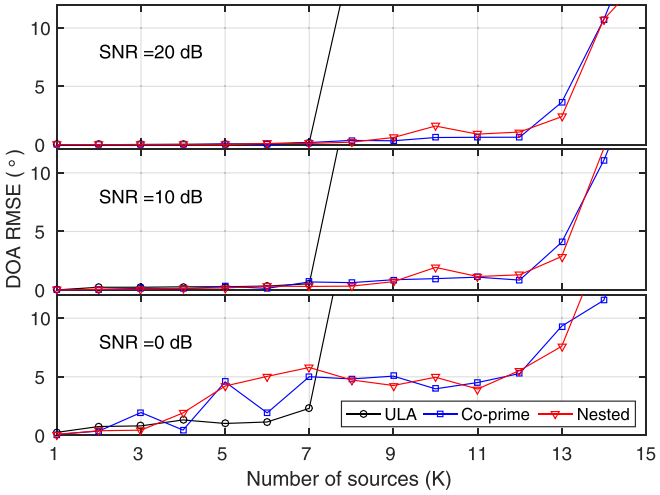

FIG. 6. (Color online) DOA RMSE versus K for ULA, co-prime, and nested arrays. $L = 100$ snapshots.

The RMSE versus SNR plot is shown in Fig. 5. The full ULA has double the number of sensors than any of the other arrays and hence has the lowest RMSE among all of them. The difference in the plots of full ULA and other arrays indicates the loss in performance due to reduced number of sensors in the corresponding array. Among other arrays, for $K = 4$, ULA has the lowest RMSE. The lower RMSE of ULA for fewer sources ($K = 4$) in Figs. 4 and 5 could be due to higher redundancy in the covariance matrix. For $K = 8$ and $K = 12$, the ULA cannot resolve all the sources whereas the nested and co-prime array can resolve all the sources and have similar errors. We also compute RMSE as a function of the number of sources (K). For this simulation, the number of snapshots $L = 100$ is fixed and SNR is selected from 0, 10, and 20 dB as seen in Fig. 6. The ULA can resolve up to 7 sources, whereas both nested and co-prime arrays can resolve even more.

D. Multi-frequency processing

When processing multiple frequencies, SBL (see Sec. III B) can avoid spatial aliasing since the location of aliased peaks changes with frequency. A demonstration of this is shown in Fig. 7. The top panel displays estimated signal power per frequency as a function of arrival angles estimated using CBF, SBL MF1, and SBL MF2. The bottom panel shows the average of the signal power across all the frequencies. Observations are simulated using a ULA of $N = 16$ sensors with spacing corresponding to a cutoff frequency of 750 Hz. $L = 100$ snapshots are generated at 10 dB SNR for each frequency in the set {500, 510,…, 1500} Hz. True sources with unit signal power are located at $-60°$, $-20°$, $20°$, and $60°$. For both CBF and SBL MF2, aliasing starts to appear around 800 Hz. For CBF, aliasing results in broadening of the peaks in the average power spectrum, whereas for SBL MF2, it results in spurious peaks in addition to the peaks at true source location. SBL MF1 is able to avoid both these problems and shows sharp peaks at correct source locations.

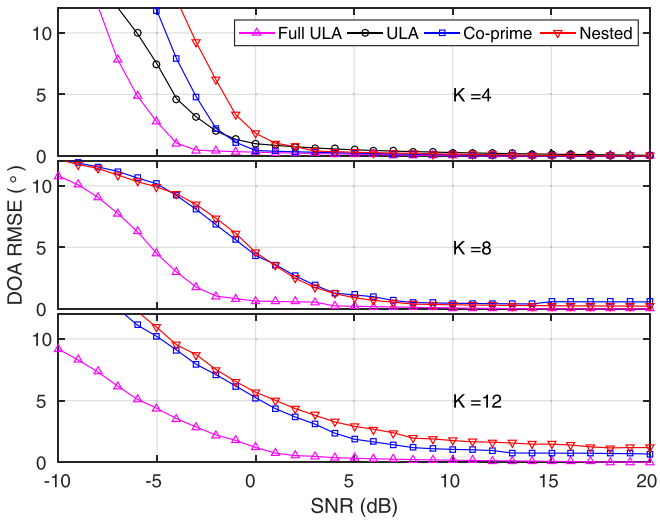

FIG. 5. (Color online) DOA RMSE versus SNR for ULA, co-prime, and nested arrays. $L = 100$ snapshots.

V. EXPERIMENT DATA

A. Data description

We use the NC09 experiment data set[33] for processing. The experiment was carried out on the Coronado Bank located south-west of Point Loma, San Diego, CA. Data was recorded on four equally configured vertical line arrays (VLAs) with VLA2 used for processing, see Fig. 8. The transect of the R/V New Horizon (source of opportunity) is easterly of the VLA bearing line and passes each VLA at port side within ∼100 m proximity. The 52 m long vessel transits at a constant speed of 2.2 m/s. Hydrophone 1 is 7 m above the seafloor and all 16 elements are used for processing. The element spacing of $d = 1$ m corresponds to a design frequency of 750 Hz at 1500 m/s sound speed.

The 30 min data is first down sampled from 25 000 Hz to a sampling frequency of 5000 Hz. Next, 4393 snapshots are computed continuously from the data with an FFT length of 4096 samples (0.8 s) and 50% overlap. Then, $L = 6$ snapshots are designated to segments of ∼ 2.9 s duration, resulting in 732 partially overlapping segments for the entire data set. The data in the segments are designated for further

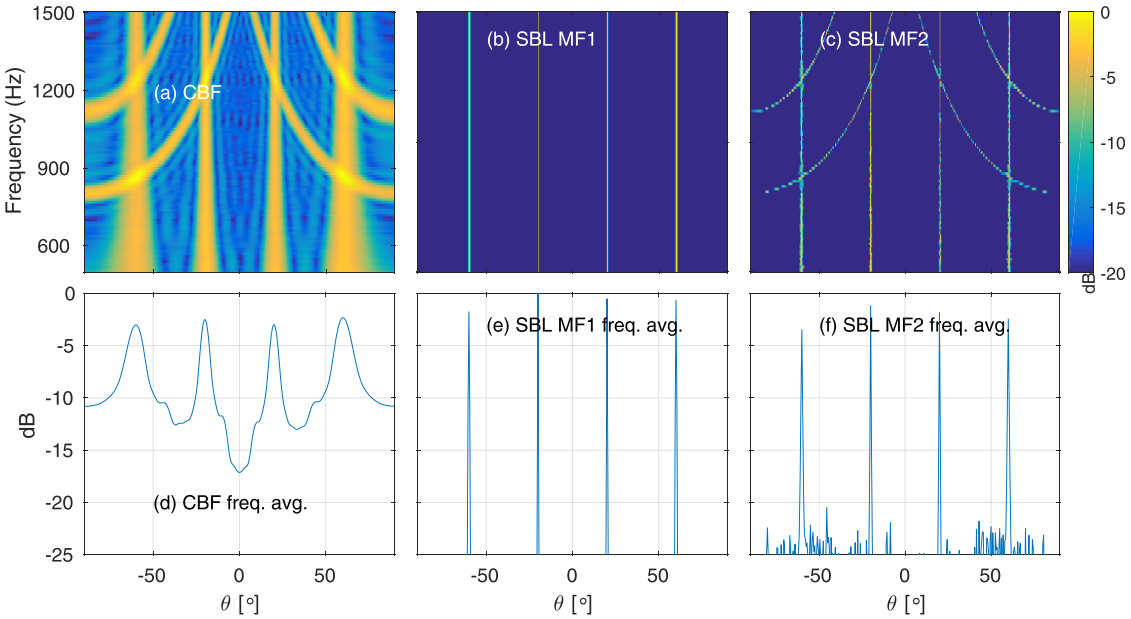

FIG. 7. (Color online) Multi-frequency processing: The signal power with changing frequency and DOA when four sources are present obtained using (a) CBF, (b) SBL MF1, and (c) SBL MF2. The bottom panel shows the plot of the average energy across the frequencies. SNR is 10 dB and $L = 100$ snapshots are used per frequency.

processing (e.g., $L = 6$ snapshots are used to construct a covariance). For conventional processing, each snapshot is windowed with a Kaiser window, where $\beta = 4.7$. The mean soundspeed over the array is 1492 m/s obtained from a nearby conductivity, temperature, and depth (CTD) cast. The mixed surface layer (1504–1506 m/s) extends to a depth of 22 m, followed by a downward refracting profile.

B. Data processing

We process data using the full 16 element VLA as well as from its subarrays shown in Table III. Specifically, we construct a co-prime array with $N_1^c = 5, N_2^c = 2$ resulting in eight sensor subarray, and a nested array with $N_1^n = 4, N_2^n = 3$ giving a seven sensor subarray. Additionally, we construct two ULAs from this array, ULA1 consists of the first 8 out of 16 elements of the original array, whereas ULA2 consists of every other sensor. Thus, both ULAs have eight sensors, but ULA2 has twice the aperture relative to ULA1. The aperture of co-prime and nested arrays is comparable to that of ULA2. The specific choice of co-prime and nested arrays is made to maximize their apertures using the given 16 element VLA. ULA1 was chosen to study the effect of a dense ULA but with limited aperture. The undersampled but uniform spaced ULA2 was chosen to demonstrate ability of multi-frequency SBL (SBL MF1) to combat spatial aliasing. All the arrays have approximately the same number of sensors and hence the same array gain.

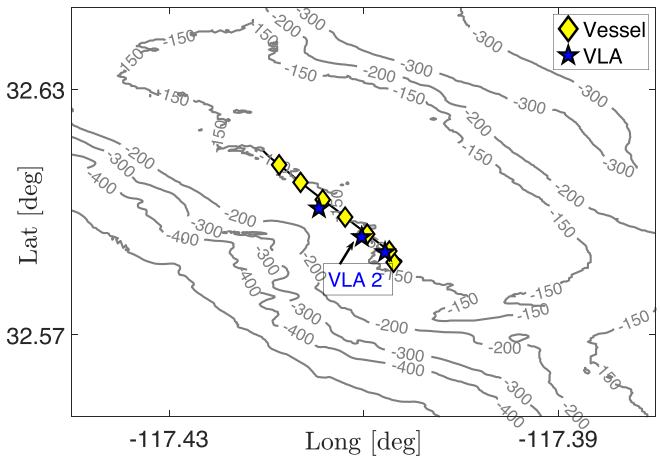

FIG. 8. (Color online) NC09 experiment map showing location of VLA2 (star) and trajectory (diamond) of the vessel over 30 min. The ship follows a path of approximately constant depth as indicated by the 150 m contour.

We use multi-frequency SBL discussed in Sec. III B to process the experiment data. The noise estimate [Eq. (19)] is used with $K = 4$ because there are always at least four relatively strong DOAs as seen from the Bellhop[43] model in Fig. 9(b). Though the number of DOAs are changing over time, we use a fixed value of K for uniformity of processing across the whole 30 min data.

The outputs of CBF, SBL, and MUSIC for the full VLA data are shown in Figs. 9(c), 9(d), and 9(e). For CBF and MUSIC, their spectrum are averaged over 650–750 Hz. For multi-frequency SBL 80–1000 Hz is used. A reduced bandwidth is used for CBF and MUSIC because for lower frequencies, the resolution is low and higher frequencies suffer from aliasing. Increasing bandwidth for CBF and MUSIC will blur the image in Fig. 9(c). SBL does

TABLE III. The four sub-arrays used for processing NC09 experiment data. The location of the sensors and the array apertures are indicated.

Array	Sensor positions	Length
ULA1	$\{1, 2, 3, 4, 5, 6, 7, 8\}d$	$7d$
ULA2	$\{1, 3, 5, 7, 9, 11, 13, 15\}d$	$14d$
Co-prime	$\{1, 3, 5, 6, 7, 9, 11, 16\}d$	$15d$
Nested	$\{1, 2, 3, 4, 5, 10, 15\}d$	$14d$

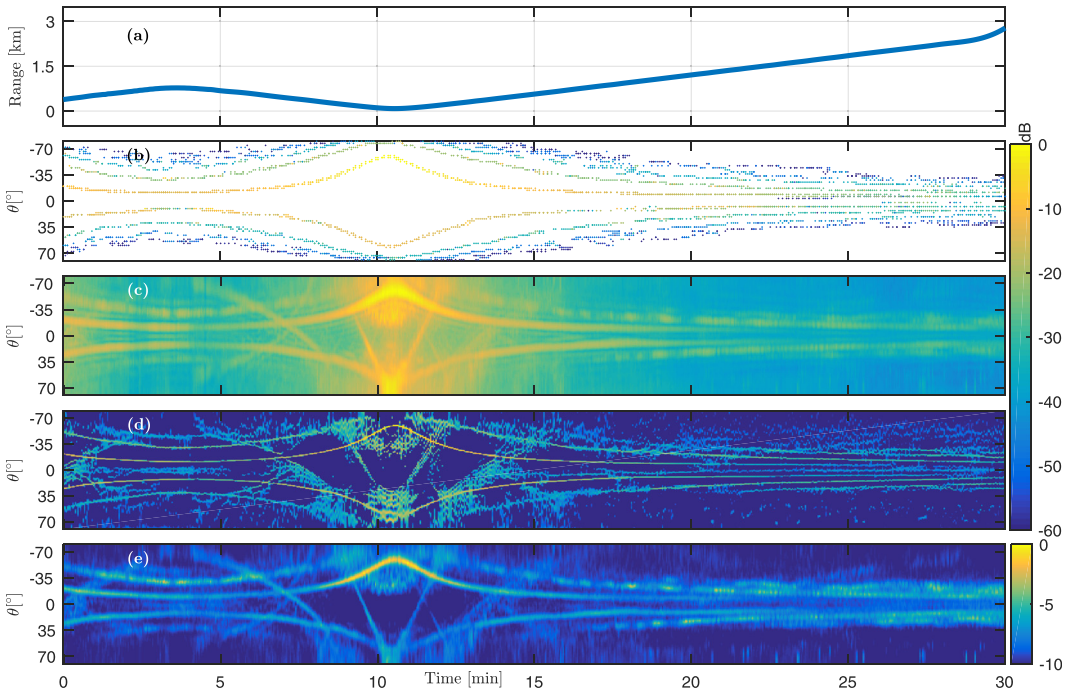

FIG. 9. (Color online) NC09: Evolution of DOAs with time as the ship moves. (a) Distance of ship from VLA. (b) Simulated DOAs using Bellhop. (c) CBF, (d) SBL, (e) MUSIC. For CBF and MUSIC, the spectrum are averaged over 650–750 Hz, while for SBL 80–1000 Hz is used.

not suffer from these issues as it simultaneously processes all frequencies giving high resolution and significantly suppresses aliasing.[12] The DOA locations obtained using SBL closely matches the Bellhop model output in Fig. 9(b).

The latter half of the event (from 15 to 30 min, see also Fig. 4 in Ref. 33) has more multipaths (DOAs) as the ship moves away from the array. We process data from the full array and the four sub-arrays (see Table III) with the estimated angular source power versus time in Fig. 10. All the sub-arrays will suffer from reduced array gain compared to the full array as the number of sensors is approximately halved. Only SBL results are shown as the resolution of CBF and MUSIC are poor due to incoherent spectrum averaging.

The sub-arrays in general are expected to have fewer DOA detections relative to the full array because of the reduced array gain. The co-prime, nested, and ULA2 sub-arrays are able to resolve the DOAs as well as the full ULA,

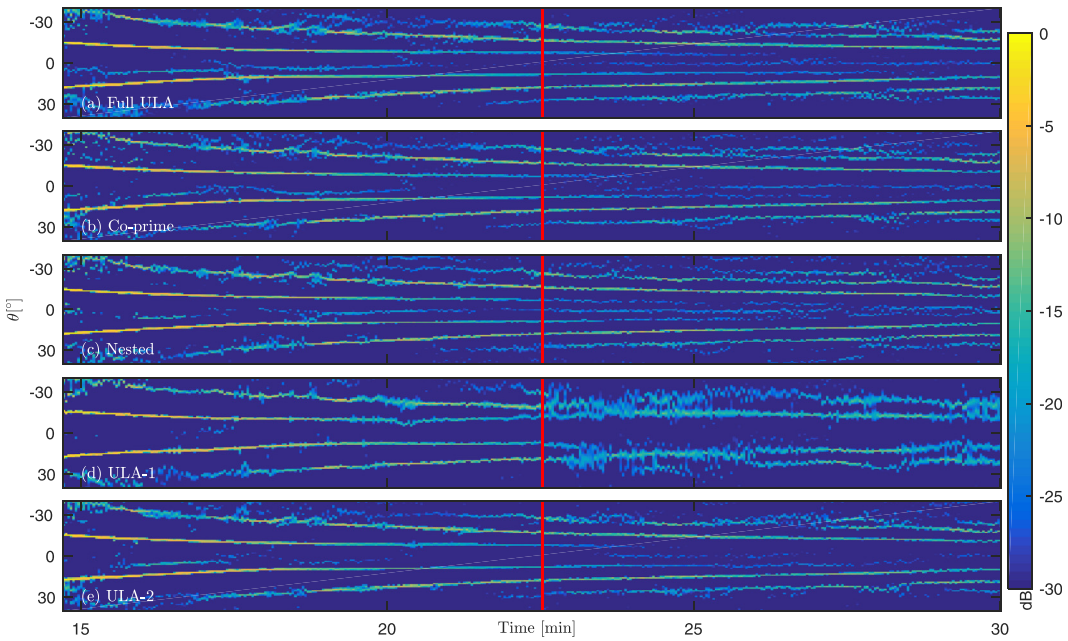

FIG. 10. (Color online) NC09: Evolution of DOAs with time (15–30 min portion from Fig. 9). The angular source powers are estimated using multi-frequency SBL (SBL MF1) for (a) full ULA, (b) co-prime, (c) nested, (d) ULA1, and (e) ULA2. The beampatterns corresponding to the vertical line around 22 min are shown in Fig. 11.

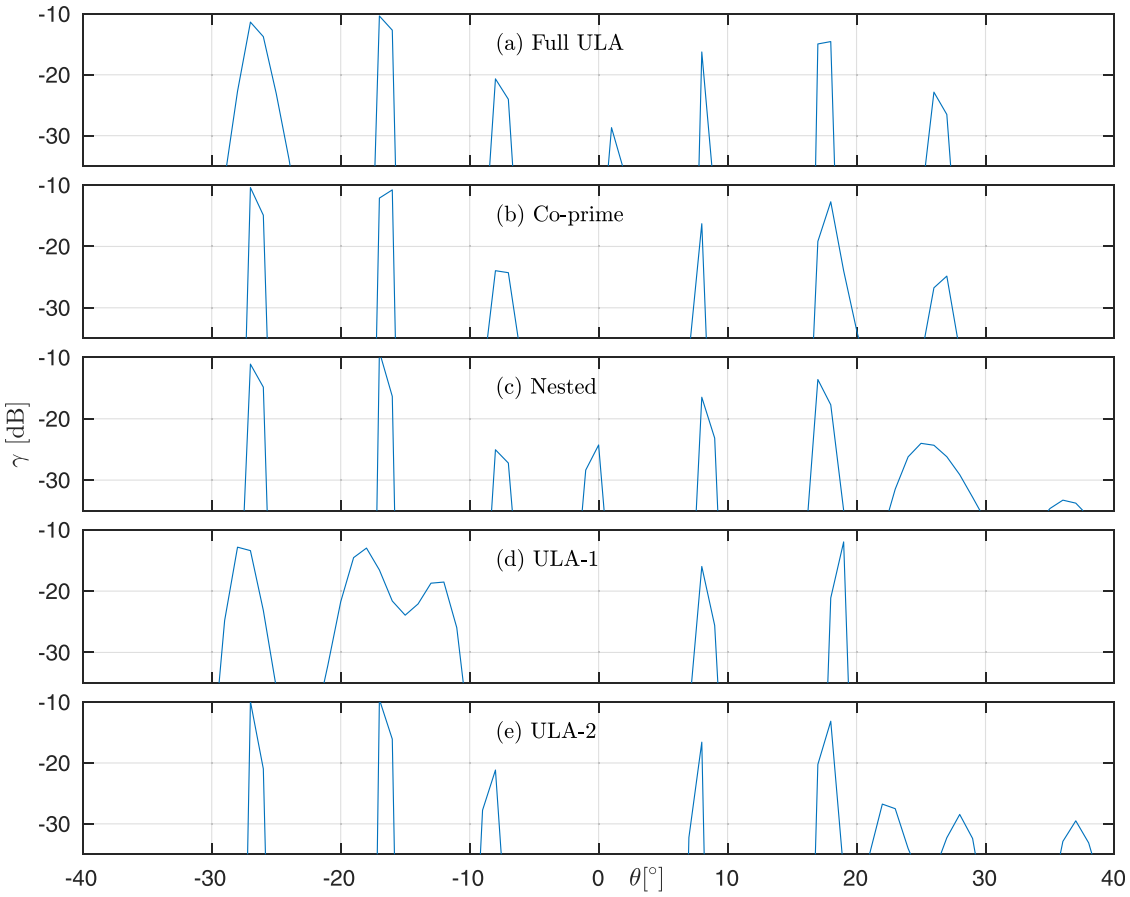

FIG. 11. (Color online) γ versus θ for the time step indicated by the vertical line around 22 min in Fig. 10.

see Fig. 10. This is highlighted around time 23 mins, indicated by the red vertical bar with the corresponding estimated γ versus angle in Fig. 11. The full ULA can identify up to 7 DOAs, whereas the co-prime, and ULA2 sub-arrays identify slightly fewer DOAs. The performance of ULA1 suffers significantly when there are more than four DOAs. This is because ULA1 has the smallest aperture of them all and its ULA geometry limits the number of sources it can identify.

In comparison, the ULA2 has twice the aperture of ULA1. Though ULA2 will suffer from aliasing for some of the higher frequencies due to the multi-frequency SBL processing, these get suppressed and only the sparse support common across all the frequencies is retained. The aliasing suppression property of multi-frequency SBL has been discussed in the literature.[12,42] Thus, multi-frequency SBL can combat spatial aliasing and provide performance similar to that of sparse arrays such as co-prime and nested arrays. However, it is important to note that the sparse ULA2 cannot identify more sources than sensors, unlike co-prime and nested arrays.

VI. CONCLUSIONS

Using beamforming simulations, we showed that SBL and co-array MUSIC are able to resolve more sources than the number of sensors when processing co-prime and nested array data. The RMSE analysis of DOA estimates showed that co-prime and nested arrays outperform an ULA with the same number of sensors, especially when the number of sources is close to the number of sensors or higher. It was seen that by making use of multi-frequency observations, SBL can significantly reduce spatial aliasing.

The NC09 experiment data was used to qualitatively compare the sparse arrays for DOA estimation using 16 element ULA data. The data was processed using multi-frequency SBL over approximately 1 kHz bandwidth for various sparse subarrays. It was observed that the sparse arrays can accurately identify DOAs while avoiding spatial aliasing.

[1]S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM Rev. **43**(1), 129–159 (2001).
[2]S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," IEEE Trans. Signal Process. **41**(12), 3397–3415 (1993).
[3]M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," J. Mach. Learn. Res. **1**, 211–244 (2001).
[4]D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005).
[5]D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," IEEE Trans. Signal Process. **55**(7), 3704–3716 (2007).
[6]K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," J. Acoust. Soc. Am. **141**(1), 92–103 (2017).
[7]P. Gerstoft, C. F. Mecklenbräuker, W. Seong, and M. Bianco, "Introduction to compressive sensing in acoustics," J. Acoust. Soc. Am. **143**(6), 3731–3736 (2018).
[8]Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning," IEEE Trans. Wireless Commun. **11**(10), 1–11 (2012).
[9]A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," J. Acoust. Soc. Am. **136**(1), 260–271 (2014).

[10]P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker, "Multiple and single snapshot compressive beamforming," J. Acoust. Soc. Am. **138**(4), 2003–2014 (2015).
[11]P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru, "Multi snapshot sparse Bayesian learning for DOA," IEEE Signal Process. Lett. **23**(10), 1469–1473 (2016).
[12]S. Nannuru, K. L. Gemba, P. Gerstoft, W. S. Hodgkiss, and C. F. Mecklenbräuker, "Sparse Bayesian learning with uncertainty models and multiple dictionaries," arXiv:1704.00436 (2017).
[13]P. Pal and P. P. Vaidyanathan, "Parameter identifiability in sparse Bayesian learning," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy (May 4–9, 2014), pp. 1851–1855.
[14]A. Koochakzadeh and P. Pal, "Cramér-Rao bounds for underdetermined source localization," IEEE Signal Process. Lett. **23**(7), 919–923 (2016).
[15]A. Koochakzadeh and P. Pal, "On saturation of the Cramér-Rao bound for sparse Bayesian learning," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, LA (March 5–9, 2017).
[16]P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," IEEE Trans. Signal Process. **58**(8), 4167–4181 (2010).
[17]P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," IEEE Trans. Signal Process. **59**(2), 573–586 (2011).
[18]Z. Weng and P. M. Djurić, "A search-free DOA estimation algorithm for coprime arrays," Digital Signal Process. **24**, 27–33 (2014).
[19]K. Adhikari, J. R. Buck, and K. E. Wage, "Extending coprime sensor arrays to achieve the peak side lobe height of a full uniform linear array," EURASIP J. Adv. Signal Process. **2014**(148) (2014).
[20]S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," IEEE Trans. Signal Process. **63**(6), 1377–1390 (2015).
[21]D. Bush and N. Xiang, "Broadband implementation of coprime linear microphone arrays for direction of arrival estimation," J. Acoust. Soc. Am. **138**(1), 447–456 (2015).
[22]N. Xiang, D. Bush, and J. E. Summers, "Experimental validation of a coprime linear microphone array for high-resolution direction-of-arrival measurements," J. Acoust. Soc. Am. **137**(4), EL261–EL266 (2015).
[23]D. Bush and N. Xiang, "n-tuple coprime sensor arrays," J. Acoust. Soc. Am. **142**(6), EL567–EL572 (2017).
[24]K. Adhikari and J. R. Buck, "Spatial spectral estimation with product processing of a pair of colinear arrays," IEEE Trans. Signal Process. **65**(9), 2389–2401 (2017).
[25]G. D. Martino and A. Iodice, "Passive beamforming with coprime arrays," IET Radar, Sonar, Navigation **11**(6), 964–971 (2017).
[26]P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proceedings of the IEEE Digital Signal Processing Workshop*, Sedona, AZ (January 4–7, 2011).
[27]P. P. Vaidyanathan and P. Pal, "Why does direct-MUSIC on sparse-arrays work?," in *Proceedings of the Asilomar Conference of Signal, Systems, Computers*, Pacific Grove, CA (November 3–6, 2013).
[28]C. L. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," IEEE Signal Process. Lett. **22**(9), 1438–1442 (2015).
[29]M. Wang, Z. Zhang, and A. Nehorai, "Performance analysis of coarray-based MUSIC in the presence of sensor location errors," IEEE Trans. Signal Process. **66**(12), 3074–3085 (2018).
[30]Y. Qin, Y. Liu, J. Liu, and Z. Yu, "Underdetermined wideband DOA estimation for off-grid sources with coprime array using sparse Bayesian learning," Sensors **18**(1), 253 (2018).
[31]M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér-Rao bound," IEEE Trans. Signal Process. **65**(4), 933–946 (2017).
[32]C.-L. Liu and P. P. Vaidyanathan, "Cramér–Rao bounds for coprime and other sparse arrays, which find more sources than sensors," Digital Signal Process. **61**, 43–61 (2017).
[33]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. Am. **144**(3), 1231–1244 (2018).
[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. Am. **141**(5), 3411–3420 (2017).
[35]Y. Liu and J. R. Buck, "High-resolution direction-of-arrival estimation in SNR and snapshot challenged scenarios using multi-frequency coprime arrays," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, LA, (March 5–9, 2017).
[36]S. Nannuru, P. Gerstoft, A. Koochakzadeh, and P. Pal, "Sparse Bayesian learning for DOA estimation using co-prime and nested arrays," in *Proceedings of the IEEE Sensor Array Multichannel Signal Processing Workshop*, Sheffield, UK, (July 8–11, 2018).
[37]C. L. Liu, P. P. Vaidyanathan, and P. Pal, "Coprime coarray interpolation for DOA estimation via nuclear norm minimization," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada (May 22–25, 2016).
[38]R. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antenn. Prop. **34**(3), 276–280 (1986).
[39]P. Pal and P. P. Vaidyanathan, "Pushing the limits of sparse support recovery using correlation information," IEEE Trans. Signal Process. **63**(3), 711–726 (2015).
[40]P. Stoica and A. Nehorai, "On the concentrated stochastic likelihood function in array signal processing," Circ. Syst. Signal Process. **14**(5), 669–674 (1995).
[41]C. D. Austin, R. L. Moses, J. N. Ash, and E. Ertin, "On the relation between sparse reconstruction and parameter estimation with model order selection," IEEE J. Sel. Topics Signal Process. **4**(3), 560–570 (2010).
[42]S. Nannuru, K. L. Gemba, and P. Gerstoft, "Sparse Bayesian learning with multiple dictionaries," in *Proceedings of the IEEE Global Conference on Signal and Information Processing*, Montreal, Canada (November 14–16, 2017).
[43]M. B. Porter and H. P. Bucker, "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am. **82**(4), 1349–1359 (1987).