

Sequential sparse Bayesian learning for time-varying direction of arrival

Yongsung Park,^{a)} Florian Meyer,^{b)} and Peter Gerstoft^{c)}

Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093-0238, USA

ABSTRACT:

This paper presents methods for the estimation of the time-varying directions of arrival (DOAs) of signals emitted by moving sources. Following the sparse Bayesian learning (SBL) framework, prior information of unknown source amplitudes is modeled as a multi-variate Gaussian distribution with zero-mean and time-varying variance parameters. For sequential estimation of the unknown variance, we present two sequential SBL-based methods that propagate statistical information across time to improve DOA estimation performance. The first method heuristically calculates the parameters of an inverse-gamma hyperprior based on the source signal estimate from the previous time step. In addition, a second sequential SBL method is proposed, which performs a prediction step to calculate the prior distribution of the current variance parameter from the variance parameter estimated at the previous time step. The SBL-based sequential processing provides high-resolution DOA tracking capabilities. Performance improvements are demonstrated by using simulated data as well as real data from the SWellEx-96 experiment.

© 2021 Acoustical Society of America. <https://doi.org/10.1121/10.0003802>

(Received 28 September 2020; revised 26 January 2021; accepted 26 February 2021; published online 29 March 2021)

[Editor: James F. Lynch]

Pages: 2089–2099

I. INTRODUCTION

Direction of arrival (DOA) estimation or direction finding aims to localize multiple sources from measurements of the wave-field provided by an array of sensors.^{1,2} It is a key signal processing task in radar, sonar, medical analytics, and seismic systems. Compressive beamforming, i.e., compressive sensing (CS)³-based DOA estimation, is a promising approach for DOA estimation that can provide source directions at a high resolution by promoting sparse solutions.^{1,2,4,5} While CS is widely used, sequential variants have received little attention. In the following, sparse Bayesian sequential estimation is derived based on sparse Bayesian learning (SBL)^{6–9} and used for DOA estimation.

The idea of CS or sparse recovery³ is to represent signals that are sparse in different domains, e.g., time domain,^{10–12} frequency domain,^{13,14} spatial domain,^{15,16} wavenumber domain,^{17–19} angular domain,^{1,2,4,5,7,8,20,21} and recover the signals by using sparse recovery techniques.^{6,22–28} This corresponds to obtaining sparse solutions to underdetermined systems of linear equations. Since the sparse solution needs a combinatorial search, which is impractical, CS is typically based on suboptimum algorithms that perform convex relaxation (e.g., l_1 minimization), greedy search, iterative thresholding, or SBL.²⁹

Often, a sequence of measurements is collected, and each measurement provides information on a sparse time-

varying signal. For more robust estimation, a sequential reconstruction method that uses information from previous steps is used.^{30–32} CS-based sequential reconstruction of sparse signals has been studied, focusing on image recovery^{33–35} and channel impulse response estimation.^{36–38} Many sequential reconstruction methods with CS have variant forms based on l_1 minimization.^{27,33,39,40} The methods have integrated CS with dynamics information exploiting support (the non-zero entries in the sparse signal vector) and amplitude change in sparse signals.³⁴ Adaptations of the sequential Bayesian estimation methods^{41,42} to CS have been introduced.^{27,35,38–40,43,44}

Sequential sparse signal reconstruction is achieved by incorporating statistical information related to the probability distribution of the time-varying sparse signal from the previous time step in the next time step. The l_1 minimization has the Bayesian framework interpretation.^{45,46} The solution is equivalent to the maximum *a posteriori* estimate, where the signal is modeled as a random vector with a Laplacian-like prior distribution and the measurement noise is assumed Gaussian.^{40,46} The hyperparameter related to the Laplace distribution can be used for sequential processing by propagating statistical information across time steps (iterative reweighted l_1 minimization was also proposed).^{27,40}

Similar to the l_1 minimization-based sequential reconstruction of sparse signals, statistical information in SBL can also be used for sequential sparse signal reconstruction.^{38,44} In SBL, the signal is modeled as a random vector with a Gaussian prior distribution and the Gaussian measurement noise. Sparse signal reconstruction as well as estimation of hyperparameters such as the measurement noise variance and the signal amplitude variance are performed

^{a)}Electronic mail: yongsungpark@ucsd.edu, ORCID: 0000-0001-6692-0262.

^{b)}Also at: Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093-0238, USA, ORCID: 0000-0001-6985-2250.

^{c)}ORCID: 0000-0002-0471-062X.

jointly by maximizing the evidence of the Bayesian estimation problem.^{6,7,24} In particular, in Ref. 44, an SBL-based method for sequential sparse signal reconstruction is used in the form of a hierarchical Bayesian model with an inverse-gamma hyperprior^{47,48} for the hyperparameter in SBL, and it utilizes the shape and scale parameters in the inverse-gamma hyperprior as statistical information across time.

In compressive beamforming, the number of potential source directions is typically much larger than the number of actual sources, and the problem of DOA estimation can be represented as a sparse signal reconstruction problem. We consider the case where source DOAs are time-varying and a time sequence of measurements is recorded. The resulting sequential sparse DOA estimation problem is solved in the SBL framework.

We propose a sequential sparse Bayesian learning method (S-SBL), which uses a conventional sequential state-space model to incorporate statistical information across time. S-SBL calculates a predicted variance parameter based on the state-transition model from the estimated variance of the previous time step. An updated variance estimate is obtained by incorporating the current measurement based on SBL.

We also adapt the sequential SBL approach⁴⁴ with the inverse-gamma hyperprior (SIG-SBL) to time-varying DOA estimation. SIG-SBL incorporates the estimated sparse source signal of the previous time step in the form of the shape and scale parameters of the inverse-gamma hyperprior to obtain the source signal amplitude variance in the next time step.

S-SBL uses a fixed-point update,^{7–9} and SIG-SBL uses an expectation maximization (EM) update. This fixed-point update can achieve more reliable estimation results^{15,18,49,50} and faster convergence compared to an update based on EM.^{7,9}

The two proposed sequential SBL methods are evaluated in realistic scenarios and compared to non-sequential DOA estimation. In addition, we apply them to real data from the SWellEx-96 (shallow water evaluation cell experiment 1996) experiment.^{51,52} Our evaluation suggests that S-SBL can provide more accurate estimation with faster convergence than SIG-SBL. On the other hand, SIG-SBL has more flexibility on how information from previous time steps is incorporated.

We use the following notation. We denote scalars by lowercase letters (e.g., x). Bold lowercase letters (e.g., \mathbf{x}) denote vectors, and bold uppercase letters (e.g., \mathbf{X}) denote matrices. \mathbf{X}^T and \mathbf{X}^H denote the transpose and Hermitian (conjugate transpose) of a matrix \mathbf{X} , respectively. $\hat{\mathbf{x}}$ denotes an estimate of \mathbf{x} . \mathbf{x}_t denotes \mathbf{x} at time t , and $\mathbf{x}_{1:t}$ denotes the ordered sequence $(\mathbf{x}_1 \dots \mathbf{x}_t)$.

II. SYSTEM FRAMEWORK

At time $t = 1, 2, \dots$, we consider K_t narrowband sources with complex signal amplitude $s_{t,k} \in \mathbb{C}$, $k = 1, \dots, K_t$. The sources are located at DOAs $\theta_{t,k} \in \Theta \triangleq [-90^\circ, 90^\circ]$,

$k = 1, \dots, K_t$ in the far-field of a linear array with M sensors. The observation $\mathbf{y}_t \in \mathbb{C}^M$ by the array at time t is modeled as

$$\mathbf{y}_t = \sum_{k=1}^{K_t} \mathbf{a}(\theta_{t,k}) s_{t,k} + \mathbf{v}_t, \quad (1)$$

where $\mathbf{v}_t \in \mathbb{C}^M$ is the measurement noise vector and $\mathbf{a}(\theta_{t,k}) \in \mathbb{C}^M$ is the steering vector. Measurement noise vectors \mathbf{v}_t have statistically independent, circular-symmetric, and zero-mean Gaussian entries with variance σ_t^2 and are assumed statistically independent across time t . The steering vector $\mathbf{a}(\theta_{t,k})$ is given by

$$\mathbf{a}(\theta_{t,k}) = [1 \ e^{-j(2\pi/\lambda)d_2 \sin \theta_{t,k}} \dots e^{-j(2\pi/\lambda)d_M \sin \theta_{t,k}}]^T, \quad (2)$$

where λ is the signal wavelength and d_m is the distance from sensor 1 to sensor m . $\mathbf{a}(\theta_{t,k})$ describes the phase delay of the signal from source k at the M sensors and time t . K_t as well as the $\theta_{t,k}$ and $s_{t,k}$, $k \in \{1, \dots, K_t\}$ in Eq. (1) are unknown. Due to the nonlinear transformation of $\theta_{t,k}$ in Eq. (2), the measurement model in Eq. (1) is nonlinear.

An alternative linear model can be obtained by introducing an angular search-grid of $N \gg K_t$ potential source DOAs $\bar{\theta} = [\bar{\theta}_1 \dots \bar{\theta}_N]^T \in \Theta^N$ and corresponding amplitude vector $\mathbf{x}_t \in \mathbb{C}^N$. Now the dictionary $\mathbf{A} = [\mathbf{a}(\bar{\theta}_1) \dots \mathbf{a}(\bar{\theta}_N)] \in \mathbb{C}^{M \times N}$ is preprocessed, and the alternative linear model is obtained as

$$\mathbf{y}_t = \mathbf{Ax}_t + \mathbf{v}_t. \quad (3)$$

Instead of estimating K_t as well as the $\theta_{t,k}$ and $s_{t,k}$, $k \in \{1, \dots, K_t\}$ in Eq. (1) at each time step t , we perform a sparse signal recovery by searching for a K_t -sparse vector \mathbf{x}_t in Eq. (3) together with the corresponding K_t columns of \mathbf{A} , which best express the measurement \mathbf{y}_t . We present the sparse signal recovery with CS and beamforming in the CS framework (compressive beamforming) (Sec. III) and cast SBL for compressive beamforming (Sec. IV).

By exploiting that the measurement noise vector \mathbf{v}_t has statistically independent, circular-symmetric, and zero-mean Gaussian entries with variance σ_t^2 , we directly obtain the probability density function associated with the likelihood at time t from Eq. (3),

$$p(\mathbf{y}_t | \mathbf{x}_t; \sigma_t^2) = \mathcal{CN}(\mathbf{y}_t; \mathbf{Ax}_t, \sigma_t^2 \mathbf{I}_M). \quad (4)$$

The time evolution of the dynamic source amplitudes is modeled by a linear state-transition model,

$$\mathbf{x}_t = \mathbf{Fx}_{t-1} + \mathbf{w}_t. \quad (5)$$

Here, $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the known matrix relating the state vector at time t to that at time $t - 1$, and $\mathbf{w}_t \sim \mathcal{CN}(\mathbf{w}_t; \mathbf{0}, \sigma_{w_t}^2 \mathbf{I}_N)$ is Gaussian process noise that is assumed statistically independent across time t .

For the special case where sources are stationary and \mathbf{x}_{t-1} and \mathbf{x}_t share a common sparsity profile, we have $\mathbf{F} = \mathbf{I}_N$

and $\sigma_{w_i}^2 = 0$. For the case where the DOAs of the sources are static but they can suddenly appear and disappear, i.e., sources can be “switched on/off,” we have $\mathbf{F} = \rho \mathbf{I}_N$ with $\rho \geq 0$ and $\sigma_{w_i}^2 > 0$. For dynamic sources, the locations of the non-zero entries of the source amplitude vector are time-dependent, and \mathbf{x}_t typically has non-zero “active” elements in the vicinity of the active elements in \mathbf{x}_{t-1} . Thus, a suitable choice for matrix \mathbf{F} is a banded symmetric Toeplitz matrix $\mathbf{F} = [F_{i,j}; i, j = 1, \dots, N]$, where

$$F_{i,j} = f_{|i-j|}, \quad (6)$$

with entries $f_l > 0$, $l \in \{0, \dots, L\}$, $f_l = 0$, $l \in \{L+1, \dots, N-1\}$. This matrix implements the convolution of \mathbf{x}_{t-1} with the L -lags blurring window $[f_L \ f_{L-1} \dots \ f_1 \ f_0 \ f_1 \dots \ f_{L-1} \ f_L]^T$.

The prior distribution of the source amplitudes \mathbf{x}_t is assumed multi-variate Gaussian with zero-mean and variance with DOA-dependent hyperparameter $\gamma_t \in \mathbb{R}^N$,

$$p(\mathbf{x}_t; \gamma_t) = \mathcal{CN}(\mathbf{x}_t; \mathbf{0}, \Gamma_t), \quad (7)$$

where $\Gamma_t = \text{diag}(\gamma_t) \in \mathbb{R}^{N \times N}$ and the entries of γ_t are assumed statistically independent across time t .

III. SPARSE SIGNAL RECOVERY

CS with a sparsity constraint gives improved beamforming results even in scenarios with a single time step (or single snapshot data). As CS-based compressive beamforming with a single time step is presented, we omit time t in Secs. III and IV for simplicity.

CS for beamforming solves DOA estimation as a sparse signal reconstruction problem. CS aims to reconstruct K -sparse source signal $\mathbf{x} \in \mathbb{C}^N$ from observation \mathbf{y} in Eq. (3), and the non-zero entries in \mathbf{x} represent active DOAs. As the number of look-directions N is much larger than the number of sources K ($N \gg K$), \mathbf{x} can be represented as a sparse signal. Mathematically, the l_0 -norm can be used as a sparsity measure, where l_0 -norm counts the number of non-zero entries in \mathbf{x} , i.e., $\|\mathbf{x}\|_0 = K$. The sparse signal reconstruction problem can be formulated as a constrained optimization. The objective is to find the solution \mathbf{x} that minimizes least squared error to the linear formulation [Eq. (3)] subject to the constraint that at most K elements of \mathbf{x} are non-zero,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2 \text{ subject to } \|\mathbf{x}\|_0 \leq K. \quad (8)$$

The minimization based on l_0 -norm is non-convex and needs a combinatorial search ($N C_K$), and suboptimum algorithms are proposed.²⁹

A convex relaxation technique based on l_1 minimization replaces the l_0 -norm constraint with the l_1 -norm constraint, which is a convex approximation to the l_0 -norm,

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \eta \|\mathbf{x}\|_1. \quad (9)$$

Here, η is a user-defined value that controls the balance between the data fit ($\|\mathbf{y} - \mathbf{Ax}\|_2^2$) and the sparsity level of \mathbf{x}

($\|\mathbf{x}\|_1$). This approach is known as least absolute shrinkage and selection operator (LASSO) or basis pursuit denoising (BPDN), which are closely related to each other using the method of Lagrange multipliers.²² The l_1 minimization, a convex relaxation technique, can be solved efficiently with convex optimization. The beamforming problem is solved in the CS framework in the form of Eq. (9),^{2,4} and CS-based beamforming methods have shown its superior performance.^{1,2,4,5,7,8,20,21}

The LASSO [Eq. (9)] has the Bayesian framework interpretation.^{4,46} The solution is equivalent to the maximum *a posteriori* estimate, where \mathbf{x} is modeled as a random vector with a Laplacian-like prior distribution and the measurement noise is assumed Gaussian. SBL, a method for solving sparse signal reconstruction problems, uses evidence maximization (also called type-II maximum likelihood), where \mathbf{x} is modeled as a random vector with a Gaussian prior distribution and the measurement noise is assumed Gaussian.^{6,24}

IV. SBL FOR COMPRESSIVE BEAMFORMING

The SBL framework uses the likelihood in Eq. (3) and the prior in Eq. (7). Sparse solution \mathbf{x} is desirable, and the sparsity of \mathbf{x} is achieved when the entry of $\gamma \in \mathbb{R}^N$ is zero, and the location of non-zero elements corresponds to the corresponding DOAs.

The SBL uses evidence maximization, and the evidence $p(\mathbf{y}; \gamma, \sigma^2)$ is calculated by the product of the likelihood [Eq. (3)] and the prior [Eq. (7)] integrated over the complex source amplitudes \mathbf{x} ,

$$\begin{aligned} p(\mathbf{y}; \gamma, \sigma^2) &= \int_{\mathbb{R}^{2N}} p(\mathbf{y}, \mathbf{x}; \gamma, \sigma^2) d\mathbf{x} \\ &= \int_{\mathbb{R}^{2N}} p(\mathbf{y}|\mathbf{x}; \sigma^2) p(\mathbf{x}; \gamma) d\mathbf{x} = \frac{e^{-\mathbf{y}^H \Sigma_y^{-1} \mathbf{y}}}{\pi^M \det \Sigma_y}, \end{aligned} \quad (10)$$

where $d\mathbf{x} = \prod_n^N \text{Re}(dx_n) \text{Im}(dx_n)$, and the measurement covariance is given as

$$\Sigma_y = \mathbf{A} \Gamma \mathbf{A}^H + \sigma^2 \mathbf{I}_M. \quad (11)$$

Maximum-likelihood hyperparameter estimation $\{\hat{\gamma}, \hat{\sigma}^2\}$ can be performed by maximizing the logarithm of the evidence [Eq. (10)],

$$\begin{aligned} (\hat{\gamma}, \hat{\sigma}^2) &= \arg \max_{\gamma \geq 0, \sigma^2 > 0} \log p(\mathbf{y}; \gamma, \sigma^2) \\ &= \arg \min_{\gamma \geq 0, \sigma^2 > 0} \log \det \Sigma_y + \mathbf{y}^H \Sigma_y^{-1} \mathbf{y}. \end{aligned} \quad (12)$$

The derivatives of Eq. (12) with respect to $\log(\gamma^{-1})$ are obtained as

$$\frac{\partial \mathcal{L}}{\partial \log(\gamma_n^{-1})} = -1 + \frac{\Sigma_{nn} + \mu_n^2}{\gamma_n}, \quad (13)$$

where

$$\boldsymbol{\mu} = \boldsymbol{\Gamma} \mathbf{A}^H \boldsymbol{\Sigma}_y^{-1} \mathbf{y} \in \mathbb{C}^N, \quad (14)$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma} \mathbf{A}^H \boldsymbol{\Sigma}_y^{-1} \boldsymbol{\Lambda} \boldsymbol{\Gamma} \in \mathbb{C}^{N \times N}. \quad (15)$$

By setting Eq. (13) to zero, we obtain the following γ and σ^2 updates for the EM-based SBL:^{6,24}

$$\gamma_n^{\text{new}} = \Sigma_{nn} + \mu_n^2, \quad (16)$$

$$(\sigma^2)^{\text{new}} = \frac{\|\mathbf{y} - \mathbf{A}\boldsymbol{\mu}\|_2^2 + \text{tr}[\mathbf{A}^H \mathbf{A} \boldsymbol{\Sigma}]}{M}. \quad (17)$$

Since the EM-based SBL is typically slow in practical applications, different update equations have been introduced to achieve faster convergence.^{6,7,9} For the γ update, a fixed-point update rule is proposed by taking the derivative of Eq. (13) with respect to γ and setting it to zero. The resulting γ update is given as^{7–9}

$$\gamma_n^{\text{new}} = \gamma_n^{\text{old}} \frac{\|\mathbf{y}^H \boldsymbol{\Sigma}_y^{-1} \mathbf{a}_n\|_2^2}{\mathbf{a}_n^H \boldsymbol{\Sigma}_y^{-1} \mathbf{a}_n}. \quad (18)$$

A noise variance σ^2 update according to Eq. (17) leads to seriously biased estimates. In beamforming applications, Eq. (17) results in $\sigma^2 = 0$.^{6,7,9,53} For this reason, we replace Eq. (17) with the following σ^2 update:^{7–9}

$$(\sigma^2)^{\text{new}} = \frac{\text{tr}[(\mathbf{I}_M - \mathbf{A}_M \mathbf{A}_M^+) \mathbf{y} \mathbf{y}^H]}{M - K}, \quad (19)$$

where \mathcal{M} is a given set of K active DOAs; \mathbf{A}_M is the corresponding active steering matrix; and \mathbf{A}_M^+ denotes the Moore–Penrose pseudo-inverse. The algorithm's convergence has not been proven so far, but it is suggested by various applications.^{7–9,15,18,49,54}

In DOA estimation, SBL can provide high-resolution DOA estimation performance for broadband sources⁹ as well as in both scenarios with single or multiple time steps.⁷ Existing SBL methods, which can process measurements collected at multiple time steps, rely on static DOAs.^{4,55} Thus, for scenarios with moving sources, most existing methods are infeasible.

V. SBL FOR TIME-VARYING DOA ESTIMATION

Time-varying DOA estimation aims to reconstruct source amplitude \mathbf{x}_t sequentially with time t by utilizing a time sequence of previous amplitudes $\mathbf{x}_{1:t-1}$ and measurements \mathbf{y}_t . We consider two SBL-based approaches for sequential sparse signal reconstruction (see Fig. 1).

A. Sequential SBL with inverse-gamma hyperpriors

The sequential SBL framework⁴⁴ uses an inverse-gamma hyperprior for the variance parameter γ_t ,

$$\begin{aligned} p(\gamma_t) &= \prod_{n=1}^N \mathcal{IG}(\gamma_{t,n}; \alpha_{t,n}, \beta_{t,n}) \\ &= C_t \prod_{n=1}^N (\gamma_{t,n})^{-\alpha_{t,n}-1} \exp(-\beta_{t,n}/\gamma_{t,n}), \end{aligned} \quad (20)$$

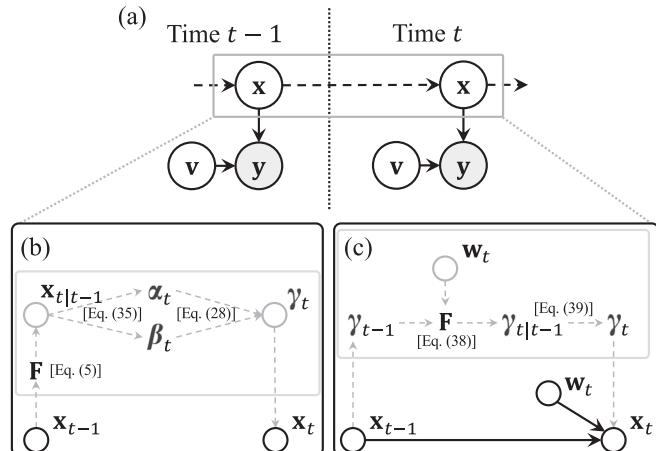


FIG. 1. Graphical model (Bayesian networks) (Ref. 56) representation of (a) sequential sparse signal reconstruction and sequential SBL approaches, (b) SIG-SBL, and (c) S-SBL. The node (circle) represents a random variable, e.g., complex source amplitudes \mathbf{x} , and the shaded node represents observations \mathbf{y} . Directed links (solid-line arrows) are added to the variables on which the distribution is conditioned. Hyperparameters and state-transition matrix \mathbf{F} are denoted without border. We have introduced gray boxes in (b) and (c), which do not follow the graphical model framework but show each procedure of the proposed approach explicitly. w_t in (c) are identical.

where C_t is a constant. Conventional non-sequential SBL methods^{6–9,18,21,24} use uniform hyperpriors. The model used by conventional SBL methods can be seen as a special case of the inverse-gamma hyperprior model that is obtained by setting shape and scale parameters to zero, i.e., $\alpha_t = \beta_t = \mathbf{0} \in \mathbb{C}^N$.

Let us consider the case of a general inverse-gamma hyperprior. The evidence $p(\mathbf{y}_t, \gamma_t; \sigma_t^2)$ can be calculated as

$$\begin{aligned} p(\mathbf{y}_t, \gamma_t; \sigma_t^2) &= \int_{\mathbb{R}^{2N}} p(\mathbf{y}_t, \mathbf{x}_t, \gamma_t; \sigma_t^2) d\mathbf{x}_t \\ &= p(\gamma_t) \int_{\mathbb{R}^{2N}} p(\mathbf{y}_t | \mathbf{x}_t; \sigma_t^2) p(\mathbf{x}_t | \gamma_t) d\mathbf{x}_t \\ &= p(\gamma_t) \frac{e^{-\mathbf{y}_t^H \boldsymbol{\Sigma}_y^{-1} \mathbf{y}_t}}{\pi^M \det \boldsymbol{\Sigma}_y}, \end{aligned} \quad (21)$$

where we introduced $p(\mathbf{x}_t | \gamma_t) = p(\mathbf{x}_t; \gamma_t)$, $d\mathbf{x}_t = \prod_{n=1}^N \text{Re}(dx_{t,n}) \text{Im}(dx_{t,n})$ and the measurement covariance

$$\boldsymbol{\Sigma}_{\mathbf{y}_t} = \mathbf{A} \boldsymbol{\Gamma}_t \mathbf{A}^H + \sigma_t^2 \mathbf{I}_M. \quad (22)$$

By using the inverse-gamma hyperprior [Eq. (20)] in Eq. (21), we can express the logarithmic evidence as

$$\begin{aligned} \log p(\mathbf{y}_t, \gamma_t; \sigma_t^2) &= \sum_{n=1}^N ((\alpha_{t,n} + 1) \log \gamma_{t,n}^{-1} - \beta_{t,n} / \gamma_{t,n}) \\ &\quad - \log \det \boldsymbol{\Sigma}_{\mathbf{y}_t} - \mathbf{y}_t^H \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-1} \mathbf{y}_t + C_t, \end{aligned} \quad (23)$$

where C_t is a normalization constant.

Maximum-likelihood hyperparameter estimation $\{\hat{\gamma}_t, \hat{\sigma}_t^2\}$ can now be performed by maximizing the logarithmic evidence [Eq. (23)],

$$(\hat{\gamma}_t, \hat{\sigma}_t^2) = \arg \max_{\gamma_t \geq 0, \sigma_t^2 > 0} \log p(\mathbf{y}_t, \gamma_t; \sigma_t^2). \quad (24)$$

By following the same procedure used for non-sequential SBL (Sec. IV), the γ and σ^2 updates for the EM-based SIG-SBL are given as

$$\gamma_{t,n}^{\text{new}} = \frac{\Sigma_{t,nn} + \mu_{t,n}^2 + \beta_{t,n}}{2 + \alpha_{t,n}}, \quad (25)$$

$$(\sigma_t^2)^{\text{new}} = \frac{\|\mathbf{y}_t - \mathbf{A}\boldsymbol{\mu}_t\|_2^2 + \text{tr}[\mathbf{A}^\text{H}\mathbf{A}\Sigma_t]}{M}, \quad (26)$$

where

$$\boldsymbol{\mu}_t = \boldsymbol{\Gamma}_t \mathbf{A}^\text{H} \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-1} \mathbf{y}_t \in \mathbb{C}^N, \quad (27)$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Gamma}_t - \boldsymbol{\Gamma}_t \mathbf{A}^\text{H} \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-1} \mathbf{A} \boldsymbol{\Gamma}_t \in \mathbb{C}^{N \times N}. \quad (28)$$

As discussed in Sec. IV, a noise variance σ^2 update according to Eq. (26) leads to seriously biased σ^2 estimates.^{6,7,9,53} For this reason, we replace Eq. (26) by the following σ^2 update:⁷

$$(\sigma_t^2)^{\text{new}} = \frac{\text{tr}[(\mathbf{I}_M - \mathbf{A}_\mathcal{M} \mathbf{A}_\mathcal{M}^+) \mathbf{y}_t \mathbf{y}_t^\text{H}]}{M - K_t}, \quad (29)$$

where \mathcal{M} is a set of K_t active DOAs; $\mathbf{A}_\mathcal{M}$ is the corresponding active steering matrix; and $\mathbf{A}_\mathcal{M}^+$ denotes the Moore–Penrose pseudo-inverse. The modified update equation [Eq. (29)] prevents the updated σ^2 from having a negative bias and vanishing and improves DOA estimation performance.^{7–9,50}

The parameters γ_t and σ_t^2 are estimated by iteratively computing Eqs. (27), (28), (25), and (29). While the algorithm's convergence has not been proven so far, it is suggested by our simulations. The convergence rate ϵ measures the relative change in γ_t ,⁷

$$\epsilon = \|\gamma_t^{\text{new}} - \gamma_t^{\text{old}}\|_1 / \|\gamma_t^{\text{old}}\|_1. \quad (30)$$

Iterative processing is stopped when $\epsilon \leq \epsilon_{\min}$. The sparse estimate of signal amplitudes $\hat{\mathbf{x}}_t$ is calculated from $\hat{\gamma}_t (= \gamma_t^{\text{new}})$ by using Eq. (27),

$$\hat{\mathbf{x}}_t = \hat{\boldsymbol{\Gamma}}_t \mathbf{A}^\text{H} \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-1} \mathbf{y}_t, \quad (31)$$

where $\hat{\boldsymbol{\Gamma}}_t = \text{diag}(\hat{\gamma}_t)$.

In contrast to the non-sequential SBL, in Eq. (25) now also the shape and scale parameters $\{\alpha_t, \beta_t\}$ of the inverse-gamma distribution are taken into account. Following Ref. 44, the parameters $\{\alpha_t, \beta_t\}$ are calculated from a predicted amplitude vector $\hat{\mathbf{x}}_{t|t-1} (= \mathbf{F}\hat{\mathbf{x}}_{t-1})$ as discussed next. To obtain the shape and scale parameters $\{\alpha_t, \beta_t\}$, the γ_t

minimizing the terms having $\{\alpha_t, \beta_t\}$ in Eq. (23), for the n th entry, is given as

$$-(\alpha_{t,n} + 1) + \frac{\beta_{t,n}}{\gamma_{t,n}} = 0 \rightarrow \gamma_{t,n} = \frac{\beta_{t,n}}{\alpha_{t,n} + 1}. \quad (32)$$

Assuming \mathbf{A} has orthogonal columns, the n th entry of γ_t , which minimizes the expected difference between the estimated $\hat{\mathbf{x}}_t$ and the predicted $\hat{\mathbf{x}}_{t|t-1}$ using the state-transition model, is given as

$$\gamma_{t,n} = \hat{x}_{t|t-1,n}^2. \quad (33)$$

As in Ref. 44, we set the shape term $\alpha_{t,n}$ equal to a constant tuning parameter $\xi - 1$ ($\xi > 1$) for all n , i.e., $\alpha_{t,n} = \xi - 1$, $n \in \{1, \dots, N\}$. ξ determines the weight of $\hat{\mathbf{x}}_{t|t-1}$ in the update step for $\hat{\mathbf{x}}_t$. Plugging Eq. (33) into Eq. (32) gives

$$\beta_{t,n} = \xi \hat{x}_{t|t-1,n}^2, \quad (34)$$

where $n \in \{1, \dots, N\}$.

The variance of the resulting inverse-gamma distribution decreases with increasing ξ . Thus, increasing the tuning parameter ξ results in a more informative prior distribution $p(\gamma_t)$. Note that SIG-SBL does not incorporate the driving noise variance $\sigma_{w_t}^2$ in Eq. (5).

SIG-SBL predicts $\hat{\mathbf{x}}_{t|t-1}$ based on the state-transition model from $\hat{\mathbf{x}}_{t-1}$ at time $t - 1$. Then the variance $\hat{\gamma}_t$ at time t is obtained in the SBL manner [Eq. (26)] by incorporating the current measurement \mathbf{y}_t and the statistical information $\{\alpha_t, \beta_t\}$ from $\hat{\mathbf{x}}_{t|t-1}$ [Eq. (34)]. The graphical model of SIG-SBL is shown in Fig. 1(b).

B. Sequential SBL with sequential state-space model

The proposed S-SBL approach is based on the complete sequential state-space model in Sec. II with fixed-point SBL update.^{15,18,49,50} S-SBL calculates predicted variances $\hat{\gamma}_{t|t-1}$ based on the state-transition model [Eq. (5)] from the variances γ_{t-1} estimated at time $t - 1$. Then the variance $\hat{\gamma}_t$ at time t is obtained by incorporating the current measurement vector \mathbf{y}_t in the SBL-based update. The graphical model of S-SBL is shown in Fig. 1(c). As demonstrated numerically in Sec. VI, it improves DOA estimation accuracy and computational effort compared to SIG-SBL.

Let us assume at time $t - 1$, we have calculated an estimate $\hat{\gamma}_{t-1}$ based on all previous measurement vectors $\mathbf{y}_{1:t-1}$. The corresponding approximation of the marginal distribution $p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1})$ is given by

$$\tilde{p}(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) = \mathcal{CN}(\mathbf{x}_{t-1}; \mathbf{0}, \hat{\boldsymbol{\Gamma}}_{t-1}), \quad (35)$$

where $\hat{\boldsymbol{\Gamma}}_{t-1} = \text{diag}(\hat{\gamma}_{t-1})$. In the prediction step, we calculate the covariance matrix $\hat{\boldsymbol{\Gamma}}_{t|t-1}$ of the predicted posterior

$$\tilde{p}(\mathbf{x}_t | \mathbf{y}_{1:t-1}) = \mathcal{CN}(\mathbf{x}_t; \mathbf{0}, \hat{\boldsymbol{\Gamma}}_{t|t-1}) \quad (36)$$

based on the state-transition model [Eq. (5)]. Since both \mathbf{x}_{t-1} and \mathbf{w}_t are complex Gaussians, we directly obtain

$$\hat{\Gamma}_{t|t-1} = \mathbf{F}\hat{\Gamma}_{t-1}\mathbf{F}^\top + \sigma_{w_t}^2 \mathbf{I}_N. \quad (37)$$

Next, the predicted variances $\hat{\gamma}_{t|t-1} = [\hat{\Gamma}_{t|t-1,11} \dots \hat{\Gamma}_{t|t-1,NN}]^\top$ are used as prior information in an SBL-based update. By plugging the likelihood [Eq. (4)] and the prior [Eq. (36) with Eq. (37)] into the SBL framework (Sec. IV), we obtain the fixed-point update for γ in the SBL, resulting in the iterative update⁷⁻⁹

$$\gamma_{t,n}^{\text{new}} = \gamma_{t,n}^{\text{old}} \frac{\|\mathbf{y}_t^\top \Sigma_{\mathbf{y}_t}^{-1} \mathbf{a}_n\|_2^2}{\mathbf{a}_n^\top \Sigma_{\mathbf{y}_t}^{-1} \mathbf{a}_n}, \quad (38)$$

where $\Sigma_{\mathbf{y}_t}$ is the data covariance, given as

$$\Sigma_{\mathbf{y}_t} = \mathbf{A}\Gamma_t^{\text{old}}\mathbf{A}^\top + \sigma_t^2 \mathbf{I}_M, \quad (39)$$

and

$$(\sigma_t^2)^{\text{new}} = \frac{\text{tr}[(\mathbf{I}_M - \mathbf{A}_M \mathbf{A}_M^\top) \mathbf{y}_t \mathbf{y}_t^\top]}{M - K_t}. \quad (40)$$

The resulting variance estimate $\hat{\gamma}_t$ represents an approximation of the marginal distribution $p(\mathbf{x}_t | \mathbf{y}_{1:t})$,

$$\tilde{p}(\mathbf{x}_t | \mathbf{y}_{1:t}) = \mathcal{CN}(\mathbf{x}_t; \mathbf{0}, \hat{\Gamma}_t), \quad (41)$$

with $\hat{\Gamma}_t = \text{diag}(\hat{\gamma}_t)$, which is used in the prediction step at time $t+1$. The proposed sequential SBL algorithm is summarized in Table I.

We can also modify the S-SBL approach with EM-based SBL by substituting Eq. (38) with Eq. (16). EM-based S-SBL provides improved DOA estimates over EM-based non-sequential SBL and similar accuracy to SIG-SBL. However, improvement in the estimates of EM-based S-SBL is limited, and the results of the EM-based S-SBL are omitted.

TABLE I. S-SBL algorithm: Input consists of data \mathbf{y}_t , the estimated variance $\hat{\gamma}_{t-1}$ of the source amplitudes at time $t-1$, and sensing matrix \mathbf{A} . Convergence is controlled by the error threshold ϵ_{\min} and maximum number of iterations N_{iter} .

S-SBL for DOA estimation

1. Input: \mathbf{y}_t , \mathbf{A} , \mathbf{F} , $\hat{\gamma}_{t-1}$
2. Parameters: $\epsilon_{\min} = 10^{-3}$, $N_{\text{iter}} = 1000$
3. Initialization: $\sigma_{w_t}^2 = \max[\hat{\gamma}_{t-1}] / 1000$, $\sigma_t^2 = 0.1$
4. Predict $\hat{\gamma}_{t|t-1}$ using Eq. (37)
5. $\gamma_t^{\text{old}} = \hat{\gamma}_{t|t-1}$
6. **for** $n_{\text{iter}} = 1$ to N_{iter}
7. Compute: $\Sigma_{\mathbf{y}_t} = \mathbf{A}\Gamma_t^{\text{old}}\mathbf{A}^\top + \sigma_t^2 \mathbf{I}_M$
8. γ_t^{new} update using Eq. (38)
9. $(\sigma_t^2)^{\text{new}}$ update using Eq. (40)
10. **If** $\|\gamma_t^{\text{new}} - \gamma_t^{\text{old}}\|_1 / \|\gamma_t^{\text{old}}\|_1 < \epsilon_{\min}$, **break**
11. $\gamma_t^{\text{old}} = \gamma_t^{\text{new}}$, $\sigma_t^2 = (\sigma_t^2)^{\text{new}}$
12. **end**
13. Output: $\hat{\gamma}_t (= \gamma_t^{\text{new}})$, $\sigma_t^2 (= (\sigma_t^2)^{\text{new}})$

VI. SIMULATION

We consider an array with $M=15$ elements and half-wavelength spacing, and 50 time steps are observed. The angle space is discretized using a grid size $N=361$ with 0.5° spacing, i.e., the potential DOAs $\theta = [-90\dots90]^\top \in \Theta^N$. The signal-to-noise ratio (SNR) is

$$\text{SNR} = 10 \log_{10} \left[\mathbb{E}\{\|\mathbf{Ax}_t\|\}_2^2 / \mathbb{E}\{\|\mathbf{v}_t\|\}_2^2 \right]. \quad (42)$$

The DOA root mean squared error (RMSE) is

$$\text{RMSE} = \sqrt{\mathbb{E} \left[\frac{1}{K_t} \sum_{k=1}^{K_t} (\hat{\theta}_{t,k} - \theta_{t,k})^2 \right]}, \quad (43)$$

where $\hat{\theta}_{t,k}$ and $\theta_{t,k}$ represent estimated and true DOA of the k th source at time t .

We study the performance of the SBL-based sequential processing, SIG-SBL [Eqs. (25) and (29)] and S-SBL [Eqs. (37), (38), and (40)], in comparison to conventional beamforming (CBF), SBL [Eqs. (18) and (19)] for a single time step,⁸ and SBL for multiple time steps (multiple measurement vectors) (MMV-SBL).⁷ For MMV-SBL, two-time step data are jointly processed, assuming stationary sources, i.e., \mathbf{y}_{t-1} and \mathbf{y}_t are processed to obtain $\hat{\mathbf{x}}_t$. We also consider the whole 50-time step data processing (whole MMV-SBL). S-SBL achieves superior DOA estimation (see Table II).

For the sequential SBL, SIG-SBL, and S-SBL, we use a rectangular 2-lags blurring window, i.e., $L=2$ and $f_0=f_1=f_2=0.2$ in Eq. (6). We set $\xi=1.5$ in Eq. (34) for SIG-SBL, which leads to favorable estimation errors.

We consider six stationary sources at DOAs $[-70^\circ, -55^\circ, -3^\circ, 2^\circ, 50^\circ, 65^\circ]$ with magnitudes [17, 17, 17, 12, 12, 22] dB at SNR = 20 dB (see Fig. 2). From sources 3 and 4, the resolution ability near broadside is examined and near endfire with same magnitude [sources 1 and 2] and different magnitude [sources 5 and 6]. The error histograms [see Figs. 2(b), 2(e), 2(f), 2(i), and 2(j)] show the error in DOA estimation for each source.

Non-sequential SBL shows low resolution near endfire when sources are close [see source 1 in Figs. 2(a) and 2(b)]. The DOA estimation is improved by jointly processing the multiple time step data (MMV-SBL) [see Figs. 2(c) and

TABLE II. RMSE for non-sequential SBL, SIG-SBL, S-SBL, MMV-SBL, and MMV-SBL for the whole 50-time step data (whole MMV-SBL) in Figs. 2-5.

	Non-sequential SBL	SIG-SBL	S-SBL	MMV-SBL	Whole MMV-SBL
Fig. 2	2.31°	0.92°	0.68°	1.37°	0.74°
Fig. 3	2.57°	0.69°	0.66°	2.42°	
Fig. 4(a)	2.05°	0.78°	0.65°		
Fig. 4(b)	2.55°	0.69°	0.62°		
Fig. 5(a)	2.99°	1.01°	0.67°		
Fig. 5(b)	2.67°	1.45°	0.67°		

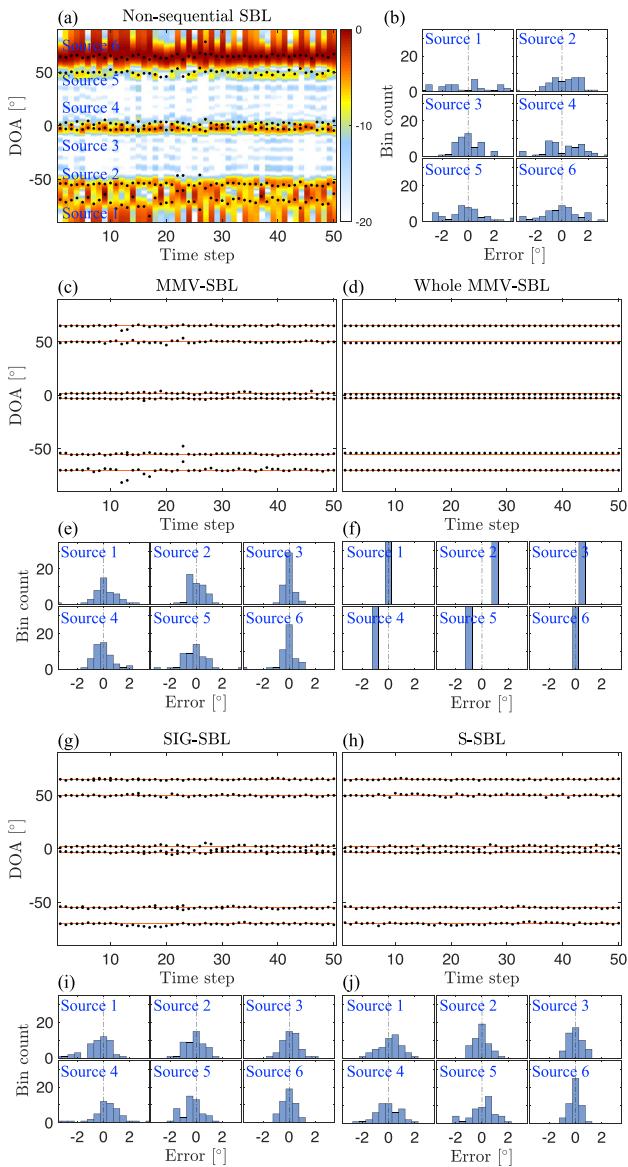


FIG. 2. (Color online) DOA estimation versus time steps with stationary sources for (a) non-sequential SBL, (c) MMV-SBL, (d) whole MMV-SBL, (g) SIG-SBL, and (h) S-SBL. DOA estimates are shown as black dots. The CBF is shown in the background (a), and the solution at each time is normalized by the maximum value. The solid lines [(c), (d), (g), and (h)] show the true source DOAs. Histograms of DOA error for each source are shown for (b) non-sequential SBL, (e) MMV-SBL, (f) whole MMV-SBL, (i) SIG-SBL, and (j) S-SBL.

2(e)]. The performance of MMV-SBL is improved further by processing more time steps (whole MMV-SBL) [see Figs. 2(d) and 2(f)]; however, it is infeasible for scenarios with moving sources. The SIG-SBL and S-SBL localize DOAs near endfire more accurately than non-sequential SBL and even MMV-SBL.

For a similar scenario as in Fig. 2, but with sources 3 and 4 moving, see Fig. 3. MMV-SBL fails to obtain accurate DOAs for moving sources as the stationary source assumption is violated [see sources 3 and 4 in Figs. 3(c)–3(f)]. S-SBL has lower RMSE than any other method.

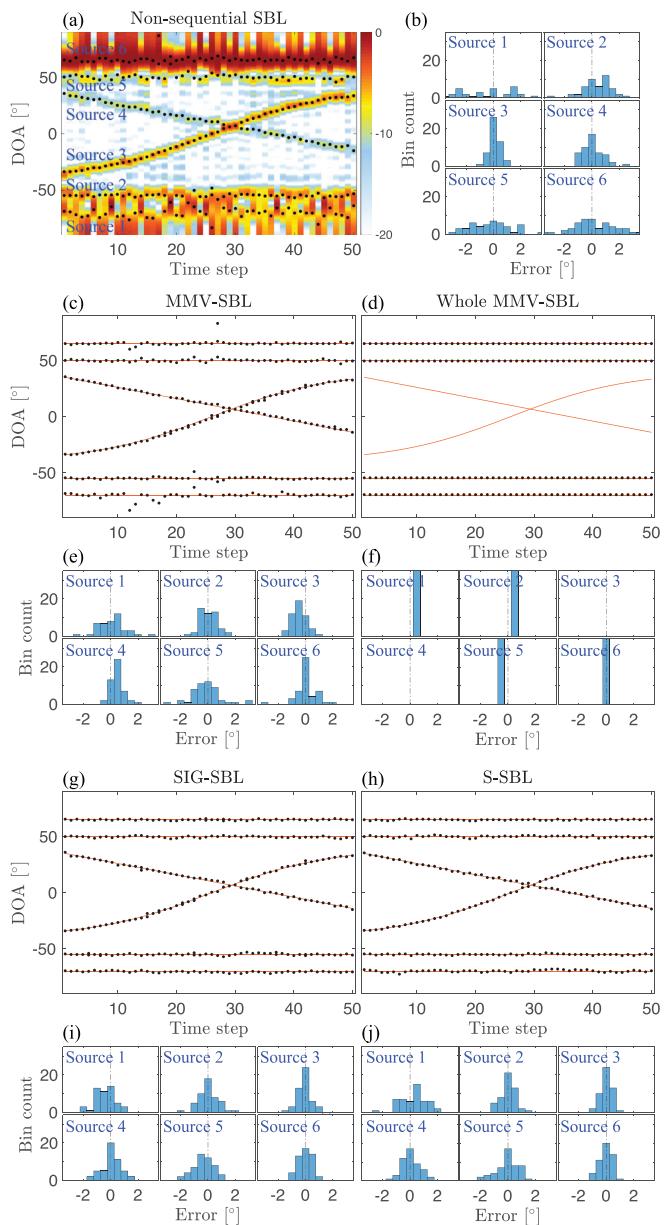


FIG. 3. (Color online) As in Fig. 2, but with moving sources.

Sources can suddenly appear and disappear, and a scenario where sources 1 and 2, 3 and 4, and 5 and 6 in Figs. 2 and 3 are deactivated is simulated (see Fig. 4). Since the update in sequential SBL incorporates current measurements that do not include deactivated DOAs, the current DOAs are estimated based on the previous estimated DOAs, but the deactivated DOAs in the current time are filtered. CBF shows a strong source for time steps 21–30, as each time is normalized by the maximum of the source strength and the strong source 6 is deactivated.

The time-varying system involves the time-varying motion of sources but also the time-varying magnitude of sources. We examine the time-varying source magnitude [see Fig. 5]. The source magnitude varies [0,20] dB for each source at SNR = 20 dB. SIG-SBL and S-SBL improve the

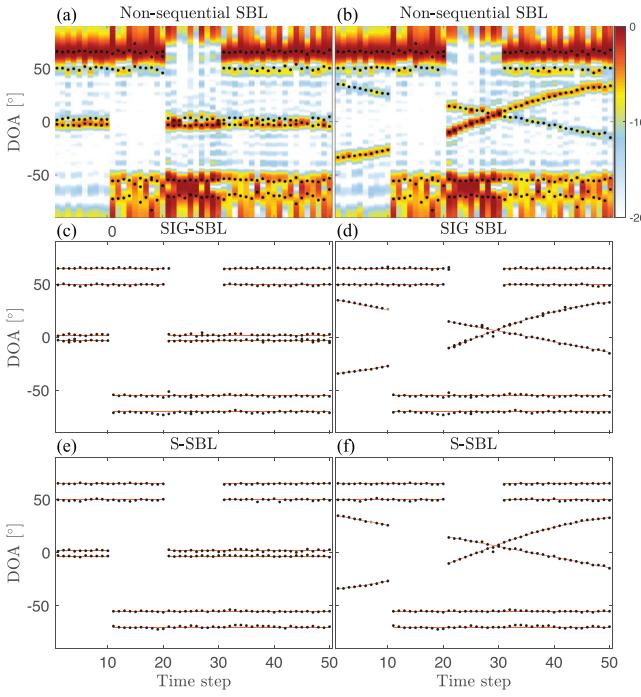


FIG. 4. (Color online) As in Fig. 3, but sources 1 and 2, 3 and 4, and 5 and 6 are deactivated sequentially: (a), (c), and (e) stationary sources; (b), (d), and (f) moving sources.

DOA estimation for both scenarios. S-SBL localizes DOAs more accurately than any other method.

We examine the effects of changing the ξ parameter for SIG-SBL in Fig. 6 and the effects of changing the state noise variance $\sigma_{w_t}^2$ for S-SBL in Fig. 7 using the scenarios in

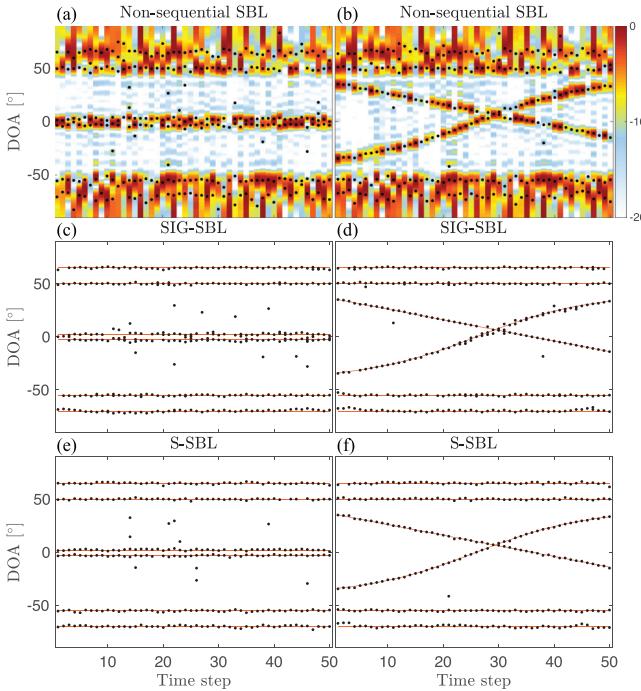


FIG. 5. (Color online) As in Fig. 3, but with time-varying magnitude of sources: (a), (c), and (e) stationary sources; (b), (d), and (f) moving sources.

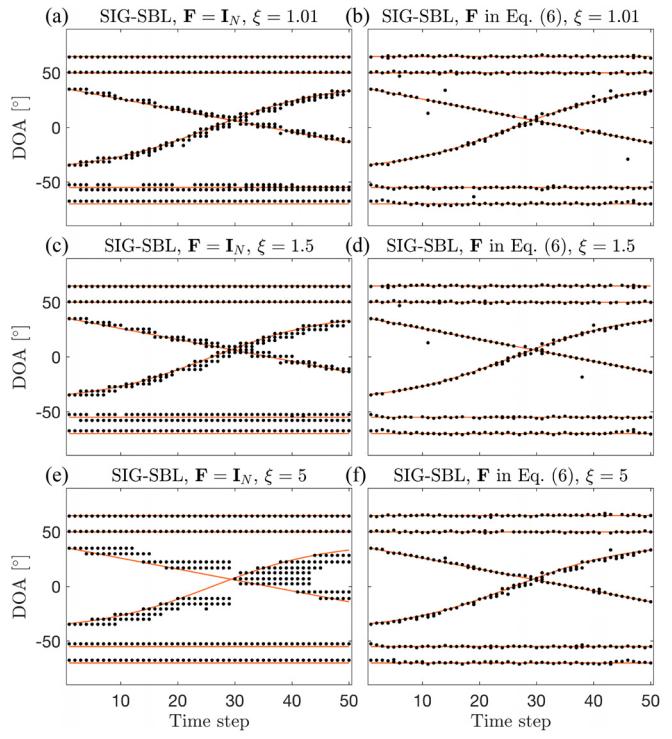


FIG. 6. (Color online) Effect of varying the parameter ξ for SIG-SBL using the scenario in Fig. 5(b) with state-transition models $\mathbf{F} = \mathbf{I}_N$ [(a), (c), and (e)] and \mathbf{F} [(b), (d), and (f)] in Eq. (6) ($L=2$ and $f_0=f_1=f_2=0.2$) for ξ 1.01 [(a) and (b)], 1.5 [(c) and (d)], and 5 [(e) and (f)].

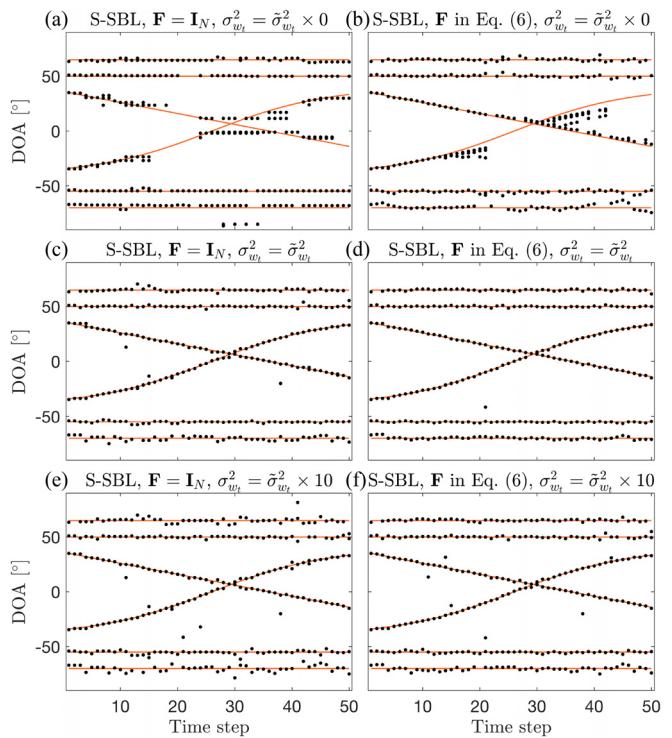


FIG. 7. (Color online) Effect of varying the state noise variance $\sigma_{w_t}^2$ for S-SBL using the scenario in Fig. 5(b) with state-transition models $\mathbf{F} = \mathbf{I}_N$ [(a), (c), and (e)] and \mathbf{F} [(b), (d), and (f)] in Eq. (6) ($L=2$ and $f_0=f_1=f_2=0.2$) for $\sigma_{w_t}^2=0$ [(a) and (b)], $\tilde{\sigma}_{w_t}^2$ [(c) and (d)], and $10\tilde{\sigma}_{w_t}^2$ [(e) and (f)], where $\tilde{\sigma}_{w_t}^2 = \max[\hat{y}_{t-1}] / 1000$.

Fig. 5(b). We use two matrices \mathbf{F} in the linear state-transition model [Eq. (5)]: $\mathbf{F} = \mathbf{I}_N$ and \mathbf{F} with a rectangular 2-lags blurring window, i.e., $L = 2$ and $f_0 = f_1 = f_2 = 0.2$ in Eq. (6). \mathbf{F} with a blurring window takes into account the source dynamics due to both source motion and measurement noise and gives more stable DOAs for both SIG-SBL and S-SBL. Increasing ξ for SIG-SBL enforces the prior information more strongly, which gives inaccurate estimation for moving sources [see Fig. 6]. For S-SBL, decreasing $\sigma_{w_t}^2$ enforces the prior information more strongly, which is inappropriate for moving sources [see Fig. 7].

The DOA performance is evaluated with the RMSE versus SNR in Fig. 8. We examine the scenarios for stationary sources (as Fig. 2) and moving sources (as Fig. 3), but six sources have equal magnitudes [10, 10, 10, 10, 10, 10] dB. SIG-SBL and S-SBL improve the DOA performance compared to non-sequential SBL. The improvement decreases at lower SNR, and the performance becomes similar.

Note that non-sequential SBL even outperforms SIG-SBL with $\mathbf{F} = \mathbf{I}_N$, which is explained by the fact that non-sequential SBL using a fixed-point update [Eq. (18)] outperforms non-sequential SBL using an EM update [Eq. (16)].

Sequential processing leads to much faster central processing unit (CPU) time. This is here demonstrated for one full realization of Fig. 2 with a 2.4 GHz Intel 8-core i9 processor: non-sequential SBL with fixed-point update (1.7 s) or with EM update (208.4 s), S-SBL (1.3 s) or with EM update (72.9 s), and SIG-SBL (48.4 s). An informative prior for each time step for sequential processing leads to faster convergence.

VII. EXPERIMENTAL RESULTS

The time-varying DOA performance of the sequential S-SBL is validated with experimental data, and it is compared with CBF. Non-sequential SBL is omitted, since sequential processing does not substantially improve the DOA performance. The data set is from the shallow water evaluation cell experiment 1996 (SWellEx-96)^{51,52} Event S5 recorded at 23:15–00:30 GMT on 10–11 May 1996 west of Point Loma, CA. The data are collected by a vertical uniform linear array, and two sources, a shallow and a deep,

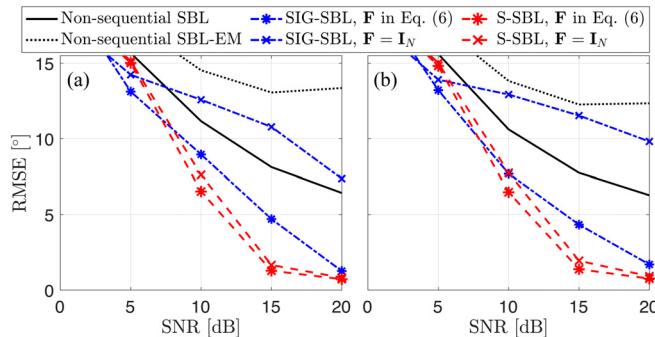


FIG. 8. (Color online) RMSE [°] versus SNR for (a) stationary and (b) moving sources. Each RMSE is averaged over 100 trials (50 time steps \times 100 trials, total 5000 time steps).

were towed simultaneously from 9 km southwest to 3 km northeast of the array with 5 kn (2.5 m/s). The array has $M = 64$ sensors with spacing $d = 1.875$ m and was deployed at a water depth of 94.125–212.25 m. (Element 43 was corrupted and thus excluded.)

We use the shallow source towed at 9 m depth at frequency 232 Hz. The data have a sampling frequency of 1500 Hz, and the record at 23:19–00:22 GMT is divided into non-overlapping 350 time steps. Each time step measurement is Fourier transformed with 2^{13} samples. Full 64-sensor array data (see Fig. 9) and a subset (32 elements, aperture: as the full array) with non-uniform intersensor spacing (see Fig. 10) are used for processing. The acoustic field was simulated using the Kraken normal mode model.⁵⁷ CBF results with the simulated field [see Figs. 9(a) and 10(a)] match well with those with the experimental data [Figs. 9(b) and 10(b)].

The source range varies from 8.3 km [time step 1] to 0.9 km [time step 305]. The strong source at around 20° corresponds to the direct path [time steps 250–350 in Fig. 9]. Due to the channel characteristics, the source cannot reach the receiver via the direct path, and a bottom-reflected path from -20° dominates (time steps 220–250 in Fig. 9). Similarly, multi-paths, which include surface-reflected paths and surface-bottom-reflected paths, alternately have significant strength [time steps 1–200 in Fig. 9]. S-SBL [Fig. 9(c)] results in improved source tracking resolution by promoting sparsity, capturing time-varying source dynamics, and reducing artifacts that CBF has.

We can encounter various array configurations,^{5,8,58–61} and a non-uniformly (or randomly) configured array measurement can outperform a uniformly configured array provided the same number of elements of the array.^{8,58} We

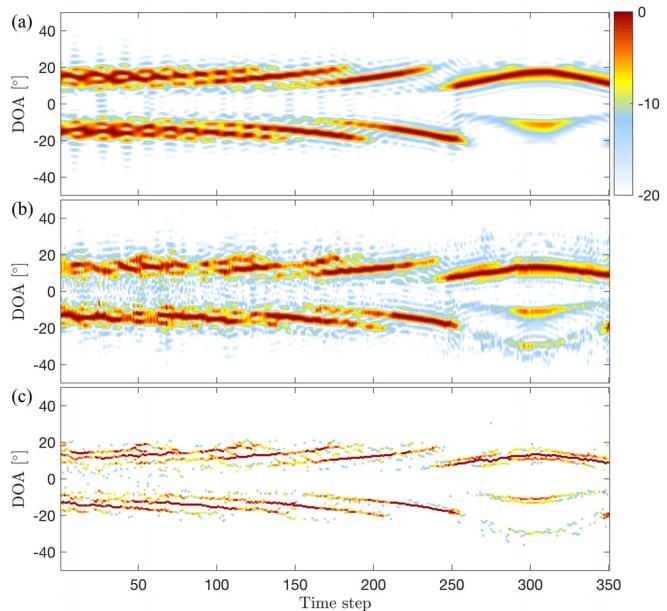


FIG. 9. (Color online) DOA estimation versus time for (a) CBF using simulated data from Kraken model and for (b) CBF and (c) S-SBL using acoustic data from the SWellEx-96 experiment. Colored dots [(c) S-SBL] show the SBL solution with source magnitudes.

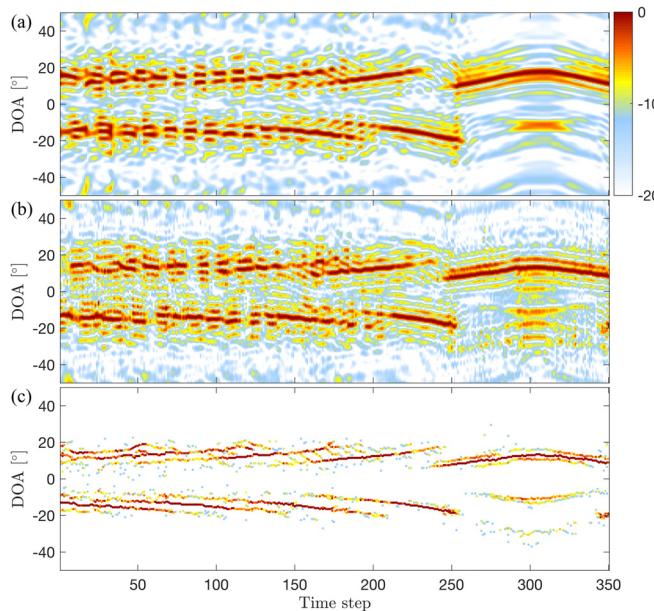


FIG. 10. (Color online) As in Fig. 9, but with non-uniformly spaced 32-element data. The array aperture is the same as that of the full 64-sensor array.

consider retaining only 1/2 of the sensors with a non-uniform configuration. 32-element data (elements 1, 3, 8, 9, 13, 14, 15, 17, 18, 20, 23, 26, 30, 31, 33, 34, 35, 38, 40, 41, 42, 44, 45, 47, 49, 51, 52, 55, 56, 57, 63, 64) are randomly chosen. CBF suffers from low resolution and artifacts, as in Fig. 10. S-SBL [Fig. 10(c)] results in improved resolution without a significant reconstruction degradation and provides the time-varying DOA tracking capability.

VIII. CONCLUSION

Estimation of time-varying DOAs was cast as a sequential sparse signal reconstruction problem. In particular, we developed two SBL approaches to reconstruct time-varying source amplitude variance parameters jointly with the noise power. In our sparse approach, the non-zero amplitude variance represents an active DOA.

To estimate the time-varying amplitude variance parameters, we considered two SBL-based methods. The first method, SIG-SBL, utilizes an inverse-gamma distribution for the variance prior and propagates the statistical information with time. The second method, S-SBL, uses a conventional sequential state-space model to predict the prior distribution for the current variance parameter from the previous time and estimates the current variance using SBL update rules.

An informative prior for sequential processing gives more robust estimation and faster convergence for both methods. S-SBL achieved here more accurate DOAs and faster convergence compared to SIG-SBL. On the other hand, SIG-SBL has a higher-level hierarchical Bayesian model, which has more flexibility on how information from previous time steps is incorporated.

Both methods are independent of array configuration and thus useful for a variety of realistic scenarios. Simulations showed that these sequential SBL methods provide improved DOA estimation accuracy for time-varying dynamic sources using much less CPU time. On real data, the good performance was demonstrated.

ACKNOWLEDGMENTS

This research was supported by Office of Naval Research Grant No. N00014-18-1-2118 and by the University of California San Diego.

- ¹D. Malioutov, M. Cetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Trans. Signal Process.* **53**(8), 3010–3022 (2005).
- ²A. Xenaki, P. Gerstoft, and K. Mosegaard, “Compressive beamforming,” *J. Acoust. Soc. Am.* **136**(1), 260–271 (2014).
- ³D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
- ⁴P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker, “Multiple and single snapshot compressive beamforming,” *J. Acoust. Soc. Am.* **138**(4), 2003–2014 (2015).
- ⁵Y. Park, W. Seong, and P. Gerstoft, “Block-sparse two-dimensional off-grid beamforming with arbitrary planar array geometry,” *J. Acoust. Soc. Am.* **147**(4), 2184–2191 (2020).
- ⁶D. 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).
- ⁷P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru, “Multisnapshot sparse Bayesian learning for DOA,” *IEEE Signal Process. Lett.* **23**(10), 1469–1473 (2016).
- ⁸S. Nannuru, A. Koochakzadeh, K. L. Gemba, P. Pal, and P. Gerstoft, “Sparse Bayesian learning for beamforming using sparse linear arrays,” *J. Acoust. Soc. Am.* **144**(5), 2719–2729 (2018).
- ⁹S. Nannuru, K. L. Gemba, P. Gerstoft, W. S. Hodgkiss, and C. F. Mecklenbräuker, “Sparse Bayesian learning with multiple dictionaries,” *Signal Process.* **159**, 159–170 (2019).
- ¹⁰C. Yardim, P. Gerstoft, W. S. Hodgkiss, and J. Traer, “Compressive geoaoustic inversion using ambient noise,” *J. Acoust. Soc. Am.* **135**(3), 1245–1255 (2014).
- ¹¹Y. Park, W. Seong, and Y. Choo, “Compressive time delay estimation off the grid,” *J. Acoust. Soc. Am.* **141**(6), EL585–EL591 (2017).
- ¹²A. Xenaki and Y. Pailhas, “Compressive synthetic aperture sonar imaging with distributed optimization,” *J. Acoust. Soc. Am.* **146**(3), 1839–1850 (2019).
- ¹³M. A. Badiu, T. L. Hansen, and B. H. Fleury, “Variational Bayesian inference of line spectra,” *IEEE Trans. Signal Process.* **65**(9), 2247–2261 (2017).
- ¹⁴J. Zhu, Q. Zhang, P. Gerstoft, M. A. Badiu, and Z. Xu, “Grid-less variational Bayesian line spectral estimation with multiple measurement vectors,” *Signal Process.* **161**, 155–164 (2019).
- ¹⁵K. L. Gemba, S. Nannuru, and P. Gerstoft, “Robust ocean acoustic localization with sparse Bayesian learning,” *IEEE J. Sel. Top. Signal Process.* **13**(1), 49–60 (2019).
- ¹⁶E. Fernandez-Grande, A. Xenaki, and P. Gerstoft, “A sparse equivalent source method for near-field acoustic holography,” *J. Acoust. Soc. Am.* **141**(1), 532–542 (2017).
- ¹⁷A. Drémeau, F. Le Courtois, and J. Bonnel, “Reconstruction of dispersion curves in the frequency-wavenumber domain using compressed sensing on a random array,” *IEEE J. Ocean. Eng.* **42**(4), 914–922 (2017).
- ¹⁸H. Niu, P. Gerstoft, E. Ozanich, Z. Li, R. Zhang, Z. Gong, and H. Wang, “Block sparse Bayesian learning for broadband mode extraction in shallow water from a vertical array,” *J. Acoust. Soc. Am.* **147**(6), 3729–3739 (2020).
- ¹⁹Y. Park, P. Gerstoft, and W. Seong, “Grid-free compressive mode extraction,” *J. Acoust. Soc. Am.* **145**(3), 1427–1442 (2019).

- ²⁰Z. Yang, L. Xie, and C. Zhang, “Off-grid direction of arrival estimation using sparse Bayesian inference,” *IEEE Trans. Signal Process.* **61**(1), 38–43 (2013).
- ²¹P. Gerstoft, S. Nannuru, C. F. Mecklenbräuker, and G. Leus, “DOA estimation in heteroscedastic noise,” *Signal Process.* **161**, 63–73 (2019).
- ²²R. Tibshirani, “Regression shrinkage and selection via the LASSO,” *J. Roy. Statist. Soc. B* **58**(1), 267–288 (1996).
- ²³I. F. Gorodnitsky and B. D. Rao, “Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm,” *IEEE Trans. Signal Process.* **45**(3), 600–616 (1997).
- ²⁴M. E. Tipping, “Sparse Bayesian learning and the relevance vector machine,” *J. Mach. Learn. Res.* **1**, 211–244 (2001).
- ²⁵D. Shutin and B. H. Fleury, “Sparse variational Bayesian SAGE algorithm with application to the estimation of multipath wireless channels,” *IEEE Trans. Signal Process.* **59**(8), 3609–3623 (2011).
- ²⁶Z. Yang and L. Xie, “On gridless sparse methods for line spectral estimation from complete and incomplete data,” *IEEE Trans. Signal Process.* **63**(12), 3139–3153 (2015).
- ²⁷A. S. Charles, A. Balavoine, and C. J. Rozell, “Dynamic filtering of time-varying sparse signals via ℓ_1 minimization,” *IEEE Trans. Signal Process.* **64**(21), 5644–5656 (2016).
- ²⁸A. Koochakzadeh, H. Qiao, and P. Pal, “On fundamental limits of joint sparse support recovery using certain correlation priors,” *IEEE Trans. Signal Process.* **66**(17), 4612–4625 (2018).
- ²⁹P. Gerstoft, C. F. Mecklenbräuker, W. Seong, and M. Bianco, “Introduction to compressive sensing in acoustics,” *J. Acoust. Soc. Am.* **143**(6), 3731–3736 (2018).
- ³⁰C. Yardim, Z. H. Michalopoulou, and P. Gerstoft, “An overview of sequential Bayesian filtering in ocean acoustics,” *IEEE J. Ocean. Eng.* **36**(1), 71–89 (2011).
- ³¹A. Koochakzadeh, P. Sarangi, and P. Pal, “A sequential approach for sparse support recovery using correlation priors,” in *Proceedings of the 2019 53rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA (November 3–6, 2019), pp. 586–590.
- ³²F. Meyer, Y. Park, and P. Gerstoft, “Variational Bayesian estimation of time-varying DOAs,” in *Proceedings of the 2020 IEEE International Conference on Information Fusion (FUSION)*, Rustenburg, South Africa (July 6–9, 2020), pp. 1–6.
- ³³M. S. Asif and J. Romberg, “Sparse recovery of streaming signals using ℓ_1 -homotopy,” *IEEE Trans. Signal Process.* **62**(16), 4209–4223 (2014).
- ³⁴N. Vaswani and J. Zhan, “Recursive recovery of sparse signal sequences from compressive measurements: A review,” *IEEE Trans. Signal Process.* **64**(13), 3523–3549 (2016).
- ³⁵J. Ziniel and P. Schniter, “Dynamic compressive sensing of time-varying signals via approximate message passing,” *IEEE Trans. Signal Process.* **61**(21), 5270–5284 (2013).
- ³⁶J. H. Yoo, S. H. Lim, B. Shim, and J. W. Choi, “Estimation of dynamically varying support of sparse signals via sequential Monte-Carlo method,” *IEEE Trans. Signal Process.* **68**, 4135–4147 (2020).
- ³⁷W. Li and J. C. Preisig, “Estimation of rapidly time-varying sparse channels,” *IEEE J. Ocean. Eng.* **32**(4), 927–939 (2007).
- ³⁸R. Prasad, C. R. Murthy, and B. D. Rao, “Joint approximately sparse channel estimation and data detection in OFDM systems using sparse Bayesian learning,” *IEEE Trans. Signal Process.* **62**(14), 3591–3603 (2014).
- ³⁹A. Carmi, P. Gurfil, and D. Kanevsky, “Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms,” *IEEE Trans. Signal Process.* **58**(4), 2405–2409 (2010).
- ⁴⁰C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, and M. Viberg, “Sequential Bayesian sparse signal reconstruction using array data,” *IEEE Trans. Signal Process.* **61**(24), 6344–6354 (2013).
- ⁴¹B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications* (Artech House, Norwood, MA, 2004), pp. 1–34.
- ⁴²F. Meyer, T. Kropfleiter, J. L. Williams, R. A. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, “Message passing algorithms for scalable multitarget tracking,” *Proc. IEEE* **106**(2), 221–259 (2018).
- ⁴³S. Farahmand, G. B. Giannakis, G. Leus, and Z. Tian, “Tracking target signal strengths on a grid using sparsity,” *EURASIP J. Adv. Signal Process.* **2014**(1), 1–17 (2014).
- ⁴⁴M. R. O’Shaughnessy, M. A. Davenport, and C. J. Rozell, “Sparse Bayesian learning with dynamic filtering for inference of time-varying sparse signals,” *IEEE Trans. Signal Process.* **68**, 388–403 (2020).
- ⁴⁵R. Giri and B. Rao, “Type I and Type II Bayesian methods for sparse signal recovery using scale mixtures,” *IEEE Trans. Signal Process.* **64**(13), 3418–3428 (2016).
- ⁴⁶A. Xenaki, E. Fernandez-Grande, and P. Gerstoft, “Block-sparse beamforming for spatially extended sources in a Bayesian formulation,” *J. Acoust. Soc. Am.* **140**(3), 1828–1838 (2016).
- ⁴⁷J. Antoni, C. Vanwynsbergh, T. L. Magueresse, S. Bouley, and L. Gilquin, “Mapping uncertainties involved in sound source reconstruction with a cross-spectral-matrix-based Gibbs sampler,” *J. Acoust. Soc. Am.* **146**(6), 4947–4961 (2019).
- ⁴⁸A. C. Barros and P. J. Gendron, “A computational Bayesian approach for localizing an acoustic scatterer in a stratified ocean environment,” *J. Acoust. Soc. Am.* **146**(3), EL245–EL250 (2019).
- ⁴⁹G. Ping, E. Fernandez-Grande, P. Gerstoft, and Z. Chu, “Three-dimensional source localization using sparse Bayesian learning on a spherical microphone array,” *J. Acoust. Soc. Am.* **147**(6), 3895–3904 (2020).
- ⁵⁰A. Xenaki, J. B. Boldt, and M. G. Christensen, “Sound source localization and speech enhancement with sparse Bayesian learning beamforming,” *J. Acoust. Soc. Am.* **143**(6), 3912–3921 (2018).
- ⁵¹N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D’Spain, and J. J. Murray, “Source localization with broad-band matched-field processing in shallow water,” *IEEE J. Ocean. Eng.* **21**(4), 402–412 (1996).
- ⁵²G. L. D’Spain, J. J. Murray, W. S. Hodgkiss, N. O. Booth, and P. W. Schey, “Mirages in shallow water matched field processing,” *J. Acoust. Soc. Am.* **105**(6), 3245–3265 (1999).
- ⁵³Z. 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).
- ⁵⁴S. Nannuru and P. Gerstoft, “2D beamforming on sparse arrays with sparse Bayesian learning,” in *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK (May 12–17, 2019), pp. 4355–4359.
- ⁵⁵S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Trans. Signal Process.* **53**(7), 2477–2488 (2005).
- ⁵⁶C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006), pp. 359–422.
- ⁵⁷M. B. Porter, *The KRAKEN Normal Mode Program* (SACLANT Undersea Research Centre, La Spezia, Italy, 1991).
- ⁵⁸A. Xenaki and P. Gerstoft, “Grid-free compressive beamforming,” *J. Acoust. Soc. Am.* **137**(4), 1923–1935 (2015).
- ⁵⁹P. Gerstoft, W. S. Hodgkiss, W. A. Kuperman, H. Song, M. Siderius, and P. L. Nielsen, “Adaptive beamforming of a towed array during a turn,” *IEEE J. Ocean. Eng.* **28**(1), 44–54 (2003).
- ⁶⁰Z. Zheng, T. C. Yang, P. Gerstoft, and X. Pan, “Joint towed array shape and direction of arrivals estimation using sparse Bayesian learning during maneuvering,” *J. Acoust. Soc. Am.* **147**(3), 1738–1751 (2020).
- ⁶¹K. Mahata and M. M. Hyder, “Grid-less T.V. minimization for DOA estimation,” *Signal Process.* **132**, 155–164 (2017).