

Received 15 December 2025, accepted 2 January 2026, date of publication 14 January 2026, date of current version 26 January 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3654100



RESEARCH ARTICLE

DOA Tracking Combining Compressive Sampling and Kalman Filtering for MIMO Radars Operating in Cluttered Environment

SOHEIL SALARI¹, (Member, IEEE), AND FRANCOIS CHAN², (Senior Member, IEEE)

¹Calian Group, Ottawa, ON K2V 1C8, Canada

²Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada

Corresponding author: Francois Chan (chan-f@rmc.ca)

ABSTRACT Taking into account the off-grid effect and clutter, a tailored compressive sampling (CS)-based algorithm is developed to estimate the Direction of Arrival (DOA) in radars that use Multiple-Input Multiple-Output (MIMO) technology and operate at sub-Nyquist rates. To enhance the performance of DOA estimation without resorting to denser sampling grids, which often entail expensive computational costs and robust sparse signal recovery issues, we incorporate the DOA estimates derived from the CS-based method into the state space of a Kalman filter. This integration eliminates the need for dense sampling grids, enhances computational efficiency, and improves robustness against clutter. Numerical experiments demonstrate that the proposed method consistently outperforms state-of-the-art techniques, achieving lower DOA error and superior receiver operating characteristic (ROC) performance.

INDEX TERMS Clutter, compressive sampling (CS), direction of arrival (DOA), Kalman filter, multiple-input multiple-output (MIMO) radar, sparse Bayesian learning (SBL).

I. INTRODUCTION

Radar employing Multiple-Input Multiple-Output (MIMO) technology have recently gained significant attention because of their enhanced performance and capabilities over traditional radar systems. MIMO radars typically operate in two main antenna configurations: widely separated [1], [2] and collocated [3], [4]. Widely separated MIMO radars observe targets from different directions simultaneously, leveraging spatial diversity for enhanced performance. On the other hand, collocated MIMO radars exploit phase differences introduced by various transmit and receive antenna pairs. This configuration provides superior resolution and performance in target detection compared to traditional phased-array radars. This work specifically aims to investigate the capability of the collocated MIMO radar in estimating the directions of arrival (DOAs) of several unidentified sources.

Compressive sampling (CS) [5], [6], [7] has emerged as a powerful framework for signal acquisition, enabling the reconstruction of signals from far fewer samples than

The associate editor coordinating the review of this manuscript and approving it for publication was Jolanta Mizera-Pietraszko .

required by the Nyquist criterion. According to CS theory, a signal $\mathbf{s} \in \mathbb{R}^N$ that is K -sparse can be accurately recovered from a small number of linear measurements

$$\mathbf{r} = \Theta \mathbf{s} + \mathbf{n}, \quad (1)$$

where $\Theta \in \mathbb{R}^{M \times N}$ represents the dictionary (sensing) matrix, $\mathbf{n} \in \mathbb{R}^M$ is noise, and $K < M \ll N$ [8]. This sampling paradigm has proven particularly effective in radar signal processing tasks, including DOA estimation, because of its capacity to overcome the constraints imposed by Nyquist sampling theory. In CS methods for DOA estimation, the spatial domain is discretized into a grid, and estimates are restricted to the nearest grid points [9]. Physically, the sparse vector \mathbf{s} represents the spatial reflectivity of the scene sampled on this discrete angular grid, with nonzero entries indicating directions that carry signal energy from unknown targets. Consequently, the support of the estimated vector $\hat{\mathbf{s}}$ directly corresponds to the DOAs. From a radar-practical perspective, any error in recovering the support or amplitudes of \mathbf{s} results in missed or false detections and higher DOA errors.

In CS formulations, achievable angular accuracy depends on the grid resolution, the mutual coherence of the sensing matrix, and the effective SNR. While finer grids reduce discretization bias and are essential for accurately resolving closely spaced targets, they also increase column correlation, which can degrade sparse-support recovery and amplify sensitivity to noise. Denser grids further introduce higher computational cost and longer processing times, posing challenges for real-time implementations [10]. To enhance DOA estimation performance without relying on denser sampling grids, various approaches have been proposed. Recently, a grid refinement strategy was implemented. Initially, a coarse grid was employed to estimate unknown DOAs, followed by the construction of a finer grid around DOA estimates to improve accuracy, as discussed in [9] and [11], and the references therein. Furthermore, off-grid DOA estimation techniques have been developed to address basis mismatch problems, as discussed in [12] and [13], and related works. Specifically, [12] proposes a framework that models off-grid deviations using a first-order Taylor expansion to enhance CS-based DOA estimation, while [13] introduces a CS-based algorithm that simultaneously estimates DOAs and compensates for unknown mutual coupling in MIMO radar systems. The limitations of [12] primarily stem from its reliance on first-order approximations, which can cause modeling errors when DOAs deviate significantly from grid points; it is also sensitive to the initial grid design and tends to underperform in the presence of clutter and non-Gaussian noise. Similarly, [13] suffers from high computational complexity due to matrix-intensive operations, sensitivity to grid resolution, and reduced reliability in low SNR conditions.

The limitations of conventional grid-based DOA estimation have motivated ongoing research into methods that are more accurate, computationally efficient, and robust. A key challenge remains the off-grid distance—the mismatch between actual DOAs and their nearest grid points—which can severely degrade performance, particularly when DOAs fall between adjacent grid points. This problem is especially pronounced in practical scenarios where target locations are unknown and optimal grid design is infeasible. Moreover, off-grid DOA estimation has not been thoroughly explored for radar systems operating in cluttered environments.

In this work, we address these challenges by integrating sparse Bayesian learning (SBL) with Kalman filtering to develop an advanced algorithm that eliminates the reliance on dense sampling grids. This integration enhances computational efficiency and significantly improves DOA estimation accuracy, particularly in dynamic and complex scenarios where the true DOAs may lie between predefined grid points.

While SBL and Kalman filtering are well-established techniques whose combination has been studied in the dynamic compressed sensing literature, the central contribution of our work is not the isolated use of these methods, nor a mere reiteration of their combination. Instead, our originality lies in the careful adaptation and extension of the SBL–Kalman

framework for DOA tracking in MIMO radar systems operating under clutter—a setting that introduces nontrivial methodological challenges. In particular, we explicitly tackle *off-grid discretization errors*, a primary source of angular bias in grid-based DOA estimation, while simultaneously accounting for the additional complexity of *clutter*, which is both practically significant and technically demanding. To the best of our knowledge, this scenario has not been systematically addressed in prior work and requires novel modeling and algorithmic strategies.

It is worth mentioning that several previous works have combined Kalman filtering with compressed sensing for tasks such as video reconstruction, streaming signal recovery, and dynamic sparse signal tracking; see [14], [15], and the references therein. However, these methods have not been applied to DOA estimation and generally overlook the off-grid scenario, in which the true target angles deviate from the discretized grid. This mismatch introduces angular bias and can substantially degrade estimation performance.

While CS and Kalman filtering have been independently explored in the dynamic CS literature, we emphasize that the contribution of this work is not the generic combination of these tools. Rather, the novelty lies in a radar-tailored SBL–KF framework that: (i) jointly estimates the DOAs, noise variance, and full clutter covariance matrix; (ii) models clutter as a structured and dynamically varying interference component instead of merging it with noise; (iii) mitigates off-grid discretization errors by embedding the SBL posterior into a physically meaningful Kalman state model linked to target kinematics; and (iv) achieves sub-grid angular resolution without requiring dense grids or multi-stage refinement. To the best of our knowledge, no prior work has developed an SBL–KF framework that simultaneously handles sub-Nyquist sampling, correlated clutter, and off-grid DOA tracking in collocated MIMO radar. This tailored integration produces an inference pipeline fundamentally different from existing CS-based DOA or dynamic CS approaches, yielding substantial improvements in accuracy, robustness, and ROC performance under challenging cluttered conditions where existing methods typically degrade.

The primary contributions of this paper are:

- We employ sub-Nyquist samples collected from MIMO radars operating in cluttered environments for DOA estimation. Sub-Nyquist sampling enables us to capture signals at lower rates than the traditional Nyquist rate, which reduces data processing requirements and computational load.
- We initially frame the DOA estimation as a sparse signal recovery problem within the CS framework, while explicitly accounting for the presence of clutter. To better capture the practical scenario in MIMO radar systems, we extend the standard CS model by introducing an additional term in equation (1) to represent the clutter component. This modification allows the compressed measurements to more accurately reflect

the combined contributions of the target signal, clutter, and noise. Solving the sparse recovery problem in (1) is known to be computationally intractable (NP-hard) [7], [16], [17], [18]. The challenge becomes even greater when unknown clutter is present, as it introduces additional uncertainty without prior information. A common workaround is to either treat interference and noise as a single aggregated disturbance or to neglect the weaker component—typically the noise. However, these simplistic treatments often degrade estimation accuracy. In this paper, we demonstrate that a probabilistic modeling approach offers a more effective and principled solution for handling the combined impact of noise and interference.

- Subsequently, a SBL-based algorithm is specifically designed to simultaneously estimate the DOAs of unknown targets along with the noise variance and the clutter covariance matrix. The sparse vector \mathbf{s} encodes target DOAs on a discrete angular grid, but grid-based discretization inherently causes off-grid errors. While finer grids reduce bias, they increase correlation and noise sensitivity, making recovery unstable. Existing methods therefore struggle, particularly when DOAs fall between adjacent grid points.
- To address the challenges of off-grid DOA estimation and tracking, we tightly integrate SBL with a Kalman filter. The Kalman update propagates a predictive prior and fuses it with the SBL posterior, (i) imposing temporal dynamics on the DOAs, (ii) interpolating between grid points, and (iii) accumulating measurements over time to reinforce accuracy. By directly incorporating SBL-based DOA estimates into the Kalman filter's state space, this synergy mitigates off-grid errors, produces robust and continuous DOA tracks, and enhances overall estimation performance. Beyond methodological novelty, this framework delivers clear operational benefits in realistic MIMO radar scenarios: higher detection probability at fixed false-alarm rates, lower DOA errors, and more stable tracks in dynamic, cluttered environments. Compared with conventional methods that rely on dense sampling grids, our approach offers a more efficient, accurate, and reliable solution for off-grid DOA estimation and tracking.
- We conduct extensive simulations across a range of operating conditions to evaluate and compare our method against existing off-grid DOA estimation techniques. The results demonstrate that our proposed approach provides a robust and computationally efficient solution for off-grid DOA estimation and tracking. Notably, it achieves superior accuracy compared to state-of-the-art methods, even in scenarios with substantial clutter.

The remainder of this paper unfolds as follows. Section II lays the groundwork by introducing the system model and reformulating the DOA estimation task as a sparse signal

recovery problem within the CS framework. Section III builds on this by detailing a probabilistic model and developing a SBL algorithm capable of jointly inferring source directions, noise variance, and clutter covariance. To further improve estimation accuracy, the SBL-derived DOA estimates are seamlessly integrated into a Kalman filtering framework, and an enhanced variant of the algorithm is proposed to tackle real-world implementation challenges. Section IV reports the algorithmic complexity of the proposed method in closed form, compares its asymptotic cost with representative baseline techniques, and discusses memory requirements alongside practical low-complexity implementation strategies. Section V presents simulation results that validate the algorithm's performance, followed by concluding remarks in Section VI.

Notation: Bold uppercase letters (e.g., \mathbf{A}) denote matrices and bold lowercase letters (e.g., \mathbf{a}) denote vectors. The operators $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ represent transpose, Hermitian (conjugate) transpose, and complex conjugation, respectively. All other symbols are defined when they are first introduced.

II. MIMO RADAR MODEL FOR MOVING TARGETS

This section models how each moving target contributes a structured space–time signal at the MIMO radar receivers, and how noise and clutter are incorporated. The resulting formulation captures both angular information (DOAs) and temporal pulse-to-pulse dynamics, making it suitable for CS-based sparse recovery and subsequent tracking.

In this paper, we investigate a collocated MIMO radar system consisting of P transmitting antennas and Q receiving antennas. The probing signals in MIMO radars typically consist of a train of D equally spaced pulses [19], i.e., $x_p(t) = \sum_{d=1}^D \tilde{x}_p(t - (d-1)T)$, $0 \leq t \leq DT$, where $x_p(t)$ ($p = 1, \dots, P$) represents the transmitted signal by p -th transmitting antenna and T denotes the pulse repetition interval (PRI).

With K far-field targets in the area of observation, the signal received at the q -th antenna can be formulated as [3] and [20]:

$$y_q(t) = \sum_{k=1}^K \sum_{d=1}^D \gamma_{k,d} \exp(j2\pi f_c \tilde{\tau}_q(\theta_{k,d})) \mathbf{a}^T(\theta_{k,d}) \tilde{\mathbf{x}}(t - (d-1)T), \\ q = 1, \dots, Q, \quad (2)$$

where $\gamma_{k,d}$ represents the complex amplitude, which is proportional to the radar cross section (RCS) of the k -th target with respect to the d -th pulse, f_c is the carrier frequency of the radar, and $\tilde{\tau}_q(\theta_{k,d})$ indicates the propagation time delay between the k -th target, relative to the d -th pulse, and the q -th receive antenna. Also, $\mathbf{a}(\theta_{k,d}) \triangleq [\exp(j2\pi f_c \tau_1(\theta_{k,d})), \dots, \exp(j2\pi f_c \tau_P(\theta_{k,d}))]^T_{P \times 1}$, where $\tau_p(\theta_{k,d})$ denotes the propagation time delay between the p -th transmit antenna and the k -th target with respect to the d -th pulse, and $\tilde{\mathbf{x}}(t - (d-1)T) \triangleq [\tilde{x}_1(t - (d-1)T), \dots, \tilde{x}_P(t - (d-1)T)]^T_{P \times 1}$. Notice that $\theta_{k,d}$ represents

the DOA of the k -th moving target during the d -th pulse. It is assumed that $\gamma_{k,d}$ and $\theta_{k,d}$ stay fixed during the PRI T but vary independently from one pulse to the next.

The received data vector $\mathbf{y}(t) \triangleq [y_1(t), \dots, y_Q(t)]_{Q \times 1}^T$ can be represented as:

$$\mathbf{y}(t) = \sum_{k=1}^K \sum_{d=1}^D \gamma_{k,d} \mathbf{b}(\theta_{k,d}) \mathbf{a}^T(\theta_{k,d}) \tilde{\mathbf{x}}(t - (d-1)T), \quad (3)$$

where $\mathbf{b}(\theta_{k,d}) \triangleq [\exp(j2\pi f_c \tilde{\tau}_1(\theta_{k,d}), \dots, \exp(j2\pi f_c \tilde{\tau}_Q(\theta_{k,d}))]_{Q \times 1}^T$. Suppose that each receiving antenna collects LT_s -spaced samples within d -th ($d = 1, \dots, D$) pulse. Then, the collected samples during the d -th pulse, i.e., $\mathbf{Y}_d = [\mathbf{y}((d-1)T + 0T_s), \dots, \mathbf{y}((d-1)T + (L-1)T_s)]_{Q \times L}$, can be expressed as¹:

$$\mathbf{Y}_d = \sum_{k=1}^K \gamma_{k,d} \mathbf{b}(\theta_{k,d}) \mathbf{a}^T(\theta_{k,d}) \mathbf{X}, \quad d = 1, \dots, D, \quad (4)$$

where $\mathbf{X} = [\tilde{\mathbf{x}}(0), \dots, \tilde{\mathbf{x}}((L-1)T_s)]_{P \times L}$.

In real-world applications, the received signal is affected by noise and clutter, which greatly hinder the performance of practical radar systems. Let $\mathbf{A} \otimes \mathbf{B}$ represent the Kronecker product of the matrices \mathbf{A} and \mathbf{B} . Taking into account the noise (denoted by \mathbf{e}_d) and clutter (denoted by \mathbf{w}_d) components, the output vector can be formed as [4]

$$\mathbf{y}_d = \sum_{k=1}^K \gamma_{k,d} (\mathbf{X}^T \mathbf{a}(\theta_{k,d})) \otimes \mathbf{b}(\theta_{k,d}) + \mathbf{w}_d + \mathbf{e}_d, \\ d = 1, \dots, D, \quad (5)$$

where \mathbf{y}_d is a vector of size $QL \times 1$, $\text{vec}[\mathbf{Y}_d] = \sum_{k=1}^K \gamma_{k,d} (\mathbf{X}^T \mathbf{a}(\theta_{k,d})) \otimes \mathbf{b}(\theta_{k,d})$, $\mathbf{w}_d \sim \mathcal{CN}(\mathbf{0}_{QL}, \Sigma_w)$, and $\mathbf{e}_d \sim \mathcal{CN}(\mathbf{0}_{QL}, \sigma^2 \times \mathbf{I}_{QL})$. Here, $\mathcal{CN}(\mu, \Sigma)$ represents the circularly symmetric complex Gaussian distribution with mean μ and covariance Σ . Assuming Gaussian noise is standard in CS-based MIMO radar research, justified by the Central Limit Theorem, which approximates the combined effect of multiple independent noise sources as Gaussian. This assumption enables analytical tractability, computational efficiency, and provides a common framework for performance comparison.

To adapt (5) for the context of CS, we define $\mathbf{g} \triangleq [g_1, g_2, \dots, g_N]$, where $N \gg K$, and sparse vector $\mathbf{s}_d \triangleq [s_{1,d}, \dots, s_{N,d}]_{N \times 1}^T$ such that [20]

$$s_{n,d} = \begin{cases} \gamma_{k,d}, & \text{if the } k\text{-th target is located at } g_n \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Also,

$$\Psi \triangleq [\mathbf{X}^T \mathbf{a}(g_1) \otimes \mathbf{b}(g_1), \dots, (\mathbf{X}^T \mathbf{a}(g_N)) \otimes \mathbf{b}(g_N)]_{QL \times N}. \quad (7)$$

Substituting (6) and (7) in (5) results in the following equation [4], [19]:

$$\mathbf{y}_d = \Psi \mathbf{s}_d + \mathbf{w}_d + \mathbf{e}_d, \quad d = 1, \dots, D. \quad (8)$$

¹Notice that the sampling interval T_s satisfies the Nyquist requirement.

Finally, exploiting the sparsity of \mathbf{s}_d enables us to express the CS measurements as:

$$\mathbf{r}_d = \Phi \mathbf{y}_d = \Theta \mathbf{s}_d + \Phi \mathbf{w}_d + \mathbf{n}_d, \quad d = 1, \dots, D, \quad (9)$$

in which, $\Theta \triangleq \Phi \Psi$ is a matrix with dimensions $M \times N$ ($K < M \ll N$), where Φ denotes a $M \times QL$ measurement matrix ($M < QL$), and $\mathbf{n}_d = \Phi \mathbf{e}_d$.

III. DOA ESTIMATION STRATEGY

Before presenting the detailed SBL-based derivations, we provide a brief high-level overview to help readers follow the probabilistic modeling steps introduced in this section.

This section introduces the core SBL probabilistic model: a Gaussian likelihood for the measurements and a Gaussian prior with adjustable variances to promote sparsity. These elements form the foundation for deriving the posterior distribution of the sparse vector and for enabling EM-based hyperparameter learning.

Denote the Euclidean norm of a vector \mathbf{a} by $\|\mathbf{a}\|_2$. We express the probability density function (pdf) of the random variable S as $p(s)$, distinguishing between random variables and deterministic parameters in the pdf using a semicolon notation. Specifically, we write $p(s; \boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ is a deterministic vector. Adopting the standard assumption of white (circularly symmetric) complex Gaussian noise, we reframe the signal model given in (9) within the SBL framework, as described in [21]:

$$p(\mathbf{r}_d | \mathbf{s}_d; \Sigma_w, \beta) = \frac{1}{(\pi\beta)^M} \exp\left(-\frac{1}{\beta} \|\mathbf{r}_d - \Theta \mathbf{s}_d - \Phi \mathbf{w}_d\|_2^2\right), \\ d = 1, \dots, D, \quad (10)$$

where β denotes the variance of the noise and Σ_w represents the covariance matrix of clutter. Estimating unknown parameters using ML from (10) can lead to significant overfitting challenges [22]. To mitigate this, we adopt a prior distribution for the sparse vector \mathbf{s}_d , following the strategy commonly used in SBL methodologies:

$$p(\mathbf{s}_d; \boldsymbol{\alpha}) = \prod_{n=1}^N \frac{1}{\pi \alpha_n} \exp\left(-\frac{|s_{n,d}|^2}{\alpha_n}\right), \quad d = 1, \dots, D, \quad (11)$$

where the vector $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_N]_{N \times 1}^T$ consists of independent hyperparameters that govern the sparsity characteristics of (9). Let $\text{diag}(\boldsymbol{\alpha})$ denote a diagonal matrix whose diagonal elements are given by the components of $\boldsymbol{\alpha}$. Defining $\mathbf{A} \triangleq \text{diag}(\boldsymbol{\alpha})$, (11) can be rewritten as:

$$p(\mathbf{s}_d; \boldsymbol{\alpha}) = \frac{1}{\pi^N |\mathbf{A}|} \exp\left(-\mathbf{s}_d^H \mathbf{A}^{-1} \mathbf{s}_d\right), \quad d = 1, \dots, D, \quad (12)$$

where $|\mathbf{A}|$ represents the determinant of the matrix \mathbf{A} . The probabilistic prior above plays a key role in promoting sparsity by allowing each coefficient to adaptively shrink through its corresponding hyperparameter α_n .

The conditional posterior distribution of \mathbf{s}_d given \mathbf{r}_d is modeled as a complex Gaussian distribution [23], expressed as:

$$p(\mathbf{s}_d|\mathbf{r}_d; \boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta) = \mathcal{CN}(\boldsymbol{\mu}_d, \boldsymbol{\Sigma}), \quad d = 1, \dots, D, \quad (13)$$

in which, $\boldsymbol{\mu}_d \triangleq E[\mathbf{s}_d|\mathbf{r}_d; \boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta] = \mathbf{A}\boldsymbol{\Theta}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r}_d$ ($E[\cdot]$ denotes the statistical expectation) and $\boldsymbol{\Sigma} \triangleq Cov[\mathbf{s}_d|\mathbf{r}_d; \boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta] = \mathbf{A} - \mathbf{A}\boldsymbol{\Theta}^H\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Theta}\mathbf{A}$, where $\boldsymbol{\Sigma}_r \triangleq \boldsymbol{\Theta}\boldsymbol{\Theta}^H + \boldsymbol{\Phi}\boldsymbol{\Sigma}_w\boldsymbol{\Phi}^H + \beta\mathbf{I}_M$.

Similar to typical SBL framework, we choose $\boldsymbol{\mu}_d$ as the sparse vector estimate, i.e.,

$$\hat{\mathbf{s}}_d \triangleq \boldsymbol{\mu}_d = \mathbf{A}\boldsymbol{\Theta}^H(\boldsymbol{\Sigma}_r)^{-1}\mathbf{r}_d, \quad d = 1, \dots, D. \quad (14)$$

Additionally, the SBL framework proposes estimating the unknown parameters by maximizing the marginal likelihood $p(\mathbf{r}_d; \boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta)$ [24], i.e.,

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\Sigma}}_w, \hat{\beta}) = \arg \max_{\boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta} \int p(\mathbf{r}_d|\mathbf{s}_d; \boldsymbol{\Sigma}_w, \beta)p(\mathbf{s}_d; \boldsymbol{\alpha})d\mathbf{s}_d, \\ d = 1, \dots, D, \quad (15)$$

which is central to the evidence maximization strategy. However, obtaining a closed-form solution for (15) is highly challenging. To address this, we develop an EM algorithm that iteratively solves (15) through the following two steps:

$$\begin{aligned} & \text{1-E-step: Compute } Q(\boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta | \boldsymbol{\alpha}^{(i)}, \boldsymbol{\Sigma}_w^{(i)}, \beta^{(i)}) \\ &= E_{\mathbf{u}|\mathbf{r}_d; \boldsymbol{\alpha}^{(i)}, \boldsymbol{\Sigma}_w^{(i)}, \beta^{(i)}} [\log p(\mathbf{u}; \boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta)] \end{aligned} \quad (16)$$

$$\begin{aligned} & \text{2-M-step: Solve } (\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\Sigma}_w^{(i+1)}, \beta^{(i+1)}) \\ &= \arg \max_{\boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta} Q(\boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta | \boldsymbol{\alpha}^{(i)}, \boldsymbol{\Sigma}_w^{(i)}, \beta^{(i)}), \end{aligned} \quad (17)$$

where i represents the EM iteration number and $\mathbf{u} = [\mathbf{s}_d^H, \mathbf{w}_d^H, \mathbf{n}_d^H]^H$ for $d = 1, \dots, D$. The EM algorithm alternates between estimating the hidden variables (E-step) and updating the model parameters (M-step). This iterative refinement enables SBL to jointly learn the sparse DOA coefficients, clutter covariance, and noise variance directly from the data, without requiring closed-form optimization.

Considering the Bayes' rule and recalling the independence of signal, noise, and clutter terms, the E-step is expressed as:

$$\begin{aligned} & Q(\boldsymbol{\alpha}, \boldsymbol{\Sigma}_w, \beta | \boldsymbol{\alpha}^{(i)}, \boldsymbol{\Sigma}_w^{(i)}, \beta^{(i)}) \\ &= Q(\boldsymbol{\alpha} | \boldsymbol{\alpha}^{(i)}) + Q(\boldsymbol{\Sigma}_w | \boldsymbol{\Sigma}_w^{(i)}) + Q(\beta | \beta^{(i)}), \end{aligned} \quad (18)$$

where

$$Q(\boldsymbol{\alpha} | \boldsymbol{\alpha}^{(i)}) = E_{\mathbf{s}_d|\mathbf{r}_d; \boldsymbol{\alpha}^{(i)}} [\log p(\mathbf{s}_d; \boldsymbol{\alpha})] \quad (19)$$

$$Q(\boldsymbol{\Sigma}_w | \boldsymbol{\Sigma}_w^{(i)}) = E_{\mathbf{w}_d|\mathbf{r}_d; \boldsymbol{\Sigma}_w^{(i)}} [\log p(\mathbf{w}_d; \boldsymbol{\Sigma}_w)] \quad (20)$$

$$Q(\beta | \beta^{(i)}) = E_{\mathbf{n}_d|\mathbf{r}_d; \beta^{(i)}} [\log p(\mathbf{n}_d; \beta)]. \quad (21)$$

Based on (18)–(21), the M-step (17) can be simplified as:

$$\boldsymbol{\alpha}^{(i+1)} = \arg \max_{\boldsymbol{\alpha}} Q(\boldsymbol{\alpha} | \boldsymbol{\alpha}^{(i)}) \quad (22)$$

$$\boldsymbol{\Sigma}_w^{(i+1)} = \arg \max_{\boldsymbol{\Sigma}_w} Q(\boldsymbol{\Sigma}_w | \boldsymbol{\Sigma}_w^{(i)}) \quad (23)$$

$$\beta^{(i+1)} = \arg \max_{\beta} Q(\beta | \beta^{(i)}). \quad (24)$$

These E/M-step expressions summarize the full probabilistic refinement process: the E-step computes expected signal, clutter, and noise contributions, while the M-step updates the hyperparameters that shape sparsity and statistical structure. This interpretation is provided to guide readers through the otherwise dense mathematical derivations.

In the appendices, we derive the relevant learning rules for (22) to (24). Let $trace(A)$ denote the trace of the matrix \mathbf{A} . By combining the learning rules outlined in the appendices, we can estimate the unknown parameters effectively. The algorithmic representation of the developed strategy is presented below.

Description of Algorithm I

1. Initialization:

Set $i = 0$ and initialize $\boldsymbol{\alpha}^{(0)}$, $\beta^{(0)}$, and $\boldsymbol{\Sigma}_w^{(0)}$.

2. Iteration:

2.1) E-step: Let $\mathbf{A}^{(i)} = diag(\boldsymbol{\alpha}^{(i)})$. Then, compute

$$\boldsymbol{\Sigma}_r^{(i)} = \boldsymbol{\Theta}\boldsymbol{\Theta}^H + \boldsymbol{\Phi}\boldsymbol{\Sigma}_w^{(i)}\boldsymbol{\Phi}^H + \beta^{(i)}\mathbf{I}_M \quad (25)$$

$$\mathbf{I}^{(i)} = (\boldsymbol{\Sigma}_r^{(i)})^{-1}\mathbf{r}_d\mathbf{r}_d^H(\boldsymbol{\Sigma}_r^{(i)})^{-1} - (\boldsymbol{\Sigma}_r^{(i)})^{-1} \quad (26)$$

2.2) M-step: Update the learning rules:

$$\boldsymbol{\alpha}_n^{(i+1)} = \boldsymbol{\alpha}_n^{(i)} + (\boldsymbol{\alpha}_n^{(i)})^2\boldsymbol{\Theta}_{.n}\boldsymbol{\Pi}^{(i)}\boldsymbol{\Theta}_{.n}, \quad (27)$$

$$\boldsymbol{\Sigma}_w^{(i+1)} = \boldsymbol{\Sigma}_w^{(i)} + \boldsymbol{\Sigma}_w^{(i)}\boldsymbol{\Phi}^H\boldsymbol{\Pi}^{(i)}\boldsymbol{\Phi}\boldsymbol{\Sigma}_w^{(i)}, \quad (28)$$

$$\beta^{(i+1)} = \beta^{(i)} + \frac{1}{M}(\beta^{(i)})^2trace(\boldsymbol{\Pi}^{(i)}). \quad (29)$$

2.3) Set $\boldsymbol{\alpha}_n^{(i+1)} = 0$, $n = 1, \dots, N$, if $\boldsymbol{\alpha}_n^{(i+1)} < \varepsilon_p$.

2.4) Increase i by one and proceed to step 2.1 until $\|\boldsymbol{\alpha}^{(i+1)} - \boldsymbol{\alpha}^{(i)}\|_2^2 < \varepsilon$ or until i reaches i_{max} .

Note: The values ε , ε_p and i_{max} are specified by the user.

3. Final approximation:

$$\hat{\mathbf{s}}_d = \mathbf{A}^{(i+1)}\boldsymbol{\Theta}^H(\boldsymbol{\Sigma}_r^{(i+1)})^{-1}\mathbf{r}_d, \quad d = 1, \dots, D.$$

The effectiveness of the proposed method in Table 1 heavily depends on the selection of initial values for the parameters $\boldsymbol{\alpha}^{(0)}$, $\beta^{(0)}$, and $\boldsymbol{\Sigma}_w^{(0)}$. According to [25], setting $\boldsymbol{\alpha}^{(0)}$ to a uniform vector often yields robust results. Nevertheless, when dealing with realistic environments containing noise and clutter, accurate initialization of $\beta^{(0)}$, and $\boldsymbol{\Sigma}_w^{(0)}$ becomes more challenging. In such cases, reliable estimates of the underlying noise and clutter statistics—typically derived from measurements taken in the absence of signal—are essential to achieve satisfactory performance.

Through comprehensive simulations, we observe that in an ideal, noise- and clutter-free environment, the majority of the estimated hyperparameters $\boldsymbol{\alpha}_n$, $n = 1, \dots, N$, rapidly converge toward zero within a limited number of EM iterations. However, under practical conditions involving measurement noise and environmental clutter, the estimated $\boldsymbol{\alpha}_n$, $n = 1, \dots, N$, values asymptotically approach small but nonzero magnitudes due to residual uncertainty and

model mismatch. To enforce sparsity and enhance numerical stability—particularly in avoiding ill-conditioning of covariance matrices—we introduce a hard-thresholding operation between successive EM updates. Specifically, any α_n , $n = 1, \dots, N$, satisfying $\alpha_n^{(i+1)} < \varepsilon_p$, where ε_p is a predefined sparsity threshold, is truncated to zero. This pruning heuristic functions as an implicit regularizer, yielding improved convergence characteristics without compromising estimation fidelity.

Based on our simulation results, setting the threshold ε_p approximately equal to the noise variance enables the system to achieve a high detection probability.

To enhance the accuracy of Algorithm 1, increasing the grid size N is an option. However, generating dense grids is computationally expensive. To address this issue, we introduce a modified version, Algorithm 2, which utilizes the KF. Algorithm 2 achieves lower computational cost compared to building a dense grid, while also delivering improved performance. Additionally, it mitigates errors associated with off-grid targets.

Before presenting the detailed KF-based formulation, we provide a brief high-level summary to clarify its purpose. Algorithm 2 integrates Kalman filtering to exploit the temporal evolution of DOAs across pulses, enabling accurate tracking without requiring a densely sampled angle grid. This significantly reduces computational cost while mitigating off-grid errors that affect purely grid-based sparse recovery.

Using the same approach as that given in [26], the state of the k -th moving target at the d -th pulse is described as $\theta_{k,d} \triangleq [\theta_{k,d}, \dot{\theta}_{k,d}]^T_{2 \times 1}$, where $\dot{\theta}_{k,d}$ represents the angular velocity (rate of change of DOA). The Kalman filter assumes the true state at time instant $d + 1$, $\theta_{d+1} \triangleq [\theta_{1,d+1}^T, \dots, \theta_{K,d+1}^T]^T_{2K \times 1}$ is evolved from the state at time instant d , $\theta_d \triangleq [\theta_{1,d}^T, \dots, \theta_{K,d}^T]^T_{2K \times 1}$, according to:

$$\theta_{d+1} = \mathbf{F}\theta_d + \mathbf{u}_d, \quad (30)$$

where $\mathbf{F} = \mathbf{I}_K \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$ represents the state transition model which is applied to the previous state θ_d and $\mathbf{u}_d = \mathcal{N}(\mathbf{0}, \mathbf{Q}_d)$ is the noise process.² Suppose that a new DOA estimate $\tilde{\theta}_{d+1} \triangleq [\tilde{\theta}_{1,d+1}, \dots, \tilde{\theta}_{K,d+1}]^T_{K \times 1}$ arrives from Algorithm 1 at time instant $d + 1$. The Kalman filter assumes a measurement model of the form [28]:

$$\tilde{\theta}_{d+1} = \mathbf{H}\theta_{d+1} + \mathbf{v}_{d+1}, \quad (31)$$

where $\mathbf{H}_{K \times 2K} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ & \dots & & \\ 0 & \dots & 1 & 0 \end{bmatrix}$ serves as the observation model, which translates the true state space into the observed space and $\mathbf{v}_{d+1} = \mathcal{N}(\mathbf{0}, \mathbf{R}_{d+1})$ is the observation noise. Notice that \mathbf{u}_d and \mathbf{v}_{d+1} are independent Gaussian random variables.

²For further details, we direct interested readers to [26], [27], and the references cited within those works.

The Kalman filter operates in two sequential steps over time: the prediction step and the update step [27], [28]. During the prediction phase, the estimate $\hat{\theta}_{d+1|d}$ is computed, accompanied by an uncertainty measurement represented by the prediction error covariance $\hat{\mathbf{P}}_{d+1|d}$. Following this, the update step takes place when a new estimate, $\tilde{\theta}_{d+1}$, becomes available. This new information is used to refine $\hat{\theta}_{d+1|d}$ and generate the updated estimate $\hat{\theta}_{d+1}$. Concurrently, $\hat{\mathbf{P}}_{d+1|d}$ is adjusted to yield $\hat{\mathbf{P}}_{d+1}$, reflecting the uncertainty associated with $\tilde{\theta}_{d+1}$. The key steps of the modified scheme are summarized below.

Algorithm II: Modified Algorithm

1. **Initialization:** Same as step 1 in Algorithm 1.
2. **Iteration:** Same as step 2 in Algorithm 1.
3. **Final Approximation:** Same as step 3 in Algorithm 1.
4. **Coarse DOA Estimation:** Determine non-zero elements of $\hat{\mathbf{s}}_d$ to obtain $\tilde{\theta}_d \triangleq [\tilde{\theta}_{1,d}, \dots, \tilde{\theta}_{K,d}]^T$.
5. **KF Prediction:**

$$\hat{\theta}_{d+1|d} = \mathbf{F}\theta_d, \quad (32)$$

$$\hat{\mathbf{P}}_{d+1|d} = \mathbf{Q}_d + \mathbf{F}\hat{\mathbf{P}}_d\mathbf{F}^T. \quad (33)$$

6. KF Correction:

$$\mathbf{G}_{d+1} \triangleq \hat{\mathbf{P}}_{d+1|d} \mathbf{H}^T (\mathbf{H}\hat{\mathbf{P}}_{d+1|d}\mathbf{H}^T + \mathbf{R}_{d+1})^{-1}, \quad (34)$$

$$\theta_{d+1} \triangleq \hat{\theta}_{d+1|d} + \mathbf{G}_{d+1}(\tilde{\theta}_d - \mathbf{H}\hat{\theta}_{d+1|d}), \quad (35)$$

$$\hat{\mathbf{P}}_{d+1} \triangleq \hat{\mathbf{P}}_{d+1|d} - \mathbf{G}_{d+1}\mathbf{H}\hat{\mathbf{P}}_{d+1|d}. \quad (36)$$

Increment d and go to step 1 until d reaches D .

7. **Fine DOA Estimation:** State Extraction: $\mathbf{H}\theta_D$
-

A. PRACTICAL INTERPRETATION AND RADAR IMPLICATIONS

Physically, the sparse vector \mathbf{s}_d represents the spatial reflectivity of the scene sampled on a discrete angular grid: nonzero entries correspond to directions that carry signal energy from unknown targets. Consequently, the support of the estimated vector $\hat{\mathbf{s}}_d$ directly maps to the DOAs. From a radar-practical perspective, any error in recovering the support or amplitudes of \mathbf{s}_d translates immediately into missed or false detections and angular bias. In compressed sensing formulations, the achievable angular accuracy depends on the grid resolution, the mutual coherence of the sensing matrix, and the effective SNR; very fine grids reduce discretization bias but increase column correlation (coherence), which degrades sparse-support recovery and heightens sensitivity to noise.

In our algorithm, the EM iterations implement a Type-II (SBL) estimation strategy: the E-step produces posterior means and covariances for \mathbf{s}_d , providing a soft detection with uncertainty information, while the M-step updates the hyperparameters and the clutter/noise covariance by maximizing the expected complete-data log-likelihood. From a radar-practical viewpoint, this alternation has two important consequences. First, the M-step's explicit estimation of

the clutter covariance and noise variance provides adaptive whitening and environment adaptation, increasing the effective SNR for subsequent sparse recovery and reducing bias caused by correlated interference. Second, the SBL hyperpriors concentrate energy on a small number of angular locations while providing posterior variances that can be leveraged to control false alarms.

Remaining discretization (off-grid) errors are a primary source of angular bias in grid-based methods: when a true DOA falls between grid points, its estimate is biased toward the nearest grid columns, and increasing grid density aggravates mutual coherence. To mitigate this, we couple SBL with a Kalman filter that imposes simple temporal dynamics on the DOAs. The Kalman update propagates a predictive prior and fuses it with the SBL posterior, effectively interpolating between grid points, reducing off-grid bias, and smoothing estimates across pulses. Operationally, this integration improves track continuity, enhances track stability, and reduces spurious track initiations and dropouts, leading to higher probability of detection at a fixed false alarm rate and lower DOA error.

IV. COMPUTATIONAL COMPLEXITY AND SCALABILITY

This section presents the algorithmic complexity of Algorithms 1 and 2 in closed form, compares their asymptotic cost with representative baseline methods such as MUSIC [29] and Orthogonal Matching Pursuit (OMP) [30], and discusses memory requirements along with practical low-complexity implementation strategies.

A. ASYMPTOTIC COST

Before proceeding, let P denote the number of transmit antennas, Q the number of receive antennas, K the number of targets, D the number of pulses, L the number of samples, N the grid size (angular dictionary), and M the number of compressed measurements per snapshot ($M < QL$). Throughout this section, the notation $\mathcal{O}(\cdot)$ is used to denote the asymptotic order of computational complexity with respect to the problem parameters.

Algorithm 1 (SBL-EM): A naïve implementation requires inversion of an $N \times N$ posterior covariance at each EM iteration, with complexity $\mathcal{O}(IN^3)$. Using the Woodbury identity reduces this to an $M \times M$ inversion, giving a per-iteration cost of

$$\mathcal{O}(M^2N + M^3),$$

dominated by M^2N when $N \gg M$. Over I EM iterations, the per-pulse complexity becomes

$$\mathcal{O}\left(I(M^2N + M^3)\right),$$

with memory cost $\mathcal{O}(MN)$ from storing the sensing dictionary.

Algorithm 2 (SBL + Kalman Filtering): Adding a Kalman recursion for K targets incurs $\mathcal{O}(K^3)$ per pulse. Since $K \ll N$ in practice, this overhead is negligible relative to the

SBL update, yielding

$$\mathcal{O}\left(I(M^2N + M^3) + K^3\right).$$

MUSIC: Eigen-decomposition of the $PQ \times PQ$ covariance costs $\mathcal{O}((PQ)^3)$, while the angular search adds $\mathcal{O}(N(PQ)^2)$.

OMP (Greedy CS): Each iteration requires correlations of cost $\mathcal{O}(MN)$, and with $\sim K$ non-zero elements in the sparse vector, the total cost is $\mathcal{O}(KMN)$.

Table 1 summarizes the asymptotic complexity and memory requirements. SBL-based methods, when implemented in measurement space, scale linearly in N and cubically in M , with memory $\mathcal{O}(MN)$. In contrast, MUSIC scales cubically with PQ and requires quadratic memory, which becomes prohibitive for large arrays. Greedy CS (OMP) scales as $\mathcal{O}(KMN)$ and is efficient for small K , but its performance degrades in cluttered or highly correlated scenarios where SBL remains robust.

TABLE 1. Asymptotic complexity (big- \mathcal{O}).

Method	Complexity	Memory Cost
SBL-EM	$\mathcal{O}(I(M^2N + M^3))$	$\mathcal{O}(MN)$
SBL + KF	$\mathcal{O}(I(M^2N + M^3) + K^3)$	$\mathcal{O}(MN)$
MUSIC	$\mathcal{O}((PQ)^3 + N(PQ)^2)$	$\mathcal{O}((PQ)^2)$
OMP	$\mathcal{O}(KMN)$	$\mathcal{O}(MN)$

B. SCALABILITY TRENDS

- Increasing P or Q raises M and sensing matrix coherence. Woodbury-based SBL remains feasible since the required inversions are of size $M \times M$, not $N \times N$.
- Increasing N improves resolution but increases the computational cost linearly with M^2N . Active-set pruning reduces this to $\mathcal{O}(M^2N_{act})$, with $N_{act} \propto K$.
- The number of targets K mainly affects the $\mathcal{O}(K^3)$ Kalman update, typically negligible compared to SBL.

C. IMPLEMENTATION NOTES

Scalability is further improved by:

- 1) Woodbury reformulation to avoid $N \times N$ inversions,
- 2) active-set pruning of the dictionary, and
- 3) GPU/FPGA acceleration for dense linear algebra.

These strategies make the methods scalable for realistic radar settings with large P , Q , L , and dense grids N .

V. SIMULATION RESULTS

This section details the numerical examples used to evaluate the performance of our proposed algorithm.³ The performance metrics are averaged over 1,000 independent Monte Carlo simulations. In each simulation, the elements of the

³The present study is limited to numerical experiments and does not include field measurements; incorporating real measurements in future work could serve to confirm the consistent performance gains our algorithm achieves over baseline methods.

measurement matrix Φ are randomly sampled from a Gaussian distribution.⁴ Additionally, the covariance matrix Σ_w for clutter is randomly drawn at each trial. It is important to highlight that the proposed algorithm is capable of estimating DOAs using significantly fewer samples—only $\frac{M}{QL} \times 100\%$ of the total—compared to conventional approaches. The specific values for the simulation parameters are provided in Table 2.

In all simulations, each performance point is averaged over $N = 1000$ independent Monte–Carlo trials. For the DOA RMSE curves, the resulting 95% confidence intervals are extremely small—far below the scale of the plotted axes. Accordingly, these intervals are omitted from the figures for clarity, which is consistent with standard practice in the signal processing and radar literature.

To quantify statistical reliability where it is most informative, we report 95% binomial confidence intervals for the probability of detection P_D of the proposed method in the ROC curve. For a given operating point, if k out of $N = 1000$ trials result in successful detection of all targets, the empirical detection probability is $\hat{p} = k/N$, and the corresponding 95% confidence interval is

$$\text{CI}_{95} = \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}. \quad (37)$$

These intervals are shown as vertical error bars in Fig. 6, while CIs for the baseline methods (MUSIC and OMP) are omitted to avoid clutter.

To evaluate DOA estimation accuracy, we use the well-established expression from [12]:

$$\sqrt{\frac{1}{KD} \sum_{d=1}^D \sum_{k=1}^K (\theta_{k,d} - \hat{\theta}_{k,d})^2}, \quad (38)$$

where $\theta_{k,d}$ is the true DOA of the k -th target at pulse d , and $\hat{\theta}_{k,d}$ is its estimate.

Following [31], interference strength is expressed in terms of the signal-to-interference ratio (SIR):

$$\text{SIR} = \frac{P_s}{P_c + P_n} = \frac{\text{SNR}}{1 + P_c/P_n}, \quad (39)$$

where P_s , P_n , and P_c denote signal, noise, and clutter powers, and $\text{SNR} = P_s/P_n$.

We also evaluate detection performance through the probability of detection (P_D) and the probability of false alarm (P_{FA}). As in [8] and [32], P_D denotes the fraction of trials in which all targets are successfully detected, and P_{FA} denotes the fraction producing any false target. The clutter-to-noise ratio is fixed at $P_c/P_n = 10$ dB unless otherwise stated.

Before presenting the main simulation examples, we examine two factors that strongly influence the behavior of the proposed algorithm: (i) the number of EM iterations needed

⁴For CS to achieve successful signal recovery, it is essential for the measurement matrix Φ to be incoherent with the basis matrix Ψ . According to the recommendations in [5], one straightforward way to ensure this incoherence is to choose Φ as a randomly generated matrix.

TABLE 2. Simulation parameters.

Parameter	Value	Comment
P	10	Num. of transmit antennas
Q	10	Num. of receive antennas
K	3	Num. of targets
D	2	Num. of pulses
L	64	Num. of samples
N	40	Grid size
M	<QL	Num. of CS measurements

for convergence, and (ii) the impact of the angular grid resolution.

Our simulation results show that the proposed algorithm typically converges within a fixed number of EM iterations. Figure 1 plots DOA estimation error versus iteration index i for multiple SIR levels. Convergence is reliably achieved within 20–25 iterations for all cases. Thus, all subsequent simulations terminate when either $\|\alpha^{(i+1)} - \alpha^{(i)}\|^2 < 10^{-8}$ or the iteration count reaches 25.

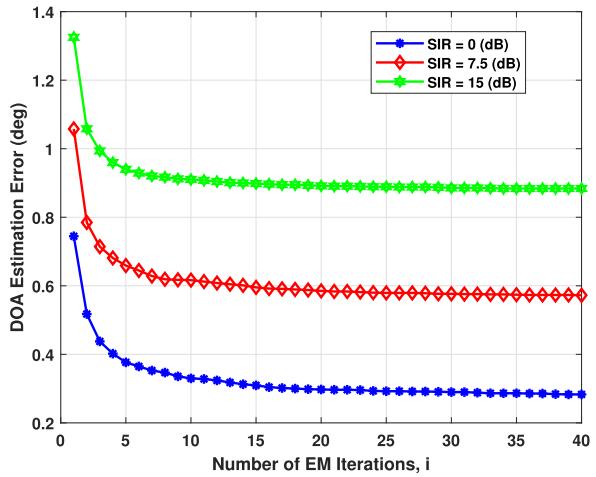


FIGURE 1. DOA estimation performance versus EM iteration index i for different SIR values.

Next, we assess the impact of the angular grid resolution on DOA estimation accuracy. Figure 2 shows the DOA estimation error as a function of the grid spacing for different SIR levels. In these simulations, the angular grid spacing is varied over

$$\Delta\theta \in \{0.25^\circ, 0.5^\circ, 1^\circ, 1.5^\circ, 2^\circ\},$$

which correspond to grid sizes ranging from $N = 160$ down to $N = 16$. For each value of $\Delta\theta$, the dictionary grid is constructed as $g = [-20^\circ, -20^\circ + \Delta\theta, \dots, 20^\circ]$, while all other simulation parameters remain unchanged.

Figure 2 shows that when $\Delta\theta < 0.5^\circ$, the dictionary becomes highly coherent, degrading sparse reconstruction, while for $\Delta\theta > 2^\circ$ the off-grid mismatch dominates and the DOA error increases. A spacing of $\Delta\theta = 1^\circ$ offers the most favorable trade-off between accuracy, dictionary size, and computational load.

Based on this analysis, a grid spacing of $\Delta\theta = 1^\circ$ is used in all subsequent simulation examples, as it provides a reliable and computationally efficient operating point. We now proceed to the main performance evaluations.

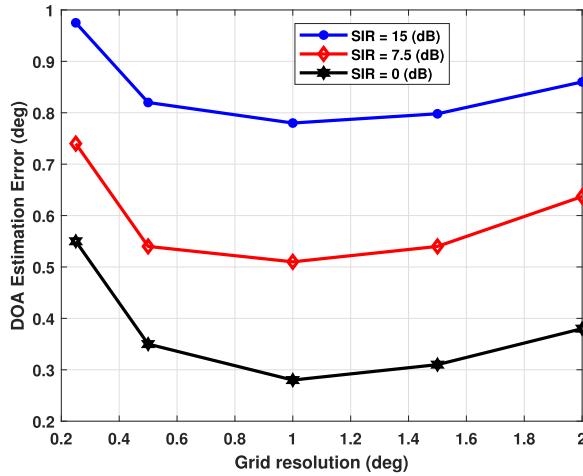


FIGURE 2. DOA estimation error versus grid resolution $\Delta\theta$.

Example 1: This experiment involves a performance comparison between Algorithm 2 and Algorithm 1. We assume that three moving targets with constant velocity are present in the area of interest, located at angles of -6.5° , 0° , 7.5° (two off-grid targets and one on-grid target).

Fig. 3 illustrates the DOA estimation performance against SIR for various percentages of M/QL values. Remarkably, Algorithm 2 demonstrates better accuracy compared to Algorithm 1 across the entire SIR range and for all M/QL values. This improvement of approximately 5 dB for 90 percent of samples is attributed to the utilization of the Kalman Filter in conjunction with the SBL technique. Additionally, Table 3 presents a comparison of the average running times of Algorithms 1 and 2 for different percentages of M/QL values. The implementation cost of Algorithm 2 is nearly the same, with only a slight increase compared to Algorithm 1 while offering better performance. This result is consistent with the computational complexity reported in Section IV (Table 1).

TABLE 3. Comparison of average running times.

$\frac{M}{QL}$ (%)	50%	70%	90%
Alg. I	6.57 s	10.52 s	15.81 s
Alg. II	6.96 s	11.13 s	16.67 s

Example 2: To further validate the effectiveness of Algorithm 2, its performance is compared with that of the proposed DOA estimator presented in [11], since both algorithms are designed to operate in cluttered environments. In Fig. 4, we present the DOA estimation performance against SIR for $M/QL = 75\%$. It is seen that our proposed scheme demonstrates superior accuracy compared to [11] across the entire SIR range. The performance improvement of almost

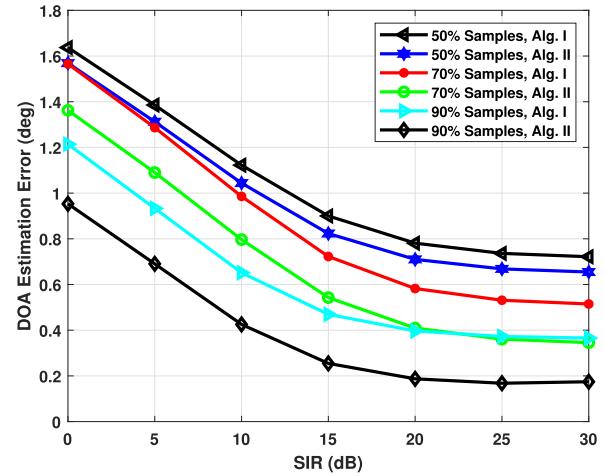


FIGURE 3. DOA estimation performance for different sampling ratios M/QL .

10 dB occurs because existing algorithms, such as [11], produce the best DOA estimate as the grid point nearest to the true DOA (e.g., 7.77° is mapped to 8°). In contrast, our scheme achieves a resolution smaller than the off-grid distance by leveraging the Kalman Filter to mitigate off-grid errors. Both methods are shown to perform well in cluttered environments. However, [11] struggles with the off-grid distance—the discrepancy between actual DOAs and their nearest grid points—which can significantly degrade performance, particularly when the true DOA lies exactly midway between two adjacent grid points (e.g., -6.5° being mapped to either -6° or -7°). This condition is especially relevant in real-world scenarios, where the precise target location is unknown and optimal grid placement is impractical. Remarkably, our proposed technique maintains significantly better performance even under this worst-case condition.

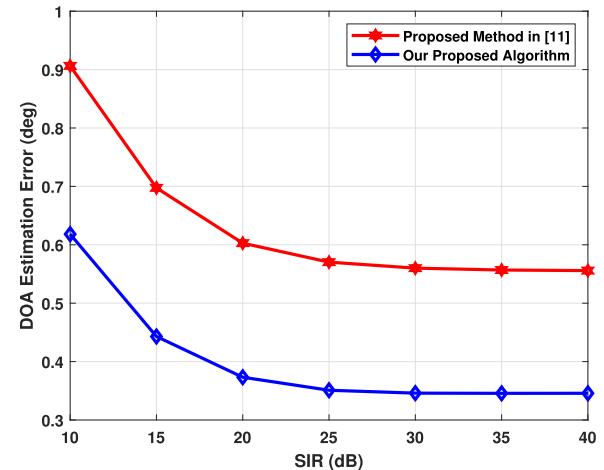


FIGURE 4. DOA estimation performance for off-grid targets in a cluttered environment.

Example 3: As a performance benchmark, we also compare our proposed algorithm with the method developed in [12],

as shown in Fig. 5. It is important to note that the other related works such as [12] and [13] limited their investigations to an ideal, clutter-free environment. Consequently, their accuracy experiences a significant decrease in the presence of clutter. To ensure a fair comparison, we eliminate the effect of clutter by setting $P_c = 0$ in (39), while keeping all other parameters identical to those used in *Example 2*. Under these conditions, Fig. 5 illustrates that our proposed algorithm achieves consistently higher DOA estimation accuracy than [12] across the entire SNR range. This result highlights the effectiveness of our approach in handling off-grid targets, even in scenarios where conventional methods like [12] typically perform well.

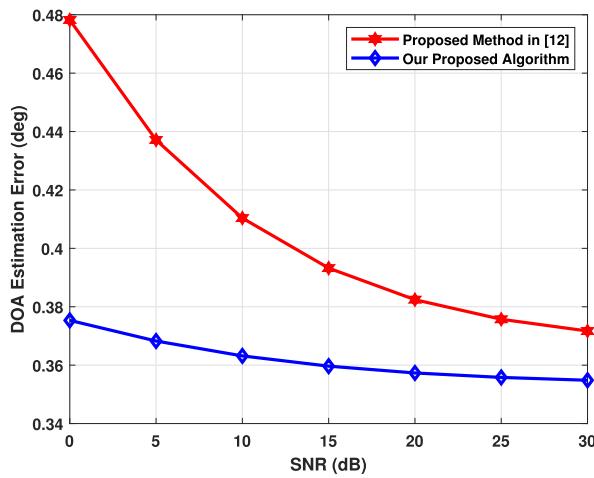


FIGURE 5. DOA estimation performance comparison between the proposed algorithm and the off-grid approach presented in [12].

Finally, to validate the effectiveness of the proposed algorithm, Fig. 6 presents the ROC performance in comparison with the MUSIC [29] and OMP [30] algorithms. The proposed scheme achieves a consistently higher probability of detection across the entire range of false alarm probabilities, as confirmed by the 95% binomial confidence intervals (vertical bars) computed over 1000 Monte-Carlo trials. This performance gain arises from the fact that the proposed algorithm explicitly estimates both the clutter covariance matrix and the noise variance, thereby providing a more accurate characterization of the interference environment. By exploiting this additional knowledge of the clutter statistics, the method significantly reduces the impact of structured interference, which explains the noticeable performance gap relative to MUSIC and OMP, both of which rely on less comprehensive interference modeling. Another key advantage is that the proposed approach enables reliable target detection from a markedly reduced number of samples compared to standard CS (Nyquist-sampled) techniques.

Example 4: We next examine the robustness of the proposed method under different clutter conditions by varying the clutter-to-noise ratio P_c/P_n , which directly controls the strength of structured interference in the received data. This analysis complements the previous examples by evaluating

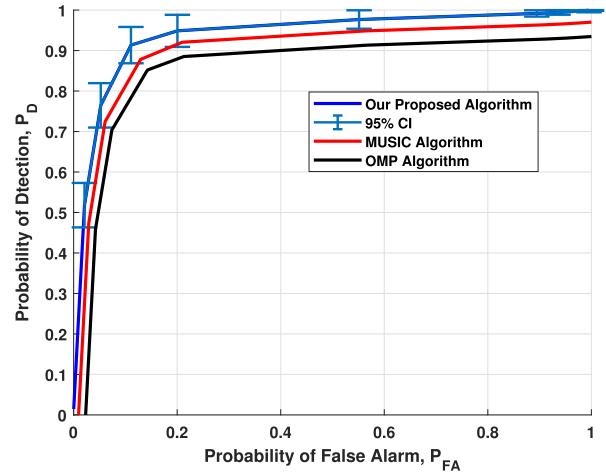


FIGURE 6. ROC performance of the proposed algorithm versus MUSIC [29] and OMP [30]. The proposed method includes 95% binomial confidence intervals (vertical bars) computed over 1000 Monte-Carlo trials, while CIs for the baseline methods are omitted for clarity.

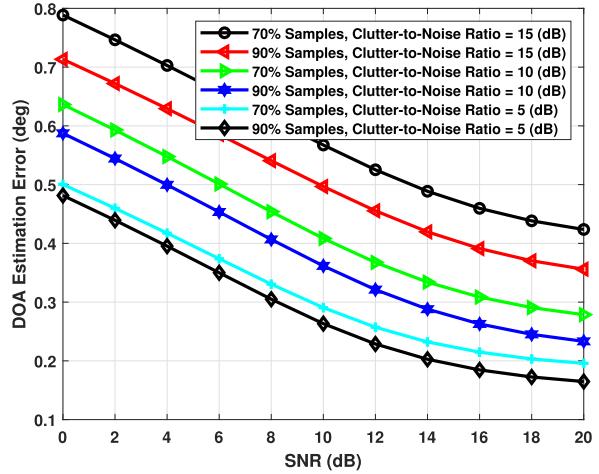


FIGURE 7. DOA estimation error versus SNR for different clutter-to-noise and sampling ratios.

whether the performance advantages observed earlier persist when the clutter environment becomes more or less severe.

Figure 7 plots the DOA estimation error versus SNR for three representative clutter levels, $P_c/P_n \in \{5, 10, 15\}$ dB, and for two sampling ratios, $M/QL = 70\%$ and 90% . As expected, increasing the clutter power leads to higher estimation error since a larger fraction of the received energy is dominated by interference. This behavior is consistent with practical radar scenarios, where clutter strength may vary significantly due to terrain, weather, foliage, or platform motion.

An important feature of the proposed framework is its operational flexibility. The sampling ratio M/QL can be adjusted based on clutter severity: in heavy clutter, using a larger sampling ratio improves robustness by providing more informative measurements, while in mild clutter the operator may safely lower M/QL to reduce sampling load

and computational cost without sacrificing accuracy. This adaptability offers a clear advantage over existing approaches that typically require full-rate sampling.

Even under the most challenging condition ($P_c/P_n = 15$ dB), Figure 7 shows that the proposed algorithm maintains stable and accurate DOA estimates over all SNR values and sampling ratios. These results demonstrate strong resilience to structured interference and confirm the effectiveness of the proposed method in both moderate and clutter-intensive environments.

VI. CONCLUSION

This work introduces an efficient framework for estimating the DOAs of potential targets in cluttered environments using CS-based MIMO radars operating at sub-Nyquist rates. We first employ a SBL approach tailored to jointly estimate the DOAs of unknown targets together with the noise variance and clutter covariance matrix. The estimated DOAs are then incorporated into the state space of a Kalman filter, enabling dynamic tracking. The integration of SBL with Kalman filtering offers a robust solution to the off-grid target problem in DOA estimation. From an operational standpoint, this joint strategy enhances track continuity and stability while suppressing spurious initiations and track dropouts. Consequently, it achieves a higher probability of detection at a fixed false alarm rate and yields lower DOA error. Numerical experiments confirm that our proposed approach offers an efficient and robust framework for off-grid DOA estimation and tracking. It consistently outperforms state-of-the-art methods in estimation accuracy, achieves superior ROC performance, and maintains robustness even in the presence of significant clutter. In future work, we plan to extend our analysis to non-Gaussian interference models and confirm the clear performance advantages our method demonstrates over baseline techniques.

APPENDIX A

UPDATE RULE FOR α

We begin by substituting (11) in (22), i.e.,

$$\boldsymbol{\alpha}^{(i+1)} = \arg \min_{\boldsymbol{\alpha}} E_{\mathbf{s}_d | \mathbf{r}_d; \boldsymbol{\alpha}^{(i)}} \left[\sum_{n=1}^N \left(\log \pi + \log \alpha_n + \alpha_n^{-1} |\mathbf{s}_{n,d}|^2 \right) \right]. \quad (\text{A.1})$$

The derivative of (A.1) concerning α_n can be formulated as:

$$\frac{\partial Q(\boldsymbol{\alpha} | \boldsymbol{\alpha}^{(i)})}{\partial \alpha_n} = \frac{1}{\alpha_n} - \frac{1}{\alpha_n^2} E_{\mathbf{s}_d | \mathbf{r}_d; \boldsymbol{\alpha}^{(i)}} \left[|\mathbf{s}_{n,d}|^2 \right], \quad n = 1, \dots, N. \quad (\text{A.2})$$

The learning rule can be obtained by setting (A.2) to zero, i.e,

$$\alpha_n^{(i+1)} = \alpha_n^{(i)} + (\alpha_n^{(i)})^2 \boldsymbol{\Theta}_{n,n}^H \boldsymbol{\Pi}^{(i)} \boldsymbol{\Theta}_{n,n}, \quad n = 1, \dots, N, \quad (\text{A.3})$$

where

$$\boldsymbol{\Pi}^{(i)} \triangleq (\boldsymbol{\Sigma}_r^{(i)})^{-1} \mathbf{r}_d \mathbf{r}_d^H (\boldsymbol{\Sigma}_r^{(i)})^{-1} - (\boldsymbol{\Sigma}_r^{(i)})^{-1}. \quad (\text{A.4})$$

APPENDIX B

UPDATE RULE FOR Σ_w

Using (20) and (23) and recalling that $\mathbf{w}_d \sim \mathcal{CN}(\mathbf{0}_{QL}, \boldsymbol{\Sigma}_w)$, it can be proven that:

$$\boldsymbol{\Sigma}_w^{(i+1)} = \arg \min_{\boldsymbol{\Sigma}_w} \left(\log |\boldsymbol{\Sigma}_w| + \text{trace} \left(\boldsymbol{\Sigma}_w^{-1} \left[\boldsymbol{\Sigma}_w^{(i)} + \boldsymbol{\Sigma}_w^{(i)} \boldsymbol{\Phi}^H \boldsymbol{\Pi}^{(i)} \boldsymbol{\Phi} \boldsymbol{\Sigma}_w^{(i)} \right] \right) \right), \quad (\text{B.1})$$

where $\mathbf{w}_d^H \boldsymbol{\Sigma}_w^{-1} \mathbf{w}_d = \text{trace}(\boldsymbol{\Sigma}_w^{-1} \mathbf{w}_d \mathbf{w}_d^H)$. Applying Lemma 1 of [33], the learning rule presented in (B.1) can be articulated as:

$$\boldsymbol{\Sigma}_w^{(i+1)} = \boldsymbol{\Sigma}_w^{(i)} + \boldsymbol{\Sigma}_w^{(i)} \boldsymbol{\Phi}^H \boldsymbol{\Pi}^{(i)} \boldsymbol{\Phi} \boldsymbol{\Sigma}_w^{(i)}. \quad (\text{B.2})$$

APPENDIX C

UPDATE RULE FOR β

Considering (24) and recalling that $\mathbf{n}_d \sim \mathcal{CN}(\mathbf{0}_M, \beta \mathbf{I}_M)$ results in

$$\beta^{(i+1)} = \arg \min_{\beta} E_{\mathbf{n}_d | \mathbf{r}_d; \beta^{(i)}} M \left[\log \pi + \log \beta + \frac{\beta^{-1}}{M} \mathbf{n}_d^H \mathbf{n}_d \right]. \quad (\text{C.1})$$

To find the derivative of (C.1) concerning β , that is,

$$\frac{\partial Q(\beta | \beta^{(i)})}{\partial \beta} = \frac{M}{\beta} - \frac{1}{\beta^2} E_{\mathbf{n}_d | \mathbf{r}_d; \beta^{(i)}} \left[\mathbf{n}_d^H \mathbf{n}_d \right]. \quad (\text{C.2})$$

By equating (C.2) to zero, we arrive at the following learning rule:

$$\beta^{(i+1)} = \beta^{(i)} + \frac{1}{M} (\beta^{(i)})^2 \text{trace}(\boldsymbol{\Pi}^{(i)}). \quad (\text{C.3})$$

Notice that (C.3) was derived based on the fact that $\mathbf{n}_d^H \mathbf{n}_d = \text{trace}(\mathbf{n}_d \mathbf{n}_d^H)$.

REFERENCES

- [1] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, and R. A. Valenzuela, "Spatial diversity in radars—Models and detection performance," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 823–838, Mar. 2006.
- [2] A. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, May 2008.
- [3] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of MIMO radar," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 968–971, Dec. 2007.
- [4] L. Xu, J. Li, and P. Stoica, "Target detection and parameter estimation for MIMO radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 927–939, Jul. 2008.
- [5] R. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [6] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling (A sensing/sampling paradigm that goes against the common knowledge in data acquisition)," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [7] S. Gleichman and Y. C. Eldar, "Blind compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6958–6975, Oct. 2011.
- [8] Y. Yu, A. P. Petropulu, and H. V. Poor, "CSSF MIMO RADAR: Compressive-sensing and step-frequency based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1490–1504, Apr. 2012.
- [9] S. Salari, I.-M. Kim, F. Chan, and S. Rajan, "Blind compressive-sensing-based electronic warfare receiver," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 4, pp. 2014–2030, Aug. 2017.

- [10] S. Salari, F. Chan, Y.-T. Chan, and R. Guay, "DOA estimation using compressive sampling-based sensors in the presence of interference," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4395–4405, Dec. 2020.
- [11] S. Salari, F. Chan, Y.-T. Chan, I.-M. Kim, and R. Cormier, "Joint DOA and clutter covariance matrix estimation in compressive sensing MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 1, pp. 318–331, Feb. 2019.
- [12] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.
- [13] P. Chen, Z. Cao, Z. Chen, and X. Wang, "Off-grid DOA estimation using sparse Bayesian learning in MIMO radar with unknown mutual coupling," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 208–220, Jan. 2019.
- [14] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5270–5284, Nov. 2013.
- [15] K. Naruka and O. P. Sahu, "An improved speech enhancement approach based on combination of compressed sensing and Kalman filter," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, Dec. 2015, pp. 1–5.
- [16] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [17] M. Niu, S. Salari, I.-M. Kim, F. Chan, and S. Rajan, "Recovery probability analysis for sparse signals via OMP," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 4, pp. 3475–3479, Oct. 2015.
- [18] R. Prasad and C. R. Murthy, "Cramér-Rao-type bounds for sparse Bayesian learning," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 622–632, Feb. 2013.
- [19] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Spatial compressive sensing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 419–430, Jan. 2014.
- [20] L. Ding, W. Chen, W. Zhang, and H. V. Poor, "MIMO radar imaging with imperfect carrier synchronization: A point spread function analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 2236–2247, Jul. 2015.
- [21] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, Jul. 2007.
- [22] K. Hwang and S. Choi, "Blind equalization method based on sparse Bayesian learning," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 315–318, Apr. 2009.
- [23] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 912–926, Sep. 2011.
- [24] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, no. 7, pp. 211–244, Jan. 2001.
- [25] R. 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.*, vol. 62, no. 14, pp. 3591–3603, Jul. 2014.
- [26] D. Kong and J. Chun, "A fast DOA tracking algorithm based on the extended Kalman filter," in *Proc. IEEE Nat. Aerosp. Electron. Conf.*, Oct. 2000, pp. 235–238.
- [27] R. Mahler, *Statistical Multisource-multitarget Information Fusion*, vol. 11. Norwood, MA, USA: Artech House, 2007, p. 65.
- [28] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear Bayesian estimation: From Kalman filtering to a broader horizon," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 2, pp. 401–417, Mar. 2018.
- [29] W. Wu, B. Tang, and R. Tao, "Performance bound optimization for MIMO radar direction finding with MUSIC," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8845–8858, Dec. 2023.
- [30] W. Xing, C. Zhou, and C. Wang, "Modified OMP method for multi-target parameter estimation in frequency-agile distributed MIMO radar," *J. Syst. Eng. Electron.*, vol. 33, no. 5, pp. 1089–1094, Oct. 2022.
- [31] P. B. Tuuk and S. L. Marple, "Compressed sensing radar amid noise and clutter using interference covariance information," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 887–897, Apr. 2014.
- [32] Y. Yu, A. P. Petropulu, and H. V. Poor, "Measurement matrix design for compressive sensing-based MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5338–5352, Nov. 2011.
- [33] M. Hurtado, C. H. Muravchik, and A. Nehorai, "Enhanced sparse Bayesian learning via statistical thresholding for signals in structured noise," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5430–5443, Nov. 2013.



SOHEIL SALARI (Member, IEEE) received the B.Sc. degree in electrical engineering from Shahid Bahonar University, Kerman, Iran, in 1998, and the M.Sc. and Ph.D. degrees in electrical engineering from the K. N. Toosi University of Technology, Tehran, Iran, in 2002 and 2007, respectively. He has served several teaching/research appointments with the related departments in K. N. Toosi University of Technology, University of Ontario Institute of Technology (UOIT), University of

Toronto, and Queen's University. He also collaborated in several industrial projects. His research interests include wireless communications, compressed sensing, digital signal processing, machine learning, and optimization theory.



FRANCOIS CHAN (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from McGill University, Montreal, Canada, and the M.Sc.A. and Ph.D. degrees in electrical engineering from the Ecole Polytechnique de Montréal, Canada. He is currently a Professor with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON, Canada. He was a Visiting Researcher with the University of California, Irvine, from 2002 to 2005. His research interests include digital communications, wireless communications, and digital signal processing.