Digital Signal Processing 113 (2021) 103028

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Maximum-likelihood direction of arrival estimation under intermittent jamming

Şafak Bilgi Akdemir *, Çağatay Candan

Department of Electrical and Electronics Engineering, Middle East Technical University (METU), 06800, Ankara, Turkey

ARTICLE INFO

Article history:
Available online 19 March 2021

Keywords:
Direction of arrival estimation
Jamming
Intermittent interference
Expectation-maximization algorithm
Cramer-Rao bound
Modified Cramer-Rao bound

ABSTRACT

Multiple-snapshot maximum-likelihood (ML) direction of arrival (DOA) estimation problem is studied for the intermittent jamming scenario. The intermittent jamming modality is based on the assumption that only a subset of the collected snapshots are contaminated by the jammer while the others are jammer-free; but the receiver does not know which is which. This type of jamming is frequently encountered in practice either inadvertently, say due to the sporadic activity of a non-hostile system; or intentionally, say due to the activity of an adversary sweeping the operational bandwidth of the receiver. Exact maximum likelihood solution for the problem is analytically intractable and an expectation maximization (EM) method based solution is developed for coherent and non-coherent signal models. Coherent signal model assumes that the phase difference between the coefficients of two consecutive snapshots are known a-priori which is an assumption compatible with the Swerling-1/3 target models in the radar signal processing literature. Non-coherent signal model does not have such an assumption and it is suitable for Swerling-2/4 targets. The suggested EM based solution is shown to yield an important estimation accuracy improvement over conventional maximum-likelihood solution which ignores the intermittency of jammer and also over the atomic norm based high resolution estimation techniques. Cramer-Rao type performance lower bounds for the problem is also provided to illustrate the efficacy of the suggested estimator.

1. Introduction

Direction of arrival (DOA) estimation problem is a fundamental problem of statistical signal processing with close connections to frequency estimation, spectral estimation and signal modeling [1]. The DOA estimation under intermittent jamming, in spite of its importance in theory and practice, has attracted limited attention in open literature. The intermittent jamming model assumes that each one of the collected snapshot vectors is jammer corrupted with some probability. Stated differently, a subset of collected snapshots are contaminated by jammer and the others are jammer-free; but the receiver does not know which is which. The main goal of this work is to study the DOA estimation problem under intermittent jamming, that is to derive the performance lower bounds for this setting; to develop a maximum likelihood DOA estimation algorithm and to evaluate its performance by comparisons with the conventional maximum likelihood estimator which ignores the intermittent behavior and also with the sparse signal representation based high resolution estimators.

Direction of arrival (DOA) literature can be organized into two categories [2]. The methods in the first category are called classical methods, namely amplitude comparison [3], time-difference of arrival, interferometer [4] and mono-pulse [5] methods. These methods generate a single DOA estimate. The methods in the second category, such as MUSIC [6], ESPRIT [7] and their extensions [8–14], generate multiple DOA estimates for multiple targets. In fact, methods of both categories are reduced complexity versions of the maximum likelihood estimators under Gaussian noise/interference model for single and multiple targets [1,15–18]. Expectation Maximization method (EM) is another approach for the complexity reduction of multiple DOA estimation problem [19,20]. EM introduces some latent random variables to the problem setting and aims to maximize the complete-data likelihood instead of the marginal distribution likelihood [21]. In DOA literature, EM is applied to generate DOA estimates for the multiple superposed signals by defining each signal term in super-

* Corresponding author.
E-mail addresses: safak.bilgi@gmail.com (Ş. Bilgi Akdemir), ccandan@metu.edu.tr (Ç. Candan).

https://doi.org/10.1016/j.dsp.2021.103028
1051-2004/

position as a latent signal. A more recent application of EM is its utilization in the non-uniform array DOA estimation problem where a hypothetical uniform array output is chosen as the complete data [22].

In the last two decades, several high resolution techniques have been suggested for the DOA estimation in the context of sparse signal representation [23–27]. Different from the subspace-based high resolution techniques, such as MUSIC and ESPRIT; recently developed methods are based on convex optimization, such as ℓ_1-SVD [23], convex-regularized optimization [28] and sparse Bayesian estimation [29–33]. For instance, an approach exploiting the properties of covariance matrix under the sparsity constraint in the application of sparse Bayesian learning methods is given in [34]. This approach can be interpreted as establishing a connection between the subspace methods and sparse Bayesian learning methods. Solutions exploiting sparsity in Bayesian learning have also been developed for the coherent multipath DOA estimation problem in [35,36]. In [36], a deterministic method which uses alternating convex search for DOA estimation is formulated. The majority of these methods operate with a fixed grid for the unknown target angular location and several methods have been suggested to improve the performance due to the grid mismatch with the true target location, the off-grid target problem [29,32,37–40]. For example, [32] proposes two off-grid DOA estimation methods, one for narrowband signals and one for wideband signals, which are the extended versions of sparse Bayesian learning based relevance vector machine (SBLRVM) algorithm [41]. For the solution of off-grid target problem, the method of Candés and Fernandez-Granda [38] stands out with its gridless, i.e. discretization-free, framework that utilizes the total variation norm which is the continuous analog of ℓ_1 norm. This work has been extended by the atomic norm formulation induced by the complex exponentials in [42–44]. More recently, the atomic norm formulation for DOA estimation has been generalized to multiple snapshots and the problem is renamed as joint sparse frequency recovery problem [45,46].

Atomic norm based estimators is one of most important recent breakthroughs in frequency estimation [47]. These estimators use the classical relation between the auto-correlation sequence and power/energy spectral density in the problem definition and optimize over the auto-correlation sequence [42–44]. The change of domains from the continuum of frequency domain (energy spectrum) to the discrete-time sequences (auto-correlation sequence) allows a feasible optimization problem with finitely many unknowns. Once the optimization over the auto-correlation sequence (also called the dual polynomial) is completed, the maxima locations in the Fourier spectrum of the optimized sequence yield the spatial/temporal frequency estimates [43,44]. The initial formulation of the atomic norm based estimators requires semi-definite programming (SDP) solvers with linear matrix inequalities (LMI) constraints which are not efficient to implement for large dimensional problems. In [44,46], a solution based on alternating direction method of multipliers (ADMM) method have been suggested for computational efficiency. For more information on atomic norm based estimators, readers are invited to consult [47].

In addition to the theoretical aspects of the problem, the array imperfections such as mutual coupling, gain/phase uncertainty, sensor location errors can cause performance degradations in practical systems. Due to their impact on DOA estimation accuracy, self calibration methods have also been studied [48–51]. [52] approaches this problem by sparse Bayesian perspective in a unified framework for both array calibration and DOA estimation.

Another classification can be given according to the signal of interest (SOI) bandwidth. The narrowband DOA estimation methods that have been mentioned so far can be extended to the wideband DOA estimation setting by different approaches. In the noncoherent processing approach, DOAs of each SOI frequency component is estimated by narrow-band processing methods and then the estimates are combined. The coherent processing approach uses a focusing matrix to transform the covariance matrices at different frequencies to a common frequency. Coherent Signal Subspace Method (CSSM) [53] and Weighted Average of Signal Subspaces (WAVES) [54] are well known examples for the coherent approach. Test of Orthogonality of Projected Subspaces (TOPS) [55] is another wideband DOA estimation method that is based on the orthogonality relation between the signal and the noise subspaces for multiple frequency components. Wideband Covariance Matrix Sparse Representation (W-CMSR) [56] exploits a priori information on the signal spectrum and it is shown to be capable of detecting higher number of sources than the number of sensors for some array geometries.

In many applications, intentional or unintentional interference sources occupy the spectrum together with the desired signal. Radar/communication jammers are examples of intentional interferers. Several methods have been developed to detect jamming and also to improve DOA estimation accuracy under jamming. Solutions for the detection and angle tracking under interference for monopulse radars have been studied in [57,58]. Similarly, main lobe and side lobe jammer cancelers for monopulse DOA estimation is studied in [59]. In [60], several techniques for jammer and clutter cancellation in adaptive airborne radar are proposed. More recent works related to detection and estimation under jamming are given in [61,62]. In [61], an adaptive beamformer orthogonal rejection test (ABORT)-like detector that can blank coherent ECM signals is proposed. In [62], the invariant detectors have been developed in a multiple pulsed jamming environment. An unintentional jamming scenario, inter-radar interference between automotive FMCW radars, is examined in [63] and an adaptive interference suppression windowing width control is suggested for the detection of interference. Jamming is also of major concern for many modern civilian or military communication systems. For example, [64] deals with the estimation of fast fading and/or frequency-selective fading channels problem under interference in orthogonal frequency division multiplexing (OFDM) multiple-input multiple-output (MIMO) systems and it uses compressive sensing approach to circumvent the effects of unknown multiple asynchronous narrow-band interference (NBI). In [65], the detection of jammers in direct sequence spread spectrum (DSSS) wireless communication systems is studied and the statistics of jamming-free symbols is utilized to discriminate jammed packets. Similarly, the interference due to the Narrowband Internet-of-Things (NB-IoT) on existing broadband systems is also examined in [66].

We would like to note that the running assumption in virtually all DOA estimation methods is the coexistence of multiple targets or target- jammer/interference signals in *all* snapshots. Different from these studies, we assume that jamming is intermittent; hence, not all but some snapshots are contaminated. The main purpose of this study is to examine the amount of performance gain when the intermittency is taken into account in the problem formulation.

We would like to mention that the intermittent jamming problem is a highly practical problem of electronic warfare (EW) and electronic support measure (ESM) receivers [3]. As an example, one may consider a wideband ESM system receiver that collects signal samples from a frequency-hopping target whose instantaneous bandwidth overlaps with jammer's relatively narrow operational bandwidth at random times, as illustrated in Fig. 1. Another example can be a narrow band jammer sweeping a wide range of frequencies whose bandwidth overlaps with the operational bandwidth of the target from time to time. As an example of unintentional interference, one may consider a communication system that is interfered by a friendly short packet communication system turning on and off at random time instants [67].

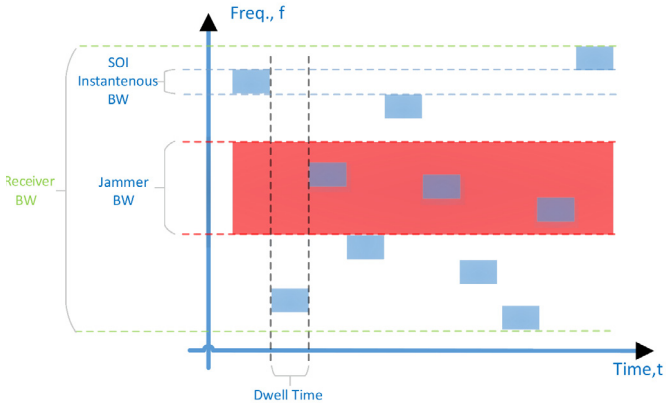


Fig. 1. An illustration for a frequency-hopping receiver whose operational bandwidth is partially jammed.

A related jamming scenario is studied by Besson et al. where the performance of ESPRIT algorithm under intermittent jamming is compared with the Cramer-Rao bound [68]. Different from this study, we do not only examine the performance of well known methods under intermittent jamming; but develop the maximum likelihood estimator for the problem.

The main contributions of this study can be listed as the development of EM-based maximum-likelihood estimator under intermittent jamming for two different target fluctuation models (coherent and non-coherent target signal models [69]), development of modified Cramer-Rao performance bound (MCRB) and a performance comparison with alternative techniques to illustrate the value of the suggested method. To retain the practicality of the suggested method and fairness in comparisons, we treat the jamming power and jamming probability as unknown nuisance parameters of the problem.

Notation: Lowercase, bold lowercase and bold uppercase symbols are used to represent scalars, vectors and matrices respectively. The superscripts $(\cdot)^T$ and $(\cdot)^H$ are used to respectively denote transpose and conjugate transpose of a vector or a matrix. We use $\boldsymbol{A}^{-1}$ to show inverse of a matrix. The notation $\mathcal{CN}(x;\mu,\sigma^2)$ denotes circularly symmetric Gaussian density of scalar random variable x, with mean μ and variance σ^2. $\mathcal{CN}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes circularly symmetric Gaussian density of vector random variable $\boldsymbol{x}$, with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The identity matrix is represented by the symbol $\boldsymbol{I}$. The symbol I_i is used to represent a binary valued random variable taking values 0 and 1.

2. Signal model and background

Consider the following set of L vectors,

$$\boldsymbol{x}_i = \underbrace{\gamma_{t,i}\boldsymbol{a}(\theta_t)}_{\text{target}} + \underbrace{I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)}_{\text{jammer}} + \underbrace{\boldsymbol{n}_i}_{\text{noise}}, \quad i=\{1,\ldots,L\}, \tag{1}$$

where $\boldsymbol{x}_i$ is an $N\times 1$ vector denoting the i'th snapshot vector, $i=\{1,\ldots,L\}$. The vector $\boldsymbol{x}_i$ is composed of three components which are signal of interest (target signal), jamming signal and noise.

The noise component $\boldsymbol{n}_i$ in (1) is a complex-valued vector whose entries are independent and identically distributed (iid) with zero mean, unit variance circularly symmetric complex Gaussian distribution, $\boldsymbol{n}_i \sim \mathcal{CN}(\boldsymbol{0},\boldsymbol{I})$. The entries of the noise vector are assumed to model the electronic noise or other unmodeled phenomena affecting the output of sensors. With this definition, the noise is assumed to be independent from sensor to sensor (elements of the vector $\boldsymbol{x}_i$) and also from snapshot to snapshot.

The jamming component in (1) is expressed as $I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)$ which is a product of three terms: The first term, $I_i=\{0,1\}$, is a binary valued random variable taking values 0 and 1. This variable indicates the presence or absence of jamming signal. The probability of jamming on the i'th snapshot is given as $P(I_i=1)=\alpha_1$. The complementary event (no jammer activity in the i'th snapshot) has the probability $P(I_i=0)=\alpha_0=1-\alpha_1$. The second term, $\gamma_{j,i}$, is a zero mean complex Gaussian random variable with variance σ^2 and represents the jammer complex amplitude for the i'th snapshot, $\gamma_{j,i}\sim\mathcal{CN}(\gamma_{j,i};0,\sigma^2)$. The third term is the vector $\boldsymbol{a}(\theta_j)$ which is the array manifold vector for the jammer located at the angular position of θ_j. For a uniform linear array, the array manifold vector becomes $\boldsymbol{a}(\theta)=[1,e^{-jkd\sin(\theta)},\cdots,e^{-jkd\sin(\theta)(N-1)}]^T$ where d is the inter-sensor spacing and $k=2\pi/\lambda$ is the wavenumber. The jammer-to-noise ratio is defined as the ratio of the jamming signal power to the noise power at a receive element. Since the amplitude of the jammer signal is Gaussian distributed with a variance of σ^2 and the noise power is assumed to be unity at each receive element, the jammer-to-noise ratio (JNR) coincides with σ^2, i.e. JNR $=\sigma^2$. Jammer power σ^2 and jamming probability α_1 are among the unknowns of the problem.

The target component in (1) is expressed as $\gamma_{t,i}\boldsymbol{a}(\theta_t)$ where $\gamma_{t,i}$ is the non-random complex-valued target signal amplitude of the i'th snapshot and θ_t is the angular position of the target. Both $\gamma_{t,i}$ and θ_t are unknowns of the problem. Target signal is assumed to be present in all snapshots. If the target amplitudes $\gamma_{t,i}$ are independent variables, that is if we have a set of L non-random unknown variables for L snapshots, we denote the target model as *non-coherent target model*. If the target amplitudes $\gamma_{t,i}$ are deterministically related, then we denote the target model as *coherent target model*. In the coherent model, one may assume that the phase difference between the phases of complex numbers $\gamma_{t,i}$ and $\gamma_{t,i+1}$ is deterministically known, i.e. pre-determined, which is the case for the snapshots collected with the pulse-Doppler radars for Swerling-1/3 targets [69]. Without any loss of generality, we assume that the deterministic phase progression at complex target amplitudes are undone, by multiplying both sides of equation (1) with the conjugate of the known phase sequence to get $\gamma_t=\gamma_{t,1}=\gamma_{t,2}=\ldots=\gamma_{t,L}$. Hence for the coherent target model, γ_t is the sole unknown that models the complex target amplitude.

In this problem, the values of the indicator variables for jammer activity I_i, $i=\{1,\ldots,L\}$ are not known to the receiver. In addition, the probability of jamming (α_1) is also an unknown of the problem. The jammer angular location θ_j is assumed to be known. This is an

acceptable assumption for reconnaissance systems, since once the jammer is detected, the jammer angular location is flagged and subsequent signal processing operations such as target detection, direction finding etc. is done with the knowledge of jammer angular location. Furthermore, in the special case of unintentional jammer, say a friendly radar system jamming DOA estimation receiver intermittently, the angular localization of jammer can be known precisely.

On conventional estimator: Conventional DOA estimators ignore the intermittency of the intercepted signals. Hence, we may assume that $I_i = 1$ for $i = \{1, \dots, L\}$ in (1), i.e. jammer is present in all snapshots, for the development of conventional estimators. For this problem, the maximum likelihood estimators for both target models are [70]:

$$\hat{\theta}_t = \arg\max_{\theta} \frac{\sum_{i=1}^{L} \left| \boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \boldsymbol{x}_i \right|^2}{\boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \boldsymbol{a}(\theta)} \quad \text{(non-coherent target model)} \tag{2}$$

$$\hat{\theta}_t = \arg\max_{\theta} \frac{\left| \boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \sum_{i=1}^{L} \boldsymbol{x}_i \right|^2}{\boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \boldsymbol{a}(\theta)} \quad \text{(coherent target model)} \tag{3}$$

where $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$ is the jammer plus noise covariance matrix. For the intermittent jamming set-up, $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ matrix becomes $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \alpha_1 \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$; since the jammer is present with probability α_1. In practice, $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ is not available to the receiver and has to be estimated from target-free auxiliary data. Estimation of the covariance matrix is also an important and difficult problem on its own; but, we assume that $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ is exactly known by the conventional estimator in this study. Availability of exact $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ extends an advantage to the conventional estimators, which is obviously not present in practice, in the performance comparisons.

3. Proposed method

The unknown parameters $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$ appearing in the observation model $\boldsymbol{x}_i = \gamma_{t,i}\boldsymbol{a}(\theta_t) + I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i, \quad i = \{1, \dots, L\}$, are non-random variables to be determined by the likelihood maximization. The probability density function of the observation vector $\boldsymbol{x}_i$ is $p(\boldsymbol{x}_i \mid \boldsymbol{\theta}) = \alpha_0 \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I}) + \alpha_1 \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j))$.

Remembering the fact that the snapshot $\boldsymbol{x}_i$ is independent from $\boldsymbol{x}_m$, $i \neq m$; the log likelihood of L snapshots becomes

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L; \boldsymbol{\theta}) = \sum_{i=1}^{L} \Bigg[\ln \Bigg(& \alpha_0 \frac{1}{\pi^N} \exp\left(- \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right)^H \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right) \right) + \\ & \alpha_1 \frac{1}{\pi^N \left| \boldsymbol{I} + \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H \right|} \exp\left(- \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right)^H \left(\boldsymbol{I} + \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H\right)^{-1} \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right) \right) \Bigg) \Bigg]. \end{aligned} \tag{4}$$

Since the argument of the logarithm function in (4) involves summation of two terms, it is not possible to have an exact analytical solution, say in the form of (3), to the likelihood maximization problem. We use Expectation Maximization (EM) algorithm to iteratively maximize the likelihood via the latent variables [21].

We denote the set of observations and latent variables as $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L\}$ and $\boldsymbol{Y} = \{I_1, \gamma_{j,1}, I_2, \gamma_{j,2}, \cdots, I_L, \gamma_{j,L}\}$, respectively. EM methodology assumes an initial estimate for the unknown variables denoted as $\boldsymbol{\theta}^{\text{old}}$. In the expectation phase, the posterior density $p(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\theta}^{\text{old}})$, which is the density of latent variables $\boldsymbol{Y}$ given the observations $\boldsymbol{X}$ and $\boldsymbol{\theta}^{\text{old}}$, is computed and then the complete-data $\{\boldsymbol{X}, \boldsymbol{Y}\}$ log-likelihood function is ensemble averaged with respect to the posterior density to get $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) = E_{\boldsymbol{Y}|\boldsymbol{X}}\{\ln p(\boldsymbol{X}, \boldsymbol{Y})|\boldsymbol{X}, \boldsymbol{\theta}^{\text{old}}\}$. In the maximization phase, $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ is maximized with respect to $\boldsymbol{\theta}$ to obtain the updated non-random parameter estimates, say $\boldsymbol{\theta}^{\text{new}}$, to be utilized in the next iteration. In this study, the random variables associated with the jamming activity, which are $\{I_i\}_{i=1}^{L}$ and $\{\gamma_{j,i}\}_{i=1}^{L}$, are chosen as the latent variables, that is $\boldsymbol{Y} = \{I_1, \gamma_{j,1}, I_2, \gamma_{j,2}, \cdots, I_L, \gamma_{j,L}\}$. For the sake of notational simplicity, we also denote the latent variable pair for the i'th snapshot as $\boldsymbol{y}_i = \{I_i, \gamma_{j,i}\}$. Hence, we have $\boldsymbol{Y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_L\}$ and $\{\boldsymbol{X}, \boldsymbol{Y}\} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_L\}$ for the latent variables and complete-data set, respectively.

Complete-data log-likelihood: The joint pdf for the i'th snapshot variables $\{\boldsymbol{x}_i, \boldsymbol{y}_i\} = \{\boldsymbol{x}_i, I_i, \gamma_{j,i}\}$ can be written as

$$\begin{aligned} p(\boldsymbol{x}_i, \boldsymbol{y}_i \mid \boldsymbol{\theta}) &= p(\boldsymbol{x}_i, I_i, \gamma_{j,i} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_i \mid I_i, \gamma_{j,i}, \boldsymbol{\theta}) p(I_i \mid \boldsymbol{\theta}) p(\gamma_{j,i} \mid \boldsymbol{\theta}) \\ &= \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t) + I_i\gamma_{j,i}\boldsymbol{a}(\theta_j), \boldsymbol{I}) \alpha_1^{I_i} \alpha_0^{1-I_i} \mathcal{CN}(\gamma_{j,i}; 0, \sigma^2). \end{aligned} \tag{5}$$

Hence, the joint pdf of $\{\boldsymbol{X}, \boldsymbol{Y}\}$ becomes

$$p(\boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{\theta}) = \prod_{i=1}^{L} \left[\frac{1}{\pi^N} \exp\left(- \left\| \boldsymbol{x}_i - \gamma_{t,i} a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) \right\|^2 \right) \alpha_1^{I_i} \alpha_0^{1-I_i} \frac{1}{\pi\sigma^2} \exp\left(- \frac{|\gamma_{j,i}|^2}{\sigma^2} \right) \right]. \tag{6}$$

The logarithm of (6), so-called complete-data log-likelihood, is then

$$\begin{aligned} \ln p(\boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{\theta}) = & -(N+1)L\ln(\pi) - \sum_{i=1}^{L} \left\| \boldsymbol{x}_i - \gamma_{t,i} a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) \right\|^2 \\ & - L\ln(\sigma^2) - \sum_{i=1}^{L} \frac{|\gamma_{j,i}|^2}{\sigma^2} + \ln(\alpha_1) \sum_{i=1}^{L} I_i + \ln(\alpha_0) \sum_{i=1}^{L} (1 - I_i). \end{aligned} \tag{7}$$

Posterior density calculation: For the derivation of the $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$, we need to calculate the posterior density of the latent variables given observations and the current values of unknowns, i.e., $p(\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\theta}^{\text{old}})$. With the application of Bayes' theorem, the posterior density $p(\boldsymbol{y}_i \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}})$ can be written as

$$\begin{aligned} p(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) &= p(I_i, \gamma_{j,i}, \boldsymbol{x}_i | \boldsymbol{\theta}^{\text{old}}) / p(\boldsymbol{x}_i \mid \boldsymbol{\theta}^{\text{old}}) \\ &= \frac{\left(\alpha_1^{\text{old}}\right)^{I_i} \left(\alpha_0^{\text{old}}\right)^{1-I_i} \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}) + I_i \gamma_{j,i} \boldsymbol{a}(\theta_j), \boldsymbol{I}) \mathcal{CN}(\gamma_{j,i}; 0, \sigma^{2,\text{old}})}{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,old} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right) + \alpha_0^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I}\right)}. \end{aligned} \tag{8}$$

Setting $I_i = 0$ in (8), yields the following simplified expression for $p(I_i = 0, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}})$

$$p(I_i = 0, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) = M_{0,i} \mathcal{CN}(\gamma_{j,i}; 0, \sigma^{2,\text{old}}) \tag{9}$$

where $M_{0,i}$ is given as

$$M_{0,i} = \frac{\alpha_0^{\text{old}} \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}), \boldsymbol{I})}{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right) + \alpha_0^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I}\right)} \quad . \tag{10}$$

When $I_i = 1$ is inserted in (8), the numerator of the ratio in (8) becomes $p(I_i = 1, \boldsymbol{x}_i, \gamma_{j,i} \mid \boldsymbol{\theta}^{\text{old}}) = p(I_i = 1 \mid \boldsymbol{\theta}^{\text{old}}) p(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$. Note that $p(I_i = 1 \mid \boldsymbol{\theta}^{\text{old}})$ is simply α_1^{old}. The other term $p(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ can be written, by considering the marginal distributions and the correlation of the random variables $\boldsymbol{x}_i$ and $\gamma_{j,i}$ under the assumption of active jammer, as follows:

$$p\left(\boldsymbol{x}_i, \gamma_{j,i}, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}}\right) = \mathcal{CN}\left(\begin{bmatrix} \boldsymbol{x}_i \\ \gamma_{j,i} \end{bmatrix}; \begin{bmatrix} \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}) \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j) & \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \\ \sigma^{2,\text{old}} \boldsymbol{a}^H(\theta_j) & \sigma^{2,\text{old}} \end{bmatrix}\right). \tag{11}$$

To further simplify $p(\boldsymbol{x}_i, \gamma_{j,i}, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ in (11), we express $p(\boldsymbol{x}_i, \gamma_{j,i}, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ as the product of $p(\boldsymbol{x}_i, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ and $p(\gamma_{j,i}, \mid \boldsymbol{x}_i, I_i = 1, \boldsymbol{\theta}^{\text{old}})$. The term $p(\boldsymbol{x}_i, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ can be explicitly written as

$$p(\boldsymbol{x}_i, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}}) = \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right). \tag{12}$$

The term $p(\gamma_{j,i}, \mid \boldsymbol{x}_i, I_i = 1, \boldsymbol{\theta}^{\text{old}})$ can be written as

$$p(\gamma_{j,i}, \mid \boldsymbol{x}_i, I_i = 1, \boldsymbol{\theta}^{\text{old}}) = \mathcal{CN}\left(\gamma_{j,i}; \frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}} \boldsymbol{a}^H(\theta_j) \left(\boldsymbol{x}_i - \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right)\right), \frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}}\right) \tag{13}$$

considering the relation for the posterior density calculation for jointly Gaussian distributed random variables [71].

Finally, the density $p(I_i = 1, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}})$, which is nothing but the product of $p(I_i = 1 \mid \boldsymbol{\theta}^{\text{old}}) = \alpha_1^{\text{old}}$, $p(\boldsymbol{x}_i, \mid I_i = 1, \boldsymbol{\theta}^{\text{old}})$ and $p(\gamma_{j,i}, \mid \boldsymbol{x}_i, I_i = 1, \boldsymbol{\theta}^{\text{old}})$ divided by $p(\boldsymbol{x}_i \mid \boldsymbol{\theta}^{\text{old}})$, can be calculated from (12) and (13) as

$$p(I_i = 1, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) = M_{1,i} \mathcal{CN}\left(\gamma_{j,i}; \frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}} \boldsymbol{a}^H(\theta_j) \left(\boldsymbol{x}_i - \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right)\right), \frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}}\right) \tag{14}$$

where $M_{1,i}$ is defined as

$$M_{1,i} = \frac{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right)}{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right) + \alpha_0^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I}\right)} \quad . \tag{15}$$

As a final note, we can easily verify from (9) and (14), by marginalization operation, that $p(I_i = 0 \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) = M_{0,i}$ and $p(I_i = 1 \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) = M_{1,i}$. As expected, we have $M_{0,i} + M_{1,i} = 1$.

Expectation phase: Using the fact that random variables $\{\boldsymbol{x}_i, \boldsymbol{y}_i\}$ are independent and identically distributed, the expected value of complete-data log-likelihood function can be written as

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) &= \boldsymbol{E}_{\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\theta}^{\text{old}}} \{\ln p(\boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{\theta}) \mid \boldsymbol{X}, \boldsymbol{\theta}^{\text{old}}\} \\ &= \sum_{i=1}^{L} E_{\boldsymbol{y}_i \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}} \{\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i | \boldsymbol{\theta}) \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\}. \end{aligned} \tag{16}$$

The expectation operation is over the posterior density of the latent variables $\boldsymbol{y}_i = \{I_i, \gamma_{j,i}\}$, given in (9) and (14):

$$E_{\boldsymbol{y}_i \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}} \{\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i | \boldsymbol{\theta})\} = \sum_{I_i=0}^{1} \int_{\gamma_{j,i}} \ln p(\boldsymbol{x}_i, \boldsymbol{y}_i | \boldsymbol{\theta}) p\left(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\right) d\gamma_{j,i}. \tag{17}$$

The expression for the argument of the expectation operation, $\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i|\boldsymbol{\theta})$, that is the complete log-likelihood function for the i'th snapshot can be written from (7) as

$$\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i|\boldsymbol{\theta}) \stackrel{c}{=} -\left\|\boldsymbol{x}_i - \gamma_{t,i}a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2 + \ln(\alpha_1)I_i + \ln(\alpha_0)(1-I_i) - \ln(\sigma^2) - \frac{|\gamma_{j,i}|^2}{\sigma^2}$$

where $\stackrel{c}{=}$ refers the equality of both sides apart from constant terms not affecting subsequent calculations.

The evaluation of $E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i|\boldsymbol{\theta})\}$ requires the calculation of three expectations operations, $E_1 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{I_i\}$, $E_2 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i}|^2\}$ and $E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{x}_i - \gamma_{t,i}a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\}$. The first expectation is straightforward to calculate

$$E_1 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{I_i\} = p(I_i = 1 \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}) = M_{1,i} \tag{18}$$

where $M_{1,i}$ is given in (15).

The second expectation E_2 can be evaluated from (9) and (14) as

$$\begin{aligned} E_2 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i}|^2\} &= \sum_{I_i=0}^{1}\int_{\gamma_{j,i}} |\gamma_{j,i}|^2 p\left(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\right) d\gamma_{j,i} \\ &= M_{0,i}\sigma^{2,\text{old}} + M_{1,i}(P^{\text{old}} + |\bar{\gamma}_i^{\text{old}}|^2) \end{aligned} \tag{19}$$

where $\bar{\gamma}_i^{\text{old}}$ and P^{old} are the mean and variance of the Gaussian density given in (14), whose explicit expressions are

$$\bar{\gamma}_i^{\text{old}} = \left(\frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}\right)\boldsymbol{a}^H(\theta_j)\left(\boldsymbol{x}_i - \gamma_{t,i}^{\text{old}}\boldsymbol{a}(\theta_t^{\text{old}})\right), \qquad P^{\text{old}} = \frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}. \tag{20}$$

The third expectation, $E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{x}_i - \gamma_{t,i}a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\}$, can be compactly written as $E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\}$ by introducing $\boldsymbol{c}_i = \boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)$. The vector $\boldsymbol{c}_i$ can be treated as a non-random vector for the sake of this calculation and E_3 can be written as:

$$\begin{aligned} E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\} &= \sum_{I_i=0}^{1}\int_{\gamma_{j,i}} \left\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2 p\left(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\right) d\gamma_{j,i} \\ &= M_{0,i}\left\|\boldsymbol{c}_i\right\|^2 + M_{1,i}E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\}. \end{aligned} \tag{21}$$

Hence, the evaluation of E_3 in (21) reduces to the calculation of $E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\}$, which is an expectation over the complex Gaussian random variable $\gamma_{j,i}$ given the observation vector $\mathbf{x}_i$ under the assumption that jamming is active, i.e. $I_i = 1$. Using expression (14), we can express this density as $\mathcal{CN}(\gamma_{j,i}; \bar{\gamma}_i^{\text{old}}, P^{\text{old}})$ where $\bar{\gamma}_i^{\text{old}}$ and P^{old} are as stated in (20) and evaluate the expectation result as follows:

$$\begin{aligned} E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\} &= E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j) - (\gamma_{j,i} - \bar{\gamma}_i^{\text{old}})\boldsymbol{a}(\theta_j)\right\|^2\} \\ &= \left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}|^2\}\left\|\boldsymbol{a}(\theta_j)\right\|^2 \\ &\quad -2\text{real}\left[(\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j))^H E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}\}\boldsymbol{a}(\theta_j)\right] \\ &= \left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + P^{\text{old}}N. \end{aligned} \tag{22}$$

Note that, we use $E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}\} = 0$ in the second line of (22). By inserting (22) into (21), we finalize the calculation of the third expectation:

$$E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\} = M_{0,i}\left\|\boldsymbol{c}_i\right\|^2 + M_{1,i}(\left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + P^{\text{old}}N). \tag{23}$$

Once the expressions for three expectations, given by equations (18), (19) and (23), are inserted in (16), the E-phase of EM algorithm is completed:

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \stackrel{c}{=} \ln(\alpha_0)\sum_{i=1}^{L} M_{0,i}^{\text{old}} + \ln(\alpha_1)\sum_{i=1}^{L} M_{1,i}^{\text{old}} - L\ln\left(\sigma^2\right) - \frac{1}{\sigma^2}\sum_{i=1}^{L}\left[M_{0,i}^{\text{old}}\sigma^{2,\text{old}} + M_{1,i}^{\text{old}}\left(P^{\text{old}} + \left|\bar{\gamma}_i^{\text{old}}\right|^2\right)\right] \\ - \sum_{i=1}^{L} M_{0,i}^{\text{old}}\left\|\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right\|^2 - \sum_{i=1}^{L} M_{1,i}^{\text{old}}\left(\left\|\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t) - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + P^{\text{old}}N\right). \end{aligned} \tag{24}$$

Maximization phase: In the maximization phase, we optimize the expected complete log-likelihood with respect to the unknown parameters $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$. We start with the optimization of σ^2 (jamming power). By taking the derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ with respect to σ^2 and equating the result to 0, the updated $\hat{\sigma}^2$ estimate becomes:

$$\hat{\sigma}^2 = \frac{1}{L}\sum_{i=1}^{L}\left[M_{0,i}^{\text{old}}\sigma^{2,\text{old}} + M_{1,i}^{\text{old}}\left(P^{\text{old}} + \left|\bar{\gamma}_i^{\text{old}}\right|^2\right)\right]. \tag{25}$$

Next, we optimize over the jamming probability α_1. The optimization with respect to $\alpha_m, m = \{0, 1\}$ is a constrained optimization problem, since $\alpha_0 + \alpha_1 = 1$. Forming the Lagrangian $\Lambda = Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right) + \lambda\left(\alpha_0 + \alpha_1 - 1\right)$ and taking its derivative with respect to $\alpha_m, m = \{0, 1\}$ leads to

$$\hat{\alpha}_0 = \frac{1}{L}\sum_{i=1}^{L} M_{0,i}^{\text{old}} \quad \text{and} \quad \hat{\alpha}_1 = \frac{1}{L}\sum_{i=1}^{L} M_{1,i}^{\text{old}}. \tag{26}$$

Next, we optimize over the target amplitude variables $\gamma_{t,i}$ $i = \{1, 2, \ldots, L\}$. We present two different target signal models, namely non-coherent and coherent target models.

Target amplitude and angle estimation for noncoherent target signal model: This model assumes that $\gamma_{t,i}$ for $i = \{1, 2, \ldots, L\}$ are L independent non-random variables. This model is valid for Swerling-2/Swerling-4 target fluctuation models in radar signal processing literature [69].

By taking derivative of $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$ in (24) with respect to $\gamma_{t,i}^*$,

$$\frac{\partial Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)}{\partial \gamma_{t,i}^*} = -N\gamma_{t,i} + \boldsymbol{a}^H(\theta_t)\boldsymbol{x}_i - M_{1,i}^{\text{old}}\boldsymbol{a}^H(\theta_t)\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j) \tag{27}$$

and equating the result to zero, the updated estimate $\hat{\gamma}_{t,i}$ can be given as

$$\hat{\gamma}_{t,i} = \frac{1}{N}\boldsymbol{a}^H(\theta_t)\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right), \quad i = \{1, 2, \ldots, L\}. \tag{28}$$

The only remaining unknown to determine is the target angular position θ_t. When the updated target amplitudes are inserted in (24), we get the compressed objective function $Q_c(\cdot, \cdot)$ as

$$Q_c(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \overset{c}{=} \frac{1}{N}\sum_{i=1}^{L}\boldsymbol{a}^H(\theta_t)\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)^H\boldsymbol{a}(\theta_t). \tag{29}$$

The angular position θ_t maximizing the compressed objective function is the updated DOA estimate for the target signal, which is

$$\hat{\theta}_t = \arg\max_{\theta}\sum_{i=1}^{L}\left|\boldsymbol{a}^H(\theta)\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)\right|^2. \quad \text{(Non-coherent target model)} \tag{30}$$

Target amplitude and angle estimation for coherent target signal model: This model assumes that the target amplitudes $\gamma_{t,i}$ for $i = \{1, 2, \ldots, L\}$ are identically the same, after a deterministic phase correction, as mentioned before. This model is valid for Swerling-1/Swerling-3 target fluctuation models in radar signal processing literature [69].

To simplify the notation, we denote the common target amplitude with γ_t, and substitute $\gamma_t = \gamma_{t,1} = \gamma_{t,2} = \ldots = \gamma_{t,L}$ in (24). After the substitution, we get,

$$Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right) \overset{c}{=} -\sum_{i=1}^{L} M_{0,i}^{\text{old}}\left\|\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t)\right\|^2 - \sum_{i=1}^{L} M_{1,i}^{\text{old}}\left\|\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t) - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2. \tag{31}$$

By taking the derivative of $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$ with respect to γ_t^*

$$\frac{\partial Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)}{\partial \gamma_t^*} = -NL\gamma_t + \boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right) \tag{32}$$

and equating the result to zero, the updated target amplitude $\hat{\gamma}_t$ can be given as

$$\hat{\gamma}_t = \frac{1}{NL}\boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right). \tag{33}$$

Inserting the optimized $\hat{\gamma}_t$ value in (31), we get the compressed objective function in the coherent case as

$$Q_c(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \overset{c}{=} \hat{\gamma}_t\sum_{i=1}^{L} M_{0,i}^{\text{old}}\boldsymbol{x}_i^H\boldsymbol{a}(\theta_t) + \hat{\gamma}_t\sum_{i=1}^{L} M_{1,i}^{\text{old}}\left(\boldsymbol{x}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)^H\boldsymbol{a}(\theta_t). \tag{34}$$

Maximizing the compressed objective function over the unknown target angle yields the updated DOA estimate for the coherent target signal model as:

$$\hat{\theta}_t = \arg\max_{\theta}\left|\boldsymbol{a}^H(\theta)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)\right|^2. \quad \text{(Coherent target model)} \tag{35}$$

This completes the derivation of M-step of EM algorithm. The steps of the proposed estimator for both target models are given in Algorithm 1 listing.

Algorithm 1 ML DOA estimation under intermittent jamming.

Initialization: Initialize $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$
Expectation phase:
  compute $M_{0,i}$ from (10)
  compute $M_{1,i}$ from (15)
  compute P^{old} and $\bar{\gamma}_i^{\text{old}}$ from (20)
Maximization phase:
  compute $\hat{\sigma}^2$ from (25)
  compute $\hat{\alpha}_0$ and $\hat{\alpha}_1$ from (26)
  if Signal model = non-coherent **then**
    compute $\hat{\gamma}_{t,i}$, $i = \{1, 2, \ldots, L\}$ from (28)
    compute $\hat{\theta}_t$ from (30)
  end if
  if Signal model = coherent **then**
    compute $\hat{\gamma}_t$ from (33)
    compute $\hat{\theta}_t$ from (35)
  end if
Termination check:
  if the convergence criteria is satisfied **then**
    return $\hat{\sigma}^2$, $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}_{t,i}(\hat{\gamma}_t)$ and $\hat{\theta}_t$
  else
    Go to Expectation Phase
  end if

4. Numerical comparisons

We present three sets of Monte Carlo simulation results for performance comparisons. In all comparisons, a uniform linear array of $N = 10$ elements with an interelement spacing of $\lambda/2$ is utilized. Noise variance is taken as unity. The target signal complex amplitude is denoted with $\gamma_{t,i}$ where $\gamma_{t,i} = \left|\gamma_{t,i}\right| e^{\phi_{t,i}}$. For coherent target model, $\left|\gamma_{t,i}\right|$ and $\phi_{t,i}$ are identically the same for $i = \{1, \ldots, L\}$. For noncoherent target model, $\left|\gamma_{t,i}\right|$ is held constant for all snapshots whereas phase is independently sampled from uniform distribution at each snapshot. For both target models SNR is defined as $\text{SNR} = \left|\gamma_{t,i}\right|^2$. The jammer existence probability (α_1) is fixed to 0.5 in the first and the third numerical experiments and it is varied in the second experiment. The jammer-to-noise-ratio (JNR) is identical to the jammer power σ^2, since the noise variance is taken as unity.

Under the intermittent interference model, the interference plus noise covariance matrix becomes $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \alpha_1 \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$. It is assumed that the conventional method utilizes the exact $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ matrix. Hence, the conventional method has the exact knowledge of jammer power and its angular position, but is not aware that jamming is intermittent. The suggested method assumes that jamming is intermittent, but is not aware of jamming probability and jamming power. The target and jammer angular locations are taken as at 0° and 10°, respectively. It should be noted that the 3 dB beamwidth of the uniform linear array with 10 elements is approximately 11 degrees. Hence, the jammer lies at the edge of the beamwidth which makes it difficult to suppress without a significant SNR loss via conventional methods.

In all figures, we present the Cramer-Rao bound (CRB) for the cases of jammer always absent/present and Modified Cramer-Rao Bound (MCRB). The performance metric used for angle estimation is root mean square error, $\text{RMSE} = \sqrt{\frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} \left(\hat{\theta}_i - \theta_t\right)^2}$ where N_{runs} is the number of Monte Carlo runs, which is 100,000 for the first two experiments and 10,000 for the third one.

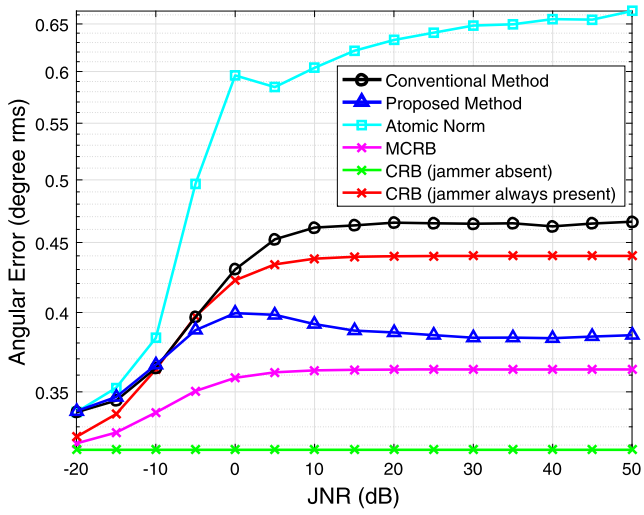
Experiment 1. Fig. 2 shows the RMS error on the target angular position estimate as JNR increases. This experiment is conducted with $L = 20$ snapshots, jamming probability of $\alpha_1 = 0.5$ and SNR = 0 dB.

The sub-figures of Fig. 2 present the results for coherent and non-coherent signal models. The red and green colored curves indicates the CRB when the jammer is always present ($\alpha_1 = 1$) and absent ($\alpha_1 = 0$), respectively. The MCRB for $\alpha_1 = 0.5$ is shown with the purple line.

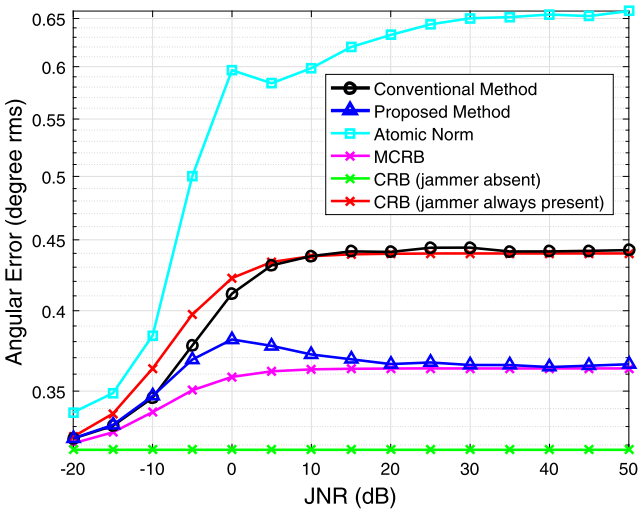
It can be noted from Fig. 2 that the proposed and conventional method perform almost identically for sufficiently small JNR values (weak jamming signal). As expected, this operational regime is dominated by the noise, not by the jammer activity; hence the performance of the proposed method, which models the intermittent nature of the jammer, is identical to the conventional one. As JNR increases, the performance of the conventional estimator, i.e. the maximum likelihood estimator with the assumption of $\alpha_1 = 0.5$, degrades rapidly after a threshold value. The threshold is around JNR $= -5$ dB in Fig. 2. The proposed method yields a better performance over the conventional estimator above this JNR threshold.

Atomic norm based technique shown in Fig. 2 estimates both jammer and target locations via convex optimization by processing L snapshots [46]. Atomic norm method, like other high resolution methods such as MUSIC or ESPRIT, does not utilize the statistical information of the observations and have no dependence on number of targets in this formulation. We observe from Fig. 2 that atomic norm based technique yields a good performance for the estimation of multiple DOA's, but its performance is far more worse than conventional and suggested estimator which utilize the specifics of the problem. The lack of sensitivity to the problem specifics of atomic norm estimator is also reflected by almost identical performance for both target models, while other estimators utilize the target signal model to further improve the performance.

Fig. 3 shows the accuracy of other parameters estimated in Experiment 1. Sub-figures (a)-(b) of Fig. 3 show the estimate for the jammer existence probability. As expected, as JNR increases, the jammer existence probability converges to the true value of 0.5. The mean and standard deviation for α_1 estimates is almost identical for both target models. Sub-figures (c)-(d) of Fig. 3 shows the estimate for the jammer power. The true jammer power is also indicated by the dashed red-line in Figs. 3(c) and 3(d). The standard deviation of the jammer power estimate is also indicated by the error bars.

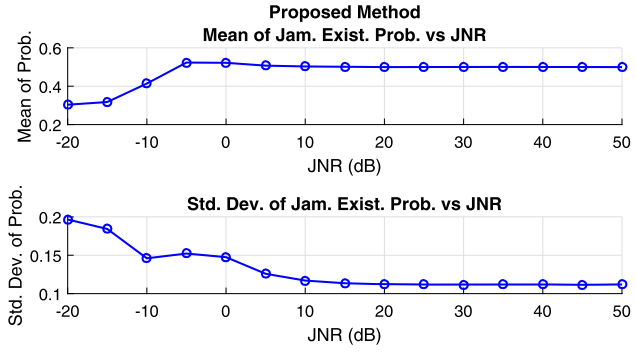


(a) Non-coherent case

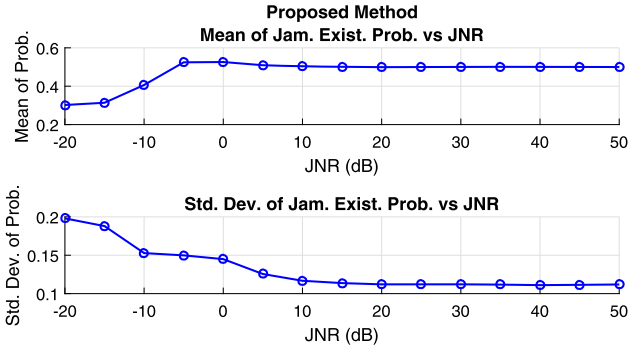


(b) Coherent case

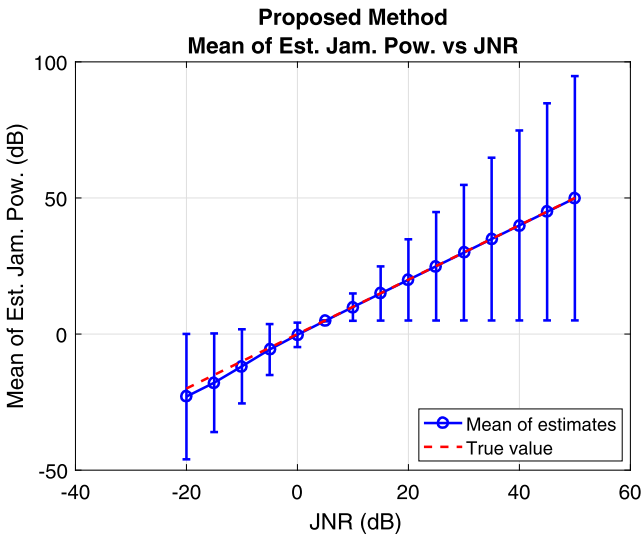
Fig. 2. Experiment 1 - angle estimation accuracy ($\theta_t = 0$, $\theta_j = 10°$, $L = 20$, $\alpha_1 = 0.5$, $N = 10$, SNR $= 0$ dB). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)



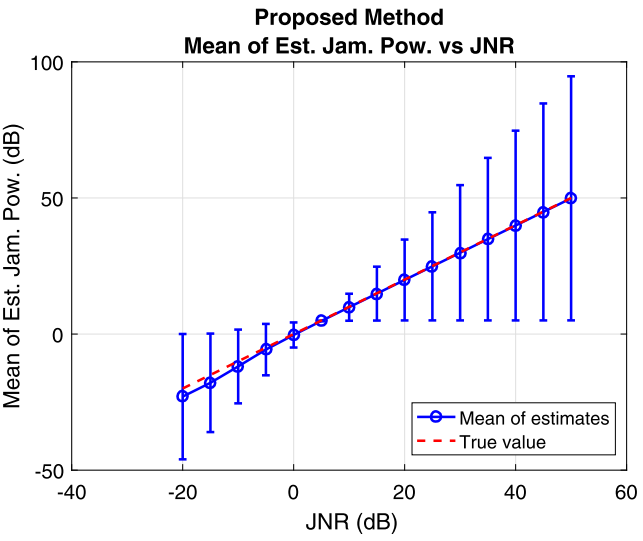
(a) Non-coherent case



(b) Coherent case

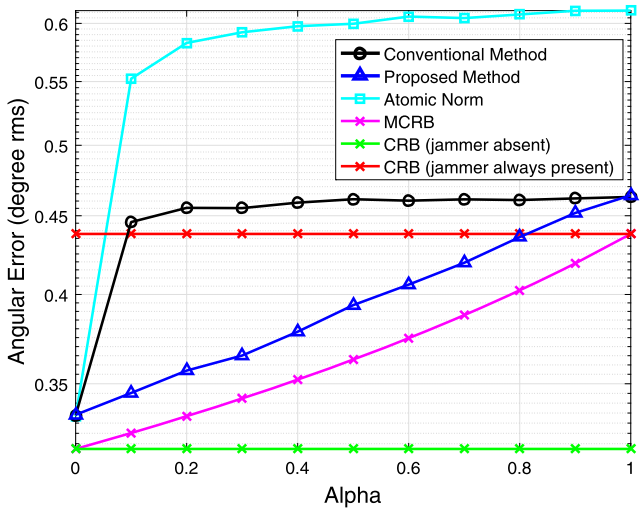


(c) Non-coherent case

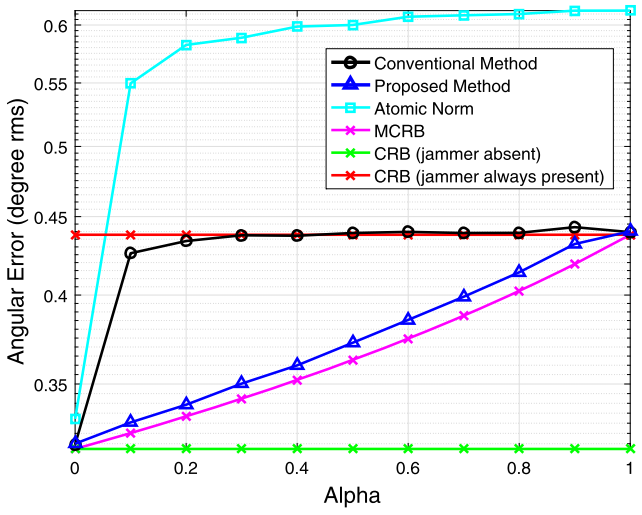


(d) Coherent case

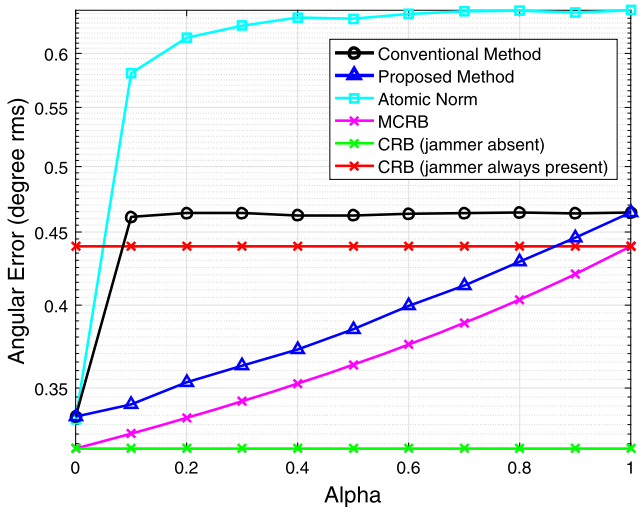
Fig. 3. Experiment 1 - sub-figures (a)-(b): Mean and standard deviation of jammer existence probability estimate, sub-figures (c)-(d): Mean value of jammer power estimates (error bars indicate one standard deviation of estimates from the mean value).



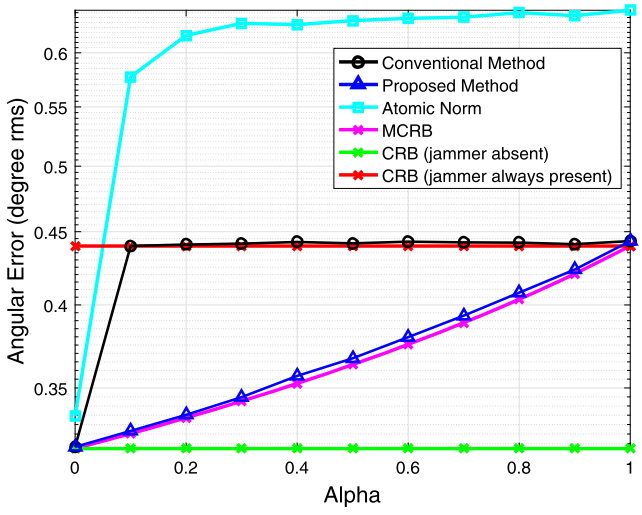
(a) JNR = 10 dB (non-coherent case)



(b) JNR = 10 dB (coherent case)



(c) JNR = 20 dB (non-coherent case)



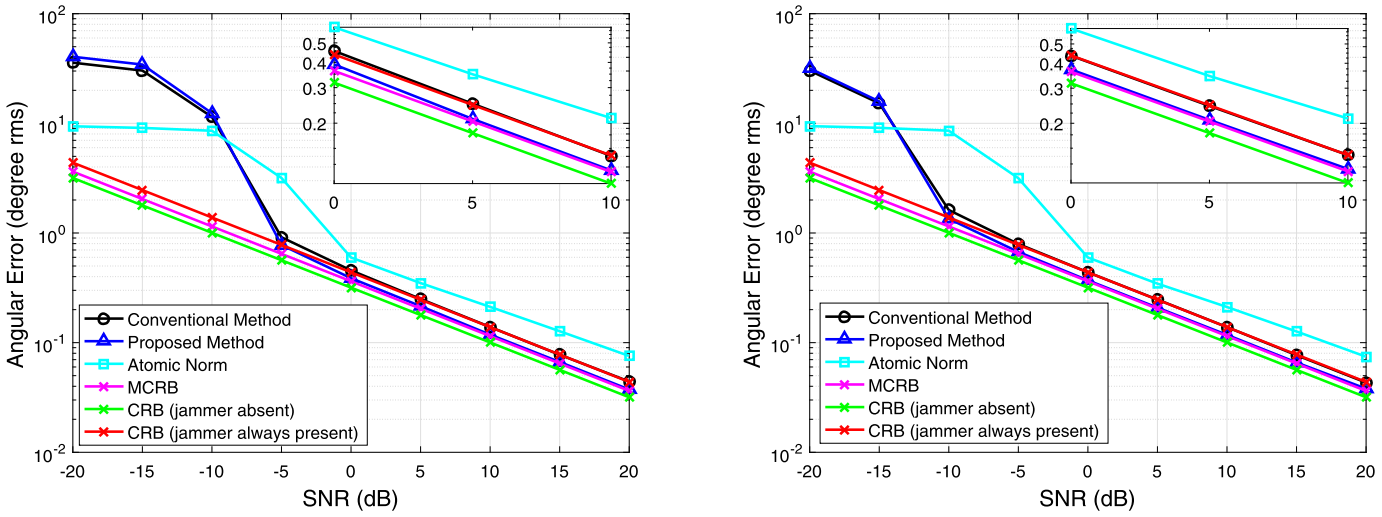
(d) JNR = 20 dB (coherent case)

Fig. 4. Experiment 2 - angle estimation accuracy ($\theta_t = 0$, $\theta_j = 10°$, $L = 20$, $N = 10$, SNR = 0 dB).

Experiment 2. JNR value is fixed in this experiment, but the jammer existence probability α_1 is varied. The other experiment parameters are identical to the ones in Experiment 1. JNR is fixed to two values, 10 and 20 dB, to illustrate the performance under medium and heavy jamming. It can be observed from Fig. 4 that conventional and atomic norm estimator suffer from a significant performance loss even at $\alpha_1 = 0.1$. This is essentially due to the ignorance of signal intermittency in the problem formulation. The performance of the suggested method is identical to the conventional method for the extreme cases of jammer always absent ($\alpha_1 = 0$) and always present ($\alpha_1 = 1$) and tracks the MCRB lower bound for all α_1 values. It is interesting to note that the case of coherent signal model tracks the MCRB very closely in the heavy jamming case. We would like to underline that the suggested method is unaware of true α_1 value and estimates this parameter in the processing chain. In spite of this, the performance of the suggested method is in close vicinity of the performance bounds for all α_1 values. The estimation accuracy of other parameter estimates ($\hat{\sigma}^2$ and $\hat{\alpha}_1$) is similar to the results in Fig. 3 and is not provided.

The success of the suggested method can be attributed to the successful "soft classification" of snapshots in terms of jammer contamination level. Another observed feature of the suggested method is the graceful degradation of estimation accuracy with increasing jamming probability. As can be most easily seen from Fig. 4(d), the suggested method works as if jammer is absent for small jammer existence probability values, while the conventional estimator suffers from significant performance losses and works as if jammer is always present.

Experiment 3. This experiment shows the accuracy of the angle estimate for a fixed JNR as SNR changes. The experiment conditions are identical to the earlier experiments except JNR is fixed to 10 dB and jammer existence probability is set to $\alpha_1 = 0.5$. Results in Fig. 5 are consistent with earlier experiments. In this experiment, atomic norm method can not estimate the target angular location at low SNR



(a) Non-coherent case (b) Coherent case

Fig. 5. Experiment 3 - angle estimation accuracy ($\theta_t = 0$, $\theta_j = 10°$, $L = 20$, $\alpha_1 = 0.5$, $N = 10$, JNR $= 10$ dB).

and presents a single estimate (instead of two) which is the estimate for the jammer location whose true value is 10 degrees. Suggested method performs close to MCRB in the asymptotic region (high SNR region of Fig. 5) and outperforms other alternatives.

5. Summary and conclusions

The main goal of this study is to examine the direction of arrival estimation problem under intermittent jamming. Intermittent jamming is a frequently encountered problem especially in electronic support systems [67]. In spite of its importance and practicality, DOA estimation under this jamming modality has received little attention in open literature.

In this study, we present the maximum likelihood angle estimator specific to the intermittent jamming modality for two different target fluctuation models. Non-coherent signal model is suitable for sensing or radar signal processing applications; while the coherent model is applicable to both communications and radar signal processing applications. In the performance comparisons, the suggested scheme is compared with the conventional estimator, atomic norm based estimator and performance bounds. The conventional estimator is assumed to have perfect knowledge of jammer parameters (jammer angular position and power), while the suggested scheme is unaware of these parameters and estimates them through the maximum likelihood framework. In spite of the additional information provided to the conventional estimator, the suggested scheme is shown to perform better in all scenarios. We note that the performance of the conventional estimator suddenly degrades with increasing α_1 (jammer existence probability) for strong jammers. This is not the case for the suggested estimator. (This is most clearly observed from Fig. 4(d) where the conventional estimator performance rapidly degrades towards the jammer-always-present CRB with increasing α_1; while the suggested estimator presents a graceful degradation closely tracking the MCRB performance bound. At the operating point of $\alpha_1 = 0.1$ in Fig. 4(d), the conventional estimator operates as if jammer is always present, while the suggested estimator operates as if jammer is always absent or with a minor degradation.) The atomic norm estimator and other high resolution estimators operate independent of the number of target signals and do not explicitly take into account the statistical description of the observation model such as signal intermittency. This leads to poorer performance for these generic estimators in comparison to the maximum likelihood estimators developed specifically for the problem. As a final note on the study, we can state that the suggested DOA estimation method for the intermittent jamming scenario significantly benefits from the inclusion of problem specifics into estimator design via the EM-formulation and this leads to significant estimation accuracy improvements over conventional or off-the-shelf estimators.

CRediT authorship contribution statement

Şafak Bilgi Akdemir: Conceptualization, Formal analysis, Software, Visualization, Writing – original draft. **Çağatay Candan:** Conceptualization, Formal analysis, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Cramer-Rao lower bound

The derivations for the Cramer-Rao Lower Bound (CRLB) and modified CRLB are presented. We remind that the Fisher information matrix (FIM) for L independent snapshot is the sum of individual FIM for each snapshot. For a Gaussian vector, the entries of FIM matrix can be written as [71],

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{k,l} = \mathrm{tr}\left[\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}_{\boldsymbol{x}}(\boldsymbol{\theta})}{\partial\theta_k}\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}_{\boldsymbol{x}}(\boldsymbol{\theta})}{\partial\theta_l}\right] + 2\,\mathrm{Re}\left[\frac{\partial \boldsymbol{\mu}^H(\boldsymbol{\theta})}{\partial\theta_k}\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_l}\right]$$

where θ_k and θ_l are k'th and l'th unknown parameters to be estimated. The set of unknown parameters change with the target signal model.

Non-coherent target signal model: This model includes $2L+1$ parameters to be estimated:

$$\boldsymbol{\theta} = \left[\begin{array}{cccccc}\theta_t & \mathrm{Re}\{\gamma_{t,1}\} & \cdots & \mathrm{Re}\{\gamma_{t,L}\} & \mathrm{Im}\{\gamma_{t,1}\} & \cdots \quad \mathrm{Im}\{\gamma_{t,L}\}\end{array}\right].$$

For this model, the non-zero entries of FIM matrix can be given as:

$$\begin{aligned}
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,1} &= \sum_{i=1}^{L} 2\,\mathrm{Re}\left[|\gamma_{t,i}|^2 (2\pi/\lambda)^2 \cos^2(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,l} &= \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[-j\gamma_{t,i}^*(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right], \quad l=\{2,\ldots,L+1\},\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,l} &= \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[\gamma_{t,i}^*(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right], \quad l=\{L+2,\ldots,2L+1\},\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{l,l} &= 2\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right], \quad l=\{2,\ldots,2L+1\}
\end{aligned}$$

where $\boldsymbol{D}$ is the diagonal matrix with the diagonal entries $[0\ d\ \cdots\ (N-1)d]$.

Coherent target signal model: This model includes 3 parameters to be estimated:

$$\boldsymbol{\theta} = \left[\begin{array}{ccc}\theta_t & \mathrm{Re}\{\gamma_t\} & \mathrm{Im}\{\gamma_t\}\end{array}\right].$$

The elements of FIM can be written as:

$$\begin{aligned}
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,1} &= 2L\,\mathrm{Re}\left[|\gamma_t|^2 (2\pi/\lambda)^2\cos^2(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,2} &= 2\,\mathrm{Re}\left[-j\gamma_t^*(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{1,3} &= 2L\,\mathrm{Re}\left[\gamma_t^*(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{2,1} &= 2L\,\mathrm{Re}\left[j\gamma_t(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{2,2} &= 2\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{2,3} &= 0,\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{3,1} &= 2\,\mathrm{Re}\left[\gamma_t(2\pi/\lambda)\cos(\theta)\,\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{3,2} &= 0,\\
\left[\boldsymbol{I}_{I_i}(\boldsymbol{\theta})\right]_{3,3} &= 2L\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right].
\end{aligned}$$

For both target models, CRLB on the DOA estimation error is (1,1) element of the inverse of the FIM.

Appendix B. Modified Cramer Rao bound

As defined in [72] and [73], the (i,j)th entry of the modified FIM $\boldsymbol{I}_{MCRB}(\boldsymbol{\theta})$ is given by

$$[\boldsymbol{I}_{MCRB}(\boldsymbol{\theta})]_{i,j} = -E_{\boldsymbol{x},\boldsymbol{\beta}}\left\{\frac{\partial^2 \ln p(\boldsymbol{x}\mid\boldsymbol{\beta};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\}$$

where the expectation is over both the observation vector $\mathbf{x}$ and the nuisance parameters $\boldsymbol{\beta}$.

In the examined problem, the indicator variable for the jammer presence I_i, $i=\{1,\cdots,L\}$ are the random nuisance parameters. With this definition, modified FIM can be written as

$$[\boldsymbol{I}_{MCRB}(\boldsymbol{\theta})]_{i,j} = \sum_{i=1}^{L} E_{\boldsymbol{x}_i,I_i}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x}_i\mid I_i;\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\} = \sum_{i=1}^{L} E_{I_i}\left\{E_{\boldsymbol{x}_i|I_i}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x}_i\mid I_i;\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\}\right\}$$

leading to

$$\boldsymbol{I}_{MCRB}(\boldsymbol{\theta}) = \alpha_0\boldsymbol{I}_{I_i=0}(\boldsymbol{\theta}) + \alpha_1\boldsymbol{I}_{I_i=1}(\boldsymbol{\theta}) \tag{B.1}$$

where $\boldsymbol{I}_{I_i=0}(\boldsymbol{\theta})$ and $\boldsymbol{I}_{I_i=1}(\boldsymbol{\theta})$ are previously given FIM for the cases of jammer absent and present, respectively.

References

[1] P. Stoica, R. Mosses, Spectral Analysis of Signals, 1st edition, Prentice Hall, Upper Saddle River, NJ, USA, 2012.
[2] E. Tuncer, B. Friedlander, Classical and Modern Direction-of-Arrival Estimation, 1st edition, Academic Press, 30 Corporate Drive, MA, USA, 2009.
[3] R.A. Poisel, Electronic Warfare Target Location Methods, 2nd edition, Artech House, Canton Street, MA, USA, 2012.
[4] E.V. Stansfield, Accuracy of an interferometer in noise, IEE Proc. Radar Sonar Navig. 143 (4) (1996) 217–226, https://doi.org/10.1049/ip-rsn:19960356.
[5] A. De Martino, Introduction to Modern EW Systems, 1st edition, Artech House, Canton Street, MA, USA, 2012.
[6] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas Propag. 34 (3) (1986) 276–280, https://doi.org/10.1109/TAP.1986.1143830.
[7] R. Roy, T. Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech Signal Process. 37 (7) (1989) 984–995, https://doi.org/10.1109/29.32276.
[8] A. Barabell, Improving the resolution performance of eigenstructure-based direction-finding algorithms, Proc. IEEE Int. Conf. Acoust. Speech Signal Proc. 8 (1983) 336–339, https://doi.org/10.1109/ICASSP.1983.1172124.
[9] B.D. Rao, K.V.S. Hari, Performance analysis of Root-MUSIC, IEEE Trans. Acoust. Speech Signal Process. 37 (12) (1989) 1939–1949, https://doi.org/10.1109/29.45540.
[10] P. Stoica, A. Nehorai MUSIC, maximum likelihood, Cramer-Rao bound, IEEE Trans. Acoust. Speech Signal Process. 37 (5) (1989) 720–741, https://doi.org/10.1109/29.17564.
[11] P. Stoica, A. Nehorai, MUSIC, maximum likelihood, and Cramer-Rao bound: further results and comparisons, IEEE Trans. Acoust. Speech Signal Process. 38 (12) (1990) 2140–2150, https://doi.org/10.1109/29.61541.
[12] P. Stoica, A. Nehorai, Performance comparison of subspace rotation and MUSIC methods for direction estimation, IEEE Trans. Signal Process. 39 (2) (1991) 446–453, https://doi.org/10.1109/78.80828.
[13] B. Porat, B. Friedlander, Analysis of the asymptotic relative efficiency of the MUSIC algorithm, IEEE Trans. Acoust. Speech Signal Process. 36 (4) (1988) 532–544, https://doi.org/10.1109/29.1557.
[14] H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach, IEEE Signal Process. Mag. 13 (4) (1996) 67–94, https://doi.org/10.1109/79.526899.
[15] Y. Bresler, A. Macovski, Exact maximum likelihood parameter estimation of superimposed exponential signals in noise, IEEE Trans. Acoust. Speech Signal Process. 34 (5) (1986) 1081–1089, https://doi.org/10.1109/TASSP.1986.1164949.
[16] P. Stoica, K.C. Sharman, Maximum likelihood methods for direction-of-arrival estimation, IEEE Trans. Acoust. Speech Signal Process. 38 (7) (1990) 1132–1143, https://doi.org/10.1109/29.57542.
[17] P. Stoica, A. Nehorai, Performance study of conditional and unconditional direction-of-arrival estimation, IEEE Trans. Acoust. Speech Signal Process. 38 (10) (1990) 1783–1795, https://doi.org/10.1109/29.60109.
[18] J. Sheinvald, M. Wax, A.J. Weiss, On maximum-likelihood localization of coherent signals, IEEE Trans. Signal Process. 44 (10) (1996) 2475–2482, https://doi.org/10.1109/78.539032.
[19] M. Feder, E. Weinstein, Parameter estimation of superimposed signals using the EM algorithm, IEEE Trans. Acoust. Speech Signal Process. 36 (4) (1988) 477–489, https://doi.org/10.1109/29.1552.
[20] M.I. Miller, D.R. Fuhrmann, Maximum-likelihood narrow-band direction finding and the EM algorithm, IEEE Trans. Acoust. Speech Signal Process. 38 (9) (1990) 1560–1577, https://doi.org/10.1109/29.60075.
[21] B.C. Levy, Principles of Signal Detection and Parameter Estimation, 1st edition, Springer Science+Business Media, LLC, Spring Street, NY, USA, 2008.
[22] J. Yang, Y. Yang, B. Lei, J. Lu, L. Yang, Nonuniform linear array DOA estimation using EM criterion, Digit. Signal Process. 86 (2019) 36–41, https://doi.org/10.1016/j.dsp.2018.12.010.
[23] D. Malioutov, M. Cetin, A.S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (8) (2005) 3010–3022, https://doi.org/10.1109/TSP.2005.850882.
[24] J. Chen, X. Huo, Theoretical results on sparse representations of multiple-measurement vectors, IEEE Trans. Signal Process. 54 (12) (2006) 4634–4643, https://doi.org/10.1109/TSP.2006.881263.
[25] M.M. Hyder, K. Mahata, Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation, IEEE Trans. Signal Process. 58 (9) (2010) 4646–4655, https://doi.org/10.1109/TSP.2010.2050477.
[26] Z. Yang, J. Li, P. Stoica, L. Xie, Sparse methods for direction-of-arrival estimation, in: R. Chellappa, S. Theodoridis (Eds.), Library in Signal Processing, vol. 7, Academic Press, Berlin, Heidelberg, 2017, pp. 509–581, Ch. 11.
[27] J. Zhang, T. Qiu, S. Luan, An efficient real-valued sparse Bayesian learning for non-circular signal's DOA estimation in the presence of impulsive noise, Digit. Signal Process. 106 (2020) 102838.
[28] Q. Liu, H.C. So, Y. Gu, Off-grid DOA estimation with nonconvex regularization via joint sparse representation, Signal Process. 140 (2017) 171–176, https://doi.org/10.1016/j.sigpro.2017.05.020.
[29] L. Hu, Z. Shi, J. Zhou, Q. Fu, Compressed sensing of complex sinusoids: an approach based on dictionary refinement, IEEE Trans. Signal Process. 60 (7) (2012) 3809–3822, https://doi.org/10.1109/TSP.2012.2193392.
[30] Z. Liu, Z. Huang, Y. Zhou, An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning, IEEE Trans. Wirel. Commun. 11 (10) (2012) 1–11, https://doi.org/10.1109/TWC.2012.090312.111912.
[31] D.P. Wipf, B.D. Rao, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem, IEEE Trans. Signal Process. 55 (7) (2007) 3704–3716, https://doi.org/10.1109/TSP.2007.894265.
[32] A. Das, T.J. Sejnowski, Narrowband and wideband off-grid direction-of-arrival estimation via sparse Bayesian learning, IEEE J. Ocean. Eng. 43 (1) (2018) 108–118, https://doi.org/10.1109/JOE.2017.2660278.
[33] M. Carlin, P. Rocca, G. Oliveri, F. Viani, A. Massa, Directions-of-arrival estimation through Bayesian compressive sensing strategies, IEEE Trans. Antennas Propag. 61 (7) (2013) 3828–3838, https://doi.org/10.1109/TAP.2013.2256093.
[34] Z. Liu, Z. Huang, Y. Zhou, Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors, IEEE Trans. Wirel. Commun. 12 (8) (2013) 1–12, https://doi.org/10.1109/TWC.2013.071113.121305.
[35] Z.-M. Liu, L. Zheng, D.-W. Feng, Z.-T. Huang, Direction-of-arrival estimation for coherent sources via sparse Bayesian learning, Int. J. Antennas Propag. 2014 (2014) 1–8, https://doi.org/10.1155/2014/959386.
[36] A. Das, Deterministic and Bayesian sparse signal processing algorithms for coherent multipath directions-of-arrival (DOAs) estimation, IEEE J. Ocean. Eng. 44 (4) (2019) 1150–1164, https://doi.org/10.1109/JOE.2018.2851119.
[37] Z. Yang, L. Xie, C. Zhang, Off-grid direction of arrival estimation using sparse Bayesian inference, IEEE Trans. Signal Process. 61 (1) (2013) 38–43, https://doi.org/10.1109/TSP.2012.2222378.
[38] E.J. Candès, C. Fernandez-Granda, Towards a mathematical theory of super-resolution, Commun. Pure Appl. Math. 67 (6) (2014) 906–956, https://doi.org/10.1002/cpa.21455.
[39] A. Das, Theoretical and experimental comparison of off-grid sparse Bayesian direction-of-arrival estimation algorithms, IEEE Access 5 (2017) 18075–18087, https://doi.org/10.1109/ACCESS.2017.2747153.
[40] P. Chen, Z. Cao, Z. Chen, C. Yu, Sparse off-grid DOA estimation method with unknown mutual coupling effect, Digit. Signal Process. 90 (2019) 1–9, https://doi.org/10.1016/j.dsp.2019.04.001.
[41] A. Das, W.S. Hodgkiss, P. Gerstoft, Coherent multipath direction-of-arrival resolution using compressed sensing, IEEE J. Ocean. Eng. 42 (2) (2017) 494–505, https://doi.org/10.1109/JOE.2016.2576198.
[42] G. Tang, B.N. Bhaskar, P. Shah, B. Recht, Compressed sensing off the grid, IEEE Trans. Inf. Theory 59 (11) (2013) 7465–7490, https://doi.org/10.1109/TIT.2013.2277451.
[43] B.N. Bhaskar, G. Tang, B. Recht, Atomic norm denoising with applications to line spectral estimation, IEEE Trans. Signal Process. 61 (23) (2013) 5987–5999, https://doi.org/10.1109/TSP.2013.2273443.
[44] G. Tang, B.N. Bhaskar, B. Recht, Near minimax line spectral estimation, IEEE Trans. Inf. Theory 61 (1) (2015) 499–512, https://doi.org/10.1109/TIT.2014.2368122.
[45] Z. Yang, L. Xie, Exact joint sparse frequency recovery via optimization methods, IEEE Trans. Signal Process. 64 (19) (2016) 5145–5157, https://doi.org/10.1109/TSP.2016.2576422.

[46] Y. Li, Y. Chi, Off-the-grid line spectrum denoising and estimation with multiple measurement vectors, IEEE Trans. Signal Process. 64 (5) (2016) 1257–1269, https://doi.org/10.1109/TSP.2015.2496294.
[47] Y. Chi, M. Ferreira Da Costa, Harnessing sparsity over the continuum: atomic norm minimization for superresolution, IEEE Signal Process. Mag. 37 (2) (2020) 39–57, https://doi.org/10.1109/MSP.2019.2962209.
[48] F. Sellone, A. Serra, A novel online mutual coupling compensation algorithm for uniform and linear arrays, IEEE Trans. Signal Process. 55 (2) (2007) 560–573, https://doi.org/10.1109/TSP.2006.885732.
[49] Z. Ye, C. Liu, On the resiliency of MUSIC direction finding against antenna sensor coupling, IEEE Trans. Antennas Propag. 56 (2) (2008) 371–380, https://doi.org/10.1109/TAP.2007.915461.
[50] Youming Li, M.H. Er, Theoretical analyses of gain and phase error calibration with optimal implementation for linear equispaced array, IEEE Trans. Signal Process. 54 (2) (2006) 712–723, https://doi.org/10.1109/TSP.2005.861892.
[51] A. Liu, G. Liao, C. Zeng, Z. Yang, Q. Xu, An eigenstructure method for estimating DOA and sensor gain-phase errors, IEEE Trans. Signal Process. 59 (12) (2011) 5944–5956, https://doi.org/10.1109/TSP.2011.2165064.
[52] Z. Liu, Y. Zhou, A unified framework and sparse bayesian perspective for direction-of-arrival estimation in the presence of array imperfections, IEEE Trans. Signal Process. 61 (15) (2013) 3786–3798, https://doi.org/10.1109/TSP.2013.2262682.
[53] E.D. di Claudio, R. Parisi, WAVES: weighted average of signal subspaces for robust wideband direction finding, IEEE Trans. Signal Process. 49 (10) (2001) 2179–2191, https://doi.org/10.1109/78.950774.
[54] H. Wang, M. Kaveh, Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources, IEEE Trans. Acoust. Speech Signal Process. 33 (4) (1985) 823–831, https://doi.org/10.1109/TASSP.1985.1164667.
[55] Yeo-Sun Yoon, L.M. Kaplan, J.H. McClellan, TOPS: new DOA estimator for wideband signals, IEEE Trans. Signal Process. 54 (6) (2006) 1977–1989, https://doi.org/10.1109/TSP.2006.872581.
[56] Z. Liu, Z. Huang, Y. Zhou, Direction-of-arrival estimation of wideband signals via covariance matrix sparse representation, IEEE Trans. Signal Process. 59 (9) (2011) 4256–4270, https://doi.org/10.1109/TSP.2011.2159214.
[57] A.D. Seifer, Monopulse-radar angle tracking in noise or noise jamming, IEEE Trans. Aerosp. Electron. Syst. 28 (3) (1992) 622–638, https://doi.org/10.1109/7.256285.
[58] W.D. Blair, G.A. Watson, T. Kirubarajan, Y. Bar-Shalom, Benchmark for radar allocation and tracking in ECM, IEEE Trans. Aerosp. Electron. Syst. 34 (4) (1998) 1097–1114, https://doi.org/10.1109/7.722694.
[59] Kai-Bor Yu, D.J. Murrow, Adaptive digital beamforming for angle estimation in jamming, IEEE Trans. Aerosp. Electron. Syst. 37 (2) (2001) 508–523, https://doi.org/10.1109/7.937465.
[60] D.J. Rabideau, Clutter and jammer multipath cancellation in airborne adaptive radar, IEEE Trans. Aerosp. Electron. Syst. 36 (2) (2000) 565–583, https://doi.org/10.1109/7.845243.
[61] F. Bandiera, A. Farina, D. Orlando, G. Ricci, Detection algorithms to discriminate between radar targets and ECM signals, IEEE Trans. Signal Process. 58 (12) (2010) 5984–5993, https://doi.org/10.1109/TSP.2010.2077283.
[62] A. De Maio, D. Orlando, Adaptive radar detection of a subspace signal embedded in subspace structured plus Gaussian interference via invariance, IEEE Trans. Signal Process. 64 (8) (2016) 2156–2167, https://doi.org/10.1109/TSP.2015.2507544.
[63] M. Umehira, T. Okuda, X. Wang, S. Takeda, H. Kuroda, An adaptive interference detection and suppression scheme using iterative processing for automotive FMCW radars, in: 2020 IEEE Radar Conference, RadarConf20, 2020, pp. 1–5.
[64] A. Gomaa, N. Al-Dhahir, A sparsity-aware approach for NBI estimation in MIMO-OFDM, IEEE Trans. Wirel. Commun. 10 (6) (2011) 1854–1862, https://doi.org/10.1109/TWC.2011.040411.101118.
[65] M. Spuhler, D. Giustiniano, V. Lenders, M. Wilhelm, J.B. Schmitt, Detection of reactive jamming in DSSS-based wireless communications, IEEE Trans. Wirel. Commun. 13 (3) (2014) 1593–1603, https://doi.org/10.1109/TWC.2013.013014.131037.
[66] S. Liu, F. Yang, J. Song, Z. Han, Block sparse bayesian learning-based NB-IoT interference elimination in LTE-advanced systems, IEEE Trans. Commun. 65 (10) (2017) 4559–4571, https://doi.org/10.1109/TCOMM.2017.2723572.
[67] L.B. VanBrunt, Applied ECM, vol. 2, 1st edition, EW Engineering, Inc., USA, 1982.
[68] O. Besson, P. Stoica, Y. Kamiya, Direction finding in the presence of an intermittent interference, IEEE Trans. Signal Process. 50 (7) (2002) 1554–1564, https://doi.org/10.1109/TSP.2002.1011196.
[69] M.A. Richards, Fundamentals of Radar Signal Processing, 1st edition, McGraw-Hill, Two Penn Plaza, NY, USA, 2005.
[70] B. Ottersten, M. Viberg, P. Stoica, A. Nehorai, Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing, in: S. Haykin, J. Litva, T.J. Shepherd (Eds.), Radar Array Processing, Springer, Berlin, Heidelberg, 1993, pp. 99–151.
[71] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, 1st edition, Prentice Hall, Upper Saddle River, NJ, USA, 1993.
[72] F. Gini, R. Reggiannini, U. Mengali, The modified Cramer-Rao bound in vector parameter estimation, IEEE Trans. Commun. 46 (1) (1998) 52–60, https://doi.org/10.1109/26.655403.
[73] F. Gini, A radar application of a modified Cramer-Rao bound: parameter estimation in non-Gaussian clutter, IEEE Trans. Signal Process. 46 (7) (1998) 1945–1953, https://doi.org/10.1109/78.700966.

Şafak Bilgi Akdemir received her B.S. and M.S.degrees all in electrical engineering from Middle East Technical University, Ankara in 2003 and 2010 respectively and is currently working toward the Ph.D. degree with the Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara. She is currently a full time senior researcher at The Scientific And Technological Research Council Of Turkey (TUBITAK). Her research interests include statistical signal processing and its applications in array and radar signal processing.

Çağatay Candan is a professor at the Department of Electrical and Electronics Engineering in Middle East Technical University, Ankara, Turkey. He received his B.S., M.S., and Ph.D. degrees, all in electrical engineering, from Middle East Technical University, Ankara, Turkey (1996), Bilkent University, Ankara, Turkey (1998) and Georgia Institute of Technology, Atlanta, USA (2004), respectively. His research interests include statistical signal processing and its applications in array signal processing, radar/sonar signal processing.