Compressive Sensing for Radar Signals

Part II: EW Receivers

Prepared by:
[1]Soheil Salari, [1]Il-Min Kim, [2]Francois Chan
[1.] Department of Electrical and Computer Engineering, Queen's University
[2.] Department of Electrical and Computer Engineering, RMC

Contractor Name and Address: Prof. Francois Chan, Royal Military College, PO Box 17000, Station Forces, 13 General Crerar Crescent, Kingston, ON, K7K 7B4
Contract Number: 2009-0302-SLA
Contract Scientific Authority: Sreeraman Rajan, Senior Defence Scientist

Contract Report
DRDC-RDDC-2015-C194
January 2015

COMPRESSIVE SENSING FOR RADAR SIGNALS: PART II: EW RECEIVERS

Prepared by [1]Soheil Salari, [1]Il-Min Kim, and [2]Francois Chan

1. Department of Electrical and Computer Engineering, Queen's University

2. Department of Electrical and Computer Engineering, RMC

Contractor name and address:

Francois Chan, RMC,

13 General Crerar Crescent, Kingston, ON, K7K 7B4

Contract number: 2009-0302-SLA

Contract Scientific Authority's name, title, and phone number:

Dr. Sreeraman Rajan, Senior Defence Scientist

613-991-4138

I. INTRODUCTION

A radar system usually consist of a transmitter (sender) and a receiver. The sender emits probing pulses of electro-magnetic waves in predetermined directions. The properties of these waves are changed when they are reflected by the targets towards the receiver. This enables the radar to detect and locate the unknown targets (threats). This type of radars are sometimes referred to as the active radars. On the other hand, passive radars which are essentially receiver-only radars have gained considerable research interest over the past decades. In passive radars, illuminators of opportunity which are often radio and television (TV) stations [1], are used for target detection.

In certain military applications, the electronic warfare (EW) receivers are preferred for immediate threat recognition and electronic surveillance purposes. In such receivers, neither reflected probing pulses nor reflected illuminator signals are needed. More specifically, EW receivers are listening to the electro-magnetic radiations of the potential threats instead of weak reflected radar (illuminator) signals. Therefore, the detection range of typical EW receivers is much higher than that of (passive) radars, while they can remain electronically silent and undetectable. Consequently, agile sensing of the modern combat environments and quick threat recognition require design and development of sophisticated EW receivers [2], [3]. The EW receivers should be able to quickly detect, locate, and identify adversary emitters as well as extract and analysis the characteristics of threat waveforms [4]. Thus, such EW receivers should carry out high amount of computations in real-time. To that end, some advanced signal processing techniques are needed. One of the possible approaches is to use the technique of compressed sensing (CS) which is an emerging method in the field of signal processing.

CS has been successfully applied to a large number of applications, such as image processing [5], wireless communications [6], direction of arrival (DOA) estimation [7], and radar detection [8]. Mathematically, CS focuses on how to reconstruct the sparse signal $\mathbf{s} \in \mathbb{R}^N$ from measurements $\mathbf{r} \in \mathbb{R}^M$ generated by following model:

$$\mathbf{r} = \mathbf{\Theta}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{\Theta}$ is the dictionary (sensing) matrix of size $M \times N$ with $M \ll N$, and $\mathbf{n}$ represents an unknown noise vector of size $M \times 1$. Since $\mathbf{\Theta}$ has much fewer rows than columns, it is non-invertible and underdetermined, rendering the CS problem ill-posed [9]. However, many suboptimal algorithms exist to solve this kind of problems. It has been shown that $\mathbf{s}$ can be accurately recovered from $\mathbf{r}$ under different criteria [10], [11]. This concept also provides an alternative paradigm to the Shannon-Nyquist sampling theorem [12], [13].

In principle, CS recovery requires perfect knowledge of sensing matrix $\mathbf{\Theta}$. According to the CS formulation,

$\boldsymbol{\Theta}$ can be represented as the multiplication of two matrixes $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$, i.e. $\boldsymbol{\Theta} \triangleq \boldsymbol{\Phi}\boldsymbol{\Psi}$, where $\boldsymbol{\Phi}$ stands for the measurement matrix that is fixed and does not depend on the signal, and $\boldsymbol{\Psi}$ denotes the basis matrix that spans the space in which the signal is sparse [14]. A random measurement matrix typically leads to a good performance in most cases [15]. In many practical scenarios, the basis matrix is determined. For example, in a multiple-input multiple-output (MIMO) radar system, the transmitted waveforms are known at each receive antenna. This information enables the receive antennas to construct the basis matrix locally [13], [15]. Unlike most of the existing work in the literature, however, our goal in this paper is to recover the sparse vector $\mathbf{s}$ from measurements $\mathbf{r}$, where perfect knowledge of the basis matrix is not available. The importance of this goal can be best put into perspective by realizing that such situations happen in many applications. Consider, for example, passive radar in which transmitters are not part of the radar system, and hence the transmitted signals are not known a priori [1]. Similarly, in the case of automatic (blind) modulation classification, perfect knowledge of the transmitted signal is not available at the receiver [16]. Adversary radar emitter identification and threat waveform recognition is another example for the case where transmitted signals are completely unknown [4], [17]. This motivates us to study the problem of blind CS in which the basis matrix $\boldsymbol{\Psi}$ is (partially) unknown.

More specifically, this paper is our effort to establish an advanced CS-based signal processing scheme for EW receivers. Since the main challenge in designing EW receivers is the lack of a priori knowledge on the characteristics of received signals, (adversary) radar waveform recognition has gained considerable research interest recently. In general, the purpose of the typical radar waveform recognition system is to identify the intercepted radar signals based on the pulse compression waveform. A block diagram of such a system is depicted in [17] to clarify its operation in practical EW receivers. Up to now, different approaches has been applied in the related literature for radar signal interception and waveform classification; see [4], [17]–[20] and the references therein. However, to the best of our knowledge, none of them employed the emerging CS-based technique to deal with the problem of radar waveform recognition.

Sparse Bayesian Learning (SBL) [21]–[27] is a favorable approach for the sparse signal recovery in the context of CS, due to its interesting attractions. For example, as stated in [24], the global minimum of the popular ℓ_1 minimization based sparse recovery algorithms is generally not the sparsest solution, while the obtained global minima through SBL approach is always the sparsest one [22]. Furthermore, compare with some classic algorithms such as the FOCUSS family [28], SBL-based methods offer much fewer local minima [24]. In the SBL framework, the sparse signal reconstruction problem is formulated from a Bayesian aspect [7], while the sparsity information is incorporated by assigning a prior distribution to the sparse vector $\mathbf{s}$ as a regularization [9]. However, none of the existing works has provided the SBL-based approach for the blind CS problem, which is the focus of this paper. To the best of our knowledge, this is the first paper that aims to

reconstruct the sparse vector $\mathbf{s}$ from measurements of the form $\mathbf{r} = \mathbf{\Phi\Psi s} + \mathbf{n}$, where $\mathbf{\Psi}$ is (partially) unknown. This is in sharp contrast to the conventional SBL assumption on availability of $\mathbf{\Psi}$ matrix [11].

In this paper, resorting to the SBL recovery approach, we derive a novel CS-based method to obtain the angle and Doppler information of potential adversary radar emitters. First, exploiting the sparsity of radar signals in various spaces, we formulate the problem of angle-Doppler estimation as that of recovery of a sparse vector using blind CS. Unlike existing literature in which basis matrix is determined, here, since the adversary radar waveforms are not known, EW receiver is not able to construct the $\mathbf{\Psi}$ matrix. Hence, we develop a SBL-based framework for the problem of adversary radar identification and threat waveform recognition that is blind, i.e. the knowledge of $\mathbf{\Psi}$ is not available. Throughout the paper, we will refer to this method as blind-SBL approach. As we shall see, the proposed scheme iteratively provides the angle-Doppler estimate as well as the radar waveform estimation, via the expectation-maximization (EM) algorithm. Although our blind-SBL approach is designed for adversary radar and waveform recognition scenario, it is also capable to accommodate any other practical scenario in which the basis matrix is (partially) unknown. Next, we incorporate a pruning mechanism between consecutive iterations of our blind-SBL method to reduce the computational complexity and improve the convergence speed, namely, pruned blind-SBL. We also analyze both the global minima and the local minima of the introduced algorithms' cost function. Our theoretical analysis confirms that the proposed framework offers desirable guarantees on the convergence and accuracy of the obtained solution. The efficiency of the proposed scheme is well demonstrated by our extensive numerical experiments.

The rest of this paper is organized as follows. In Section II, we first present the signal model and then define the problem of adversary radar identification and threat waveform recognition as that of recovery of a sparse vector using blind CS. In Section III, we reformulate the defined problem based on a SBL perspective. Next, we derive a novel blind-SBL approach to jointly estimate the angle-Doppler of the adversary radar and the threat waveform. Finally, we develop a modified version of the proposed method which is suitable for practical purposes. Section IV is devoted to the analysis of the global and local minima properties of the proposed blind-SBL frameworks' cost function. Numerical results are presented in Section V to demonstrate the efficiency of the proposed method, and conclusions are drawn in Section VI.

Notation: We make use of the following notational convention throughout this paper. Bold upper case symbols denote matrices and bold lower case symbols denote vectors. The l-th element of a vector $\mathbf{a}$ is written as $\mathbf{a}(l)$, and the (i,j)-th element of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}(i,j)$. Tr($\mathbf{A}$) represents the trace of the matrix $\mathbf{A}$, while $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$. In addition, $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of matrixes $\mathbf{A}$ and $\mathbf{B}$. Furthermore, $\|\mathbf{a}\|$ stands for the Euclidean norm of the vector $\mathbf{a}$ and $\mathrm{diag}(\mathbf{a})$ represents a diagonal matrix with the elements of the vector $\mathbf{a}$ as its diagonal entries. Also, $\mathbf{1}_N$ denote an $N \times 1$ vector of all elements being

one, while $\mathbf{I}_N$ is an $N \times N$ identity matrix. Subscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^\dagger$ stand for complex conjugate, transpose, Hermitian, and Moore-Penrose pseudoinverse, respectively. Also, we use $\mathrm{E}[\cdot]$ to denote the statistical expectation and $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to represent the circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. The pdf of the random variable S is represented as $p(s)$ and the random variables and deterministic parameters in the pdf are separated using a semicolon. Finally, $\Re e\{z\}$ represents the real part of the complex number z.

II. Signal Model for CS-Based EW Receiver

A. Received Signal Model

The purpose of an EW receiver is to detect, locate, and identify the adversary radar. Let (d_e, α) denotes the location in polar coordinates of the EW receiver, where d_e denote the distance between the EW receiver and the origin, while α stands for the azimuth angle of the EW receiver. Also, suppose that an adversary radar lies on the same plane. The adversary radar is at azimuth angle θ and moves with constant radial speed v. So, its range equals $d_r(t) = d_r(0) - vt$, where $d_r(0)$ is the distance between the adversary radar and the origin at time instant zero. Under the far field assumption [13], i.e., $d_r(t) \gg d_e$, the distance between the adversary radar and the EW receiver, $d(t)$, can be approximated as:

$$d(t) \approx d_r(t) - \eta(\theta) = d_r(0) - vt - \eta(\theta), \tag{2}$$

where $\eta(\theta) = d_e cos(\theta - \alpha)$.

Similar to [17] and [4], the EW receiver presented in this paper is designed to identify pulse compression radar waveforms. In this setting, the adversary radar employs a continuous-time phase-modulated prototype pulse $c(t)$ of duration τ. Such a pulse can be modeled as a collection of L contiguous subpulses $c_l(t)$ of duration $\tau_c = \tau/L$ each with the same frequency but a (possibly) different phase, i.e.,

$$c(t) = \begin{cases} \sum_{l=0}^{L-1} c_l(t - l\tau_c), & 0 \le t \le \tau \\ 0, & \text{elsewhere.} \end{cases} \tag{3}$$

Individual subpulses

$$c_l(t) = \begin{cases} e^{j\phi_l}, & 0 \le t \le \tau_c \\ 0, & \text{elsewhere} \end{cases} \tag{4}$$

are often referred to as chips [29] while $\left\{e^{j\phi_l}\right\}_{l=0}^{L-1}$ is called pulse compression (modulating) code sequence. There are two main categories for the modulating code sequences: biphase codes and polyphase codes. The

first one has only two possible choices for the phase state ϕ_l, while the second one offers more than two phase states (see, e.g., [29] and the references therein). In this study, ϕ_l is selected from a set consisting of q different phase states, and hence, $e^{j\phi_l}$ belongs to a q-point constellation $\mathcal{S}$.

In this approach, the treat radar generates a train of P equally spaced pulses given by:

$$x(t) = \sum_{p=0}^{P-1} c(t - pT), \quad 0 \leq t \leq PT, \tag{5}$$

where the pulse-to-pulse delay T is referred to as the pulse repetition interval (PRI) [30]. Let f_c be the carrier frequency. Then, the adversary radar transmits continues-time waveform $x(t)e^{j2\pi f_c t}$ towards the area of interest. Throughout the paper, we assume that $x(t)$ is a narrow-band signal.

Considering a clutter-free environment, the observed signal by the EW receiver can be expressed as follows:

$$y(t) = \kappa x\left(t - \frac{d(t)}{c}\right) e^{j2\pi f_c\left(t - \frac{d(t)}{c}\right)} + w(t), \tag{6}$$

where κ denotes the path loss value, c represents the speed of light, and $w(t)$ is a zero-mean white Gaussian noise with variance σ^2. The EW receiver first detects the received signal energy and estimates its carrier frequency. As in the related literature (e.g., [17], [4] and the references therein) it is assumed that the signal detection and carrier frequency estimation have already been accomplished. Therefore, substituting (2) into (6) and applying the narrow-band assumption, the baseband received signal can be approximated as:

$$y(t) \approx \kappa x(t) e^{-j2\pi f_c \frac{d_r(0) - vt - \eta(\theta)}{c}} + w(t). \tag{7}$$

Defining $f_d = \frac{f_c v}{c}$ as the Doppler shift caused by the movement of the adversary radar, (7) can be rewritten as:

$$y(t) = \gamma e^{j\frac{2\pi}{\lambda}\eta(\theta)} e^{j2\pi f_d t} x(t) + w(t), \tag{8}$$

where $\lambda = c/f_c$ is the transmitted signal wavelength and $\gamma \triangleq \kappa e^{-j\frac{2\pi}{\lambda}d_r(0)}$. In the remainder of the paper, we assume that the Doppler shift is small, i.e., $f_d \tau \ll 1$. Thus, the Doppler shift change within a pulse is negligible as compared to the change between pulses as in [8].

The pulse processor in the EW receiver first identifies the leading and falling edges of the received pulses through the envelope detection, and hence, allows them to be segmented from the train into a stack of P independent pulse observations [4]. Then, pulse observations are sampled at the chips (subpulse) frequency. The cyclostationarity properties of the phase-modulated pulses can be utilized in the subpulse rate estimation (for more insight, see [17] and [31]). We leave out the details to avoid redundancy and refer the interested

reader to the mentioned references.

The collected samples during the p-th pulse (segment) can be expressed in matrix form as:

$$\mathbf{y}_p = \gamma e^{j\frac{2\pi}{\lambda}\eta(\theta)} e^{j2\pi f_d(p-1)T}\mathbf{C}\mathbf{d}(f_d) + \mathbf{w}_p, \qquad p = 1, ..., P. \tag{9}$$

In (9), $\mathbf{y}_p = [y((p-1)T + 0T_s), ..., y((p-1)T + (L-1)T_s)]^T_{L\times 1}$ denotes the baseband equivalent of the p-th received pulse, where $T_s = \tau/L$ is the sampling interval that is assumed to satisfy the Nyquist requirement. In addition, $\mathbf{C} \triangleq \mathrm{diag}(\mathbf{c})$, where $\mathbf{c} = [c((p-1)T+0T_s), ..., c((p-1)T+(L-1)T_s)]^T_{L\times 1}$ represents the modulating code sequence[1] $\left\{e^{j\phi_l}\right\}_{l=0}^{L-1}$ employed by the adversary radar for the pulse compression purpose. Furthermore, $\mathbf{d}(f_d) = [e^{j2\pi f_d 0T_s}, ..., e^{j2\pi f_d(L-1)T_s}]^T_{L\times 1}$ and $\mathbf{w}_p = [w((p-1)T + 0T_s), ..., w((p-1)T + (L-1)T_s)]^T_{L\times 1}$.

The ultimate goal of this work is to estimate the azimuth angle θ and movement speed v of the adversary radar, while the knowledge of the threat waveform $\mathbf{C}$ is not available at EW receiver. To this end, exploiting the sparsity of radar signals in various spaces, we will transform (9) to the CS context in the following subsection.

B. Problem Definition

We first discretize the angle-Doppler space on a fine grid:

$$\mathbf{a} = [(a_1, b_1), (a_2, b_2), ..., (a_N, b_N)]\,. \tag{10}$$

Next, we define the matrix

$$\mathbf{\Psi}_p \triangleq \left[e^{j\frac{2\pi}{\lambda}\eta(a_1)} e^{j2\pi b_1(p-1)T}\mathbf{C}\mathbf{d}(b_1), ..., e^{j\frac{2\pi}{\lambda}\eta(a_N)} e^{j2\pi b_N(p-1)T}\mathbf{C}\mathbf{d}(b_N)\right]_{L\times N}, \tag{11}$$

which can be expressed as the multiplication of two matrixes, i.e., $\mathbf{\Psi}_p \triangleq \mathbf{C}\hat{\mathbf{\Psi}}_p$. Here, we define $\hat{\mathbf{\Psi}}_p = \left[e^{j\frac{2\pi}{\lambda}\eta(a_1)} e^{j2\pi b_1(p-1)T}\mathbf{d}(b_1), ..., e^{j\frac{2\pi}{\lambda}\eta(a_N)} e^{j2\pi b_N(p-1)T}\mathbf{d}(b_N)\right]_{L\times N}$. We also define the vector

$$\mathbf{s} \triangleq [s_1, ..., s_N]^T_{N\times 1} \tag{12}$$

such that

$$s_n = \begin{cases} \gamma, & \text{if the adversary radar is at } (a_n, b_n) \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

It is clear that $\mathbf{s}$ is a sparse vector in the angle-Doppler plane, and thus, the location of its non-zero element provides information on the angle-speed of the adversary radar. Finally, employing (11) and (12) enables us

[1]For convenience, we will sometimes use $\mathbf{c}$ and $\mathbf{C}$ interchangeably.

to provide the following sparse representation for the signal model (9):

$$\mathbf{y}_p = \mathbf{\Psi}_p \mathbf{s} + \mathbf{w}_p, \qquad p = 1, ..., P. \tag{14}$$

Using the predefined $M \times L$ $(M < L)$ measurement matrix $\mathbf{\Phi}$, our proposed EW receiver measures linear projections of $\mathbf{y}_p$ as:

$$\mathbf{r}_p = \mathbf{\Phi}\mathbf{y}_p = \mathbf{\Phi}\mathbf{\Psi}_p \mathbf{s} + \mathbf{n}_p, \qquad p = 1, ..., P, \tag{15}$$

where $\mathbf{n}_p = \mathbf{\Phi}\mathbf{w}_p$ for $p = 1, ..., P$. Note that since threat waveforms are not known, the knowledge of $\mathbf{C}$ is not available, and hence EW receiver is not able to construct $\mathbf{\Psi}_p = \mathbf{C}\hat{\mathbf{\Psi}}_p$. Therefore, the signal model (15) is a blind CS problem. In particular, our aim is to recover the sparse vector $\mathbf{s}$ from measurements of the form (15), where $\mathbf{\Psi}_p$ is unknown.

In the case of multiple measurement vectors, it is more convenient to formulate the basic model (15) in matrix form as:

$$\begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_P \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi C} & & 0 \\ & \ddots & \\ 0 & & \mathbf{\Phi C} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Psi}}_1 \\ \vdots \\ \hat{\mathbf{\Psi}}_P \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_P \end{bmatrix} \tag{16}$$

or, equivalently, as:

$$\mathbf{r} = \mathbf{M}\mathbf{\Psi}\mathbf{s} + \mathbf{n}, \tag{17}$$

where $\mathbf{r} = [\mathbf{r}_1^T, ..., \mathbf{r}_P^T]_{MP\times 1}^T$, $\mathbf{M}_{MP\times LP} = \mathbf{I}_P \otimes \mathbf{\Phi C}$, $\mathbf{\Psi} = [\hat{\mathbf{\Psi}}_1^T, ..., \hat{\mathbf{\Psi}}_P^T]_{LP\times N}^T$, and $\mathbf{n} = [\mathbf{n}_1^T, ..., \mathbf{n}_P^T]_{MP\times 1}^T$. As before mentioned, EW receiver is not aware of threat waveform, $\mathbf{C}$. Therefore, equation (17) contains two unknown parameters: the matrix $\mathbf{M}$ and the sparse vector $\mathbf{s}$. Since $\mathbf{\Phi}$ is a determined (random) matrix, once $\mathbf{C}$ is estimated, $\mathbf{M}$ can be obtained through $\mathbf{M} = \mathbf{I}_P \otimes \mathbf{\Phi C}$. Hence, in the sequel of this study, we consider the problem of $\mathbf{C}$ estimation rather than that of $\mathbf{M}$ estimation.

III. Blind Sparse Bayesian Learning Framework

In this Section, we first reformulate the introduced problem in Section III based on a SBL perspective, which is a popular approach for the sparse signal recovery in the context of compress sampling theory. Then, we derive an iterative algorithm to jointly estimate the angle-Doppler of the adversary radar and the threat waveform. Finally, making some improvements, we develop a modified version of the proposed algorithm which is suitable for practical purposes.

A. SBL Formulation

Here, to enhance the readability of the paper, we mathematically formulate our problem in SBL framework, and delay the derivation of the proposed algorithms to the next subsection.

The SBL has the conventional assumption that $\mathbf{n}$ in (17) is modeled probabilistically as white complex Gaussian noise with variance β. In most existing literature (see, e.g., [32], [23], and the references therein), the parameter β is assumed to be known which is a useful starting point because the analysis is more straightforward and there is no ambiguity involved regarding how the parameter β should be chosen. In practice, the knowledge of the variance β may not be perfectly available, and hence, β should be estimated from the data. Here, we assume that the variance β is unknown. Such an assumption implies the following multivariate Gaussian likelihood for the sparse vector $\mathbf{s}$:

$$p(\mathbf{r}|\mathbf{s};\mathbf{M},\beta) = (\pi\beta)^{-MP}\exp\Big(-\frac{1}{\beta}\big\|\mathbf{r}-\mathbf{M}\boldsymbol{\Psi}\mathbf{s}\big\|_2^2\Big). \tag{18}$$

As stated in [32], the maximum likelihood estimation of parameters from (18) leads to severe over-fitting problem. To overcome this issue, SBL impose an additional constraint on the parameters. In this context, similar as [22], [23], we aim to complement the likelihood function in (18) by employing the following prior distribution on the sparse vector $\mathbf{s}$:

$$p(\mathbf{s};\boldsymbol{\alpha}) = \prod_{i=1}^{N}(2\pi\alpha_i)^{-\frac{1}{2}}\exp\Big(-\frac{s_i^2}{2\alpha_i}\Big), \tag{19}$$

where $\boldsymbol{\alpha} = [\alpha_1,...,\alpha_N]_{N\times 1}^T$ is the vector of independent hyperparameters which is responsible for the sparsity properties of the model. Defining $\mathbf{A} = \mathrm{diag}(\boldsymbol{\alpha})$, (19) can be expressed in matrix form as:

$$p(\mathbf{s};\boldsymbol{\alpha}) = (2\pi)^{-\frac{N}{2}}|\mathbf{A}|^{-\frac{1}{2}}\exp\Big(-\frac{1}{2}\mathbf{s}^H\mathbf{A}^{-1}\mathbf{s}\Big). \tag{20}$$

Combining the likelihood and the prior within the Bayes rule, the posterior distribution of $\mathbf{s}$ is obtained. It is easy to show that the calculated posterior has a complex Gaussian distribution, i.e.,

$$p(\mathbf{s}|\mathbf{r};\boldsymbol{\alpha},\mathbf{M},\beta) = \mathcal{CN}(\boldsymbol{\mu},\boldsymbol{\Sigma}) \tag{21}$$

with mean and covariance given by

$$\boldsymbol{\mu} \triangleq E\big[\mathbf{s}|\mathbf{r};\boldsymbol{\alpha},\mathbf{M},\beta\big] = \beta^{-1}\boldsymbol{\Sigma}\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{r} \tag{22}$$

and

$$\boldsymbol{\Sigma} \triangleq \mathrm{Cov}\big[\mathbf{s}|\mathbf{r};\boldsymbol{\alpha},\mathbf{M},\beta\big] = \big(\beta^{-1}\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi} + \mathbf{A}^{-1}\big)^{-1}. \tag{23}$$

In a typical SBL framework, the posterior mean $\boldsymbol{\mu}$ is chosen as the sparse vector estimate. Therefore, we have:

$$\hat{\mathbf{s}} \triangleq \boldsymbol{\mu} = \big(\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi} + \beta\mathbf{A}^{-1}\big)^{-1}\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{r}. \tag{24}$$

Note that our ultimate goal is to estimate the unknown parameters $\boldsymbol{\alpha}$, $\mathbf{C}$, and β which are required for the sparse vector recovery. These parameters can be found from the measurements $\mathbf{r}$ via maximizing the marginalized pdf given by:

$$\begin{aligned} p(\mathbf{r};\boldsymbol{\alpha},\mathbf{M},\beta) &= \int p(\mathbf{r}|\mathbf{s};\mathbf{M},\beta)p(\mathbf{s};\boldsymbol{\alpha})d\mathbf{s} \\ &= (2\pi)^{-MP}|\boldsymbol{\Sigma}_r|^{-\frac{1}{2}}\exp\Big(-\frac{1}{2}\mathbf{r}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r}\Big), \end{aligned} \tag{25}$$

where $\boldsymbol{\Sigma}_r \triangleq \beta\mathbf{I}_{MP} + \mathbf{M}\boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Psi}^H\mathbf{M}^H$. This approach is called evidence maximization strategy or type-II maximum likelihood procedure [21]–[23]. In summary, we aim to solve the following optimization problem:

$$\big(\hat{\boldsymbol{\alpha}},\hat{\mathbf{C}},\hat{\beta}\big) = \arg\max_{\boldsymbol{\alpha},\mathbf{C},\beta}\ p(\mathbf{r};\boldsymbol{\alpha},\mathbf{M},\beta). \tag{26}$$

It is very difficult to find any closed form solutions for the optimization problem (26). As stated in the literature [21]–[23], there are two different strategies to optimize such an optimization problem. In the first method, which is called MacKay approach [33], the optimization problem can be minimized by taking the derivative with respect to the intended unknown parameter, setting the derivative to zero, rearranging, and forming fixed-point equation. The second approach, called EM-based method, resorts to the EM algorithm to minimize such a problem. In what follows, applying second approach, we will consider the sparse vector $\mathbf{s}$ as the hidden variable and maximize $E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha},\mathbf{C},\beta}\big[\log p(\mathbf{r},\mathbf{s};\boldsymbol{\alpha},\mathbf{M},\beta)\big]$, where $p(\mathbf{r},\mathbf{s};\boldsymbol{\alpha},\mathbf{M},\beta) = p(\mathbf{r}|\mathbf{s};\mathbf{M},\beta)p(\mathbf{s};\boldsymbol{\alpha})$ represents the likelihood of the complete data set $\{\mathbf{s},\mathbf{r}\}$. Notice that by virtue of the monotonicity property of the EM-based SBL approach [22], the convergence of our method is guaranteed. This statement will be validated through solid theoretical justifications in Section IV. Moreover, since both $\mathbf{C}$ and β estimations are packaged with the hyperparameters during evidence maximization (Section III-B), this guarantee is also applicable to both of them. Numerical results in Section V empirically ensures the validate of this claim.

B. Blind-SBL Algorithm

The basic idea of the EM algorithm is to solve the optimization problem (26) iteratively according to the following two steps:

$$\text{1-E-step: Compute } Q(\boldsymbol{\alpha}, \mathbf{C}, \beta|\boldsymbol{\alpha}^{(k)}, \mathbf{C}^{(k)}, \beta^{(k)}) = E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\big[\log p(\mathbf{r}, \mathbf{s}; \boldsymbol{\alpha}, \mathbf{M}, \beta)\big] \tag{27}$$

$$\text{2-M-step: Solve } \big(\boldsymbol{\alpha}^{(k+1)}, \mathbf{C}^{(k+1)}, \beta^{(k+1)}\big) = \arg\max_{\boldsymbol{\alpha},\mathbf{C},\beta}\ Q(\boldsymbol{\alpha}, \mathbf{C}, \beta|\boldsymbol{\alpha}^{(k)}, \mathbf{C}^{(k)}, \beta^{(k)}). \tag{28}$$

Note that since $Q(\boldsymbol{\alpha}, \mathbf{C}, \beta|\boldsymbol{\alpha}^{(k)}, \mathbf{C}^{(k)}, \beta^{(k)})$ is a nondecreasing function of k [34], the convergence of the EM algorithm is guaranteed. It is well known that the accuracy of the EM algorithm is closely related to the quality of the unknown parameters $\boldsymbol{\alpha}^{(0)}$, $\mathbf{c}^{(0)}$, and $\beta^{(0)}$. We show that the random initialization for $\mathbf{C}$ and a non-negative random initialization for $\boldsymbol{\alpha}$ and β leads to accurate solutions (see Section VI).

In the E-step, the expectation is taken with respect to the hidden variable $\mathbf{s}$ conditioned on $\mathbf{r}$, $\boldsymbol{\alpha}^{(k)}$, $\mathbf{C}^{(k)}$, and $\beta^{(k)}$. Using (18) and (20), it can be shown that:

$$\begin{aligned} Q(\boldsymbol{\alpha}, \mathbf{C}, \sigma^2|\boldsymbol{\alpha}^{(k)}, \mathbf{C}^{(k)}, \beta^{(k)}) &= E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\big[\log p(\mathbf{r}|\mathbf{s}; \mathbf{M}, \beta) + \log p(\mathbf{s}; \boldsymbol{\alpha})\big] \\ &= -E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[MP\log\beta + \frac{1}{\beta}\big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\mathbf{s}\big\|_2^2\Big] \\ &\quad -\frac{1}{2}E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[\log|\mathbf{A}| + \mathbf{s}^H\mathbf{A}^{-1}\mathbf{s}\Big] + const, \end{aligned} \tag{29}$$

where terms without dependency on $\boldsymbol{\alpha}^{(k)}$, $\mathbf{C}^{(k)}$, and $\beta^{(k)}$ are discarded. The first expectation in (29) is independent of $\boldsymbol{\alpha}$ while the second one does not rely on $\mathbf{C}$ and β. Therefore, the E-step can be rewritten as:

$$Q(\boldsymbol{\alpha}, \mathbf{C}, \beta|\boldsymbol{\alpha}^{(k)}, \mathbf{C}^{(k)}, \beta^{(k)}) = Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)}) + Q(\mathbf{C}, \sigma^2|\mathbf{C}^{(k)}, \beta^{(k)}), \tag{30}$$

where $Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)})$ and $Q(\mathbf{C}, \beta|\mathbf{C}^{(k)}, \beta^{(k)})$ are respectively given by:

$$Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)}) = -\frac{1}{2}E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[\log|\mathbf{A}| + \mathbf{s}^H\mathbf{A}^{-1}\mathbf{s}\Big] \tag{31}$$

and

$$Q(\mathbf{C}, \sigma^2|\mathbf{C}^{(k)}, \beta^{(k)}) = -E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[MP\log\beta + \frac{1}{\beta}\big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\mathbf{s}\big\|_2^2\Big]. \tag{32}$$

Based on (30)-(32), the M-step can be decoupled into two independent optimization problems. In what follows, we first optimize $Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)})$ to find $\boldsymbol{\alpha}^{(k)}$ and then optimize $Q(\mathbf{C}, \beta|\mathbf{C}^{(k)}, \beta^{(k)})$ to obtain $\mathbf{C}^{(k)}$ and $\beta^{(k)}$.

The update rule for $\boldsymbol{\alpha}$ is obtained by setting the derivative of $Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)})$ over $\boldsymbol{\alpha}$ to zero. The derivative of (31) with respect to $\boldsymbol{\alpha}(n)$ can be expressed as follows:

$$\begin{aligned}\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(k)})}{\partial \boldsymbol{\alpha}(n)} &= -\frac{1}{2\boldsymbol{\alpha}(n)} + \frac{1}{2\boldsymbol{\alpha}^2(n)} E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[|\mathbf{s}(n)|^2\Big] \\ &= -\frac{1}{2\boldsymbol{\alpha}(n)} + \frac{1}{2\boldsymbol{\alpha}^2(n)}\big(|\boldsymbol{\mu}(n)|^2 + \boldsymbol{\Sigma}(n,n)\big). \end{aligned} \tag{33}$$

Hence, we find the following update rule for $\boldsymbol{\alpha}$ during the M-step:

$$\boldsymbol{\alpha}^{(k+1)}(n) = |\boldsymbol{\mu}(n)|^2 + \boldsymbol{\Sigma}(n,n), \qquad n = 1, ..., N. \tag{34}$$

It is worthy to mention that upon convergence, $\hat{\boldsymbol{\alpha}}(n) = 0$ forces the posterior to satisfy $p(\mathbf{s}(n) = 0|\mathbf{r}; \hat{\boldsymbol{\alpha}}(n) = 0, \mathbf{M}, \beta) = 1$. In other words, estimating the sparsity profile of $\mathbf{s}$ is converted to estimating the hyperparameter vector $\boldsymbol{\alpha}$ with the correct location of the nonzero element.

On the other hand, the update rules for $\mathbf{C}$ and β can be found by maximising $Q(\mathbf{C}, \sigma^2|\mathbf{C}^{(k)}, \beta^{(k)})$, e.g.,

$$\begin{aligned}\big(\mathbf{C}^{(k+1)}, \beta^{(k+1)}\big) &= \arg\max_{\mathbf{C},\beta}\ Q(\mathbf{C}, \beta|\mathbf{C}^{(k)}, \beta^{(k)}) \\ &= \arg\min_{\mathbf{C},\beta}\ \Big\{MP\log\beta + \frac{1}{\beta}\text{trace}\big(\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^H)\big) - \frac{2}{\beta}\Re e\big(\mathbf{r}^H\mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu}\big)\Big\}, \end{aligned} \tag{35}$$

where (35) follows from the fact that:

$$\begin{aligned} E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[\big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\mathbf{s}\big\|_2^2\Big] &= E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[\big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu} + \mathbf{M}\boldsymbol{\Psi}(\boldsymbol{\mu} - \mathbf{s})\big\|_2^2\Big] \\ &= \big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu}\big\|_2^2 + E_{\mathbf{s}|\mathbf{r};\boldsymbol{\alpha}^{(k)},\mathbf{C}^{(k)},\beta^{(k)}}\Big[(\boldsymbol{\mu} - \mathbf{s})^H\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi}(\boldsymbol{\mu} - \mathbf{s})\Big] \\ &= \big\|\mathbf{r} - \mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu}\big\|_2^2 + \text{trace}\big(\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi}\boldsymbol{\Sigma}\big) \\ &= \mathbf{r}^H\mathbf{r} - 2\Re e\big(\mathbf{r}^H\mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu}\big) + \text{trace}\big(\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^H)\big). \end{aligned} \tag{36}$$

Since $\mathbf{M} = (\mathbf{I}_P \otimes \boldsymbol{\Phi})(\mathbf{I}_P \otimes \mathbf{C})$, we have:

$$\mathbf{r}^H\mathbf{M}\boldsymbol{\Psi}\boldsymbol{\mu} = \text{trace}\big((\mathbf{I}_P \otimes \mathbf{C})\boldsymbol{\Psi}\boldsymbol{\mu}\mathbf{r}^H(\mathbf{I}_P \otimes \boldsymbol{\Phi})\big) = \sum_{p=1}^{P}\text{trace}\big(\mathbf{C}\hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi}\big). \tag{37}$$

Further, let $\beta^{(k-1)}$ stands for the estimated β in the previous iteration. Then using (23) results in:

$$\boldsymbol{\Psi}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\Psi} = \beta^{(k-1)}\big(\boldsymbol{\Sigma}^{-1} - \mathbf{A}^{(k)}\big). \tag{38}$$

Substituting (37) and (38) into (35) leads us to the following Q function:

$$Q(\mathbf{C},\beta|\mathbf{C}^{(k)},\beta^{(k)}) = MP\log\beta \quad + \quad \frac{\beta^{(k-1)}}{\beta}\text{trace}\big((\boldsymbol{\Sigma}^{-1}-\mathbf{A}^{(k)})(\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^H)\big) - \quad \frac{2}{\beta}\sum_{p=1}^{P}\Re e\big\{\text{trace}\big(\mathbf{C}\hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi}\big)\big\}. \tag{39}$$

The derivative of (39) over β is given by:

$$\frac{\partial Q(\mathbf{C},\sigma^2|\mathbf{C}^{(k)},\beta^{(k)})}{\partial\beta} = \frac{MP}{\beta} \quad - \quad \frac{\beta^{(k-1)}}{\beta^2}\text{trace}\big((\boldsymbol{\Sigma}^{-1}-\mathbf{A}^{(k)})(\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^H)\big) + \quad \frac{2}{\beta^2}\sum_{p=1}^{P}\Re e\big\{\text{trace}\big(\mathbf{C}\hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi}\big)\big\}, \tag{40}$$

which results in the following update rule for β during the M-step:

$$\beta^{(k+1)} = \frac{\beta^{(k)}\text{trace}\big((\boldsymbol{\Sigma}^{-1}-\mathbf{A}^{(k)})(\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^H)\big) - 2\sum_{p=1}^{P}\Re e\big\{\text{trace}\big(\mathbf{C}\hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi}\big)\big\}}{MP}. \tag{41}$$

From (39), it is seen that only the last term of $Q(\mathbf{C},\sigma^2|\mathbf{C}^{(k)},\beta^{(k)})$ is related on $\mathbf{C}$ and other terms are independent from $\mathbf{C}$. Therefore, the learning rule for $\mathbf{C}=\text{diag}(\mathbf{c})$ is given by:

$$\mathbf{c}^{(k+1)} = \arg\max_{\mathbf{c}\in\mathcal{Q}}\sum_{p=1}^{P}\Re e\big\{\text{trace}(\mathbf{C}\hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi})\big\}. \tag{42}$$

In (42), $\mathcal{Q}$ denotes an L-dimensional space that contains q^L different vectors. Considering the fact that $\mathbf{C}$ is a diagonal matrix, the update rule in (42) can be decoupled into L independent maximization problems as:

$$\mathbf{c}^{(k+1)}(l) = \arg\max_{x\in\mathcal{S}}\sum_{p=1}^{P}\Re e\big\{x\mathbf{D}_p(l,l)\big\}, \qquad l=1,...,L, \tag{43}$$

where $\mathbf{D}_p \triangleq \hat{\boldsymbol{\Psi}}_p\boldsymbol{\mu}\mathbf{r}_p^H\boldsymbol{\Phi}$. The crucial difference between (42) and (43) is that, instead of an exhaustive search in a L-dimensional space, (43) only needs examine an one-dimensional space. This reduces the number of possible solutions from q^L to only qL options. Therefore, incorporating (43) significantly decreases the computational complexity of the proposed EM-based method.

In the following, we will refer the method using the learning rules (34), (41) and (43) as the Blind-SBL algorithm.

C. Modified Algorithm: Pruned Blind-SBL

In modern combat environments, it is very important that the EW receiver be able to quickly identifies any possible threats, which requires employing fast algorithms. This motivates us to further reduce the computational

complexity and improve the convergence speed of the proposed blind-SBL algorithm. In this subsection, we aim to develop a modified version of blind-SBL algorithm which is suitable for practical scenarios.

Recalling that the main computational complexity of each iteration of the Blind-SBL algorithm comes from the matrix inversion in (23), which is required for updating $\mathbf{\Sigma}$. On the other hand, after a few iterations, most of the estimated $\boldsymbol{\alpha}(n)$'s goes to zero, e.g., $\boldsymbol{\alpha}(n) \rightarrow 0$ $(n = 1, ..., N)$ [21]. Therefore, during the learning procedure when any $\boldsymbol{\alpha}(n)(n = 1, ..., N)$ becomes less than a small predefined threshold (e.g., 10^{-4}), then it can be pruned from the model along with the associated $\mathbf{M\Psi}$ matrix column. In other words, $\mathbf{A}^{(k+1)}$ can be initialized with a pruned version of its previous estimate, $\mathbf{A}^{(k)}$. Hence, applying a pruning step between consecutive iterations, decreases the computational complexity of our algorithm and improves its convergence speed.

It is also worthy to mention that we terminate EM iteration when $\|\boldsymbol{\alpha}^{(k+1)} - \boldsymbol{\alpha}^{(k)}\|_2^2$ becomes lower than a specified threshold ε or when k attains a maximum number of iterations k_{max}, where both ε and k_{max} values are set by user.

In the following, we will refer the modified method as the Pruned Blind-SBL algorithm. Recalling that the EW receiver observes $\mathbf{r}$ and knows $\mathbf{\Phi}$ and $\mathbf{\Psi}$, the steps of the proposed Pruned Blind-SBL algorithm can be summarized as follows:

1. Initialization:

Set the iteration count k to zero, $\boldsymbol{\alpha}^{(0)} = \mathbf{1}_L$, $\beta^{(0)} = \beta_{init}$, and a random initialization for $\mathbf{c}^{(0)}$.

2. Iteration:

a) E-step: Compute the posterior sufficient statistics $\mathbf{\Sigma}$ and $\boldsymbol{\mu}$ using the following equations:

$$\mathbf{\Sigma} = \left(\beta^{-1^{(k)}} \mathbf{\Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi}\mathbf{C}^{(k)})^H (\mathbf{I}_P \otimes \mathbf{\Phi}\mathbf{C}^{(k)})\mathbf{\Psi} + \mathbf{A}^{-1^{(k)}}\right)^{-1} \tag{44}$$

$$\boldsymbol{\mu} = \beta^{-1^{(k)}} \mathbf{\Sigma}\mathbf{\Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi}\mathbf{C}^{(k)})^H \mathbf{r} \tag{45}$$

b) M-step: Update the weights for all $n = 1, ..., N$ using

$$\boldsymbol{\alpha}^{(k+1)}(n) = |\boldsymbol{\mu}(n)|^2 + \mathbf{\Sigma}(n, n) \tag{46}$$

and estimate the code for all $l = 1, ..., L$ using

$$\mathbf{c}^{(k+1)}(l) = \arg\max_{x \in \mathcal{S}} \sum_{p=1}^{P} \Re e\{x \mathbf{D}_p(l, l)\}, \qquad l = 1, ..., L, \tag{47}$$

where $\mathbf{D}_p = \hat{\mathbf{\Psi}}_p \boldsymbol{\mu} \mathbf{r}_p^H \mathbf{\Phi}$. Obtain the estimate of β through the following update rule:

$$\beta^{(k+1)} = \frac{\left\|\mathbf{r} - (\mathbf{I}_P \otimes \mathbf{\Phi}\mathbf{C}^{(k)})\mathbf{\Psi}\boldsymbol{\mu}\right\|_2^2 + \beta^{(k)}\text{trace}\big((\mathbf{\Sigma}^{-1} - \mathbf{A}^{(k)})\mathbf{\Sigma}\big)}{MP}. \tag{48}$$

c) Pruning operation: Performing a pruning mechanism based on choosing $\boldsymbol{\alpha}^{(k+1)}(n)(n = 1, ..., N)$ which are smaller than a pre-defined threshold.

d) Increment k and go to a) until convergence ($\|\boldsymbol{\alpha}^{(k+1)} - \boldsymbol{\alpha}^{(k)}\|_2^2 < \varepsilon$) or if k reaches k_{max}.

3. Final approximation:

$\hat{\mathbf{s}} = \boldsymbol{\mu}$ determines the azimuth angle θ and movement speed v of the adversary radar while $\hat{\mathbf{c}} = \mathbf{c}^{(k)}$ determines the pulse compression code used by the adversary radar.

IV. Analysis of Global and Local Minima

In the context of CS, most of the existed sparse recovery method, typically offer some guarantees on the signal recovery accuracy. For example, the basic SBL framework has guaranteed convergence to the sparsest solution when the noise variance β is zero, and convergence to a sparse local minimum irrespective the presence of noise [22]. The focus of this section is to guarantee that the proposed blind-SBL framework is able to provide a good solution to the optimization problem (26).

We first note that the optimization problem (26) is equivalent to the following optimization problem:

$$\arg\min_{\boldsymbol{\alpha},\mathbf{C},\beta} \; -2\log p(\mathbf{r}; \boldsymbol{\alpha}, \mathbf{M}, \beta) \equiv \arg\min_{\boldsymbol{\alpha},\mathbf{C},\beta} \; \mathcal{L}, \tag{49}$$

where based on (25), cost function $\mathcal{L}$ can be defined as:

$$\mathcal{L} = \log|\mathbf{\Sigma}_r| + \mathbf{r}^H \mathbf{\Sigma}_r^{-1} \mathbf{r}. \tag{50}$$

We now provide a theoretical justification about the global and local minima of the cost function (50).

A. Analysis of the Global Minima

Some preliminary definitions are required to proceed. Let $\mathbf{s}_0$ denotes the maximally sparse solution to $\mathbf{r} = \mathbf{M}\mathbf{\Psi}\mathbf{s}$. In such a noiseless environment, the obtained expressions for $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ in (22) and (23) can be rewritten as:

$$\boldsymbol{\mu} = \mathbf{A}^{1/2}\big(\mathbf{M}\mathbf{\Psi}\mathbf{A}^{1/2}\big)^{\dagger}\mathbf{r} \tag{51}$$

$$\mathbf{\Sigma} = \big(\mathbf{I} - \mathbf{A}^{1/2}\big(\mathbf{M}\mathbf{\Psi}\mathbf{A}^{1/2}\big)^{\dagger}\mathbf{M}\mathbf{\Psi}\big)\mathbf{A}, \tag{52}$$

where (51) and (52) results from linear algebra. Furthermore, let N_0 stands for the number of nonzero elements in $\mathbf{s}_0$ and $\| \mathbf{s}_0 \|_2 < \infty$.

We now present a theorem that ensures the global minimum property of the introduced cost function in (50) for the noise-free case.

Theorem 1. In the limit as $\beta \to 0$ (noiseless case) and assuming $N_0 < M$, for the cost function (50) the global minimum $\hat{\boldsymbol{\alpha}}$ produces the sparse solution $\hat{\mathbf{s}}$ which is equal to $\mathbf{s}_0$ regardless of the estimated $\hat{\mathbf{C}}$; where $\hat{\mathbf{s}}$ is calculated using (51).

Proof. See Appendix A. □

It is important to note that the performance of the proposed blind-SBL framework is closely related to the quality of the estimated $\hat{\mathbf{C}}$, because $\hat{\boldsymbol{\alpha}}$ is a function of $\hat{\mathbf{C}}$. However, this theorem stated that even when the estimated $\hat{\mathbf{C}}$ is different from the true $\mathbf{C}$, the estimated $\hat{\mathbf{s}}$ is equal to the maximally sparse solution $\mathbf{s}_0$ at the global minimum of the cost function (50). As our experiments show later, this does not mean that $\hat{\mathbf{C}}$ has no effect on the performance of the proposed algorithms.

B. Analysis of the Local Minima

Since the cost function (50) can potentially have some local minima, this subsection is devoted to study the local minimum property of $\mathcal{L}$. To do so, we first provide two required lemmas for the first and second terms of the cost function (50), respectively.

Lemma 1. For any fixed $\mathbf{C}$, the term $\log|\boldsymbol{\Sigma}_r| = \log\left|\beta\mathbf{I}_{MP} + (\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})\boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Psi}^H(\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})^H\right|$ is concave with respect to $\mathbf{A}$ (or, equivalently, $\boldsymbol{\alpha}$).

Proof. The desired result can be achieved by applying the composition property of concave functions [24]. We leave out the details to avoid redundancy. □

Lemma 2. For any fixed $\mathbf{C}$, the term $\mathbf{r}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r} = \mathbf{r}^H\left(\beta\mathbf{I}_{MP} + (\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})\boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Psi}^H(\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})^H\right)^{-1}\mathbf{r}$ is equal to a constant z as long as the following linear constraints hold for all $\boldsymbol{\alpha}$:

$$\mathbf{v} = \mathbf{W}\boldsymbol{\alpha}. \tag{53}$$

In (53), $\mathbf{v} \triangleq \mathbf{r} - \beta\mathbf{u}$ and $\mathbf{W} \triangleq (\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})\boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\Psi}^H(\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})^H\mathbf{u})$, where $\mathbf{u}$ is any fixed vector such that $\mathbf{r}^H\mathbf{u} = z$.

Proof. See [22] for the proof. □

We now present the following theorem on the convergence of the cost function (50) to a sparse local minima, which is applicable to both noisy and noiseless conditions.

Theorem 2. Every local minimum of the cost function (50) occurs at a sparse solution regardless of the values of β and $\mathbf{C}$.

Proof. See Appendix B. □

As will be discussed later, increasing P mitigates the effects of the noise. More importantly, increasing P improves the chance that the proposed algorithms converge to the global minimum of the cost function (50) rather than its local minimum, regardless of the presence or absence of noise. In other words, the globally minimizing Blind-SBL hyperparameters guarantee convergence to the sparsest solution, and increasing the value of P raises the probability of finding these hyperparameters.

V. NUMERICAL SIMULATIONS

The goal of this section is to demonstrate the performance of the proposed Blind-SBL framework and analyse its properties through a set of Monte Carlo simulations. The results that we show are obtained by averaging the corresponding quantity over 10^4 independent simulation runs.

A. Simulation Setup

In practice, based on the desired application and sensing equipment, EW receivers are tuned in frequency to a specific band. Consequently, multiple the EW receivers are typically needed for surveillance of the whole spectrum. In summery, EW receivers first detect the signal energy and estimate carrier frequency. Then, the pulse processors identify pulse parameters. Here, similar to existed literature, it is assumed that the required preprocessing has been accomplished and the unknown threat signal has been initially identified as a phase-modulated waveform (for more insight, please see [17], [4], and references therein). Throughout our simulations, no other assumptions are made about the input threat signal. For the simulation purposes, we generate random pulse compression code sequences of length $L = 32$ and the carrier frequency is considered to be 5 GHz.

Computational complexity of the proposed algorithms depends on the number of grid points, N. Specifically, complexity increases with the size of the grid, N, or equivalently, as the discretization step decreases. Thus, restricting the candidate angle-Doppler space to a small interval significantly reduces the computation time. This can be done by first providing an initial angle-Doppler estimation employing a coarse grid, and then constructing a finer grid around the initial obtained values to improve the estimation accuracy [8], [13]. In our

simulations, the azimuth angle and Doppler frequency are randomly generated within a grid of size $N = 64$. Moreover, it is assumed that the targets fall on the grid points.

The goal of the proposed blind SBL framework is to estimate the unknown pulse modulating code sequence $\mathbf{C}$, as well as finding the azimuth angle θ and movement speed v of the target. Note that the sparsity profile of vector $\hat{\mathbf{s}}$ provides target azimuth angle θ and movement speed v. Therefore, we use two different metrics to evaluate the performance of the proposed framework. One is the *Failure Rate* (*Detection Rate*) defined in [23] and [24], which presents the percentage of failed trials in the total trials. A failed trial is recognized if the index of the non-zero element in estimated sparse vector $\hat{\mathbf{s}}$ is different from the index of the non-zero element in true sparse vector $\mathbf{s}$. In addition, we employ *Code Sequence Error Rate* as the second performance metric, which is defined as the ratio of the number of incorrectly estimated $\hat{\mathbf{C}}$ to the number of total simulation runs.

In the context of CS theory, sparse vector recovery requires that measurement matrix is incoherent with basis matrix. It is also well known that any random measurement matrix with independent and identically distributed elements is almost incoherent with any fixed basis matrix [35]. This property is simply achieves by selecting the measurement matrix elements from a Gaussian distribution [15]. Hence, throughout simulations, we generate random measurement matrices with Gaussian distribution.

B. Simulation Results

First, we test the effect of the number of EM iterations on the convergence characteristics of the proposed blind-SBL framework. In Fig. 1, we depict the Failure Rate curves of the proposed algorithm versus the number of EM iterations when SNR is set to 4 (dB). The Failure Rate is averaged for each P over 10^3 independent Monte-Carlo runs. As it can be seen, the Failure Rate performance converges after ten iterations for almost all P values. Therefore, throughout simulations, the stopping criterion is chosen as $\|\alpha^{(k+1)} - \alpha^{(k)}\|_2^2 < 10^{-8}$ or $k_{\max} = 20$ iterations, whichever is achieved earlier. In addition, Fig. 1 also reveals that increasing P significantly improves the probability of convergence to the global minimum (sparsest solution).

Next, we study the influence of the variance β estimation error on the performance of proposed Blind-SBL framework. In Fig. 2, we present the Detection Rate performance as a function of P for SNRs of 2 and 14 (dB). We also show the case where the knowledge of the variance β is perfectly available at the EW receiver, as a benchmark. Fig. 2 indicates that although the β update rule in (41) may not be optimal at least in low SNR values, the proposed Blind-SBL algorithm achieves near-optimal performance in all ranges of SNR values. Therefore, the proposed Blind-SBL algorithm is a reliable solution for practical scenarios where β is an unknown parameter.

We also investigate the performance of the proposed estimation rule for $\mathbf{C}$ in (43). Fig. 3 shows the Code

Sequence Error Rate versus P. Here, SNR varies from -6 (dB) to 10 (dB) while other parameters keep fixed. From Fig. 3, we observe that increasing P significantly improves the pulse modulating code estimation accuracy. Moreover, it is obvious that increasing P offers more performance improvement compare with increasing the SNR value.

Here, we demonstrate how proposed algorithms benefit from increasing the number of measurements, P. Fig. 4 presents the Detection Rate versus the number of measurements when SNR varies from -2 (dB) to 8 (dB), while Fig. 5 illustrates the Detection Rate against SNR for different P values. From Fig. 4, it is easily seen that increasing P significantly improves the probability of angle-Doppler detection and at the same time, mitigates the effect of the noise. In addition, the behaviors of ploted curves in Fig. 5 confirms that increasing the number of measurements leads to much more performance improvement rather than increasing the SNR.

Finally, we compare the performance of the proposed blind-SBL and pruned blind-SBL algorithms. To this end, in Fig. 6, we depict the Detection Rate curves versus the number of measurements for SNR of -2 (dB). Here, we present the detection rate performance of the pruned blind-SBL algorithm at three different threshold values, i.e. 0.005, 0.001, and 0.0005. Fig. 6 reveals that with large threshold value (0.005), the performance gap between blind-SBL and pruned blind-SBL algorithms increases. In contrast, the computational complexity as well as the convergence speed improves. So, due to the performance trade-off between accuracy and complexity, one can set the desired threshold value for each practical scenario.

APPENDIX A
PROOF OF THEOREM 1

Let $\hat{\mathbf{s}}$ is computed using $\mathbf{A}^{1/2}\big(\mathbf{M}\boldsymbol{\Psi}\mathbf{A}^{1/2}\big)^{\dagger}\mathbf{r}$ according to (24). For given $\hat{\mathbf{C}}$, the hyperparameters $\hat{\boldsymbol{\alpha}} = [\hat{\alpha_1}, ..., \hat{\alpha_N}]$ is obtained by globally minimizing the following cost function:

$$\mathcal{L} = \log|\boldsymbol{\Sigma}_r| + \mathbf{r}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r}. \tag{A.1}$$

It is clear that $\mathbf{r}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r}$ is strictly greater than zero for all $\boldsymbol{\alpha}$, $\mathbf{C}$, and β. If we reduce the value of $|\boldsymbol{\Sigma}_r|$ to zero, the cost function $\mathcal{L}$ achieves its minimum (minus infinity). In other word, the minimum of cost function $\mathcal{L}$ occurs whenever $|\boldsymbol{\Sigma}_r| = 0$ while maintaining some finite bound z such that $0 \leq \mathbf{r}^H\boldsymbol{\Sigma}_r^{-1}\mathbf{r} \leq z$. It is worthy to mention that we fix $\hat{\mathbf{C}}$ because it does not affect these conditions at all.

Based on the Theorem 1 in [22], one can show that $\hat{\boldsymbol{\alpha}}$ satisfies the following conditions:

$$|\boldsymbol{\Sigma}_r| = \big|\beta\mathbf{I}_{MP} + (\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})\boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Psi}^H(\mathbf{I}_P \otimes \boldsymbol{\Phi}\mathbf{C})^H\big| = 0 \tag{A.2}$$

and

$$0 \leq \mathbf{r}^H \left(\beta \mathbf{I}_{MP} + (\mathbf{I}_P \otimes \mathbf{\Phi C}) \mathbf{\Psi A \Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi C})^H\right)^{-1} \mathbf{r} \leq z. \tag{A.3}$$

Therefore, we finish the proof.

APPENDIX B
PROOF OF THEOREM 2

The proof results from the Theorem 2 in [22] and using our Lemma 1 and Lemma 2.

The cost function $\mathcal{L}$ consists of two terms: $\log|\mathbf{\Sigma}_r|$ and $\mathbf{r}^H \mathbf{\Sigma}_r^{-1} \mathbf{r}$. We aim to minimize $\log|\mathbf{\Sigma}_r|$ while keeping $\mathbf{r}^H \mathbf{\Sigma}_r^{-1} \mathbf{r}$ constant to some z. To move forward in this direction, we define the following optimization problem:

$$\begin{aligned} \min \quad & g(\boldsymbol{\alpha}) \\ \text{subject to} \quad & \mathbf{W}\boldsymbol{\alpha} = \mathbf{v} \\ & \boldsymbol{\alpha} \geq 0, \end{aligned} \tag{B.1}$$

where $g(\boldsymbol{\alpha}) \triangleq \log\left|\beta \mathbf{I}_{MP} + (\mathbf{I}_P \otimes \mathbf{\Phi C}) \mathbf{\Psi A \Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi C})^H\right|$ is a concave function with respect to $\boldsymbol{\alpha}$ (Lemma 1). Furthermore, the constraints $\mathbf{W}\boldsymbol{\alpha} = \mathbf{v}$ and $\boldsymbol{\alpha} \geq 0$ hold $\mathbf{r}^H \mathbf{\Sigma}_r^{-1} \mathbf{r} = \mathbf{r}^H \left(\beta \mathbf{I}_{MP} + (\mathbf{I}_P \otimes \mathbf{\Phi C}) \mathbf{\Psi A \Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi C})^H\right)^{-1} \mathbf{r}$ constant on a closed, bounded convex polytope (Lemma 2).

It is clear that any local minimum of cost function $\mathcal{L}$, e.g., $\hat{\boldsymbol{\alpha}}$, $\hat{\mathbf{C}}$, and $\hat{\beta}$, must also be a local minimum of the optimization problem (B.1) with $z = \mathbf{r}^H \mathbf{u} = \mathbf{r}^H \left(\hat{\beta} \mathbf{I}_{MP} + (\mathbf{I}_P \otimes \mathbf{\Phi}\hat{\mathbf{C}}) \mathbf{\Psi}\hat{\mathbf{A}}\mathbf{\Psi}^H (\mathbf{I}_P \otimes \mathbf{\Phi}\hat{\mathbf{C}})^H\right)^{-1} \mathbf{r}$. According to Theorem 6.5.3 in [36], all minima of (B.1) are accrued at extreme points. In addition, based on the definition of basic feasible solution (BFS) [24], since $\boldsymbol{\alpha}$ satisfies $\mathbf{W}\boldsymbol{\alpha} = \mathbf{v}$, then it is a BFS to $\mathbf{W}\boldsymbol{\alpha} = \mathbf{v}$. Considering the equivalence between extreme points and BFS [36], we conclude that all local minima of $\mathcal{L}$ must be achieved at sparse solutions.

REFERENCES

[1] C. Berger, B. Demissie, J. Heckenbach, P. Willett, and S. Zhou, "Signal processing for passive radar using ofdm waveforms," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 1, pp. 226–238, Feb 2010.

[2] C.-H. Cheng, D. Lin, L. Liou, and J. Tsui, "Electronic warfare receiver with multiple fft frame sizes," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 48, no. 4, pp. 3318–3330, October 2012.

[3] J. Grajal, R. Blazquez, G. Lopez-Risueno, J. Sanz, M. Burgos, and A. Asensio, "Analysis and characterization of a monobit receiver for electronic warfare," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 39, no. 1, pp. 244–258, Jan 2003.

[4] B. Rigling and C. Roush, "Acf-based classification of phase modulated waveforms," in *Proc. IEEE Radar Conf.*, May 2010, pp. 287–291.

[5] J. Romberg, "Imaging via compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 14–20, March 2008.

[6] Y. Wang, Z. Tian, and C. Feng, "Sparsity order estimation and its application in compressive spectrum sensing for cognitive radios," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 6, pp. 2116–2125, June 2012.

[7] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse bayesian inference," *Signal Processing, IEEE Transactions on*, vol. 61, no. 1, pp. 38–43, Jan 2013.

[8] Y. Yu, A. Petropulu, and H. Poor, "Cssf mimo radar: Compressive-sensing and step-frequency based mimo radar," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 48, no. 2, pp. 1490–1504, APRIL 2012.

[9] R. Prasad and C. Murthy, "Cramer-rao-type bounds for sparse bayesian learning," *Signal Processing, IEEE Transactions on*, vol. 61, no. 3, pp. 622–632, Feb 2013.

[10] E. Candes and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, Dec 2005.

[11] S. Gleichman and Y. Eldar, "Blind compressed sensing," *Information Theory, IEEE Transactions on*, vol. 57, no. 10, pp. 6958–6975, Oct 2011.

[12] E. Candes and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, March 2008.

[13] Y. Yu, A. Petropulu, and H. Poor, "Mimo radar using compressive sampling," *IEEE Jour. Select. Topics Sig. Process.*, vol. 4, no. 1, pp. 146–163, Feb 2010.

[14] R. Baraniuk, "Compressive sensing [lecture notes]," *Signal Processing Magazine, IEEE*, vol. 24, no. 4, pp. 118–121, July 2007.

[15] Y. Yu, A. Petropulu, and H. Poor, "Measurement matrix design for compressive sensing based mimo radar," *Signal Processing, IEEE Transactions on*, vol. 59, no. 11, pp. 5338–5352, Nov 2011.

[16] A. Ramezani-Kebrya, I.-M. Kim, D. I. Kim, F. Chan, and R. Inkol, "Likelihood-based modulation classification for multiple-antenna receiver," *Communications, IEEE Transactions on*, vol. 61, no. 9, pp. 3816–3829, September 2013.

[17] J. Lunden and V. Koivunen, "Automatic radar waveform recognition," *IEEE Jour. Selec. Top. Sig. Process.*, vol. 1, no. 1, pp. 124–136, June 2007.

[18] G. Lopez-Risueno and J. Grajal, "Multiple signal detection and estimation using atomic decomposition and em," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 1, pp. 84–102, Jan 2006.

[19] G. Lopez-Risueno, J. Grajal, and A. Sanz-Osorio, "Digital channelized receiver based on time-frequency analysis for signal interception," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 41, no. 3, pp. 879–898, July 2005.

[20] H. Jiang, W. Guan, and L. Ai, "Specific radar emitter identification based on a digital channelized receiver," in *Image and Signal Processing (CISP), 2012 5th International Congress on*, Oct 2012, pp. 1855–1860.

[21] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Machine Learning Res.*, vol. 1, no. 7, pp. 211–244, 2001.

[22] D. Wipf and B. Rao, "Sparse bayesian learning for basis selection," *IEEE Trans. Sig. Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[23] ——, "An empirical bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Sig. Process.*, vol. 55, no. 7, pp. 3704–3716, Jul. 2007.

[24] Z. Zhang and B. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," *IEEE Jour. Selec. Sig. Process.*, vol. 5, no. 5, pp. 912–926, Sept. 2011.

[25] D. Shutin, T. Buchgraber, S. Kulkarni, and H. Poor, "Fast variational sparse bayesian learning with automatic relevance determination for superimposed signals," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 6257–6261, Dec 2011.

[26] O. Balkan, K. Kreutz-Delgado, and S. Makeig, "Localization of more sources than sensors via jointly-sparse bayesian learning," *Signal Processing Letters, IEEE*, vol. 21, no. 2, pp. 131–134, Feb 2014.

[27] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse bayesian learning for recovery of block-sparse signals," *Signal Processing, IEEE Transactions on*, vol. 63, no. 2, pp. 360–372, Jan 2015.

[28] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 53, no. 7, pp. 2477–2488, July 2005.

[29] M. A. Richards, *Fundamentals of Radar Signal Processing, Second Edittion*. McGraw-Hill, New York, 2014.

[30] O. Bar-Ilan and Y. Eldar, "Sub-nyquist radar via doppler focusing," *IEEE Trans. Sig. Process.*, vol. 62, no. 7, pp. 1796–1811, Apr. 2014.

[31] P. Ciblat, P. Loubaton, E. Serpedin, and G. Giannakis, "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Trans. Informa. Theo.*, vol. 48, no. 7, pp. 1922–1934, Jul 2002.

[32] K. Hwang and S. Choi, "Blind equalization method based on sparse bayesian learning," *IEEE Sig. Process. Let.*, vol. 16, no. 4, pp. 315–318, Apr. 2009.

[33] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 5, pp. 415–447, Jul. 1992.

[34] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. Wiley, New York, 1997.

[35] E. J. Candes 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.

[36] D. G. Luenberger, *Linear and Nonlinear Programming, Second ed. Reading*. MA: Addison-Wesley, 1984.

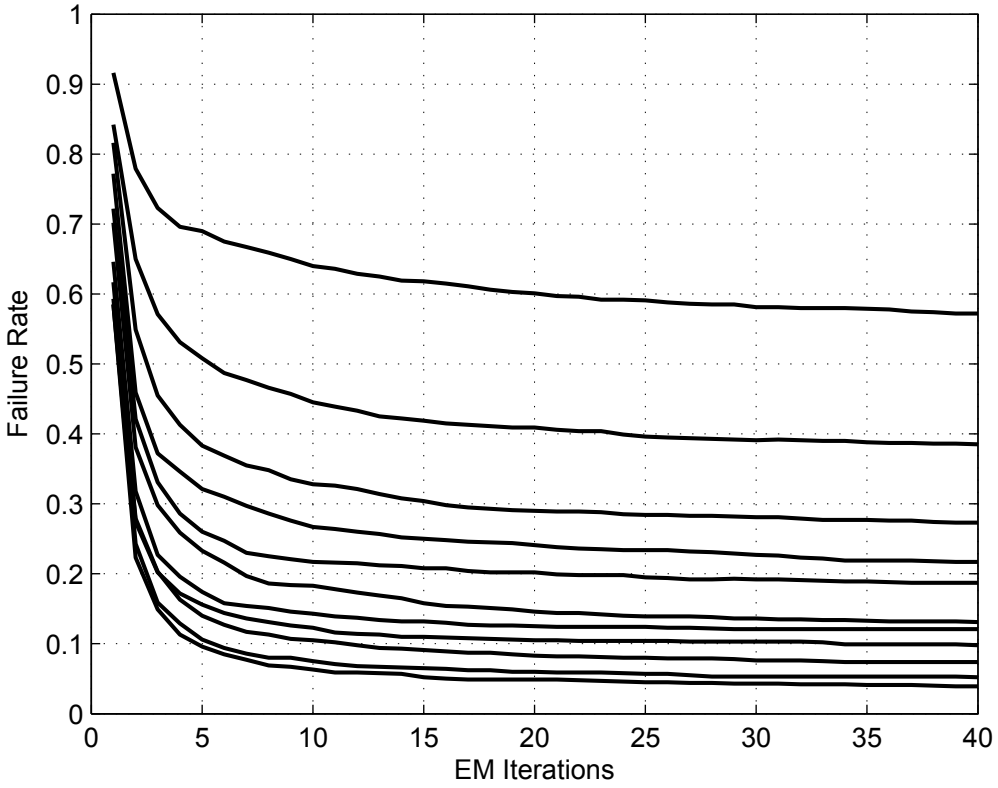

Fig. 1. Failure Rate performance versus the number of EM iterations when SNR = 4 (dB).

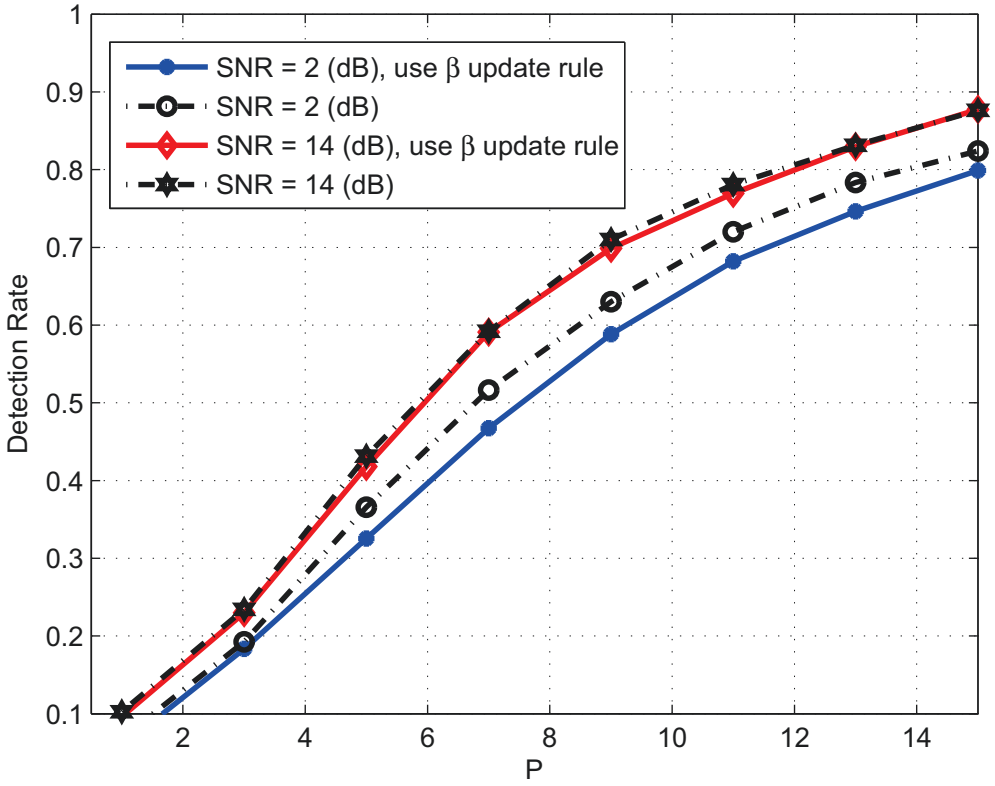

Fig. 2. Detection Rate performance as a function of P.

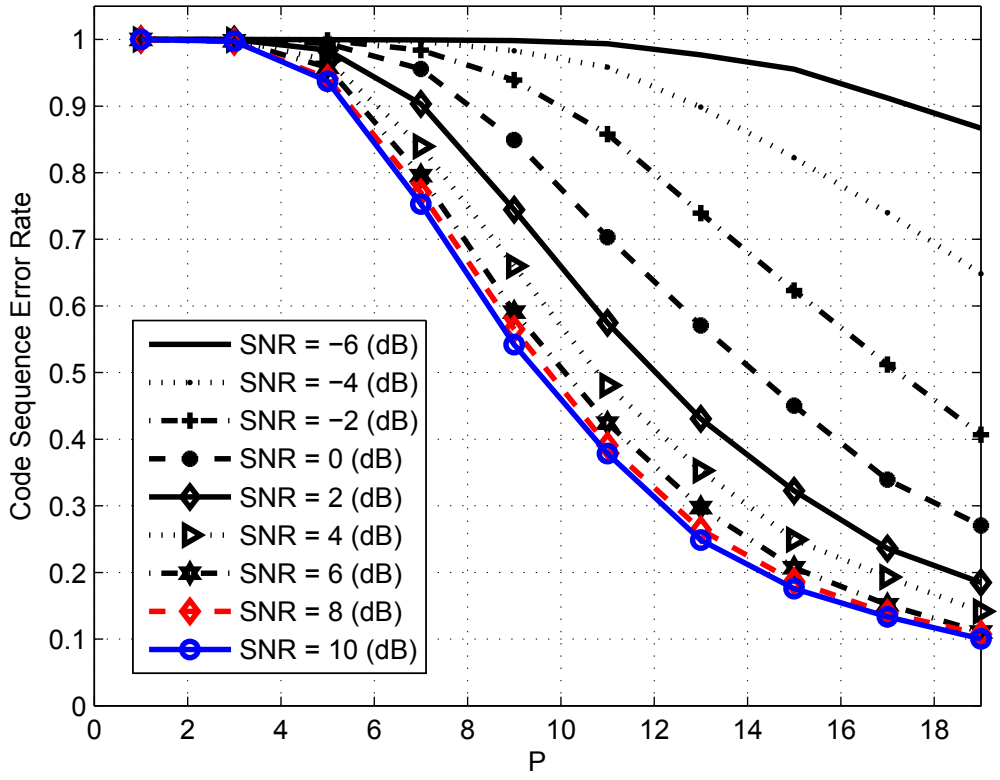

Fig. 3. Code Sequence Error Rate performance as a function of P.

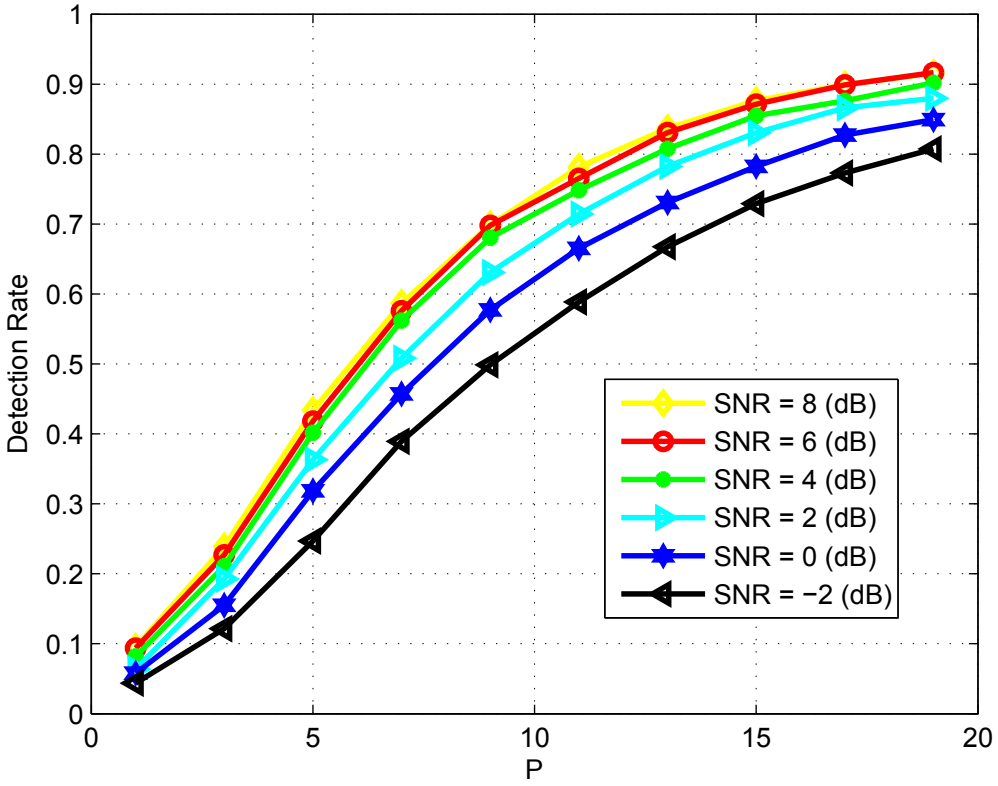

Fig. 4. Detection Rate performance of the proposed Blind-SBL framework as a function of P.

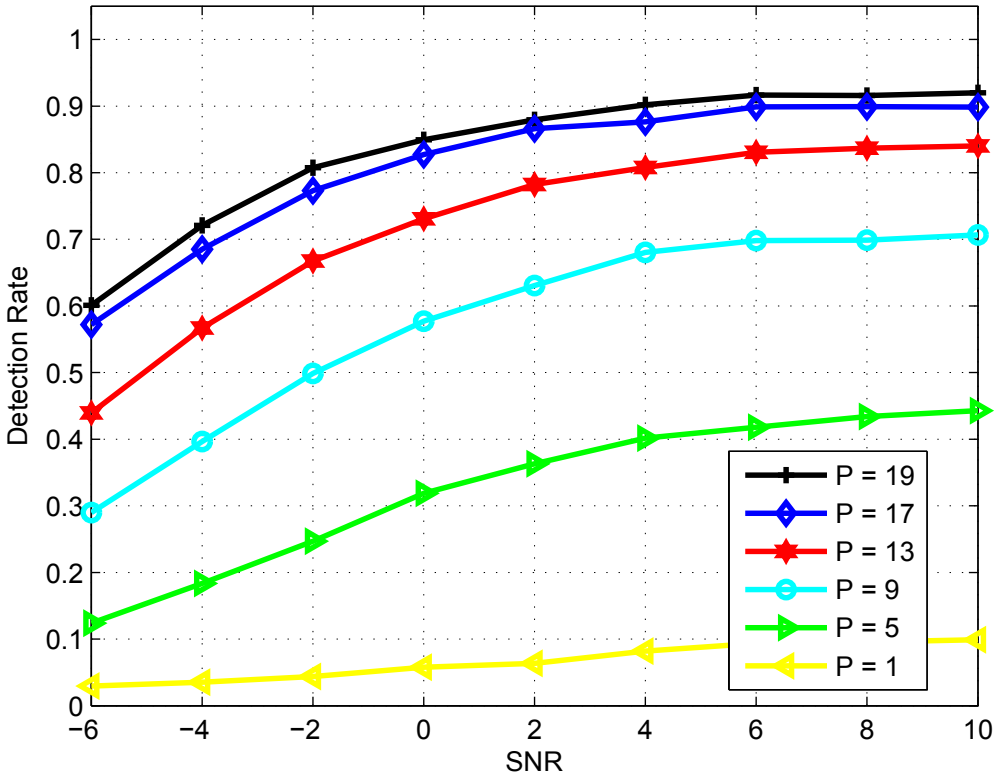

Fig. 5. Detection Rate performance of the proposed Blind-SBL framework versus SNR.

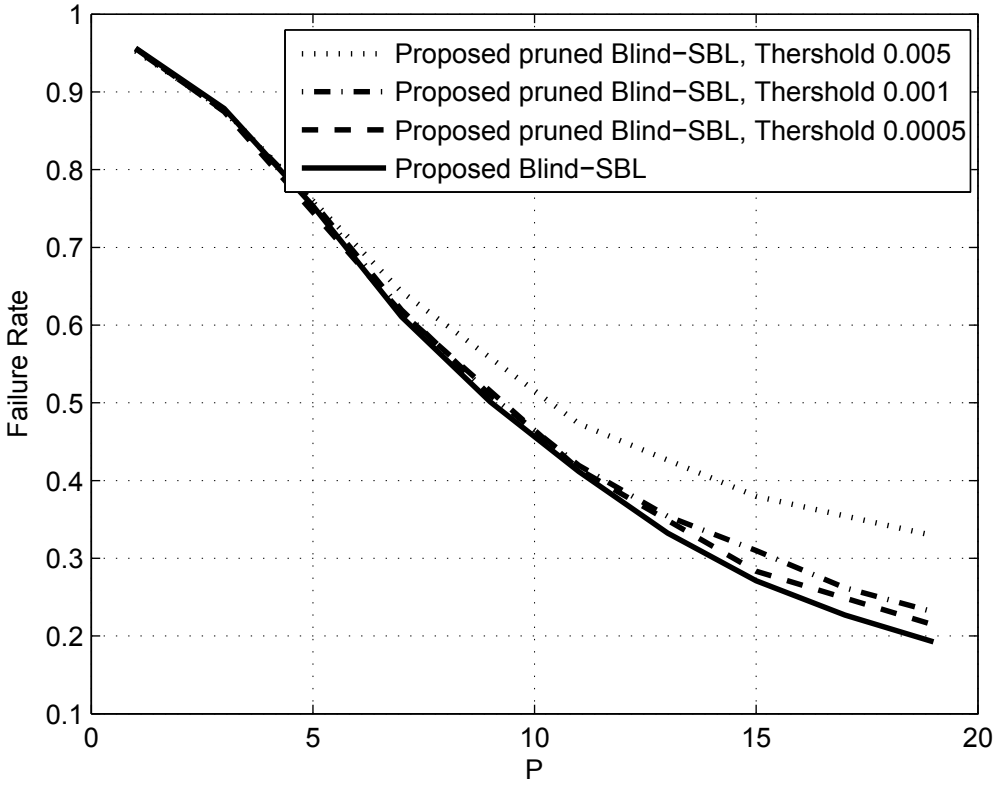

Fig. 6. Performance comparison of Blind-SBL and Pruned Blind-SBL algorithms for SNR = -2 (dB).