Maximum a posteriori based approach for target detection in MTI radar

Md Mashud Hyder and Kaushik Mahata

***Abstract*—We propose a sparse recovery approach to detect moving targets in clutter. In presence of clutter, the target space is not sparse. We propose a simple way to estimate the clutter region. We then enforce sparsity by modeling the clutter as a single extended cluster of nonzero components. This done by solving a sparse signal recovery problem with partially known support within a maximum a posteriori (MAP) estimation framework. The resulting algorithm is applied in angle-Doppler imaging for moving target indication (MTI) in an airborne radar. Our approach has a number of advantages including improved robustness to noise and increased resolution with limited data.**

***Index Terms*—Sparse representation, airborne radar, maximum a posteriori estimation, angle-Doppler imaging**

I. Introduction

Detection of targets in presence of non-stationary clutter and noise is a major problem in airborne Ground-Moving-Target-Indicator (GMTI) radar systems [1], [2]. A space-time adaptive processing (STAP) [2] can, in principle, be used to mitigate the clutter. STAP relies on the clutter-and-noise covariance matrix, and thus, requires sufficiently long record of target-free training data, see *e.g.* [1], [2]. Moreover, since the target statistics are not considered in the covariance estimation, STAP is suboptimal in presence of multiple targets. These issues have motivated knowledge-aided STAP [3], [4]. However, the clutter return is often non-stationary in complex propagation conditions and highly variable backscatter. This non-stationarity results performance degradation of STAP.

Methods [5]–[8] aimed at mitigating above issues have their own limitations. The joint-domain localized method [5] suffers from broad main beam (smearing) and high side-lobe level (leakage) problems. Spatially adaptive detectors [6] can handle spatially structured interferences, but can not perform well in Doppler spread clutter. Maximum-likelihood (ML) based adaptive temporal processing [7] requires multiple snapshots of the radar data, and thus, gets affected by non-stationarity of clutter return across the range of interest. In [8] an abruptly time-varying autoregressive (ATVAR) model is used to accommodate both slowly varying and rapidly varying clutter statistics during an intra-coherent processing interval (CPI). The targets are detected via a generalized likelihood ratio test (GLRT) based on an approximate ML estimate of the clutter covariance. However, this approach assumes certain clutter signal features, which are often quite restrictive in practice [9] due to uncertainties in clutter motion.

Some high resolution angle-Doppler imaging algorithms process the data in each range bin of interest (also known as primary data) independently [10], [11], and do not need training data to extract clutter statistics. The iterative adaptive approach (IAA) [10], [12] can handle both stationary and non-stationary clutter. However, IAA's performance degrades when the target signal-to-noise ratio (SNR) is low. Moreover, IAA does not do well in presence of steering vector error. This issue is addressed in [13], [14], although these methods are not directly applicable to angle-Doppler imaging. Sparse learning via iterative minimization (SLIM) [15] has been applied for MIMO radar imaging [15], and target detection in presence of clutter [16]. The idea here is to remove the clutter ridge from the data, and subsequently use SLIM iterations for target detection. In our experience (see Figure 3), SLIM's performance degrades when signal clutter to noise ratio (SCNR) is small.

Sparse signal recovery approaches like [17], [18] are not useful in presence of a large clutter, as the target space is no longer sparse. To deal with this issue, we estimate the clutter region via a simple method. Subsequently, we detect the targets by solving a "sparse recovery problem with partially known support" [19]. Here we use our crude estimate of the clutter region as the "partialy known support". The prior knowledge about the support allows us to reconstruct a sparse signal from significantly less number of samples [19]–[21]. However, the previous algorithms [19]–[21] need fairly accurate knowledge of the support. Since we don't have such accurate prior estimate of the clutter ridge, an improved algorithm is needed, see Figure 3 for details. We acomplish this by using a MAP estimation approach. Suppose, the unknown signal is x. We set up *a priori* density function for x, such that a component of x within the known support is of large magnitude with a high probability, and a component of x outside the known support is of large magnitude with a very small probability. We propose numerical strategies to solve the resulting optimization problem. Subsequently, we give a new method to correct steering vector errors in practical MTI radar. The correction approach is equally applicable to other conventional algorithms like, IAA [10]. In the experiments, the proposed approach exhibits a number of advantages over other angle-Doppler imaging techniques, which include increased resolution, and improved robustness to noise. In addition, it does not require an accurate initialization.

Authors are with the Department of Electrical Engineering, University of Newcastle, Callaghan, NSW-2308, Australia. (email: md.hyder@uon.edu.au, kaushik.mahata@newcastle.edu.au).

Research is supported by the Australian research council. Part of this work has been included in the PhD thesis of Md Mashud Hyder: http://ogma.newcastle.edu.au:8080/vital/access/manager/Index

II. Problem Formulation

Consider the standard setting [1], [2], where an airborne radar uses a M element sensor array. It transmits N pulses during a coherent processing interval (CPI) at a pulse repetition interval H. The transmitted signals are reflected by the targets, and the reflections are received. Let the whole range of interest be divided into R uniform range-bins. For a fixed elevation angle, the location of the t-th target can be specified by its range index i, azimuth angle $\bar{f}_t$, and normalized Doppler frequency $\bar{\omega}_t$. In the following, we assume that the data due to the return for each range bin is processed independently. Hence we drop the range bin index. Then the t-th target is indicated by an angle-Doppler pair $(\bar{f}_t, \bar{\omega}_t)$. We assume there is no range and Doppler frequency ambiguities. This is a common assumption [10].

We use an overcomplete representation of the signal received by the radar antenna array. We set up a uniform grid over the whole azimuth angle range consisting of T points. Similarly, we set up a L point uniform grid over the Doppler frequency range. Let $X \in \mathbb{C}^{T\times L}$ be a matrix such that $X_{t,\ell}$ (*i.e.* the element at t-th row and ℓ-th column) denotes the complex amplitude of the returned signal due to the reflection from a target located at (f_t, ω_ℓ). Every $X_{t,\ell}$ is assumed to remain constant during a CPI and is proportional to the radar-cross-section of the corresponding target. The data vector $\mathbf{y}(j)$ obtained by matched filtering and digitizing the received signal due to the j-th pulse can be modeled as [10], [11]

$$\mathbf{y}(j) = \sum_{t=1}^{T}\sum_{\ell=1}^{L} \phi(f_t) X_{t,\ell} \mathrm{e}^{\mathrm{i}2\pi(j-1)\omega_\ell} + \mathbf{e}(j), \qquad (1)$$

where $\mathbf{y}(j) \in \mathbb{C}^{M\times 1}$, $\mathbf{e}(j)$ accounts for additive noise, and $\phi(f_t) \in \mathbb{C}^{M\times 1}$ denotes the array steering vector corresponding to an azimuth angle f_t, which depends on the array geometry. Defining $\Phi = [\ \phi(f_1) \quad \cdots \quad \phi(f_T)\ ]$, we get

$$Y = [\ \mathbf{y}(1) \quad \cdots \quad \mathbf{y}(N)\ ] = \Phi X \Theta + E. \qquad (2)$$

Here $E = [\ \mathbf{e}(1) \quad \cdots \quad \mathbf{e}(N)\ ]$; the matrix Θ is given as

$$[\Theta]_{\ell,j} = \mathrm{e}^{\mathrm{i}2\pi\omega_\ell(j-1)}: \quad j = 1, \cdots, N,$$

and $\theta(\omega_\ell) = [1 \quad \mathrm{e}^{\mathrm{i}2\pi\omega_\ell} \cdots \mathrm{e}^{\mathrm{i}2\pi\omega_\ell(N-1)}]$ indicates one row corresponding to Doppler ω_ℓ. One uses X to visualize the azimuth angle and Doppler shifts of the targets in form of an image, with each $X_{t,\ell}$ representing a pixel. A non-zero $X_{t,\ell}$ represents a target at the t-th azimuth angle and the ℓ-th Doppler bin. As we design the grid, the matrices Φ, Θ are known and do not depend on the actual target location $(\bar{f}_t, \bar{\omega}_t)$.

Note that for any given t, the true attribute $(\bar{f}_t, \bar{\omega}_t)$ may not lie exactly on a grid-point of our azimuth-Doppler grid. In other words, there may not exist i and j satisfying $\bar{f}_t = f_i$ and $\bar{\omega}_t = \omega_j$ exactly. Nevertheless, by making the grid dense enough one can ensure $\bar{f}_t \approx f_i$ and $\bar{\omega}_t \approx \omega_j$ for some i and j, and the remaining model error is absorbed by the residual term E. Suppose there are P scatterers. Then only P pixels of X are nonzero, i.e. $X_{i,j} \neq 0$ only if $f_i \approx \bar{f}_t$ and $\omega_j \approx \bar{\omega}_t$ for some $t \in \{1, 2, \ldots, P\}$. In effect, (2) allows us to pose the problem of estimating $\{(\bar{f}_k, \bar{\omega}_t)\}_{t=1}^{P}$ as that of estimating the sparse matrix X.

III. Proposed Reconstruction Approach

In presence of a big clutter consisting of numerous point scatterers, X can no longer be considered sparse, and methods like [11] cannot be used. We propose a two step algorithm to deal with this issue. First we approximately identify the clutter location. Then we apply a MAP based sparse recovery approach to locate the moving targets.

A. Step 1 - Approximate Localization of the Clutter Spread

Consider a stationary scatterer within the clutter located at an azimuth angle f relative to the array. Then for a side-looking airborne radar and small carb angle, the normalized Doppler shift ω induced on the scatterer is given by [2]

$$\omega = \frac{2v_{air}T}{\gamma}\sin(f) =: \beta\hat{f}, \qquad (3)$$

where $\hat{f} = \sin(f)$; v_{air} is the ragar-speed; γ is the operating wavelength. However, clutter has intrinsic motion resulting in random Doppler spreads causing the clutter energy to "leak" in both sides of the line $\omega = \beta\hat{f}$ along the ω direction.

Given the a priori information in form of (3), we use the principles of MAP to device a fast algorithm to roughly estimate the clutter spread. Since the clutter to signal ratio is generally above 30 dB in practice [22], at this stage we assume that the non-zero components in X are due to the clutter only. We assume that the a priori density of $X_{t,\ell}$ is a Gaussian density with mean zero and variance $\sigma^2_{t,\ell}$, and X_{t_1,ℓ_1} is independent of X_{t_2,ℓ_2} if $t_1 \neq t_2$ and/or $\ell_1 \neq \ell_2$. The fact that the clutter is concentrated around the line $\omega = \beta\hat{f}$ is used to specify the value $\sigma^2_{t,\ell}$. We set

$$\sigma^2_{t,\ell} = \mathrm{e}^{-\kappa(\omega_\ell - \beta\hat{f}_t)^2}.$$

where κ is a constant whose value depends on the clutter spread. In practice, the variance of the Doppler spread due to intrinsic motion in the clutter is known from physical experiments [23], [24]. With the assumption on the *a priori* density of X, the MAP estimate of X is given by [25]

$$X^{(0)} = \arg\min_{X} \sum_{t=1}^{T}\sum_{\ell=1}^{L} \frac{|X_{t,\ell}|^2}{2\sigma^2_{t,\ell}} \quad \text{subject to } Y = \Phi X \Theta. \qquad (4)$$

Here we assume the variance of the elements of E are negligible compared to the clutter energy, which is quite appropriate in practice. The solution to (4) can be given analytically. Let

$$\Sigma = \mathrm{diag}\left\{ \mathrm{vec}\left(\begin{bmatrix} \sigma^2_{1,1} & \cdots & \sigma^2_{1,L} \\ \vdots & & \vdots \\ \sigma^2_{T,1} & \cdots & \sigma^2_{T,L} \end{bmatrix} \right) \right\},$$

$\mathbf{x}^{(0)} = \mathrm{vec}(X^{(0)})$, $\mathbf{y} = \mathrm{vec}(Y)$, $\Psi = \Theta' \otimes \Phi$. Then

$$\mathbf{x}^{(0)} = \Sigma\Psi^*(\Psi\Sigma\Psi^*)^{-1}\mathbf{y}. \qquad (5)$$

Since the values of $\sigma^2_{t,\ell}$ in (4) are just guesses, one would expect somewhat better estimation performance if we repeat the estimation with better estimates of $\sigma^2_{t,\ell}$. This suggests an

iterative procedure where we compute the i-th iterate $\hat{\mathbf{x}}^{(i)} = \mathrm{vec}(X^{(i)}), i > 0$ as

$$\hat{\mathbf{x}}^{(i)} = \Sigma_i \Psi^* (\Psi \Sigma_i \Psi^*)^{-1} \mathbf{y}, \tag{6}$$

where for $i \geq 0$ we define

$$\Sigma_{i+1} = \mathrm{diag}\left\{ \mathrm{vec}\left(\begin{bmatrix} |X_{1,1}^{(i)}|^2 & \cdots & |X_{1,L}^{(i)}|^2 \\ \vdots & & \vdots \\ |X_{T,1}^{(i)}|^2 & \cdots & |X_{T,L}^{(i)}|^2 \end{bmatrix} \right) \right\}. \tag{7}$$

The iterative approaches like above are quite popular in signal processing, and are often called the iteratively re-weighted least squares (IRLS) [26]. Also the use of (7) is quite common, see *e.g.* [12]. In particular, the FOCal Underdetermined System Solver (FOCUSS) [27] for sparse recovery, and SLIM use a similar approach. In general, the $t+1$ th iterate $\boldsymbol{x}^{(t+1)}$ in FOCUSS [27] or SLIM [15] has the form

$$\boldsymbol{x}^{(t+1)} = \mathbf{W}_{t+1} \Psi^* \left[\Psi \mathbf{W}_{t+1} \Psi^* + \varsigma \mathbf{I} \right]^{-1} \boldsymbol{y}, \tag{8}$$

where $\mathbf{W}_{t+1} = \mathrm{diag}\{|\boldsymbol{x}_n^{(t)}|^{2-p}\}$ and $0 < p \leq 1$. There are few differences between the proposed IRLS and FOCUSS/SLIM. Here we use $p = 0$ and $\varsigma = 0$. Also we take the advantage of initial approximation of $\boldsymbol{x}$ in (5). Finally, unlike FOCUSS/SLIM, we do not wait for convergence of the proposed IRLS algorithm. Our goal is to get an approximate knowledge of the clutter support, and in our experience we get a result of desired accuracy in about three iterations. Figure 1 illustrates a typical result. The angle-Doppler image of a typical clutter (considering $\beta = 0.5$ in (3)) is shown in Figure 1 (a). The recovered angle-Doppler image estimated after three IRLS iterations is shown in Figure 1(b).

We need a criterion to determine the clutter spread from the angle-Doppler image (Figure 1(b)) obtained using the MAP/IRLS approach described above. In that goal, let us define the cumulative distribution function

$$g(\delta) := \sum_{t=1}^{T} \left\{ \sum_{\ell \in \mathbb{L}_\delta(t)} |X_{t,\ell}^{(3)}|^2 \right\}, \tag{9}$$

where $\mathbb{L}_\delta(t) = \{\ell : \omega_\ell < \beta \hat{f}_t + \delta\}$. Figure 1(c) shows $g(\delta)$ as a function of δ for angel-Doppler images obtained by different algorithms. As expected, $g(\delta)$ increases rapidly around the clutter region i.e., around $\delta = 0$. Note that IRLS provides a closer estimation of clutter energy compared to DAS [5]. Now to approximate the clutter region we have to determine a band around the line $\omega = \beta \hat{f}$ so that most of the energy lie within this band. Hence we set a small $\varepsilon > 0$, and then find the clutter region as

$$\mathbb{I}_\varepsilon = \left\{ (t,\ell) : \varepsilon \leq \frac{g(\omega_\ell - \beta f_t)}{\sum_{t=1}^{T} \sum_{\ell=1}^{L} |X_{t,\ell}^{(3)}|^2} \leq 1 - \varepsilon \right\}. \tag{10}$$

In the experiments, we set $\varepsilon = 0.02$.

B. Step 2 - Sparse Recovery Algorithm for Target Detection

1) MAP Formulation: In absence of clutter, X is a sparse matrix, with non-zero entries associated with a small number of targets. This has been the main motivation behind several radar imaging algorithms based in compressed sensing [17], [18]. However in presence of a clutter, so far there is no known way to exploit the sparsity of the targets. In the following we propose a statistical framework to address this issue.

The matrix X has two types of elements. If $(t,\ell) \in \mathbb{I}_\varepsilon$, then we know that this point lies within the clutter support. Hence there is a point scatterer in this point with a high probability, and hence $|X_{t,\ell}|$ is large with high probability. On the other hand if $(t,\ell) \notin \mathbb{I}_\varepsilon$, then we know that this point lies outside the clutter support. Hence there is a target in this point with a quite low probability, and hence $|X_{t,\ell}|$ is large with a low probability. Such a-priori information is often used by modeling the a priori density function of X as a Gaussian mixture model. To characterize the two types of elements in X, we use two positive numbers ϱ and $\hat{\mu}$ with $\varrho \gg \hat{\mu}$. Then we assume that $X_{t,\ell}$ is a random variable which draws its value either from a complex Gaussian density of zero mean and standard deviation[1] ϱ with probability $p_{(t,\ell)}$; or from another complex Gaussian density of zero mean and standard deviation $\hat{\mu}$ with probability $1 - p_{(t,\ell)}$. In other words, the probability density function of $X_{t,\ell}$ is given by

$$h(X_{t,\ell}, p_{t,\ell}) := p_{t,\ell} \mathcal{N}(X_{t,\ell}, \varrho) + (1 - p_{t,\ell}) \mathcal{N}(X_{t,\ell}, \hat{\mu}), \tag{11}$$

where

$$\mathcal{N}(x, \varpi) = \frac{1}{2\pi\varpi^2} \exp\left\{ -\frac{|x|^2}{2\varpi^2} \right\}. \tag{12}$$

In order to distinguish $\mathbb{I}_\varepsilon$ from its complement, we assign

$$p_{(t,\ell)} = \begin{cases} \tau_1, & (t,\ell) \in \mathbb{I}_\varepsilon \\ \tau_2, & (t,\ell) \notin \mathbb{I}_\varepsilon. \end{cases} \tag{13}$$

with $\tau_1 \gg \tau_2$. This reflects the fact that $|X_{t,\ell}|$ is large with a high probability if $(t,\ell) \in \mathbb{I}_\varepsilon$. Otherwise, it is large with a very small probability. Since τ_2 is the probability of finding a target at a point, its value is determined from the prior knowledge about the experimental setting. On the other hand τ_1, denotes the probability of finding a scatterer in the clutter, which is very close to 1. For instance, $\tau_1 = 0.99$ and $\tau_2 = 0.01$ is considered for the experimental settings in the sequel. Note that the values of τ_1 and τ_2 need not be accurate, see [28] for details.

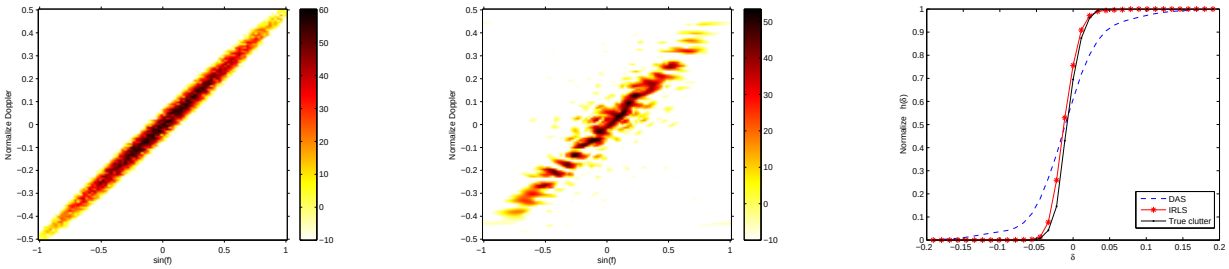
As before, we assume that X_{t_1,ℓ_1} is independent of X_{t_2,ℓ_2} if $t_1 \neq t_2$ and/or $\ell_1 \neq \ell_2$. Hence the negative log likelihood of X is given by

$$-\sum_{t=1}^{T} \sum_{\ell=1}^{L} \ln[h(X_{t,\ell}, p_{t,\ell})].$$

Assume that the elements of E in (2) are mutually independent and identically distributed as zero mean Gaussian distribution with a variance $1/\lambda$. Then using (2) the joint negative log-likelihood of X and Y is given by

$$\wp(X,Y) = \frac{\lambda}{2} ||Y - \Phi X \Theta||_F^2 - \sum_{t=1}^{T} \sum_{\ell=1}^{L} \ln[h(X_{t,\ell}, p_{t,\ell})], \tag{14}$$

[1]When we say a complex valued random variable $x = x_1 + \mathrm{i}x_2$ is complex Gaussian with a standard deviation ϱ, we mean that x_1 and x_2 are independent and each of them has a variance ϱ^2.



(a) (b) (c)

Fig. 1. Clutter detection. (a) True clutter, (b) IRLS, (c) Normalize $h(\delta)$ vs δ of the angle-Doppler image by different algorithms.

where $\|.\|_F$ indicates Frobenius norm. In order to estimate X, we propose to minimize $\wp(X,Y)$ with respect to X. This approach is same as the maximum a posteriori estimation idea. Since, we minimize (14) with respect to X, we write $\wp(X)$ instead of $\wp(X,Y)$ for simplicity.

2) Optimization algorithm: Using (12) and (11) we have

$$-\ln[h(X_{t,\ell},p_{t,\ell})] = \frac{X_{t,\ell}^2}{2\varrho^2} - \ln\left[1 + r_{t,\ell}\exp\left\{-\frac{X_{t,\ell}^2}{2\hat{\sigma}^2}\right\}\right] \tag{15}$$

up to a constant independent of X, where

$$r_{t,\ell} = \frac{(1-p_{t,\ell})\varrho^2}{p_{t,\ell}\ \hat{\mu}^2}, \qquad \frac{1}{\hat{\sigma}^2} = \frac{1}{\hat{\mu}^2} - \frac{1}{\varrho^2}. \tag{16}$$

Using (15) and (14) we get

$$\wp(X) = \sum_{t=1}^{T}\sum_{\ell=1}^{L}\left\{\frac{|X_{t,\ell}|^2}{2\varrho^2} - \ln\left[1 + r_{(t,\ell)}\exp\left(-\frac{|X_{t,\ell}|^2}{2\hat{\sigma}^2}\right)\right]\right\} + \frac{\lambda}{2}\|Y-\Phi X\Theta\|_F^2. \tag{17}$$

Like many optimization problems encountered in CS literature [29], [30], (17) is a nonconvex problem, and there is no known numerical algorithm with guaranteed convergence properties. To deal with the nonconvex optimization problem we adopt a sequential optimization approach where we construct a sequence of functions $\wp_j(X), j = 0,1,2,\ldots,w$ such that

- $\wp_0(X)$ is quadratic;
- $|\wp_{j+1}(X) - \wp_j(X)|$ is small in the neighbourhood of the minimizer $\hat{X}^{(j)}$ of $\wp_j(X)$;
- $\wp_w(X) = \lambda\wp(X)$ for a user chosen integer w.

Because $\wp_0(X)$ is quadratic we shall be able to compute $\hat{X}^{(0)}$ using the standard analytical expression. Then as $|\wp_1(X) - \wp_0(X)|$ is small in the neighborhood of $\hat{X}^{(0)}$, by initializing a numerical algorithm to minimize $\wp_1(X)$ at $\hat{X}^{(0)}$ one has a high probability of converging to $\hat{X}^{(1)}$. If we continue this process of initializing the numerical algorithm to optimize $\wp_{j+1}(X)$ at our estimate of $\hat{X}^{(j)}$ obtained by numerically optimizing $\wp_j(X)$, then one can expect that $\hat{X}^{(w)}$ is likely to be the minimizer of $\wp_w(X)$.

We construct $\{\wp_j(X)\}_{j=0}^{w}$ as follows. We choose appropriate real numbers $\hat{\sigma}_1$ and λ_1 (more details on how the choices are made will follow shortly), and define

$$\begin{aligned}\wp_0(X) &= \frac{1}{2}\|Y-\Phi X\Theta\|_F^2;\\ \wp_j(X) &= \tilde{\wp}_j(X) + \frac{1}{2}\|Y-\Phi X\Theta\|_F^2, \qquad j = 1,2,\ldots,w,\end{aligned}$$

where

$$\tilde{\wp}_j(X) = \frac{1}{\lambda_j}\sum_{t=1}^{T}\sum_{\ell=1}^{L}\left\{\frac{|X_{t,\ell}|^2}{2\varrho^2} - \ln\left[1 + r_{(t,\ell)}\mathrm{e}^{|X_{t,\ell}/\hat{\sigma}_j|^2/2}\right]\right\}, \tag{18}$$

where $\hat{\sigma}_j = (\hat{\sigma}/\hat{\sigma}_1)^{j/w}\hat{\sigma}_1$, and $\lambda_j = (\lambda/\lambda_1)^{j/w}\lambda_1$. We take $\hat{\sigma}_1$ and λ_1 quite large positive numbers. By taking λ_1 large, we ensure that $|\wp_1(X) - \wp_0(X)|$ is small. The subsequent changes $|\wp_{j+1}(X) - \wp_j(X)|$ for $j \geq 1$ depends on w.

The parameter $\hat{\sigma}_j$ controls the degree of nonconvexity in $\tilde{\wp}_j$ (and $\wp_j$). If we take $\hat{\sigma}_1 \to \infty$, then the logarithmic term in (18) tends to $\ln(1 + r_{(t,\ell)})$, making $\tilde{\wp}_1(x) = \sum_{t=1}^{T}\sum_{\ell=1}^{L}[|X_{t,\ell}|^2/(2\varrho^2) - \ln(1 + r_{(t,\ell)})]$, a quadratic function in X. In practice, we take

$$\hat{\sigma}_1 \geq 5\max_{(t,\ell)}|\hat{X}_{t,\ell}^{(0)}|,$$

This ensures $\exp\{-|\hat{X}_{t,\ell}^{(0)}|^2/(2\hat{\sigma}_1^2)\} \geq 0.99$ for all t,ℓ. Consequently, $\exp\{-|X_{t,\ell}|^2/(2\hat{\sigma}_1^2)\} \approx 1$ for all X satisfying $||X - \hat{X}^{(0)}||_F < ||\hat{X}^{(1)} - \hat{X}^{(0)}||_F$ [31].

By increasing j form 1 to w, we gradually transform $\wp_j$ from a convex function $\wp_0$ to the likelihood function $\wp_w$. If w is sufficiently large, then the change from $\wp_{j-1}$ to $\wp_j$ is small, and so is the change from $\hat{X}^{(j-1)}$ to $\hat{X}^{(j)}$. A Gauss-Newton type convex-concave procedure is used to minimize $\wp_j$ for a fixed $\hat{\sigma}_j$ which builds on the following Lemma.

Lemma 1. *Define the mapping* $\zeta_j : \mathbb{C}^{T\times L} \to \mathbb{C}^{T\times L}$ *such that*

$$\mathrm{vec}\{\zeta_j(X)\} = \lambda_j\left[\Upsilon_j/\hat{\sigma}_j^2 + \lambda_j\Psi^*\Psi\right]^{-1}\Psi^*\mathbf{y}, \tag{19}$$

and Υ_j *is an* $S\times S$ *diagonal matrix defined as*

$$\Upsilon = \mathrm{diag}\left\{\mathrm{vec}\left(\begin{bmatrix}\xi_j(1,1) & \cdots & \xi_j(1,L)\\ \vdots & & \vdots\\ \xi_j(T,1) & \cdots & \xi_j(T,L)\end{bmatrix}\right)\right\},$$

TABLE I
MAP BASED ALGORITHM FOR TARGET DETECTION

Initialization
1. Set $\text{vec}(X) = \Psi^*(\Psi\Psi^*)^{-1}\mathbf{y}$.
2. Set $\hat{\sigma}_0 = 5\max_{t,\ell}|X_{t,\ell}|$.
3. Set $j = 1$, choose $\nu \in (0, 1]$, $\beta = 0.5$ and w.

repeat
4. Set $\kappa = 1$.
5. while $\wp_j[\kappa\zeta_j(X) + (1-\kappa)X] > \wp_j(X)$
   $\kappa = \beta\kappa$.
   end
6. $X_o = X$, and $X = \kappa\zeta_j(X) + (1-\kappa)X$
7. If $\frac{\|X - X_o\|_F}{\|X_o\|_2} < \nu$
   $j = j + 1$.
   end

while $j \leq w$.

where

$$\xi_j(t,\ell) = \frac{1}{\varrho^2} + \frac{r_{(t,\ell)}\exp\left(-\frac{|X_{t,\ell}|^2}{2\hat{\sigma}_j^2}\right)}{\hat{\sigma}_j^2\left[1 + r_{(t,\ell)}\exp\left(-\frac{|X_{t,\ell}|^2}{2\hat{\sigma}_j^2}\right)\right]}. \quad (20)$$

Then $\hat{X}^{(j)} = \zeta\{\hat{X}^{(j)}\}$. Furthermore, for any $X \in \mathbb{C}^{T\times L}$ there exists a real-valued scalar $\alpha \geq 0$ such that

$$L_\sigma[\alpha\zeta(X) + (1-\alpha)X] \leq L_\sigma(X). \quad (21)$$

Proof. The proof is similar to that of Lemma 1 in [32]. □

The major fraction of the computation is involved in computing $\zeta_j(X)$ in (19). However, using the matrix inversion lemma in (19), one can verify that

$$\text{vec}\{\zeta_j(X)\} = \Upsilon_j^{-1}\Psi^*\left[I/\lambda_j + \Psi\Upsilon_j^{-1}\Psi^*\right]^{-1}\mathbf{y}. \quad (22)$$

Using (22) one can reduce the computation time significantly.

The optimization strategy is summarized in Table I. The algorithm assumes that ϱ and $\hat{\mu}$ are known. In each iteration, along the decent-direction $\zeta_j(X) - X$ we find the step-length κ using the standard backtracking strategy (step 4-5) [33]. We set $\beta = 0.5$. The inner-iteration for updating X for a given j terminates when the relative change in the magnitude of X is below ν, see step 7. Upon convergence of each inner iteration we increment j (step 7). We found that in MAP, we can choose w between 15 and 40. Note that a smaller η and a larger w will increase the reliability in the cost of computation time. Our experimental study suggests that choosing $w = 20$, and $\nu = 0.02$ makes a good tradeoff. Upon convergence at $j = w$, MAP algorithm stops its iteration.

C. Steering Vector Error

So far we have assumed that the matrix Φ is known exactly. In this section, we consider the error in Φ due to imperfections in sensor positioning. We consider a generic 3 dimensional arrangement of an arbitrary sensor array. The nominal location of the i-th sensor is indicated by $\mathbf{z}^{(i)} = (z_1^{(i)}, z_2^{(i)}, z_3^{(i)})$. However, due to manufacturing errors, the actual sensor location may not be the same as the nominal locations. Let the coordinates of actual sensor locations be given by

$$\bar{\mathbf{z}}^{(i)} = \mathbf{z}^{(i)} + \boldsymbol{\delta}^{(i)}; \quad \text{for } i = 1, \cdots, M-1, \quad (23)$$

where $\boldsymbol{\delta}^{(i)} = (\delta_1^{(i)}, \delta_2^{(i)}, \delta_3^{(i)})$ denote the unknown error in positioning the i th sensor. Note that without any loss of generality we can set $\bar{\mathbf{z}}^{(0)} = \mathbf{z}^{(0)} = 0$.

In the following we distinguish between the true manifold matrix $\mathring{\Phi}$ associated with the true (but unknown) sensor positions $\{\bar{\mathbf{z}}^{(i)}\}_{i=0}^{M-1}$, and the nominal manifold matrix Φ corresponding to the nominal sensor positions $\{\mathbf{z}^{(i)}\}_{i=0}^{M-1}$, and characterize $\tilde{\Phi} := \mathring{\Phi} - \Phi$ in terms of $\{\boldsymbol{\delta}^{(i)}\}_{i=1}^{M-1}$.

The i-th row of the array manifold matrix in (2) is a function of the position of i-th sensor. In particular, we know the function γ so that we can write

$$\begin{aligned}\mathring{\Phi} &= [\ \gamma(\bar{\mathbf{z}}^{(0)}) \quad \cdots \quad \gamma(\bar{\mathbf{z}}^{(M-1)})\ ]^\intercal,\\ \Phi &= [\ \gamma(\mathbf{z}^{(0)}) \quad \cdots \quad \gamma(\mathbf{z}^{(M-1)})\ ]^\intercal. \end{aligned} \quad (24)$$

Now for small $||\boldsymbol{\delta}^{(i)}||_2$ we can neglect the higher order terms of the Taylor's series expansion of $\gamma(\bar{\mathbf{z}}^{(i)})$ in the neighborhood of $\mathbf{z}^{(i)}$ to get

$$\gamma(\bar{\mathbf{z}}^{(i)}) = \gamma(\mathbf{z}^{(i)}) + \delta_1^{(i)}\gamma_1(\mathbf{z}^{(i)}) + \delta_2^{(i)}\gamma_2(\mathbf{z}^{(i)}) + \delta_3^{(i)}\gamma_3(\mathbf{z}^{(i)}) \quad (25)$$

where, $\gamma_t(\mathbf{z}) = \frac{\partial\gamma(\mathbf{z})}{\partial z_t}$, $t = 1, 2, 3$. Now let

$$\Phi_t = [\ 0 \quad \gamma_t(\mathbf{z}^{(1)}) \quad \cdots \quad \gamma_t(\mathbf{z}^{(M-1)})\ ]^\intercal, \quad t = 1, 2, 3.$$

Then by (24) and (25) we have

$$\tilde{\Phi} = \mathring{\Phi} - \Phi = \Delta_1\Phi_1 + \Delta_2\Phi_2 + \Delta_3\Phi_3$$

where, $\Delta_t = \text{diag}(0 \ \ \delta_t^{(1)} \cdots \delta_t^{(M-1)})$. Consequently, we have a modified model

$$\begin{aligned} Y &= \mathring{\Phi}X\Theta + E = \{\Phi + \tilde{\Phi}\}X\Theta + E\\ &= \Phi X\Theta + \tilde{E} \end{aligned} \quad (26)$$

where, $\tilde{E} = \tilde{\Phi}X\Theta + E$. Thus if we knew $\tilde{E}$, then we could estimate $\{\Delta_t\}_{t=1}^3$ by solving a least squares problem

$$\min_{\Delta_1,\Delta_2,\Delta_3} \ ||\tilde{E} - (\Delta_1\Phi_1 + \Delta_2\Phi_2 + \Delta_3\Phi_3)X\Theta||_F^2. \quad (27)$$

In the following we assume $\mathsf{E}\{||E||_F^2\} = \mu^2$, where E is the mathematical expectation operator. Furthermore, we assume the sensor errors $\{\boldsymbol{\delta}^{(i)}\}_{i=1}^{M-1}$ are independent of E, and are mutually independent and identically distributed zero mean random vectors with a known covariance matrix $\nu^2 I$. Then it can be shown that

$$\chi^2 := \mathsf{E}\{||\tilde{E}||_F^2\} = \mu^2 + \nu^2\text{Tr}\{\Phi_1\Xi\Phi_1^* + \Phi_2\Xi\Phi_2^* + \Phi_3\Xi\Phi_3^*\} \quad (28)$$

where $\Xi = X\Theta\Theta^*X^*$. Using the observations made so far we now present an iterative approach which is able to correct the sensor positioning error while estimating X. This approach relies on tuning the parameter λ in (18). If we neglect the sensor positioning error then according to the concept in Section III-B2 we should set

$$\lambda = \chi^{-2}, \quad (29)$$

TABLE II
STEERING VECTOR ERROR APPROXIMATION

Initialization

1. Set $i = 0$, $\lambda^{(i)} = 1/\chi^2$, construct $\Phi, \{\Phi_t\}_{t=1}^3$ using nomimal $\{\mathbf{z}^{(i)}\}_{i=1}^{M-1}$.

repeat

2. Estimate X using algorithm in Table I with $\lambda = \lambda^{(i)}$, denote the estimate by $\hat{X}^{(i)}$.
3. Calculate the residual $\tilde{E}^{(i)} = Y - \Phi\hat{X}^{(i)}\Theta$.
4. Solve $\min_{\Delta_1,\Delta_2,\Delta_3} ||\tilde{E}^{(i)} - \sum_{t=1}^3 \Delta_t\Phi_t\hat{X}^{(i)}\Theta||_F^2$ to obtain estimates $\hat{\boldsymbol{\delta}}_i^{(1)} \cdots \hat{\boldsymbol{\delta}}_i^{(M-1)}$, and denote the associated estimate of $\tilde{\Phi}$ by $\tilde{\Phi}_i$.
5. Set $\mathbf{z}^{(j)} = \mathbf{z}^{(j)} + \hat{\boldsymbol{\delta}}_i^{(j)}$, for $j = 0, \cdots M-1$, and update $\Phi, \{\Phi_t\}_{t=1}^3$ accordingly.
6. Calculate χ^2 using (28) with $X = \hat{X}^{(i)}$.
7. Set $\lambda^{(i+1)} = 1/\chi^2$ and $i = i + 1$.

until $\sum_{j=1}^3 ||\boldsymbol{\delta}_i^{(j)}||^2 >$ a predefined threshold.

which is the starting value of λ used in our algorithm. Note that we need X to compute χ. For this purpose we use the solution obtained via the IRLS approach discussed in Section III-A. Using the value of λ in (29) we can obtain an initial estimate $\hat{X}^{(1)}$ which can be used to compute an estimate of $\tilde{E}$. This estimate of $\tilde{E}$ can be used to solve the optimization problem in (27) to estimate the sensor positioning errors. Next we can update the manifold matrix Φ by incorporating the estimates of $\{\boldsymbol{\delta}^{(i)}\}_{i=1}^{M-1}$, and the procedure can be repeated. The resulting iterative approach is outlined in Table II. The iterative algorithm in Table II is basically refines the estimates of X and $\{\boldsymbol{\delta}^{(j)}\}_{j=0}^{M-1}$ in each iteration.

IV. Simulation Results

We consider a radar with $M = 10$ element uniform linear array with half wavelength spacing between sensors. In each CPI, the radar transmits $N = 32$ pulses. The pulse repletion frequency (PRF) is 1 KHz. We mainly consider cases with significant intrinsic clutter motion (ICM) [23], [24]. In all simulations, we take $\beta = 0.5$ in (3). The noise vectors $\{\mathbf{e}(n)\}_{n=1}^N$ are mutually independent, complex Gaussian distributed with zero-mean, and covariance matrix $\eta^2 I$. The SNR of the p-th target is defined as

$$\text{SNR} = 10\log_{10}(\vartheta_p^2/\eta^2), \tag{30}$$

where ϑ_p^2 is the power of p-th target and $\eta = 0.5$. We set $T = 90$ and $L = 128$. For an arrangement having P targets, we say the targets are detected correctly when an algorithm produces P spikes outside the true clutter region, and the maximum absolute difference between any actual Doppler and recovered Doppler is $\pi/30$ radian. We say there is a spike at a particular location in the angle-Doppler image if the image has a local-maximum at that location with an absolute magnitude above a threshold 10^{-4}. We consider three techniques: (i) a CS based approach [34], (ii) the iterative adaptive approach (IAA) [10], and (iii) MAP. IAA is an iterative algorithm based on a weighted least squares criterion [12]. IAA usually converges in 10 to 14 iterations [10], [12], and thus, in our simulations we consider 14 iterations (See Appendix-A for a description of the IAA implementation considered in the sequel). The CS approach solves

$$\min_x \quad \|x\|_1 \quad \text{subject to } \|y - \Psi x\|_2 \le \eta. \tag{31}$$

The SPGL1 [34] framework is used to optimize (31) numerically.

The signal-to-clutter-and-noise ratio (SCNR) is defined as [10]

$$\text{SCNR} = 10\log_{10}\left(\frac{1}{P}\frac{\sum_{p=1}^P \text{tr}[\vartheta_p^2\Omega(f_p,\omega_p)\Omega^*(f_p,\omega_p)]}{\text{tr}[R_{CN}]}\right) \tag{32}$$

where R_{CN} is the true clutter-and-noise covariance matrix which is defined as [10]: $R_{CN} = \Psi Q \Psi^*$, where $Q \in \mathbb{R}^{S\times S}$ is the diagonal matrix with the clutter scatterer powers corresponding to the scanning points on the diagonal. Also f_p and ω_p are the azimuth angle and normalize Doppler frequency of p-th target and $\Omega(f_p,\omega_p) = \theta(\omega_\ell)' \otimes \phi(f_t)$,. For reference, to resolve two targets with SNR=2dB, SCNR=-65.22 dB, $T = 90$, and $L = 128$, the computation time required for CS, IAA, and MAP are $35, 29$ and 22 sec, respectively.

We compare the resolution performances of different algorithms in Figure 2. We consider two targets with SNR 2 dB, and an average SCNR -65 dB. At first we investigate the Doppler resolution achievable by the algorithms (Figure 2(a)). Both targets have azimuth angle -42°, whereas the first target has a fixed Doppler shift -0.015. The Doppler shift of the second target is gradually decreased from 0.05, and consequently, the difference in Doppler shifts between the two targets becomes smaller. We define the Doppler resolution of an algorithm as the minimum difference in Doppler shifts for which it can resolve the targets in 90% cases. The Doppler resolutions of IAA and MAP are 0.035 and 0.025, respectively.

We consider the angle resolution performance in Figure 2(b). The Doppler shifts of two targets are -0.015 and -0.005. The azimuth angle of first target is fixed to -42°, while the azimuth angle of second target is brought closer to -42°. Figure 2(b) reveals that the angle resolution of IAA is 14°, while that of MAP is 16°.

Finally in Figure 2(c) we investigate what happens when a target is close to the clutter. The first target has azimuth angle and Doppler shift are -42° and -0.015 respectively. The second target has a fixed Doppler shift -0.035, while its azimuth angle is gradually decreasing from 30°, so that it comes closer to the clutter. Clearly MAP algorithm performs better than IAA. Figure 3 shows typical angle-Doppler images when azimuth angle of second target is 21°. Note that CS fails to locate targets in presence of large clutter. While IAA can approximately locate the first target at -42° azimuth, it is unable to locate the second target. MAP detects both targets clearly. Note that by clustering the approximate clutter locations, we have several advantages. First, the clutter region is separate from the whole area which does not produce any ripples. Second, by enforcing the whole clutter as an

independent component the number of active components in the whole angle-Doppler image are reduced significantly. This is very helpful in sparse recovery techniques. Finally, as observed in the simulations, the clustering concept helps MAP algorithm to become robust in the noisy environment. Note that the actual targets in Figure 3 (a) are not very bright due to low SNR. However, MAP produces very clear image of the targets[2]. For illustration purposes, we consider another CS based algorithm called modified-CS [19], which can utilize the clutter information in its optimization process. Figure 3(e) shows that Modified-CS is unable to detect the targets in the setting considered. We also consider the SLIM algorithm [16] in Figure 3(f). Typically, SLIM converges in about 15 iterations [15]. However, we vary the iterations of SLIM between 10 and 25, change the value of p in (8) between 0.3 and 0.9 to obtain the best possible output. However, as can be seen in Figure 3(f), the SLIM always generates ripples in the area of interest, and often SLIM is unable to locate the targets. In the following, we do not include the results for modified-CS and SLIM.

We investigate robustness of the algorithms in different environments in Figure 4. At first we show effect of noise variance η (see (30)) on different algorithms in Figure 4(a). We consider two targets. We fix target SNR to 2 dB. Clearly SCNR changes with changing η. We fix the clutter power such that average SCNR remains -65 dB when $\eta = 0.5$. The detection of an algorithm is considered successful when it resolves all targets. The result is shown in Figure 4(a). MAP algorithm outperforms IAA. It is interesting to note that IAA performs well for $0.3 \leq \eta \leq 0.6$, and its performance degrades slightly when η below 0.3.

To illustrate the effect of clutter power on the detection performance we again consider two targets. The target SNR is 2 dB and $\eta = 0.5$. We then increase clutter power and try to resolve the targets using different algorithms (Figure 4(b)). Note that MAP algorithm can resolve targets efficiently even if the clutter SCNR remains above -75 dB.

In Figure 4(c) we investigate the maximum number of targets that can be resolved by various algorithms. When placing targets around the clutter, we ensure that the minimum separation between the targets is more than resolution limit of different algorithms, see Figure 2. Note that MAP algorithm can resolve upto 8 targets properly.

We claim in Section III-A that our algorithm needs only an approximate knowledge of the true clutter. We justify this in Figure 5 (a). Here the clutter is unevenly distributed around the diagonal of the angle-Doppler image. Also clutter is not distributed over the whole diagonal area (similar to the simulations in [10]). Nevertheless, as can be seen in Figure 5(d), MAP can still locate the targets. IAA can locate only one target approximately.

Figure 6 illustrates the utility of the steering vector error correction algorithm presented in Table II. We set $\nu = 0.01$ to generate $\boldsymbol{\delta}$ as in (23). We consider two targets with SNR 2 dB and SCNR = -64.8 dB. Figure 6 (a) shows the range-Doppler images generated by MAP, without error correction. Note that the image is not satisfactory. We then use algorithm of Table II to reproduce Φ. The new output using modified Φ is shown in Figures 6(b) and 6(c). Note that MAP resolves the targets clearly. Although, IAA also yields an improved image, its performance is unsatisfactory in presence of a large clutter. The situation improves when we use our steering vector correction in IAA. For instance, in Figure 7 we consider SCNR = -56.8 dB and target SNR = 2 dB. As can be seen in Figure 7, IAA locates the targets fairly.

Figure 8 shows frequency of resolving targets as a function of ν. The target-arrangement is same as in Figure 6, and average SCNR = -65 dB. We found that until $\nu \leq 0.015$, MAP algorithm can resolve the targets reliably.

Figure 9 shows the receiver operating characteristic (ROC) of MAP algorithm. We consider two radar setup. In first setup, the radar is without steering vector error as considered in Figure 3. In second setup, we consider a radar system with steering vector error similar to Figure 6. In both cases, the target-arrangements are similar to the Figures, while the target SNRs are varied. Let $\hat{X}$ be estimate of X for a particular SNR. We then select a threshold τ_i . Any local maximum of $\hat{X}$ with an absolute value is greater than τ_i is considered as a target. We vary τ_i within an interval $[\tau_s, \tau_e]$ and for each value of τ_i we record whether all actual targets are detected and whether any false targets are detected. After repeating the experiment 100 times, this procedure yields the empirical probability p_a of detecting an actual target, and the empirical probability p_f of detecting any false target for different values of τ_i. The sensitivity of the algorithm is defined as the smallest SNR for which $p_f \leq 0.1$ and $p_a \geq 0.85$. In Figure 9(a) we show the ROC in absence of any steering vector error, while ROC in presence of steering vector error is shown in 9(b). Note that in both cases the sensitivity is -1 dB, and the value corresponding to the threshold is 10^{-4}.

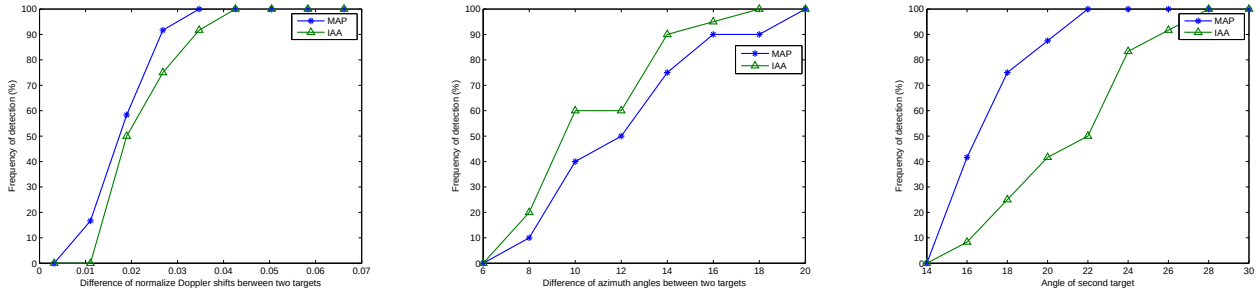
V. Conclusions

The angle-Doppler imaging problem in presence of clutter is posed as a sparse signal reconstruction problem. However, conventional sparse recovery approaches are not suitable for this framework. We find that under certain reasonable assumptions one can approximately estimate the clutter region. Using the clutter information, the problem of estimating targets can be posed to the MAP formulation. We compare the algorithm with other known imaging algorithms.

Appendix A
The IAA Algorithm [10]

- **Initialization:** Set $P_{t,\ell} = \frac{1}{(NM)^2}|(\phi(f_t) \otimes \theta(\omega_\ell))^* \boldsymbol{y}|^2$, for $t = 1, 2, \cdots, T$, and $\ell = 1, 2, \cdots L$.
- **repeat**
  1. $\mathbf{R}_{\text{IAA}} = \Psi P \Psi^*$.
  2. for $t = 1, 2, \cdots, T$
  3. for $\ell = 1, 2, \cdots L$
  4. $P_{t,\ell} = \left| \frac{(\phi(f_t) \otimes \theta(\omega_\ell))^* \mathbf{R}_{\text{IAA}}^{-1} \boldsymbol{y}}{(\phi(f_t) \otimes \theta(\omega_\ell))^* \mathbf{R}_{\text{IAA}}^{-1} (\phi(f_t) \otimes \theta(\omega_\ell))} \right|^2$
  - end

[2]Figure 3(d) might appear to be a binary version of the actual image recovered by MAP. However that is not true.



(a) (b) (c)

Fig. 2. Resolution performance of algorithms. Average SCNR $=-65$ dB. (a) Doppler resolution, (b) Angle resolution, (c) Clutter resolution.

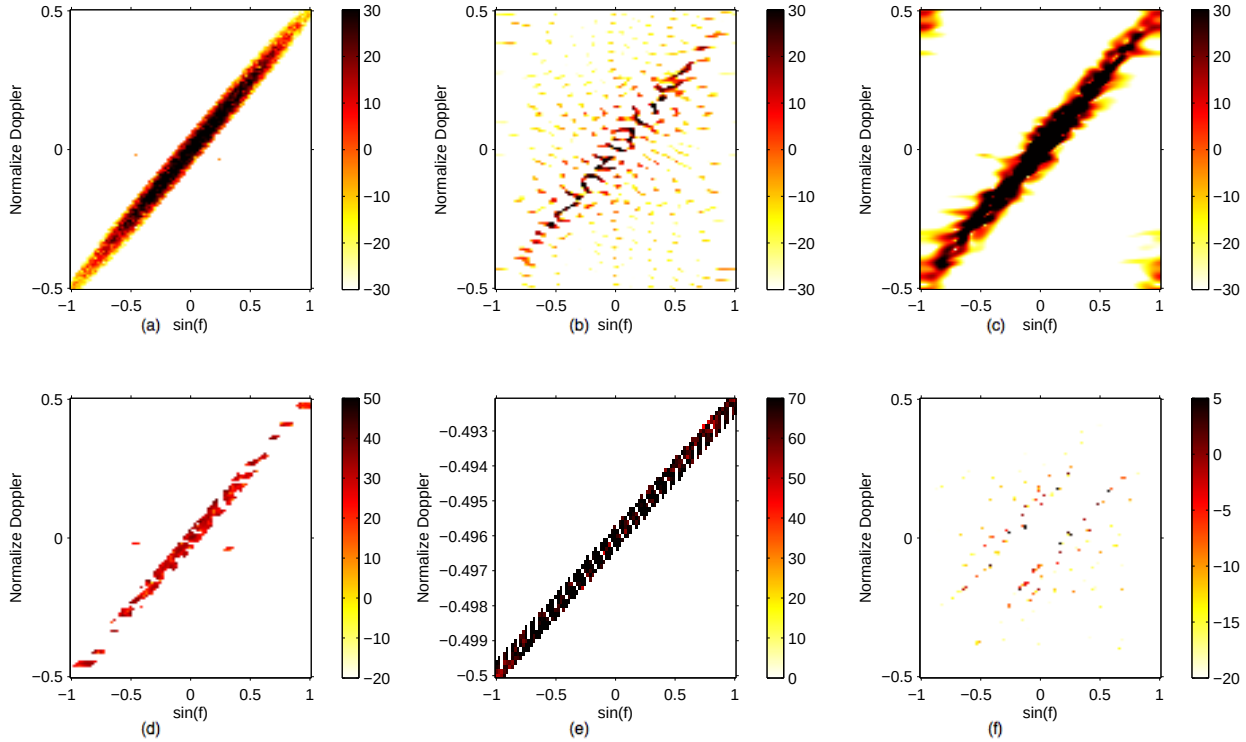
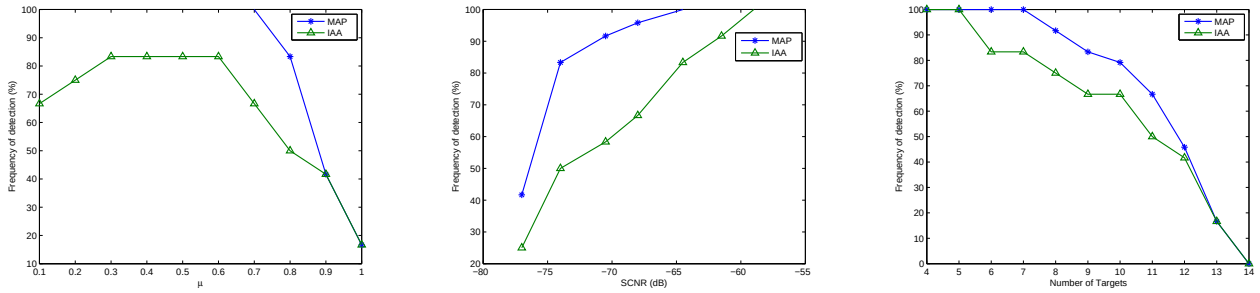


Fig. 3. Angle-Doppler images. SNRs of two targets are 2 dB and 2 dB. SCNR=-66.12 dB. Azimuth angles of the targets are -42° and 21°. Normalize Doppler shifts are -0.015 and -0.035. (a) true image, (b) CS, (c) IAA, (d) MAP algorithm, (e) Modified-CS [19], (f) SLIM [16].



(a) (b) (c)

Fig. 4. Robustness of algorithms. (a) Robustness against noise power. SNRs of two targets are 2 dB and 2 dB. Azimuth angles of the targets are $\{-42^\circ, 24^\circ\}$ and Doppler shifts are $\{-0.015, -0.035\}$. (b) Robustness against clutter power. SNRs of two targets are 2 dB, 2 dB and $\eta = 0.5$. Azimuth angles of the targets are $\{-42^\circ, 24^\circ\}$ and Doppler shifts are $\{-0.015, -0.035\}$. (c) Multiple targets detection.

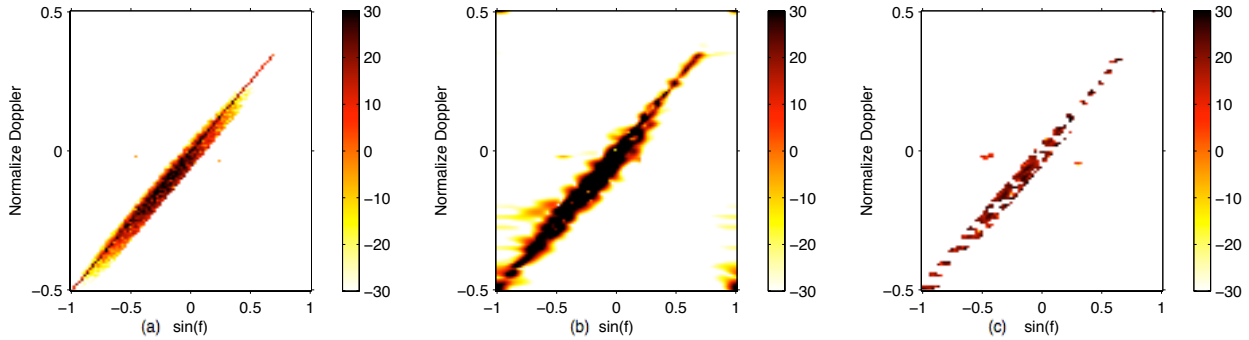


Fig. 5. Angle-Doppler images. SNRs of two targets are 2 dB and 2 dB. SCNR=-58.22 dB. Azimuth angles of the targets are $\{-42°, 21°\}$ and Doppler shifts are $\{-0.015, -0.035\}$. (a) True image, (b) IAA, (c) MAP algorithm.

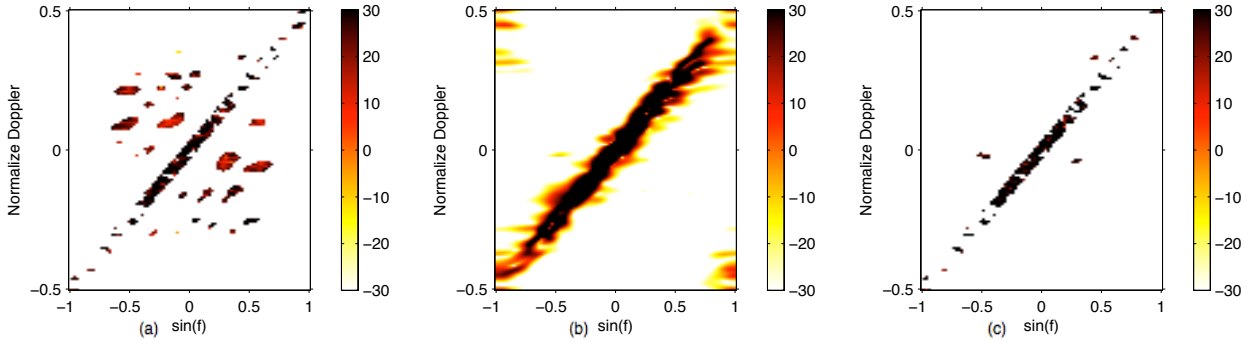


Fig. 6. Angle-Doppler images in presence of steering vector error. SNRs of two targets are 2 dB and 2 dB. SCNR = -64.8 dB. Azimuth angles of the targets are $\{-42°, 24°\}$ and Doppler shifts are $\{-0.015, -0.035\}$. (a) MAP algorithm with steering vector error. (b) IAA after steering vector error correction. (c) MAP algorithm after error correction.

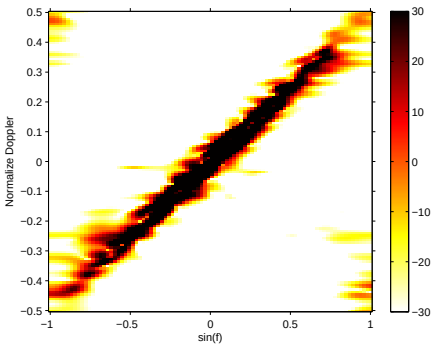


Fig. 7. Angle-Doppler image recovered by IAA after steering vector error correction. SNRs of two targets are 2 dB. SCNR=-56.8 dB. Azimuth angles of the targets are $\{-42°, 24°\}$ and Doppler shifts are $\{-0.015, -0.035\}$.

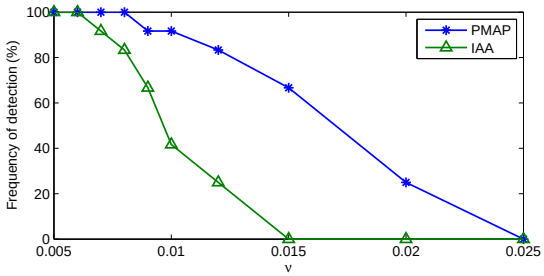


Fig. 8. Frequency of detecting targets as a function of ν. Two targets with SNR =2 dB are placed at azimuth angle $\{-42°, 24°\}$ and their Doppler shifts are $\{-0.015, -0.0354\}$. Average SCNR =-65 dB.

– end

- **until** a certain number of iterations is reached.

REFERENCES

[1] J. Guerci, *Space-Time Adaptive Processing for Radar*. Artech House, 2003.

[2] J. Ward, "Space-time adaptive processing for airborne radar," *Lincoln Lab., Mass. Inst. Techno., Lexington, MA, Tech. Pep. 1015*, Dec. 1994.

[3] J. Guerci and E. Baranoski, "Knowledge-aided adaptive radar at DARPA: an overview," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 41 – 50, jan. 2006.

[4] P. Stoica, J. Li, X. Zhu, and J. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Transactions on Signal Processing*, vol. 56, pp. 2598 – 2602, June. 2008.

[5] H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 660 –670, jul. 1994.

[6] G. Fabrizio, A. Gershman, and M. Turley, "Robust adaptive beamform-

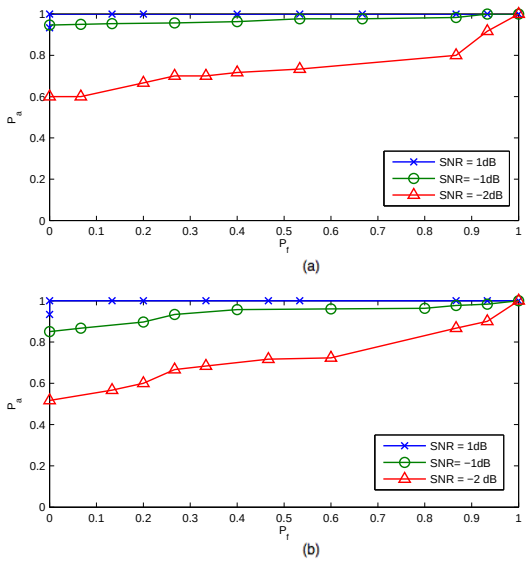


Fig. 9. Receiver operating characteristic (ROC) for MAP algorithm. (a) Without steering vector error, (b) With steering vector error.

ing for HF surface wave over-the-horizon radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 2, pp. 510 – 525, apr. 2004.

[7] J. Wang, A. Dogandzic, and A. Nehorai, "Maximum likelihood estimation of compound-gaussian clutter and target parameters," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3884 –3898, oct. 2006.

[8] D. Ramakrishnan and J. Krolik, "Adaptive radar detection in doubly nonstationary autoregressive doppler spread clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 484 –501, apr. 2009.

[9] D. Musicki and D. Nesic, "Extremum seeking control of ill-defined exponential process," in *Control and Decision Conference (CCDC), 2011 Chinese*, may 2011, pp. 2698 –2703.

[10] J. Li, X. Zhu, P. Stoica, and M. Rangaswamy, "High resolution angle-doppler imaging for MTI radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1544 –1556, jul. 2010.

[11] S. Maria and J.-J. Fuchs, "Application of the global matched filter to stap data an efficient algorithmic approach," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.*, vol. 4, may. 2006.

[12] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. Baggeroer, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 425 –443, jan. 2010.

[13] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust Capon beamformer," *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2407 – 2423, sep. 2004.

[14] Y. Zhang, J. Lie, B. P. Ng, and C. See, "Robust minimum ℓ_1-norm adaptive beamformer against intermittent sensor failure and steering vector error," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 5, pp. 1796 –1801, may. 2010.

[15] X. Tan, W. Roberts, J. Li, and P. Stoica, "Sparse learning via iterative minimization with application to MIMO radar imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1088 –1101, march 2011.

[16] M. Xue, W. Roberts, J. Li, X. Tan, and P. Stoica, "MIMO radar sparse angle-Doppler imaging for ground moving target indication," in *2010 IEEE Radar Conference*, may 2010, pp. 553 –558.

[17] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2275 –2284, june 2009.

[18] T. Strohmer and B. Friedlander, "Compressed sensing for MIMO radar - algorithms and performance," in *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, nov. 2009, pp. 464 –468.

[19] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4595 –4607, sept. 2010.

[20] M. Amin Khajehnejad, W. Xu, A. Salman Avestimehr, and B. Hassibi, "Weighted ℓ_1 minimization for sparse recovery with prior information," in *IEEE International Symposium on Information Theory, 2009. ISIT 2009*, Jul. 2009, pp. 483 –487.

[21] N. Vaswani, "LS-CS-residual (LS-CS): Compressive sensing on least squares residual," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4108 –4120, aug. 2010.

[22] B. R. Mahafza and H. Alabama, *MATLAB Simulations for Radar Systems Design*. CRC press, 2003.

[23] P. Techau, J. Bergin, and J. Guerci, "Effects of internal clutter motion on STAP in a heterogeneous environment," in *Proceedings of the 2001 IEEE Radar Conference, 2001.*, 2001, pp. 204 –209.

[24] W. Zeng, B. Tang, Y. Zeng, and W. Peng, "Robust space-time adaptive processing in heterogeneous environment with internal clutter motion," in *2010 International Conference on Communications, Circuits and Systems (ICCCAS)*, jul. 2010, pp. 457 –460.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[26] C. L. Lawson, *Contributions to the theory of linear least maximum approximations*. PhD thesis, UCLA, 1961.

[27] B. Rao, K. Engan, S. Cotter, J. Palmer, and K. Kreutz-Delgado, "Subset selection in noise based on diversity measure minimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 760 – 770, march 2003.

[28] M. Hyder and K. Mahata, "Maximum a posteriori estimation approach to sparse recovery," in *2011 17th International Conference on Digital Signal Processing (DSP)*, july 2011, pp. 1 –6.

[29] J. Trzasko and A. Manduca, "Relaxed conditions for sparse signal recovery with general concave priors," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4347 –4354, nov. 2009.

[30] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707 –710, Oct. 2007.

[31] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[32] M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4646 –4655, sept. 2010.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[34] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, http://www.cs.ubc.ca/labs/scl/spgl1.