

Sparsity-Aware STAP Algorithms With L_1 -norm Regularization For Airborne Radar

Zhaocheng Yang, Rodrigo C. de Lamare, *Senior Member, IEEE* and Xiang Li, *Member IEEE*

Abstract—This article proposes novel sparsity-aware space-time adaptive processing (SA-STAP) algorithms with l_1 -norm regularization for airborne phased-array radar applications. The proposed SA-STAP algorithms suppose that a number of samples of the full-rank STAP data cube are not meaningful for processing and the optimal full-rank STAP filter weight vector is sparse, or nearly sparse. The core idea of the proposed method is imposing a sparse regularization (l_1 -norm type) to the minimum variance (MV) STAP cost function. Under some reasonable assumptions, we firstly propose a l_1 -based sample matrix inversion (SMI) to compute the optimal filter weight vector. However, it is impractical due to its matrix inversion, which requires a high computational cost when using a large phased-array antenna. In order to compute the STAP parameters in a cost-effective way, we devise low-complexity algorithms based on conjugate gradient (CG) techniques. A computational complexity comparison with the existing algorithms and an analysis of the proposed algorithms are conducted. Simulation results with both simulated and the Mountain Top data demonstrate that fast signal-to-interference-plus-noise-ratio (SINR) convergence and good performance of the proposed algorithms are achieved.

Index Terms— l_1 regularization, Sparsity-aware Space-time adaptive processing, Conjugate gradient techniques, Airborne radar, Mountain Top data.

I. INTRODUCTION

Space-time adaptive processing (STAP) is an efficient tool for detection of slow targets by airborne or spaceborne radar systems in hostile environments, such as strong clutter and jammers [1]–[4]. However, the full-rank adaptive STAP based on linearly constrained minimum variance (LCMV) criterion gives rise to two of the major limitations in practical applications of radar [2], [4]. First, the computational load required to solve the interference matrix inversion is quite high. In addition, the number of training data samples required for an accurate estimate of the interference covariance matrix can become impractical for high-dimensional problems, particularly in heterogeneous environments. It is therefore desirable to develop STAP techniques with low computational complexity and that can provide high performance in small-sample support situations.

The diagonal loading sample matrix inversion (LSMI) technique is considered to be a simple and robust approach for both homogeneous and heterogeneous environments [5], but has a

high computational cost. Reduced-rank techniques have been investigated for solving the previously discussed problems in the last decades [6]–[12], [14]–[22]. One of the most important reduced-rank techniques is the class of the Krylov subspace methods, which includes the auxiliary-vector filters (AVF) [7], [8], the multistage Wiener filter (MWF) [9]–[13] and the conjugate gradient (CG) algorithm [14]–[17]. These methods project the observation data onto a lower-dimensional Krylov subspace and can obtain an improved convergence and tracking performance. The main differences amongst them lie in the computational cost, the structure of adaptation and the ease of implementation. Knowledge-aided (KA) STAP techniques have currently gained significant attention as an effective STAP algorithm to mitigate the effects of the heterogeneity in the secondary data by exploiting a priori knowledge [18], [19], [23], [24]. However, the exact form of prior knowledge is still problem-dependent and hard to be derived. More recently, several authors have considered sparse recovery (SR) ideas for moving target indication (MTI) and STAP problems [25]–[31]. These works based on SR techniques rely on the recovery of the clutter spectrum in the angle-Doppler plane, which is usually carried out via two steps: first, recovering the clutter angle-Doppler profile by some SR algorithms; second, estimating the covariance matrix based on the result obtained in the first step, and computing the Capon’s optimal filter. Although some fast sparse recovery algorithms are proposed, e.g., the fast iterated shrinkage/thresholding (FISTA) algorithm [27], and the focal underdetermined system solution (FOCUSS) based algorithm [30], it is more computationally expensive than conventional STAP because the Capon’s optimal filter requires a matrix inversion, and the recovery procedure needs additional computations.

In airborne radar systems, most interference suppression problems are rank deficient in nature [2]–[4], i.e., they require less adaptive degrees of freedom (DOFs) than the full DOFs provided by the array. In this case, the total adaptive DOFs provided by the array will be much great than the number that needed to suppress the interference. Motivated by this, the authors in [32] proposed a sequential approach to obtain a sparse solution for the transformation matrix that selects the “best” DOFs to be retained in a partially adaptive beamformer. Moreover, the property described above suggests that there is a high degree of sparsity of the filter weight vector. Hence, in our prior work, an l_1 type regularization to the generalized sidelobe canceler (GSC) STAP processor using the l_1 -based online coordinate gradient (OCD) method [33] and the l_1 -based recursive least squares method [34] is introduced to exploit the sparsity of the received data and filter weights,

Z. Yang and X. Li are with Research Institute of Space Electronics, Electronics Science and Engineering School, National University of Defense Technology, Changsha, 410073, China. e-mail: yangzhaocheng@gmail.com, lixiang01@vip.sina.com.

R. C. de Lamare is with Communications Research Group, Department of Electronics, University of York, YO10 5DD, UK. e-mail: rcd1500@ohm.york.ac.uk

resulting in an improvement in both convergence rate and steady-state signal-to-interference-plus-noise ratio (SINR) performance. In this paper, we extend the work presented in [33] and [34] to the direct filter STAP processor (DFP) **and focus on a computationally efficient algorithm for implementation. By adding the sparsity constraint (l_1 -norm regularization) to the MV cost function, the proposed STAP technique is able to automatically shrink the least relevant coefficients of the filter to zero, thereby using only the most important DOFs and creating a sparse filter.** We derive the l_1 -regularized optimal filter under some reasonable assumptions, and then propose a sparsity-aware (SA) adaptive STAP strategy for airborne radar systems. **One approach to computing the parameters of the filter** is to use the l_1 -based SMI recursion algorithm to compute the filter weights. However, it requires the matrix inversion operation, which prevents its use in practice. The CG method has a low computational complexity and is the simplest Krylov subspace method since it only needs the forward stage, unlike the MWF that requires both forward and backward stages. Therefore, low complexity l_1 -based CG type algorithms are devised. The simulations are conducted using both simulated and measured data, which show that the proposed algorithms exhibit improved performance as compared to existing techniques.

This paper is organized as follows. Section II introduces the STAP signal model for airborne radar. In Section III, we first introduce the strategy of the SA-STAP algorithm. Then l_1 -based SMI and l_1 -based CG type algorithms are developed and their computational complexity is also shown. Furthermore, we conduct an analysis of the proposed algorithms. In Section IV, some examples of performance of the proposed algorithms with both simulated and the Mountain Top data are shown. Finally, the conclusions are given in Section V.

Notation: In this paper, scalar quantities are denoted with italic typeface. Lowercase boldface quantities denote vectors and uppercase boldface quantities denote matrices. The operations of transposition, complex conjugation, and conjugate transposition are denoted by superscripts T , $*$, and H , respectively. The symbols \otimes represents the Kronecker product and \odot denotes the Hadamard matrix product. Finally, the symbol $E\{\cdot\}$ denotes the expected value of a random quantity, operator $\Re[\cdot]$ selects the real part of argument, and the symbol $\|\cdot\|_p$ denotes the l_p -norm operation of a vector.

II. SIGNAL MODEL AND PROBLEM STATEMENT

The system under consideration is a pulsed Doppler radar residing on an airborne platform. The radar antenna is a **uniform linear array** (ULA) which consists of M elements. The platform is at altitude h_p and moving with constant velocity v_p . The chosen coordinate system is shown in Fig.1(a). The angle variables ϕ and θ refer to elevation and azimuth. The radar transmits a coherent burst of pulses at a constant pulse repetition frequency (PRF) $f_r = 1/T_r$, where T_r is the pulse repetition interval (PRI). The transmitter carrier frequency is $f_c = c/\lambda_c$, where c is the propagation velocity and λ_c is the wavelength. The coherent processing interval (CPI) length is equal to NT_r . For each PRI, K time samples are collected

to cover the range interval. After matched filtering to the radar returns from each pulse, the received data set for one CPI comprises KNM complex baseband samples, which is referred to as the radar datacube shown in Fig.1(b). The data are then processed at one range of interest, which corresponds to a slice of the CPI datacube. The slice is an $M \times N$ matrix which consists of $M \times 1$ spatial snapshots for pulses at the range of interest. It is convenient to stack the matrix column-wise to form the $NM \times 1$ vector $\mathbf{x}[k]$, termed a space-time snapshot, where k is the range sample index and $1 \leq k \leq K$ [2]–[4].

Target detection in airborne radar systems can be formulated into a binary hypothesis problem, where the hypothesis H_0 corresponds to target absence and the hypothesis H_1 corresponds to target presence, given as

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{x}_u \\ H_1 : \mathbf{x} &= \alpha_s \mathbf{s} + \mathbf{x}_u, \end{aligned} \quad (1)$$

where α_s is a complex gain and the vector \mathbf{s} , which is the $NM \times 1$ normalized space-time steering vector in the space-time look-direction, defined as

$$\mathbf{s} = \frac{\mathbf{s}_t(f_d) \otimes \mathbf{s}_s(f_s)}{\|\mathbf{s}_t(f_d) \otimes \mathbf{s}_s(f_s)\|_2}, \quad (2)$$

where $\mathbf{s}_t(f_d)$ denotes the $N \times 1$ temporal steering vector at the target Doppler frequency f_d and $\mathbf{s}_s(f_s)$ denotes the $M \times 1$ spatial steering vector in the direction provided by the target frequency f_s . The vector \mathbf{x}_u encompasses any undesired interference or noise component of the data including clutter \mathbf{x}_c , jamming \mathbf{x}_j and thermal noise \mathbf{x}_n . Generally, we assume the thermal noise is spatially and temporally uncorrelated, and the jamming is temporally uncorrelated but spatially strongly correlated. As for the clutter, a general model for the clutter space-time snapshot is given by [36]

$$\begin{aligned} \mathbf{x}_c[k] &= \sum_{m=1}^{N_r} \sum_{n=1}^{N_c} \sigma_{c/k;m,n} (\boldsymbol{\alpha}_t(k;m,n) \odot \mathbf{s}_t(f_{d/k;m,n})) \\ &\quad \otimes (\boldsymbol{\alpha}_s(k;m,n) \odot \mathbf{s}_s(f_{s/k;m,n})), \end{aligned} \quad (3)$$

where N_r is the number of range ambiguities, N_c is the number of independent clutter patches that are evenly distributed in azimuth about the radar, $\boldsymbol{\alpha}_t(k;m,n)$ is a vector describing the normalized pulse-to-pulse voltages, and $\boldsymbol{\alpha}_s(k;m,n)$ accounts for spatial decorrelation. $\sigma_{c/k;m,n}$ describes the average voltage for the m th clutter patch and k th range. The clutter-jammer-noise (for short, calling interference in the following part) covariance matrix \mathbf{R} can be expressed as

$$\mathbf{R} = E\{\mathbf{x}_u \mathbf{x}_u^H\} = \mathbf{R}_c + \mathbf{R}_j + \mathbf{R}_n, \quad (4)$$

where $\mathbf{R}_c = E\{\mathbf{x}_c \mathbf{x}_c^H\}$, $\mathbf{R}_j = E\{\mathbf{x}_j \mathbf{x}_j^H\}$ and $\mathbf{R}_n = E\{\mathbf{x}_n \mathbf{x}_n^H\}$, denote clutter, jammer and thermal noise covariance matrix, respectively.

Generally, the space-time processor linearly combines the elements of the data snapshot, yielding the scalar output [4]

$$y = \mathbf{w}^H \mathbf{x}, \quad (5)$$

where \mathbf{w} is the $NM \times 1$ weight vector. The idea behind LCMV approach is to minimize the STAP output power whilst

constraining the gain in the direction of the desired signal. This leads to the following power minimization with constraints

$$\min_{\mathbf{w}} J(\mathbf{w}) = E \left\{ \|\mathbf{w}^H \mathbf{x}\|_2^2 \right\} \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{s} = 1. \quad (6)$$

Using the method of Lagrange multipliers, the optimal full-rank LCMV STAP weights are given by [1]

$$\mathbf{w}_{\text{LCMV}} = \frac{\mathbf{R}^{-1} \mathbf{s}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}. \quad (7)$$

III. SA-STAP WITH L_1 -NORM REGULARIZATION

In this section, we detail the design of the proposed SA-STAP strategy, derive the l_1 -based SMI recursion algorithm and the l_1 -based CG type algorithms and detail their complexity. Finally, the analysis of the proposed SA-STAP algorithms is shown.

A. SA-STAP Strategy

In airborne radar systems, most interference suppression problems are rank deficient in nature, which means that they require less adaptive DOFs than those offered by the array, the additional DOFs that are not required can be discarded so that only those that are important are retained, which has been exploited by several techniques including reduced-dimension [35], reduced-rank [6]–[12], [14]–[19] and partially STAP technique [32]. Furthermore, full DOFs will lead to slow convergence, i.e. requiring many snapshots to train the filter, which is difficult to obtain especially in non-homogeneous clutter environments. As a result, the total adaptive DOFs provided by the array will be much greater than those required to suppress the interference. **In other words, the required length of the filter weight vector (required DOFs) to suppress the clutter is much less than NM , which corresponds to the number of nonzero elements or significant elements in the full filter weight vector is much less than NM . From this point of view, there is a high degree of sparsity to be exploited in the filter weight vector.** However, in practice it is not easy to estimate the required DOFs related to the sparsity and to decide which DOFs are the most important ones. The authors in [33], [34] proposed an l_1 regularized STAP algorithm for GSC structure to exploit the sparsity of the received data and filter weights. In this paper, we extend this work to a more general framework for airborne radar systems, by employing the sparse regularization to the MV STAP cost function, which is described as the following optimization problem

$$\min_{\mathbf{w}} E \left\{ \|\mathbf{w}^H \mathbf{x}\|_2^2 \right\} + 2\lambda \Gamma(\mathbf{w}) \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{s} = 1, \quad (8)$$

where λ is a positive scalar which provides a trade-off between the sparsity and the output interference power. The larger the chosen λ , the more components are shrunk to zero [37]. The sparse regularization is usually conducted by the l_0 -norm constraint [38]–[40]. However, since this kind of optimization problem is known to be NP-hard, one of the approximation algorithms, called l_1 -norm, is considered for the convexity and simple complexity [39]. In the following, we adopt the l_1 -norm regularization, i.e., $\Gamma(\mathbf{w}) = \|\mathbf{w}\|_1$. Now, the question that arises is how to effectively solve the l_1 regularized MV STAP.

Albeit convex, the cost function $J_1(\mathbf{w})$ is non-differentiable which leads to a difficulty with the use of the method of Lagrange multipliers directly. Thus, inspired by [38], we propose an approximation to the regularization term, which is given by

$$\Gamma(\mathbf{w}) = \|\mathbf{w}\|_1 \approx \mathbf{w}^H \mathbf{\Lambda} \mathbf{w}, \quad (9)$$

where

$$\mathbf{\Lambda} = \text{diag} \left\{ \frac{1}{|w_1| + \epsilon}, \frac{1}{|w_2| + \epsilon}, \dots, \frac{1}{|w_{NM}| + \epsilon} \right\}, \quad (10)$$

where ϵ is a small positive constant (e.g., $\epsilon = 0.01 \sim 0.1$ is acceptable), and $w_i, i = 1, 2, \dots, NM$ are the entries of the filter weight vector \mathbf{w} . Thus the regularization term $\mathbf{w}^H \mathbf{\Lambda} \mathbf{w}$ has a quadratic structure, if we assume that the diagonal matrix $\mathbf{\Lambda}$ is fixed. The minimization can be performed iteratively by assuming that the term $\mathbf{\Lambda}$ is fixed and computed with the current solution \mathbf{w} [38]. By fixing the term $\mathbf{\Lambda}$, we compute the partial derivative of (9) with respect to \mathbf{w}^* , which is given as follows

$$\frac{\partial \|\mathbf{w}\|_1}{\partial \mathbf{w}^*} \approx \frac{\partial (\mathbf{w}^H \mathbf{\Lambda} \mathbf{w})}{\partial \mathbf{w}^*} = \mathbf{\Lambda} \mathbf{w}. \quad (11)$$

The above constrained optimization problem described by (8) can be transformed into an unconstrained optimization problem by the method of Lagrange multipliers, whose cost function becomes

$$\mathcal{L} = E \left\{ \|\mathbf{w}^H \mathbf{x}\|_2^2 \right\} + 2\lambda \|\mathbf{w}\|_1 + 2\Re \{ \kappa^* (\mathbf{w}^H \mathbf{s} - 1) \}, \quad (12)$$

where κ is a complex Lagrange multiplier. Computing the gradient terms of (12) with respect to \mathbf{w}^* and κ^* , we get

$$\begin{aligned} \nabla \mathcal{L}_{\mathbf{w}^*} &= \mathbf{R} \mathbf{w} + \lambda \mathbf{\Lambda} \mathbf{w} + \kappa^* \mathbf{s} \\ \nabla \mathcal{L}_{\kappa^*} &= \mathbf{w}^H \mathbf{s} - 1. \end{aligned} \quad (13)$$

By equating the above gradient terms to zero, we obtain the filter weight vector

$$\mathbf{w} = \frac{[\mathbf{R} + \lambda \mathbf{\Lambda}]^{-1} \mathbf{s}}{\mathbf{s}^H [\mathbf{R} + \lambda \mathbf{\Lambda}]^{-1} \mathbf{s}}. \quad (14)$$

By inspecting (14), we verify that there is an additional term $\lambda \mathbf{\Lambda}$ in the inverse of the interference covariance matrix \mathbf{R} , which is due to the l_1 -norm regularization. One should note that the filter weight vector expression in (14) is not a closed-form solution since $\mathbf{\Lambda}$ is a function of \mathbf{w} . Thus it is necessary to develop an iterative procedure to compute the filter weight vector, which will be shown in the following parts.

B. L_1 -Based SMI Recursion Algorithm

In practice, because the interference covariance is unknown to us, it is most common to compute the interference covariance matrix estimate as [2]–[4]

$$\hat{\mathbf{R}} = \frac{1}{L} \sum_{n=1}^L \mathbf{x}[n] \mathbf{x}^H[n], \quad (15)$$

where $\{\mathbf{x}[n]\}_{n=1}^L$ are known as the secondary or training data. In our following derivation, to develop an iterative procedure, we add an exponential weighting factor to the interference

covariance matrix, which may allow the STAP algorithms to accommodate possible non-stationarities in the input. We write the $\hat{\mathbf{R}}[k]$ as

$$\hat{\mathbf{R}}[k] = \sum_{n=1}^i \beta^{k-n} \mathbf{x}[n] \mathbf{x}^H[n] = \beta \hat{\mathbf{R}}[k-1] + \mathbf{x}[i] \mathbf{x}^H[i], \quad (16)$$

where β is the forgetting factor, and $\hat{\mathbf{R}}[0] = \delta \mathbf{I}$, where δ is a small positive quantity and \mathbf{I} is the identity matrix. Since $\Lambda[k]$ is a function of $\mathbf{w}[k]$, we assume that the filter weight values do not change significantly in a single snapshot step, which is reasonable because we want the instantaneous error of the filter weight vector to change slowly [41]. Hence, $\Lambda[k]$ can be approximated by

$$\Lambda[k] \approx \Lambda[k-1] = \text{diag} \left\{ \frac{1}{|w_1[k-1]| + \epsilon}, \dots, \frac{1}{|w_{NM}[k-1]| + \epsilon} \right\}. \quad (17)$$

However, we note that the computational complexity of the l_1 -based SMI recursion algorithm is proportional to $O((NM)^3)$, which is not practical, especially in large antenna arrays. In the next section, we will develop some low complexity algorithms.

C. L_1 -Based CG Algorithms

In order to reduce the computational complexity of the l_1 -based SMI recursion algorithm, we introduce low complexity adaptive algorithms based on CG techniques to iteratively compute the filter weights. There are two different basic strategies for using the CG method. One is the conventional CG (CCG) [15], [16], which executes several iterations per sample and runs the reset periodically for convergence. The other is the modified CG (MCG) [14], [16], [17], which operates only one iteration per sample. CCG has a faster convergence than MCG, but a higher computational complexity. In the following, we detail the derivation of the l_1 -based SA-STAP algorithms based on these two strategies, called l_1 -based CCG algorithm and l_1 -based MCG algorithm. For simplicity, we first introduce an auxiliary vector given by

$$\mathbf{v}[k] = \left[\hat{\mathbf{R}}[k] + \lambda \Lambda[k] \right]^{-1} \mathbf{r}_t. \quad (18)$$

Then the STAP filter weight vector can be described as $\mathbf{w}[k] = \mathbf{v}[k] / (\mathbf{s}^H \mathbf{v}[k])$. The solution of $\mathbf{v}[k]$ described by (18) is also the solution of the following minimization problem [16]:

$$\min \mathcal{J}(\mathbf{v}) = \mathbf{v}^H \left[\hat{\mathbf{R}} + \lambda \Lambda \right] \mathbf{v} - 2 \Re \{ \mathbf{v}^H \mathbf{s} \}. \quad (19)$$

Then the CG-based weight vector is expressed by

$$\mathbf{v}[k] = \mathbf{v}[k-1] + \alpha[k] \mathbf{p}[k], \quad (20)$$

where $\mathbf{p}[k]$ is the direction vector, $\alpha[k]$ is the corresponding adaptive step size.

For the l_1 -based CCG algorithm, the iteration procedure for the CG-based weight vector \mathbf{v} is executed per sample. For the k th sample, it assumes constant $\hat{\mathbf{R}}[k] + \lambda \Lambda[k]$ within the internal iterations, and D internal iterations are performed

per input data sample. The main difference between the l_1 -based CCG algorithm and the existing CCG algorithm after the derivation is that we add an additional term $\lambda \Lambda[k]$ to the estimated interference covariance matrix $\hat{\mathbf{R}}[k]$. A summary of the algorithm is shown in Table I.

The l_1 -based CCG algorithm operates multiple iterations per sample and runs the reset periodically for convergence, which increases the computational load in the sample-by-sample update. In the following, we detail the derivations of the l_1 -based MCG algorithm with one iteration per sample for STAP. From [14], one way to realize the conjugate gradient method with one iteration per snapshot is the application of the degenerated scheme, which means that the residual vector $\mathbf{g}[k]$ will not be completely orthogonal to the subspace spanned by the direction vectors $\{\mathbf{p}[0], \mathbf{p}[1], \dots, \mathbf{p}[k-1]\}$. Under this condition, the adaptive step size $\alpha[k]$ has to fulfill the convergence bound given by

$$0 \leq |\mathbf{p}^H[k] \mathbf{g}[k]| \leq 0.5 |\mathbf{p}^H[k] \mathbf{g}[k-1]|, \quad (21)$$

where $\mathbf{g}[k]$ is the negative gradient vector of $\mathcal{J}(\mathbf{v})$ in (19). Thus, $\mathbf{g}[k]$ can be written as

$$\mathbf{g}[k] = -\nabla \mathcal{J}(\mathbf{v})_{\mathbf{v}^*} = - \left[\hat{\mathbf{R}}[k] + \lambda \Lambda[k] \right] \mathbf{v}[k] + \mathbf{s}, \quad (22)$$

which can be calculated recursively by

$$\mathbf{g}[k] = (1 - \beta) \mathbf{s} + \beta \mathbf{g}[k-1] - \alpha[k] \left[\hat{\mathbf{R}}[k] + \lambda \Lambda[k] \right] \mathbf{p}[k] - \left\{ \lambda [1 - \beta] \Lambda[k] + \mathbf{x}[k] \mathbf{x}^H[k] \right\} \mathbf{v}[k-1]. \quad (23)$$

In the previous equation, we use the approximation that $\Lambda[k-1] \approx \Lambda[k]$. Premultiplying (23) by $\mathbf{p}^H[k]$, taking the expectation of both sides and considering $\mathbf{p}[k]$ uncorrelated with \mathbf{s} , $\mathbf{x}[k]$ and $\mathbf{v}[k-1]$ [14], we obtain

$$\begin{aligned} E[\mathbf{p}^H[k] \mathbf{g}[k]] &\approx \beta E[\mathbf{p}^H[k] \mathbf{g}[k-1]] \\ &\quad - \beta E[\mathbf{p}^H[k] \mathbf{s}] \\ &\quad - E[\alpha[k] \mathbf{p}^H[k] \left(\hat{\mathbf{R}}[k] + \lambda \Lambda[k] \right) \mathbf{p}[k]]. \end{aligned} \quad (24)$$

Here, it is assumed that the algorithm converges with the assumption that $E[\mathbf{v}[k-1] - \mathbf{v}_{opt}] \approx 0$, $E[\mathbf{x}[k] \mathbf{r}^H[k] \mathbf{v}[k-1]] \approx \mathbf{s}$, and $E[\lambda [1 - \beta] \Lambda[k] \mathbf{v}[k-1]] \approx 0$. Making a rearrangement of (25) and following the convergence bound (21), we obtain

$$\alpha[k] = \left[\mathbf{p}^H[k] \left(\hat{\mathbf{R}}[k] + \lambda \Lambda[k] \right) \mathbf{p}[k] \right]^{-1} \left\{ \beta [\mathbf{p}^H[k] \mathbf{g}[k-1] - \mathbf{p}^H[k] \mathbf{s}] - \mu \mathbf{p}^H[k] \mathbf{g}[k-1] \right\}. \quad (25)$$

where $0 \leq \mu \leq 0.5$. The direction vector is a linear combination from the previous direction vector and the negative gradient, which is described as

$$\mathbf{p}[k] = \mathbf{g}[k-1] + \nu[k] \mathbf{p}[k], \quad (26)$$

where $\nu[k]$ is computed for avoiding the reset procedure by employing the Polak-Ribiere approach, which often has an improved performance [14], [16], and is stated as

$$\nu[k] = \frac{[\mathbf{g}[k] - \mathbf{g}[k-1]]^H \mathbf{g}[k]}{\mathbf{g}^H[k-1] \mathbf{g}[k-1]}. \quad (27)$$

The proposed l_1 -based MCG STAP algorithm is summarized in Table II.

From the above discussions, two aspects **should be noted**: First, the performance of our proposed algorithms (both l_1 -based SMI and l_1 -based CG-type algorithms) depends on **the** regularization parameter λ . An approach to choose λ is introduced in [34], which can be easily extended to our proposed algorithms, but is not discussed in this paper due to space limitations. Second, the convergence analysis in [16] is suitable to our proposed CG-type algorithms, where the convergence is governed by

$$\|\varsigma_{i+1}[k]\|_{\mathbf{G}[k]} \leq 2 \left(\frac{\sqrt{\tau_{\max}/\tau_{\min}} - 1}{\tau_{\max}/\tau_{\min} + 1} \right)^i \|\varsigma_0[k]\|_{\mathbf{G}[k]}, \quad (28)$$

where $\varsigma_i[k] = \mathbf{v}_{\text{opt}}[k] - \mathbf{v}_i[k]$ is the CG-based weight vector error at the i th iteration for the k th snapshot, $\mathbf{v}_{\text{opt}}[k]$ is the optimal solution at the k th snapshot, τ_{\max} and τ_{\min} are the maximal and minimal eigenvalues with respect to $\mathbf{G}[k] = \hat{\mathbf{R}}[k] + \lambda \mathbf{\Lambda}[k]$, and $\|\varsigma_i[k]\|_{\mathbf{G}[k]} = \varsigma_i^H[k] \mathbf{G}[k] \varsigma_i[k]$. From the above equation, we note that the convergence behavior of the proposed algorithms is related to the CG-based weight vector error $\varsigma_0[k]$ and the condition number τ_{\max}/τ_{\min} .

D. Complexity Analysis

In this section, we detail the computational complexity in terms of complex additions and complex multiplications of the proposed l_1 -based SMI, l_1 -based CG type algorithms, and other existing STAP algorithms, namely the LSMI, the AVF, the MWF and the conventional CG type algorithms, as shown in Table III. It should be noted that the rank D may not be equal to the clutter rank, and can be smaller than that. This is because the principle of the Krylov subspace approach is different from that of the eigen-decomposition approach. An eigen-decomposition approach would usually require an SVD on the full-rank covariance matrix and the selection of the D eigenvectors associated with the D largest eigenvalues, which is high related to the clutter rank. In contrast to that, the Krylov-based approach does not require an eigen-decomposition and selects the D basis vectors which minimize the desired cost function and will form the projection matrix, where D can be decreased without significantly degrading the SINR [12]. In the table, D is the rank for the CCG type, the AVF and the MWF algorithms, and $L = NM$ is the system size. Seen from the table, the computational complexity of l_1 -based SMI is similar to the conventional LSMI algorithm, both requiring one to calculate the matrix inversion. With respect to the proposed l_1 -based CG type algorithms, the computational complexity is nearly the same as the conventional CG type algorithms. Note that the complexity of the CCG type, the AVF and the MWF algorithms is dependent on the rank D . This is a tradeoff between complexity and performance. We found that the rank of the proposed l_1 -based CCG algorithm with $D = 7$ works well (while the best rank for AVF and MWF is much larger), as will be verified in the following simulations. The low-rank characteristic will bring computational savings. The computational complexity of all algorithms is shown in Fig.2, where we use the best rank obtained from the simulations for

these algorithms ($D = 7$ for CCG type, $D = 18$ for AVF and $D = 14$ for MWF). We see that the proposed CG type algorithms have much lower complexity than **the** AVF and MWF algorithms.

Furthermore, the filter weights need to be repeatedly computed for target detection in airborne radar systems, especially in heterogeneous **environments**. In this case, our proposed algorithms can work in an iterative way and do not need to recompute all the filter weights, which can lead to significant computational savings. Usually, secondary data of the sliding window are used in detection procedures, where the parameter that defines the length of the sliding window is K . Assume $\mathbf{R}_i[K]$ denotes the estimated interference covariance matrix according to (16) and $\mathbf{w}_i[K]$ denotes the filter weight vector at the cell under test (CUT) of the i th range bin, respectively. Consider the case of the $i + 1$ th CUT, we first remove the impact of $i + 1$ th CUT, given by

$$\beta \mathbf{R}_{i+1}[K - 1] = \mathbf{R}_i[K] - \mathbf{x}[i + 1] \mathbf{x}^H[i + 1]. \quad (29)$$

Since an exponentially decaying data window is used, we do not need to remove the first snapshot used to compute the filter weights. Then, similarly, we consider the case of adding snapshots. Two snapshots, one is at the primary i th CUT and another is the new snapshot \mathbf{x}_{new} which was not included in the sliding window before, should be added to the $i + 1$ th CUT secondary data. The procedure can be written as

$$\mathbf{R}_{i+1}[K] = \beta \mathbf{R}_{i+1}[K - 1] + \beta \mathbf{x}_{\text{new}} \mathbf{x}_{\text{new}}^H + \mathbf{x}[i + 1] \mathbf{x}^H[i + 1]. \quad (30)$$

As for the filter weight vector $\mathbf{w}_{i+1}[K]$ at the $i + 1$ th CUT, it can be updated using the new interference covariance matrix and the filter weight vector $\mathbf{w}_i[K - 1]$.

In addition, the proposed algorithms adopt an adaptive filtering **approach**, which can obtain a near optimum interference rejection at a low cost [42]. The advantage of this approach is that filtering can be accomplished in a pipeline mode as the echo pulses come in. The required number of calculations for filtering can be realized easily with nowadays digital technology [41].

E. Analysis of the SA-STAP Algorithm

At this point, we have finished the derivation of the SA-STAP algorithms. The following simulation results will show that the proposed SA-STAP algorithms have a faster SINR convergence speed and better SINR steady-state performance than the conventional algorithms. This translates into a superior detection performance. However, why do the SA-STAP algorithms work is an interesting question. This section will try to explain that from two points of view.

First, to understand the behavior, we write the filter weight vector using the eigenvalue decomposition (EVD) of $\hat{\mathbf{R}}$. We assume that the eigenvalues of the estimated interference covariance matrix are $\hat{\gamma}_n$ with the corresponding eigenvectors denoted by \mathbf{u}_n , $n = 1, 2, \dots, NM$. The eigenvalues are ordered as,

$$\hat{\gamma}_1 \geq \hat{\gamma}_2 \geq \dots \geq \hat{\gamma}_{NM} = \hat{\gamma}_{\min}. \quad (31)$$

Thus, through the EVD, the estimated interference covariance matrix can be written as

$$\hat{\mathbf{R}} = \sum_{n=1}^{NM} \hat{\gamma}_n \mathbf{u}_n \mathbf{u}_n^H. \quad (32)$$

Substituting (32) into (14), the filter weight vector of the SA-STAP algorithm can be written as

$$\mathbf{w}_{SA} = \varsigma_{SA} \left\{ \mathbf{r}_t - \sum_{n=1}^{NM} \frac{\hat{\gamma}_n + \Delta_n - \hat{\gamma}_{min}}{\hat{\gamma}_n + \delta_{min}^{SA} + \Delta_n} (\mathbf{u}_n^H \mathbf{r}_t) \mathbf{u}_n \right\}, \quad (33)$$

where $\delta_{min}^{SA} = \min \left(\frac{\lambda}{|w_n| + \epsilon} \right)$, $n = 1, 2, \dots, NM$, Δ_n is the difference between $\frac{\lambda}{|w_n| + \epsilon}$ and δ_{min}^{SA} , and ς_{SA} is a scalar quantity, which does not affect the SINR.

By inspecting (33), we observe that the SA-STAP belongs to the class of diagonal loading STAP techniques in a sense. Moreover, it is equivalent to an adaptive diagonal loading technique, which will apply to each eigenbeam of the interference covariance matrix different weights and exploit the sparsity of the filter weights and the received data.

Second, we examine the relationship between the SINR performance and the l_1 -norm-sum quantity of the filter weights. Assume the scene is the same as the one with homogeneous environment introduced in the next section. We compute the SINR loss and the l_1 -norm-sum quantity of the filter weights against the number of snapshots using the SMI algorithm. The results are plotted in Fig.3. From the figure, we find that the better the SINR performance, the smaller the l_1 -norm-sum quantity of the filter weights. From this point of view, a constraint on the l_1 -norm-sum quantity of the filter weights can lead to fast convergence, which in fact exploits the sparsity of the received data and filter weights.

IV. PERFORMANCE ASSESSMENT

In this section, we assess the proposed SA-STAP algorithms using both simulated and measured data and compare them with the existing algorithms, such as the conventional CCG, MCG, MWF, AVF and LSMI algorithms. We measure the SINR, the SINR loss and the probability of detection curves, where the SINR and the SINR loss are defined as follows [4], respectively.

$$\text{SINR} = \frac{|\hat{\mathbf{w}}^H \mathbf{x}|^2}{|\hat{\mathbf{w}}^H \hat{\mathbf{R}} \hat{\mathbf{w}}|}, \quad (34)$$

$$\text{SINR}_{\text{loss}} = \frac{|\hat{\mathbf{w}}^H \mathbf{x}|^2}{|\hat{\mathbf{w}}^H \hat{\mathbf{R}} \hat{\mathbf{w}}| |\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}|}, \quad (35)$$

where \mathbf{R} is the exact interference covariance matrix at the detection range bin and $\hat{\mathbf{w}}$ is the estimated filter weights using the neighbor secondary data.

A. Simulated Data

Consider a monostatic sidelooking radar with $M = 10$ antenna elements and $N = 8$ pulses in one CPI, giving a space-time steering vector of length $L = 80$. We assume

a simulated scenario with the following parameters: half-wavelength spaced antennas, uniform transmit pattern, carrier frequency 450MHz, PRF set to 300Hz, platform velocity of 50m/s and height of 9000m, the clutter uniformly distributed from azimuth $-\pi/2$ to $\pi/2$ with clutter-to-noise-ratio (CNR) of 40dB, two jammer located at -45 and 60 with jammer-to-noise-ratio (JNR) of 40dB, the target located at 0° azimuth with Doppler frequency of 100Hz and signal-to-noise-ratio (SNR) of 0dB, and the thermal noise power is 0.01W. We consider the inner clutter motion (ICM) in simulated data. One common model, referred to as the Billingsley model, was developed by Billingsley of MIT Lincoln Laboratory [3]. The only parameters required to specify the clutter Doppler power spectrum are essentially the shape parameter b and the wind speed parameter ω . In this paper, we assume $b = 3.8$ and $\omega = 51.45$ miles per hour (mph). All presented results are averages over 100 independent Monte Carlo runs.

In our first example, we consider the SINR performance versus the rank D of the proposed l_1 -based CCG algorithm, the conventional CCG algorithm, the AVF algorithm and the MWF algorithm. A total of $K = 160$ snapshots are considered. The results in Fig.4 show that our proposed l_1 -based CCG algorithm can obtain its best performance when the rank is larger than $D = 7$. It requires a much lower rank to obtain its best performance than that of AVF ($D = 18$) and MWF ($D = 14$) algorithms. The low-rank characteristic will bring considerable computational savings, which is very important for STAP in radar systems. One should note that the performance of the conventional CCG algorithm will degrade when the rank is too large, while our proposed l_1 -based CCG can always keep good performance resulting in further robustness. Since the SINR performance is much worse when the rank is lower than the best rank, thus, we will use $D = 7$ for CCG type algorithms, $D = 18$ for the AVF algorithm and $D = 14$ for the MWF algorithm in the following examples.

In the second example, we examine the effect of the regularization parameter λ on the SINR performance of our proposed l_1 -based SMI and l_1 -based CG type algorithms. The results with simulated data are shown in Fig.5(a), (b) and (c). From the figures, we verify that too small or too large values of the regularization parameter λ will lead to performance degradation. This is because of the following conditions. When λ is too small, the proposed algorithms behave like the conventional algorithms, which results in slower convergence. And when λ is too large, many elements of the STAP filter weights will be shrunk to zero or be very close to zero, which leads to performance degradation due to decreasing number of DOFs to suppress the interference. Moreover, there is a range of values of λ in the proposed algorithms which can improve the SINR steady-state performance and also the convergence speed. Hence, the approach to choose λ introduced by [34] can be easily extended to our proposed algorithms. Due to space limitations, strategies to automatically choose λ will not be discussed in this paper and the reader is referred to [34].

In the next example, we evaluate the SINR loss performance against the number of snapshots of the proposed algorithms with the existing algorithms, as depicted in Fig.6. The curves show that: (1) the SINR performance of the proposed l_1 -based

SMI algorithm is a suboptimal algorithm, but exhibits the best performance compared with other algorithms. (2) l_1 -based CG type algorithms outperform conventional CG type algorithms in terms of convergence rate and steady-state performance; (3) the SINR performance of the l_1 -based CCG algorithm is better than the AVF and MWF algorithms. (4) Although the l_1 -based MCG algorithm shows slower SINR convergence than the MWF algorithm, we can obtain a better SINR performance when the number of snapshots is larger than 100. One should note that the proposed CG type algorithms have a much lower computational complexity than LSMI, AVF and MWF algorithms.

In the fourth example, we present the probability of detection P_d versus SNR with the target injected at the azimuth of 0° and Doppler frequency 100Hz in Fig.7. We assume the false alarm rate P_{fa} is set to 10^{-6} and the number of the secondary data is $K = 110$. The plots illustrate a similar trend to the SINR loss performance in the second example. Note that we obtain a performance gain of about 1dB in terms of SNR for l_1 -based CG type algorithms, as compared with conventional CG type algorithms.

Fig.8 shows the SINR performance against the target Doppler frequency at the azimuth of 0° with a total of $K = 100$ snapshots. Here, we suppose the potential Doppler frequency space is from -100Hz to 100Hz . The parameters of all algorithms are the same as the second example. The curves in the figure demonstrate a similar trend to the results of previous examples. Additionally, the l_1 -based SMI algorithm displays much better performance to the slow targets than other algorithms.

B. Measured Data

In this section, we apply the proposed algorithms to the Mountain-Top data set. This data set was collected from commanding sites (mountain tops) and radar motion is emulated using a technique developed at Lincoln Laboratories [6], [43]. The sensor consists of 14 elements and the data are organized in CPIs of 16 pulses. Here, we use the data file *t38pre01v1* CPI6, which could be obtained from the internet [44]. The pulse PRF was 625Hz and the instance bandwidth after pulse compression was 500kHz. There are 403 independent range samples available for the training data support. The clutter was located around 245° azimuth and the target was at 275° with a Doppler frequency 156Hz. All the data processed following are through pulse compression firstly. Note that the clutter and target have the same Doppler frequency, hence separation is impossible in the Doppler domain but possible in the spatial domain. The estimated clutter and target spectrum using all 403 samples is given in Fig.9, which shows a serious heterogeneity.

Fig.10(a) and Fig.10(b) display the STAP output power of all algorithms in the range of 147-162 km. Here, the interference covariance matrix is estimated using a symmetric sliding window with a total of 20 snapshots for Fig.10(a) and 40 snapshots for Fig.10(b). For each CUT, the snapshots do not include the 6 snapshots around the CUT. In the figures, we also give the unadapted weight vector, which equals the steering

vector $\mathbf{w} = \mathbf{s}$. We see that the target is clearly not detectable without adaptive processing. To have a clear comparison amongst different algorithms, we show the differences between the output power at the range bin of the target (154 km) and the next highest power peak in Table IV, where "-" presents the target not detectable. Here, 6 range bins around the range bin of the target is not used for comparison since they are the guide cells. Seen from the table, we find that: (1) the proposed l_1 -based SMI algorithm obtains the best detection performance in both situations, which is the same conclusion as that using simulated data; (2) the proposed l_1 -based CG type algorithms obtain better performance than the conventional CG type algorithms (although the proposed l_1 -based MCG algorithm has a pseudo target at the range 153km, when the secondary data record is 20 snapshots, the conventional MCG algorithm can not detect the target at all.); (3) the proposed l_1 -based CCG algorithm outperforms the AVF and MWF algorithms in both situations. Hence, we can conclude that our proposed algorithms show a robust performance in heterogeneous environments.

V. CONCLUSIONS

In this paper, we have proposed novel SA-STAP algorithms with l_1 -norm regularization for target detection in airborne radar systems. The proposed SA-STAP algorithms employed a sparse regularization to the MV cost function to exploit the sparsity of the received data and filter weights. To solve this kind of optimization problem, an l_1 -based SMI algorithm was directly developed, but it required a matrix inversion resulting in a high computational cost. Accordingly, we have proposed low-complexity SA-STAP algorithms based on CG techniques. A detailed analysis of the computational complexity and the performance of the SA-STAP algorithms were carried out. Simulation results with both simulated and measured data have shown that the proposed algorithms outperformed conventional algorithms and exhibited a robust performance in heterogeneous environments.

REFERENCES

- [1] Brennan L. E., and Reed L. S.: 'Theory of adaptive radar', *IEEE Trans. Aerosp. Electron. Syst.*, 1973, **9**, (2), pp.237-252
- [2] Ward J.: 'Space-time adaptive processing for airborne radar'. Technical Report 1015, MIT Lincoln laboratory, Lexington, MA, vol. Dec. 1994
- [3] Guerci J. R.: 'Space-time adaptive processing for radar' (Artech House, 2003)
- [4] Melvin W. L.: 'A stap overview', *IEEE Aerosp. Electron. Syst. Mag.*, 2004, **19**, (1), pp.19-35
- [5] Carlson B. D.: 'Covariance matrix estimation errors and diagonal loading in adaptive arrays', *IEEE Trans. Aerosp. Electron. Syst.*, 1988, **24**, (4), pp.397-401
- [6] Peckham C. D., Haimovich A. M., Ayoub T. F., Goldstein J. S., and Reed I. S.: 'Reduced-rank STAP performance analysis', *IEEE Trans. Aerosp. Electron. Syst.*, 2000, **36**, (2), pp.664-676
- [7] Pados D. A. and Karystinos G. N.: 'An iterative algorithm for the computation of the MVDR filter', *IEEE Trans. Sig. Proc.*, 2001, **49**, (2), pp.290-300
- [8] Pados D. A., Karystinos G. N., Batalama S. N., and Matyjas J. D.: 'Short-data-record adaptive detection', *Proc. IEEE Radar Conf.*, Apr. 2007, pp.357-361
- [9] Goldstein J. S., and Reed I. S.: 'Theory of partially adaptive radar', *IEEE Trans. Aerosp. Electron. Syst.*, 1997, **33**, (4), pp.1309-1325
- [10] Goldstein J. S., Reed I. S., and Zulch P. A.: 'Multistage Partially Adaptive STAP CFAR Detection Algorithm', *IEEE Trans. Aerosp. Electron. Syst.*, 1999, **35**, (2), pp.645-661

- [11] Scharf L. L., Chong E. K. P., Zoltowski M. D., Goldstein J. S., and Reed I. S.: ‘Subspace expansion and the equivalence of conjugate direction and multistage Wiener filters’, *IEEE Trans. Sig. Proc.*, 2008, **56**, (10), pp.5013-5019
- [12] Honig M. L., and Goldstein J. S.: ‘Adaptive reduced-rank interference suppression based on the multistage Wiener filter’, *IEEE Trans. Commun.*, 2002, **50**, (6), pp.986-994.
- [13] R. C. de Lamare, M. Haardt and R. Sampaio-Neto, Blind Adaptive Constrained Reduced-Rank Parameter Estimation based on Constant Modulus Design for CDMA Interference Suppression, *IEEE Transactions on Signal Processing*, vol. 56., no. 6, June 2008.
- [14] Chang P. S. and Willson Jr. A. N.: ‘Analysis of conjugate gradient algorithms for adaptive filtering’, *IEEE Trans. Sig. Proc.*, 2000, **48**, (2), pp.409-418
- [15] Jiang C., Li H., and Rangaswamy M.: ‘Conjugate gradient parametric adaptive matched filter’, Proc. IEEE Radar Conf., 2010, pp.740-745
- [16] Wang L., and de Lamare R.C.: ‘Constrained adaptive filtering algorithms based on conjugate gradient techniques for beamforming’, *IET Sig. Proc.*, 2010, **4**, (6), pp.686-697
- [17] Fa R., de Lamare R. C., and Wang L.: ‘Reduced-rank STAP schemes for airborne radar based on switched joint interpolation, decimation and filtering algorithm’, *IEEE Trans. Sig. Proc.*, 2010, **58**, (8), pp.4182-4194
- [18] Fa R., and de Lamare R. C.: ‘Knowledge-aided reduced-rank STAP for MIMO radar based on based on joint iterative constrained optimization of adaptive filters with multiple constraints’, Proc. IEEE Int. Conf. Acoust. Speech and Sig. Proc., 2010, pp.2762-2765
- [19] Fa R., de Lamare R. C., and Nascimento V. H.: ‘Knowledge-Aided STAP Algorithm Using Convex Combination of Inverse Covariance Matrices for Heterogeneous Clutter’, Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc., Sep. 2010, pp.2742-2745.
- [20] R. C. de Lamare, L. wang and R. Fa, ‘Adaptive Reduced-Rank LCMV Beamforming Algorithm Based on Joint Iterative Optimization of Filters: Design and Analysis’” *Elsevier Journal on Signal Processing*, Vol. 90, Issue 2, pp. 640-652, Feb 2010.
- [21] R. C. de Lamare and R. Sampaio-Neto, ‘Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering,’ *IEEE Transactions on Signal Processing*, vol. 57, no. 7, July 2009, pp. 2503 - 2514.
- [22] R.C. de Lamare, R. Sampaio-Neto, M. Haardt, ‘Blind Adaptive Constrained Constant-Modulus Reduced-Rank Interference Suppression Algorithms Based on Interpolation and Switched Decimation,’ *IEEE Trans. on Signal Processing*, vol.59, no.2, pp.681-695, Feb. 2011.
- [23] Melvin W. L., and Showman G. A.: ‘An approach to knowledge-aided covariance estimation’, *IEEE Trans. Aerosp. Electron. Syst.*, 2006, **42**, (3), pp.1021-1042
- [24] Guerci J. R., and Baranoski E. J.: ‘Knowledge-Aided adaptive radar at DARPA: an overview’, *IEEE Sig. Proc. Mag.*, 2006, **23**, (1), pp.41-50
- [25] Maria S., and Fuchs J. J.: ‘Application of the global matched filter to STAP data an efficient algorithmic approach’, Proc. IEEE Int. Conf. Acoust. Speech and Sig. Proc., 2006, pp. 14-19
- [26] Parker J. T., and Potter L. C.: ‘A Bayesian perspective on sparse regularization for STAP post-processing’, Proc. IEEE Radar Conf., May 2010, pp.1471-1475
- [27] Selesnick I. W., Pillai S. U., Li K. Y. and Himed B.: ‘Angle-Doppler processing using sparse regularization’, Proc. IEEE Int. Conf. Acoust. Speech and Sig. Proc., 2010, pp.2750-2753
- [28] Sun K., Zhang H., Li G., Meng H., and Wang X.: ‘A novel STAP algorithm using sparse recovery technique’, IGARSS, 2009, pp.336-339
- [29] Sun K., Zhang H., Li G., Meng H., and Wang X.: ‘Airborne radar STAP using sparse recovery of clutter spectrum’, 2010[Online]. Available: <http://arxiv.org/abs/1008.4185>
- [30] Sun K., Meng H., Wang Y., and Wang X.: ‘Direct data domain STAP using sparse representation of clutter spectrum’, *Sig. Proc.*, 2011, **91**, (9), pp.2222-2236
- [31] Sun K., Meng H., Lapierre F. D., and Wang X.: ‘Registration-based compensation using sparse representation in conformal-array STAP’, *Sig. Proc.*, 2011, **91**, (10), pp.2268-2276
- [32] Scott I., and Mulgrew B.: ‘Sparse LCMV beamformer design for suppression of ground clutter in airborne radar’, *IEEE Trans. Sig. Proc.*, 1995, **43**, (12), pp. 2843-2851
- [33] Yang Z., de Lamare R. C., and Li X.: ‘ L_1 regularized STAP algorithm with a generalized sidelobe canceler architecture for airborne radar’, Proc. IEEE Workshop on Stat. Sig. Proc., Nice, France, 2011, pp.329-332
- [34] Yang Z., de Lamare R. C., and Li X.: ‘ L_1 -regularized STAP algorithms with a generalized sidelobe canceler architecture for airborne radar’, *IEEE Trans. Signal Process.*, 2012, **60**, (2), pp.674-686
- [35] Wang H., and Cai L.: ‘On adaptive spartial-temporal processing for airborne surveillance radar systems’, *IEEE Trans. Aerosp. Electron. Syst.*, 1994, **30**, (3), pp.660-670
- [36] Melvin W. L.: ‘Space-time adaptive radar performance in heterogeneous clutter’, *IEEE Trans. Aerosp. Electron. Syst.*, 2000, **36**, (2), pp.621-633
- [37] Angelosante D., Bazerque J. A., and Giannakis G. B.: ‘Online adaptive estimation of sparse signals: where RLS meets the l_1 -norm’, *IEEE Trans. Sig. Proc.*, 2010, **58**, (7), pp. 3436-3446
- [38] Elad M., ‘Why simple shrinkage is still relevant for redundant representations?’ *IEEE Trans. Inf. Theory*, **52**, (12), 2006, pp. 5559-5569
- [39] Zibulevsky M., and Elad M.: ‘ L_1 - L_2 optimization in signal and image processing’, *IEEE Sig. Proc. Mag.*, 2010, **27**, (3), pp. 76-88
- [40] Jin J., Gu Y., and Mei S.: ‘A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework’, *IEEE Journal of Selected Topics in Sig. Proc.*, 2010, **4**, (2), pp.409-420
- [41] Van Trees H. L.: ‘Optimal array processing, part IV of detection, estimation and modulation theory’ (New York: John Wiley & Sons, Inc., 2002)
- [42] Klemm R.: ‘Introduction to space-time adaptive processing’, Proc. IEEE Colloquium on Space-Time Adaptive Proc., 1998
- [43] Titi G. W., and Marshall D. F.: ‘The ARPA/NAVY mountaintop program: adaptive signal processing for airborne early warning radar’, Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc., May 1996, pp.1165-1168
- [44] http://spib.rice.edu/spib/mtn_top.html

TABLE I
THE l_1 -BASED CCG ALGORITHM

Initialization:

$$\hat{\mathbf{R}}[0] = \delta \mathbf{I}, \mathbf{v}[0] = \mathbf{s}, \eta_{CCG},$$

Recursion: For each snapshot $k = 1, \dots, L$

STEP 1: Start:

$$\hat{\mathbf{R}}[k] = \beta \hat{\mathbf{R}}[k-1] + \mathbf{x}[k] \mathbf{x}^H[k],$$

$$\mathbf{\Lambda}[k] = \text{diag} \left\{ \frac{1}{|w_1[k-1]| + \epsilon}, \dots, \frac{1}{|w_{NM}[k-1]| + \epsilon} \right\},$$

$$\mathbf{G}[k] = \hat{\mathbf{R}}[k] + \lambda \mathbf{\Lambda}[k],$$

$$\mathbf{g}_0[k] = \mathbf{s} - \mathbf{G}[k] \mathbf{v}_0[k], \mathbf{p}_1[k] = \mathbf{g}_0[k], \rho_0[k] = \mathbf{g}_0^H[k] \mathbf{g}_0[k],$$

STEP 2: For $d = 1, \dots, D$ **and** $\rho_{d-1}[k] > \eta_{CCG}$

$$\mathbf{z}_d[k] = \mathbf{G}[k] \mathbf{p}_d[k],$$

$$\alpha_d[k] = [\mathbf{p}_d^H[k] \mathbf{z}_d[k]]^{-1} \rho_{d-1}[k],$$

$$\mathbf{v}_d[k] = \mathbf{v}_{d-1}[k-1] + \alpha_d[k] \mathbf{p}_d[k],$$

$$\mathbf{g}_d[k] = \mathbf{g}_{d-1}[k] - \alpha_d[k] \mathbf{z}_d[k],$$

$$\rho_d[k] = \mathbf{g}_d^H[k] \mathbf{g}_d[k],$$

$$\nu_d[k] = \frac{\rho_d[k]}{\rho_{d-1}[k]},$$

$$\mathbf{p}_{d+1}[k] = \mathbf{g}_d[k] + \nu_d[k] \mathbf{p}_d[k],$$

STEP 3: After end STEP 2

$$\mathbf{v}_0[k+1] = \mathbf{v}_d[k],$$

$$\mathbf{w}[k] = \frac{\mathbf{v}_0[k]}{\mathbf{s}^H \mathbf{v}_0[k]},$$

Final output:

$$y[k] = \mathbf{w}^H[k] \mathbf{x}[k].$$

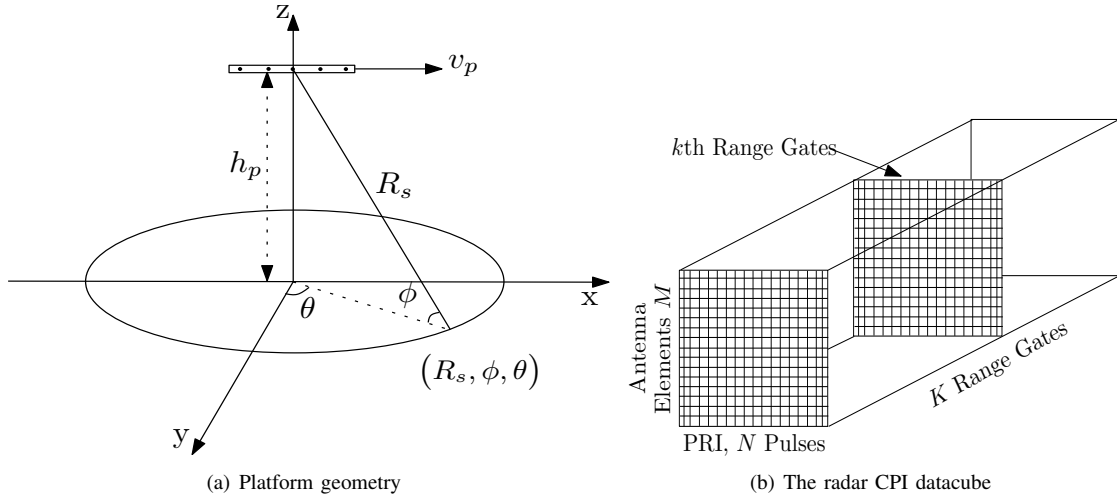


Fig. 1. The radar platform geometry and the radar CPI datacube

TABLE II
THE l_1 -BASED MCG ALGORITHM

Initialization:

$$\hat{\mathbf{R}}[0] = \delta \mathbf{I}, \mathbf{v}[0] = \mathbf{s}, \mathbf{w}[0] = \mathbf{s}, \mathbf{g}[0] = \mathbf{s}, \mathbf{p}[1] = \mathbf{s},$$

Recursion: For each snapshot $k = 1, \dots, L$

$$\hat{\mathbf{R}}[k] = \beta \hat{\mathbf{R}}[k-1] + \mathbf{x}[k] \mathbf{x}^H[k],$$

$$\Lambda[k] = \text{diag} \left\{ \frac{1}{|w_1[k-1]| + \epsilon}, \dots, \frac{1}{|w_{NM}[k-1]| + \epsilon} \right\},$$

$$\mathbf{G}[k] = \hat{\mathbf{R}}[k] + \lambda \Lambda[k],$$

$$\alpha[k] = [\mathbf{p}^H[k] \mathbf{G}[k] \mathbf{p}[k]]^{-1} \times \{ \beta [\mathbf{p}^H[k] \mathbf{g}[k-1] - \mathbf{p}^H[k] \mathbf{s}] - \mu \mathbf{p}^H[k] \mathbf{g}[k-1] \},$$

$$\mathbf{v}[k] = \mathbf{v}[k-1] + \alpha[k] \mathbf{p}[k],$$

$$\mathbf{g}[k] = (1 - \beta) \mathbf{s} + \beta \mathbf{g}[k-1] - \alpha[k] \mathbf{G}[k] \mathbf{p}[k] - \{ \lambda [1 - \beta] \Lambda[k] + \mathbf{x}[k] \mathbf{x}^H[k] \} \mathbf{v}[k-1],$$

$$\nu[k] = \frac{[\mathbf{g}[k] - \mathbf{g}[k-1]]^H \mathbf{g}[k]}{\mathbf{g}^H[k-1] \mathbf{g}[k-1]},$$

$$\mathbf{p}[k+1] = \mathbf{g}[k-1] + \nu[k] \mathbf{p}[k],$$

$$\mathbf{w}[k] = \frac{\mathbf{v}[k]}{\mathbf{s}^H \mathbf{v}[k]},$$

Output:

$$y[k] = \mathbf{w}^H[k] \mathbf{x}[k].$$

TABLE III
COMPARISON OF THE COMPUTATIONAL COMPLEXITY

Algorithm	Additions	Multiplications
LSMI	$O(L^3) + O(L^2)$	$O(L^3) + O(L^2)$
l_1 -based SMI	$O(L^3) + O(L^2)$	$O(L^3) + O(L^2)$
MWF	$DL^2 + (4D - D^2)L + \frac{D^3}{3} - \frac{3D^2}{2} - \frac{D}{3}$	$2DL^2 + (5D - D^2)L + \frac{2D^3}{3} - 2D^2 + \frac{16D}{3}$
AVF	$(2D+1)L^2 + (4D+1)L - 4D - 1$	$2(D+1)L^2 + 7DL + L$
CCG	$(D+2)L^2 + (4D+2)L - 2D - 2$	$(D+3)L^2 + 5DL + 3L$
MCG	$3L^2 + 10L - 4$	$4L^2 + 13L + 2$
l_1 -based CCG	$(D+3)L^2 + (4D+3)L - 2D - 2$	$(D+3)L^2 + 7DL + L$
l_1 -based MCG	$5L^2 + 11L - 4$	$4L^2 + 13L + 2$

TABLE IV
RESULTS OF MOUNTAIN TOP DATA

Algorithms	20 snapshots	40 snapshots
unadapted	-	-
LSMI	7.1dB	12.4dB
l_1 -based SMI	9.3dB	15.1dB
MWF	4.6dB	10.0dB
AVF	5.6dB	11.9dB
CCG	5.2dB	11.9dB
MCG	-	6.2dB
l_1 -based CCG	7.1dB	13.0dB
l_1 -based MCG	0dB	9.5dB

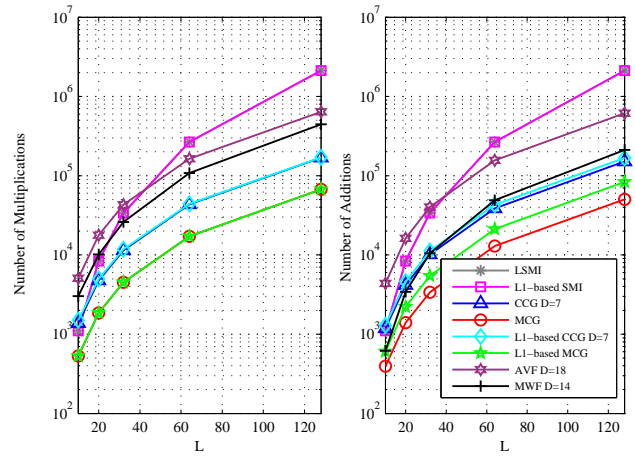


Fig. 2. The computational complexity per snapshot.

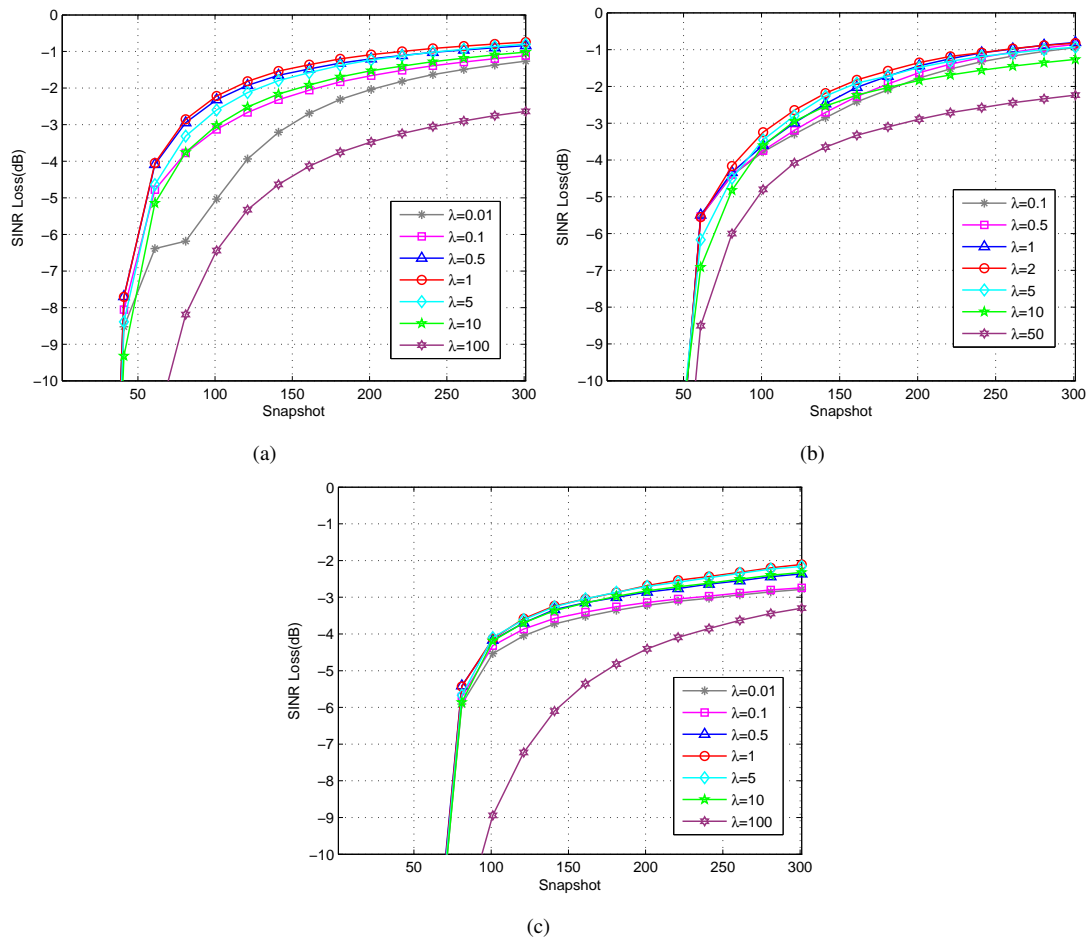


Fig. 5. Impact of regularization parameter λ on the SINR performance against the number of snapshots. Parameters: $\beta = 0.9998$, $\eta_{CCG} = 10^{-5}$ and $\mathbf{R}[0] = 0.001\mathbf{I}$ for CG type algorithms. (a) SINR performance of the l_1 -based SMI algorithm with different λ ; (b) SINR performance of the l_1 -based CCG algorithm with different λ ; (c) SINR performance of the l_1 -based MCG algorithm with different λ .

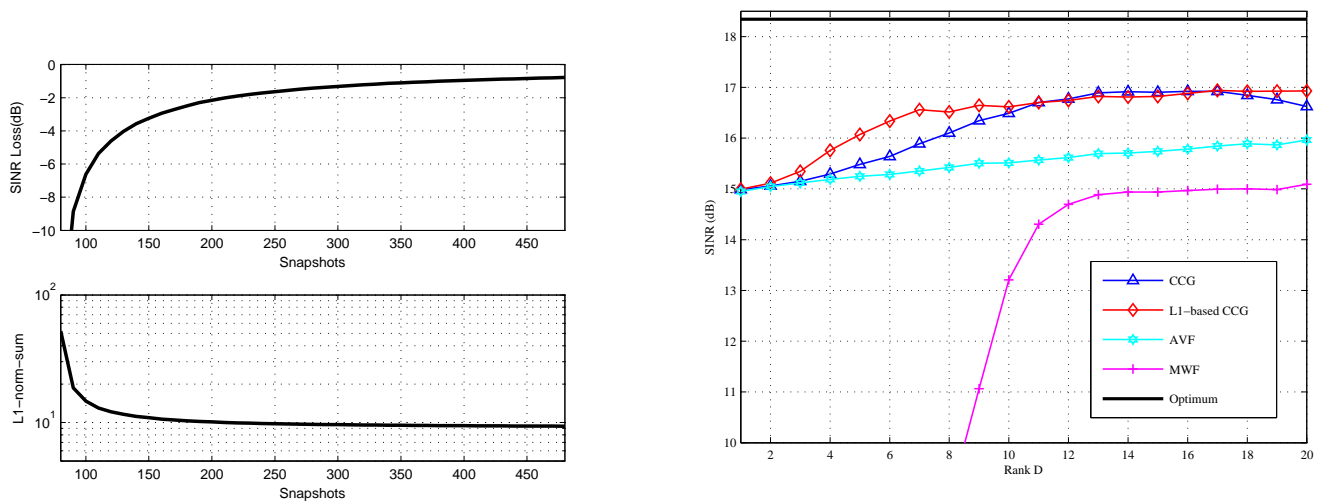


Fig. 3. The relationship between the SINR performance and the l_1 -norm-sum quantity of the filter weights.

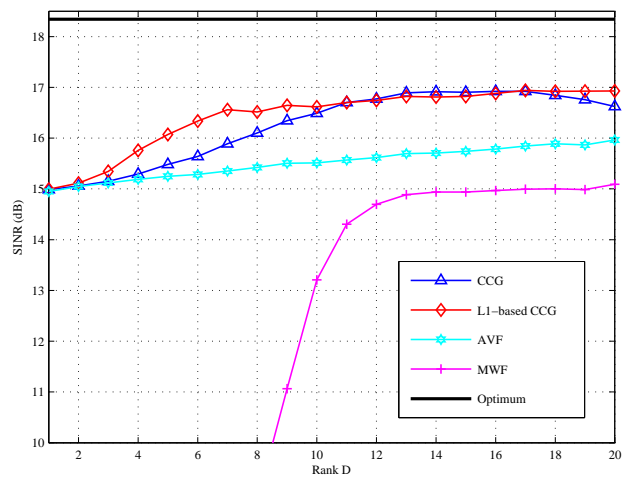


Fig. 4. The SINR performance versus the rank D . Parameters: the diagonal loading factor for AVF and MWF algorithms is 10dB to the thermal noise power; $\beta = 0.9998$, $\eta_{CCG} = 10^{-5}$ and $\mathbf{R}[0] = 0.001\mathbf{I}$ for CCG type algorithms; $\lambda = 2$ for l_1 -based CCG algorithms.

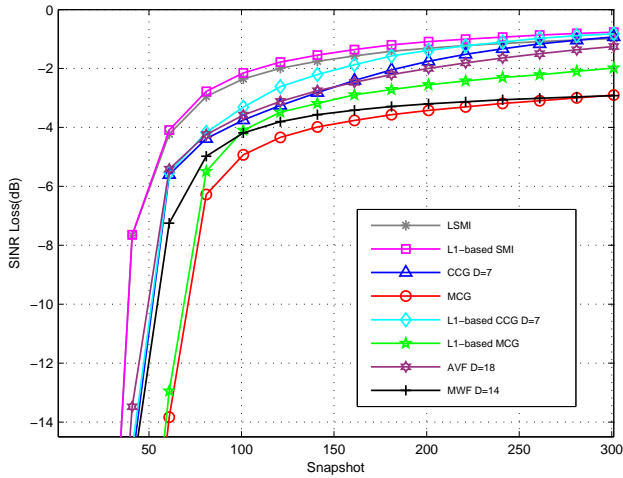


Fig. 6. The SINR loss performance against the number of snapshots. Parameters: the diagonal loading factor for LSMI, AVF and MWF algorithms is 10dB to the thermal noise power; $\beta = 0.9998$ and $\mathbf{R}_l[0] = 0.001\mathbf{I}$ for CG type algorithms; $\lambda = 1$ for the l_1 -based SMI, $\lambda = 2$ for the l_1 -based CCG algorithm and $\lambda = 1$ for the l_1 -based MCG algorithm; $\eta_{CCG} = 10^{-5}$.

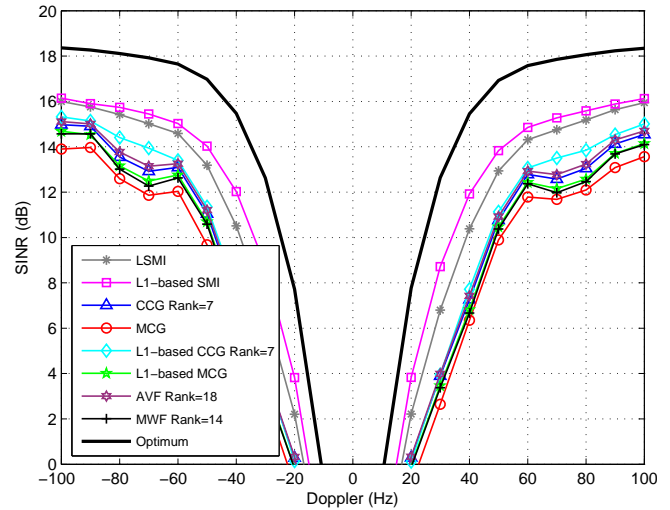


Fig. 8. SINR performance against Doppler frequency with $K = 100$ snapshots and Doppler frequency space from -100 to 100 Hz. The other parameters are the same as the second example.

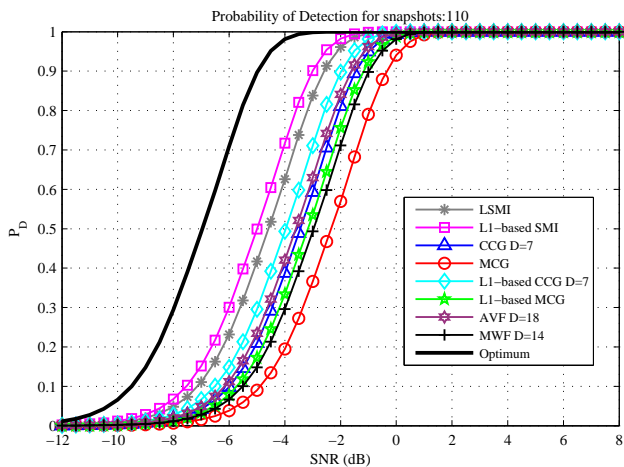


Fig. 7. Probability of detection performance versus SNR with $K = 110$ snapshots. $P_{fa} = 10^{-6}$ and the other parameters are the same as the second example.

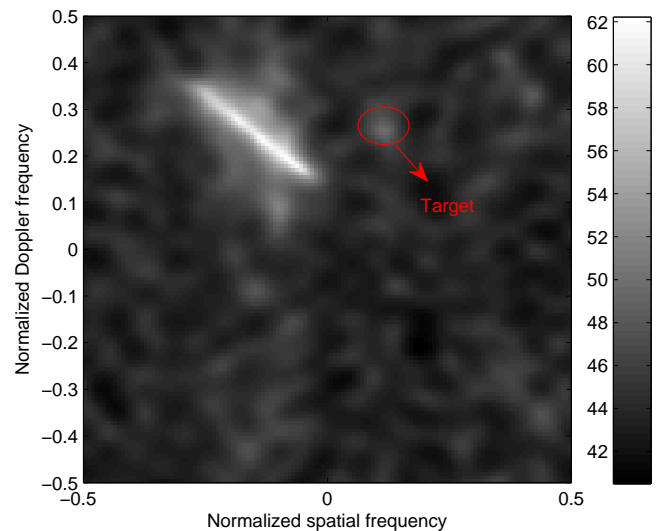
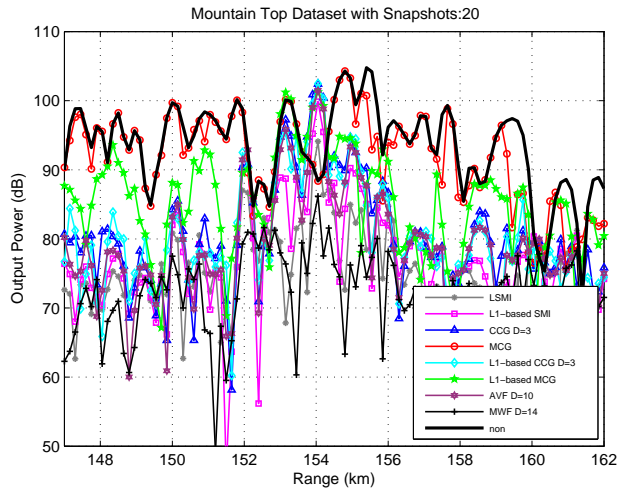
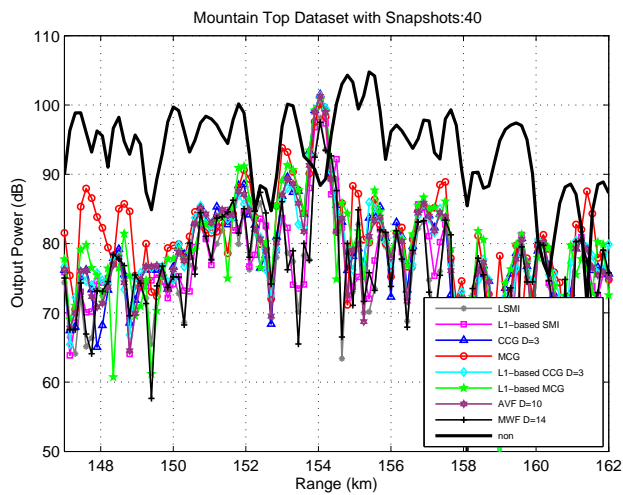


Fig. 9. The estimated clutter and target spectrum using all 403 samples. The target is located at azimuth of 15° referred to the boresight with normalized Doppler frequency of 0.25.



(a) The output power after STAP using 20 snapshots



(b) The output power after STAP using 40 snapshots

Fig. 10. Range plots of Mountain Top data using different STAP algorithms. $D = 3$ for CCG type algorithms, $D = 18$ for the AVF algorithm and $D = 14$ for the MWF algorithm. "Non" presents the case using the unadapted weight vector.