TUDelft

Delft University of Technology

Fast ℓ_1-regularized space-Time adaptive processing using alternating direction method of multipliers

Qin, Lilong; Wu, Manqing; Wang, Xuan; Dong, Zhen

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Fast ℓ_1-regularized space-time adaptive processing using alternating direction method of multipliers

Lilong Qin
Manqing Wu
Xuan Wang
Zhen Dong

SPIE.

Fast ℓ_1-regularized space-time adaptive processing using alternating direction method of multipliers

Lilong Qin,[a,b,]* Manqing Wu,[c] Xuan Wang,[d] and Zhen Dong[a]
[a]National University of Defense Technology, School of Electronic Science and Engineering, Changsha, China
[b]Aalto University, Department of Signal Processing and Acoustics, Espoo, Finland
[c]China Electronics Technology Group Corporation,
China Academy of Electronics and Information Technology, Beijing, China
[d]Delft University of Technology, Electrical Engineering,
Mathematics and Computer Science, Delft, The Netherlands

Abstract. Motivated by the sparsity of filter coefficients in full-dimension space-time adaptive processing (STAP) algorithms, this paper proposes a fast ℓ_1-regularized STAP algorithm based on the alternating direction method of multipliers to accelerate the convergence and reduce the calculations. The proposed algorithm uses a splitting variable to obtain an equivalent optimization formulation, which is addressed with an augmented Lagrangian method. Using the alternating recursive algorithm, the method can rapidly result in a low minimum mean-square error without a large number of calculations. Through theoretical analysis and experimental verification, we demonstrate that the proposed algorithm provides a better output signal-to-clutter-noise ratio performance than other algorithms. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10 .1117/1.JRS.11.026004]

1 Introduction

Space-time adaptive processing (STAP) can effectively suppress strong ground/sea clutter and improve the moving target indication performance for airborne/spaceborne radar systems.[1] In full-dimension STAP algorithms, however, a large number of independent and identically distributed (I.I.D.) training snapshots are required to yield an average signal-to-clutter-noise ratio (SCNR) loss of ∼3 dB.[2] Moreover, full-dimension STAP algorithms have a high system complexity and require many memory elements.[3] In practical applications, it is generally difficult to satisfy these requirements.

To date, many algorithms have been proposed to overcome the drawbacks of full-dimension STAP algorithms. Reduced-rank STAP algorithms can reduce the clutter space while maintaining the performance of fully STAP algorithms.[4,5] Consequently, the required number of snapshots can be reduced. However, eigenvalue decomposition is used, which is computationally expensive. To reduce the computational expense and the number of training snapshots simultaneously, some typical reduced-dimension STAP algorithms have been proposed, such as the joint domain localized approach, auxiliary channel processing, etc.[6–8] However, the nonadaptive selection of the reduced-dimension projection matrix, which relies on intuitive experience, results in a performance degradation to a certain extent.[2]

The sparsity of the filter coefficients in STAP has recently been studied, and the theoretical framework for sparsity-based STAP algorithms using the ℓ_1-regularized constraint, which is the

---

*Address all correspondence to: Lilong Qin, E-mail: qinlilong13@nudt.edu.cn

so-called least absolute shrinkage and selection operator (LASSO), has been established.[9–12] The classical algorithms for solving the LASSO problem adopt convex optimization, e.g., the interior point algorithm, to obtain a sparse solution. The complexity of the algorithms can be very high when the size of the problem is large, which is not pragmatic in practice. To effectively solve the optimization problem, the ℓ_1-regularized recursive least-squares STAP (RLS-STAP) algorithm,[13] the ℓ_1-regularized least-mean-square STAP algorithm,[14] and the homotopy-STAP algorithm[15] have been proposed. Compared with conventional STAP methods, sparsity-based STAP techniques have been shown to provide high resolution and exhibit better performance than conventional STAP algorithms.[16]

The alternating direction method of multipliers (ADMM) is a technique used to combine the decomposability of dual ascent with the rapid convergence speed of the method of multipliers.[17,18] This technique is well suited for solving the optimization problems of the ℓ_1 constraint, particularly large-scale problems.[19] The ADMM technique can converge within a few tens of iterations, which is acceptable in practical use.[20] In this study, according to the optimal criterion of minimizing the mean-square error, we propose an algorithm based on the ADMM technique to solve the ℓ_1-regularized STAP problem. The proposed method provides better performance with a small number of I.I.D. training snapshots and without a large number of calculations.

The reminder of this paper is organized as follows. The system model of the generalized side-lobe canceler (GSC) form of the sparsity-based STAP is introduced in Sec. 2. In Sec. 3, the theory of the ADMM algorithm is introduced, and the ℓ_1-regularized ADMM-STAP algorithm is proposed. The associated optimization problem is formulated and solved analytically. The performance improvement of the proposed algorithm is shown in Sec. 4. Section 5 provides the conclusion.

Notation: In this paper, a variable, a column vector, and a matrix are represented by a lowercase letter, a lowercase bold letter, and a capital bold letter, respectively. The operations of transposition, complex conjugation, and conjugate transposition are denoted by $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$, respectively. The symbol $\otimes$; denotes the Kronecker product, and the symbol $\|\cdot\|_n$ denotes the ℓ_n-norm operator. $E(\mathbf{x})$ denotes the expected value of $\mathbf{x}$, $|x|$ indicates the absolute value of x, and $(x)_+ \triangleq \max(0, x)$. $\mathrm{sign}(\cdot)$ is the component-wise sign function.[13]

2 Background and Problem Formulation

2.1 *System Model*

The STAP technique is known for its ability to suppress clutter energy interference while detecting moving targets. Consider an airborne radar system equipped with a uniform linear array (ULA) consisting of N receiving elements, as shown in Fig. 1. The radar transmits K
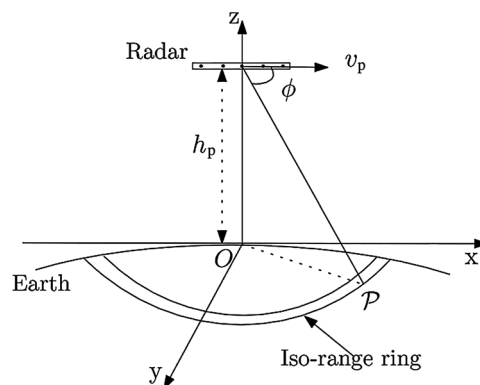

Fig. 1 Radar platform flies at speed v_p along the azimuth direction (x-axis). Without loss of generality, the center of elements is defined as the origin of coordinates. h_p is the flight height, and ϕ represents the AOA of the clutter patch in the isorange clutter ring.

identical pulses at a constant pulse repetition frequency (PRF) $f_r \triangleq 1/T_r$ during a coherent processing interval (CPI), where T_r is the pulse repetition interval. The received signal from the range bin of interest is represented as $\mathbf{x} = \mathbf{x}_t + \mathbf{x}_c + \mathbf{n}$, where $\mathbf{x}_t$ is the target vector, $\mathbf{x}_c$ is the clutter vector, and $\mathbf{n}$ is the thermal noise vector with noise power σ_n^2 on each channel and pulse. The space-time clutter vector can be represented as[21]

$$\mathbf{x}_c = \sum_{n=1}^{N_c} \sigma_{c,n} \mathbf{v}(f_{d,n}, f_{s,n}), \tag{1}$$

where N_c denotes the number of clutter patches in the range bin of interest and $\sigma_{c,n}$ denotes the random complex reflection coefficient. $f_{d,n} \triangleq 2v_p T_r \sin\phi_n/\lambda$ and $f_{s,n} \triangleq d \sin\phi_n/\lambda$ are the Doppler frequency and spatial frequency for the n'th clutter patch, respectively, where λ is the wavelength and d is the innersensor spacing of the ULA. $\mathbf{v}(f_{d,n}, f_{s,n}) \in \mathbb{C}^{NK \times 1}$ is the space-time steering vector, which is defined as a Kronecker product of the temporal and spatial steering vectors, i.e., $\mathbf{v}(f_d, f_s) = \mathbf{v}_d(f_d) \otimes \mathbf{v}_s(f_s)$, where

$$\begin{aligned} \mathbf{v}_d(f_d) &= [\, 1 \quad \exp(j2\pi f_d) \quad \cdots \quad \exp[j2\pi f_d(K-1)]\,]^T \\ \mathbf{v}_s(f_s) &= [\, 1 \quad \exp(j2\pi f_s) \quad \cdots \quad \exp[j2\pi f_s(N-1)]\,]^T. \end{aligned} \tag{2}$$

The target vector is $\mathbf{x}_t = \sigma_t \mathbf{v}(f_{d,t}, f_{s,t})$, where $f_{d,t} \triangleq 2v_p T_r \sin\phi_t/\lambda + 2v_t T_r/\lambda$ and $f_{s,t} \triangleq d \sin\phi_t/\lambda$. v_t is the radial velocity of the moving target, and ϕ_t represents the angle of arrival (AOA) of the target. Note that in the following, $\mathbf{v}(f_{d,t}, f_{s,t})$ is rewritten as $\mathbf{v}_t$ for convenience.

To clearly illustrate how the STAP method works, the GSC form of the STAP method is shown in Fig. 2. $\mathbf{B} \in \mathbb{C}^{NK \times (NK-1)}$ is the signal blocking matrix, which satisfies $\mathbf{B}^H \mathbf{v}_t = \mathbf{0}$ and $\mathbf{B}\mathbf{B}^H = \mathbf{I}$. Generally, $\mathbf{B}$ can be obtained by singular value decomposition (SVD):

$$[\,\mathbf{U} \quad \mathbf{S} \quad \mathbf{V}\,] = \text{svd}(\mathbf{v}_t^H), \qquad \mathbf{B} = \mathbf{V}(:, 2:NK). \tag{3}$$

After the transformation by $\mathbf{b} = \mathbf{B}^H \mathbf{x} \in \mathbb{C}^{(NK-1) \times 1}$, $NK - 1$ clutter data are available. In the full-dimension STAP, all the data are selected to cancel the clutter. The output is

$$y = d_0 - \boldsymbol{\omega}_b^H \mathbf{b}, \tag{4}$$

where $d_0 = \mathbf{v}_t^H \mathbf{x}$ and $\boldsymbol{\omega}_b = \mathbf{R}_b^{-1} \mathbf{r}_{bd}$. $\mathbf{R}_b = E(\mathbf{b}\mathbf{b}^H)$ is the clutter covariance matrix, and $\mathbf{r}_{bd} = E(\mathbf{b}d_0^*)$ is the cross-correlation vector between d_0 and $\mathbf{b}$. The output clutter power can be computed as

$$P = \mathbf{v}_t^H \mathbf{R}_x \mathbf{v}_t - \mathbf{r}_{bd}^H \mathbf{R}_b^{-1} \mathbf{r}_{bd}, \tag{5}$$

where $\mathbf{R}_x = E(\mathbf{x}\mathbf{x}^H)$ is the input covariance matrix. The output SCNR can be expressed as

$$\xi = \frac{NM|\alpha|^2}{\mathbf{v}_t^H \mathbf{R}_x \mathbf{v}_t - \mathbf{r}_{bd}^H \mathbf{R}_b^{-1} \mathbf{r}_{bd}}. \tag{6}$$
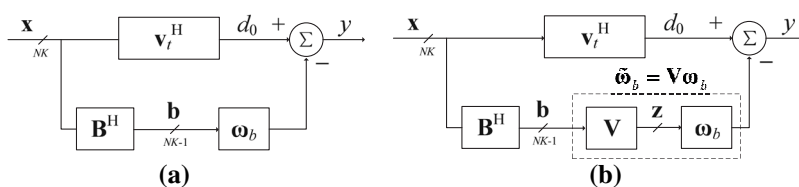

Fig. 2 (a) GSC form of the conventional STAP and (b) GSC form of the sparsity-based STAP.

Maximizing the output SCNR is equivalent to maximizing the detection probability. However, $\mathbf{R}_b$ and $\mathbf{r}_{bd}$ are unknown in practice, and the secondary training snapshots are required to estimate these parameters.[15] The best performance can be achieved if there are sufficient I.I.D. training snapshots. However, in many practical cases, it is impossible to obtain sufficient snapshots, and the performance degrades significantly.

2.2 *Sparsity-Based STAP*

According to the STAP theory, it has been shown that the rank of clutter covariance is far lower than the DOFs of the system.[22,23] Consequently, some RR-STAP and RD-STAP algorithms have been used to reduce the filter length, i.e., the filter coefficient vector obtained by full-dimension STAP is sparse.[14] Hence, in the GSC form of the sparsity-based STAP algorithm (see Fig. 2), the filter coefficient vector $\boldsymbol{\omega}_b$ can be replaced by $\tilde{\boldsymbol{\omega}}_b = \mathbf{V}\boldsymbol{\omega}_b$, where $\mathbf{V} \overset{\Delta}{=} \mathrm{diag}(\mathbf{v})$ and $\mathbf{v} \in \mathbb{C}^{NK-1}$ denote a sparse vector. Then, we obtain

$$\mathbf{z} = \mathbf{V}^{\mathrm{H}}\mathbf{b} \in \mathbb{C}^{(NK-1)\times 1}. \tag{7}$$

The output of the sparsity-based STAP is

$$y_{\mathrm{r}} = d_0 - \boldsymbol{\omega}_b^{\mathrm{H}}\mathbf{z} = \begin{bmatrix} 1 & \boldsymbol{\omega}_b^{\mathrm{H}} - \boldsymbol{\omega}_b^{\mathrm{H}}\mathbf{V}^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} y \\ \mathbf{b} \end{bmatrix}. \tag{8}$$

Hence, the output clutter power for the sparsity-based STAP can be computed as

$$P_{\mathrm{r}} = \mathbf{v}_{\mathrm{t}}^{\mathrm{H}}\mathbf{R}_x\mathbf{v}_{\mathrm{t}} - \mathbf{r}_{bd}^{\mathrm{H}}\mathbf{R}_b^{-1}\mathbf{r}_{bd} + \boldsymbol{\varepsilon}^{\mathrm{H}}\mathbf{R}_b\boldsymbol{\varepsilon}, \tag{9}$$

where $\boldsymbol{\varepsilon} = \boldsymbol{\omega}_b - \tilde{\boldsymbol{\omega}}_b$ is the weight error vector caused by the sparsity constraint. Note that the target signal power is not affected by the sparsity constraint. The output SCNR can be expressed as

$$\xi_{\mathrm{r}} = \frac{NM|\alpha|^2}{\mathbf{v}_{\mathrm{t}}^{\mathrm{H}}\mathbf{R}_x\mathbf{v}_{\mathrm{t}} - \mathbf{r}_{bd}^{\mathrm{H}}\mathbf{R}_b^{-1}\mathbf{r}_{bd} + \boldsymbol{\varepsilon}^{\mathrm{H}}\mathbf{R}_b\boldsymbol{\varepsilon}}. \tag{10}$$

Hence, the aim is to minimize the mean-square error $\boldsymbol{\varepsilon}^{\mathrm{H}}\mathbf{R}_b\boldsymbol{\varepsilon}$. The objective function of the minimization problem can be rewritten as

$$\boldsymbol{\varepsilon}^{\mathrm{H}}\mathbf{R}_b\boldsymbol{\varepsilon} = \mathbf{r}_{bd}^{\mathrm{H}}\mathbf{R}_b^{-1}\mathbf{r}_{bd} - \mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b. \tag{11}$$

$\tilde{\boldsymbol{\omega}}_b$ is sparse, i.e., most of its elements are considerably smaller than the others. Hence, the minimization problem can be expressed as

$$\min \quad -\mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b + \lambda\|\tilde{\boldsymbol{\omega}}_b\|_0, \tag{12}$$

where λ is the regularization parameter for regulating the sparseness of $\tilde{\boldsymbol{\omega}}_b$. However, the ℓ_0-norm problem is nonconvex. Consequently, it is intractable even for optimization problems with a moderate size. Equation (12) can be further programmed as an LASSO algorithm

$$\min \quad -\mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b + \lambda\|\tilde{\boldsymbol{\omega}}_b\|_1. \tag{13}$$

In contrast to Eq. (12), Eq. (13) is convex and can be solved by convex optimization algorithms, such as the interior point method (IPM). The complexity of IPM-STAP can be very high when the size of the problem is large, which is not pragmatic in practice.

3 Proposed ℓ_1-Regularized STAP Algorithm

3.1 *Variable Splitting*

In general, the ADMM algorithm can converge rapidly when a modest-accuracy result is acceptable. Fortunately, this is the case for the parameter estimation problem in the STAP application that we are considering. For statistical problems, solving a parameter estimation problem to a very high accuracy often yields little improvement.[19] The ADMM-STAP algorithm is based on the algorithm of variable splitting, i.e., we split the variable $\tilde{\boldsymbol{\omega}}_b$ into a pair of variables, say, $\tilde{\boldsymbol{\omega}}_b$ and $\mathbf{z}$, and add a constraint that the two variables are equal. Moreover, the objective function is split as the sum of two functions, and then we minimize the sum of the two functions. Explicitly, Eq. (13) can be rewritten in the ADMM form

$$\min_{\tilde{\boldsymbol{\omega}}_b, \mathbf{z}} \mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b + \lambda\|\mathbf{z}\|_1 \quad \text{s.t. } \tilde{\boldsymbol{\omega}}_b = \mathbf{z}. \tag{14}$$

The problems of Eqs. (13) and (14) are clearly equivalent. In many cases, it is easier to solve the constrained problem Eq. (14) than the original unconstrained problem. As in the method of multipliers, the augmented Lagrangian function is formed as[19,20]

$$\mathcal{L}_\rho(\tilde{\boldsymbol{\omega}}_b, \mathbf{z}, \mathbf{y}) = -\mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b + \lambda\|\mathbf{z}\|_1 + (\rho/2)\|\tilde{\boldsymbol{\omega}}_b - \mathbf{z}\|_2^2 + \mathbf{y}^{\mathrm{H}}(\tilde{\boldsymbol{\omega}}_b - \mathbf{z}), \tag{15}$$

where $\rho > 0$ is the augmented Lagrangian parameter and $\mathbf{y}$ is a vector of Lagrange multipliers.

3.2 ℓ_1-*Regularized ADMM-STAP*

Define the residual and the scaled dual variable as $\mathbf{r} = \tilde{\boldsymbol{\omega}}_b - \mathbf{z}$ and $\mathbf{d} = (1/\rho)\mathbf{y}$, respectively. Then, we have

$$(\rho/2)\|\tilde{\boldsymbol{\omega}}_b - \mathbf{z}\|_2^2 + \mathbf{y}^{\mathrm{H}}(\tilde{\boldsymbol{\omega}}_b - \mathbf{z}) = (\rho/2)\|\mathbf{r}\|_2^2 + \mathbf{y}^{\mathrm{H}}\mathbf{r} \quad = (\rho/2)\|\mathbf{r} + \mathbf{d}\|_2^2 - (\rho/2)\|\mathbf{d}\|_2^2. \tag{16}$$

Subsequently, the ADMM-STAP algorithm can be rewritten in a convenient form

$$\tilde{\boldsymbol{\omega}}_b^{(k+1)} = \arg\min_{\tilde{\boldsymbol{\omega}}_b}(-\mathbf{r}_{bd}^{\mathrm{H}}\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^{\mathrm{H}}\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b + (\rho/2)\|\tilde{\boldsymbol{\omega}}_b - \mathbf{z}^{(k)} + \mathbf{d}^{(k)}\|_2^2),$$

$$\mathbf{z}^{(k+1)} = \arg\min_{\mathbf{z}}(\lambda\|\mathbf{z}\|_1 + (\rho/2)\|\tilde{\boldsymbol{\omega}}_b^{(k+1)} - \mathbf{z} + \mathbf{d}^{(k)}\|_2^2),$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} + \mathbf{r}^{(k+1)}, \tag{17}$$

where $\mathbf{r}^{(k)} = \tilde{\boldsymbol{\omega}}_b^{(k)} - \mathbf{z}^{(k)}$ is the residual at the k'th iteration and $\mathbf{d}^{(k)} = \mathbf{d}^{(0)} + \sum_{j=1}^{k}\mathbf{r}^{(j)}$ is the summation of the residuals. In the first line of Eq. (17), the objective is to minimize a strictly convex quadratic function, and the solution can be easily obtained as

$$\tilde{\boldsymbol{\omega}}_b^{(k+1)} = (\mathbf{R}_b + \rho\mathbf{I})^{-1}[\mathbf{r}_{bd} + \rho(\mathbf{z}^{(k)} - \mathbf{d}^{(k)})]. \tag{18}$$

As mentioned, $\mathbf{R}_b$ and $\mathbf{r}_{bd}$ are unknown in practice, and they can be estimated as $\mathbf{R}_b = \sum_{l=1}^{L}\mathbf{b}(l)\mathbf{b}^{\mathrm{H}}(l)/L$ and $\mathbf{r}_{bd} = \sum_{l=1}^{L}\mathbf{b}(l)d_0^*(l)/L$, where L denotes the number of snapshots that are used. Moreover, $\mathbf{b}(l) = \mathbf{B}^{\mathrm{H}}\mathbf{x}(l)$ and $d_0(l) = \mathbf{v}_t^{\mathrm{H}}\mathbf{x}(l)$, where $\mathbf{x}(l)$ denotes the l'th space-time snapshot.[13–15]

The solution of Eq. (18) can be obtained directly, i.e., noniteratively. However, it is impractical because the inversion of $(\mathbf{R}_b + \rho\mathbf{I})$ has a high computational complexity of $\mathcal{O}[(NK - 1)^3]$. Note that, according to Fig. 3, the clutter covariance matrix constructed by the training snapshots with regard to the current detecting snapshot can be written as

$$\mathbf{R}_b = \widehat{\mathbf{R}}_b + \sum_{m=1}^{4}\frac{(-1)^m\mathbf{b}_m\mathbf{b}_m^{\mathrm{H}}}{L}, \tag{19}$$
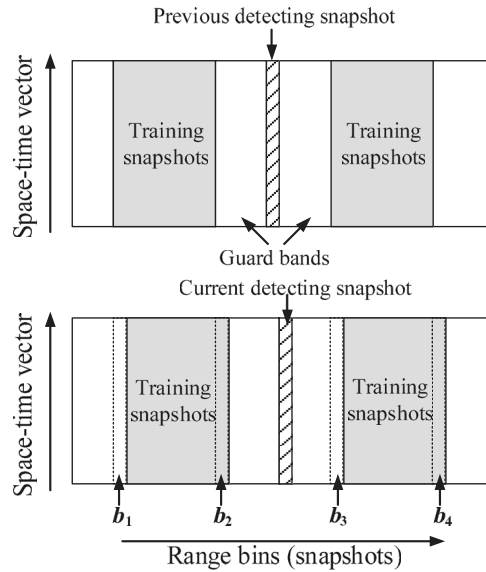
Fig. 3 Selection of I.I.D. training snapshots. The guard bands are used to guarantee that the training snapshots contain no components of the moving target.

where $\widehat{\mathbf{R}}_b$ is constructed by the training snapshots with regard to the previous detecting snapshot. Denote $\mathbf{P}^{(0)} = (\widehat{\mathbf{R}}_b + \rho\mathbf{I})^{-1}$; then, according to the matrix inversion lemma,[24] we obtain

$$\mathbf{P}^{(m)} = \mathbf{P}^{(m-1)} - \frac{\mathbf{P}^{(m-1)}\mathbf{b}_m\mathbf{b}_m^{\mathrm{H}}\mathbf{P}^{(m-1)}}{\frac{L}{(-1)^m} + \mathbf{b}_m^{\mathrm{H}}\mathbf{P}^{(m-1)}\mathbf{b}_m}, \qquad m = 1,2,3,4. \tag{20}$$

It is clear that $\mathbf{P}^{(4)} = (\mathbf{R}_b + \rho\mathbf{I})^{-1}$. Hence, the computational complexity can be reduced to $\mathcal{O}[8(NK-1)^2]$. A full analysis of the computational complexity is presented in Table 1.

In the second line of Eq. (17), the $\mathbf{z}$-update can be represented as

$$z_i^{(k+1)} = \arg\min_{z_i}[\lambda|z_i| + (\rho/2)(z_i - \tilde{w}_i^{(k+1)} - d_i^{(k)})^2]. \tag{21}$$

Although the absolute value function is not differentiable, a simple closed-form solution can easily be obtained. Explicitly, the solution is

$$z_i^{(k+1)} = S_{\lambda/\rho}(\tilde{w}_i^{(k+1)} + d_i^{(k)}), \tag{22}$$

where $S_\lambda(z)$ is the soft-thresholding operator. The soft-thresholding operator is essentially a shrinkage operator, which moves a point toward zero.

In the ADMM-STAP algorithm, $\tilde{\boldsymbol{\omega}}_b$ and $\mathbf{z}$ are updated alternately, which accounts for the term alternating direction. The reasonable stopping criteria are that the primal and dual residuals must be small,

Table 1 Computational complexity.

Algorithm	Complex multiplications	Complex additions
SMI-STAP	$\mathcal{O}[(NK-1)^3]$	$\mathcal{O}[(NK-1)^3]$
RLS-STAP	$[4(NK)^2 - 2\,NK - 1]L$	$[3(NK)^2 - 3\,NK]L$
OCD-STAP	$[4(NK)^2 - 5\,NK + 2]L$	$[3(NK)^2 - 4\,NK + 1]L$
ADMM-STAP	$(M+8)(NK-1)^2 + 4\,NK - 4$	$(M+8)(NK-1)^2 + 4M(NK-1) - 4$

Initialization:
Select $\rho > 0, \lambda > 0$.
$\begin{bmatrix} \mathbf{U} & \mathbf{S} & \mathbf{V} \end{bmatrix} = \mathrm{svd}\left(\mathbf{v}_t^{\mathrm{H}}\right), \mathbf{B} = \mathbf{V}(:, 2 : NK)$.
Set $\mathbf{z}^0 = \mathbf{0}$, and $\mathbf{d}^0 = \mathbf{0}$.

For m=1,2,3,4

$$\mathbf{P}^{(m)} = \mathbf{P}^{(m-1)} - \frac{\mathbf{P}^{(m-1)}\mathbf{b}_m \mathbf{b}_m^{\mathrm{H}}\mathbf{P}^{(m-1)}}{\dfrac{L}{(-1)^m} + \mathbf{b}_m^{\mathrm{H}}\mathbf{P}^{(m-1)}\mathbf{b}_m}$$

End for

For k=1,2,...

   1. $\tilde{\boldsymbol{\omega}}_b^{(k+1)} = \mathbf{P}^{(4)}\left[\mathbf{r}_{bd} + \rho\left(\mathbf{z}^{(k)} - \mathbf{d}^{(k)}\right)\right]$

   2. $\mathbf{z}^{(k+1)} = S_{\lambda/\rho}\left(\tilde{\boldsymbol{\omega}}_b^{(k+1)} + \mathbf{d}^{(k)}\right)$

   3. $\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} + \tilde{\boldsymbol{\omega}}_b^{(k+1)} - \mathbf{z}^{(k+1)}$

   4. Break if stopping criteria are satisfied

End for

Output: $\tilde{\boldsymbol{\omega}}_b^{(k+1)}$
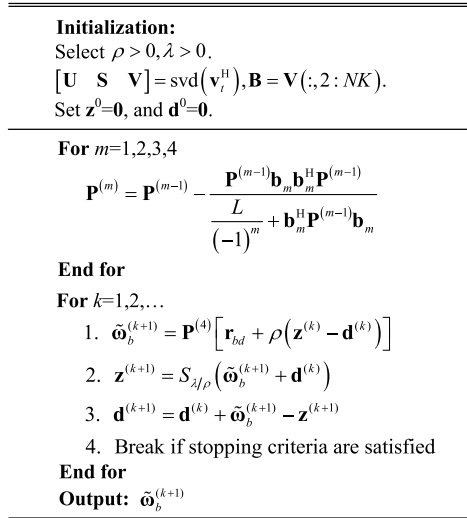
Fig. 4 The detailed iterative procedure of ADMM-STAP.

$$\left\|\tilde{\boldsymbol{\omega}}_b^{(k)} - \mathbf{z}^{(k)}\right\|_2 \le \varepsilon_{\mathrm{pri}} \quad \text{and} \quad \left\|\rho(\mathbf{z}^{(k)} - \mathbf{z}^{(k-1)})\right\|_2 \le \varepsilon_{\mathrm{dual}}, \tag{23}$$

where $\varepsilon_{\mathrm{pri}}$ and $\varepsilon_{\mathrm{dual}}$ are thresholds that are chosen by absolute and relative criteria

$$\varepsilon_{\mathrm{pri}} = \sqrt{p}\varepsilon_{\mathrm{abs}} + \varepsilon_{\mathrm{rel}}\max\{\|\tilde{\boldsymbol{\omega}}_b^{(k)}\|_2, \|\mathbf{z}^{(k)}\|_2\},$$
$$\varepsilon_{\mathrm{dual}} = \sqrt{n}\varepsilon_{\mathrm{abs}} + \varepsilon_{\mathrm{rel}}\|\mathbf{y}^{(k)}\|_2. \tag{24}$$

A reasonable value for $\varepsilon_{\mathrm{rel}}$ is $10^{-4} - 10^{-3}$, and the choice of $\varepsilon_{\mathrm{abs}}$ depends on the scale of the typical variable values. The detailed iterative procedure of ADMM-STAP is shown in Fig. 4.

3.3 Analysis of Convergence

A proof of the convergence result is presented in this section. First, we begin our proof by presenting the following theorem.

   Theorem 1 (Eckstein–Bertsekas):[25] Consider the problem

$$\min_{\mathbf{u}} f_1(\mathbf{u}) + f_2(\mathbf{v}) \quad \text{s.t. } \mathbf{v} = \mathbf{Gu}, \tag{25}$$

in the case where the functions $f_1(\cdot)$ and $f_2(\cdot)$ are closed, proper, and convex and $\mathbf{G}$ has a full column rank. Let $\{\eta_k \ge 0, k = 0, 1, \cdots\}$ and $\{\gamma_k \ge 0, k = 0, 1, \cdots\}$ be two sequences such that

$$\sum_{k=0}^{\infty} \eta_k < 0 \quad \text{and} \quad \sum_{k=0}^{\infty} \gamma_k < 0. \tag{26}$$

Assume that there are three sequences $\{\mathbf{u}_k, k = 0, 1, \cdots\}$, $\{\mathbf{v}_k, k = 0, 1, \cdots\}$, and $\{\mathbf{t}_k, k = 0, 1, \cdots\}$ that satisfy

$$\eta_k \ge \left\|\mathbf{u}_{k+1} - \arg\min_{\mathbf{u}}\{f_1(\mathbf{u}) + (\rho/2)\|\mathbf{Gu} - \mathbf{v}_k - \mathbf{t}_k\|_2^2\}\right\|$$

$$\gamma_k \ge \left\|\mathbf{v}_{k+1} - \arg\min_{\mathbf{v}}\{f_2(\mathbf{v}) + (\rho/2)\|\mathbf{Gu}_{k+1} - \mathbf{v} - \mathbf{t}_k\|_2^2\}\right\|$$

$$\mathbf{t}_{k+1} = \mathbf{t}_k - (\mathbf{Gu}_{k+1} - \mathbf{v}_{k+1}). \tag{27}$$

Then, if Eq. (25) has an optimal solution $\mathbf{u}^\dagger$, the sequence $\{\mathbf{u}_k\}$ converges to this solution, i.e., $\mathbf{u}_k \to \mathbf{u}^\dagger$.

First, since Eq. (14) is a particular instance when $\mathbf{G} = \mathbf{I}$, the full-rank condition in Theorem 1 can be satisfied. Second, it is clear that $f_1(\tilde{\boldsymbol{\omega}}_b) = -\mathbf{r}_{bd}^H\tilde{\boldsymbol{\omega}}_b - \tilde{\boldsymbol{\omega}}_b^H\mathbf{r}_{bd} + \tilde{\boldsymbol{\omega}}_b^H\mathbf{R}_b\tilde{\boldsymbol{\omega}}_b$ and $f_2(\mathbf{z}) = 2\lambda\|\mathbf{z}\|_1$ in Eq. (14) are closed, proper, and convex. Moreover, the sequences $\{\tilde{\boldsymbol{\omega}}_b^{(k)}\}$, $\{\mathbf{z}^{(k)}\}$, and $\{\mathbf{u}^{(k)}\}$ generated by Eq. (17) satisfy the conditions of Eq. (27) in a strict sense ($\eta_k = \gamma_k = 0$). Hence, the convergence is guaranteed.

3.4 Analysis of Computational Complexity

A comparison of the computational complexities of four STAP algorithms, namely, the conventional sample matrix inversion (SMI) STAP,[2] ℓ_1-regularized RLS-STAP,[14] ℓ_1-regularized online coordinate descent (OCD) STAP,[26] and the proposed ADMM-STAP algorithms, is presented in Table 1. The computational complexity is measured by the number of complex multiplications and additions. As shown in Table 1, the ADMM-STAP algorithm has a computational complexity of $\mathcal{O}[(M + 8)(NK - 1)^2]$, where M is the number of iterations. According to the simulation in Sec. 4, the algorithm can converge to an acceptable solution within a few tens of iterations, i.e., $M + 8$ would be less than $4L$ and $NK - 1$. Hence, the ADMM-STAP algorithm has the lowest level of computational complexity.

4 Simulation Results

The simulation parameters for the ground moving target indication application are listed in Table 2: a radar system equipped with a side-looking ULA is employed, and the elements are spaced half a wavelength apart, i.e., $d = \lambda/2$. Additive noise is modeled as spatially and temporally independent complex Gaussian noise with zero mean and unit variance. $f_r = 4v_p/\lambda$; hence, $\beta = 2v_pT_r/d = 1$. All the results are obtained from the average of 100 independent Monte–Carlo simulations.

4.1 Setting of Regularization Parameter

The regularization parameter provides a tradeoff between the SCNR steady-state performance and the convergence speed. Although it is clear that the value of λ should be proportional to the noise power and be inversely proportional to the rank of the clutter covariance matrix, it is still difficult to determine the optimal value. Adjusting the regularization parameter adaptively is an interesting research area (e.g., Refs. 13 and 14). However, this area is not the main focus of our paper. In this paper, the regularization parameter is selected from a fixed set $\Omega = \{0.1, 1, 10, 50\}$.

Table 2 Simulation parameters for airborne radar.

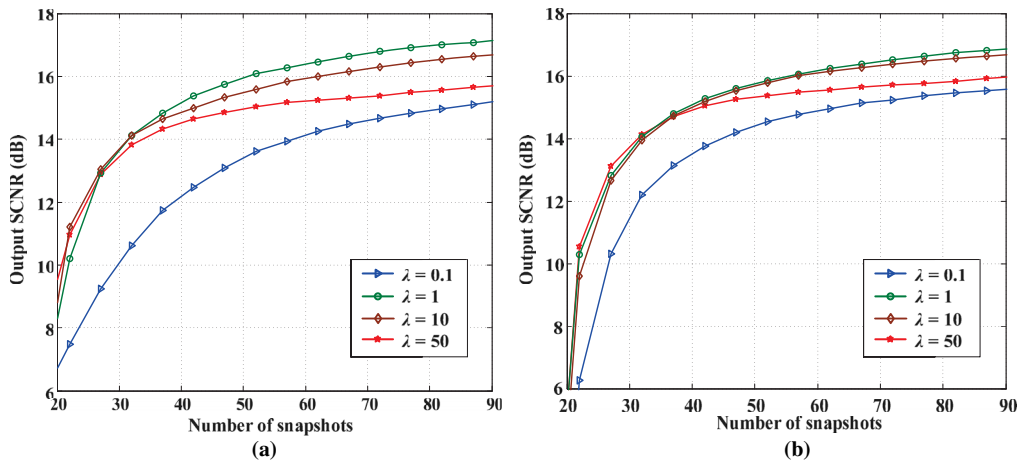| Parameter	Notation	Value	Unit
Antenna array spacing	d	$\lambda/2$	m
Pulse repetition frequency	f_r	2314	Hz
Carrier frequency	f_c	1.24	GHz
Array element number	N	10	—
CPI pulse number	K	10	—
Bandwidth	—	10	MHz
Platform velocity	v_p	140	m/s
Platform height	h_p	8000	m
Signal-to-noise ratio	SNR	0	dB

Fig. 5 Output SCNR versus the number of used snapshots with different regularization parameters. (a) CNR = 20 dB and (b) CNR = 40 dB.

The output SCNR versus the number of snapshots that are used with different values of the regularization parameter λ is shown in Fig. 5. In this simulation, we assume that the signal of the moving target impinges the array from a DOA of 90 deg and that the radial velocity of the moving target v_t is 28 m/s (the Doppler frequency of the moving target is nearly 231 Hz). The results in Fig. 5 indicate that (i) the value of λ is crucial to the output SCNR performance, and there is a reasonable range of values, i.e., $1 \leq \lambda \leq 10$, that can improve the convergence speed and the output SCNR steady-state performance simultaneously; (ii) the output SCNR is degraded when λ is too large since the filter weight vector is shrunk to zero; and (iii) the output SCNR performance is not considerably improved when λ is too small. In this case, the output SCNR performance is nearly similar to that of the conventional STAP algorithm.

The output SCNR performance versus the Doppler frequency of the moving target at a DOA of 90 deg is shown in Fig. 6. The range of potential Doppler frequency is from $-500-$ to 500 Hz, and 60 snapshots are used to optimize the filter vector. The same conclusion can be obtained. This figure shows that the ADMM-STAP algorithm with $1 \leq \lambda \leq 10$ provides a satisfactory output SCNR performance.
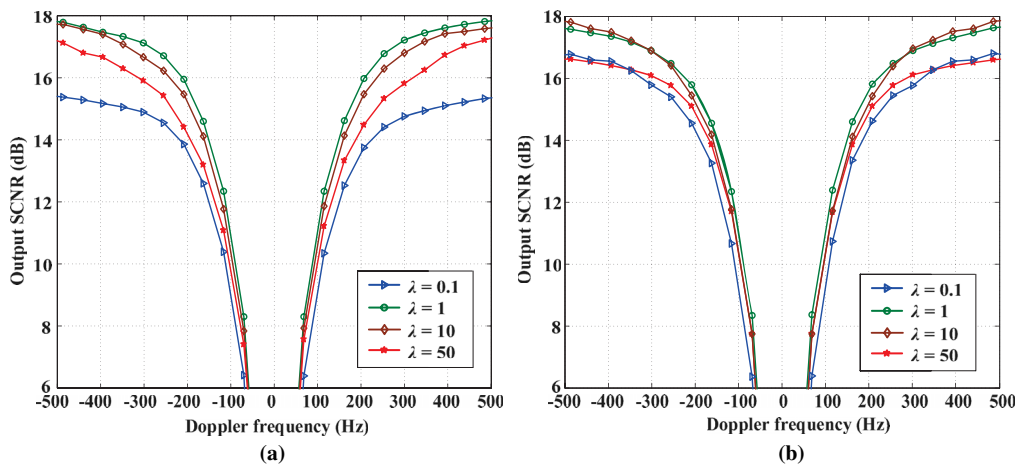

Fig. 6 Output SCNR performance versus Doppler frequency with different regularization parameters, and the range of Doppler frequency is from -500 to 500 Hz. (a) CNR = 20 dB and (b) CNR = 40 dB.
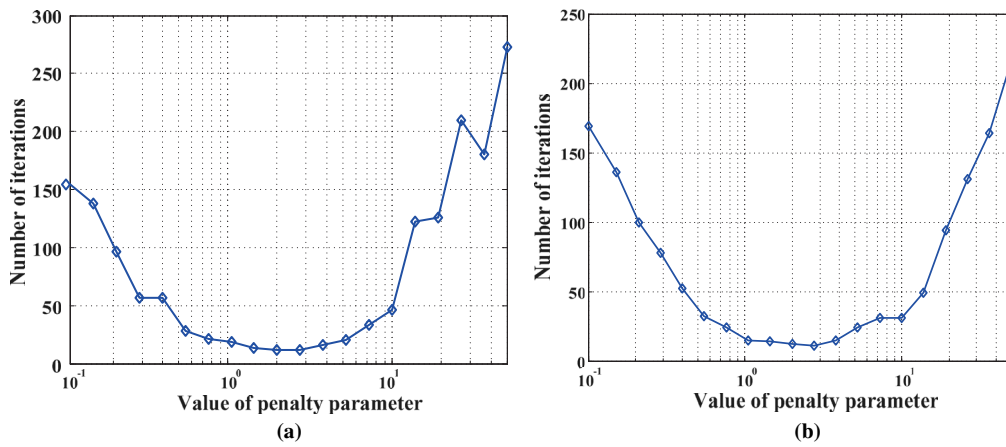
Fig. 7 Number of iterations versus the value of λ. The radial velocity of the moving target is 28 m/s, and 60 snapshots are used in the simulation. (a) CNR = 20 dB and (b) CNR = 40 dB.

The number of iterations with different values of λ is shown in Fig. 7. As shown, if we choose λ from an appropriate range ($0.5 \leq \lambda \leq 10$), then the ADMM-STAP algorithm can converge rapidly within a few tens of iterations, which is acceptable in practice. Otherwise, the number of iterations increases significantly, and the iteration output cannot converge to the optimal solution leading, to a performance degradation to a certain extent.

4.2 *Comparison with Other Algorithms*

In this section, we will compare the output SCNR performance of our proposed algorithm with that of IPM-STAP, OCD-STAP, and RLS-STAP algorithms. The regularization parameter λ is set to 1 for all the algorithms, and the other parameters are the same as in the previous simulations. The output SCNR performances versus the number of used snapshots and the target Doppler frequency are compared in Figs. 8 and 9. As shown in these figures, we can see that (i) the output SCNR performance of the IPM-STAP algorithm is superior to that of the RLS-STAP and OCD-STAP algorithms. However, it is achieved at a high computational cost and (ii) the output SCNR performance of the ADMM-STAP algorithm can outperform that of the IPM-STAP algorithm, which supports our previous conclusion that optimizing the problem of parameter estimation to a high accuracy generally yields no improvement.
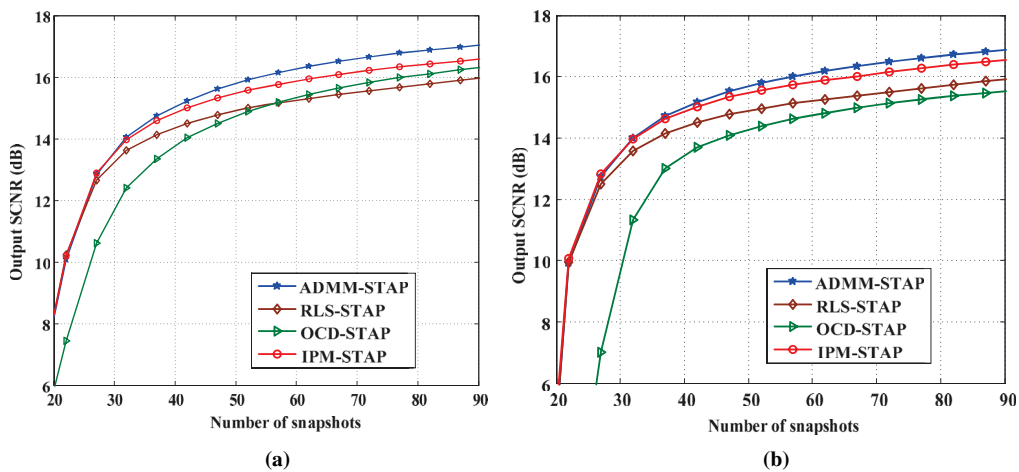

Fig. 8 Output SCNR versus the number of used snapshots when the radial velocity of the moving target is 28 m/s. (a) CNR = 20 dB and (b) CNR = 40 dB.
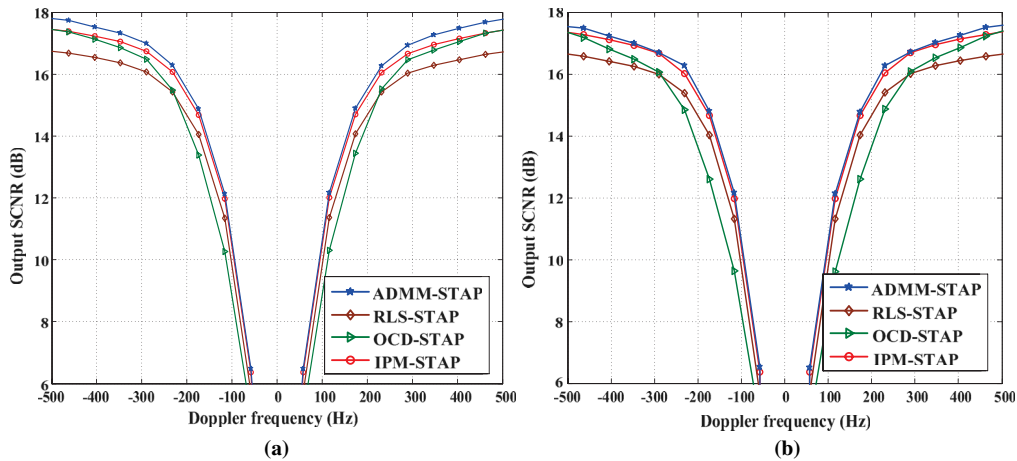
Fig. 9 Output SCNR performance versus Doppler frequency with 60 snapshots, and the range of Doppler frequency is from −500 to 500 Hz. (a) CNR = 20 dB and (b) CNR = 40 dB.

5 ℓ_1-Regularized STAP with Mountaintop Data

The performance of the ℓ_1-regularized STAP approaches is verified here using the Mountaintop data set (data No. t38pre01v1) acquired with the experimental radar system RSTER (radar surveillance technology experimental radar) sponsored by the Advanced Research Projects Agency. The Mountaintop program is devoted to supporting the mission requirements of next-generation airborne early warning platforms and to supporting the evaluation of STAP algorithms. The antenna for the system is a 5-m wide by 10-m high horizontally polarized array composed of 14 column elements. The CPI pulse number is 16, the antenna array spacing is 0.333 m, the PRF is 625 Hz, the carrier frequency is 435 MHz, and the bandwidth is 500 kHz. The transmit beam is steered to illuminate a mountain range (a large clutter scatter).

The data set is divided into two subsets in our experiment. The first subset, including 100 snapshots, is used to train the STAP filters. The second subset, including 100 snapshots, is used to test the performance. Two simulated moving targets are added to the test data subset. The signal of the first target impinges the array from a DOA of −25 deg, and the Doppler frequency is 62.5 Hz. The signal of the second target impinges the array from a DOA of 20 deg, and the Doppler frequency is 187.5 Hz. Hence, the first target can essentially be regarded as a ground moving vehicle in the mountain, and the second target can be regarded as an aircraft near the mountain. The minimum variance distortionless response (MVDR) spectra of the two subsets are shown in Fig. 10.

The improvement factor (IF) performance, which is defined as the ratio of the output SCNR to the input SCNR, is investigated in Fig. 11. The regularization parameter λ is set to 1 for all the
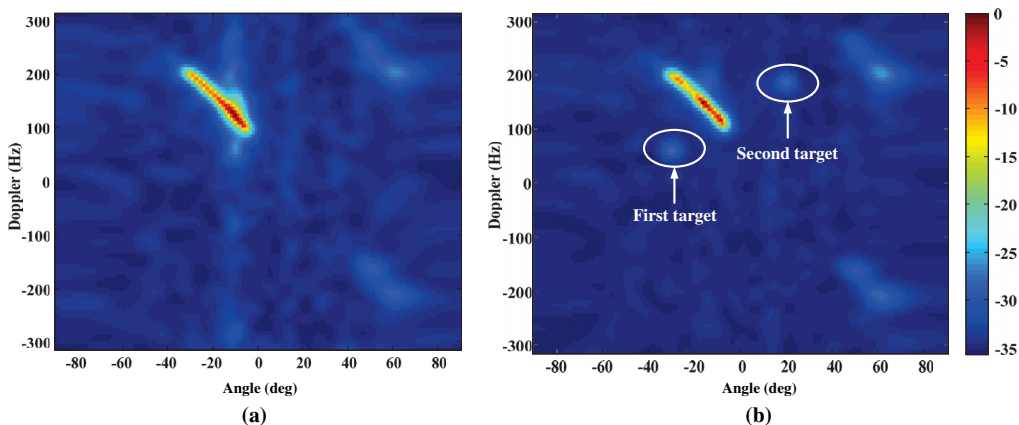

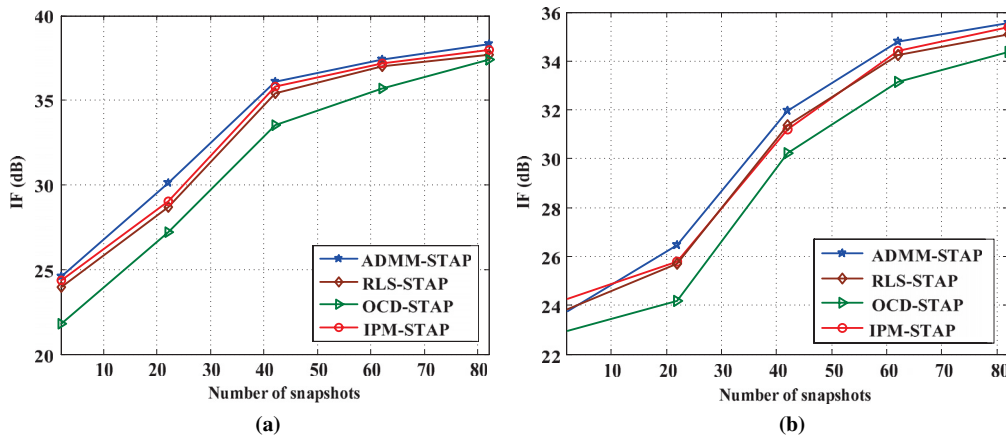**Fig. 10** (a) MVDR spectrum of the training subset and (b) MVDR spectrum of the test subset.

Fig. 11 (a) IF performance versus the number of used snapshots for the first target and (b) IF performance versus the number of used snapshots for the second target.

algorithms. As shown, the IF performance of the proposed ADMM-STAP approach substantially outperforms that of the other approaches. Hence, the effectiveness of the proposed approach is confirmed by an experimental multichannel radar system RSTER.

6 Conclusions

In this paper, we proposed a sparsity-based approach based on an ℓ_1-regularized constraint to accelerate the convergence speed of STAP. The optimization problem with an additional ℓ_1-regularized constraint was solved using the ADMM, and the detailed iterative procedure of ADMM-SATP was derived. Through the examples, it was demonstrated that the proposed method can effectively decrease the required number of secondary snapshots and provide better performance than the ℓ_1-regularized OCD-STAP and ℓ_1-regularized RLS-STAP methods.

Acknowledgments

References

1. H. Wang et al., "Robust waveform design for MIMO-STAP to improve the worst-case detection performance," *EURASIP J. Adv. Signal Process.* **2013**(1), 1–8 (2013).
2. W. L. Melvin, "A STAP overview," *IEEE Aerosp. Electron. Syst. Mag.* **19**(1), 19–35 (2004).
3. X. Guo et al., "Modified reconstruction algorithm based on space-time adaptive processing for multichannel synthetic aperture radar systems in azimuth," *J. Appl. Remote Sens.* **10**(3), 035022 (2016).
4. R. Fa and R. C. De Lamare, "Reduced-rank STAP algorithms using joint iterative optimization of filters," *IEEE Trans. Aerosp. Electron. Syst.* **47**(3), 1668–1684 (2011).
5. R. Fa, R. C. de Lamare, and L. Wang, "Reduced-rank STAP schemes for airborne radar based on switched joint interpolation, decimation and filtering algorithm," *IEEE Trans. Signal Process.* **58**(8), 4182–4194 (2010).
6. R. Li et al., "Reduced-dimension space-time adaptive processing based on angle-Doppler correlation coefficient," *EURASIP J. Adv. Signal Process.* **2016**(97), 1–9 (2016).
7. W. Zhang et al., "Multiple-input multiple-output radar multistage multiple-beam beamspace reduced-dimension space-time adaptive processing," *IET Radar Sonar Navig.* **7**(3), 295–303 (2013).
8. W. Zhang et al., "A method for finding best channels in beam-space post-Doppler reduced-dimension STAP," *IEEE Trans. Aerosp. Electron. Syst.* **50**(1), 254–264 (2014).
9. K. Sun, H. Meng, and F. Lapierre, "Registration-based compensation using sparse representation in conformal-array STAP," *Signal Process.* **91**(10), 2268–2276 (2011).

10. K. Sun, H. Meng, and Y. Wang, "Direct data domain STAP using sparse representation of clutter spectrum," *Signal Process.* **91**(9), 2222–2236 (2011).
11. S. Sen, "OFDM radar space-time adaptive processing by exploiting spatio-temporal sparsity," *IEEE Trans. Signal Process.* **61**(1), 118–130 (2013).
12. Z. Yang, X. Li, and H. Wang, "Space-time adaptive processing based on weighted regularized sparse recovery," *Prog. Electromagnet. Res. B* **42**, 245–262 (2012).
13. Z. Yang, R. C. De Lamare, and X. Li, "L1-regularized STAP algorithm with a generalized side-lobe canceler architecture for airborne radar," *IEEE Trans. Signal Process.* **60**(2), 674–686 (2012).
14. Z. Gao et al., "L1-regularised joint iterative optimisation space-time adaptive processing algorithm," *IET Radar Sonar Navig.* **10**(3), 435–441 (2016).
15. Z. Yang et al., "Sparsity-based space-time adaptive processing using complex-valued homotopy technique for airborne radar," *IET Signal Process.* **8**(5), 552–564 (2014).
16. M. Shen et al., "An efficient moving target detection algorithm based on sparsity-aware spectrum estimation," *Sensors* **14**(9), 17055–17067 (2014).
17. J. Qin, I. Yanovsky, and W. Yin, "Efficient simultaneous image deconvolution and upsampling algorithm for low-resolution microwave sounder data," *J. Appl. Remote Sens.* **9**(1), 095035 (2015).
18. H. Zhai et al., "Reweighted mass center based object-oriented sparse subspace clustering for hyperspectral images," *J. Appl. Remote Sens.* **10**(4), 046014 (2016).
19. S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**(1), 1–122 (2010).
20. M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.* **19**(9), 2345–2356 (2010).
21. Z. Yang et al., "On clutter sparsity analysis in space-time adaptive processing airborne radar," *IEEE Geosci. Remote Sens. Lett.* **10**(5), 1214–1218 (2013).
22. G. M. Herbert, "Clutter modeling for space-time adaptive processing in airborne radar," *IET Radar Sonar Navig.* **4**(2), 178–186 (2010).
23. H. Sun et al., "Estimation of the ocean clutter rank for HF/VHF radar space-time adaptive processing," *IET Radar Sonar Navig.* **4**(6), 755–763 (2010).
24. B. Chen et al., "Quantized kernel recursive least squares algorithm," *IEEE Trans. Neural Netw. Learn. Syst.* **24**(9), 1484–1491 (2013).
25. J. Eckstein and D. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.* **55**(1), 293–318 (1992).
26. D. Angelosante, J. A. Bazerque, and G. B. Giannakis, "Online adaptive estimation of sparse signals: where RLS meets the L1-norm," *IEEE Trans. Signal Process.* **58**(7), 3436–3447 (2010).

Lilong Qin is working toward his PhD at the National University of Defense Technology, Changsha, China, and is working with Aalto University, Espoo, Finland. He received his BS degree in information engineering and his MS degree in circuit and system from the Electronic Engineering Institute, Hefei, China, in 2010 and 2013, respectively. His current research interests include synthetic aperture radar and adaptive beamforming.

Manqing Wu received his MS degree from the National University of Defense Technology, Changsha, China, in 1990. Currently, he is a professor with the China Electronics Technology Group Corporation, Beijing, China, and is a member of the Chinese Academy of Engineering. His research field is radar signal processing.

Xuan Wang received her PhD in signal and information processing from Beijing Institute of Technology, Beijing, China, in 2016. Currently, she is working at Delft University of Technology, Delft, the Netherlands. Her current research interest is synthetic aperture radar.

Zhen Dong received his PhD from the National University of Defense Technology, Changsha, China, in 2001. Currently, he is a professor with the National University of Defense Technology, and his research field includes synthetic aperture radar interferometry and array radar.