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

**Fast 1-regularized space-Time adaptive processing using alternating direction method of multipliers**

*Published in: JOURNAL OF APPLIED REMOTE SENSING*

**DOI:**
10.1117/1.JRS.11.026004

Published: 01/04/2017

*Document Version*
Publisher's PDF, also known as Version of record

*Please cite the original version:*
Fast $\ell_1$-regularized space-time adaptive processing using alternating direction method of multipliers

Lilong Qin
Manqing Wu
Xuan Wang
Zhen Dong

Fast $\ell_1$-regularized space-time adaptive processing using alternating direction method of multipliers

Lilong Qin,a,b,* Manqing Wu,c Xuan Wang,d and Zhen Donga

aNational University of Defense Technology, School of Electronic Science and Engineering, Changsha, China
bAalto University, Department of Signal Processing and Acoustics, Espoo, Finland
cChina Electronics Technology Group Corporation, China Academy of Electronics and Information Technology, Beijing, China
dDelft 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 $\ell_1$-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]

Keywords: space-time adaptive processing; alternating direction method of multipliers; generalized side-lobe canceler; sparse representation; recursive least-squares.

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 $\sim$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 $\ell_1$-regularized constraint, which is the
so-called least absolute shrinkage and selection operator (LASSO), has been established.\textsuperscript{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 $\ell_1$-regularized recursive least-squares STAP (RLS-STAP) algorithm,\textsuperscript{13} the $\ell_1$-regularized least-mean-square STAP algorithm,\textsuperscript{14} and the homotopy-STAP algorithm\textsuperscript{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.\textsuperscript{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.\textsuperscript{17,18} This technique is well suited for solving the optimization problems of the $\ell_1$ constraint, particularly large-scale problems.\textsuperscript{19} The ADMM technique can converge within a few tens of iterations, which is acceptable in practical use.\textsuperscript{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 $\ell_1$-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 $\ell_1$-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 $\left(\cdot\right)^T$, $\left(\cdot\right)^*$, and $\left(\cdot\right)^H$, respectively. The symbol $\otimes$ denotes the Kronecker product, and the symbol $\left\| \cdot \right\|_n$ denotes the $\ell_n$-norm operator. $E\left(\mathbf{x}\right)$ denotes the expected value of $\mathbf{x}$, $|x|$ indicates the absolute value of $x$, and $(x)_+ \triangleq \max(0, x)$. sign($\cdot$) is the component-wise sign function.\textsuperscript{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 $K$
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 $x = x_t + x_c + n$, where $x_t$ is the target vector, $x_c$ is the clutter vector, and $n$ is the thermal noise vector with noise power $\sigma_n^2$ on each channel and pulse. The space-time clutter vector can be represented as

$$x_c = \sum_{n=1}^{N_c} \sigma_{c,n} v(f_{d,n}, f_{s,n}),$$

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 $\lambda$ is the wavelength and $d$ is the inner-sensor spacing of the ULA. $v(f_{d,n}, f_{s,n}) \in \mathbb{C}^{NK \times 1}$ is the space-time steering vector, which is defined as a Kronecker product of the temporal and spatial steering vectors, i.e., $v(f_{d,n}, f_{s,n}) = v_d(f_d) \otimes v_s(f_s)$, where

$$v_d(f_d) = [1 \exp(j2\pi f_d) \cdots \exp(j2\pi f_d(K-1))]^T,$$

$$v_s(f_s) = [1 \exp(j2\pi f_s) \cdots \exp(j2\pi f_s(N-1))]^T.$$

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

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

$$[U \quad S \quad V] = \text{svd}(v_t^H), \quad B = V(:, 2 : NK).$$

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

$$y = d_0 - \omega_b^H b,$$

where $d_0 = v_t^H x$ and $\omega_b = R_b^{-1} r_{bd}$. $R_b = E(bb^H)$ is the clutter covariance matrix, and $r_{bd} = E(bd_b^H)$ is the cross-correlation vector between $d_0$ and $b$. The output clutter power can be computed as

$$P = v_t^H R_x v_t - r_{bd}^H R_b^{-1} r_{bd},$$

where $R_x = E(xx^H)$ is the input covariance matrix. The output SCNR can be expressed as

$$\xi = \frac{NM[d_t]^2}{v_t^H R_x v_t - r_{bd}^H R_b^{-1} r_{bd}}.$$

**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, \( \mathbf{R}_s \) and \( \mathbf{r}_{bd} \) are unknown in practice, and the secondary training snapshots are required to estimate these parameters. 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. 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. Hence, in the GSC form of the sparsity-based STAP algorithm (see Fig. 2), the filter coefficient vector \( \omega_b \) can be replaced by \( \tilde{\omega}_b = \mathbf{V}\omega_b \), where \( \mathbf{V} = \text{diag}(\mathbf{v}) \) and \( \mathbf{v} \in \mathbb{C}^{NK-1} \) denote a sparse vector. Then, we obtain

\[
\mathbf{z} = \mathbf{V}^{\dagger}\mathbf{b} \in \mathbb{C}^{(NK-1) \times 1}.
\]  

The output of the sparsity-based STAP is

\[
y_r = d_0 - \omega_b^{\dagger}\mathbf{z} = \left[ 1 \quad \omega_b^{\dagger} - \omega_b^{\dagger}\mathbf{V}^{\dagger}\right] \begin{bmatrix} \mathbf{y} \\ \mathbf{b} \end{bmatrix}.
\]  

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

\[
\mathcal{P}_c = \mathbf{v}_t^{\dagger}\mathbf{R}_c \mathbf{v}_t - \mathbf{r}_{bd}^{\dagger}\mathbf{R}_b^{-1}\mathbf{r}_{bd} + \epsilon^{\dagger}\mathbf{R}_b \epsilon,
\]

where \( \epsilon = \omega_b - \tilde{\omega}_b \) 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

\[
\xi_r = \frac{NM|\alpha|^2}{\mathbf{v}_t^{\dagger}\mathbf{R}_c \mathbf{v}_t - \mathbf{r}_{bd}^{\dagger}\mathbf{R}_b^{-1}\mathbf{r}_{bd} + \epsilon^{\dagger}\mathbf{R}_b \epsilon}.
\]

Hence, the aim is to minimize the mean-square error \( \epsilon^{\dagger}\mathbf{R}_b \epsilon \). The objective function of the minimization problem can be rewritten as

\[
\epsilon^{\dagger}\mathbf{R}_b \epsilon = \mathbf{r}_{bd}^{\dagger}\mathbf{R}_b^{-1}\mathbf{r}_{bd} - \tilde{\omega}_b^{\dagger}\mathbf{r}_{bd} - \tilde{\omega}_b^{\dagger}\mathbf{R}_b \tilde{\omega}_b + \tilde{\omega}_b^{\dagger}\mathbf{R}_b \tilde{\omega}_b.
\]

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

\[
\min_{\tilde{\omega}_b} - \mathbf{r}_{bd}^{\dagger}\tilde{\omega}_b - \omega_b^{\dagger}\mathbf{r}_{bd} + \tilde{\omega}_b^{\dagger}\mathbf{R}_b \tilde{\omega}_b + \lambda \|\tilde{\omega}_b\|_0,
\]

where \( \lambda \) is the regularization parameter for regulating the sparseness of \( \tilde{\omega}_b \). However, the \( \ell_0 \)-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

\[
\min_{\tilde{\omega}_b} - \mathbf{r}_{bd}^{\dagger}\tilde{\omega}_b - \omega_b^{\dagger}\mathbf{r}_{bd} + \tilde{\omega}_b^{\dagger}\mathbf{R}_b \tilde{\omega}_b + \lambda \|\tilde{\omega}_b\|_1.
\]

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 \( \ell_1 \)-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.\(^9\) The ADMM-STAP algorithm is based on the algorithm of variable splitting, i.e., we split the variable \( \hat{\theta}_b \) into a pair of variables, say, \( \hat{\theta}_b \) and \( z \), 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

\[
\min_{\hat{\theta}_b, z} \mathbf{r}_{bd}^H \hat{\theta}_b - \hat{\theta}_b^H \mathbf{r}_{bd} + \hat{\theta}_b^H \mathbf{R}_b \hat{\theta}_b + \lambda \| z \|_1 \quad \text{s.t.} \quad \hat{\theta}_b = z.
\] \hspace{1cm} (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 Lagrangian function is formed as\(^19,20\)

\[
\mathcal{L}_\rho(\hat{\theta}_b, z, y) = -\mathbf{r}_{bd}^H \hat{\theta}_b - \hat{\theta}_b^H \mathbf{r}_{bd} + \hat{\theta}_b^H \mathbf{R}_b \hat{\theta}_b + \lambda \| z \|_1 + (\rho/2)\| \hat{\theta}_b - z \|^2_2 + y^H (\hat{\theta}_b - z),
\] \hspace{1cm} (15)

where \( \rho > 0 \) is the augmented Lagrangian parameter and \( y \) is a vector of Lagrange multipliers.

3.2 \( \ell_1 \)-Regularized ADMM-STAP

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

\[
(\rho/2)\| \hat{\theta}_b - z \|^2_2 + y^H (\hat{\theta}_b - z) = (\rho/2)\| \mathbf{r} \|^2_2 + y^H \mathbf{r} = (\rho/2)\| \mathbf{r} + \mathbf{d} \|^2_2 - (\rho/2)\| \mathbf{d} \|^2_2.
\] \hspace{1cm} (16)

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

\[
\hat{\theta}_b^{(k+1)} = \arg \min_{\hat{\theta}_b} (-\mathbf{r}_{bd}^H \hat{\theta}_b - \hat{\theta}_b^H \mathbf{r}_{bd} + \hat{\theta}_b^H \mathbf{R}_b \hat{\theta}_b + (\rho/2)\| \hat{\theta}_b - z^{(k)} + \mathbf{d}^{(k)} \|^2_2),
\]

\[
z^{(k+1)} = \arg \min_{z} \lambda \| z \|_1 + (\rho/2)\| \hat{\theta}_b^{(k+1)} - z + \mathbf{d}^{(k)} \|^2_2,
\]

\[
d^{(k+1)} = \mathbf{d}^{(k)} + \mathbf{r}^{(k+1)},
\] \hspace{1cm} (17)

where \( \mathbf{r}^{(k)} = \hat{\theta}_b^{(k)} - z^{(k)} \) 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

\[
\hat{\theta}_b^{(k+1)} = (\mathbf{R}_b + \rho \mathbf{I})^{-1} [\mathbf{r}_{bd} + \rho (z^{(k)} - d^{(k)})].
\] \hspace{1cm} (18)

As mentioned, \( \mathbf{R}_b \) and \( \mathbf{r}_{bd} \) are unknown in practice, and they can be estimated as \( \mathbf{R}_b = \sum_{l=1}^{\ell} \mathbf{b}(l) \mathbf{b}(l)^H / L \) and \( \mathbf{r}_{bd} = \sum_{l=1}^{\ell} \mathbf{b}(l) d_0(l) / L \), where \( L \) denotes the number of snapshots that are used. Moreover, \( \mathbf{b}(l) = \mathbf{B}^H \mathbf{x}(l) \) and \( d_0(l) = \mathbf{v}_l^H \mathbf{x}(l) \), 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 \( O((NK - 1)^3) \). 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

\[
\mathbf{R}_b = \bar{\mathbf{R}}_b + \sum_{m=1}^{4} (-1)^m \mathbf{b}_m \mathbf{b}_m^H / L.
\] \hspace{1cm} (19)
where \( \bar{R}_b \) is constructed by the training snapshots with regard to the previous detecting snapshot. Denote \( P(0) = (\bar{R}_b + \rho I)^{-1} \); then, according to the matrix inversion lemma, we obtain

\[
P(m) = P(m-1) - \frac{P(m-1) b_m b_m^H P(m-1)}{b_m^H b_m} b_m^H P(m-1), \quad m = 1, 2, 3, 4.
\]

(20)

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

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

\[
z_i^{(k+1)} = \arg \min_{z_i} |z_i| + \frac{\rho}{2} (z_i - w_i^{(k+1)} - d_i^{(k)})^2.
\]

(21)

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

\[
z_i^{(k+1)} = S_{\lambda}(w_i^{(k+1)} + d_i^{(k)}),
\]

(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, \( \tilde{o}_b \) and \( z \) 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>
<thead>
<tr>
<th>Algorithm</th>
<th>Complex multiplications</th>
<th>Complex additions</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMI-STAP</td>
<td>( O((NK - 1)^2) )</td>
<td>( O((NK - 1)^3) )</td>
</tr>
<tr>
<td>RLS-STAP</td>
<td>( [4(NK)^2 - 2 NK - 1]L )</td>
<td>( [3(NK)^2 - 3 NK]L )</td>
</tr>
<tr>
<td>OCD-STAP</td>
<td>( [4(NK)^2 - 5 NK + 2]L )</td>
<td>( [3(NK)^2 - 4 NK + 1]L )</td>
</tr>
<tr>
<td>ADMM-STAP</td>
<td>( (M + 8)(NK - 1)^2 + 4 NK - 4 )</td>
<td>( (M + 8)(NK - 1)^2 + 4M(NK - 1) - 4 )</td>
</tr>
</tbody>
</table>

Table 1 Computational complexity.
\[ \omega_0(k) - z(k) \leq \varepsilon_{\text{pri}} \quad \text{and} \quad \rho(z(k) - z(k-1)) \leq \varepsilon_{\text{dual}}, \]  

where \( \varepsilon_{\text{pri}} \) and \( \varepsilon_{\text{dual}} \) are thresholds that are chosen by absolute and relative criteria.

\[ \varepsilon_{\text{pri}} = \sqrt{p\varepsilon_{\text{abs}}} + e_{\text{rel}} \max \left\{ \| \omega_0(k) \|_2, \| z(k) \|_2 \right\}, \]

\[ \varepsilon_{\text{dual}} = \sqrt{n\varepsilon_{\text{abs}}} + e_{\text{rel}}\| y(k) \|_2. \]  

A reasonable value for \( e_{\text{rel}} \) is \( 10^{-4} \) to \( 10^{-3} \), and the choice of \( \varepsilon_{\text{abs}} \) 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

\[ \min_u f_1(u) + f_2(v) \quad \text{s.t.} \quad v = Gu, \]  

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

\[ \sum_{k=0}^{\infty} \eta_k < 0 \quad \text{and} \quad \sum_{k=0}^{\infty} \gamma_k < 0. \]  

Assume that there are three sequences \( \{ u_k, k = 0, 1, \cdots \} \), \( \{ v_k, k = 0, 1, \cdots \} \), and \( \{ t_k, k = 0, 1, \cdots \} \) that satisfy

\[ \eta_k \geq \| u_{k+1} - \arg \min_u \{ f_1(u) + (\rho/2) \| Gu - v_k - t_k \|_2^2 \} \| \]

\[ \gamma_k \geq \| v_{k+1} - \arg \min_v \{ f_2(v) + (\rho/2) \| Gu_{k+1} - v - t_k \|_2^2 \} \| \]

\[ t_{k+1} = t_k - (Gu_{k+1} - v_{k+1}). \]  

Then, if Eq. (25) has an optimal solution \( u^* \), the sequence \( \{ u_k \} \) converges to this solution, i.e., \( u_k \to u^* \).
First, since Eq. (14) is a particular instance when $G = I$, the full-rank condition in Theorem 1 can be satisfied. Second, it is clear that $f_1(\tilde{\omega}_b) = -r_H^T\tilde{\omega}_b - \tilde{\omega}_b^Hr_H + \tilde{\omega}_b^H R_b \tilde{\omega}_b$ and $f_2(z) = 2\lambda||z||_1$ in Eq. (14) are closed, proper, and convex. Moreover, the sequences \{\tilde{\omega}^{(k)}_b\}, \{z^{(k)}\}, and \{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,\textsuperscript{2} \ell_1-regularized RLS-STAP,\textsuperscript{14} \ell_1-regularized online coordinate descent (OCD) STAP,\textsuperscript{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 $O((M + 8)(NK - 1)^2)$, 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, $\beta = 2v_p T_r / d = 1$. 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 $\lambda$ 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>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna array spacing</td>
<td>$d$</td>
<td>$\lambda/2$</td>
<td>m</td>
</tr>
<tr>
<td>Pulse repetition frequency</td>
<td>$f_r$</td>
<td>2314</td>
<td>Hz</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>$f_c$</td>
<td>1.24</td>
<td>GHz</td>
</tr>
<tr>
<td>Array element number</td>
<td>$N$</td>
<td>10</td>
<td>—</td>
</tr>
<tr>
<td>CPI pulse number</td>
<td>$K$</td>
<td>10</td>
<td>—</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>—</td>
<td>10</td>
<td>MHz</td>
</tr>
<tr>
<td>Platform velocity</td>
<td>$v_p$</td>
<td>140</td>
<td>m/s</td>
</tr>
<tr>
<td>Platform height</td>
<td>$h_p$</td>
<td>8000</td>
<td>m</td>
</tr>
<tr>
<td>Signal-to-noise ratio</td>
<td>SNR</td>
<td>0</td>
<td>dB</td>
</tr>
</tbody>
</table>

*Qin et al.: Fast $\ell_1$-regularized space-time…*

The output SCNR versus the number of snapshots that are used with different values of the regularization parameter $\lambda$ 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 $\lambda$ 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 $\lambda$ is too large since the filter weight vector is shrunk to zero; and (iii) the output SCNR performance is not considerably improved when $\lambda$ 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. 5 Output SCNR versus the number of used snapshots with different regularization parameters. (a) CNR = 20 dB and (b) CNR = 40 dB.

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.
The number of iterations with different values of $\lambda$ is shown in Fig. 7. As shown, if we choose $\lambda$ 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 $\lambda$ 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. 7 Number of iterations versus the value of $\lambda$. 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.

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.
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 $\lambda$ is set to 1 for all the targets.

**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 $\lambda$ is set to 1 for all the targets.
Qin et al.: Fast $\ell_1$-regularized space-time…

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 $\ell_1$-regularized constraint to accelerate the convergence speed of STAP. The optimization problem with an additional $\ell_1$-regularized constraint was solved using the ADMM, and the detailed iterative procedure of ADMM-STAP 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 $\ell_1$-regularized OCD-STAP and $\ell_1$-regularized RLS-STAP methods.

Acknowledgments

The authors thank the National Natural Science Foundation of China under Grant No. 61101178 and the China Scholarship Council for their support.

References


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.