remote sensing

MDPI

Article

A Novel Fast Sparse Bayesian Learning STAP Algorithm for Conformal Array Radar

Bing Ren and Tong Wang *

National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China; 18021110245@stu.xidian.edu.cn
* Correspondence: twang@mail.xidian.edu.cn

Abstract: Space-time adaptive processing (STAP) is an important method of clutter suppression that requires adequate training samples. For an airborne conformal array radar, conventional STAP methods do not have enough training samples to acquire good performance due to the range dependent clutter caused by geometry and the problem of polarization. Sparse-recovery-based STAP (SR-STAP) methods have garnered significant attention in the past few decades because they only require a small number of training samples. Sparse Bayesian Learning (SBL) methods have seen increasing amounts of development due to its robust, self-regularizing nature and because it is not sensitive to user parameters, but it converges slowly. In this paper, a novel fast SBL (NFSBL) method is put forward to increase the rate of convergence. To minimize the SBL penalty function, the proposed method introduces the conjugate function to construct a surrogate function. Additional solution sparsity will be achieved through iteratively minimizing the surrogate function. Then, from the proposed method, we could obtain a more accurate clutter plus noise covariance matrix. Numerical simulation results express that this method could acquire better performance of STAP and improvement in convergence and computational complexity for a conformal array.

Keywords: sparse Bayesian learning; space-time adaptive processing; airborne radar; conformal array

Citation: Ren, B.; Wang, T. A Novel Fast Sparse Bayesian Learning STAP Algorithm for Conformal Array Radar. *Remote Sens.* **2023**, *15*, 2824. https://doi.org/10.3390/rs15112824

Academic Editor: Lorenzo Capineri

Received: 14 April 2023
Revised: 24 May 2023
Accepted: 24 May 2023
Published: 29 May 2023

1. Introduction

Conformal array radar has attracted extensive attention because of the advantages which contain the underlying larger effective aperture, minimum payload weight, rise scanning range without cumbersome mechanical couplings and avoiding signal modulation from rotary antenna [1].

Space-time adaptive processing (STAP) is a meaningful method because of its excellent clutter suppression performance [2]. The optimum weight vector aims to maximize the output signal-interference-noise-ratio (SINR) [3]. Usually, the clutter plus noise covariance matrix (CNCM) should be obtained by using the training samples with the adjacent to range cell under test (CUT) to estimate [4]. The Reed–Mallett–Brennan rule (RMB) [5] shows that the training samples must fulfill these conditions to maintain a 3 dB output SINR loss: (a) it should be independent and identically distributed (IID); (b) it is required to have no target; (c) the number is required to be more than twice the degrees of freedom. However, in the complex structures, such as conformal arrays, there is nonlinearity and range-dependence between the Doppler frequencies and spatial frequencies. Meanwhile, clutter has no stationarity over the range [6]. For conformal arrays, it is more difficult to acquire the needed IID training samples.

In order to deal with influence caused by range dependence, many methods based on compensation have been proposed. The angle-Doppler compensation method (ADC) [7] deterministically aligns maximum angle-Doppler responses over a range to improve the training set. However, the deterministic compensation needs a certain degree of prior information. The adaptive angle-Doppler compensation method (A^2 DC) [8,9] adaptively estimates features of dominant subspace from data and aligns range-varying responses to CUT. Some related methods [10,11] intend to apply compensation algorithms to the conformal array or other geometry. The methods mentioned above make the angle-Doppler characteristic of training samples similar to CUT. However, these methods are only suitable for antenna beam-patterns with high directivity [12].

In the last few decades, to decrease the requirement of the IID secondary samples and obtain an excellent clutter suppression performance, different effective methods have been represented. The reduced-dimension methods (RD) attempt to introduce a reduced dimension matrix to decrease the requirements of IID training samples [13–15]. The reduced-rank (RR) methods try to decrease the number of needed training samples by using a data-dependent transformation matrix [16]. However, the training samples needed by these methods is still hard to satisfy in a nonhomogeneous environment. The named direct data domain STAP method [17] solves the shortcoming of inadequate training samples by using the CUT only. Unfortunately, the DDD method introduces the cost of the decreased degrees of freedom, which causes STAP performance degradation. Recently, the knowledge-aided (KA) methods have been introduced in the STAP methods [18,19] to enhance the clutter suppression performance in nonhomogeneous environments, which utilize the prior information of the array geometry, array system parameters, or measured data to obtain a relatively accurate CNCM. However, when errors exist in the prior knowledge, the KA-STAP methods will result in significant performance loss.

With the progress of compressed sensing technology, sparse recovery (SR) methods have been introduced into STAP methods and obtained good development [20,21]. SR-STAP methods aim to acquire the precise recovery of the clutter spectrum by exploiting the clutter sparse property. These methods can estimate an accurate CNCM with a few training samples by making use of the sparse property of the clutter spectrum [22,23]. Unfortunately, these SR-STAP methods are sensitive to user parameters or have high computational burden. A sparse representation RBC method (SR-RBC) was represented to further obtain an excellent STAP performance of conformal array [24] that uses the more accurate CNCM estimated by SR method to design the transform matrix by making the training samples more stationary, but the performance of SR-RBC method is still influenced by the number of training samples. KA sparse iterative covariance-based estimation method (KASPICE) has been proposed [25] with exploiting the aided knowledge of the clutter spectrum and covariance fitting criteria, which can acquire excellent performance, and it will not be affected by user parameters. The KASPICE method requires accurate prior knowledge because it utilizes the idea of KA. A sparse Bayesian learning (SBL) algorithm for the single measurement vectors case (SMV) [26] and the algorithm for the multiple measurement vectors case (MMV) [27] was proposed due to its excellent robustness and performance. However, SBL has heavy computational complexity and converges slowly. A SBL with the fast-convergence method (FCSBL) is proposed [28] to increase the convergence speed with the same framework but start with a modified signal model.

In this work, a novel fast SBL-STAP method is proposed. The proposed method creates a novel surrogate function for the SBL cost function. Then, the surrogate function will be iteratively minimized and the updated formula for the SMV condition will be obtained. The extension to the MMV condition is straightforward. Finally, the proposed method improves the sparsity of the solution and estimates the clutter power, which calculates a relatively accurate CNCM. This method also resolves the drawbacks mentioned above and has the lower computational complexity than the SBL method and FCSBL method. The proposed method can provide satisfactory performance.

The main contributions are listed as:

(1) The concept of conjugate function will be introduced to make up a surrogate function for the SBL cost function. The STAP performance is similar to the conventional SBL method, and converges faster. Meanwhile, the proposed method has lower computational burden than the SBL and FCSBL methods.

(2) The proposed method obtains the hyper-parameter by iteratively minimizing the surrogate function. For each minimizing step, a close-form solution can be achieved, which will guarantee the convergence.
(3) The extension of the novel SBL method to the multiple measurement vector (MMV) condition is rather straightforward.
(4) Detailed comparison of clutter suppression performance and Capon spectrum between the proposed method and other STAP algorithms are expressed.

The remaining section is as follows. In Section 2, the clutter signal model and SR-STAP based on conformal array are introduced. In Section 3, the proposed method is expressed. In Section 4, the performance of the various methods will be evaluated with simulated data. In Section 5, the conclusions are given.

Notations: Vector, matrix and scalar quantity are represented by boldface lower case, boldface upper case and italic typeface. The matrix inverse, transpose and conjugate transpose are represented by $[\cdot]^{-1}$ $[\cdot]^{T}$ and $[\cdot]^{H}$ respectively. $\|\cdot\|_0$ is the l_0 norm. $|\cdot|$ represents the absolute value. $\|\cdot\|_1$ expresses the l_1 norm. $\|\cdot\|_2$ stands for l_2 norm or l_F norm. $\odot$ expresses the Hadamard product. $\otimes$ represents the Kronecker product. $Diag(\cdot)$ denotes transforming a vector into a diagonal. $\mathbf{I}_N$ is an identity matrix with $N \times N$. $\mathcal{CN}(\mu, \sigma)$ represents the complex Gaussian distribution where μ denotes mean and σ stands for covariance.

2. Signal Model and SR-STAP Model and Principle

2.1. Signal Model

The signal model of conformal array which considers polarization has been proposed in our previous work [29]. In this section, the signal model is to be expressed briefly. The radar system is shown in Figure 1. The $\mathbf{v}$ can be expressed as

$$\mathbf{v}= [v\cos\psi, v\sin\psi, 0]^{T} \tag{1}$$

where v represents the platform velocity and ψ stands for crab angle.

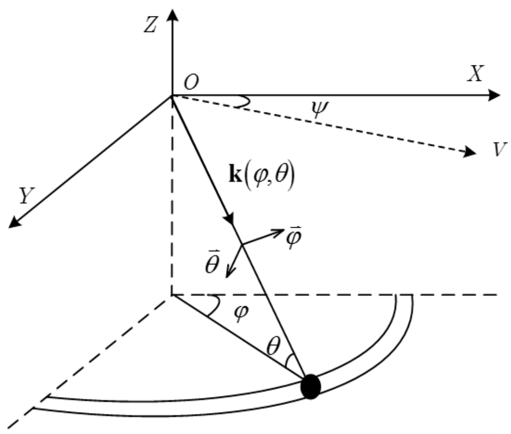


Figure 1. Abridged general view of airborne radar system.

The unit vector $\mathbf{k}(\varphi, \theta)$ points to propagate direction and could be expressed as

$$\mathbf{k}(\varphi, \theta) = [\cos\varphi\cos\theta, \sin\varphi\cos\theta, \sin\theta] \tag{2}$$

where φ stands for the azimuth angle and θ represents elevation angle.

Figure 2 exhibits the configuration of a circular arc array. The conformal array has N elements. Wavelength is denoted by λ. The element spacing is $\lambda/2$. The direction vector to the ith element is expressed by $\mathbf{r}_i = [x_i, y_i, z_i]^{T}$.

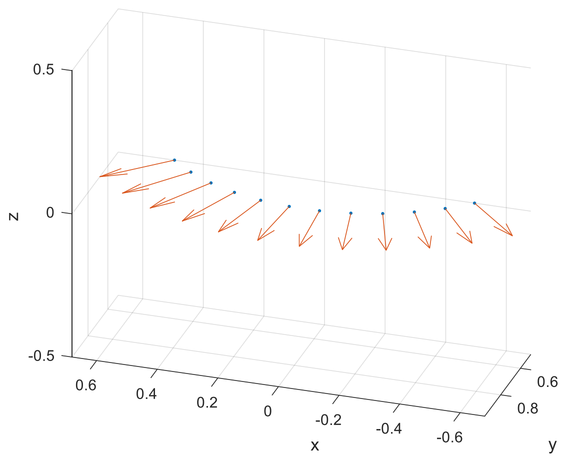


Figure 2. Geometry configuration.

The normalized array factor from (φ, θ) can be written as

$$\mathbf{s}_0(\varphi, \theta) = [e^{j\frac{2\pi}{\lambda}\mathbf{k}\cdot\mathbf{r}_1}, e^{j\frac{2\pi}{\lambda}\mathbf{k}\cdot\mathbf{r}_2}, \cdots, e^{j\frac{2\pi}{\lambda}\mathbf{k}\cdot\mathbf{r}_N}]^T \tag{3}$$

In this work, consider the individually-polarized dipole conformal array to be placed tangentially to circumference. $\mathbf{g}_i$ represents the element factor of the ith element and could be divided into two polarizations which are mutually orthogonal

$$\mathbf{g}_i = g_{i\vec{\varphi}}(\varphi, \theta)\vec{\varphi} + g_{i\vec{\theta}}(\varphi, \theta)\vec{\theta} \tag{4}$$

where $g_{i\vec{\varphi}}(\varphi, \theta)$ and $g_{i\vec{\theta}}(\varphi, \theta)$ stand for the $\vec{\varphi}$ polarized component and $\vec{\theta}$ polarized component, respectively.

Element factor is expressed in the local coordinate, and $\tilde{\mathbf{g}}_i(\tilde{x}, \tilde{y}, \tilde{z})$ is known. The accurate formula of the conformal array space steering vector in global coordinates should be obtained. For this purpose, the Euler rotation matrix [30,31] would be used as the transformed tool between the local coordinate system and the global coordinate system for the element factor.

The transformation matrix $(x, y, z) \Rightarrow (\tilde{x}, \tilde{y}, \tilde{z})$ could be expressed as $\mathbf{T} = \mathbf{T}_x\mathbf{T}_y\mathbf{T}_z$, where

$$\mathbf{T}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\sigma_x & -\sin\sigma_x \\ 0 & \sin\sigma_x & \cos\sigma_x \end{bmatrix} \tag{5}$$

$$\mathbf{T}_y = \begin{bmatrix} \cos\sigma_y & 0 & \sin\sigma_y \\ 0 & 1 & 0 \\ -\sin\sigma_y & 0 & \cos\sigma_y \end{bmatrix} \tag{6}$$

$$\mathbf{T}_z = \begin{bmatrix} \cos\sigma_z & -\sin\sigma_z & 0 \\ \sin\sigma_z & \cos\sigma_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

Utilizing the transformation matrix, the following could be obtained

$$\tilde{\mathbf{g}}_i(\tilde{x}, \tilde{y}, \tilde{z}) = \mathbf{T} \cdot \mathbf{g}_i(x, y, z) \tag{8}$$

Then, the coordinate transformation could be written as

$$\mathbf{g}_i(x, y, z) = \mathbf{T}^{-1} \cdot \tilde{\mathbf{g}}_i(\tilde{x}, \tilde{y}, \tilde{z}) \tag{9}$$

The expression of $\mathbf{g}_i$ in the spherical coordinate system $g_{i\vec{\varphi}}(\varphi, \theta)$ and $g_{i\vec{\theta}}(\varphi, \theta)$ could be obtained using the Cartesian coordinate system by means of the following transformation formula

$$\begin{bmatrix} \vec{\varphi} \\ \vec{\theta} \end{bmatrix} = \begin{bmatrix} -\sin\varphi & \cos\varphi & 0 \\ \sin\theta\cos\varphi & \sin\theta\sin\varphi & -\cos\theta \end{bmatrix} \begin{bmatrix} \vec{x} \\ \vec{y} \\ \vec{z} \end{bmatrix} \tag{10}$$

Now, the conformal array exact space steering vector could be written as follows

$$\mathbf{S}_s(\varphi, \theta) = [\mathbf{G}_\varphi(\varphi, \theta) \odot \mathbf{s}_0(\varphi, \theta), \mathbf{G}_\theta(\varphi, \theta) \odot \mathbf{s}_0(\varphi, \theta)], \mathbf{S}_s \in \mathbb{C}^{N\times 2} \tag{11}$$

where

$$\mathbf{G}_\varphi(\varphi, \theta) = [g_{1\vec{\varphi}}(\varphi, \theta), g_{2\vec{\varphi}}(\varphi, \theta), \cdots, g_{N\vec{\varphi}}(\varphi, \theta)]^T \tag{12}$$

$$\mathbf{G}_\theta(\varphi, \theta) = [g_{1\vec{\theta}}(\varphi, \theta), g_{2\vec{\theta}}(\varphi, \theta), \cdots, g_{N\vec{\theta}}(\varphi, \theta)]^T \tag{13}$$

It is assumed that the array emits M pulses. f_r stands for pulse repetition frequency. The temporal steering vector could be represented as

$$\mathbf{s}_d(\varphi, \theta) = [1, e^{j\frac{2\pi}{\lambda f_r}2\mathbf{k}\cdot\mathbf{v}}, \cdots, e^{j\frac{2\pi}{\lambda f_r}2\mathbf{k}\cdot\mathbf{v}(M-1)}]^T \tag{14}$$

Then, the space-time steering vector for conformal array could be represented by

$$\mathbf{S}_{st}(\varphi, \theta) = \mathbf{s}_d(\varphi, \theta) \otimes \mathbf{S}_s(\varphi, \theta), \mathbf{S}_{st} \in \mathbb{C}^{NM\times 2} \tag{15}$$

In this paper, it is supposed that the wave is the completely polarized wave. The Jones vector is used to denote the completely polarized wave, and it can be written as

$$\mathbf{e}_p = \begin{bmatrix} \cos\mu \\ \sin\mu e^{j\beta} \end{bmatrix} \tag{16}$$

where $\mu \in [0, \pi/2]$ denotes the polarization angle and $\beta \in [-\pi, \pi]$ is the polarization phase difference [32].

Finally, the clutter data $\mathbf{c} \in \mathbb{C}^{NM\times 1}$ from (φ, θ) can be represented as

$$\mathbf{c} = \alpha\mathbf{S}_{st}(\varphi, \theta) \cdot \mathbf{e}_p = \alpha\mathbf{s}(\varphi, \theta) \tag{17}$$

where $\mathbf{e}_p$ expresses the polarized state of the echo and α denotes the complex amplitude of the clutter scatter.

The received data from the CUT could be denoted as

$$\mathbf{x} = \sum_{l=1}^{N_l}\sum_{c=1}^{N_c} \alpha_{l,c}\mathbf{s}_{l,c} + \mathbf{n} \tag{18}$$

where N_l stands for the number of ambiguous range cells and $\mathbf{s}_{l,c} = \mathbf{S}_{st}(\varphi_c, \theta_l) \cdot \mathbf{e}_{p,(l,c)}$. A clutter range gate could be considered as a superposition of N_c clutter scatters. $\mathbf{n} \in \mathbb{C}^{NM\times 1}$ represents the Gaussian noise vector.

It is supposed that clutter scatters are mutually independent. The ideal CNCM $\mathbf{R} \in \mathbb{C}^{NM\times NM}$ could be written as

$$\mathbf{R} = \sum_{l=1}^{N_l}\sum_{c=1}^{N_c} \sigma_{l,c}^2 \mathbf{s}_{l,c} \cdot \mathbf{s}_{l,c}{}^H + \sigma_n^2\mathbf{I}_{NM} \tag{19}$$

where $\sigma_{l,c}^2$ and σ_n^2 represent the power of the l,cth clutter scatter and the noise power, respectively.

The STAP weight vector could be acquired through the linearly constrained minimum variance criterion:

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{s}_t}{\mathbf{s}_t^H\mathbf{R}^{-1}\mathbf{s}_t} \tag{20}$$

where the target space-time steering vector could be denoted by $\mathbf{s}_t = \mathbf{S}_{st}(\varphi_0, \theta_0) \cdot \mathbf{e}_{p,t}$.

Due to the unknown interference environment, the CNCM should be estimated from IID training samples, i.e., $\hat{\mathbf{R}} = \frac{1}{L}\sum_{l}^{L}\mathbf{x}_l\mathbf{x}_l^H$, where L stands for the number of the IID training samples. It is often difficult to acquire enough IID training samples, leading to bad STAP performance.

2.2. SR-STAP Formulation

In conventional SR-STAP methods, the angle-Doppler plane is often uniformly divided into $K = N_sN_f$ grid points, and $N_f = \rho_f M\left(\rho_f > 1\right)$, $N_s = \rho_s N(\rho_s > 1)$ denotes the number of normalized Doppler frequencies grids and normalized spatial frequencies grids, respectively. The space-time steering vector from these grid points form the dictionary matrix. However, in the case of conformal array, the formula of the spatial steering vector is different from the planar array. It cannot be expressed in terms of spatial frequency. Thus, the form of the dictionary matrix for conformal array is also different.

The spatial steering vectors of the conformal array are decided by azimuth angle φ, essentially due to the fixed elevation angle θ within a certain range cell. Therefore, we could divide the azimuth angle to produce spatial grids. In this paper, the angle and Doppler frequencies are also divided into $K = N_sN_d$ grids. Let $\varphi_i = 2\pi i/N_s(1 \le i \le N_s)$ and $f_{d,j} = j/N_d(-N_d/2 \le j \le N_d/2)$ represent the uniformly divided azimuth angle and Doppler frequency, respectively. The corresponding time steering vector is

$$\mathbf{s}_d\left(f_{d,j}\right) = \left[1, e^{j2\pi f_{d,j}}, \cdots, e^{j(M-1)2\pi f_{d,j}}\right]^T \tag{21}$$

The SR-STAP formulation for the SMV condition could be

$$\mathbf{x} = \mathbf{D}\mathbf{y} + \mathbf{n} \tag{22}$$

where $\mathbf{D}$ has following formula:

$$\begin{aligned}\mathbf{D} =& [\mathbf{s}_d(f_{d,1}) \otimes \mathbf{S}_s(\varphi_1, \theta_k), \cdots, \mathbf{s}_d(f_{d,1}) \otimes \mathbf{S}_s(\varphi_{N_s}, \theta_k), \\ & \cdots, \mathbf{s}_d\left(f_{d,N_d}\right) \otimes \mathbf{S}_s(\varphi_{N_s}, \theta_k)]\ \mathbb{C}^{NM \times 2N_sN_d}\end{aligned} \tag{23}$$

$\mathbf{y} \in \mathbb{C}^{2N_sN_d \times 1}$ can be written as

$$\mathbf{y} = \left[\alpha_1 \cos\mu_1, \alpha_1 \sin\mu_1 e^{j\beta_1}, \cdots, \alpha_{N_sN_d}\cos\mu_{N_sN_d}, \alpha_{N_sN_d}\sin\mu_{N_sN_d}e^{j\beta_{N_sN_d}}\right]^T \tag{24}$$

Here, $\mathbf{D}$ is a dictionary matrix including space-time steering vectors from all grid points. $\mathbf{y}$ denotes sparse coefficient vector

In following subsection, we remark that $K = 2N_sN_d$.

In sparse recovery methods, the SR-STAP problem could be solved by the following

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{\text{argmin}}\|\mathbf{y}\|_0, \text{ s.t. } \|\mathbf{x} - \mathbf{D}\mathbf{y}\| \le \varepsilon_1 \tag{25}$$

where ε_1 represents the fitting error tolerance. We cannot solve (25) directly due to the NP-hard problem. Hence, l_1 norm were used in SR-STAP problem.

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{y}\|_1, \text{ s.t. } \|\mathbf{x} - \mathbf{D}\mathbf{y}\| \leq \varepsilon_2 \tag{26}$$

where ε_2 is a user parameter.

3. The Proposed Method

3.1. Derivation of the Proposed Method

In this framework, we adopt the Bayesian perspective to address the linear problem [26]. This will determine the likelihood function of the complex data $\mathbf{x}$ and the posterior distribution function of the complex amplitude $\mathbf{y}$.

If it is supposed that $\mathbf{n}$ in (22) denotes a complex Gaussian vector, the likelihood function of the received data $\mathbf{x}$ could be represented as

$$p(\mathbf{x}|\mathbf{y}, \eta) = \frac{1}{(\pi\eta)^{NM}} e^{-\frac{1}{\eta}\|\mathbf{x}-\mathbf{D}\mathbf{y}\|^2} \tag{27}$$

Assuming that $\mathbf{y}$ is considered to subject to a zeros-mean complex Gaussian prior distribution,

$$\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}) \tag{28}$$

where $\boldsymbol{\Gamma} = \text{Diag}(\boldsymbol{\gamma})$. $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \cdots, \gamma_K)^T$ is a variance vector.

The following is the formula for $\mathbf{y}$:

$$p(\mathbf{y}|\boldsymbol{\Gamma}) = \pi^{-K} |\boldsymbol{\Gamma}|^{-1} e^{-\mathbf{y}^H \boldsymbol{\Gamma}^{-1} \mathbf{y}} \tag{29}$$

Given the likelihood for the radar array received data $\mathbf{x}$ (27) and the prior (29), the posterior probability density function (PDF) of the complex amplitude $\mathbf{y}$ could be expressed as following based on the Bayesian rule:

$$p(\mathbf{y}|\mathbf{x}, \boldsymbol{\Gamma}, \eta) = \frac{p(\mathbf{x}|\mathbf{y}, \eta) p(\mathbf{y}|\boldsymbol{\Gamma})}{\int p(\mathbf{x}|\mathbf{y}, \eta) p(\mathbf{y}|\boldsymbol{\Gamma}) d\mathbf{y}} \tag{30}$$

Assuming that the $\boldsymbol{\gamma}^{(t)}$ and $\eta^{(t)}$ are known in the tth iteration step, the PDF of $\mathbf{y}$ in the $(t+1)$th iteration could be further represented as:

$$\begin{aligned} p\left(\mathbf{y}^{(t+1)} \middle| \mathbf{x}, \boldsymbol{\Gamma}^{(t)}, \eta^{(t)}\right) &= \frac{p\left(\mathbf{x}, \mathbf{y}^{(t+1)} \middle| \boldsymbol{\Gamma}^{(t)}, \eta^{(t)}\right)}{p\left(\mathbf{x}, \middle| \boldsymbol{\Gamma}^{(t)}, \eta^{(t)}\right)} \\ &= \pi^{-K} \left|\boldsymbol{\Sigma}^{(t+1)}\right|^{-1} e^{-\left(\mathbf{y}^{(t+1)} - \boldsymbol{\mu}^{(t+1)}\right)^H \left|\boldsymbol{\Sigma}^{(t+1)}\right|^{-1} \left(\mathbf{y}^{(t+1)} - \boldsymbol{\mu}^{(t+1)}\right)} \end{aligned} \tag{31}$$

where

$$\boldsymbol{\mu}^{(t+1)} = \boldsymbol{\Gamma}^{(t)} \mathbf{D}^H \boldsymbol{\Lambda}^{-1} \mathbf{x} \tag{32}$$

$$\boldsymbol{\Sigma}^{(t+1)} = \boldsymbol{\Gamma}^{(t)} - \boldsymbol{\Gamma}^{(t)} \mathbf{D}^H \boldsymbol{\Lambda}^{-1} \mathbf{D} \boldsymbol{\Gamma}^{(t)} \tag{33}$$

where

$$\boldsymbol{\Lambda} = \mathbf{D} \boldsymbol{\Gamma}^{(t)} \mathbf{D}^H + \eta^{(t)} \mathbf{I} \tag{34}$$

The $\boldsymbol{\gamma}^{(t)}$ could be estimated by a type-II maximum likelihood, which stands for maximizing the marginal likelihood function or other form of it with regard to $\boldsymbol{\gamma}$. The function is the product of (27) and (29) integrated over the complex amplitudes of $\mathbf{y}$

$$\begin{aligned} p(\mathbf{x}|\boldsymbol{\Gamma},\eta) &= \int p(\mathbf{x}|\mathbf{y},\eta)p(\mathbf{y}|\boldsymbol{\Gamma})d\mathbf{y} \\ &= \pi^{-NM}|\mathbf{C}|^{-1}e^{-\mathbf{x}^H\mathbf{C}^{-1}\mathbf{x}} \end{aligned} \tag{35}$$

The marginal log-likelihood function is

$$\begin{aligned} \mathcal{R}(\boldsymbol{\gamma}) &= \ln p(\mathbf{x}|\boldsymbol{\Gamma},\eta) \\ &= -\ln|\mathbf{C}| - \mathbf{x}^H\mathbf{C}^{-1}\mathbf{x} + \text{constant} \end{aligned} \tag{36}$$

where

$$\mathbf{C} = \mathbf{D}\boldsymbol{\Gamma}\mathbf{D}^H + \eta\mathbf{I} \tag{37}$$

The hyper-parameters $\boldsymbol{\gamma}$ will be obtained by maximizing the log function.

$$\begin{aligned} \hat{\boldsymbol{\gamma}} &= \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \mathcal{R}(\boldsymbol{\gamma}) \\ &= \underset{\boldsymbol{\gamma}}{\operatorname{argmin}} \mathcal{L}(\boldsymbol{\gamma}) \triangleq \ln|\mathbf{C}| + \mathbf{x}^H\mathbf{C}^{-1}\mathbf{x} \end{aligned} \tag{38}$$

Next, considering that the log-determinant in (38) is concave, a continuous surrogate function will be constructed as the cost function of the optimization problem and the closed solution will be obtained.

The log-determinant term $\ln|\mathbf{C}|$ of $\mathcal{L}(\boldsymbol{\gamma})$

$$\ln|\mathbf{C}| = \ln\left|\mathbf{D}\boldsymbol{\Gamma}\mathbf{D}^H + \eta\mathbf{I}\right| = \ln\left|\mathbf{D}diag(\boldsymbol{\gamma})\mathbf{D}^H + \eta\mathbf{I}\right| \tag{39}$$

By defining new variables $\tau_i = \ln(\gamma_i)$ for $i = 1,2,\cdots,K$, the non-convex term $\ln\left|\mathbf{D}diag(\boldsymbol{\gamma})\mathbf{D}^H + \eta\mathbf{I}\right|$ can be written as

$$\begin{aligned} \ln\left|\mathbf{D}diag(\boldsymbol{\gamma})\mathbf{D}^H + \eta\mathbf{I}\right| &= \ln\left|\sum_{i=1}^{K}\gamma_i\mathbf{d}_i{\mathbf{d}_i}^H + \eta\mathbf{I}\right| \\ &= \ln\left|\sum_{i=1}^{K}e^{\tau_i}\mathbf{d}_i{\mathbf{d}_i}^H + \eta\mathbf{I}\right| \end{aligned} \tag{40}$$

where $\mathbf{d}_i$ stands for the ith column of $\mathbf{D}$.

So, utilizing the concept of the conjugate function, the conjugate function of $\ln|\mathbf{C}|$ can be obtained by

$$h^*(\mathbf{z}) = \sup_{\boldsymbol{\tau}} \mathbf{z}^H\boldsymbol{\tau} - \ln|\mathbf{C}| \tag{41}$$

Then, based on Fenchel inequality, the following can be obtained

$$\ln|\mathbf{C}| + h^*(\mathbf{z}) \geq \mathbf{z}^H\boldsymbol{\tau} \tag{42}$$

With the above, the (38) can be expressed as

$$\hat{\boldsymbol{\gamma}} = \min_{\gamma} \max_{\mathbf{z}>0} \mathbf{z}^H\boldsymbol{\tau} - h^*(\mathbf{z}) + \mathbf{x}^H\mathbf{C}^{-1}\mathbf{x} \tag{43}$$

From [26], it can be known that the $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can also be expressed as

$$\boldsymbol{\Sigma} = \left(\eta^{-1}\mathbf{D}^H\mathbf{D} + \boldsymbol{\Gamma}^{-1}\right)^{-1} \tag{44}$$

$$\boldsymbol{\mu} = \eta^{-1}\boldsymbol{\Sigma}\mathbf{D}^H\mathbf{x} \tag{45}$$

Based on matrix inversion lemma and (44) and (45), the following formula can be obtained

$$\begin{aligned}\mathbf{C}^{-1} &= \left(\eta\mathbf{I} + \mathbf{D}\boldsymbol{\Gamma}\mathbf{D}^H\right)^{-1} \\ &= \eta^{-1}\mathbf{I} - \eta^{-1}\mathbf{I}\mathbf{D}\left(\boldsymbol{\Gamma}^{-1} + \eta^{-1}\mathbf{D}^H\mathbf{D}\right)^{-1}\mathbf{D}^H\eta^{-1}\mathbf{I} \\ &= \eta^{-1}\mathbf{I} - \eta^{-1}\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^H\eta^{-1}\end{aligned} \tag{46}$$

The second term of (38) can be further derived

$$\begin{aligned}\mathbf{x}^H\mathbf{C}^{-1}\mathbf{x} &= \eta^{-1}\mathbf{x}^H\mathbf{x} - \eta^{-1}\mathbf{x}^H\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^H\eta^{-1}\mathbf{x} \\ &= \eta^{-1}\mathbf{x}^H(\mathbf{x} - \mathbf{D}\boldsymbol{\mu}) \\ &= \eta^{-1}\|\mathbf{x} - \mathbf{D}\boldsymbol{\mu}\|_2^2 + \eta^{-1}\boldsymbol{\mu}^H\mathbf{D}^H(\mathbf{x} - \mathbf{D}\boldsymbol{\mu}) \\ &= \eta^{-1}\|\mathbf{x} - \mathbf{D}\boldsymbol{\mu}\|_2^2 + \boldsymbol{\mu}^H\boldsymbol{\Gamma}^{-1}\boldsymbol{\mu}\end{aligned} \tag{47}$$

With (43), and discarding irrelevant items, the following expression can be obtained

$$\hat{\boldsymbol{\gamma}} = \min_{\gamma} \max_{\mathbf{z}>0} \mathbf{z}^H\boldsymbol{\tau} - h^*(\mathbf{z}) + \sum_{i=1}^{K} \frac{|\mu_i|^2}{\gamma_i} \tag{48}$$

It leads to a so-called iterative min–max process. The lower-bounding is repeatedly minimized, which then tightens the bound. For any fixed $\boldsymbol{\tau}^*$, given the property of the conjugate function, the '=' of the (42) will be true when $\mathbf{z}$ satisfies the following formula.

$$z_i^{(t+1)} = \left.\frac{\partial \ln|\mathbf{C}|}{\partial \tau_i}\right|_{\tau_i^*} \tag{49}$$

The following expression can be obtained

$$\begin{aligned}\frac{\partial \ln|\mathbf{C}|}{\partial \gamma_i} &= \frac{\partial \ln|\mathbf{C}|}{\partial \ln(\gamma_i)}\frac{\partial \ln(\gamma_i)}{\partial \gamma_i} \\ &= \mathbf{d}_i^H(\mathbf{C})^{-1}\mathbf{d}_i \\ &= \mathbf{d}_i^H\left(\mathbf{D}diag(\boldsymbol{\gamma})\mathbf{D}^H + \eta\mathbf{I}\right)^{-1}\mathbf{d}_i\end{aligned} \tag{50}$$

Let $\gamma_i^{(t)}$ denote the value of γ_i in the tth iteration. With (49) and (50), we can obtain the $\mathbf{z}$ in the $t+1$th iteration

$$z_i^{(t+1)} = \left.\frac{\partial \ln|\mathbf{C}|}{\partial \tau_i}\right|_{\tau_i^{(t)}} = \left.\frac{\partial \ln|\mathbf{C}|}{\partial \ln(\gamma_i)}\right|_{\gamma_i^{(t)}} = \gamma_i^{(t)}\mathbf{d}_i^H\left(\mathbf{D}diag\left(\boldsymbol{\gamma}^{(t)}\right)\mathbf{D}^H + \eta^{(t)}\mathbf{I}\right)^{-1}\mathbf{d}_i \tag{51}$$

Then, for the fixed $\mathbf{z}^{(t+1)}$, the (48) can be written as

$$\begin{aligned}\hat{\boldsymbol{\gamma}} &= \min_{\gamma}\left(\mathbf{z}^{(t+1)}\right)^H\boldsymbol{\tau} - h^*\left(\mathbf{z}^{(t+1)}\right) + \sum_{i=1}^{K}\frac{|\mu_i|^2}{\gamma_i} \\ &= \min_{\gamma}\left(\mathbf{z}^{(t+1)}\right)^H\ln(\gamma) - h^*\left(\mathbf{z}^{(t+1)}\right) + \sum_{i=1}^{K}\frac{|\mu_i|^2}{\gamma_i}\end{aligned} \tag{52}$$

The derivative of (52) with respect to γ_i can be obtained by

$$\frac{\partial\left(\left(\mathbf{z}^{(t+1)}\right)^H\ln(\boldsymbol{\gamma}) - h^*\left(\mathbf{z}^{(t+1)}\right) + \sum\limits_{i=1}^{K}\frac{|\mu_i|^2}{\gamma_i}\right)}{\partial \gamma_i} = \frac{z_i^{(t+1)}}{\gamma_i} - \frac{|\mu_i|^2}{\gamma_i^2} \tag{53}$$

Set (53) to zero, and the following update rule can be acquired

$$\gamma_i^{(t+1)} = \left(z_i^{(t+1)}\right)^{-1}\left|\mu_i^{(t+1)}\right|^2 \tag{54}$$

3.2. The Estimation of η

In the above, the value of η is assumed to be known; here, we need to estimate η. This procedure has been researched in previous work [28]. For convenience, the final result will be presented directly. In each iteration, η is updated by

$$\eta^{(t+1)} = \frac{\left\|\mathbf{x} - \mathbf{D}\boldsymbol{\mu}^{(t+1)}\right\|_2^2 + \eta^{(t)}\sum_{i=1}^{K}\gamma_i^{(t)}\mathbf{d}_i^H\left(\mathbf{C}^{(t)}\right)^{-1}\mathbf{d}_i}{NM} \tag{55}$$

Since the proposed method is an iterative process, initializing the hyper-parameter $\gamma_i^{(0)} = \left|\mathbf{d}_i^H\mathbf{x}\right|^2 / \left|\mathbf{d}_i^H\mathbf{d}_i\right|^2$, for $i = 1, 2, \cdots, K$, and η could be initialized with a positive scalar. We use $\eta^{(0)} = 1$.

The convergence criterion of the iteration is followed

a. The number of iterations reaches the upper limit.
b. The estimate of hyper-parameter $\boldsymbol{\gamma}$ meet $\left\|\boldsymbol{\gamma}^{(t)} - \boldsymbol{\gamma}^{(t-1)}\right\|_2 / \left\|\boldsymbol{\gamma}^{(t)}\right\|_2 < \delta$, where δ is a small positive number.

By the proposed method, the relatively precise estimate of $\boldsymbol{\mu}$ and η can be obtained. The estimated CNCM expresses as follows:

$$\mathbf{R} = \sum_{i=1}^{K}\left(|\mu_i|^2\right)\mathbf{d}_i\mathbf{d}_i^{\mathrm{H}} + \eta\mathbf{I} \tag{56}$$

The pseudo-code for the NFSBL method is given in Algorithm 1.

Algorithm 1: NFSBL Method

Step 1	Input data $\mathbf{x}$ and dictionary matrix $\mathbf{D}$
Step 2	Initialize $\boldsymbol{\gamma}$ and η
Step 3	Update the mean vector $\boldsymbol{\mu}$ using (32)
Step 4	Update $\boldsymbol{\gamma}$ and η using (54) and (55), respectively
Step 5	Repeat step 3–4 until the convergence criterion is met
Step 6	Calculate CNCM $\mathbf{R} = \sum_{i=1}^{K}\left(|\mu_i|^2\right)\mathbf{d}_i\mathbf{d}_i^{\mathrm{H}} + \eta\mathbf{I}$
Step 7	Compute STAP weight $\mathbf{w}$

3.3. Extension to the MMV Case

With above discussion, it has been known that the (23) is the dictionary matrix which contains the steering vectors of kth range cell. By carrying out some calculations, a conclusion can be drawn that the elevation angle from the neighboring range cells are only slightly different from the elevation angle of kth range cell. In this work, the proposed method can approximatively consider that the dictionary matrix in (23) also contains the steering vectors of neighboring range cells. Then, the proposed method could be developed to the MMV case.

The SR formulation for MMV case is

$$\mathbf{X} = \mathbf{D}\mathbf{Y} + \mathbf{N} \tag{57}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_L]$ and $\mathbf{Y}, \mathbf{N}$ have similar structures.

In the MMV case, the estimate of the $\mathbf{Y}$ is denoted by the matrix form of (32).

$$\boldsymbol{\Psi}^{(t+1)} = \boldsymbol{\Gamma}^{(t)}\mathbf{D}^{H}\boldsymbol{\Lambda}^{-1}\mathbf{X} \tag{58}$$

where $\boldsymbol{\Psi} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_L]$.

Similarly, (54) and (55) could be directly developed to the MMV case and the following can be obtained

$$\gamma_i^{(t+1)} = \left(z_i^{(t+1)}\right)^{-1}\frac{1}{L}\sum_{l=1}^{L}\left|\mu_{l,i}^{(t+1)}\right|^2 \tag{59}$$

$$\eta^{(t+1)} = \frac{\frac{1}{L}\sum\limits_{l=1}^{L}\left\|\mathbf{x} - \mathbf{D}\boldsymbol{\mu}_l^{(t+1)}\right\|_2^2}{NM - \sum\limits_{i=1}^{K}\gamma_i^{(t)}\mathbf{d}_i^H\left(\mathbf{C}^{(t)}\right)^{-1}\mathbf{d}_i} \tag{60}$$

The hyper-parameter $\boldsymbol{\gamma}$ could be initialized as $\gamma_i^{(0)} = \frac{1}{L}\sum\limits_{l=1}^{L}\left|\mathbf{d}_i^H\mathbf{x}_l\right|^2 / \left|\mathbf{d}_i^H\mathbf{d}_i\right|^2$.

The estimated CNCM expresses as follows

$$\mathbf{R} = \frac{1}{L}\sum_{l=1}^{L}\sum_{i=1}^{K}\left(\left|\mu_{l,i}\right|^2\right)\mathbf{d}_i\mathbf{d}_i^{\mathrm{H}} + \eta\mathbf{I} \tag{61}$$

The proposed method is summarized in Algorithm 2.

Algorithm 2: M-NFSBL Method

Step 1	Input data $\mathbf{X}$ and dictionary matrix $\mathbf{D}$
Step 2	Initialize $\boldsymbol{\gamma}$ and η
Step 3	Update the mean matrix $\boldsymbol{\Psi}$ using (58)
Step 4	Update $\boldsymbol{\gamma}$ and η using (59) and (60), respectively
Step 5	Repeat steps 3–4 until the convergence criterion is met
Step 6	Calculate CNCM $\mathbf{R} = \frac{1}{L}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\left(\left|\mu_{l,i}\right|^2\right)\mathbf{d}_i\mathbf{d}_i^{\mathrm{H}} + \eta\mathbf{I}$
Step 7	Compute STAP weight $\mathbf{w}$

4. Discussion

4.1. Complexity Analysis

In this work, the computational complexity will be researched in the MMV condition. In order to obtain better readability, some definitions of important parameters will be re-emphasized here. K is the number of the space-time vector of the dictionary matrix. L stands for the number of IID training samples. M represents the pulse number and N expresses the element number. The number of complex multiplications is used as the evaluation metric of the computational complexity. For convenience, we ignore the low-order multiplications. From the above analysis, the proposed method contains the calculation of $\boldsymbol{\Psi}$, $\boldsymbol{\gamma}$, and η in an iteration. Therefore, it is found that the computational complexity of M-NFSBL method can be denoted as $o\left(KL + (3K + 2KL + L)MN + 3K(MN)^2 + (MN)^3\right)$. The computational complexity of other SR methods are shown in Table 1.

Table 1. The computational complexity of various methods.

Algorithm	Computational Load for an Iteration
M-CVX	$o\left((KL)^3\right)$
M-SBL	$o\left(K + K^2 + KL + (2K + 2KL + K^2 + L)MN + 2K(MN)^2 + (MN)^3\right)$
M-FCSBL	$o\left((4K + 2KL + L)MN + 5K(MN)^2 + (MN)^3\right)$
M-NFSBL	$o\left(KL + (3K + 2KL + L)MN + 3K(MN)^2 + (MN)^3\right)$

Computational complexity could be viewed as a function of the number of pulses. Figure 3 express the computational complexity of the various methods. If it is supposed that $K = 1600$, $N = 8$, and $L = 5$, it is concluded that the computational complexity of the M-NFSBL method is more advantageous.

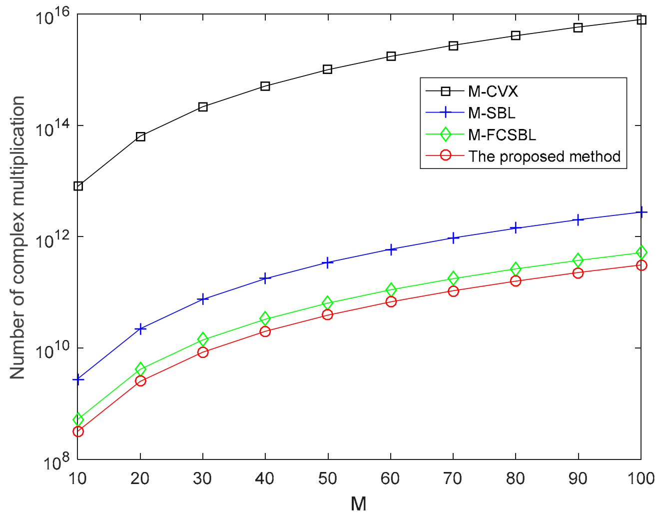


Figure 3. Computational complexity of different methods.

4.2. Convergence Analysis

The above discussion shows that the cost function could be written as $\mathcal{L}(\boldsymbol{\gamma}) = (\mathbf{z})^H \ln(\boldsymbol{\gamma}) - h^*(\mathbf{z}) + \mathbf{x}^H\mathbf{C}^{-1}\mathbf{x}$ in the every iteration so long as $\mathbf{z}$ satisfied the (51), which is a function with respect to $\boldsymbol{\gamma}$. We could obtain that $\mathcal{L}\left(\boldsymbol{\gamma}^{(t+1)}\right) \leq \mathcal{L}\left(\boldsymbol{\gamma}^{(t)}\right)$ from (53) and (54). According to [33], we know that $\mathcal{L}(\boldsymbol{\gamma})$ has a bound. It is therefore a monotonically decreasing function and indicates that the proposed method converges.

The following simulation experiments demonstrate that the proposed method converges faster than SBL.

5. Simulation Results

The STAP performance of the proposed method will be assessed in this section. Given a side-looking conformal array radar, such as shown in Figure 2, the direction of the desired signal is $\varphi_0 = 90^{\circ}$, $\theta_0 = 0^{\circ}$ and the polarization angle and polarization phase difference of the desired signal are $\mu = \pi/2$ and $\beta = 0$. The relevant radar system parameters are provided in Table 2. The 300th range cell has been selected to be CUT. Then, the IF is used as the evaluation metric of STAP performance:

$$\mathrm{IF} = \frac{\left|\mathbf{w}^{\mathrm{H}}\mathbf{s}\right|^2}{\mathbf{w}^{\mathrm{H}}\mathbf{R}\mathbf{w}} \frac{\mathrm{tr}(\mathbf{R})}{\mathbf{s}^{\mathrm{H}}\mathbf{s}} \tag{62}$$

where $\mathbf{R}$ represents the ideal CNCM.

Table 2. Radar parameters.

Parameter	Value	Unit
Pulse number	16	-
Element number	12	-
Platform velocity	200	m/s
Wavelength	0.2	m
Bandwidth	5	MHz
CNR	60	dB
Distance between elements	0.1	m
Pulse repetition frequency	5000	Hz
Platform height	3000	m

As mentioned in Section 3.3, the elevation angle corresponding to the neighboring range cells are only slightly different from the elevation angle of the kth range cell. So, in the MMV case, a small number of range cells are used as the data for multiple snapshots, and the number is set as $L = 5$.

Figure 4 illustrates the IF curves of the proposed method in the SMV condition and the MMV condition. It has been shown that the M-NFSBL method has slight advantage (about 0.2 dB). When the Doppler frequency is close to ± 0.5 Hz, the M-NFSBL method shows obvious superior performance.

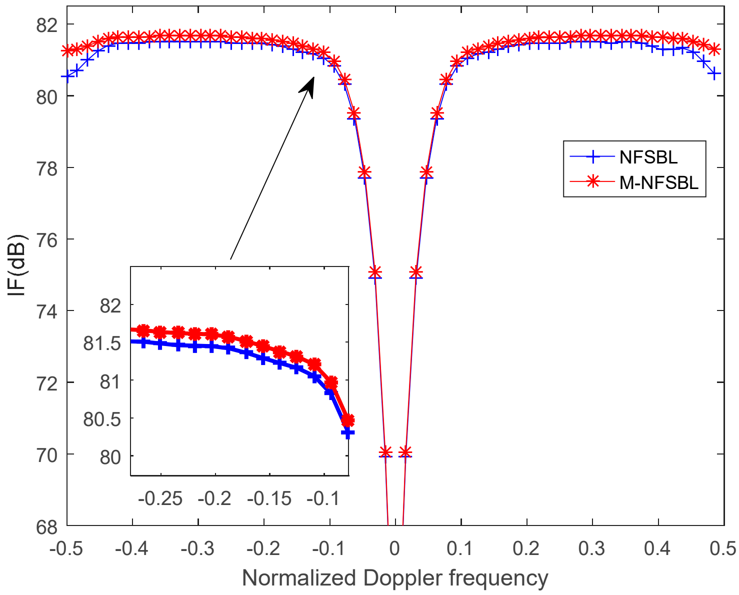


Figure 4. IF curves with SMV and MMV.

So, in the following discussion, the simulation experiments exhibit the performance with different methods in the MMV case. These methods include the sample matrix inversion method (SMI), the registration-based compensation method with sparse recovery (SR-RBC), a multiple sparse Bayesian learning with fast-convergence method (M-FCSBL), and a multiple sparse Bayesian learning method (M-SBL), and the proposed method will be evaluated with the simulated data in the ideal and non-ideal cases.

The cost function $\ln|\mathbf{C}| + \mathrm{Tr}\left(\mathbf{C}^{-1}\mathbf{R}_{ML}\right)$ will be utilized to assess the convergence performance of different methods. $\mathbf{R}_{ML} = \mathbf{X}\mathbf{X}^H/L$. Three methods will be considered here: M-SBL, M-FCSBL, and the proposed method.

As shown in Figure 5, it could be found that the M-FCSBL and the proposed method converge faster and need about 20 iterations to obtain a steady state. The M-SBL need more iterations to converge to the steady state.

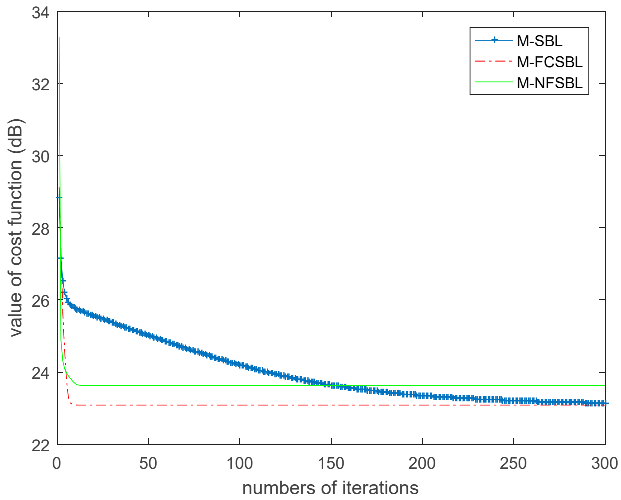
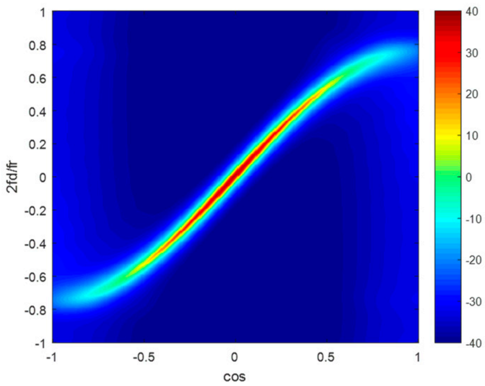


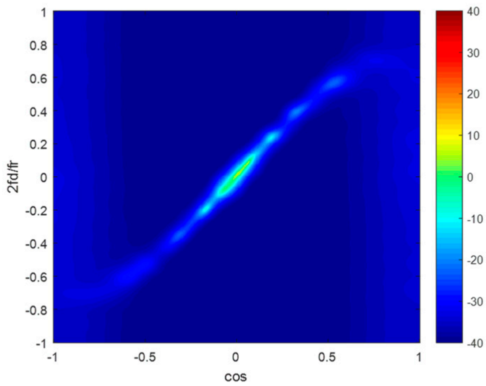
Figure 5. Value of the cost function.

5.1. Ideal Condition

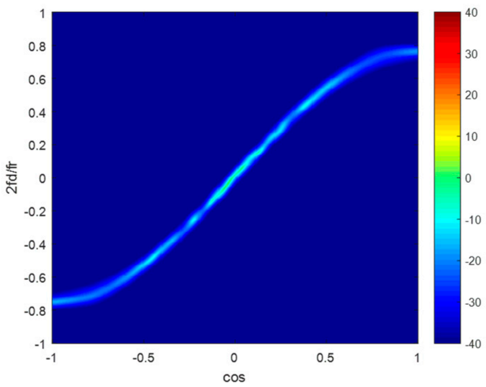
In this subsection, the clutter Capon spectrums of various methods have been given. In order to avoid the matrix singularity caused by inadequate training samples, the diagonal load method has been applied in the SMI and SR-RBC methods. From Figure 6a–f, we can know that the recovered clutter spectrums of the SMI method and the SR-RBC method are disappointing due to the inadequate training samples. It can be known that the more accurate Capon spectrums are acquired by the M-SBL method, the M-FCSBL method, and the proposed method.



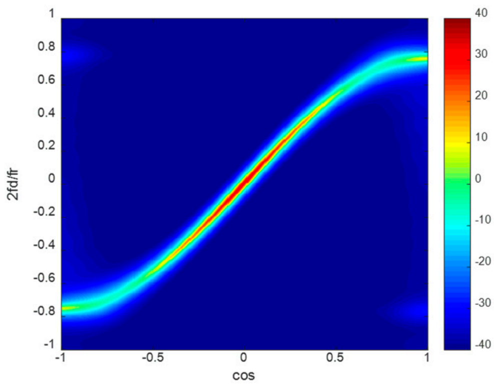
(**a**)



(**b**)



(**c**)



(**d**)

Figure 6. *Cont.*

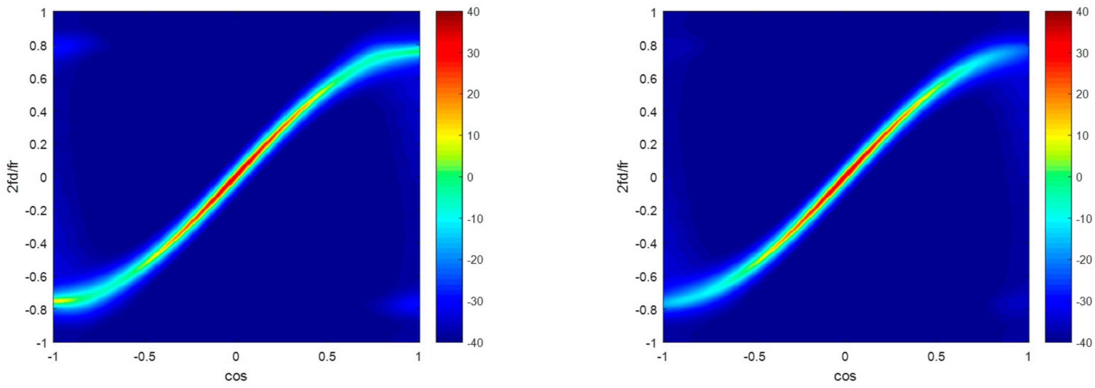


Figure 6. Clutter Capon spectrums of various methods: (**a**) exact CNCM; (**b**) SMI; (**c**) SR-RBC; (**d**) M-SBL; (**e**) M-FCSBL; (**f**) the proposed method.

Figure 7 depicts the IF comparison by different methods. It can be observed that excellent performance could be acquired by the M-SBL method, the M-FCSBL method, and the proposed method. This further indicates that the proposed method does not change the clutter suppression performance of the original M-SBL method with the faster converge rate and lower computational complexity.

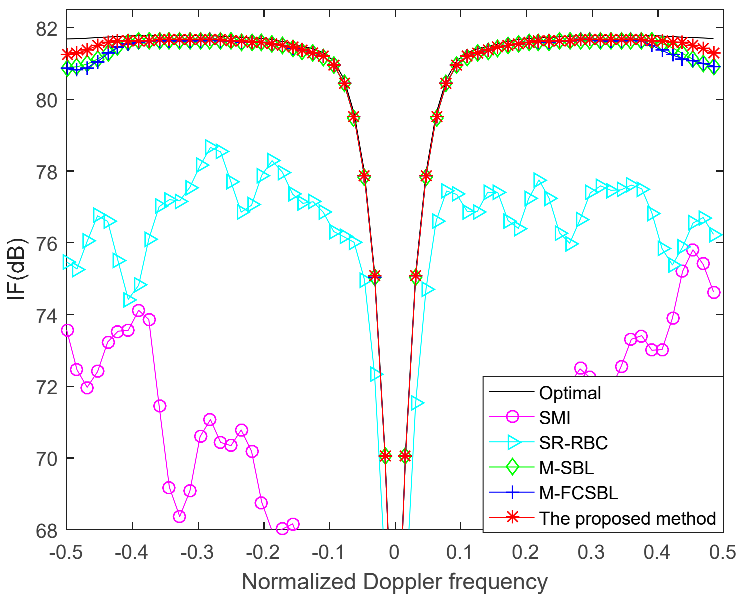


Figure 7. The IF comparison of different methods in ideal condition.

5.2. Non-Ideal Condition

In this subsection, a non-ideal condition with the presence of gain and phase error will be taken into account. This will lead to mismatch between the dictionary matrix and the real space-time steering vector. The estimated CNCM will be inaccurate.

As shown in Figure 8, the non-ideal case is introduced with 0.02 gain error and 0.2° phase error in these methods. From Figure 8a–f, it can also be found that the estimated clutter Capon spectrums from the M-SBL method, the M-FCSBL method, and the proposed method are more similar to the Capon spectrum estimated by exact CNCM.

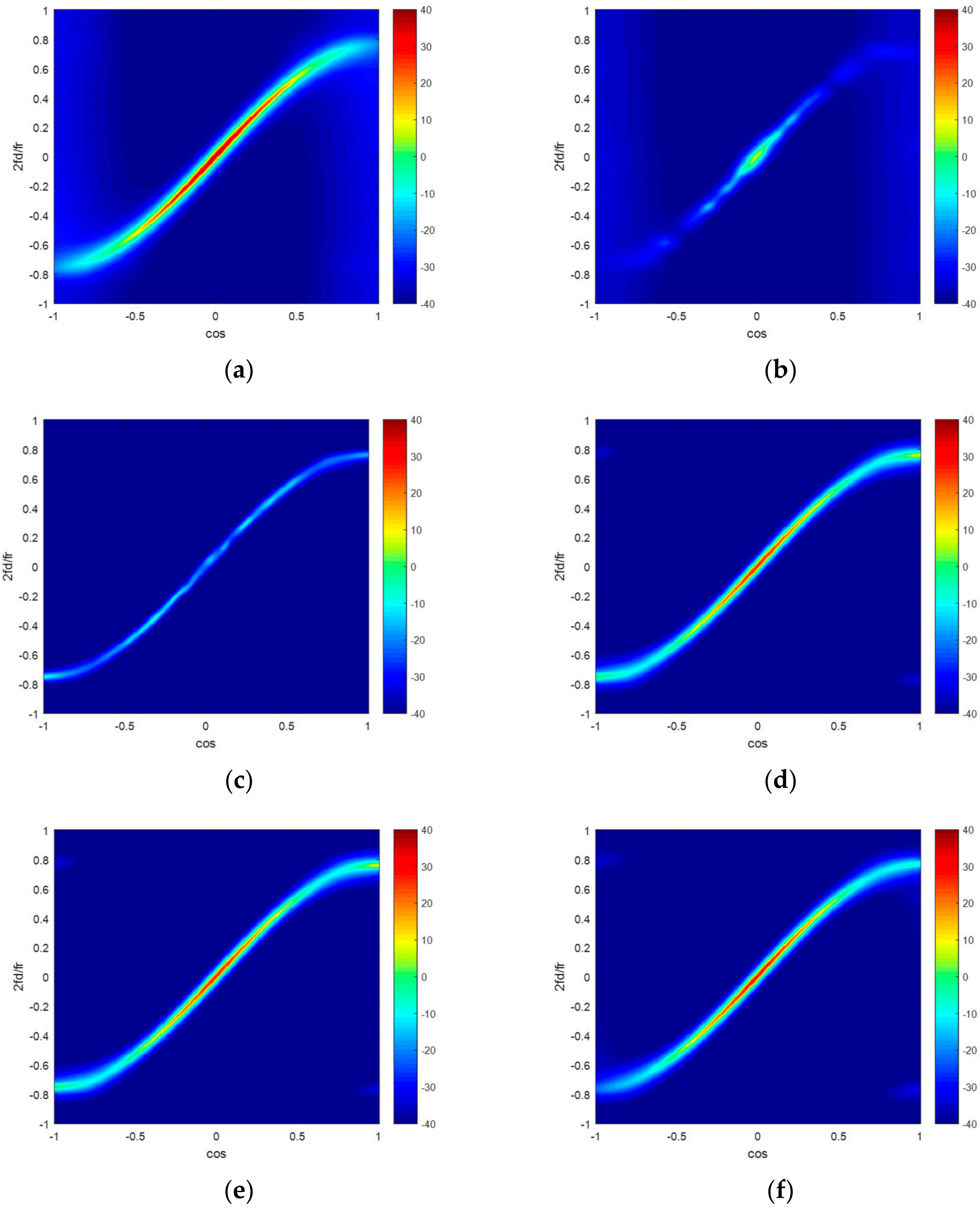


Figure 8. Clutter Capon spectrums of various methods with gain and phase error: (**a**) exact CNCM; (**b**) SMI; (**c**) SR-RBC; (**d**) M-SBL; (**e**) M-FCSBL; (**f**) the proposed method.

Figure 9 expresses the IF comparison by different methods with gain and phase error. Compared with the optimal IF curve, the M-SBL method, the M-FCSBL method, and the proposed method have little performance degradation and the main lobe notch has been slightly widened. However, the performance of these methods is still superior to other traditional methods.

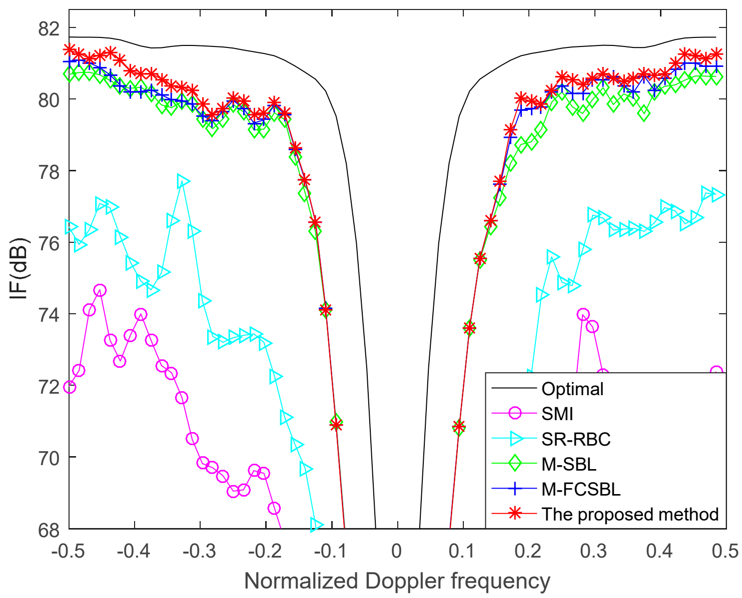


Figure 9. The IF comparison of different methods in non-ideal condition.

6. Conclusions

In this paper, a novel fast sparse Bayesian learning method named NFSBL was presented with the goal of improving the performance of suppressing clutter for conformal arrays. The novel method introduces the concept of the conjugate function to construct a surrogate function for the SBL cost function. This method improves the convergence speed of the SBL method. The proposed NFSBL method in the MMV condition has also been developed; it expresses a lower computational load compared with other SR-STAP methods. All of the simulation experiments have demonstrated that the M-NFSBL method has a superior performance in STAP.

Author Contributions: Methodology, B.R.; conceptualization, B.R.; validation B.R. and T.W.; writing-original draft preparation, B.R.; writing-review and editing, B.R.; software, B.R.; visualization, B.R.; project administration, T.W.; supervision, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: National Key R&D Program of China, grant number 2021YFA1000400.

Data Availability Statement: Not applicable.

Conflicts of Interest: There are no conflict of interest to be declared.

References

1. Hersey, R.K.; Melvin, W.L.; McClellan, J.H.; Culpepper, E. Adaptive ground clutter suppression for conformal array radar systems. *IET Radar Sonar Navig.* **2009**, *3*, 357–372. [CrossRef]
2. Ward, J. *Space-Time Adaptive Processing for Airborne Radar*; Technical Report, 1015; MIT Lincoln Laboratory: Lexington, MA, USA, 1994.
3. Klemm, R. *Principles of Space-Time Adaptive Processing*, 3rd ed.; IET: London, UK, 2006.
4. Wu, Y.F.; Wang, T.; Wu, J.X. Robust training samples selection algorithm based on spectral similarity for space-time adaptive processing in heterogeneous interference environments. *IET Radar Sonar Navig.* **2015**, *9*, 778–782.
5. Reed, I.S.; Mallett, J.D.; Brennan, L.E. Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *10*, 853–863. [CrossRef]
6. Klemm, R. *The Applications of Space-Time Adaptive Processing*, 2nd ed.; IEE Press: Stevenage, UK, 2004.
7. Himed, B.; Zhang, Y.; Hajjari, A. STAP with angle-Doppler compensation for bistatic airborne radars. In Proceedings of the IEEE National Radar Conference, Long Beach, CA, USA, 25 April 2002.
8. Melvin, W.; Himed, B.; Davis, M. Doubly adaptive bistatic clutter filtering. In Proceedings of the IEEE National Radar Conference, Hunstville, AL, USA, 8 May 2003.
9. Melvin, W.; Davis, M. Adaptive cancellation method for geometry-induced nonstationary bistatic clutter environment. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 651–672. [CrossRef]

10. Han, X.; Luo, T.; Yu, W.; Wu, M. Space-time adaptive processing for airborne radar with conformal array. In Proceedings of the IET International Radar Conference, Hangzhou, China, 14–16 October 2015; pp. 5–10.
11. Varadarajan, V.; Krolik, J.L. Joint space-time interpolation for distorted linear and bistatic array geometries. *IEEE Trans. Signal Process.* **2006**, *56*, 848–860. [CrossRef]
12. Ries, P.; Lapierre, F.D.; Verly, J.G. Geometry-induced range-dependence compensation for bistatic STAP with conformal arrays. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 275–294. [CrossRef]
13. Klemm, R. Adaptive airborne MTI: An auxiliary channel approach. *IEEE Proc. F Commun. Radar Signal Process.* **1987**, *134*, 269–276. [CrossRef]
14. Wang, Y.X.; Li, X.M.; Gao, W. Dimension-reduced bi-iterative space-time adaptive processing method for airborne radar. *J. Eng.* **2019**, *2019*, 8138–8141. [CrossRef]
15. Wang, X.; Yang, Z.; Yang, J.X. Reduced-dimension space-time adaptive processing for airborne radar with co-prime array. *J. Eng.* **2019**, *2019*, 5971–5974. [CrossRef]
16. Hua, Y.; Nikpour, M.; Stoica, P. Optimal reduced rank estimation and filtering. *IEEE Trans. Signal Process.* **2001**, *49*, 457–469.
17. Sarkar, T.K. A deterministic least-squares approach to space-time adaptive processing (STAP). *IEEE Trans. Antennas Propag.* **2001**, *49*, 91–103. [CrossRef]
18. Melvin, W.L.; Showman, G.A. An approach to knowledge-aided covariance estimation. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 1021–1042. [CrossRef]
19. Wicks, M.C.; Rangaswamy, M.; Adve, R.; Hale, T.B. Space-time adaptive processing: A knowledge-based perspective for airborne radar. *IEEE Signal Process. Mag.* **2006**, *23*, 51–65. [CrossRef]
20. Yang, Z.C.; Li, X.; Wang, H.Q.; Jiang, W.D. On clutter sparsity analysis in space-time adaptive processing airborne radar. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1215–1218.
21. Sen, S. Low-rank matrix decomposition and spatio-temporal sparse recovery for STAP radar. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 1510–1523. [CrossRef]
22. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* **1996**, *58*, 267–288. [CrossRef]
23. Candés, E. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Math.* **2008**, *346*, 589–592. [CrossRef]
24. Sun, K.; Meng, H.D.; Fabian, D.L. Registration-based compensation using sparse representation in conformal-array STAP. *Signal Process.* **2011**, *91*, 2268–2276. [CrossRef]
25. Tao, F.Y.; Wang, T.; Wu, J.X.; Su, Y.Y. A knowledge aided SPICE space time adaptive processing method for airborne radar with conformal array. *Signal Process.* **2018**, *152*, 54–62. [CrossRef]
26. Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn.* **2001**, *1*, 211–244.
27. Wipf, D.P.; Rao, B.D. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Trans. Signal Process.* **2007**, *55*, 3704–3716. [CrossRef]
28. Wang, Z.T.; Xie, W.C.; Duan, K.Q.; Wang, Y.L. Clutter suppression algorithm based on fast converging sparse Bayesian learning for airborne radar. *Signal Process.* **2017**, *130*, 159–168. [CrossRef]
29. Ren, B.; Wang, T. Space-time adaptive processing based on modified sparse learning via iterative minimization for conformal array radar. *Sensors* **2022**, *22*, 6917. [CrossRef] [PubMed]
30. Burger, H.A. Use of Euler-Rotation angles for generating antenna patterns. *IEEE Antennas Propag. Mag.* **1995**, *37*, 56–63. [CrossRef]
31. Milligan, T. More applications of Euler rotation angles. *IEEE Antennas Propag. Mag.* **1999**, *41*, 78–83. [CrossRef]
32. Zhuang, Z.W. *Radar Polarization Information Processing and Application*, 1st ed.; National Defense Industrial Press: Beijing, China, 1999.
33. Wipf, D.; Nagarajan, S. A new view of automatic relevance determination. In *Advances in Neural Information Processing Systems 20*; MIT Press: New York, NY, USA, 2008.