

Article

Self-Calibrating STAP Algorithm for Dictionary Dimensionality Reduction Based on Sparse Bayesian Learning

Zhiqi Gao ^{1,2}, Na Yang ^{1,2,*}, Pingping Huang ^{1,2}, Wei Xu ^{1,2} , Weixian Tan ^{1,2}  and Zhixia Wu ^{1,2}

¹ College of Information Engineering, Inner Mongolia University of Technology, Hohhot 010080, China; gzqngd@imut.edu.cn (Z.G.); hpp@imut.edu.cn (P.H.); xuwei1983@imut.edu.cn (W.X.); wxtan@imut.edu.cn (W.T.); zxwu@imut.edu.cn (Z.W.)

² Inner Mongolia Key Laboratory of Radar Technology and Application, Hohhot 010051, China

* Correspondence: 20231100112@imut.edu.cn

Abstract: Sparse recovery space–time adaptive processing (STAP) has an off-grid feature and high computational complexity. To address these shortcomings, this study proposes a self-calibrating STAP algorithm based on sparse Bayesian learning (SBL). The proposed algorithm constructs a dimensionality reduction dictionary by selecting the steering vectors corresponding to atoms with high power values. Then, a small-scale auxiliary dictionary is constructed with a stepwise search approach to calibrate the uniformly discretized dictionary. In this way, the atoms of the auxiliary dictionary can converge to the clutter ridge adaptively when off-grid occurs. The clutter plus noise covariance matrix is estimated via SBL combined with the updated dictionary. The experimental results demonstrate that the proposed algorithm can effectively suppress the clutter ridge expansion caused by the off-grid problem while reducing the computation burden significantly compared with the existing methods.

Keywords: space–time adaptive processing; sparse recovery; moving target detection; sparse Bayesian learning; dictionary correction



Academic Editor: Stefano Scanzio

Received: 11 April 2025

Revised: 2 June 2025

Accepted: 5 June 2025

Published: 8 June 2025

Citation: Gao, Z.; Yang, N.; Huang, P.; Xu, W.; Tan, W.; Wu, Z. Self-Calibrating STAP Algorithm for Dictionary Dimensionality Reduction Based on Sparse Bayesian Learning. *Electronics* **2025**, *14*, 2350. <https://doi.org/10.3390/electronics14122350>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The identification of terrestrial and aerial targets that are moving has represented a fundamental task for airborne warning radar systems in the past. However, environmental interference from surface reflections, particularly ground or sea clutter, significantly degrades detection performance. Currently, the space–time adaptive processing (STAP) method denotes a valuable solution for suppressing clutter [1–4]. Effective implementation of STAP techniques fundamentally depends on accurate calculation of the clutter plus noise covariance matrix (CCM). However, achieving an accurate CCM estimation depends upon numerous training snapshots that adhere to the independent and identically distributed (IID) conditions. Nonetheless, an adequate number of such training snapshots cannot be obtained in non-homogeneous environments, which consequently impairs the accuracy of CCM estimation and can cause a notable degradation in the performance of traditional STAP methods.

To address the above-mentioned problems, recent studies have proposed many relevant methods that can decrease the requirements for IID training samples needed for the STAP method. Some of these methods include reduced-dimension methods, reduced-rank methods, direct data domain methods, and knowledge-aided methods [5–8]. The reduced dimension methods also require a large number of IID snapshots, but this requirement remains hardly achievable in practice, particularly for systems with high degrees of freedom.

The rank-reduction methods have the advantage of reduced computational complexity, but the clutter rank cannot be accurately selected, which can severely affect the performance. The direct data domain methods can achieve a satisfactory clutter suppression effect using only data obtained from the cell under test, but their performance might degrade due to the aperture loss [9,10]. The knowledge-aided methods use prior information acquired by airborne radar to improve interference mitigation capabilities [11,12], but their performance is highly sensitive to the accuracy of the prior information. Namely, if there is a discrepancy between prior knowledge and actual data, the achievement of these approaches can significantly degrade. Therefore, solving the problem of the STAP method's performance being constrained by the number of training snapshots requires further research.

The integration of sparse representation concepts into STAP architectures has driven methodological innovations. Using the sparsity of clutter, these approaches can accomplish accurate CCM estimation under the condition of a limited number of training snapshots. However, the sparse recovery of signals is based on the l_0 norm, which is a nondeterministic polynomial (NP) hard problem and cannot be solved using conventional methods. Therefore, it is necessary to develop an equivalent solution method. Researchers have recently proposed many greedy methods, including the orthogonal matching pursuit, which is simple and has high computational efficiency [13,14]. While sparse recovery techniques demonstrate theoretical promise, their practical implementation faces limitations in consistently achieving universal optimality, potentially compromising reconstruction fidelity. An alternative approach to sparse recovery is to substitute the l_0 norm with the l_1 norm [15,16]. This substitution introduces unintended bias through the disproportionate weighting of active coefficients, ultimately degrading the covariance matrix estimation precision via excessive regularization effects. Furthermore, recent studies have applied the sparse Bayesian learning (SBL) methods to the SR-STAP algorithms. These methods allow for dynamic updating of parameter estimation using prior knowledge and data, thus reducing reliance on the hyperparameter while achieving good clutter suppression performance [17]. Reference [18] introduced the STAP method based on SBL, which employs a generalized dual Pareto prior distribution to enhance the sparsity of the coefficient matrix in the SBL model, thus improving the effectiveness of sparse recovery. An SBL-based STAP method [19], which combines SBL with off-grid self-calibration, was proposed to obtain a more accurate CCM under the conditions of insufficient snapshots and unknown parameters.

SR-STAP implementations necessitate quantizing the continuous space–time plane into discrete grids, and then an overcomplete dictionary matrix is constructed for the sparse reconstruction of snapshots [20,21]. The accomplishment of most SR-STAP methods relies on the level of discretization of the space–time plane. Although increased grid density improves theoretical reconstruction accuracy, this increases its computational burden, and adjacent dictionary atoms exhibit heightened correlation levels, which compromises the accuracy of sparse recovery [22]. In addition, the algorithms mentioned above assume that the clutter is completely allocated on the predefined space–time plane grids, whereas in real scenarios, the clutter does not perfectly align with the discretized grid points, which inevitably results in the degradation of the accuracy of sparse recovery and spectral leakage; this occasion is called the off-grid effect. In recent years, researchers have proposed several solutions to solve the aforementioned problems. For example, a low complexity off-grid STAP algorithm based on local search clutter subspace estimation (RD-LSCSE-STAP) [23], which searches for local optimal atoms based on global atoms and can avoid the impact of the off-grid effect, was proposed; however, this algorithm has a high computational cost. Reference [24] introduced a sparse Bayesian learning-based robust STAP algorithm, which constructs an error signal model using the Kronecker structure of a space–time steering

vector. By combining Bayesian inference and the expectation-maximization (EM) algorithm, this algorithm can iteratively estimate the angle-Doppler image and error parameters; this can enhance interference suppression efficacy and detection reliability in the phenomenon of array amplitude-phase errors and grid mismatch. However, significant arithmetic complexity and constrained temporal responsiveness hinder practical deployment feasibility. In addition, scholars in various fields have studied the integration of Bayesian inference with EM algorithms. For example, reference [25] proposed a channel estimation method based on the combination of Bayesian inference and EM algorithms. This method achieves efficient data estimation by sharing sparsity priors while maintaining adaptability to different system configurations and high computational efficiency.

To deal with the above-presented problem, this study offers a dimensionality reduction dictionary self-calibration STAP algorithm based on SBL. First, the proposed algorithm reduces dictionary dimensionality by selecting the steering vectors corresponding to atoms with high power values, thereby improving computational efficiency. Then, the optimal Doppler and spatial frequencies are searched in each of the time-domain and space-domain channels to obtain the steering vector corresponding to the atom with the highest power, mitigating the off-grid effect by updating the auxiliary dictionary. Finally, a hierarchical prior model is developed for complex signals, and an updated dictionary is integrated with the sparse Bayesian learning algorithm and used to update the required parameters iteratively. Furthermore, the hyperparameters are employed to estimate the clutter plus noise covariance. Simulation analyses demonstrate that the proposed algorithm can effectively overcome the off-grid effect while maintaining superior accuracy in clutter spectral estimation. Moreover, the proposed algorithm can effectively improve real-time performance compared with the existing methods while having a smaller computing burden. The overall process is shown in Figure 1.

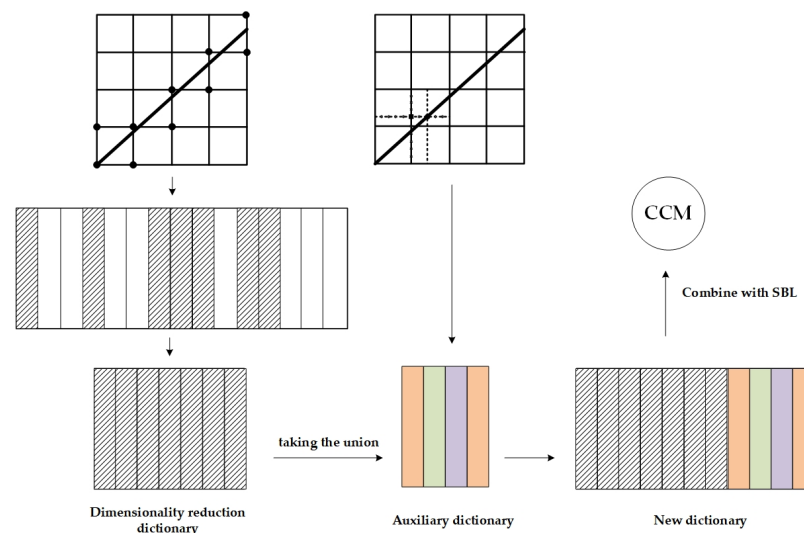


Figure 1. Distribution diagram.

This research is systematically structured across five components: Section 2 establishes fundamental radar signal modeling principles for aerial surveillance systems. Section 3 presents the method of dictionary dimension reduction and the method of auxiliary dictionary update. A performance evaluation using simulated radar data occupies Section 4, with comprehensive conclusions and research outlook detailed in the closing section.

2. Signal Modeling

2.1. STAP Signal Model

An airborne side-looking uniform linear array (ULA) radar system was used as a research object in this study. The considered ULA was composed of N elements, and the spacing between two elements, d , was set to half of the ULA's operational wavelength. The geometry of the airborne radar platform used in this study is presented in Figure 2. The radar system operates by transmitting M pulses at a constant pulse repetition frequency (PRF) f_r ; the height of the radar platform is denoted by H ; the azimuth and elevation angles of the ground reflection point are denoted by θ and φ , respectively; and the velocity of the carrier platform is V_a .

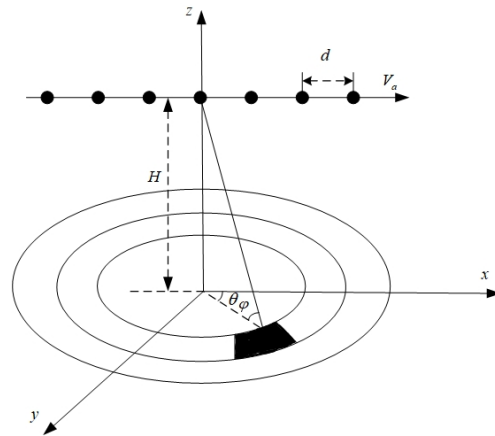


Figure 2. Illustration of the geometric configuration of an airborne radar system.

The radar-observed range cell was uniformly divided into N_c patches, and the effect of range ambiguity was neglected. Consequently, the radar echo signal model is defined as follows:

$$\begin{aligned} \mathbf{x} &= \mathbf{c} + \mathbf{n} \\ &= \sum_{k=1}^{N_c} \eta_{c,k} \mathbf{v}(f_{d,k}, f_{s,k}) + \mathbf{n} \\ &= \sum_{k=1}^{N_c} \eta_{c,k} (\mathbf{v}_t(f_{d,k}) \otimes \mathbf{v}_s(f_{s,k})) + \mathbf{n}, \end{aligned} \quad (1)$$

where $\mathbf{c}$ and $\mathbf{n}$ are the clutter and Gaussian white noise, respectively; $\eta_{c,k}$ and $\mathbf{v}(f_{d,k}, f_{s,k})$ denote the complex amplitude and space–time steering vectors of the k th clutter patch, respectively; $\mathbf{v}_t(f_{d,k})$ and $\mathbf{v}_s(f_{s,k})$ are the temporal and spatial steering vectors, respectively, and they are expressed as follows:

$$\begin{aligned} \mathbf{v}_t(f_{d,k}) &= [1, e^{j2\pi f_{d,k}}, \dots, e^{j2\pi(M-1)f_{d,k}}]^T \\ \mathbf{v}_s(f_{s,k}) &= [1, e^{j2\pi f_{s,k}}, \dots, e^{j2\pi(N-1)f_{s,k}}]^T \end{aligned} \quad (2)$$

where $f_{d,k}$ and $f_{s,k}$ represent the normalized Doppler frequency and the normalized spatial frequency of the k th clutter patch, respectively, which are defined by

$$\begin{aligned} f_{d,k} &= 2v_a \cos \varphi_n \cos \theta_n / (\lambda f_r) \\ f_{s,k} &= d \cos \varphi_n \cos \theta_n / \lambda \end{aligned} \quad (3)$$

Assuming that the clutter and noise components are uncorrelated, a clutter plus noise covariance matrix (CCM) can be defined by

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

where $E[\cdot]$ is the expectation operator and $(\cdot)^H$ is the conjugate transpose; $\mathbf{R}_c$ and $\mathbf{R}_n$ are the clutter covariance matrix and noise covariance matrix, respectively, and they are defined by

$$\begin{aligned}\mathbf{R}_c &= E[\mathbf{x}_c \mathbf{x}_c^H] = \sum_{k=1}^{N_c} |\eta_{c,k}|^2 \mathbf{v}(f_{d,k}, f_{s,k}) \mathbf{v}^H(f_{d,k}, f_{s,k}), \\ \mathbf{R}_n &= E[\mathbf{n} \mathbf{n}^H] = \sigma^2 \mathbf{I}_{MN},\end{aligned}\quad (5)$$

where σ^2 is the noise power, and $\mathbf{I}_{MN}$ is the identity matrix with a size of $MN \times MN$.

In practical scenarios, it is difficult to obtain the ideal CCM directly. Typically, adjacent samples of the cell under test are used as training snapshots, and the CCM is estimated using the maximum likelihood estimation method as follows:

$$\hat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l^H, \quad (6)$$

where L is the number of training snapshots.

The accuracy of CCM estimation is defined by the number of training snapshots, and according to the Reed–Mallett–Brennan criterion [26], the number of training snapshots should exceed twice the number of degrees of freedom of a system.

Under the criterion of maximizing the signal-to-interference-plus-noise ratio (SINR), an optimal STAP weight vector $\mathbf{w}$ can be obtained by solving the subsequent optimization problem:

$$\begin{cases} \min \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w}, \\ \text{s.t. } \mathbf{w}^H \mathbf{v}(f_{d,k}, f_{s,k}) = 1. \end{cases} \quad (7)$$

By solving the above optimal problem, the optimal weight vector of the STAP filter can be obtained by

$$\mathbf{w} = \frac{\hat{\mathbf{R}}^{-1} \mathbf{v}(f_{d,k}, f_{s,k})}{\mathbf{v}^H(f_{d,k}, f_{s,k}) \hat{\mathbf{R}}^{-1} \mathbf{v}(f_{d,k}, f_{s,k})}, \quad (8)$$

where $\mathbf{v}(f_{d,k}, f_{s,k})$ is the space–time steering vector of a target.

2.2. SR-STAP Theory and Off-Grid Problem

Based on the current research findings, the combination of sparse recovery techniques with the STAP approach can fully exploit the sparsity characteristics of clutter, enabling high-precision estimation of CCM while using a restricted set of training snapshots.

The model of a radar echo signal in the sparse recovery framework can be defined by

$$\mathbf{x} = \mathbf{V} \mathbf{A} + \mathbf{n}, \quad (9)$$

where $\mathbf{V} = [\mathbf{v}(f_{d,1}, f_{s,1}), \dots, \mathbf{v}(f_{d,N_d}, f_{s,N_s})]$ denotes the dictionary matrix with a size of $MN \times N_s N_d$; $N_s = \rho_s N$ is the number of points in the spatial domain; $N_d = \rho_d M$ is the number of points in the temporal domain; ρ_s and ρ_d indicate the resolutions in the spatial and Doppler domains, respectively; the space–time steering vectors associated with each point are termed atoms; $\mathbf{A} \in \mathbb{C}^{K \times 1}$ represents the space–time coefficients matrix and its rows' elements are either jointly negligible or non-zero; and $\mathbf{n} \in \mathbb{C}^{MN \times 1}$ is the noise matrix.

Based on the theory of sparse recovery, $\mathbf{A}$ can be obtained by solving the following optimization problem:

$$\begin{cases} \min \|\mathbf{A}\|_0, \\ \text{s.t. } \|\mathbf{x} - \mathbf{V} \mathbf{A}\|_2^2 \leq \varepsilon, \end{cases} \quad (10)$$

where $\|\cdot\|_0$ denotes the l_0 norm, $\|\cdot\|_2$ indicates the l_2 norm, and ε is the fitting error threshold.

For a side-looking ULA, the spatial and temporal frequencies of each clutter patch are coupled. When the antenna array is a side-looking array, the normalized spatial frequency f_s and normalized Doppler frequency f_d of a clutter scattering unit at any distance are directly proportional [27], that is: $f_d = \frac{2v_p}{\lambda f_r} \cos \psi$, $f_s = \frac{d}{\lambda} \cos \psi$. It can be deduced that $f_d = \beta f_s$, where $\beta = 2v_p / df_r$. Therefore, it can be seen that the slope of the clutter ridge in the angle-Doppler plane is a constant. Existing SR-STAP methods typically presume that the true clutter ridge aligns with the grid points of a discretized angle-Doppler plane. Accordingly, the clutter can be fully positioned on the grid points only when the ratio of f_s to f_d is an integer multiple of the slope of the clutter ridges in the angle-Doppler plane. However, in practical applications, a dictionary cannot be constructed to guarantee that the real clutter points are located on the grid points because accurate knowledge about the slopes of clutter ridges is not available.

To provide a more intuitive representation, this study illustrates the clutter distribution on a discretized angle-Doppler plane under the conditions of both with and without off-grid, as shown in Figure 3. The clutter distribution in the absence of grid mismatch is illustrated in Figure 3a, where the slope of the clutter ridge corresponds to the ratio of f_s to f_d , and the grid points are perfectly aligned along the clutter ridge. The clutter distribution with grid mismatch is depicted in Figure 3b, where more clutter points fall outside the grid points, yielding a wider clutter ridge and lower accuracy of clutter power estimation, which affects the CCM estimation accuracy. Although the impact of the off-grid effect can be alleviated to a certain extent by reducing the grid spacing, this can cause problems of strong coherence among dictionary atoms and increased computational load. Therefore, an effective method is urgently needed to address the grid mismatch problem.

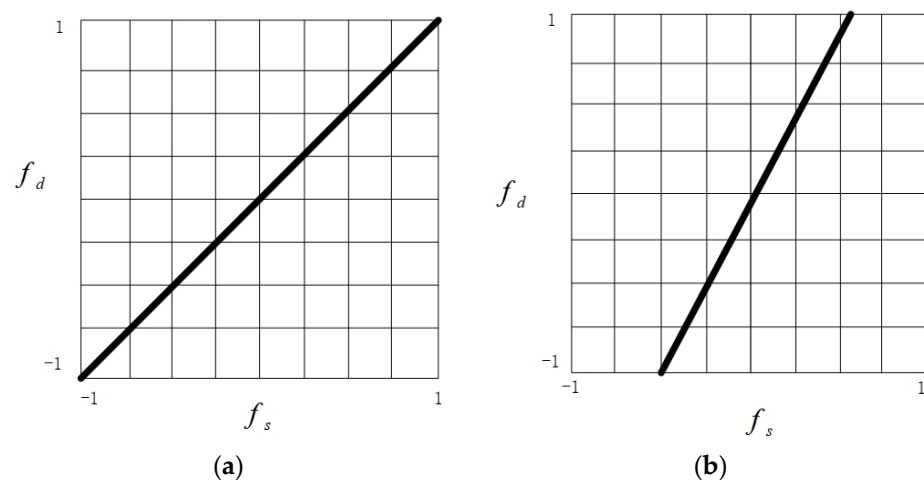


Figure 3. Distribution of the discretized space–time plane clutter: (a) clutter distribution without off-grid; (b) clutter distribution with off-grid.

3. Improved Algorithm

3.1. Dimensionality Reduction

A sparse recovery method based on Bayesian learning can achieve feature selection and model sparsification by constructing appropriate prior distributions. The advantage of this type of method is that the model that needs to be constructed depends only on a few key parameters, and various signal models can be determined using different prior distributions. However, the calculation of posterior probabilities involves complex mathematical operations, and the parameter update process requires multiple iterations, both of which can increase computational complexity. Therefore, dictionary dimensionality should be reduced to decrease computational complexity. Considering that clutter is sparse in the

angle-Doppler plane, which means that the clutter power spectrum on most discretized grid points is approximately zero, the grid points can be eliminated by removing the steering vectors corresponding to atoms with lower power values than the average, thereby enhancing the computational efficiency of the algorithm.

The algorithm initially employs the received snapshots denoted by $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L]$ to estimate a clutter covariance matrix $\mathbf{R}_x = \mathbf{X}\mathbf{X}^H/L$. Then, the power value of each clutter point is estimated using $\mathbf{R}_x$ and the steering vectors as $P(f_{d,k}, f_{s,i}) = \frac{1}{\mathbf{v}(f_{d,k}, f_{s,i})^H \mathbf{R}_x^{-1} \mathbf{v}(f_{d,k}, f_{s,i})}$.

Afterward, the estimated power values are aggregated and averaged: $\bar{P} = \frac{\sum_{i=1}^{N_s} \sum_{k=1}^{N_d} P(f_{s,i}, f_{d,k})}{N_s N_d}$, and the resultant value is defined as a threshold for dictionary atom selection. Ultimately, the power values associated with each atom are subjected to thresholding, and all grid points exceeding the threshold are retained. The retained grid points are then used to construct a reduced-dimension dictionary Ψ .

3.2. Auxiliary Dictionary Correction Methods

On the angle-Doppler plane, the SR-STAP method involves the construction of a uniformly discrete dictionary, where each discrete grid represents a clutter point. However, if the spatial and Doppler frequencies of an actual clutter point do not coincide with the uniformly discrete grid points, energy leakage can occur in the neighboring grids. This leakage can further reduce the accuracy of sparse recovery of a clutter.

To solve the off-grid problem, this study adds N_s auxiliary atoms to the original discrete dictionary and constructs an auxiliary dictionary matrix with dimensions of $MN \times N_s$. In cases where some clutter points deviate from the grid, the atoms in the auxiliary dictionary are capable of adaptively converging to the clutter ridge. Within the angle-Doppler plane, clutter points in a particular direction exhibit sparse distribution across the normalized Doppler and spatial frequencies. Consequently, the Doppler and spatial frequencies are consistently located at the positions of the atoms with the maximum power. Therefore, a search-based method is employed to update the Doppler and spatial frequencies iteratively.

The specific steps are as follows:

Step 1: The Doppler frequencies are updated. First, it is assumed that each atom has an invariant spatial frequency $f_{s,k}$ and a variable Doppler frequency $\tilde{f}_{d,k}$. The actual Doppler frequency $\tilde{f}_{d,k}$ is constrained in a local region Ξ_k , where $\Xi_k \in [f_{d,k} - \Delta f, f_{d,k} + \Delta f]$ and $\Delta f = 1/N_d$ represents the quantization interval of the normalized Doppler frequency. The interval is divided Ξ_k into P equal parts, which results in a set of Doppler frequencies denoted by $\{\tilde{f}_{d,k}^p\}_{p=1}^P$; then, the optimal solution can be obtained by

$$\hat{\tilde{f}}_{d,k} = \arg \max_{\tilde{f}_{d,k} \in \{\tilde{f}_{d,k}^p\}_{p=1}^P} \frac{\|\mathbf{v}(f_{s,k}, \tilde{f}_{d,k})^H (\mathbf{R}_x)^{-1} \mathbf{X}\|_2^2}{\mathbf{v}(f_{s,k}, \tilde{f}_{d,k})^H (\mathbf{R}_x)^{-1} \mathbf{v}(f_{s,k}, \tilde{f}_{d,k})}. \quad (11)$$

Step 2: The spatial frequencies are optimized based on the corrected Doppler frequency. First, it is assumed that each atom has a fixed Doppler frequency $\tilde{f}_{d,i}$ and a variable spatial frequency $\tilde{f}_{s,k}$. Similar to the method for updating the Doppler frequencies, the actual spatial frequency $\tilde{f}_{s,i}$ is constrained in a local domain Ω_i , where $\Omega_i \in [f_{s,i} - \Delta f', f_{s,i} + \Delta f']$, where $\Delta f' = 1/N_s$ represents the quantization interval of the normalized spatial frequency. Next, interval Ω_i is divided into O equal parts, resulting in a set of spatial frequencies denoted by $\{\tilde{f}_{s,i}^o\}_{o=1}^O$. The optimal solution is obtained by

$$\tilde{f}_{s,i} = \arg \max_{\tilde{f}_{s,i} \in \{\tilde{f}_{s,i}^0\}_{o=1}^O} \frac{\left\| v(\tilde{f}_{s,i}, \tilde{f}_{d,i})^H \mathbf{R}_x^{-1} \mathbf{X} \right\|_2^2}{v(\tilde{f}_{s,i}, \tilde{f}_{d,i})^H \mathbf{R}_x^{-1} v(\tilde{f}_{s,i}, \tilde{f}_{d,i})} \quad (12)$$

The two search processes are illustrated in Figure 4, where the atoms obtained by fixing the spatial domain channel and searching for the Doppler frequencies are represented by black square patterns. Next, the Doppler frequency corresponding to the position of the black square is fixed, and the spatial frequency is searched. In Figure 4, the actual location of the clutter power is indicated by a black circle.

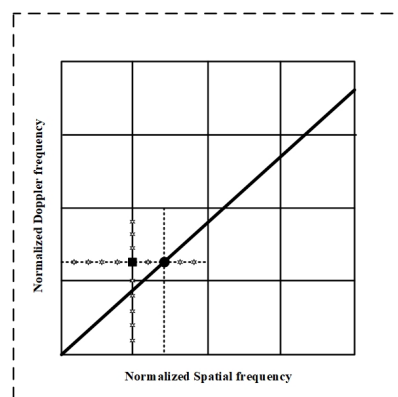


Figure 4. Schematic of the auxiliary dictionary update.

Step 3: The auxiliary dictionary is updated via the revised Doppler and spatial frequencies. Moreover, iterative calibration of the Doppler and spatial frequencies ensures that the atoms in the discrete dictionary are accurately aligned with the clutter distribution range, yielding an improved estimation accuracy of CCM.

3.3. Sparse Bayesian Learning

In the context of the sparse recovery problem, l_1 regularization techniques have been frequently employed in recent research to solve Equation (10). Nevertheless, an SR-STAP filter exhibits a heightened sensitivity to the regularization parameter and noise error tolerance. Generally, the effectiveness of this approach depends on the parameter selection; therefore, under non-ideal conditions, its performance can substantially deteriorate because of parameter blindness.

Using a combination of Bayesian learning methods with sparse priors can facilitate the avoidance of problems associated with regularization parameters when addressing the optimal problem in Equation (10). This approach ensures the robustness of sparse spectra estimation in unknown environments. However, despite the superior clutter suppression capabilities of the existing SBL methods, the previously employed models cannot significantly enhance sparsity. Therefore, a larger number of iterations are required for parameter updating, and the substantial computational burden decreases the application of these methods in actual practice. Considering all the mentioned aspects and aiming to enhance the effect of the SBL-STAP method, this study proposes an innovative SBL algorithm that uses the generalized double Pareto (GDP) prior to clutter estimation [28]. Studies have shown that the sparsity of the solution obtained by sparse Bayesian learning is proportional to the shape of the prior probability density function [28–30]. As illustrated in Figure 5, the shape of the generalized double Pareto prior distribution is sharper than that of the commonly used Laplace and Gaussian priors. Therefore, the GDP prior is more capable of enhancing the sparsity of the solution.

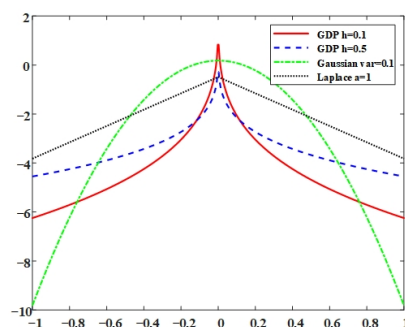


Figure 5. Distribution of different prior probability density functions.

In comparison to the standard Gaussian–gamma setting, applying gamma priors to β enables dynamic adaptation to signals of different sparsity levels while mitigating the excessive penalty associated with large coefficients and maintaining small coefficients close to zero. However, compared with higher-level generalized hyperbolic priors [31], the latter introduces more latent variables and hyperparameters for optimization, making them more flexible but also more complex. Although the prior distribution of GDP discussed in this article is slightly less flexible than the above algorithms, it can balance efficiency and accuracy, making it more suitable for the application scenarios discussed.

First, it is assumed that elements α_l ($l = 1, 2, \dots, L$) of matrix A follow independent complex Gaussian distribution, which is expressed by

$$p(A|\gamma) = \prod_{l=1}^L p(\alpha_l|\gamma) = \prod_{l=1}^L \mathcal{CN}(\alpha_l|\mathbf{0}, \Gamma). \quad (13)$$

The variances of different rows in matrix A are denoted by $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_Q]^T$ and $\Gamma = \text{diag}(\gamma)$.

Next, each hyperparameter γ_q in γ is assumed to follow a Gamma distribution with a parameter β_q , which is defined by

$$p(\gamma|\beta) = \prod_{q=1}^Q \mathcal{G}(\gamma_q; 3/2, \beta_q^2/4), \quad (14)$$

where $\beta = [\beta_1, \beta_2, \dots, \beta_Q]^T$, $\mathcal{G}(x; a, b) = \Gamma^{-1}(a) b^a x^{a-1} e^{-bx}$, and $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$.

At the same time, a hyperparameter β_q in β obeys a Gamma distribution with a parameter h , which is expressed by

$$p(\beta) = \prod_{q=1}^Q \mathcal{G}(\beta_q; h, h) \quad (15)$$

where h is a positive constant which is very small.

The marginal distribution of α_l can be obtained through integration processes which are about γ and β as follows:

$$\begin{aligned} p(\alpha_l|h) &= \prod_{q=1}^Q \int_0^\infty \int_0^\infty p(\alpha_{l,q}|\gamma) p(\gamma_q|\beta_q) p(\beta_q) d\gamma_q d\beta_q \\ &= \prod_{q=1}^Q \frac{1}{2\pi} \frac{(h+1)/h}{(|\alpha_{l,q}|/h+1)^{(h+1)+1}}. \end{aligned} \quad (16)$$

Furthermore, it is assumed that the joint probability density function of observed noise $\mathbf{n}$ is expressed as follows:

$$p(\mathbf{n}|\sigma_n^2) = \prod_{l=1}^L \mathcal{CN}(\mathbf{n}_l | \mathbf{0}_{MN \times 1}, \sigma_n^2 \mathbf{I}_{MN}). \quad (17)$$

After establishing the Bayesian hierarchical prior model, the parameters that need to be solved estimated using the Type II estimation method, which is represented as $\mathcal{L}_{\text{II}}(\gamma) = \arg \max_{\gamma} \int p(\mathbf{Y}|\mathbf{X}, \beta) p(\mathbf{X}|\gamma) p(\gamma) d\mathbf{X}$ [29,32,33]. Owing to the addition of new hyperparameters in the SBL framework in this study, the cost function of the Type II estimation method is formed as follows:

$$\mathcal{L}_{\text{II}}(\gamma, \beta, \sigma_n^2) = \arg \max_{\gamma, \beta, \sigma_n^2} \int p(\mathbf{X}|\mathbf{A}, \sigma_n^2) p(\mathbf{A}|\gamma) p(\gamma|\beta) p(\beta) d\mathbf{A}. \quad (18)$$

The EM algorithm is employed to maximize Equation (18).

In the E-step, the expectation of the log-likelihood function is obtained by

$$Q(\gamma, \beta, \sigma_n^2) = E_{p(\mathbf{A}|\gamma, \beta, \sigma_n^2, \mathbf{X})} \left\{ \ln(p(\mathbf{X}|\mathbf{A}, \sigma_n^2) p(\mathbf{A}|\gamma) p(\gamma|\beta) p(\beta)) \right\}, \quad (19)$$

where $E_p[\cdot]$ is the expectation with respect to p ; $p(\mathbf{A}|\gamma, \beta, \sigma_n^2, \mathbf{X}) \propto \prod_{l=1}^L \mathcal{CN}(\mathbf{a}_l | \boldsymbol{\mu}_l, \Sigma)$, $\boldsymbol{\mu}_l = \boldsymbol{\Gamma} \boldsymbol{\Phi}^H (\Sigma_x)^{-1} \mathbf{x}_l$, $\Sigma = \boldsymbol{\Gamma} - \boldsymbol{\Gamma} \boldsymbol{\Phi}^H (\Sigma_x)^{-1} \boldsymbol{\Phi} \boldsymbol{\Gamma}$, and $\Sigma_x = \sigma_n^2 \mathbf{I}_{MN} + \boldsymbol{\Phi} \boldsymbol{\Gamma} \boldsymbol{\Phi}^H$. The irrelevant terms of γ and β are omitted in Equation (19); then, the likelihood function of γ and β can be derived as follows:

$$Q(\gamma, \beta) = E_{p(\mathbf{A}|\gamma, \beta, \sigma_n^2, \mathbf{X})} \{ \ln p(\mathbf{A}|\gamma) p(\gamma|\beta) p(\beta) \} \\ \propto \sum_{q=1}^Q \left\{ -L \ln \gamma_q - \gamma_q^{-1} (\sum_{l=1}^L |\mu_{l,q}|^2 + L \Sigma_{q,q}) + \ln \sqrt{\gamma_q} + \ln \beta_q^3 - \beta_q^2 \gamma_q / 4 + (h-1) \ln \beta_q - h \beta_q \right\}, \quad (20)$$

where $\Sigma_{q,q}$ is the (q, q) element of a matrix Σ , and $\mu_{l,q}$ is the q th element of $\boldsymbol{\mu}_l$.

During the M-step, the parameters γ_q and β_q are updated by maximizing the likelihood function defined by Equation (20). First, the partial derivatives of the likelihood function with respect to parameters γ_q are calculated and equal to zero, yielding the following expression:

$$\frac{\partial Q(\gamma, \beta)}{\partial \gamma_q} = -\frac{L}{\gamma_q} + \frac{\sum_{l=1}^L |\mu_{l,q}|^2 + L \Sigma_{q,q}}{\gamma_q^2} + \frac{1}{2\gamma_q} - \frac{\beta_q^2}{4} = 0. \quad (21)$$

However, the convergence speed for the update of γ_q will be slow because of solving Equation (21) directly. According to the existing research results [30], fixed-point updates can accelerate the convergence rate. However, because of the specific GDP prior to the SBL framework, this part defines a new fixed-point update rule. To this end, this study defines $W_q = -1 + \frac{\Sigma_{q,q}}{\gamma_q}$ and substitutes it into Equation (21), obtaining the result by

$$\gamma_q = (\beta_q)^{-2} (1 + 2LW_q + 2\sqrt{(1/2 + LW_q)^2 + (\beta_q)^2 \sum_{l=1}^L |\mu_{l,q}|^2}). \quad (22)$$

Similarly, setting $\frac{\partial Q(\gamma, \beta)}{\partial \beta_q} = 0$ leads to the update rule of β_q , which is expressed as follows:

$$\beta_q = (\sqrt{h^2 + 2\gamma_q(h+2)} - h) / \gamma_q. \quad (23)$$

Next, the updating formula of noise power σ_n^2 is derived. First, a likelihood function $Q(\sigma_n^2)$ of a parameter σ_n^2 is solved with respect to the joint posterior distribution of the sparse coefficients, which can be expressed by

$$\begin{aligned} Q(\sigma_n^2) &= E_{p(A|\gamma, \beta, \sigma_n^2, X)} \{ \ln p(X|A, \sigma_n^2) \} \\ &= E_{\sum_{l=1}^L \mathcal{CN}(\alpha_l|\mu_l, \Sigma)} \left\{ \ln \prod_{l=1}^L \mathcal{CN}(x_l|\Phi, \sigma_n^2 I_{MN}) \right\} \\ &= \sum_{l=1}^L E_{\mathcal{CN}(\alpha_l|\mu_l, \Sigma)} \{ \ln \mathcal{CN}(x_l|\Phi, \sigma_n^2 I_{MN}) \}. \end{aligned} \quad (24)$$

Setting $\frac{\partial Q(\sigma_n^2)}{\partial \sigma_n^2} = 0$ provides the updated equation of σ_n^2 , which is expressed as follows:

$$\sigma_n^2 = \left(\sum_{l=1}^L \|x_l - \Phi \mu_l\|_2^2 + L \text{tr}((\Phi)^H \Phi \Sigma) \right) / MNL. \quad (25)$$

3.4. STAP Filter Weight Calculation

After estimating the CCM, the weight of an adaptive filter can be computed by

$$w = \frac{R^{-1}v(f_{d,t}, f_{s,t})}{v(f_{d,t}, f_{s,t})^H R^{-1}v(f_{d,t}, f_{s,t})}. \quad (26)$$

The main calculation equations and steps of the proposed RD-SBL-STAP algorithm are presented in Algorithm 1.

Algorithm 1 Proposed RD-SBL-STAP algorithm.

Initialization:

$$D^{(0)} = 0_{NM \times N_s}, \Sigma_x^{(0)} = I_{NM}, \beta^{(0)} = 1_{Q \times 1}, (\sigma_n^2)^{(0)} = 1, \gamma_q^{(0)} =$$

$$\frac{1}{L} \sum_{l=1}^L \left| \frac{v(f_{s,i}, f_{d,k})^H x_l}{v(f_{s,i}, f_{d,k})^H v(f_{s,i}, f_{d,k})} \right|^2, q = 1, 2, \dots$$

Dictionary descent definition:

$$\bar{P} = \frac{\sum_{i=1}^{N_s} \sum_{k=1}^{N_d} P(f_{s,i}, f_{d,k})}{N_s N_d}, w_p = \begin{cases} v(f_{s,i}, f_{d,k}), & P(f_{s,i}, f_{d,k}) > \bar{P} \\ \text{pass}, & P(f_{s,i}, f_{d,k}) \leq \bar{P} \end{cases}, \Psi = [w_1, w_2, \dots]$$

Auxiliary dictionary update:

$$\begin{aligned} &\text{for } k = 1 : N_d \\ &\quad \Xi_k = [f_{d,k} - \Delta f, f_{d,k} + \Delta f] \\ &\quad \hat{f}_{d,k} = \underset{\tilde{f}_{d,k} \in \{\tilde{f}_{d,k}^p\}_{p=1}^P}{\text{argmax}} \frac{\|v(f_{s,k}, \tilde{f}_{d,k})^H R_x^{-1} X\|_2^2}{v(f_{s,k}, \tilde{f}_{d,k})^H R_x^{-1} v(f_{s,k}, \tilde{f}_{d,k})} \\ &\quad \text{end} \\ &\text{for } i = 1 : N_s \\ &\quad \Omega_i = [f_{s,i} - \Delta f', f_{s,i} + \Delta f'] \\ &\quad \hat{f}_{s,i} = \underset{\tilde{f}_{s,i} \in \{\tilde{f}_{s,i}^o\}_{o=1}^O}{\text{argmax}} \frac{\|v(\tilde{f}_{s,i}, \hat{f}_{d,i})^H R_x^{-1} X\|_2^2}{v(\tilde{f}_{s,i}, \hat{f}_{d,i})^H R_x^{-1} v(\tilde{f}_{s,i}, \hat{f}_{d,i})} \\ &\quad v_i = v(\hat{f}_{s,i}, \hat{f}_{d,i}) \\ &\quad \text{end} \\ &\quad D = [v_1, v_2, \dots, v_{N_s}], \Phi = [\Psi, D] \end{aligned}$$

Algorithm 1 *Cont.*

Repeat the following iterations until the condition given in the last line is met:

Update the data covariance matrix by Σ_x : $\Sigma_x^{(k+1)} = (\sigma_n^2)^{(k)} \mathbf{I}_{MN} + \Phi^{(k)} \Gamma^{(k)} (\Phi^{(k)})^H$

Update the mean of the posterior distribution of sparse coefficients

μ_l by $\mu_l^{(k+1)} = \Gamma^{(k)} (\Phi^{(k)})^H (\Sigma_x^{(k+1)})^{-1} \mathbf{x}_l, l = 1, 2, \dots, L$

Update the covariance matrix of the posterior distribution of sparse coefficients Σ by

$\Sigma^{(k+1)} = \Gamma^{(k)} - \Gamma^{(k)} (\Phi^{(k)})^H (\Sigma_x^{(k+1)})^{-1} \Phi^{(k)} \Gamma^{(k)}$

Update the hyperparameters γ_q by

$\gamma_q^{(k+1)} = (\beta_q^{(k)})^{-2} (1 + 2LW_q^{(k+1)}) + 2\sqrt{(1/2 + LW_q^{(k+1)})^2 + (\beta_q^{(k)})^2 \sum_{l=1}^L |\mu_{l,q}^{(k+1)}|^2}$, where

$W_q^{(k+1)} = -1 + \frac{\Sigma_{q,q}^{(k+1)}}{\gamma_q^{(k)}}$

Update the hyperparameters β_q by $\beta_q^{(k+1)} = (\sqrt{h^2 + 2\gamma_q^{(k+1)}}(h+2) - h)/\gamma_q^{(k+1)}$

Update the noise power σ_n^2 by

$(\sigma_n^2)^{(k+1)} = (\sum_{l=1}^L \|\mathbf{x}_l - \Phi^{(k)} \mu_l^{(k+1)}\|_2^2 + L \text{tr}((\Phi^{(k)})^H \Phi^{(k)} \Sigma_{q,q}^{(k+1)})) / MNL$

Iteration termination condition: $\left| \frac{(\sigma_n^2)^{(k+1)} - (\sigma_n^2)^{(k)}}{(\sigma_n^2)^{(k)}} \right| \leq \zeta$

4. Analysis of Simulation Results

In this study, the effect of the proposed RD-SBL-STAP algorithm was verified by simulation tests. At the same time, the new algorithm was also compared with the SR-STAP algorithm [34], the RSBL-STAP algorithm [24], and the SBL-STAP algorithm [35]. For all algorithms, the number of training snapshots was set to 20, and the grid resolution was set to $\mu_d = 6$ and $\mu_s = 6$. The iteration-stopping threshold was set to 10^{-12} . The relevant simulation parameters are provided in Table 1.

Table 1. T Radar system parameters and their values.

Parameter	Value
Antenna array type	Uniform linear array for side-looking
Number of array elements	10
Number of pulses sent in one CPI	10
Array element spacing (m)	0.15
Wavelength (m)	0.3
Aircraft speed (m/s)	240
Aircraft altitude (m)	3000
Pulse repetition frequency (Hz)	4000

4.1. Clutter Power Spectrum Analysis

The initial analysis focused on the estimation results of the clutter power spectrum without off-grid across four STAP algorithms which are SR-STAP, SBL-STAP, RSBL-STAP, and RD-SBL-STAP, as illustrated in Figure 6. As shown in Figure 6a, the clutter spectrum generated by the SR-STAP algorithm was entirely concentrated on the clutter ridge. However, the clutter energy which is estimated was relatively low, and this was unfavorable for subsequent clutter suppression. In addition, as displayed in Figure 6b, because of the off-grid effect, the SBL-STAP algorithm showed a severe broadening of the clutter spectrum, and low-speed targets might be annihilated by the broadened clutter. Furthermore, as depicted in Figure 6c, the clutter power spectrum generated by the RSBL-STAP algorithm was concentrated on the clutter ridge, but significant broadening was still present. The

results in Figure 6d indicate that the RD-SBL-STAP algorithm achieves precise spectral energy concentration along theoretical clutter contours. Compared with the SBL-STAP and RSBL-STAP algorithms, the phenomenon concerning the broadening of the clutter ridge of the proposed RD-SBL-STAP algorithm was significantly reduced, and the clutter power spectrum value was much higher, which was beneficial for subsequent clutter suppression.

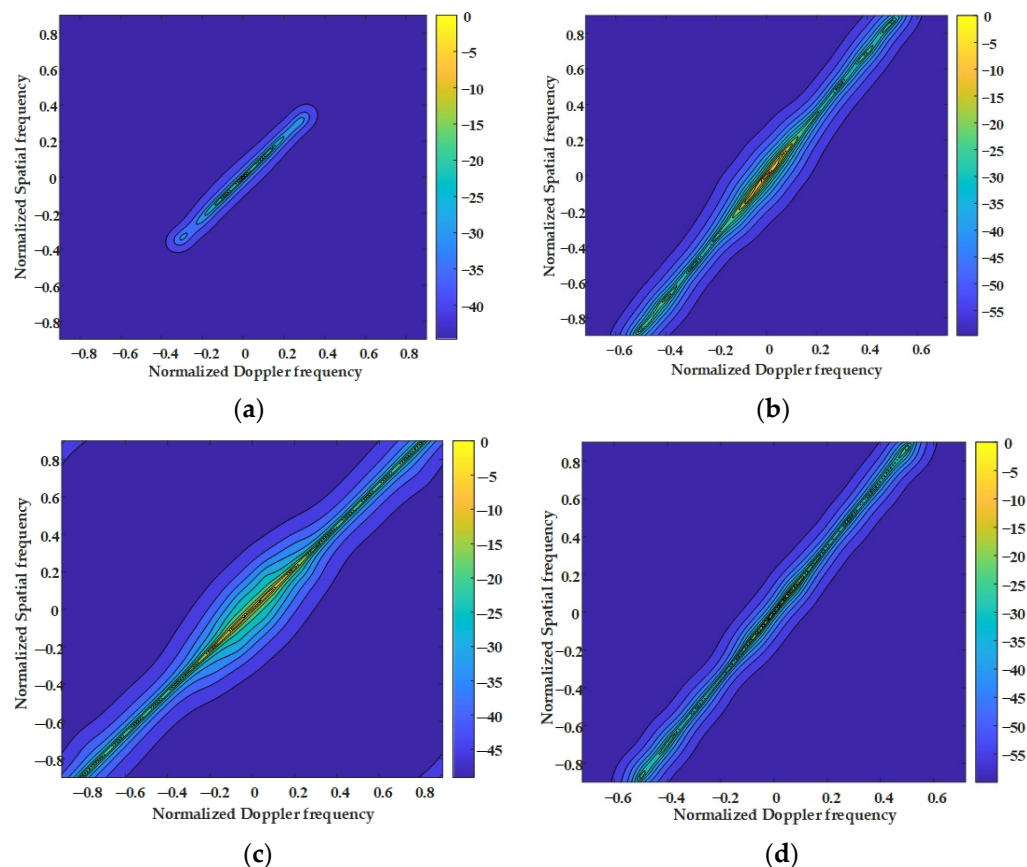


Figure 6. Clutter power spectra estimation results without off-grid: (a) clutter power spectrum obtained using SR-STAP algorithm; (b) clutter power spectrum obtained using SBL-STAP algorithm; (c) clutter power spectrum obtained using RSBL-STAP algorithm; (d) clutter power spectrum obtained using RD-SBL-STAP algorithm.

The second analysis focused on the estimation results of the clutter power spectrum on the occasion of off-grid across the STAP algorithms, which are SR-STAP, SBL-STAP, RSBL-STAP, and RD-SBL-STAP, as illustrated in Figure 7. Compared with the experimental results shown in Figure 6, the clutter power spectra of the SR-STAP and SBL-STAP algorithms, as shown in Figure 7, demonstrate significant broadening. The RSBL-STAP algorithm has undergone grid correction, and compared with the experiment without grid mismatch, its clutter ridge broadening has improved slightly. The RD-SBL-STAP clutter power spectrum shows slight broadening, with almost no significant difference compared to the result without off-grid. Therefore, the proposed algorithm is robust and insensitive to grid mismatch.

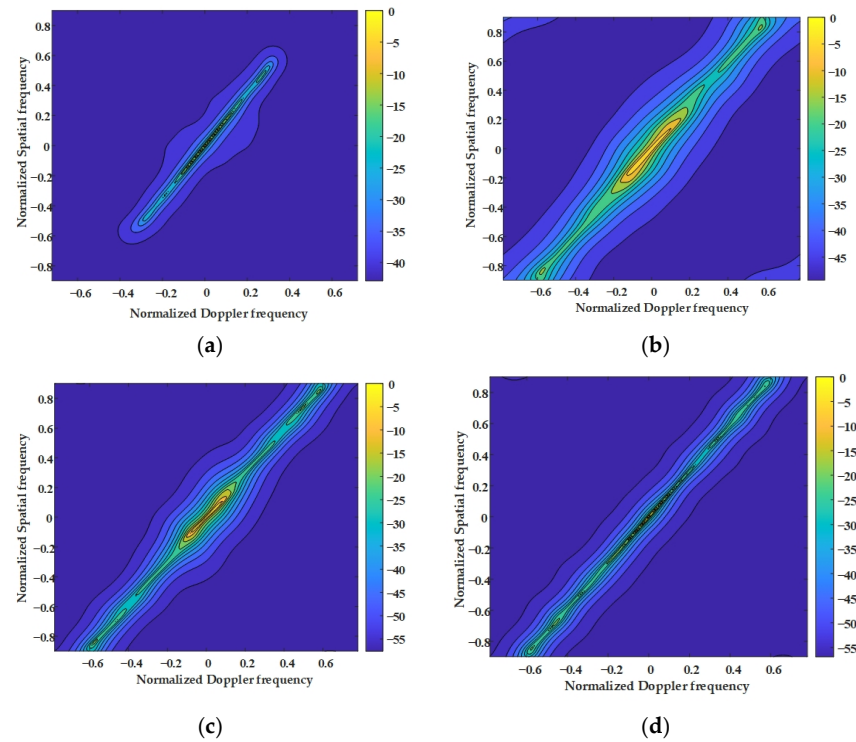


Figure 7. Clutter power spectra estimation results within off-grid: (a) clutter power spectrum obtained using SR-STAP algorithm; (b) clutter power spectrum obtained using SBL-STAP algorithm; (c) clutter power spectrum obtained using RSBL-STAP algorithm; (d) clutter power spectrum obtained using RD-SBL-STAP algorithm.

4.2. Output Signal-To-Noise Ratio Loss Analysis

The second experiment compared the performance of the SR-STAP, SBL-STAP, RSBL-STAP, and proposed RD-SBL-STAP algorithms by analyzing the output signal-to-clutter-plus-noise-ratio loss (SCNRloss), as displayed in Figure 8. The comparison consequences showed that the SR-STAP algorithm had a narrow notch but with a shallow depth. The SBL-STAP algorithm demonstrated a deeper notch, but the notch width was broad. The RSBL-STAP algorithm outperformed both the SBL-STAP and RD-SBL-STAP algorithms in the main lobe clutter region; however, its output SCNRloss value in the sidelobe clutter region was slightly inferior to that of the new algorithm. The proposed RD-SBL-STAP algorithm achieved a narrow notch with a significant depth, which indicated its superior clutter suppression performance.

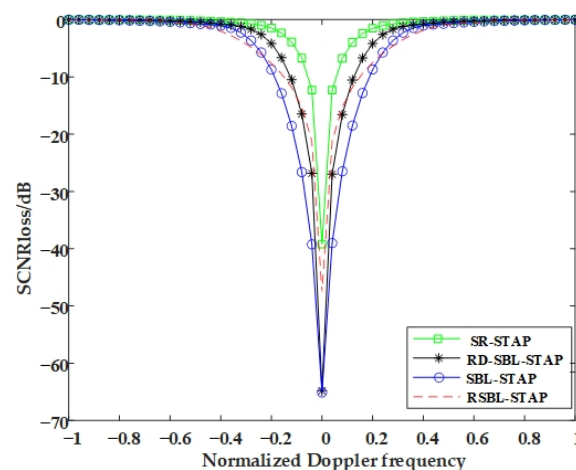


Figure 8. Signal-to-clutter-plus-noise ratio loss results of different algorithms.

4.3. Target Detection Performance Analysis

The third experiment focused on evaluating moving target detection capabilities across four algorithms: SR-STAP, SBL-STAP, RSBL-STAP, and the proposed RD-SBL-STAP framework. In this experiment, each algorithm processed 100-range-cell snapshots with the targets located at cell 151, as depicted in Figure 9, where each technique could successfully detect the location of the target. However, the residual clutter output power of the proposed RD-SBL-STAP algorithm was significantly lower than that of the other algorithms. Consequently, the fidelity of moving target detection was demonstrated to be superior to the three comparison algorithms.

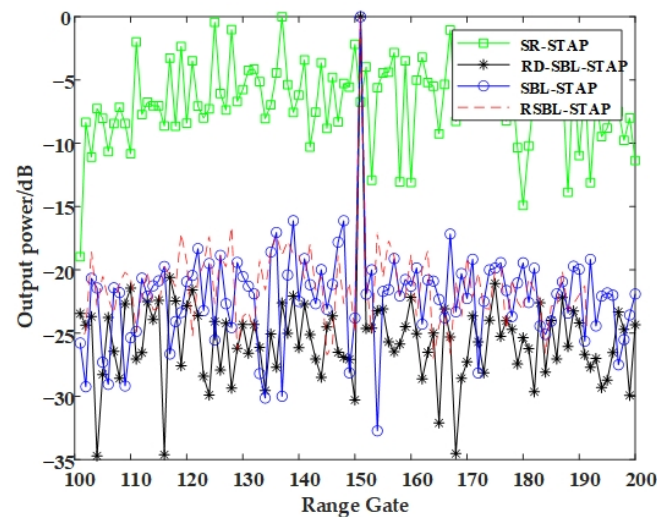


Figure 9. Output power results of the four algorithms.

4.4. Computational Complexity Analysis

This section analyzes the computational complexity of the SR-STAP, SBL-STAP, RSBL-STAP, and RD-SBL-STAP algorithms. The average computational complexity results for a single snapshot are shown in Table 2, where $J = MN$, $\tilde{J}_s = N_s N_d$, $Q = N_s(N_d + 1) = \tilde{J}_s + N_s$, and $I_{\max}[\cdot]$ denotes the number of complex multiplications necessary for a single iteration in the worst case. By substituting specific values into the data table, it can be calculated that the SR-STAP algorithm requires 54,809,600 multiplications per iteration, the SBL-STAP-OS algorithm requires 74,312,121 multiplications per iteration, the RSBL-STAP algorithm requires 700,649,200 multiplications per iteration and the RD-SBL-STAP algorithm requires 1,133,060 multiplications per iteration.

Table 2. Comparison results of average complexity of single samples.

Algorithm	Complex Multiplier
SR-STAP	$(2J^2 + 5J + 4)\tilde{J} + J^3$
SBL-STAP-OS	$(3J^2 + 4JQ + 2JL + Q^2 + 4)Q + (2J^2 + J + LJ + 1)N_s P$
RSBL-STAP	$I_{\max}\{J^2(5Q + LQN + 2J) + (2Q + 2QL + Q^2 + LN^2)J + (L + Q)Q\} + 2N_s P(J^2 + J)$
RD-SBL-STAP	$I_{\max}\{(3J^2 + 5JQ + 2JL + 2Q^2 + 7)Q\} + 2(2J^2 + J + LJ + 1)N_s P$

The RSBL-STAP algorithm required the diagonalization of each steering vector in the dictionary due to the introduction of the amplitude-phase error mismatch vectors, which significantly increased the dimensionality of the dictionary. Although dimensionality reduction was performed subsequently, the process of solving the grid mismatch parameters through multiple iterations still consumed substantial computational time. The

SBL-STAP algorithm, which employed the GDP prior with the aim to accelerate the fitting speed and improve the results, achieved experimental results with fewer iterations than the other algorithms. However, the high-dimensional dictionary matrix still contributed to the increase in computational time. In contrast to the SBL-STAP algorithm, the proposed RD-SBL-STAP algorithm incorporated a dictionary dimensionality-reduction component. Therefore, at the same number of iterations, the proposed algorithm required a lower dictionary dimensionality than the SBL-STAP algorithm, which resulted in faster computational speed. Furthermore, unlike the RSBL-STAP algorithm, the proposed method eliminated the need for steering vector diagonalization and did not require multiple iterations for updating the grid mismatch parameters. Consequently, among all the studied algorithms, the proposed RD-SBL-STAP algorithm exhibited the lowest computational complexity.

5. Conclusions

This study has proposed a self-calibrating STAP algorithm for dictionary dimensionality reduction based on SBL. First, the dimensionality of the dictionary matrix is reduced by selecting the steering vectors corresponding to the atoms with power values higher than the average value. Then, the optimal Doppler and spatial frequencies are searched in each Doppler- and space-domain channel to obtain the steering vectors corresponding to the atoms on the clutter ridge to update the auxiliary dictionaries. Finally, the updated dictionary is used to combine with the SBL algorithm to obtain the required parameters iteratively, and the exact CCM and STAP weight vectors are calculated using the solved parameters. Finally, the empirical validation through numerical simulations confirmed that the developed methodology substantially enhances computational efficiency while improving the operational effectiveness of space–time adaptive processing under non-ideal grid alignment conditions. However, this paper only considers the off-grid phenomenon and does not analyze the mismatch phenomenon caused by the array element radiation phase error. Therefore, further research is needed in the optimization of the algorithm model.

Author Contributions: Conceptualization, Z.G. and N.Y.; methodology, N.Y.; software, N.Y.; validation, Z.G. and N.Y.; formal analysis, Z.G.; investigation, Z.G., N.Y. and P.H.; resources, Z.G. and N.Y.; data curation, Z.G. and N.Y.; writing—original draft preparation, N.Y.; writing—review and editing, Z.G. and P.H.; visualization, N.Y.; supervision, W.X., W.T. and Z.W.; project administration, Z.G. and P.H.; funding acquisition, Z.G. and P.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under grant number 61761037, the Inner Mongolia Science and Technology Plan Project under grant number 2025SKYPT0044, and the Basic Scientific Research Business Cost Project of Colleges directly under the Inner Mongolia under grant number ZTY2025034.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brennan, L.E.; Reed, L. Theory of adaptive radar. *IEEE Trans. Aerosp. Electron. Syst.* **1973**, *AES-9*, 237–252. [[CrossRef](#)]
2. Duan, K.; Xu, H.; Yuan, H.; Xie, H.; Wang, Y. Reduced-DOF three-dimensional STAP via subarray synthesis for nonsidelooking planar array airborne radar. *IEEE Trans. Aerosp. Electron. Syst.* **2019**, *56*, 3311–3325. [[CrossRef](#)]
3. Guerci, J.R. *Space-Time Adaptive Processing for Radar*; Artech House: Norwood, MA, USA, 2014.
4. Klemm, R. *Principles of Space-Time Adaptive Processing*; IET: London, UK, 2002.
5. Goldstein, J.S.; Reed, I.S. Reduced-rank adaptive filtering. *Trans. Signal Process.* **1997**, *45*, 492–496. [[CrossRef](#)]

6. Goldstein, J.S.; Reed, I.S.; Zulch, P.A. Multistage partially adaptive STAP CFAR detection algorithm. *IEEE Trans. Aerosp. Electron. Syst.* **1999**, *35*, 645–661. [\[CrossRef\]](#)
7. Tong, Y.; Wang, T.; Wu, J. Improving EFA-STAP performance using persymmetric covariance matrix estimation. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 924–936. [\[CrossRef\]](#)
8. Wei, Z.; Zishu, H.; Huiyong, L.; Jun, L.; Xiang, D. Beam-space reduced-dimension space-time adaptive processing for airborne radar in sample starved heterogeneous environments. *IET Radar Sonar Navig.* **2016**, *10*, 1627–1634. [\[CrossRef\]](#)
9. Cristallini, D.; Burger, W. A robust direct data domain approach for STAP. *IEEE Trans. Signal Process.* **2011**, *60*, 1283–1294. [\[CrossRef\]](#)
10. Sarkar, T.K.; Wang, H.; Park, S.; Adve, R.; Koh, J.; Kim, K.; Zhang, Y.; Wicks, M.C.; Brown, R.D. A deterministic least-squares approach to space-time adaptive processing (STAP). *IEEE Trans. Antennas Propag.* **2001**, *49*, 91–103. [\[CrossRef\]](#)
11. Melvin, W.L.; Showman, G.A. An approach to knowledge-aided covariance estimation. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 1021–1042. [\[CrossRef\]](#)
12. Stoica, P.; Li, J.; Zhu, X.; Guerci, J.R. On using a priori knowledge in space-time adaptive processing. *IEEE Trans. Signal Process.* **2008**, *56*, 2598–2602. [\[CrossRef\]](#)
13. Mallat, S.G.; Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **1993**, *41*, 3397–3415. [\[CrossRef\]](#)
14. Pati, Y.C.; Rezaifar, R.; Krishnaprasad, P.S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 1–3 November 1993; pp. 40–44.
15. Candes, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [\[CrossRef\]](#)
16. Chartrand, R.; Yin, W. Iteratively reweighted algorithms for compressive sensing. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 3869–3872.
17. Zhang, Y.X.; Jin, Y.X.; Chen, S.J.; Wu, Y.Q.; Hao, C.P. A Robust STAP Algorithm Using Two Prior Knowledge. *J. Signal Process.* **2022**, *38*, 1367–1379.
18. Zhang, S.; Wang, T.; Liu, C.; Wang, D. A space-time adaptive processing method based on sparse Bayesian learning for maneuvering airborne radar. *Sensors* **2022**, *22*, 5479. [\[CrossRef\]](#)
19. Yuan, H.; Xu, H.; Duan, K.; Xie, W.; Liu, W.; Wang, Y. Sparse Bayesian Learning-Based Space-Time Adaptive Processing With Off-Grid Self-Calibration for Airborne Radar. *IEEE Access* **2018**, *6*, 47296–47307. [\[CrossRef\]](#)
20. Sun, K.; Meng, H.; Wang, Y.; Wang, X. Direct data domain STAP using sparse representation of clutter spectrum. *Signal Process.* **2011**, *91*, 2222–2236. [\[CrossRef\]](#)
21. Sun, K.; Zhang, H.; Li, G.; Meng, H.; Wang, X. A novel STAP algorithm using sparse recovery technique. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; pp. V-336–V-339.
22. Duan, K.; Wang, Z.; Xie, W.; Chen, H.; Wang, Y. Sparsity-based STAP algorithm with multiple measurement vectors via sparse Bayesian learning strategy for airborne radar. *IET Signal Process.* **2017**, *11*, 544–553. [\[CrossRef\]](#)
23. Li, Z.; Zhang, Y.; He, X.; Guo, Y. Low-Complexity Off-Grid STAP Algorithm Based on Local Search Clutter Subspace Estimation. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1862–1866. [\[CrossRef\]](#)
24. Li, Z.Y.; Wang, T. Sparse Bayesian learning-based robust STAP algorithm. *Syst. Eng. Electron.* **2023**, *45*, 3032–3040.
25. Xu, L.; Cheng, L.; Wong, N.; Wu, Y.-C.; Poor, H.V. Overcoming Beam Squint in Dual-Wideband mmWave MIMO Channel Estimation: A Bayesian Multi-Band Sparsity Approach. *arXiv* **2023**, arXiv:2306.11149.
26. Reed, I.S.; Mallett, J.D.; Brennan, L.E. Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *AES-10*, 853–863. [\[CrossRef\]](#)
27. Yan, X. Research on Sparse STAP Method for Clutter and Interference Suppression of Airborne Radar. Master's Thesis, Xidian University, Xi'an, China, 2022.
28. Wang, Q.; Yu, H.; Li, J.; Ji, F.; Chen, F. Sparse Bayesian Learning Using Generalized Double Pareto Prior for DOA Estimation. *IEEE Signal Process. Lett.* **2021**, *28*, 1744–1748. [\[CrossRef\]](#)
29. Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
30. 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. [\[CrossRef\]](#)
31. Cheng, L.; Chen, Z.; Shi, Q.; Wu, Y.-C.; Theodoridis, S. Towards flexible sparsity-aware modeling: Automatic tensor rank learning using the generalized hyperbolic prior. *IEEE Trans. Signal Process.* **2022**, *70*, 1834–1849. [\[CrossRef\]](#)
32. Giri, R.; Rao, B. Type I and type II Bayesian methods for sparse signal recovery using scale mixtures. *IEEE Trans. Signal Process.* **2016**, *64*, 3418–3428. [\[CrossRef\]](#)
33. Wipf, D.P.; Rao, B.D.; Nagarajan, S. Latent variable Bayesian models for promoting sparsity. *IEEE Trans. Inf. Theory* **2011**, *57*, 6236–6255. [\[CrossRef\]](#)

34. Yang, Z.; Li, X.; Wang, H.; Jiang, W. On clutter sparsity analysis in space–time adaptive processing airborne radar. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1214–1218. [[CrossRef](#)]
35. Sun, Y.; Yang, X.; Long, T.; Sarkar, T.K. Robust sparse Bayesian learning STAP method for discrete interference suppression in nonhomogeneous clutter. In Proceedings of the 2017 IEEE Radar Conference (RadarConf), Seattle, WA, USA, 8–12 May 2017; pp. 1003–1008.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.