



Article

Beam-Space Post-Doppler Reduced-Dimension STAP Based on Sparse Bayesian Learning

Junxiang Cao , Tong Wang * and Degen Wang

National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China; 22021110329@stu.xidian.edu.cn (J.C.); wangdg@stu.xidian.edu.cn (D.W.)

* Correspondence: twang@mail.xidian.edu.cn

Abstract: The space–time adaptive processing (STAP) technique can effectively suppress the ground clutter faced by the airborne radar during its downward-looking operation and thus can significantly improve the detection performance of moving targets. However, the optimal STAP requires a large number of independent identically distributed (i.i.d) samples to accurately estimate the clutter plus noise covariance matrix (CNCM), which limits its application in practice. In this paper, we fully consider the heterogeneity of clutter in real-world environments and propose a sparse Bayesian learning-based reduced-dimension STAP method that achieves suboptimal clutter suppression performance using only a single sample. First, the sparse Bayesian learning (SBL) algorithm is used to estimate the CNCM using a single training sample. Second, a novel angular Doppler channel selection algorithm is proposed with the criterion of maximizing the output signal-to-clutter-noise ratio (SCNR). Finally, the reduced-dimension STAP filter is constructed using the selected channels. Simulation results show that the proposed algorithm can achieve suboptimal clutter suppression performance in extremely heterogeneous clutter environments where only one training sample can be used.

Keywords: space–time adaptive processing; sparse Bayesian learning; reduced dimension; angular Doppler channel



Citation: Cao, J.; Wang, T.; Wang, D. Beam-Space Post-Doppler Reduced-Dimension STAP Based on Sparse Bayesian Learning. *Remote Sens.* **2024**, *16*, 307. <https://doi.org/10.3390/rs16020307>

Academic Editor: David Gomez-Ortiz

Received: 27 November 2023

Revised: 3 January 2024

Accepted: 4 January 2024

Published: 11 January 2024



Copyright: © 2024 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

Airborne radar is highly valued by many countries in the world because of its strong maneuverability and longer direct viewing distance than ground radar. However, due to the influence of ground clutter, the moving target detection performance of airborne radar will be seriously degraded when looking down. The ground clutter is not only strong but also has different speeds relative to the aircraft in different directions, which greatly broadens the clutter spectrum. How to suppress ground clutter is the key problem to be solved in airborne radar systems. Space–time adaptive processing (STAP) [1] technology has been widely used by many researchers because of its excellent ground clutter suppression performance. Traditional STAP algorithms need to use target-free samples adjacent to the cell under test (CUT) to obtain an estimate of the clutter plus noise covariance matrix (CNCM). According to the Reed–Mallett–Brennan (RMB) criterion [2], the output signal-to-clutter-noise ratio (SINR) loss after processing by the STAP algorithm is less than 3 dB only if the number of training samples used is more than two times the number of degrees of freedom (DOFs) of the system. Unfortunately, it is difficult in reality to provide such a large number of i.i.d samples for airborne radar since ground clutter is often heterogeneous due to external non-ideal factors. Therefore, the study of STAP algorithms with few samples and high performance is very important for practical applications.

In the decades since the STAP algorithm was first proposed, numerous researchers have proposed many methods that can reduce the sample requirements of the STAP to some extent. Klemm proposed an auxiliary channel method (ACP) [3], which can effectively

reduce the number of i.i.d samples required for STAP. Under ideal conditions, ACP can achieve near-optimal performance, but when non-ideal factors (such as internal clutter motion and array amplitude and phase errors) are present, the performance loss can be significant. Dipietro proposed the extended factor approach (EFA) [4], which uses all spatial domain channels for adaptive processing after Doppler filtering. As a result, EFA is robust to non-ideal factors in the spatial domain. However, when the airborne radar system has a large number of spatial DOFs, EFA still requires many training samples. Wang proposed a joint domain localized processing algorithm (JDL) [5] for space–time adaptive processing by selecting multiple channels adjacent to the channels to be detected in the angular Doppler plane. Compared with EFA, this method has the advantage of further reducing the training sample and lowering the computational complexity and the disadvantage of being more affected by non-ideal factors. In the study of STAP techniques based on spatial–temporal spectral sparsity, the literature [6–12] utilizes several training samples without target signals to reconstruct the spatial–temporal amplitude spectrum of the clutter and estimate the CNCM. This type of method has the advantage of estimating the clutter covariance matrix relatively accurately using only 4–6 i.i.d training samples, but it also suffers from high computational complexity and inevitable off-grid problems.

In this paper, an RD STAP method based on the sparse Bayesian algorithm is proposed, which has the advantages of requiring fewer samples and suitable clutter suppression performance. In the traditional RD STAP algorithm, we found the following phenomenon, i.e., when the number of samples is insufficient, the output SINR obtained using the RD matrix to process the estimated CNCM followed by adaptive processing is larger than that obtained by directly using the estimated CNCM for adaptive processing. Inspired by this, we consider estimating an inaccurate CNCM using only a single sample in the extremely heterogeneous clutter environment and then selecting the optimal RD channel to make the output SINR satisfy the target detection requirements. In our proposed algorithm, the CNCM is first estimated using the SBL algorithm when only a single sample is available. Although this estimated CNCM cannot be used directly for adaptive processing, it can be used to construct RD transformation matrices. Second, the proposed channel selection algorithm is utilized to select the optimal RD channel in the angular Doppler plane with the criterion of maximizing the output SINR. Finally, the selected channels and estimated CNCM are used to construct the RD STAP filter. The main contributions of this paper are as follows:

1. A novel RD STAP method based on the SBL algorithm is proposed, which has suboptimal clutter suppression performance in extremely heterogeneous clutter environments with only one training sample available.
2. A novel angular Doppler domain RD channel selection algorithm is proposed, which maximizes the output SINR as a criterion for selecting auxiliary channels. In general, suboptimal clutter suppression performance can be achieved by selecting 3–8 auxiliary channels using the proposed algorithm.

In order to more clearly show the superiority of the proposed algorithm in sample demand, Table 1 lists the minimum number of samples required by several related algorithms to achieve suboptimal performance. In Table 1, M refers to the number of pulses and N refers to the number of elements.

Table 1. Training samples required by several related algorithms.

Algorithm	Required Samples
EFA [4]	6N
JDL(3*3) [5]	18
SBL-STAP [9]	6
Proposed method	1

The remainder of this paper is organized according to the structure below. In Section 2, we briefly review the STAP model and processors based on the GSC architecture. In Section 3, we derive the processing flow of the proposed algorithm in detail. In Section 4, we perform extensive simulation experiments to illustrate the effectiveness of the proposed algorithm. Finally, in Section 5, we summarize some useful conclusions.

Notation: In this paper, the symbol $\otimes$ denotes the Kronecker product. The symbol $\|\cdot\|_F$ and $\|\cdot\|_0$ are used to represent the Frobenius (ℓ_F) norms and ℓ_0 pseudo-norms, respectively. Scalars are italicized, and lowercase bold and uppercase bold denote vectors and matrices, respectively. The operators $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate transpose, respectively. The operation of finding the expectation of a random variable v is denoted by $E(v)$. The symbol $A \in B$ presents set A belongs to set B . The empty set is represented by the symbol $\emptyset$.

2. STAP Model and GSC form Processor

2.1. STAP Model

The system considered is a pulse Doppler radar installed on an airborne platform. The antenna of this radar is a uniform line array (ULA) containing N array elements with a spacing d equal to half of the operating wavelength λ of the system. The airborne platform is flying at an altitude of h and a speed of v . The radar transmits a coherent burst of pulses at a fixed pulse repetition frequency (PRF) $f_r = 1/T_r$, where T_r refers to the pulse repetition time (PRT). A total of M pulses are transmitted during the coherent processing interval (CPI), so the length of the coherent processing time is MT_r . The geometric model of the airborne radar is shown in Figure 1a. Each PRT needs to be sampled L times to cover the distance interval, and LMN complex baseband samples are obtained after matched filtering the returns from each pulse within a CPI, which is referred to as radar datacube, shown in Figure 1b.

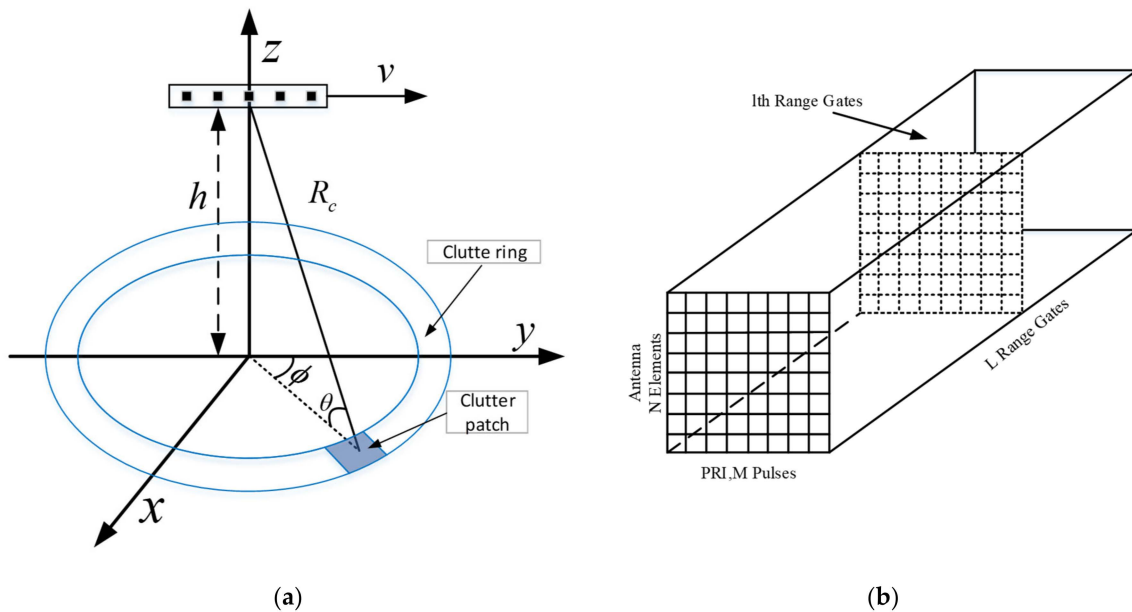


Figure 1. Airborne radar geometric configuration and datacube. (a) Airborne radar downward-looking working model; (b) the radar datacube.

According to the clutter model proposed by Ward [1], the normalized spatial and normalized Doppler frequency of the i -th clutter patch can be expressed as

$$f_{d,i} = \frac{2v}{\lambda f_r} \cos(\theta_i) \cos(\varphi_i) \quad (1)$$

$$f_{s,i} = \frac{d}{\lambda} \cos(\theta_i) \cos(\varphi_i) \quad (2)$$

where θ_i and φ_i are the elevation and azimuth angle, respectively. Then, the space–time steering vector can be expressed as

$$s(f_{d,i}, f_{s,i}) = a(f_{d,i}) \otimes b(f_{s,i}) \quad (3)$$

$$a(f_{d,i}) = [1 \exp(j2\pi f_{d,i}) \cdots \exp(j2\pi(M-1)f_{d,i})]^T \quad (4)$$

$$b(f_{s,i}) = [1 \exp(j2\pi f_{s,i}) \cdots \exp(j2\pi(N-1)f_{s,i})]^T \quad (5)$$

where $a(f_{d,i})$ and $b(f_{s,i})$ are the time and spatial steering vectors, respectively. Considering the range ambiguity, the space–time snapshot can be expressed as [1]

$$x_l = x_c + x_n = \sum_{a=1}^{N_a} \sum_{i=1}^{N_c} \alpha_{a,i} s(f_{d,a,i}, f_{s,a,i}) + x_n \quad (6)$$

where $\alpha_{a,i}$ represents the random complex amplitude of the i -th clutter patch of a -th ambiguous range. x_n is modeled as zero mean Gaussian white noise. Assuming that the clutter patches are independent of each other, the ideal CNCM can be calculated as

$$R_{c+n} = \Phi \Gamma \Phi^H + \sigma^2 I_{MN} = R_c + \sigma^2 I_{MN} \quad (7)$$

where R_c represents the clutter covariance matrix, σ^2 represents the noise power, $\Phi \in C^{MN \times N_c}$ denotes the matrix consisting of the space–time steering vectors of each clutter patch, $\Gamma \in C^{N_c \times N_c}$ is a diagonal matrix with the main diagonal elements being the power of each clutter patch. Assuming that v_t is the space–time steering vector of the target, the optimal STAP filtering weight can be obtained by solving the following optimization problem [13]:

$$\begin{aligned} \min \quad & W_{opt}^H R_{c+n} W_{opt} \\ \text{s.t.} \quad & W_{opt}^H v_t = 1 \end{aligned} \quad (8)$$

and the optimal weight is expressed as

$$W_{opt} = \frac{R_{c+n}^{-1} v_t}{v_t^H R_{c+n}^{-1} v_t} \quad (9)$$

In the traditional SR STAP algorithm, we discretized the normalized space–time plane uniformly to $K = N_s N_d$ points, where $N_s \gg N$ is the number of spatial channels and $N_d \gg M$ is the number of Doppler channels. Then, the SR signal model with single measurement vectors (SMV) is expressed as [14]

$$x = Dw + \varepsilon \quad (10)$$

where x represents the measurement vector and D is the space–time overcomplete dictionary obtained by discretizing the space–time plane. ε represents the additive zero mean Gaussian white noise vector. w is a sparse coefficient vector, which is also the parameter to be solved. The sparse coefficient solution problem in Equation (10) can be approximated as the following convex optimization problem [15]:

$$\begin{aligned} \min_w \quad & \|w\|_1 \\ \text{s.t.} \quad & \|x - Dw\|_2 \leq \rho \end{aligned} \quad (11)$$

where ρ denotes the fitting error tolerance.

2.2. GSC form Processor

As shown in Figure 2, the GSC structure processor is divided into upper and lower branches [16,17]. The upper branch is the main channel containing the target and clutter component, and the lower branch is the auxiliary channel containing only the clutter component. The GSC processor can improve the SINR in the main channel by utilizing the clutter in the auxiliary channel to cancel the clutter in the main channel. The space–time steering vector of the main channel can be expressed as

$$S_t(f_{d,t}, f_{s,t}) = a(f_{d,t}) \otimes b(f_{s,t}) \quad (12)$$

where $f_{d,t}$ and $f_{s,t}$ denote the normalized spatial and normalized Doppler frequency, respectively. And $a(f_{d,t})$ and $b(f_{s,t})$ are the time and spatial steering vector, respectively. The matrix B of the lower branch represents the blocking matrix, which consists of the angular Doppler channel other than the main channel. On the one hand, there is no dimensionality reduction in the STAP processing when B contains all angular Doppler channels other than the main channel. On the other hand, different subsets of all angular Doppler channels other than the main channel can be selected to form different RD STAP algorithms.

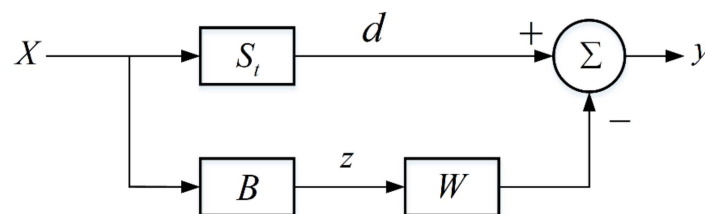


Figure 2. GSC form processor.

The GSC processor transforms the space–time adaptive processing detection structure into a standard Wiener filter. As shown in Figure 2, the output of the GSC filter without reduced dimension can be expressed as

$$y = d - W^H z \quad (13)$$

where d is the output of the upper branch and $z \in C^{(NM-1) \times 1}$ is the clutter vector that does not contain the target, given by

$$d = S_t^H X \quad (14)$$

$$z = B^H X \quad (15)$$

where $B \in C^{NM \times (NM-1)}$ is the blocking matrix, which is composed of space–time steer vectors of different angular Doppler channels except the main channel. It is well known that angular Doppler channels are perpendicular to each other. Therefore, we have

$$B^H S_t = 0 \in C^{(MN-1) \times 1} \quad (16)$$

The design philosophy of the GSC filter is to minimize the output power, which gives rise to the following optimization problem.

$$\min_W J(W) = \min_W E[\|y\|_2^2] \quad (17)$$

The optimal weight vector is calculated as

$$W_{opt} = R_z^{-1} r_{zd} \quad (18)$$

where R_z denotes the CNCM of the lower branch

$$R_z = E[zz^H] \quad (19)$$

r_{zd} is the cross-correlation vector between the upper and lower branch outputs.

$$r_{zd} = E[zd^*] \quad (20)$$

3. Proposed Algorithm

In the proposed algorithm, the CNCM of the cell under test $\hat{R}$ is first estimated by the SBL algorithm. Then, $\hat{R}$ and the proposed angular Doppler channel selection algorithm are used to design the RD transformation matrix T_r . When T_r is obtained, T_r and $\hat{R}$ are used to design the RD filter. Finally, the RD filter is used to process the data of the CUT and detect the target. The flowchart of the proposed algorithm is shown in Figure 3.

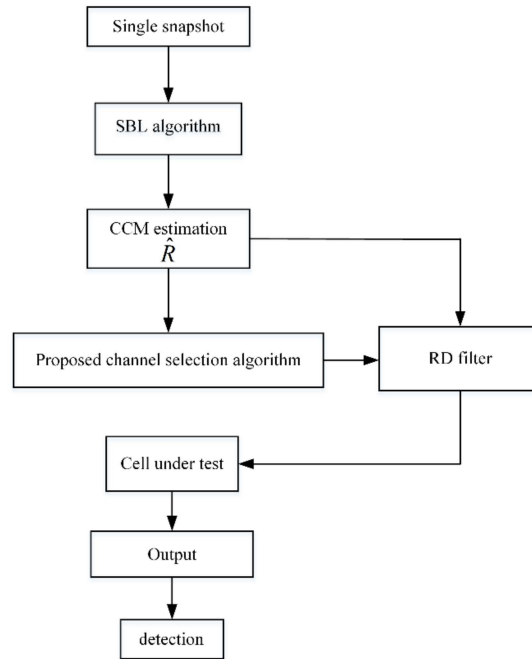


Figure 3. The flowchart of the proposed algorithm.

3.1. CNCM Estimation Method Based on SBL

In this section, we first derive the method for estimating the CNCM using SBL when only a single sample is available and then provide a pseudo-code for computing the covariance matrix. Rewrite the sparse signal model expressed in Equation (10) as follows.

$$x = Dw + \varepsilon \quad (21)$$

where $x \in C^{MN}$ is the measurement data vector, which is the single sample used in the proposed algorithm. $D \in C^{MN \times N_s N_d}$ is an overcomplete space-time dictionary obtained by discretizing the normalized space-time plane. $w \in C^{N_s N_d \times 1}$ is the sparse coefficient vector to be solved. $\varepsilon \in C^{MN \times 1}$ is the noise vector with mean 0 and variance $\sigma^2 \mathbf{I}_{MN}$, and noise power σ^2 is the unknown parameter. Based on the above assumptions, the likelihood probability density function of x has the following form:

$$p(x|w, \sigma^2) = (2\pi\sigma^2)^{-\frac{MN}{2}} \exp\left(-\frac{1}{2\sigma^2} \|x - Dw\|^2\right) \quad (22)$$

To ensure the sparsity of the model, we assign to each element of w a Gaussian prior with zero mean and α_i^{-1} variance. As a result, the joint distribution of each element in w can be expressed as

$$p(w|\alpha) = \prod_{i=1}^{N_s N_d} (2\pi)^{-\frac{1}{2}} \alpha_i^{\frac{1}{2}} \exp\left(-\frac{\alpha_i}{2} w_i^2\right) \quad (23)$$

where $\alpha = [\alpha_1^{-1}, \alpha_2^{-1}, \dots, \alpha_{N_d N_s}^{-1}]^T$ is the hyperparameters vector. According to [18], the suitable priors thereover are Gamma distributions for α and σ^2 . As a result, the joint probability density function of the elements in vector α can be expressed as

$$p(\alpha) = \prod_{i=1}^{N_s N_d} \text{Gamma}(\alpha_i | a, b) \quad (24)$$

$$p(\sigma^{-2}) = \text{Gamma}(\sigma^{-2} | c, d) \quad (25)$$

where

$$\text{Gamma}(\alpha | a, b) = \Gamma(a)^{-1} b^a \alpha^{a-1} e^{-b\alpha} \quad (26)$$

with $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$. To make these priors non-informative [18], we might fix their parameters to small values: e.g., $a = b = c = d = 10^{-6}$. Then, calculating the estimates of each parameter based on the given data can be accomplished by calculating the posterior probability density function as follows:

$$p(w, \alpha, \sigma^2 | x) = \frac{p(x | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2)}{p(x)} \quad (27)$$

To simplify the calculation, this posterior is written in another equivalent form:

$$p(w, \alpha, \sigma^2 | x) = p(w | x, \alpha, \sigma^2) p(\alpha, \sigma^2 | x) \quad (28)$$

According to the famous Bayes' theorem, the posterior distribution over the w is thus given by [18]

$$\begin{aligned} p(w | x, \alpha, \sigma^2) &= \frac{p(x | w, \sigma^2) p(w | \alpha)}{p(x | \alpha, \sigma^2)} \\ &= (2\pi)^{-\frac{N_s N_d}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(w - \mu)^T \Sigma^{-1}(w - \mu)\right) \end{aligned} \quad (29)$$

Therefore, the posterior covariance and mean are expressed respectively as [19,20]

$$\Sigma = (\sigma^{-2} D^H D + \Lambda)^{-1} \quad (30)$$

$$\mu = \sigma^{-2} \Sigma D^H x \quad (31)$$

where $\Lambda = \text{diag}([\alpha_1, \alpha_2, \dots, \alpha_{N_s N_d}])$.

Next, the expectation maximization (EM) [18] algorithm is used to update the hyperparameters α and σ^2 . First, calculate the updated formula for the hyperparameter α . Ignoring the terms in the logarithm that are not relevant to α , we can obtain the following expectation:

$$E_{w | t, \alpha, \beta}(\log(p(w | \alpha) p(\alpha))) \quad (32)$$

We maximize this expectation to obtain an iterative update formula for α . By differentiating this expectation in Equation (32) and making the derivative equal to 0, we obtain the following iterative updating equation:

$$(\alpha_i)^{new} = \frac{1}{E_{w | t, \alpha, \beta}(w_i^2)} = \frac{1}{\Sigma_{i,i} + \mu_i^2} \quad (33)$$

By following the same steps as before, we can obtain the following expectation associated only with σ^{-2} :

$$E_{w | t, \alpha, \beta}(\log(p(x | w, \beta) p(\beta))) \quad (34)$$

By maximizing the expectation in Equation (34), we can obtain an iterative update formula for σ^2 :

$$(\sigma^2)^{new} = \frac{E_{w|t,\alpha,\beta}(\|x - Dw\|^2)}{N_s N_d} = \frac{\|x - D\mu\|^2 + \text{trace}(\Sigma D^H D)}{N_s N_d} \quad (35)$$

The iterative updates of α and σ^2 does not stop until the preset constraints are satisfied. After the iterative updating of the parameters converges, the CNCM can be calculated as

$$\hat{R} = D\Pi D^H + \sigma^2 I_{MN} \quad (36)$$

where $\Pi = \text{diag}(\mu_1^2, \mu_2^2, \dots, \mu_{N_s N_d}^2)$.

The pseudo-code for the SBL-based CNCM estimation method is shown in Algorithm 1.

Algorithm 1. CNCM estimation method based on SBL.

Step1: Given the initial values $\sigma_0^2 = 0.01$, $\alpha_{0,i} = 1$, $i = 1, 2, \dots, N_s N_d$

Step2: Compute the posterior moments

$$\Sigma_{k+1} = (\sigma_k^{-2} D^H D + \Lambda_k)^{-1}$$

$$\mu_{k+1} = \sigma_k^{-2} \Sigma_{k+1} D^H x$$

Step3: Update the α and σ^2 using the EM rule

$$\alpha_{k+1,i} = \frac{1}{\Sigma_{k+1,i,i} + \mu_{k+1,i}^2}$$

$$\sigma_{k+1}^2 = \frac{\|x - D\mu_{k+1}\|^2 + \text{trace}(\Sigma_{k+1} D^H D)}{N_s N_d}$$

Step4: Continue Step 2 and 3 until convergence

Step5: Assuming $\hat{w} = \mu_{k+1}$

Step6: Compute the CNCM by

$$\hat{R} = D\hat{w}\hat{w}^H D^H + \sigma_{k+1}^2 I_{MN}$$

3.2. Channel Selection Algorithm Based on SINR Maximum Criterion

As shown in Figure 4, the matrix V in the lower branch denotes the selection matrix, which consists of some columns of the unit matrix. If K channels are selected as auxiliary channels, then the selection matrix V will consist of the K columns of the unit matrix and the DOF of the filter is reduced from $(NM - 1)$ to K . In RD STAP, the RD observation data vector is formed using the selection matrix $V \in C^{(NM-1) \times K}$:

$$b = V^H z \in C^{K \times 1} \quad (37)$$

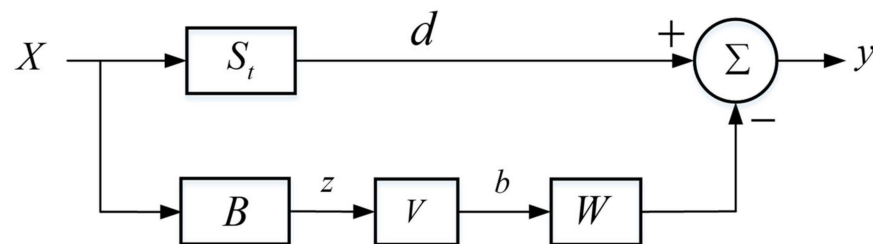


Figure 4. Reduced-dimension GSC form processor.

The corresponding $K \times K$ RD CNCM is shown below.

$$R_b = V^H R_z V \in C^{K \times K} \quad (38)$$

Now, we need to specify how to select the optimal auxiliary channel. We propose an iterative algorithm for selecting the auxiliary channel using the maximum output SINR as a criterion. When selecting the first channel, we need to choose the one that maximizes the output SINR from the $J = (MN - 1)$ candidates. When we choose the j -th channel

as the first auxiliary channel, the selection matrix $V = [e_j]$ has only one column, where e_j represents the j -th column of the unit matrix. In this case, the RD matrix can then be expressed as

$$T_r = [S_t, BV] \in \mathbb{C}^{NM \times 2} \quad (39)$$

Apply the RD matrix to the estimated CNCM $\hat{R}$ to obtain the localized CNCM $\hat{R}_{Local}$.

$$\hat{R}_{Local} = T_r^H \hat{R} T_r \in \mathbb{C}^{2 \times 2} \quad (40)$$

According to [21], we have the following relationship:

$$\hat{R}_{Local} = \begin{bmatrix} \sigma_d^2 & \hat{r}_{bd}^H \\ \hat{r}_{bd} & \hat{R}_b \end{bmatrix} \quad (41)$$

Based on the conclusions of Equation (18), the filter weight under the RD GSC structure can be expressed as

$$W = \hat{R}_{Local}^{-1} \hat{r}_{bd} \quad (42)$$

and we have

$$W_T = \begin{bmatrix} 1 \\ -W \end{bmatrix} \quad (43)$$

The output SINR when the j -th auxiliary channel is selected can be calculated by the following equation:

$$SINR_j = \frac{|W_T^H S_T|^2}{W_T^H \hat{R}_{Local} W_T}, j = 1, 2, \dots, J \quad (44)$$

The angular Doppler channel that maximizes the output SINR is the first auxiliary channel selected. The problem of finding the maximum output SINR is expressed as follows:

$$m = \max_j (SINR_j) \quad (45)$$

Then, the selection matrix $V = [e_m]$, for the sake of expression, let $v_1 = e_m$, this means that the first column of the selection matrix is e_m . Then, the selected channel is deleted from the candidate channel, and there are $(NM - 2)$ candidate channels left.

When we select the k -th auxiliary channel, there are already $(k - 1)$ channels selected and $(MN - k - 2)$ candidate channels left. The selection matrix V can be expressed as

$$V = [v_1, \dots, v_{k-1}, v_k] \quad (46)$$

where $v_1, \dots, v_{k-1}$ have been selected in the previous step, and v_k is the channel to be selected. In other words, the selection matrix V in Equation (46) has only one column to be determined, namely v_k . In this case, the RD matrix is $T_r \in \mathbb{C}^{MN \times (k+1)}$, and the localized CNCM is $\hat{R}_{Local} \in \mathbb{C}^{(k+1) \times (k+1)}$. Calculate the output SINR according to Equations (39)–(45) and select the channel with the largest output SINR as v_k . In general, selecting 3–5 channels can achieve suboptimal clutter suppression performance.

The pseudo-code of the proposed channel selection algorithm is represented in Algorithm 2. It should be emphasized that the proposed channel selection algorithm is an iterative algorithm. Before use, you can preset the number of channels to be used as the iteration stop condition. For example, if you want the algorithm to stop when C channels are selected, then the C channels selected have the largest output SINR compared to other channel combinations with the same number of channels. Step2 and Step3 in Algorithm 2 are the core iteration steps. Each time Step2 and Step3 are executed, one channel will be selected and the number of candidate channels will be reduced by one.

Algorithm 2. The pseudo-code of the proposed channel selection algorithm.

Step1: Given the initial value:
 Estimated covariance matrix $\hat{R}$
 selection matrix $V \in \emptyset$
 blocking matrix B contains all candidate auxiliary channels
 total number of candidate channels $J = (MN - 1)$

Step2: Loop: $j = 1, 2, \dots, J$
 $Q = [V, e_j]$
 $T_r = [S_t, BQ]$
 $\hat{R}_{Local} = T_r^H \hat{R} T_r$
 $W = \hat{R}_{Local}^{-1} \hat{r}_{bd}$
 $W_T = \begin{bmatrix} 1 \\ -W \end{bmatrix}$
 $SINR_j = \frac{|W_T^H S_t|^2}{W_T^H \hat{R}_{Local} W_T}$
 $SINR(j) = SINR_j$
 end loop

Step3: solve $m = \max_j (SINR(j))$
 Let $V = [V, e_m]$, and remove $B e_m$ from candidate channel
 the number of remaining candidate channels $J = J - 1$

Step4: Continue Step2 and Step3 until preset constraints are satisfied.

Output: Selection matrix V

4. Numerical Simulation

In this section, we will compare the performance of the proposed algorithm with several existing classical algorithms of the same type through simulation experiments. The performance is mainly judged by the output SINR and output SINR loss, and the formulas for calculating these two metrics are provided below [1]:

$$SINR = \frac{\sigma^2 \xi_t |W^H S_t|^2}{W^H R W} \quad (47)$$

$$LSINR = \frac{\sigma^2}{MN} \frac{|W^H S_t|^2}{W^H R W} \quad (48)$$

where R , W , and ξ_t refer to the ideal CNCM, STAP filter weight, and preset signal-to-noise ratio (SNR), respectively.

An airborne phased array radar system is considered in the simulation, and the preset simulation parameters are shown in Table 2. To ensure that the simulations are correct, all the simulation results in the paper are the average of 100 Monte Carlo experiments.

Table 2. List of simulation parameters.

Symbols	Parameters	Value
λ	wavelength	0.25 m
d	element spacing	0.125 m
v	velocity	125 m/s
H	height	8000 m
M	number of pulses	12
N	number of elements	12
f_r	PRF	2000 Hz
f_s	sampling frequency	1 MHz
θ_m	main lobe elevation angle	0°
ϕ_m	main lobe azimuth angle	90°
ξ_c	CNR	40 dB
ξ_t	SNR	10 dB
N_s	Angular discretization	4 N
N_d	Doppler discretization	4 M

4.1. Performance Analysis Based on Simulation Data

In order to clearly state the simulation results in this section, the intuitive view of the clutter Capon spectrum of the data used in the simulation is given. Figure 5a shows the known ideal Capon spectrum of the CUT. Figure 5b shows the Capon spectrum estimated by the SBL algorithm, with only a single sample available. The value of the color bar in Figure 5 represents the normalized value of the clutter distribution relative to the main lobe on the space–time two-dimensional plane. Figure 5c,d show the clutter Capon spectra obtained using 100 training samples and 10 training samples according to the ML estimation method, respectively. By comparing these four Capon spectra in Figure 5, it can be clearly seen that the accuracy of the Capon spectra estimated by the SBL algorithm is between those estimated by 100 and 10 samples, respectively.

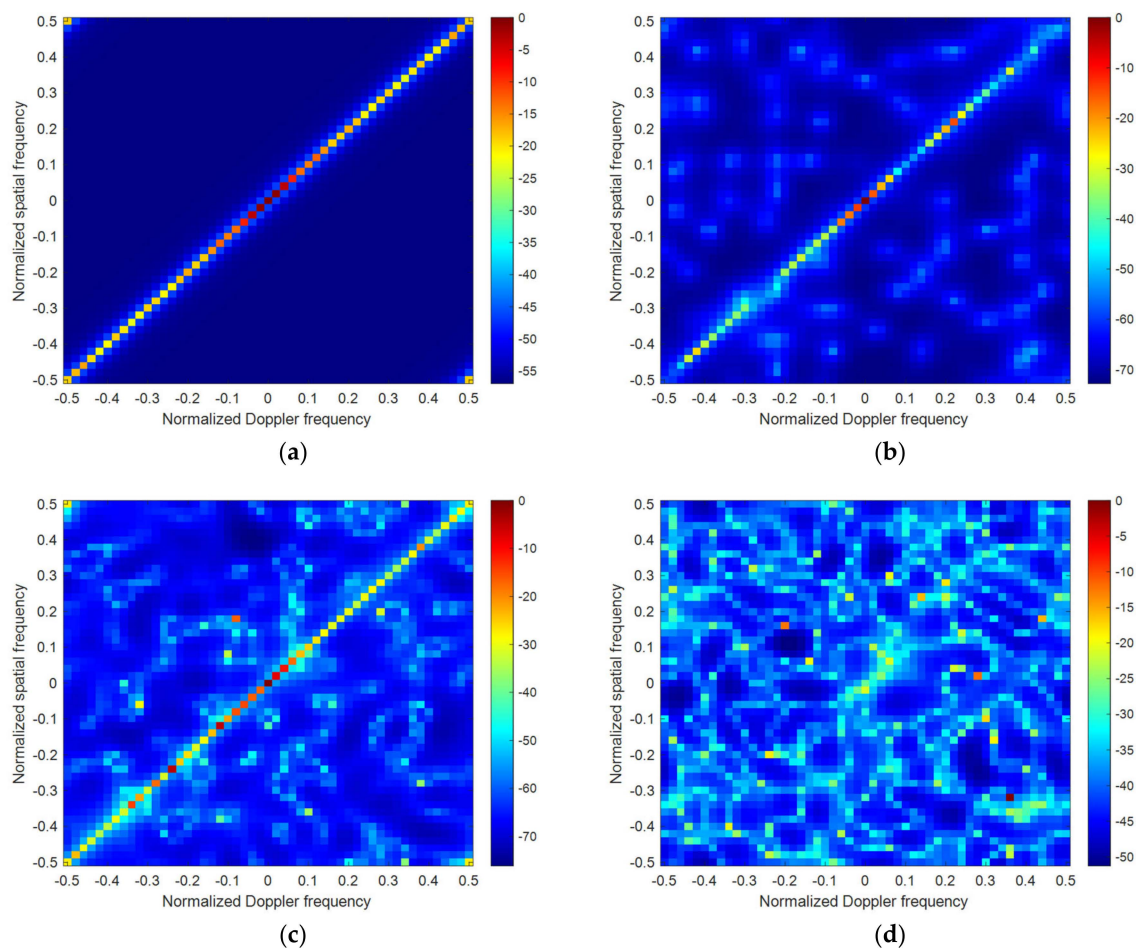


Figure 5. Comparison of estimated clutter Capon spectra. (a) Ideal CNCM; (b) SBL, a single snapshot; (c) ML, 100 snapshots; (d) ML, 10 snapshots.

It is well known that the space scanned by the eigenvectors corresponding to large eigenvalues is the clutter subspace [12,22], so the number of large eigenvalues can reflect the accuracy of CNCM estimation. According to Brennan’s criterion [1], the number of large eigenvalues in this paper should be 23. Figure 6 clearly shows that the accuracy of estimating CNCM using 100 samples is much higher than using 10 samples, and the accuracy of estimating CNCM using the SBL algorithm in the single-sample case is somewhere in between.

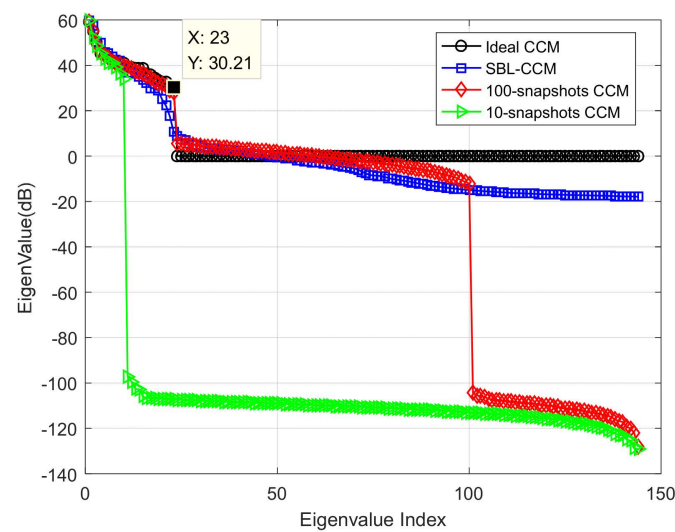


Figure 6. Comparison of eigenspectra of ideal and estimated CNCM.

In addition, the performance of the proposed algorithm varies with the number of selected auxiliary channels under different target normalized Doppler frequencies, which are presented in Figure 7a–f. From Figure 7, We can clearly see that only 3–5 auxiliary channels are needed to achieve suboptimal clutter suppression performance, and using more channels not only takes a huge computational burden but also gains little performance improvement. It can also be found that the clutter suppression performance is rather degraded when too many auxiliary channels are used, which is due to the inaccurate CNCM estimation. Therefore, from the point of view of balancing algorithmic performance and computational complexity, it is wise to choose 3–7 auxiliary channels.

It should be noted that there is a total of $(MN - 1)$ auxiliary channels to choose from. When all candidate channels are selected, the proposed algorithm degenerates into the traditional STAP algorithm based on sparse recovery (SR STAP), and naturally, there is no dimensionality reduction. The SINR of the proposed algorithm with different numbers of auxiliary channels is shown in Figure 8, where C represents the number of auxiliary channels selected. $C = 1$ indicates that the proposed algorithm selects only one auxiliary channel. From Figure 8, we can observe that the output SINR of the proposed algorithm is better than that of traditional SR STAP when only one auxiliary channel is used. Furthermore, we can observe that the SINR of the proposed algorithm with 80 auxiliary channels is approximately the same as that with 5 auxiliary channels but slightly lower than that with 7 auxiliary channels, which is consistent with the simulation results in Figure 7.

The comparison of the output SINR loss of the proposed algorithm with the best STAP [1], JDL [5] algorithm, and direct data domain (D3) algorithm [23] is shown in Figure 9. In the simulation of Figure 9, the number of auxiliary channels used by the proposed algorithm is 3, and the number of auxiliary channels used by the JDL algorithm is 8. Three scenarios for estimating the CNCM are considered: 100 samples, 10 samples, and the SBL algorithm, which uses only one sample to estimate the CNCM. In the simulation of the direct data domain algorithm, the lengths of the spatial and temporal sliding windows are 2, 4, and 6, respectively. The simulation results are shown in Figure 9, where it can be clearly seen that the JDL algorithm still has a larger output SINR loss than the proposed algorithm when using 100 samples. Moreover, the output SINR loss of the proposed algorithm is nearly 20 dB smaller than that of the direct data domain algorithm, which indicates that the proposed algorithm has better clutter suppression performance than the direct data domain algorithm.

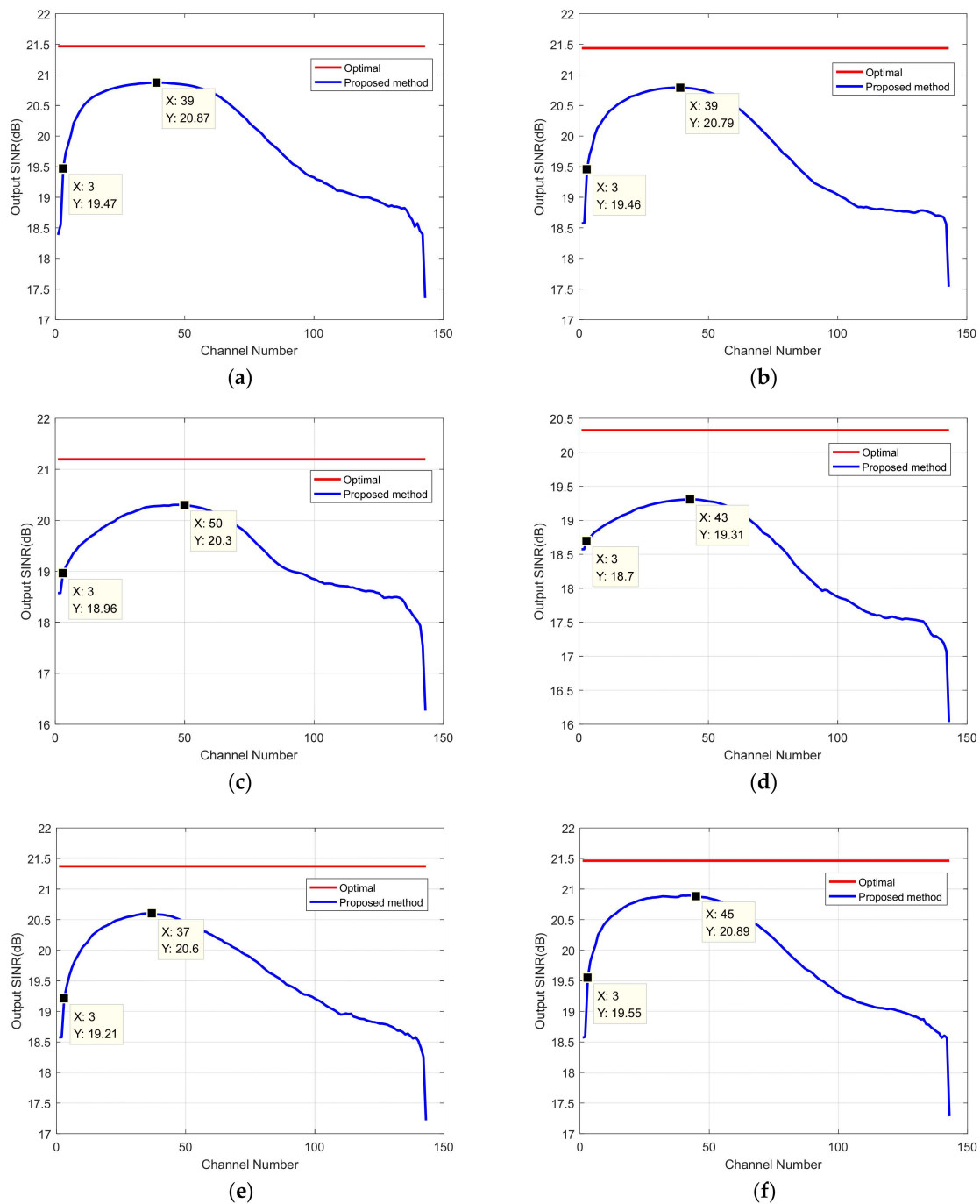


Figure 7. The SINR of the proposed algorithm varies with the number of auxiliary channels used. (a) $f_d = -0.5$; (b) $f_d = -0.3333$; (c) $f_d = -0.1667$; (d) $f_d = 0.0833$; (e) $f_d = 0.25$; (f) $f_d = 0.4167$.

To illustrate the superiority of the proposed channel selection algorithm, we compare the output SINR loss of the SBL-JDL algorithm with that of the proposed algorithm. The SBL-JDL algorithm is a single-sample algorithm that obtains the CNCM by the SBL algorithm and then selects 8 auxiliary channels according to the JDL algorithm. The only difference between the SBL-JDL algorithm and the proposed algorithm is the method of selecting auxiliary channels. As can be seen in Figure 9, although the number of auxiliary channels used is less than that of the SBL-JDL algorithm, the proposed algorithm still has a better SINR loss than the SBL-JDL algorithm.

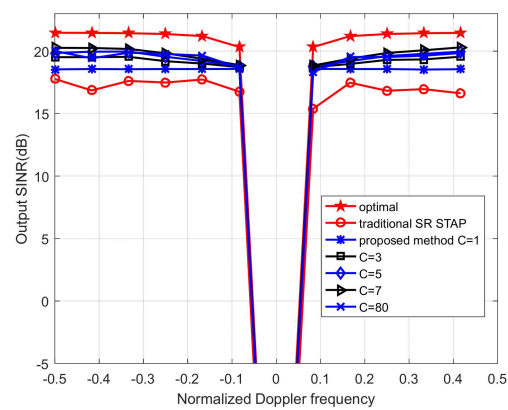


Figure 8. SINR comparison of the proposed method with $C = 1, 3, 5, 7, 80$ and the traditional SR STAP.

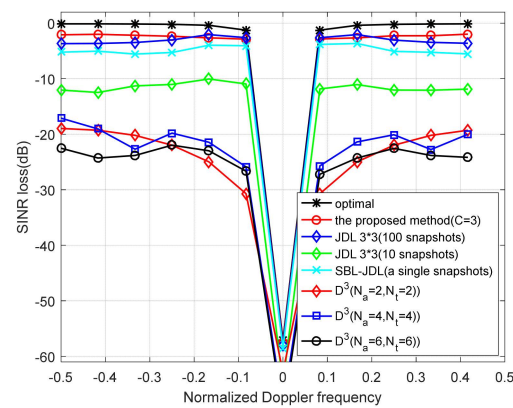


Figure 9. Clutter suppression performance comparison of several typical algorithms.

We also simulated the target detection performance of the optimum STAP, JDL algorithm (100 samples), direct data domain algorithm, and the proposed algorithm. The target is added to the 51– th range bin. The simulation results are shown in Figure 10. We can see that the proposed algorithm performs close to the optimal STAP and the JDL algorithm using 100 samples when the constant false alarm detector is used.

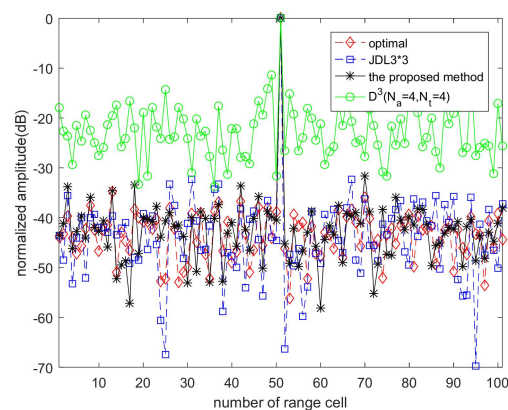


Figure 10. Comparison of adaptive beamforming output of several typical algorithms.

So far, we have been considering an ideal error-free clutter environment. In order to verify the robustness, we introduce the gain-phase (GP) error in the simulation experiments in Figure 11. According to [16,24], the clutter model with GP errors can be expressed as

$$x_{GP} = EDw + x_n \quad (49)$$

The GP error matrix E can be expressed as

$$E = I_M \otimes \text{diag}\left([e_1 \exp(j\varphi_1), e_2 \exp(j\varphi_2), \dots, e_N \exp(j\varphi_N)]^T\right) \quad (50)$$

where e_j and φ_j denote the amplitude error and phase error added to the j -th array element, respectively. In Figure 11, the average fluctuation of the added amplitude error is 3%, 5%, and 7% of the amplitude without error, respectively. The average fluctuations of the added phase error are 3° , 5° , and 10° , respectively. The number of auxiliary channels used by the proposed algorithm is 3, 5, 7, and 8, while the number of auxiliary channels used by the JDL algorithm is always 8. In addition to this, the proposed algorithm uses only one sample, and the JDL algorithm uses 100 samples.

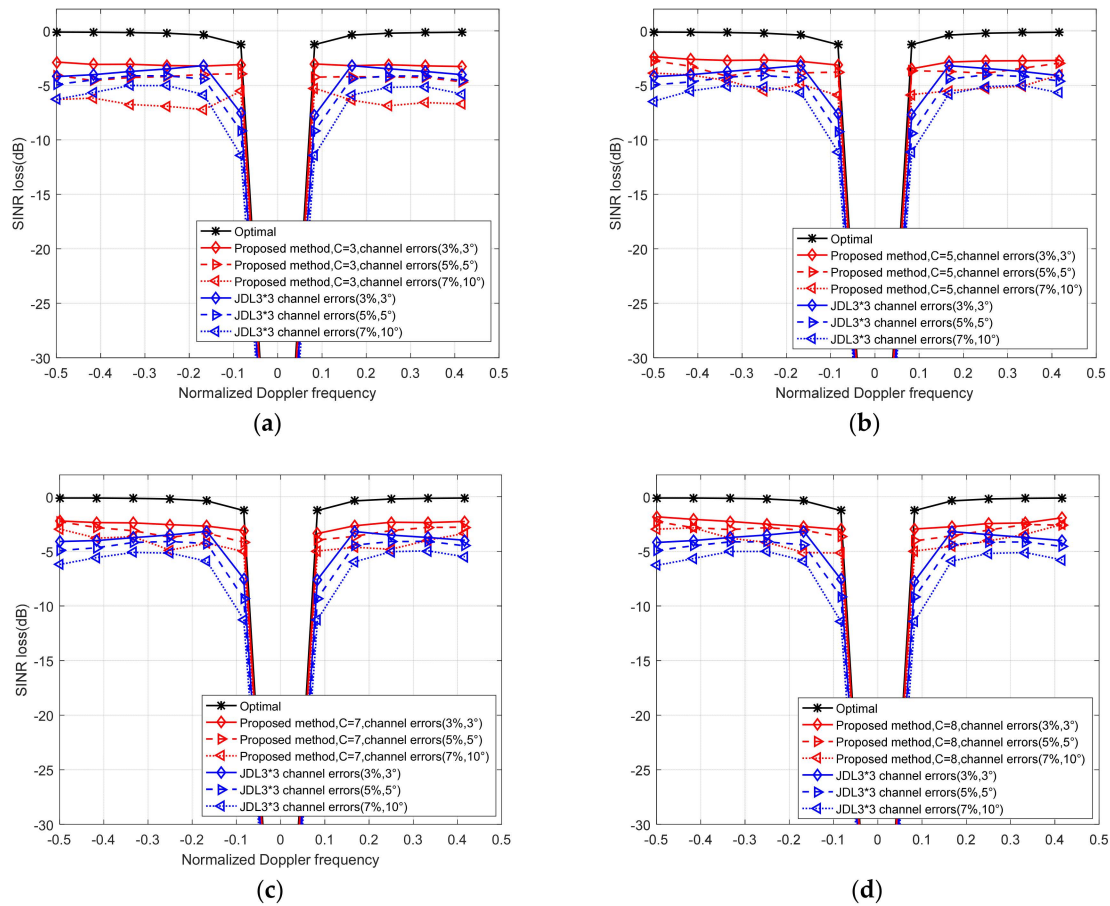


Figure 11. SINR loss of the proposed algorithm and JDL algorithm in the presence of GP error. (a) Proposed algorithm, $C = 3$; JDL, $C = 8$; (b) proposed algorithm, $C = 5$; JDL, $C = 8$; (c) proposed algorithm, $C = 7$; JDL, $C = 8$; (d) proposed algorithm, $C = 8$; JDL, $C = 8$.

The simulation results in Figure 11 show that the introduction of GP error degrades the performance of the proposed algorithm and the JDL algorithm, and the larger the GP error, the more the performance degrades. In addition, the output SINR loss of the proposed algorithm decreases as the number of auxiliary channels used increases. The proposed algorithm has no significant advantage over the JDL algorithm when only three auxiliary channels are used in the presence of GP errors, as can be seen in Figure 11a. From Figure 11b–d, gradually increasing the number of auxiliary channels used, the performance of the proposed algorithm gradually improves and outperforms the JDL algorithm. The proposed algorithm performs significantly better than the JDL algorithm when 7 or 8 auxiliary channels are used.

4.2. Performance Analysis Based on Measured Data

To further analyze the performance of the proposed algorithm, the proposed algorithm is applied to the Mountain-top data (t38pre01v1_cpi_6) in this subsection. The data were collected by the Lincoln Laboratory at the Massachusetts Institute of Technology [25]. The specific parameters of the Mountain-top dataset are shown in Table 3.

Table 3. The main parameter of Mountain-top dataset.

Parameters	Value
Array number	14
Pulses number	16
PRF	625 Hz
Bandwidth	500 kHz
Range cell number	403
Range cell of target	147
Azimuth angle of target	275°
Normalized Doppler frequency of target	0.25

This dataset contains a total of 403 range cells, and the target is located at the 147th range cell. Using the clutter data of all 403 range cells, the clutter covariance matrix is estimated as [2]

$$R_m = \frac{1}{403} \sum_{i=1}^{403} x_i x_i^H \quad (51)$$

In Formula (51), x_i represents the clutter data of the i -th range cell. Using R_m to calculate clutter Capon spectrum is shown in Figure 12.

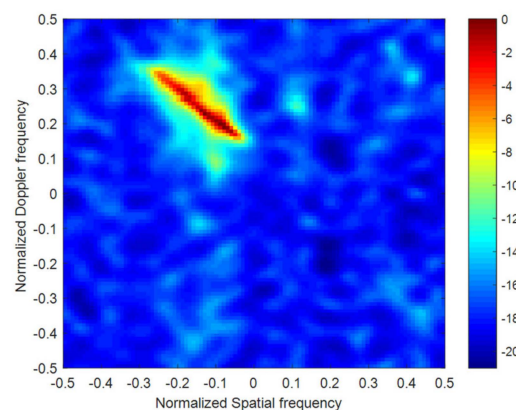


Figure 12. Estimated clutter spectrum using all 403 range cells.

When the performance of the proposed algorithm is verified using the measured dataset, the sliding window method is used to detect the presence of the target along the range cell. In order to prevent the target energy leakage caused by the movement of the target from affecting the detection performance, four range cells on the left and right sides of the range cell under test are specified as protection cells. The clutter data of the fifth range cell on the left or right adjacent to the range cell under test are used as training samples for target detection according to the proposed algorithm. The detection results along the range cell are shown in Figure 13. In Figure 13, circled by a purple dashed ellipse, is the output energy of the range cell where the target is located. It can be seen from Figure 13 that the proposed algorithm can detect targets both when 5 channels and 10 channels are used. The difference between the two cases is that when 10 channels are used, the range cell without a target has less clutter remaining. Comparing the proposed algorithm with the traditional SR STAP algorithm, it can be seen that when only one training sample is used, the traditional SR STAP not only fails to detect the target but also generates a false alarm in

the range cell without a target. It is obvious that the performance of traditional SR STAP will decrease significantly when the CCN estimation is not accurate, but the proposed algorithm can still maintain acceptable target detection performance.

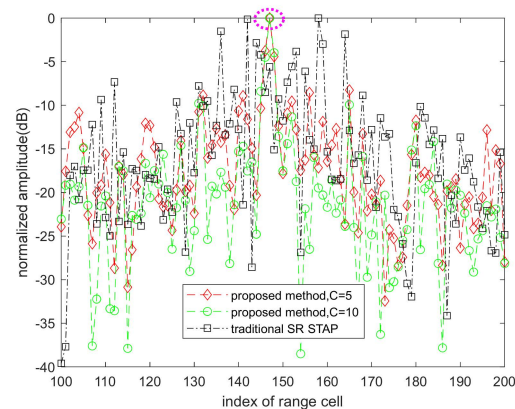


Figure 13. Output power against the range cell.

5. Conclusions

In this work, an angular Doppler domain reduced-dimension STAP algorithm based on SBL was proposed. First, the CNM is estimated by the SBL algorithm using only one sample. Second, the RD matrix is designed using the estimated CNM and the proposed angular Doppler channel selection algorithm. Finally, the estimated CNM and the designed RD matrix are used to design an RD filter to process the data of the CUT and detect the target. The experimental results demonstrate that the proposed algorithm maintains suboptimal performance in extremely heterogeneous clutter environments where only one sample is available. Due to the high computational complexity of the currently popular efficient sparse recovery algorithms, in order to improve the real-time performance of the proposed algorithms, the future research direction is mainly focused on the fast and accurate sparse recovery algorithms suitable for airborne radar.

Author Contributions: Methodology, J.C.; Software, J.C.; Validation, D.W.; Investigation, T.W. and D.W.; Writing—original draft, J.C.; Visualization, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: Key Technologies Research and Development Program, Grant Number: 2021YFA1000400.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

References

1. Ward, J. *Space-Time Adaptive Processing for Airborne Radar*; Technical Report, 1015; MIT Lincoln Laboratory: Lexington, MA, USA, 1994.
2. Reed, I.S.; Mallett, J.D.; Brennan, L.E. Rapid Convergence Rate in Adaptive Arrays. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *10*, 853–863. [\[CrossRef\]](#)
3. Klemm, R. Adaptive airborne MTI: An auxiliary channel approach. In *IEE Proceedings F (Communications, Radar and Signal Processing)*; IET Digital Library: London, UK, 1987; Volume 134, pp. 269–276.
4. Dipietro, R.C. Extended factored space-time processing for airborne radar systems. In *Proceedings of the Conference Record of the Twenty-Sixth Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, USA, 26–28 October 1992; IEEE Computer Society: Pacific Grove, CA, USA, 1992; pp. 425–430.
5. Wang, H.; Cai, L. On Adaptive Spatial-Temporal Processing for Airborne Surveillance Radar Systems. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *7*, 660–670. [\[CrossRef\]](#)
6. 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.
7. Yang, Z.; Li, X.; Wang, H.; Nie, L. Sparsity-Based Space-Time Adaptive Processing Using Complex-Valued Homotopy Technique. In *Proceedings of the 2013 IEEE Radar Conference*, Ottawa, ON, Canada, 29 April–3 May 2013.

8. Sen, S. Low-Rank Matrix Decomposition and Spatial-Temporal Sparse Recovery for STAP Radar. *IEEE J. Sel. Top. Signal Process.* **2015**, *12*, 1510–1523. [[CrossRef](#)]
9. Duan, K.; Wang, Z.; Xie, W. Sparsity-based STAP algorithm with multiple measurement vectors via sparse Bayesian learning strategy for airborne radar. *IET Signal Process.* **2016**, *9*, 544–553. [[CrossRef](#)]
10. Wang, Z.; Wang, Y.; Gao, F. Clutter nulling space-time adaptive processing algorithm based on sparse representation for airborne radar. *IET Radar Sonar Navig.* **2017**, *11*, 177–184.
11. Feng, W.; Guo, Y.; He, X. Jointly Iterative Adaptive Approach Based Space Time Adaptive Processing Using MIMO Radar. *IEEE Access.* **2018**, *6*, 26605–26616. [[CrossRef](#)]
12. Wang, D.; Wang, T.; Cui, W.; Zhang, X. A Clutter Suppression Algorithm via Enhanced Sparse Bayesian Learning for Airborne Radar. *IEEE Sens. J.* **2023**, *5*, 10900–10911. [[CrossRef](#)]
13. Fa, R.; De Lamare, R.C. Reduced-Rank STAP Algorithms using Joint Iterative Optimization of Filters. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *7*, 1668–1684. [[CrossRef](#)]
14. Sun, K.; Meng, H.; Wang, Y. Direct Data Domain STAP using Sparse Representation of Clutter Spectrum. *Signal Process.* **2011**, *4*, 2222–2236. [[CrossRef](#)]
15. Guo, Y.; Liao, G.; Feng, W. Sparse Representation Based Algorithm for Airborne Radar in Beam-Space Pose-Doppler Reduced-Dimension Space-Time Adaptive Processing. *IEEE Access* **2017**, *4*, 5896–5903. [[CrossRef](#)]
16. Yang, Z.; de Lamare, R.C.; Li, X. L1-Regularized STAP Algorithms with a Generalized Sidelobe Canceler Architecture for Airborne Radar. *IEEE Trans. Signal Process.* **2012**, *2*, 674–686. [[CrossRef](#)]
17. Zhang, W.; He, Z.; Li, J. A Method for Finding Best Channels in Beam-Space Post-Doppler Reduced-Dimension STAP. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *1*, 254–264. [[CrossRef](#)]
18. Tipping, M.E. Sparse Bayesian Learning and the Relevance Vector Machine. *J. Mach. Learn.* **2001**, *1*, 211–244.
19. Shutin, D.; Buchgraber, T.; Kulkarni, S.R. Fast Variational Sparse Bayesian Learning with Automatic Relevance Determination for Superimposed Signals. *IEEE Trans. Signal Process.* **2011**, *12*, 6257–6261. [[CrossRef](#)]
20. Zhou, W.; Zhang, H.-T.; Wang, J. An efficient Sparse Bayesian Learning Algorithm Based on Gaussian-Scale Mixtures. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *1*, 3065–3078. [[CrossRef](#)] [[PubMed](#)]
21. Zhang, W.; An, R.; He, N. Reduced Dimension STAP Based on Sparse Recovery in Heterogenous Clutter Environments. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *2*, 785–795. [[CrossRef](#)]
22. Haimovich, A.M.; Berin, M. Eigenanalysis-based space-time adaptive radar: Performance analysis. *IEEE Trans. Aerosp. Electron. Syst.* **1997**, *10*, 1170–1179. [[CrossRef](#)]
23. Sarkar, T.K.; Wang, H.; Park, S. A deterministic least-squares approach to space-time adaptive processing (STAP). *IEEE Trans. Antennas Propag.* **2001**, *2*, 91–103. [[CrossRef](#)]
24. Wang, D.; Wang, T.; Cui, W. Adaptive Support-Driven Sparse Recovery STAP Method with Subspace Penalty. *Remote Sens.* **2022**, *14*, 4463. [[CrossRef](#)]
25. Titi, G.W.; Marshall, D.F. The ARPA/NAVY mountaintop program: Adaptive signal processing for airborne early warning radar. In Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference, Atlanta, GA, USA, 9 May 1996; IEEE: Piscataway, NJ, USA, 1996; Volume 5, pp. 1165–1168.

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.