CrossMark

Sparsity-Based Direct Data Domain Space-Time Adaptive Processing with Intrinsic Clutter Motion

Zhaocheng Yang[1] · **Yuliang Qin**[2] ·
Rodrigo C. de Lamare[3] · **Hongqiang Wang**[2] ·
Xiang Li[2]

Abstract In this paper, we propose a sparsity-based direct data domain space-time adaptive processing (D3-STAP) algorithm for airborne radar that considers the intrinsic clutter motion (ICM). The proposed D3-STAP scheme models the received returns in the presence of ICM as a sparse measurement model. Then, we derive the principle of the sparsity-based D3-STAP that uses the focal underdetermined system solution (FOCUSS) method. The proposed D3-STAP algorithm estimates the clutter covariance matrix by a Hadamard product of the covariance matrix taper (CMT) and the clutter covariance matrix estimate with the FOCUSS technique. In addition, we develop a CMT adaptation approach for the proposed D3-STAP algorithm to automatically select the best CMT. Simulation results show that the proposed algorithm outperforms the

✉ Zhaocheng Yang
yangzhaocheng@gmail.com

Yuliang Qin
nudtqyliang@gmail.com

Rodrigo C. de Lamare
rcdl500@ohm.york.ac.uk

Hongqiang Wang
oliverwhq@vip.tom.com

Xiang Li
lixiang01@vip.sina.com

[1] College of Information Engineering, Shenzhen University, Shenzhen, Guangdong, China

[2] Research Institute of Space Electronics, Electronics Science and Engineering School, National University of Defense Technology, Changsha 410073, China

[3] Communications Research Group, Department of Electronics, University of York, York YO10 5DD, UK

Birkhäuser

existing D3-STAP using the least-squares technique and the sparsity-based D3-STAP algorithm without CMT.

Keywords Space-time adaptive processing · Covariance matrix taper · Intrinsic clutter motion · Spatio-temporal sparsity · Direct data domain

1 Introduction

Space-time adaptive processing (STAP) techniques usually employ the training data to design the filters, which requires that the statistics of the training data are the same as those of the test signal, such as reduced-rank and reduced-dimension methods (see [7,11,14,26] and the references therein), the parametric adaptive matched filter (PAMF) approach [20] and sparse beamformers [35,36]. However, in most real scenarios, the statistics of the training data are different from those of the test signal, resulting in significant performance degradation of conventional STAP algorithms. On the other hand, an alternative strategy that uses the received data in the cell under test (CUT) and requires no training data is developed to avoid estimation distortion caused by different statistics of the training data. Two types of approaches based on this idea were proposed: the direct data domain least-squares (D3-LS) STAP approach [24] and the maximum likelihood estimation detector (MLED) approach [1,4,5], where the first one was a deterministic approach and the latter one was a statistical processor. Unfortunately, the benefits of these approaches come at the cost of reduced system degrees of freedom (DOFs) leading to decreased performance. Hybrid detection approaches combining the STAP algorithms using the training data and those solely using the data in the CUT were subsequently put forward with improved robustness to nonhomogeneous environments [2,30]. Knowledge-aided (KA) STAP methods were developed to enhance the detection performance, especially in nonhomogeneous environments (see, e.g., [9,27] and the references therein). Specifically, the authors in [17] introduced a KA parametric covariance estimation (KAPE) scheme by blending both prior knowledge and data observations within a parameterized model to capture instantaneous characteristics of the CUT. But the performance of the KAPE approach relies on the accuracy of prior knowledge.

Recently, sparsity-based STAP techniques have been considered for applications in moving target indication (MTI) [15,16,18,21–23,25,31–34]. In [18,23], they supposed that the moving targets in the spatio-temporal plane were sparse and excluded the clutter ridge with a mask or eliminated the clutter components with conventional STAP algorithms before applying a sparse regularization (SR) to the STAP filter design. However, the performance of these methods relies on the accurate prior knowledge of the clutter ridge or the quantities of the elimination of the clutter components. The approaches in [15,16] assumed that the entire scene with the target and clutter was sparse in the spatio-temporal plane and tried to reconstruct both the clutter and the target. Then, direct target detection was followed by using the reconstructed clutter and target image. Similar to the above idea, [22,25,32,34] also first reconstructed the scene with the clutter and the target, or only the clutter, but designed the STAP filter to suppress the clutter before target detection. Considering the interference with both

jammers and clutter, the approach in [21] sequentially estimated the jammer, clutter and target from measurements on passive radar operation, measurements adjacent to the CUT and measurements in the CUT, respectively. The work in [33] performed a theoretical analysis on clutter sparsity for a side-looking uniform linear array (ULA) with constant pulse repetition frequency (PRF), constant velocity and no crab. In [31], performance analysis and parameters setting of sparsity-based STAP algorithms based on five representative fast SR techniques were conducted. Compared with conventional STAP algorithms, the above sparsity-based STAP algorithms provide high resolution of the scene and exhibit significantly better performance in a very small number of snapshots [21,25,31,32,34] or even a single snapshot in the CUT [15,16,18,22,23,33].

In this work, we develop a sparsity-based D3-STAP algorithm by formulating a more realistic sparse measurement model for airborne radar that considers the intrinsic clutter motion (ICM), whereas none of the above-mentioned sparsity-based STAP articles addresses the ICM problem. By exploiting the fact that the ICM can be modeled as a Hadamard product of a unitary modulo "tapering" by the response of clutter patches associated without the ICM [10] and the properties of the Hadamard product, the proposed algorithm can be divided into four steps. First, it applies the focal underdetermined system solution (FOCUSS) method [12] as a sparse recovery technique to obtain the target plus clutter estimate. The primary clutter covariance matrix is sequentially calculated by excluding the target signal using a prior knowledge of the assumed target spatio-temporal frequencies in the spatio-temporal domain. Then, it combines the estimated clutter covariance matrix and the received data to adaptively estimate the "covariance matrix taper" (CMT). Third, it obtains the final clutter covariance matrix estimate by a Hadamard product of the CMT and the previously estimated clutter covariance matrix. Finally, it designs the STAP filter weights using the estimated clutter covariance matrix and suppresses the clutter, followed by the target detection. We consider a different clutter matrix modeling to that of [7,10,17]. Unlike [7,10,17] which employ a conventional STAP with CMT, we employ the SR framework to model the intrinsic sparsity of the target plus clutter spectrum and to derive the proposed sparsity-based D3-STAP algorithm. The numerical results show that the proposed algorithm provides a considerable signal-to-interference-plus-noise ratio (SINR) improvement compared with the existing D3-STAP algorithms, namely the D3-LS STAP algorithm and the conventional sparsity-based D3-STAP algorithm.

The main contributions of our paper are listed as follows:

- By introducing a sparse measurement model considering the ICM, a novel sparsity-based D3-STAP algorithm is proposed. Additionally, related parameter settings are analyzed.
- In the proposed algorithm, a CMT framework is introduced to overcome the ICM problem. An efficient adaptive approach is developed to select the best CMT. Furthermore, fast implementations of the proposed algorithm are developed.
- A study and comparative analysis of the proposed algorithm with other existing STAP algorithms for radar systems is carried out.

The remainder of the paper is organized as follows. In Sect. 2, we first develop a sparse measurement model considering ICM for airborne radar systems. Then, in Sect. 3, we derive the proposed sparsity-based D3-STAP algorithm and also discuss

the performance metrics for the proposed algorithm. Simulated airborne radar data are used to evaluate the performance of the proposed algorithm in Sect. 4. Section 5 provides the summary and conclusions.

Notation: Scalar quantities are denoted with italic typeface. Lowercase boldface quantities denote vectors and uppercase boldface quantities denote matrices. The operations of transposition, complex conjugation, and conjugate transposition are denoted by superscripts T, $*$, and H, respectively. The symbols $\otimes, \odot$, tr represent the Kronecker product, Hadamard product and trace operation, respectively. The symbol $E[\cdot]$ denotes the expected value of a random quantity, $\|x\|_p$ denotes the l_p-norm operation of x, and $|\cdot|$ denotes the absolute operation. $\mathbf{1}_N$ and $\mathbf{1}_{N \times N}$ represent the $N \times 1$ vector with N unity-elements and the $N \times N$ matrix with all unity-elements, respectively. Toeplitz$(\cdot)$ denotes the Toeplitz matrix.

2 Sparse Measurement Model

Considering a pulsed Doppler side-looking airborne radar moving with constant velocity v_p, we assume that the radar antenna is a ULA which consists of M elements with inner spacing d_a. The radar transmits N coherent burst of pulses at a constant PRF f_r, and the transmitter carrier frequency is $f_c = c/\lambda_c$, where c is the propagation velocity and λ_c is the wavelength.

In airborne radar, if we discretize the whole spatio-temporal plane into $N_s N_d$ ($N_s N_d \gg NM$) grid points, where N_s and N_d are the number of grid points along the spatial and temporal/Doppler frequency axes, respectively, a nonzero element from any such grid point would suggest the presence of a scatterer at that particular spatial and Doppler frequencies [21]. Thus, the target return can be represented by

$$\mathbf{x}_t = \mathbf{\Phi} \boldsymbol{\gamma}_t, \tag{1}$$

where $\boldsymbol{\gamma}_t = [\gamma_{t;1,1}, \gamma_{t;1,2}, \ldots, \gamma_{t;N_d,N_s}]^T$ denotes the $N_d N_s \times 1$ target spatio-temporal profile with nonzero elements representing the target scatterers. The matrix $\mathbf{\Phi}$ is the $NM \times N_d N_s$ over-completed space-time steering dictionary, as given by

$$\mathbf{\Phi} = [\mathbf{v}(f_{d,1}, f_{s,1}), \ldots, \mathbf{v}(f_{d,1}, f_{s,N_s}), \ldots, \mathbf{v}(f_{d,N_d}, f_{s,N_s})], \tag{2}$$

where

$$\begin{aligned} \mathbf{v}(f_{d,i}, f_{s,k}) &= \mathbf{v}_t(f_{d,i}) \otimes \mathbf{v}_s(f_{s,k}), \\ &\quad i = 1, \ldots, N_d, \ k = 1, \ldots, N_s. \end{aligned} \tag{3}$$

is an $NM \times 1$ ideal space-time steering vector with

$$\mathbf{v}_t(f_{d,i}) = [1, \ldots, \exp(j2\pi(N-1)f_{d,i})]^T, \tag{4}$$

$$\mathbf{v}_s(f_{s,k}) = [1, \ldots, \exp(j2\pi(M-1)f_{s,k})]^T, \tag{5}$$

indicating the ideal temporal/Doppler and spatial steering vectors, respectively.

As an approximation to a continuous field of clutter, the clutter return from iso-range of interest can be modeled as the superposition of a large number N_c of independent clutter patches that are evenly distributed in angle of arrival (AOA) about the radar [26]. Then, ignoring the impact of range ambiguities, the clutter component $\mathbf{x}_c$ can be represented as

$$\mathbf{x}_c = \mathbf{\Phi}\boldsymbol{\gamma}_c, \tag{6}$$

where $\boldsymbol{\gamma}_c$ denotes the $N_d N_s \times 1$ clutter spatio-temporal profile with nonzero elements representing clutter patches.

In real applications, in the presence of the ICM, the clutter return can be manifested as a random modulation of the form [10]

$$\mathbf{x}_{ct} = \mathbf{x}_c \odot \mathbf{t} = (\mathbf{\Phi}\boldsymbol{\gamma}_c) \odot \mathbf{t}, \tag{7}$$

where $\mathbf{t} = \mathbf{t}_d \otimes \mathbf{1}_M$ and $\mathbf{t}_d$, called the temporal decorrelation term, is an $N \times 1$ vector random process which is uncorrelated with $\mathbf{x}_c$. Then, the received signal, composed of the moving target, the clutter and the receiver thermal noise, can be represented by

$$\mathbf{x} = \mathbf{x}_t + \mathbf{x}_{ct} + \mathbf{n}, \tag{8}$$

where the $NM \times 1$ thermal noise vector $\mathbf{n}$ is usually assumed to be both spatially and temporally uncorrelated. On the one hand, because the clutter responses occupy only a diagonal ridge on the spatio-temporal plane (for the side-looking ULA or a set of concentric ellipses (for the nonside-looking case) [22], and the number of targets is limited, the number of nonzero elements in the spatio-temporal profile is much smaller than the dimension $N_d N_s$ of that. On the other hand, it is proved that for the case of side-looking radar with a ULA, constant PRF, constant platform velocity and no crab angle, there is a group of space-time steering vectors (whose number is equivalent to the clutter rank) that can approximately represent the clutter subspace [33].[1] Therefore, there is a high degree of sparsity in the spatio-temporal profile, thereby calling the model in (8) as the sparse measurement model.

We assume the responses of the clutter are zero-mean, complex Gaussian random numbers; then, the clutter covariance matrix is given by

$$\mathbf{R}_{ct} = E[\mathbf{x}_{ct}\mathbf{x}_{ct}^H] = E[((\mathbf{\Phi}\boldsymbol{\gamma}_c) \odot \mathbf{t})((\mathbf{\Phi}\boldsymbol{\gamma}_c) \odot \mathbf{t})^H]$$
$$= (\mathbf{\Phi}\mathbf{\Gamma}_c\mathbf{\Phi}^H) \odot \mathbf{T}, \tag{9}$$

where $\mathbf{\Gamma}_c = E[\boldsymbol{\gamma}_c\boldsymbol{\gamma}_c^H] = \mathrm{diag}(\mathbf{P})$ is a diagonal matrix, $\mathbf{P} = E[\boldsymbol{\gamma}_c \odot \boldsymbol{\gamma}_c^*]$ denotes the powers of $\boldsymbol{\gamma}_c$ and

$$\mathbf{T} = E[\mathbf{t}\mathbf{t}^H] = E\left[(\mathbf{t}_d \otimes \mathbf{1}_M)(\mathbf{t}_d \otimes \mathbf{1}_M)^H\right] \tag{10}$$

---

[1] The clutter rank can be estimated by counting the number of resolution grids that are occupied by the significant clutter spectrum components [28], which shows a prove for the conclusion in [33].

is a CMT.[2]

3 Proposed Sparsity-Based D3-STAP Algorithm

In this section, we detail the principles of the proposed sparsity-based D3-STAP algorithm and the related issues about parameter settings. Then, it is presented an automatic approach to adjust the CMT of the proposed algorithm. Finally, the performance metrics are discussed and introduced for the proposed algorithm.

3.1 Spatio-Temporal Spectrum Estimation

As the conventional sparsity-based D3-STAP algorithm, we first estimate the spatio-temporal profile $\boldsymbol{\gamma}_c$. Because $N_d N_s \gg NM$, the space-time steering dictionary $\boldsymbol{\Phi}$ is ill-conditioned, which results in the solution $\boldsymbol{\gamma}_c$ of $\mathbf{x}$ being an underdetermined inverse problem. In this case, the LS solution usually has a large norm and a poor performance. From the above section, we note that there is a high degree of sparsity in the target and clutter spatio-temporal profiles. Motivated by the recent developing SR techniques, we can obtain the estimates of spatio-temporal profiles by solving the minimization of the following objective function:

$$
\begin{aligned}
(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) &= \arg \min_{\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c} \mathcal{L}(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) \\
&= \arg \min_{\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c} \|\mathbf{x} - \mathbf{x}_t - \mathbf{x}_{ct}\|_2^2 + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p ,
\end{aligned}
\tag{11}
$$

where $0 \leq p \leq 1$, and κ represents the regularization parameters that provide a trade-off between the approximation error and the sparsity.

Substituting (7) into (11), we have

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) &= \left\| \mathbf{x} - \boldsymbol{\Phi}\boldsymbol{\gamma}_t - (\boldsymbol{\Phi}\boldsymbol{\gamma}_c) \odot (\mathbf{t}_d \otimes \mathbf{1}_M) \right\|_2^2 \\
&\quad + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p .
\end{aligned}
\tag{12}
$$

To solve (12) more effectively, we assume the target is also affected by the temporal decorrelation $\mathbf{t}_d$. Then, (12) becomes

$$
\begin{aligned}
\mathcal{L}'(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) &= \left\| \mathbf{x} - \boldsymbol{\Phi} \left(\boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right) \odot (\mathbf{t}_d \otimes \mathbf{1}_M) \right\|_2^2 \\
&\quad + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p .
\end{aligned}
\tag{13}
$$

The above approximation assumption is justified as follows: (1) the target power is usually much lower than the clutter power, resulting in the recovery the clutter being easier than the target, which means that the target may be not recovered when in small

---

[2] Note that the above equation makes use of the fact that $(\mathbf{A} \odot \mathbf{B}) (\mathbf{C} \odot \mathbf{D})^H = \left(\mathbf{A}\mathbf{C}^H \right) \odot \left(\mathbf{B}\mathbf{D}^H \right)$, where $\mathbf{A}, \mathbf{B}$ are $g \times h$ and $\mathbf{C}, \mathbf{D}$ are $e \times h$ matrices [7].

target case and the above approximation effects can be ignored. (2) If the target power is high or comparable to the clutter power, this approximation assumption will lead to the target energy spreading to the adjacent spatio-temporal grids. But the effects of this can be mitigated by the developed approach that will be discussed when estimating the clutter covariance in the parts below.

Conventionally, the amplitudes of the elements in the temporal decorrelation vector $\mathbf{t}_d$ caused by the ICM are assumed unities,[3] such as the commonly used Billingsley model [7] and the Gaussian model [26]. Define the matrix $\mathbf{Q}$ as

$$\mathbf{Q} = \mathrm{diag}(\mathbf{t}_d \otimes \mathbf{1}_M). \tag{14}$$

Then, $\mathbf{Q}$ is an $NM \times NM$ diagonal matrix and also orthonormal matrix. Thus, (13) can be rewritten with the form of (15),

$$
\begin{aligned}
\mathcal{L}'(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) &= \left\| \mathbf{Q} \left[\mathbf{Q}^{-1}\mathbf{x} - \mathbf{Q}^{-1}\boldsymbol{\Phi}\left(\boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c\right) \odot (\mathbf{t}_d \otimes \mathbf{1}_M) \right] \right\|_2^2 + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p \\
&= \left\| \mathbf{Q} \left[\mathbf{Q}^{-1}\mathbf{x} - \boldsymbol{\Phi}\left(\boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c\right) \odot (\mathbf{Q}^{-1}(\mathbf{t}_d \otimes \mathbf{1}_M)) \right] \right\|_2^2 + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p \\
&= \left\| \mathbf{Q} \left[\mathbf{Q}^{-1}\mathbf{x} - \boldsymbol{\Phi}\left(\boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c\right) \right] \right\|_2^2 + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p .
\end{aligned}
\tag{15}
$$

By exploiting the unitary invariance property of the l_2-norm,[4] we can discard the multiplication by $\mathbf{Q}$ in the l_2-norm term in (15) and obtain [6]

$$
\begin{aligned}
\min_{\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c} \mathcal{L}'(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c) &= \min_{\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c} \left\| \mathbf{Q}^{-1}\mathbf{x} - \boldsymbol{\Phi}\left(\boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c\right) \right\|_2^2 + \kappa \left\| \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c \right\|_p \\
&= \min_{\tilde{\boldsymbol{\gamma}}} \| \tilde{\mathbf{x}} - \boldsymbol{\Phi}\tilde{\boldsymbol{\gamma}} \|_2^2 + \kappa \| \tilde{\boldsymbol{\gamma}} \|_p .
\end{aligned}
\tag{16}
$$

Here, $\tilde{\boldsymbol{\gamma}} = \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_c$ and $\tilde{\mathbf{x}} = \mathbf{Q}^{-1}\mathbf{x}$. It is found that the the minimization of the objective function $\mathcal{L}'(\boldsymbol{\gamma}_t, \boldsymbol{\gamma}_c)$ has a similar formulation as that presented by articles [15,16,18,21–23,25,32–34] when there is no subspace leakage. However, since there are unknown parameters $\mathbf{t}_d$ in $\tilde{\mathbf{x}}$, the above problem cannot be directly solved by the SR techniques.

Next, we will show how to use the FOCUSS algorithm to estimate $\tilde{\boldsymbol{\gamma}}$ with the unknown temporal decorrelation term $\mathbf{t}_d$. The principle of the FOCUSS algorithm is to use the weighted l_2-norm minimization to operate recursive adjustments to the weighting matrix until most elements of the solution are close to zero [12]. The basic form of the FOCUSS algorithm for problem (16) is composed of the following two steps:

---

[3] This assumption is also seen in CMT methods developed by Guerci in [8].

[4] This is where we exploit both the orthogonality of $\mathbf{Q}$ and the white Gaussianity of the thermal noise.

$$\text{Step1}: \quad \mathbf{W}_q = \text{diag}(|\tilde{\boldsymbol{\gamma}}_{q-1}|^{1-p/2}), \ 0 \le p \le 1 \tag{17}$$

$$\text{Step2}: \quad \tilde{\boldsymbol{\gamma}}_q = \mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\tilde{\mathbf{x}}, \tag{18}$$

where $(\mathbf{A})^{\dagger} = \mathbf{A}^H \left(\mathbf{A}\mathbf{A}^H\right)^{-1}$ denotes the pseudo-inverse operation of matrix $\mathbf{A}$. Since there is an unknown parameter $\mathbf{Q}$ in $\tilde{\mathbf{x}}$, it can not directly use the above iterations to estimate $\tilde{\boldsymbol{\gamma}}$. We rewrite $|\tilde{\boldsymbol{\gamma}}_q|$ as (19),

$$
\begin{aligned}
\left|\tilde{\boldsymbol{\gamma}}_q\right| &= \left|\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\tilde{\mathbf{x}}\right| \\
&= \left\{ \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\tilde{\mathbf{x}}\right) \odot \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\tilde{\mathbf{x}}\right)^{*} \right\}^{1/2} \\
&= \left\{ \left[\left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\right) \odot \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\right)^{*}\right] (\tilde{\mathbf{x}} \odot \tilde{\mathbf{x}}^{*}) \right\}^{1/2} \\
&= \left\{ \left[\left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\right) \odot \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\right)^{*}\right] (\mathbf{x} \odot \mathbf{x}^{*}) \right\}^{1/2} \\
&= \left\{ \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\mathbf{x}\right) \odot \left(\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\mathbf{x}\right)^{*} \right\}^{1/2} \\
&= \left|\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\mathbf{x}\right|.
\end{aligned}
\tag{19}
$$

From (19), we observe that $\mathbf{W}_q$ is independent of the temporal decorrelation term $\mathbf{t}_d$ and can be uniquely determined by the received snapshot $\mathbf{x}$ and $\boldsymbol{\Phi}$. In other words, the presented FOCUSS algorithm, termed as CMT-FOCUSS algorithm, can be operated in an iterative way with the following two main steps:

$$\text{Step1}: \quad \mathbf{W}_q = \text{diag}(|\boldsymbol{\gamma}_{q-1}|^{1-p/2}), \ 0 \le p \le 1 \tag{20}$$

$$\text{Step2}: \quad \boldsymbol{\gamma}_q = \mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\mathbf{x}, \tag{21}$$

The final solution $\tilde{\boldsymbol{\gamma}}_q$ is updated by

$$
\begin{aligned}
\tilde{\boldsymbol{\gamma}}_q &= \mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\tilde{\mathbf{x}} \\
&= (\mathbf{W}_q(\boldsymbol{\Phi}\mathbf{W}_q)^{\dagger}\mathbf{x}) \odot (\mathbf{t}_d \otimes \mathbf{1}_M) = \boldsymbol{\gamma}_q \odot (\mathbf{t}_d \otimes \mathbf{1}_M).
\end{aligned}
\tag{22}
$$

The iterative process terminates when certain criteria are satisfied, e.g., when the iteration number achieves a preset limit $q_{\max}$, or when the relative change in $\boldsymbol{\gamma}$ between consecutive iterations being sufficiently small: $\left|(\boldsymbol{\gamma}_q - \boldsymbol{\gamma}_{q-1})/\boldsymbol{\gamma}_q\right| \le \epsilon$ (where ϵ is a small positive number). Here, we keep the representation of $\tilde{\boldsymbol{\gamma}}_q$ including the unknown term $\mathbf{t}_d$. In the following, we will show that it can estimate the clutter covariance matrix by estimating the CMT instead of $\mathbf{t}_d$.

Similar as the derivation of the above proposed CMT-FOCUSS algorithm, other versions of the FOCUSS algorithms, such as regularized FOCUSS [3,19] and adaptive regularized FOCUSS [29], can also be applied to solve for problem (16). For space limitation, further details about parameters setting and implementation issues of the FOCUSS algorithm are referred to see [3,12,19,29].

3.2 Target Detection

After obtaining the spatio-temporal spectrum in the CUT with high accuracy, it is possible to make a direct amplitude detection in the spatio-temporal domain. However, this relies on many conditions, such as a high target's signal-to-noise ratio (SNR), very small off-grid errors and sufficient noncoherence of the overcomplete dictionary. In fact, the columns in the overcomplete dictionary are highly correlated because of $N_s N_d \gg NM$. Off-grid errors always exist for target and clutter estimation. It is also hard to make sure there is a high target's SNR. Therefore, we cannot guarantee a high-accuracy target estimation from a surrounding strong clutter using SR. But we can expect to recover significant clutter components, resulting in a good estimation of the clutter subspace. This is because the clutter to noise is always high and we care much more about the recovery of the whole clutter subspace than the exact positions of the clutter components. We can only extract the clutter components to form the adaptive filter to suppress the clutter and then detect the target. This extraction requires the prior knowledge of the assumed target spatio-temporal frequencies, which is available and is also used in the STAP algorithms employing the generalized sidelobe canceler (GSC-STAP) [11,35], the D3-LS STAP [24] and the MLED [1,4,5]. However, the differences among them lie that: In GSC-STAP, it uses this prior knowledge to form the signal blocking matrix; in D3-LS STAP, it uses this prior knowledge to eliminate the influence of the target signal in the direct data domain; in MLED, it uses this prior knowledge to form the estimated covariance matrix. Unlike all the algorithms mentioned above, we use this prior knowledge to exclude the target signal in the recovered spatio-temporal domain, as also done in [22].

We first determine the signal of interest area Ω_{SOI} in the spatio-temporal plane using prior knowledge of the target signal as

$$\Omega_{\text{SOI}} = \left\{ (i_{n_1}, k_{n_1}), (i_{n_2}, k_{n_2}), \ldots, (i_{n_{N_{\text{SOI}}}}, k_{n_{N_{\text{SOI}}}}) \right\}, \tag{23}$$

where $(i_{n_m}, k_{n_m}), 1 \le m \le N_{\text{SOI}}$ denote the possible indexes of the target signal in the discretized spatio-temporal plane. The size of N_{SOI} reflects the uncertainty or the energy spread along the spatial and temporal frequency axes. Then, the estimated clutter spatio-temporal profile $\tilde{\boldsymbol{\gamma}}_c$ is given by

$$\tilde{\gamma}_{c;i,k} = \begin{cases} \tilde{\gamma}_{i,k} & \text{for } (i,k) \notin \Omega_{\text{SOI}} \\ 0 & \text{for } (i,k) \in \Omega_{\text{SOI}} \end{cases}, \tag{24}$$

where $\tilde{\gamma}_{c;i,k}$ and $\tilde{\gamma}_{i,k}$ are the (i,k)th elements in $\tilde{\boldsymbol{\gamma}}_c$ and $\tilde{\boldsymbol{\gamma}}$, respectively. Through this operation, the effects caused by the assumption of the target signal in (13) can be mitigated in a sense. Then, the parameter $\boldsymbol{\Gamma}_c$ in the clutter covariance matrix $\mathbf{R}_{ct}$ can be estimated by

$$\hat{\boldsymbol{\Gamma}}_c = \text{diag}(\hat{\mathbf{P}}) = \text{diag}\left(\tilde{\boldsymbol{\gamma}}_c \odot \tilde{\boldsymbol{\gamma}}_c^*\right) = \text{diag}\left(\boldsymbol{\gamma}_c \odot \boldsymbol{\gamma}_c^*\right) \tag{25}$$

where it uses the fact that the amplitudes of the elements in $\mathbf{t}_d$ are assumed unities and $\boldsymbol{\gamma}_c$ represents the clutter spatio-temporal profile directly computed by $\mathbf{x}$, given by

$$\gamma_{c;i,k} = \begin{cases} \gamma_{i,k} & \text{for } (i,k) \notin \Omega_{\text{SOI}} \\ 0 & \text{for } (i,k) \in \Omega_{\text{SOI}} \end{cases}. \tag{26}$$

Thus, the clutter covariance matrix estimate can be calculated as

$$\hat{\mathbf{R}}_{ct} = (\mathbf{\Phi}\hat{\mathbf{\Gamma}}_c\mathbf{\Phi}^H) \odot \hat{\mathbf{T}}, \tag{27}$$

where $\hat{\mathbf{T}}$ denotes the CMT estimate caused by the temporal decorrelation. From (27), we find that the CMT only operates on the clutter components not on the target. That is to say the approximation assumption in (13) is reasonable, and target energy spreading effects are mitigated by excluding the target signal in Ω_{SOI}.

Note that in $\hat{\mathbf{R}}_{ct}$, it requires an estimate of $\hat{\mathbf{T}}$, which have been addressed in a variety of articles, such as [7,8,10,17,37]. For a land scenario, we can select the Billingsley model [7] to describe the temporal decorrelation. The only parameters required to specify the clutter Doppler power spectrum are essentially the operating wavelength and wind speed. The operating wavelength is usually known, while the wind speed should be estimated. For a water scenario, we prefer to select the Gaussian model presented by J. Ward in [26] to estimate $\hat{\mathbf{T}}$. The temporal autocorrelation of the fluctuations for this model is Gaussian in shape with the form:

$$\zeta(m) = \exp\left\{-\frac{8\pi^2\sigma_v^2 T_r^2}{\lambda_c^2}m^2\right\}, \tag{28}$$

where σ_v^2 is the variance of the clutter spectral spread in m^2/s^2. The CMT $\mathbf{T}(\sigma_v)$ is formulated as the form of

$$\mathbf{T}(\sigma_v) = \text{Toeplitz}(\zeta(0), \zeta(1)\ldots, \zeta(N-1)) \otimes \mathbf{1}_{M \times M}. \tag{29}$$

In the following simulations, we consider the CMT model of the latter one. It should be noted that the proposed algorithm firstly estimates the clutter covariance matrix in exactly the same manner as the existing sparsity-based D3-STAP without CMT, e.g., the method in [22] and then obtains the final clutter statistics by a Hadamard product of the CMT and the previously estimated clutter covariance matrix. Therefore, compared with the existing sparsity-based D3-STAP without CMT, the additional computational complexity of the proposed algorithm is due to the Hadamard product and the CMT estimation, i.e., on the order of $O((NM)^2)$, resulting in a modest complexity increase. However, the values of the estimated CMT will affect the performance of the proposed algorithm. If there is no prior knowledge of the clutter, an automatically method must be developed, which will be discussed in the following Sect. 3.4.

Finally, we design the STAP filter based on a linearly constrained minimum variance (LCMV) approach, which is to minimize the clutter plus noise output power while constraining the gain in the direction of the desired target signal. The optimal LCMV STAP filter weight vector is given by

$$\hat{\mathbf{w}} = \frac{\left(\hat{\mathbf{R}}_{ct} + \beta_L \mathbf{I}\right)^{-1}\mathbf{s}}{\mathbf{s}^H \left(\hat{\mathbf{R}}_{ct} + \beta_L \mathbf{I}\right)^{-1}\mathbf{s}}. \tag{30}$$

where $\mathbf{s}$ is the $NM \times 1$ space-time steering vector in the target direction and β_L denotes the diagonal loading factor related to the noise level σ_n^2 (which can be estimated by the receiver when the radar transmitter operates in passive mode [14]. Thus, the STAP filter output is computed by

$$y = \hat{\mathbf{w}}^H \mathbf{x}. \tag{31}$$

3.3 Parameter Settings

There are four important parameters, namely the size of the dictionary $N_s N_d$, the target of interest area Ω_{SOI}, the diagonal loading term β_L and the CMT, that must be estimated or set somehow in the proposed algorithm. Regarding the CMT, it will develop an adaptive approach to reduce the estimating difficulty in the next subsection. Regarding the diagonal loading term β_L, simulation results will show that the proposed algorithm is robust to a range of values of β_L. In the following, we focus on the discussions about the settings of $N_s N_d$ and Ω_{SOI}.

The issue of the the size of the dictionary is also called the grid sampling issue, which is critical important to the performance of the sparsity-based STAP algorithms. All the sparsity-based STAP algorithms should be careful to set that. Intuitively, the denser the gird sampling of the spatio-temporal plane, the better approximation the described sparse model but the higher the computational complexity required by sparse recovery. Thus, it is a trade-off to choose a suitable gird sampling. We will conduct discussions from two aspects.

(1) *From a point view of correlation:*

For a point with temporal frequency f_d and spatial frequency f_s, the space-time steering vector is

$$\mathbf{v}(f_d, f_s) = \mathbf{v}_t(f_d) \otimes \mathbf{v}_s(f_s). \tag{32}$$

The correlation absolute value between $\mathbf{v}(f_d, f_s)$ and $\mathbf{v}(f_{d,i}, f_{s,k})$ in the space-time steering dictionary is

$$\begin{aligned}\rho &= \frac{\left|\mathbf{v}^H(f_{d,i}, f_{s,k})\mathbf{v}(f_d, f_s)\right|}{NM} \\ &= \left|\frac{\sin \pi M(f_s - f_{s,k})}{M \sin \pi(f_s - f_{s,k})} \cdot \frac{\sin \pi N(f_d - f_{d,i})}{N \sin \pi(f_d - f_{d,i})}\right|.\end{aligned} \tag{33}$$

Without loss of generality, Fig. 1a plots the correlation absolute bias value with $f_d = 0$ and $f_s = 0$. It is shown that the correlation absolute value keeps high when constraining in a diffraction-limited resolution cell and decays quickly elsewhere. In other words, for a component with f_d and f_s, its power mostly lies in the corresponding diffraction-limited resolution cell. The smaller the bias between the true

spatio-temporal frequencies and the sampling spatio-temporal frequencies, the better the approximation using the sampling spatio-temporal frequencies. To have a further intuitive understanding, Fig. 1b plots the maximal correlation absolute bias value caused by different grid sampling rates, which is shown that the correlation absolute bias value is very small when $N_s \geq 4M$ and $N_d \geq 4N$. From this point of view, we can expect to recover the clutter or target signal approximately when $N_s \geq 4M$ and $N_d \geq 4N$.

(2) *From a point view of the proposed strategy:*

For the clutter, we care much more about the recovery of the whole clutter subspace than the exact positions of the clutter components. In another word, if one of the clutter components does not exactly lie on the grid, its main power will spread into the most nearest grids centered at its own spatio-temporal frequency. After the sparse recovery of the clutter, we do not know the exactly clutter spatio-temporal frequencies but still can have a good estimation of the clutter subspace. From this point of view, the influence of the gird-off issue of the proposed algorithm is much less important than that in the field of direction-of-arrival estimation. Regarding how to sample the grid points, Melvin and Showman pointed out that it was important to oversample relative to the diffraction-limited case [17]. Moreover, they provided a good rule of thumb based on numerical simulations that it is to set the spacing of the steering vectors to about twenty to thirty percent of the diffraction-limited resolution. That is to say we should sample the grids with N_s about 4–5 times of M and N_d about 4–5 times of N. Additionally, the influence of the grid sampling to the performance of the sparsity-based STAP algorithms has been conducted in the papers, namely [34] and [31]. It was also shown that the performance improvement was very little based on the simulations when $Ns \geq 4M$ and $N_d \geq 4N$. The reason for that is: when $N_s N_d$ is not large enough, there are serious mismatches between the sampling space-time steering vectors and those of the true clutter; when $N_s N_d$ is sufficiently large, the accuracy of the sparse recovery solution is hardly to improve because of the existence of the noise and high correlations among the sampling space-time steering vectors. Therefore, it is empirical to set N_s about 4–5 times of M and N_d about 4–5 times of N.

For the target, if the target does not lie in the grid, the power of the target spreads to the nearest spatio-temporal frequencies. In the proposed algorithm, we have considered about this situation and estimate the clutter covariance matrix by excluding spatio-temporal frequencies in the region Ω_{SOI}. From the discussions about Fig. 1a, b, we can set the size of the region Ω_{SOI} is about a diffraction-limited resolution cell. Furthermore, because the off-grid issue is an important one to the target recovery performance, especially in a strong clutter environment, we choose an alternative approach to detect the target. We consider that the recovering of the target component is not an easy task and the direct estimation of the target from the surrounding strong clutter may be unreliable [22]. Therefore, we only extract the significant clutter components and exclude the unstable target components with the prior knowledge of the assumed target spatio-temporal frequencies to form the adaptive filter. The above reasons result that the proposed algorithm is relatively robust to the off-grid issue in the target recovery.
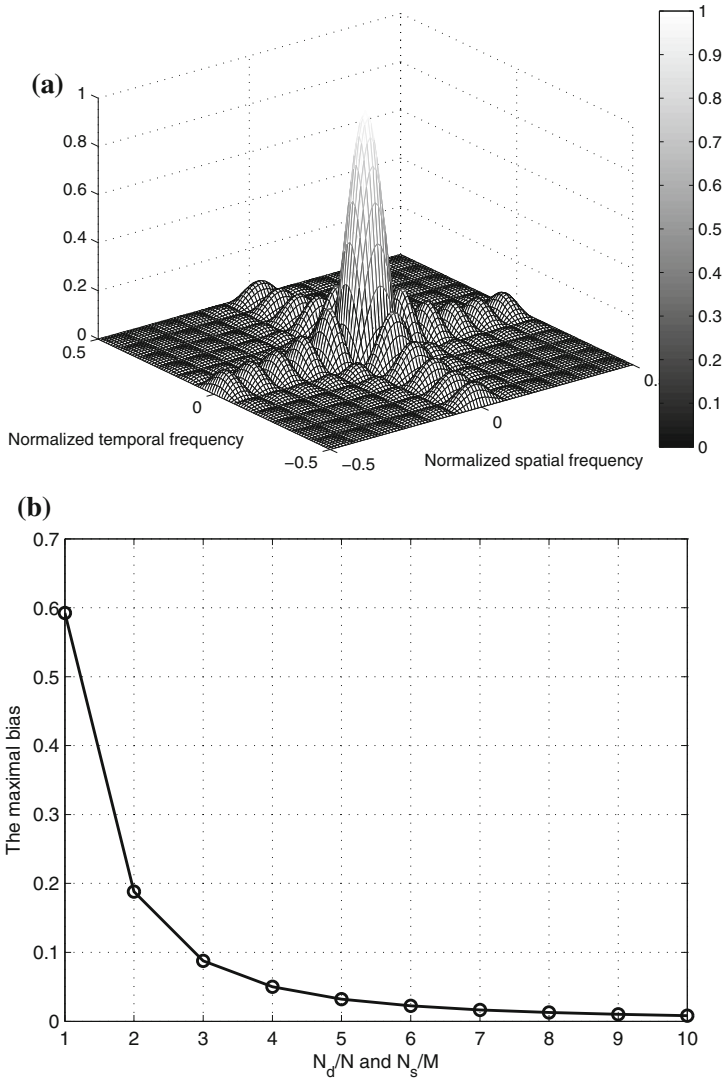
Fig. 1 Influence of the grid sampling. **a** ρ with $f_d = 0$ and $f_s = 0$ ($N = 12$, $M = 12$). **b** The maximal correlation absolute bias value

3.4 CMT Adaptation

The performance of the proposed approach to estimate the clutter covariance matrix described in the previous subsections depends on the CMT **T**. For the Gaussian model, from (29), the CMT **T** is related to the clutter spectral spread variance σ_v^2.

Here, we present an adaptation method for automatically selecting the parameter σ_v of the proposed approach. Specifically, we first constrain the parameter σ_v within a range of appropriate values (the candidates space of the parameter σ_v is presented by $\Omega_{\sigma_v} = \{\sigma_{v;k}|k = 1, 2, \ldots, K\}$). In fact, for a typical environment, the typical region of the clutter spread deviation can be roughly estimated by the knowledge of the general clutter type and wind state (for the Gaussian CMT model, the spectral standard deviation is $2\sigma_v/\lambda_c$ Hz) [26]. This task is far easier than directly estimate the value of the clutter spread deviation. Furthermore, even in the case without any knowledge about the environments, we can set a roughly large value of the region of the clutter spread deviation. But this may lead to the increase of the number of candidates in the region of the clutter spread deviation.

In order to select σ_v in a more appropriate way, we operate this procedure over a localized region by L adjacent range snapshots in the CUT. Let $\{\mathbf{x}_l\}_{l=1}^{L}$ denote the adjacent range snapshots. Then, the best selection of σ_v relies on minimizing the STAP filter output with the cost function $C(\Omega_{\sigma_v})$:

$$\sigma_{v;\text{opt}} = \arg \min_{\sigma_v} C(\Omega_{\sigma_v})$$

$$= \arg \min_{\sigma_v} \frac{1}{L} \sum_{l=1}^{L} \left| \hat{\mathbf{w}}^H(\sigma_{v;k})\mathbf{x}_l \right|^2$$

$$= \arg \min_{\sigma_v} \frac{1}{L} \sum_{l=1}^{L} \left| \frac{\mathbf{s}^H \left(\hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2\mathbf{I} \right)^{-1} \mathbf{x}_l}{\mathbf{s}^H \left(\hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2\mathbf{I} \right)^{-1} \mathbf{s}} \right|^2, \qquad (34)$$

where

$$\hat{\mathbf{R}}_{ct}' = \mathbf{\Phi}\hat{\mathbf{\Gamma}}_c\mathbf{\Phi}^H. \qquad (35)$$

The above criterion is used because the STAP filters satisfy $\hat{\mathbf{w}}^H(\sigma_{v;k})\mathbf{s} = 1$ (where $1 \leq k \leq K$) and minimizing the STAP filter output $|\hat{\mathbf{w}}^H(\sigma_{v;k})\mathbf{x}_l|^2$ is equivalent to minimizing the clutter components in the final output. Then the final clutter covariance matrix estimate can be calculated by

$$\hat{\mathbf{R}}_{ct} = \hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;\text{opt}}) = (\mathbf{\Phi}\hat{\mathbf{\Gamma}}_c\mathbf{\Phi}^H) \odot \mathbf{T}(\sigma_{v;\text{opt}}). \qquad (36)$$

Since the computational complexity of the proposed CMT adaptation approach depends on the number of candidates in Ω_{σ_v}, the larger K, the higher computational complexity. Thus, we expect to use only a few candidates to achieve good performance. In the simulations, it will demonstrate that the resulting SINR is relatively invariant to errors in the estimate of the CMT $\mathbf{T}$ and a few candidates will be enough to obtain an acceptable SINR.

In addition, it should be noted that for each $\sigma_{v;k} \in \Omega_{\sigma_v}$, a matrix inversion operation is required in (34). Below, we will develop a fast approach to compute (34), which can be divided into two steps: the first is to compute the eigenvectors and eigenvalues of $\hat{\mathbf{R}}_{ct}'$ and the second is to compute the matrix inversion of $\hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2\mathbf{I}$.

Step 1: Compute the eigenvectors and eigenvalues of $\hat{\mathbf{R}}_{ct}'$.

Because

$$\hat{\mathbf{R}}'_{ct} = \mathbf{\Phi}\hat{\mathbf{\Gamma}}_c\mathbf{\Phi}^H$$

$$= \sum_{(i,k)=(1,1)}^{N_d,N_s} \hat{P}_{i,k}\mathbf{v}(f_{d,i}, f_{s,k})\mathbf{v}^H(f_{d,i}, f_{s,k}). \tag{37}$$

Since there is a high degree of sparsity in $\hat{P}_{i,k}$ (in other words, there will be only a small number of significant elements in $\hat{P}_{i,k}$[5]), we can approximate $\hat{\mathbf{R}}'_{ct}$ by

$$\hat{\mathbf{R}}'_{ct} \approx \hat{\mathbf{R}}''_{ct} = \sum_{(i,k)=\Omega_P} \hat{P}_{i,k}\mathbf{v}(f_{d,i}, f_{s,k})\mathbf{v}^H(f_{d,i}, f_{s,k}), \tag{38}$$

where Ω_P denotes the index set of significant elements in $\hat{P}_{i,k}$. From (38), we note that the eigenvector subspace (denoted as $\mathbf{U}$) of $\hat{\mathbf{R}}''_{ct}$ is decided by vectors $\mathbf{v}(f_{d,i}, f_{s,k}), (i, k) = \Omega_P$. For further reducing complexity, we can take the Gram–Schmidt orthogonalization procedure on these vectors to compute the eigenvector subspace instead of singular value decomposition (SVD), where the implementation steps of the Gram–Schmidt orthogonalization are omitted here, and interested readers are referred to [13] for further details. Then, the corresponding eigenvalues are also can be easily calculated by $\lambda_q = \mathbf{u}_q^H\hat{\mathbf{R}}''_{ct}\mathbf{u}_q, 1 \le q \le N_{\Omega_P}$.

Step 2: Compute the matrix inversion of $\hat{\mathbf{R}}'_{ct} \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2\mathbf{I}$.

For each $\sigma_{v;k} \in \Omega_{\sigma_v}$, we first compute and prestore the eigenvalues (denotes as $\lambda_{\mathbf{T}(\sigma_{v;k});q}$) and eigenvectors (denotes as $\mathbf{u}_{\mathbf{T}(\sigma_{v;k});q}$) of $\mathbf{T}(\sigma_{v;k})$, where $1 \le q \le NM$. Note that $\mathbf{T}(\sigma_{v;k})$ has a block Toeplitz structure, which can be exploited to reduce the complexity of eigenvalue decomposition [13]. Substituting the computed eigenvalues and eigenvectors of $\hat{\mathbf{R}}''_{ct}$ and $\mathbf{T}(\sigma_{v;k})$, we have (39),

$$\hat{\mathbf{R}}'_{ct} \odot \mathbf{T}(\sigma_{v;k}) \approx \hat{\mathbf{R}}''_{ct} \odot \mathbf{T}(\sigma_{v;k})$$

$$= \left(\sum_{i=1}^{N_{\Omega_P}} \lambda_i\mathbf{u}_i\mathbf{u}_i^H\right) \odot \left(\sum_{j=1}^{N_{NM}} \lambda_{\mathbf{T}(\sigma_{v;k});j}\mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}\mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}^H\right)$$

$$= \sum_{i=1}^{N_{\Omega_P}}\sum_{j=1}^{N_{NM}} \lambda_i\lambda_{\mathbf{T}(\sigma_{v;k});j}\left(\mathbf{u}_i\mathbf{u}_i^H\right) \odot \left(\mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}\mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}^H\right)$$

$$= \sum_{i=1}^{N_{\Omega_P}}\sum_{j=1}^{N_{NM}} \lambda_i\lambda_{\mathbf{T}(\sigma_{v;k});j}\left(\mathbf{u}_i \odot \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}\right)\left(\mathbf{u}_i \odot \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}\right)^H. \tag{39}$$

Let $\tilde{u}_{i,j} = \mathbf{u}_i \odot \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}$, and we note that vectors $\tilde{u}_{i,j}$ retain the orthonormality property (orthogonal and unit norm), proved as follows:

---

[5] Significant elements are defined by those whose powers are higher than the thermal noise level.

$$
\begin{aligned}
\tilde{u}_{i,j}^H \tilde{u}_i', j' &= \left(\mathbf{u}_i \odot \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}\right)^H \left(\mathbf{u}_{i'} \odot \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j'}\right) \\
&= \left(\mathbf{u}_i^H \mathbf{u}_i'\right) \odot \left(\mathbf{u}_{\mathbf{T}(\sigma_{v;k});j}^H \mathbf{u}_{\mathbf{T}(\sigma_{v;k});j'}\right) \\
&= \delta_{ii'}' \delta_{jj}'
\end{aligned}
\tag{40}
$$

where $\delta_{ii'}$ and $\delta_{jj'}$ denote the Kronecker delta function. Therefore, the matrix inversion of $\hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;k})$ can be calculated as

$$
\begin{aligned}
&\left(\hat{\mathbf{R}}_{ct}' \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2 \mathbf{I}\right)^{-1} \\
&\approx \left(\hat{\mathbf{R}}_{ct}'' \odot \mathbf{T}(\sigma_{v;k}) + \hat{\sigma}_n^2 \mathbf{I}\right)^{-1} \\
&= \left(\sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j} \tilde{u}_{i,j} \tilde{u}_{i,j}^H + \hat{\sigma}_n^2 \mathbf{I}\right)^{-1} \\
&= \frac{1}{\hat{\sigma}_n^2}\left(\mathbf{I} - \sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \frac{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j}}{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j} + \hat{\sigma}_n^2} \tilde{u}_{i,j} \tilde{u}_{i,j}^H\right).
\end{aligned}
\tag{41}
$$

Substituting (41) into (34), we obtain

$$
C(\Omega_{\sigma_v}) = \frac{1}{L} \sum_{l=1}^{L}
$$
$$
\left| \frac{\mathbf{s}^H \mathbf{x}_l - \sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \frac{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j}}{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j} + \hat{\sigma}_n^2} \left(\tilde{u}_{i,j}^H \mathbf{s}\right)\left(\tilde{u}_{i,j}^H \mathbf{x}_l\right)}{1 - \sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \frac{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j}}{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;k});j} + \hat{\sigma}_n^2} \left|\tilde{u}_{i,j}^H \mathbf{s}\right|^2} \right|^2.
\tag{42}
$$

From the above equation, we observe that (34) requires to compute the matrix inversion for each $\sigma_{v;k} \in \Omega_{\sigma_v}$, while (42) only needs to calculate simple multiplications using the eigenvalues and eigenvectors of the matrix $\hat{\mathbf{R}}_{ct}'$ and the prestored eigenvalues and eigenvectors of $\mathbf{T}(\sigma_{v;k})$. At last, the final STAP filter weight vector $\hat{\mathbf{w}}$ in (30) can be rewritten as

$$
\hat{\mathbf{w}} = \frac{\mathbf{s} - \sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \frac{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;\text{opt}});j}}{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;\text{opt}});j} + \hat{\sigma}_n^2} \left(\tilde{u}_{i,j}^H \mathbf{s}\right) \tilde{u}_{i,j}}{1 - \sum_{i=1}^{N_{\Omega P}} \sum_{j=1}^{N_{NM}} \frac{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;\text{opt}});j}}{\lambda_i \lambda_{\mathbf{T}(\sigma_{v;\text{opt}});j} + \hat{\sigma}_n^2} \left|\tilde{u}_{i,j}^H \mathbf{s}\right|^2}.
\tag{43}
$$

3.5 Performance Metrics

Here, we detail two traditional performance metrics, i.e., SINR loss and SINR improvement factor, for analyzing the performance of the proposed sparsity-based STAP algorithm. The output SINR loss is defined as the output SINR performance relative to the matched filter SNR$_{\text{opt}}$ in an interference-free environment (where the

SNR_{opt} is equivalent to the number of system DOFs NM multiplying the target power $\sigma_t^2 = E[|\alpha_t|^2]$, i.e., $NM\sigma_t^2$) [26], given as

$$\text{SINR}_{\text{loss}} = \frac{\text{SINR}}{\text{SNR}_{\text{opt}}} = \frac{\left|\hat{\mathbf{w}}^H \mathbf{s}\right|^2}{NM\left|\hat{\mathbf{w}}^H \mathbf{R}\hat{\mathbf{w}}\right|}, \tag{44}$$

where $\mathbf{R}$ is the true clutter plus noise covariance matrix. As pointed in [21], there is no closed-form statistical characterizations for the estimated clutter covariance matrix when using the sparsity-based STAP algorithms. Therefore, we use a Monte Carlo technique to generate the SINR loss performance.

The SINR improvement factor, IF, is as a gain in SINR relative to the input SINR SINR_{in} on a single channel and a single pulse and defined as

$$\text{IF} = \frac{\text{SINR}}{\text{SINR}_{\text{in}}} = \frac{(1 + \text{CNR})\left|\hat{\mathbf{w}}^H \mathbf{s}\right|^2}{\left|\hat{\mathbf{w}}^H \mathbf{R}\hat{\mathbf{w}}\right|}, \tag{45}$$

where CNR is defined by

$$\text{CNR} = \frac{\text{tr}\,(\mathbf{R}_{ct})}{\sigma_n^2 NM}. \tag{46}$$

Additionally, we define the target's SNR as [26]

$$\text{SNR} = \frac{E[\mathbf{x}_t^H \mathbf{x}_t]}{\sigma_n^2 NM}, \tag{47}$$

which is the ratio of the input target power to the input noise. The SINR improvement factor provides not only the amount of clutter rejection but also the coherent gain on target due to spatial and temporal beamforming [26].

In addition, to study the effect of various parameters on the SINR, we employ the average value of the SINR loss or SINR improvement factor over all Doppler frequencies as composite performance metrics given by

$$\overline{\text{SINR}}_{\text{loss}} = \int_{-0.5}^{0.5} \text{SINR}_{\text{loss}}(f)df. \tag{48}$$

$$\overline{\text{IF}} = \int_{-0.5}^{0.5} \text{IF}(f)df. \tag{49}$$

Similar to the SINR loss, we use a Monte Carlo technique to generate IF, $\overline{\text{SINR}}_{\text{loss}}$ and $\overline{\text{IF}}$.

Table 1 Radar system parameters

Parameter	Value
Antenna array	Side-looking ULA
Antenna array spacing	$\lambda_c/2$
Number of elements in ULA	12
Number of coherent pulses	12
Carrier wavelength	$\lambda_c = 0.3\,\mathrm{m}$
Instantaneous bandwidth	5 MHz
Transmit taper	Uniform
Pulse repetition frequency (PRF)	4000 Hz
Platform velocity	300 m/s
Platform height	3000 m
CNR	50 dB
The target AOA	0°
The target signal-to-noise ratio (SNR)	0 dB

4 Numerical Results

In this section, we assess the output SINR and P_d performance of the proposed algorithm (shortened as CMT-FOCUSS-D3-STAP algorithm[6]) using simulated data and compare it with the conventional sample matrix inversion algorithm (SMI), the D3-LS STAP method in [24], the FOCUSS-D3-STAP algorithm in [22] and the sparsity-aware beamformer in [35]. Throughout the simulations, unless otherwise stated, the parameters of the simulated scenarios are given in Table 1. The thermal noise is modeled as a Gaussian white noise with unity power. The clutter and target powers are referred to the thermal noise power. In the simulations, for the sparsity-based STAP algorithms, we set $N_d = 4N$ and $N_s = 4M$ for the discretized spatio-temporal plane and set the diagonal loading factor β_L to the noise level in the STAP filter design. We set the l_p-norm to $p = 1$, the maximum iteration number is 500, and the stopping criterion is decided by the preset limit relative change of the solution between two adjacent iterations 10^{-4}. For the D3-LS STAP algorithm, the numbers of temporal and spatial DOFs used to design the filter weights are $N' = 7$ and $M' = 7$, respectively. For the SMI algorithm, the number of snapshots without target is $2NM = 288$. For the sparsity-aware beamformer algorithm, the number of snapshots without a target is 100 and the regularization parameter value is set to 0.3. All presented results are averaged over 300 independent Monte Carlo runs.

---

[6] Regarding the FOCUSS algorithm, the MATLAB code can be downloaded at http://dsp.ucsd.edu/~zhilin/MFCOUSS.m.
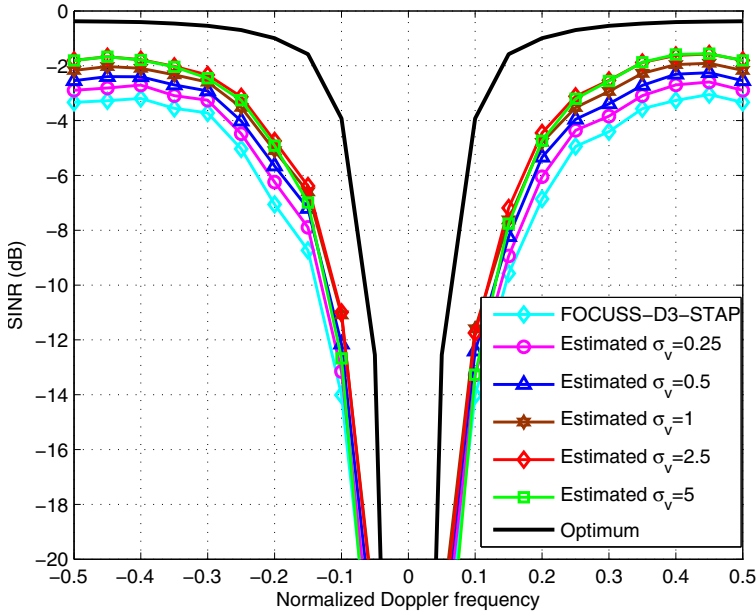
Fig. 2 SINR loss against target Doppler frequency of the CMT-FOCUSS-D3-STAP algorithm with different estimated CMT

4.1 Effects of Parameter Settings on Performance

In this subsection, we focus on the effects of parameter settings on the performance of the proposed algorithm, such as the estimated CMT, the diagonal loading term β_L and the target region Ω_{SOI}.

In the first example, we evaluate the effect of the estimated CMT on the SINR performance for the proposed algorithm. We consider the ICM with $\sigma_v = 1$ and assume the estimated CMT with six different values, i.e., $\hat{\sigma}_v = 0, 0.25, 0.5, 1, 2.5, 5$, where $\hat{\sigma}_v = 0$ corresponds to the FOCUSS-D3-STAP algorithm. Figure 2 depicts the SINR loss against normalized target Doppler frequency. The curves suggest that the proposed CMT-FOCUSS-D3-STAP algorithm outperforms the conventional FOCUSS-D3-STAP algorithm, and the resulting SINR is relatively invariant to errors in the estimate of the CMT, e.g., the proposed algorithms with $\hat{\sigma}_v = 1, 2.5, 5$ provide close performance.

In the second example, we access the effect of the assumed target region Ω_{SOI} on the performance of the proposed algorithm. We consider the ICM with $\sigma_v = 1$ and assume the estimated CMT of the proposed algorithm with $\hat{\sigma}_v = 2$. As shown in Fig. 3, we see that it provides the best performance when the assumed target region Ω_{SOI} is a diffraction-limited resolution cell, which validates the analysis in Sect. 3.3.

In the third example, we consider the sensitiveness of the diagonal loading term β_L in the filter (30) to the performance of the proposed algorithm. The assumed target region Ω_{SOI} is set to be a diffraction-limited resolution cell, and other simulation parameters are the same as those in the second example, except for β_L. The
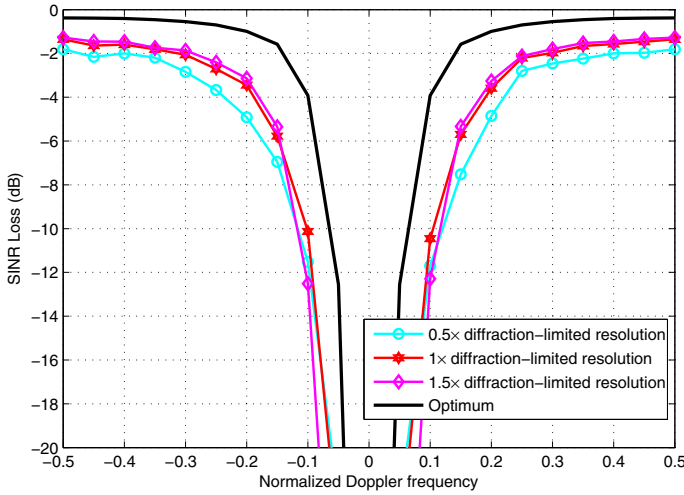
Fig. 3 SINR loss against target Doppler frequency of the CMT-FOCUSS-D3-STAP algorithm with different Ω_{SOI}
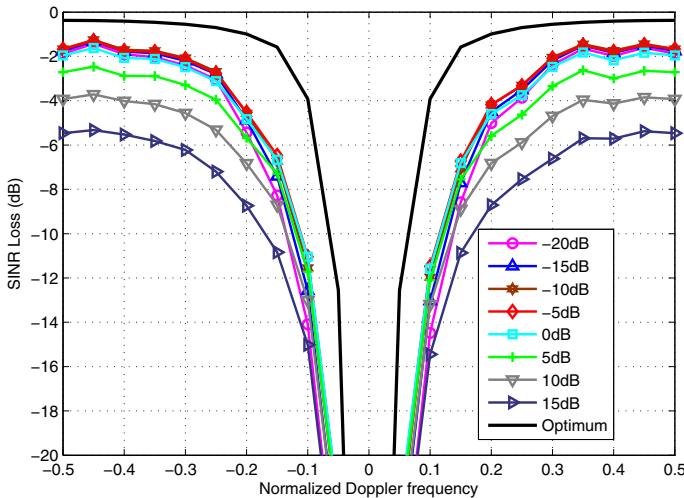

Fig. 4 SINR loss against target Doppler frequency of the CMT-FOCUSS-D3-STAP algorithm with different β_L

results are plotted in Fig. 4. The curves show that the proposed algorithm exhibits robust performance for a range of values of the diagonal loading term β_L, i.e., $\beta_L = -20, -15, -10, -5, 0\,\mathrm{dB}$ to the true noise level.

4.2 Comparisons with Existing Algorithms

In this subsection, we focus on the performance comparisons between the the proposed algorithm and other existing algorithms.
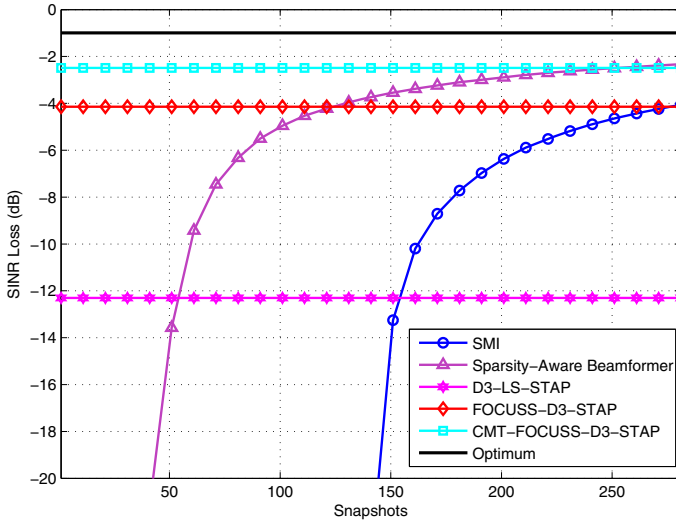
Fig. 5 SINR loss against the number of snapshots for all algorithms

To have a comparable convergence performance with the existing STAP performance using training data, we illustrate the SINR loss against the number of snapshots in Fig. 5. In this example, we consider the ICM with $\sigma_v = 1$ and assume the estimated CMT of the proposed algorithm with $\hat{\sigma}_v = 2$. In addition, we keep the target normalized Doppler frequency of 0.3 for all algorithms. Because the D3-type STAP algorithms, i.e., the D3-LS STAP algorithm, the FOCUSS-D3-STAP algorithm and the proposed CMT-FOCUSS-D3-STAP algorithm, only use the snapshot in the CUT, the SINR performance keeps invariant with the increase of snapshots. While the SINR performance of the conventional SMI algorithm and the sparsity-aware beamformer algorithm increases as the number of snapshots increases. Thus, the D3-type STAP algorithms exhibit a significant performance improvement over the standard SMI algorithm and the sparsity-aware beamformer algorithm when the number of IID snapshots is very low, e.g., in seriously nonhomogeneous environments.

In the second example, we compare the performance of the proposed CMT-FOCUSS-D3-STAP algorithm using CMT adaptation with other existing STAP algorithms in four different ICM cases, namely $\sigma_v = 0.25, 0.5, 1, 5$. In the simulation, we set the candidates space of the parameter σ_v as $\Omega_{\sigma_v} = \{\sigma_{v;k} = 0.1, 0.3, 0.9, 1.5, 2.5, 5, 6\}$. The clutter eigenspectrum for different levels of ICM is illustrated in Fig. 6. As the velocity standard deviation σ_v increases, the tails of the eigenspectrum become larger as the clutter rank increases. Figure 7a–d plots the SINR loss against target Doppler frequency of different STAP algorithms with the above-mentioned levels of ICM. From the figures, we observe that: (i) the proposed CMT-FOCUSS-D3-STAP algorithm outperforms the FOCUSS-D3-STAP algorithm and the D3-LS STAP algorithm in almost all Doppler bins; (ii) as σ_v increases, the performance of the FOCUSS-D3-STAP algorithm and the D3-LS STAP algorithm degrades significantly while the proposed CMT-FOCUSS-D3-STAP algorithm provides relative robustness
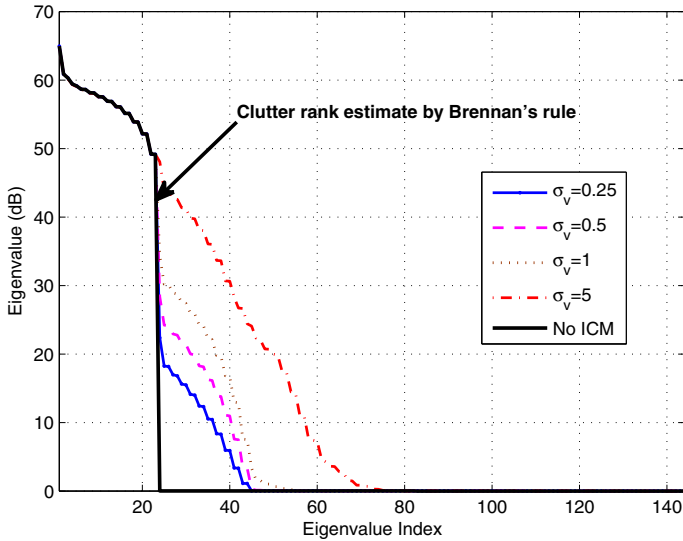
Fig. 6 Clutter eigenspectra for different levels of ICM

to the ICM. This is because the proposed algorithm employs a CMT to overcome the effects of the ICM. Furthermore, compared with the results of the proposed algorithm in Figs. 2 and 7c, it implies that the proposed CMT adaptation approach can efficiently select the best CMT to suppress the clutter. In addition, we also observe that the proposed algorithm performs worse than the SMI algorithm and the sparsity-aware beamformer algorithm in the normalized Doppler range of −0.2 to 0.2. That is because: on the one hand, some clutter components may also be excluded by using the target region Ω_{SOI} when the target is close to the clutter ridge. On the other hand, it should be noted that the numbers of snapshots used in the SMI and the sparsity-aware beamformer algorithm are 288 and 100 respectively, while the proposed algorithm only uses the snapshot in the CUT. When in seriously nonhomogeneous environments, the performance of the SMI algorithm and the sparsity-aware beamformer algorithm will be seriously affected because of the snapshot deficiency.

Because the target component will be included in the recovered spatio-temporal profile, in this example, we evaluate the effect of SNR on the P_d performance of the proposed algorithm with adaptive CMT approach, as shown in Fig. 8. In the simulations, we keep the ICM with $\sigma_v = 1$ and the CNR 50dB. The false alarm rate (P_{fa}) is set to 10^{-3}, and for simulation purposes the threshold and probability of detection estimates are based on 1000 samples. Furthermore, we consider a scenario with a target not close to the clutter ridge (normalized Doppler frequency 0.3) and another scenario with a target close to the clutter ridge (the normalized Doppler frequency 0.1). The other parameters are the same as those in the previous examples. From the figures, it is found that although the performance of the proposed algorithm degrades when the target is close to the clutter ridge, it always provides higher detection performance than the conventional FOCUSS-D3-STAP algorithm.

Fig. 7 Comparison of SINR loss against target Doppler frequency of different STAP algorithms with different levels of ICM. **a** $\sigma_v = 0.25$; **b** $\sigma_v = 0.5$; **c** $\sigma_v = 1$; and **d** $\sigma_v = 5$

(a)

(b)
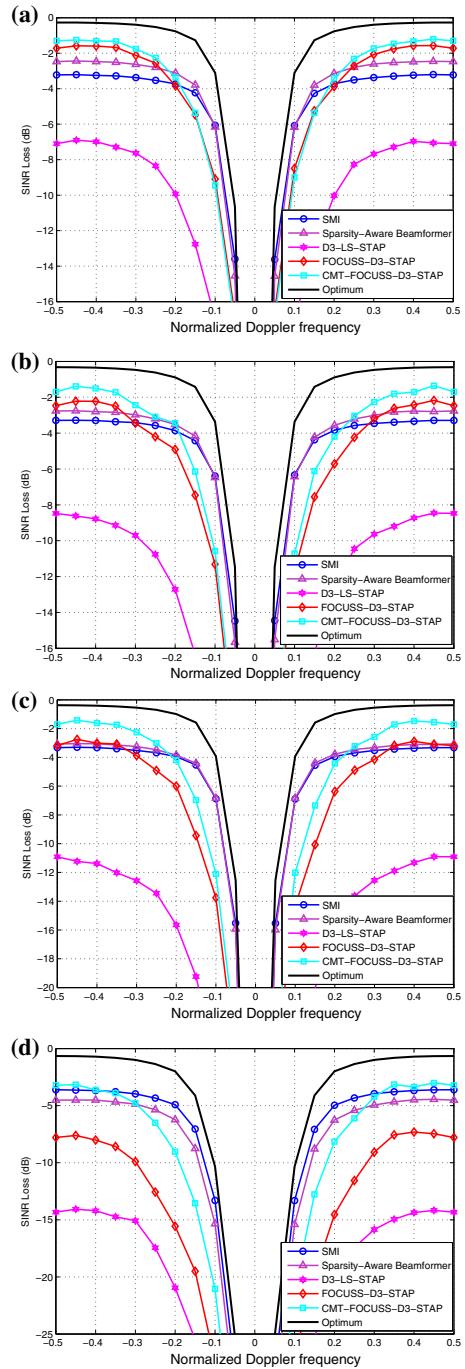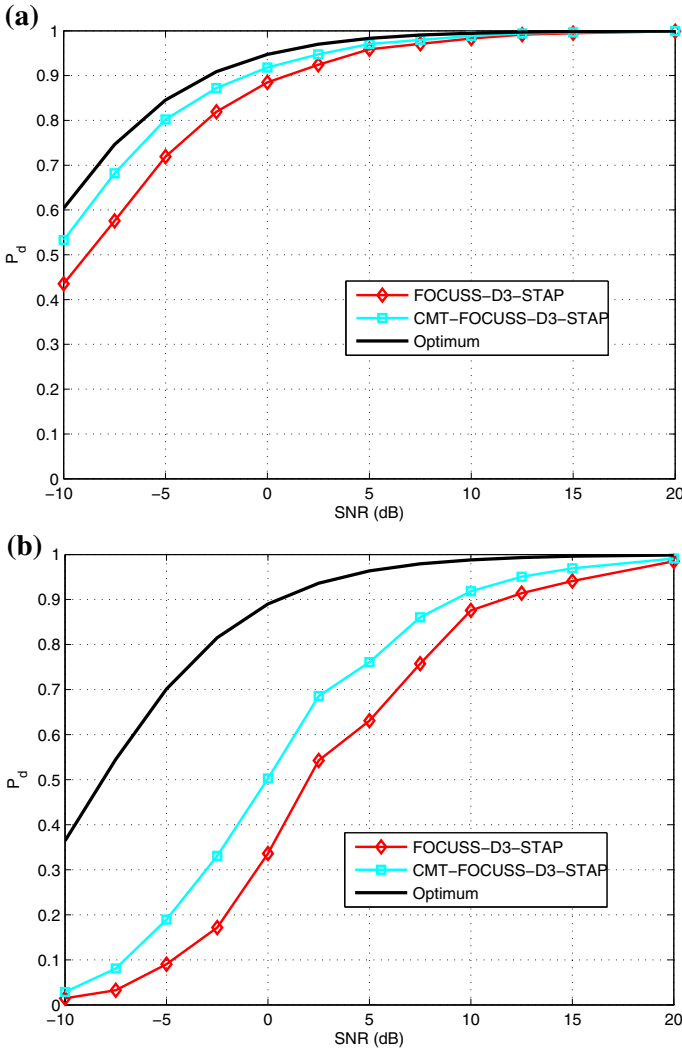
Fig. 8 P_d performance versus target's SNR

To have a further comprehensive analysis of the effect of SNR on the proposed algorithm, average SINR loss performance versus different levels of SNR for different STAP algorithms are plotted in Fig. 9. Here, we consider five different cases of SNR, i.e., SNR $= -10, 0, 10, 20, 30$ dB. The results show that when the SNR is slightly higher than the noise power but far lower than the clutter power, i.e., SNR $= 10$ dB, the SINR improvements of the proposed algorithm and the FOCUSS-D3-STAP algorithm slightly degrade. This is due to the fact that, in this situation, it is hard to accurately recover the target components resulting in some target energy spreading to other elements of the spatio-temporal profile. It is interesting that the smaller the target energy, the smaller effect the target signal to the estimated clutter covariance matrix.
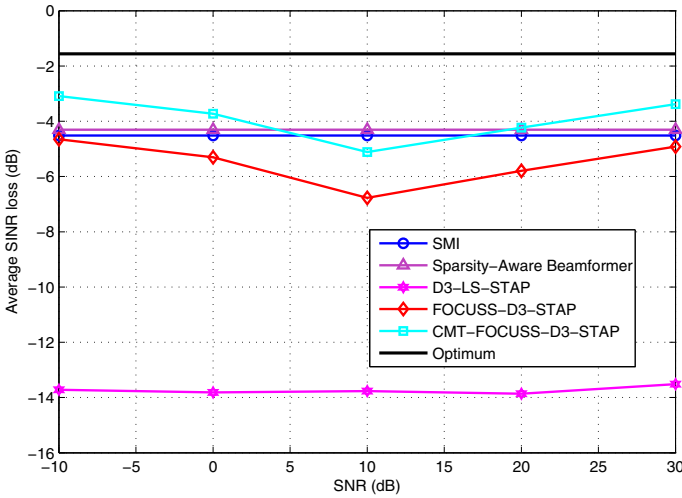
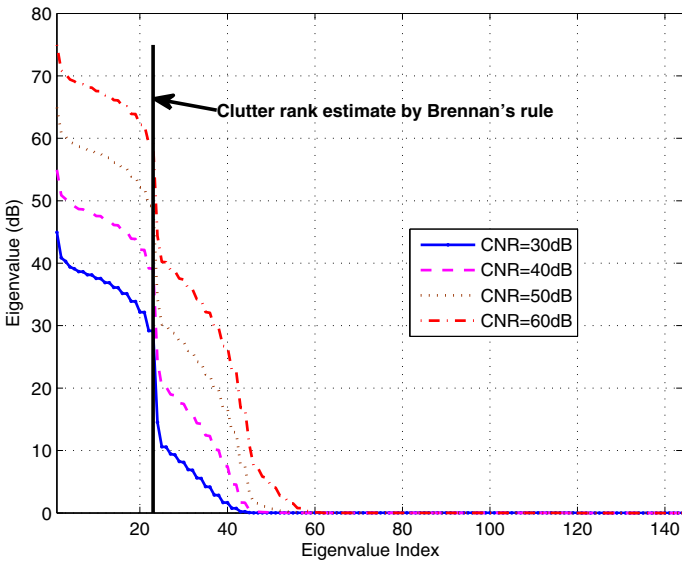**Fig. 9** Average SINR loss performance versus different levels of SNR

Fig. 10 Clutter eigenspectra for different levels of CNR

In other words, the proposed algorithm can obtain a larger SINR improvement in weak target detection.

In the final example, we assess the effect of CNR on SINR performance for the proposed algorithm. In the simulations, we keep the ICM with $\sigma_v = 1$ and the SNR 0 dB, but vary the CNR, i.e., CNR = 30, 40, 50, 60 dB. The other parameters are the same as previous examples. Figure 10 depicts the clutter eigenspectrum for different levels of CNR. Seen from the figure, it is noted that as CNR increases, the clutter

(a)

(b)
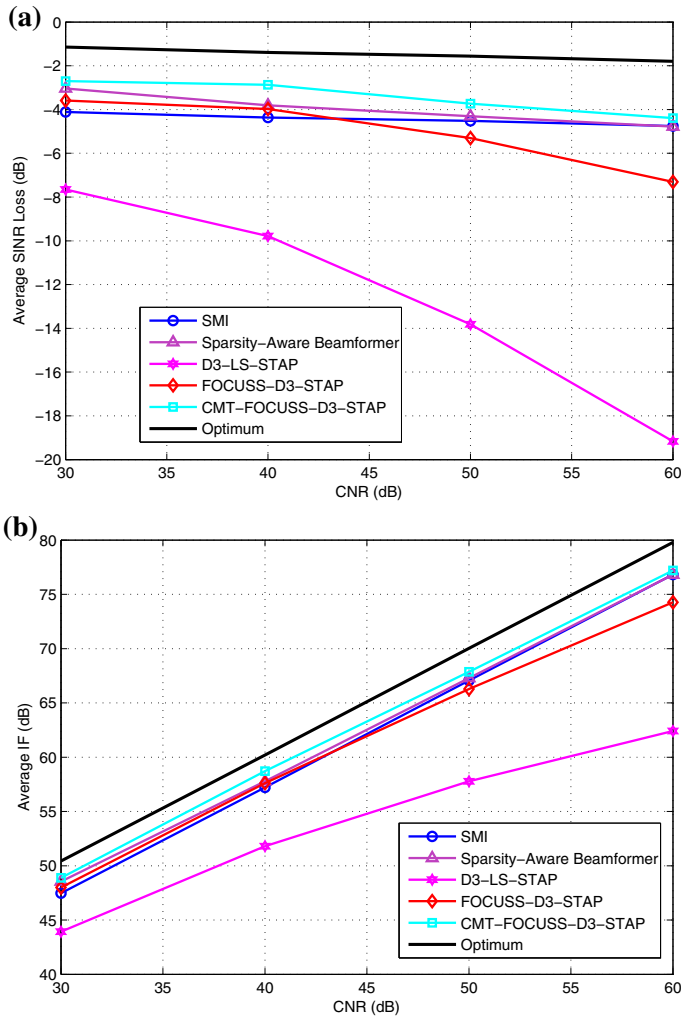
Fig. 11 Comparison of average SINR loss and average IF against different levels of CNR. **a** Average SINR loss. **b** Average IF

rank increases and the tails of the eigenspectrum become larger. Figure 11a, b plots the average SINR loss and average IF performance versus different levels of CNR for different STAP algorithms. The curves show that the proposed CMT-FOCUSS-D3-STAP algorithm exhibits relative more robustness to the CNR than the FOCUSS-D3-STAP algorithm and the D3-LS STAP algorithm. It also should be noted that the SMI algorithm and the sparsity-aware beamformer algorithm provide the best robustness performance against different levels of CNR. This is because the SMI algorithm and the sparsity-aware beamformer algorithm use many snapshots to estimate the clutter covariance matrix. However, the D3-type STAP algorithms only use the snapshot in

the CUT, resulting in some deficiency of the estimated clutter power. Through the CMT, the proposed CMT-FOCUSS-D3-STAP algorithm can mitigate the deficiency of the estimated clutter power in a sense.

5 Conclusions

In this paper, we have proposed a sparsity-based D3-STAP algorithm by formulating a more realistic sparse measurement model for airborne radar that considers the ICM. It first derived the principle of the sparsity-based D3-STAP algorithm based on the FOCUSS technique, which illustrated that the proposed algorithm estimates the clutter covariance matrix by a Hadamard product of the CMT and the clutter covariance matrix estimates with the SR technique. Furthermore, a CMT adaptation approach for the proposed algorithm was developed to automatically select the best CMT. The results with different scenarios in an airborne radar system have showed that the proposed algorithm outperforms the D3-LS STAP algorithm and the conventional sparsity-based D3-STAP algorithm and is expect to obtain much better performance than the SMI algorithm and the sparsity-aware beamformer algorithm in heterogeneous environments. In future, we will extend our model to incorporate more realistic physical effects, such as channel mismatch, antenna array misalignment and range walk.

References

1. E. Aboutanios, B. Mulgrew, A STAP algorithm for radar target detection in heterogeneous environments, in *IEEE Workshop Statist. Signal Process.*(2005), pp. 966–971
2. E. Aboutanios, B. Mulgrew, Hybrid detection approach for STAP in heterogeneous clutter. IEEE Trans. Aerosp. Electron. Syst. **46**, 1021–1033 (2010)
3. S.F. Cotter, B.D. Rao, K. Engan, K. Kreutz-Delgado, Sparse solutions to linear inverse problems with multiple measurement vectors. IEEE Trans. Signal Process. **53**, 2477–2488 (2005)
4. J.-F. Degurse, L. Savy, S. Marcos, Reduced-rank STAP for target detection in heterogeneous environments. IEEE Trans. Aerosp. Electron. Syst. **50**, 1153–1162 (2014)
5. J.-F. Degurse, L. Savy, S. Marcos, J.-P. Molini, Deterministic aided STAP for target detection in heterogeneous situations. Int. J. Antennas Propagat. (2013). doi:10.1155/2013/826935
6. M. Elad, Why simple shrinkage is still relevant for redundant representations? IEEE Trans. Inf. Theory **52**, 5559–5569 (2006)
7. J.R. Guerci, *Space-Time Adaptive Processing for Radar* (Artech House, London, 2003)
8. J.R. Guerci, Theory and application of covariance matrix tapers for robust adaptive beamforming. IEEE Trans. Signal Process. **47**, 977–985 (1999)
9. J.R. Guerci, E.J. Baranoski, Knowledge-aided adaptive radar at DARPA: an overview. IEEE Signal Process. Mag. **23**, 41–50 (2006)
10. J.R. Guerci, J.S. Bergin, Principal components, covariance matrix tapers, and the subspace leakage problem. IEEE Trans. Aerosp. Electron. Syst. **38**, 152–162 (2002)
11. J.S. Goldstein, I.S. Reed, Theory of partially adaptive radar. IEEE Trans. Aerosp. Electron. Syst. **33**, 1309–1325 (1997)
12. I.F. Gorodnitsky, B.D. Rao, Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. IEEE Trans. Signal Process. **45**, 600–616 (1997)
13. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985)

14. R. Klemm, *Principles of Space-Time Adaptive Processing* (Institute of Electrical Engineering, London, UK, 2006)
15. J. Li, X. Zhu, P. Stoica, M. Rangaswamy, High resolution angle-Doppler imaging for MTI radar. IEEE Trans. Aerosp. Electron. Syst. **46**, 1544–1556 (2010)
16. S. Maria, J. J. Fuchs, Application of the global matched filter to STAP data an efficient algorithmic approach, in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2006), pp.14–19
17. W.L. Mevin, G.A. Showman, An approach to knowledge-aided covariance estimation. IEEE Trans. Aerosp. Electron. Syst. **42**, 1021–1042 (2006)
18. J.T. Parker, L.C. Potter, A Bayesian perspective on sparse regularization for STAP post-processing, in *IEEE Radar Conference* (2010), pp. 1471–1475
19. B.D. Rao, K. Engan, S.F. Cotter, J. Palmer, K. Kreutz-Delgado, Subset selection in noise based on diversity measure minimization. IEEE Trans. Signal Process. **51**, 760–770 (2003)
20. J.R. Roman, M. Rangaswamy, D.W. Davis, Q. Zhang, B. Himed, J.H. Michels, Parametric adaptive matched filter for airborne radar applications. IEEE Trans. Aerosp. Electron. Syst. **36**, 677–692 (2000)
21. S. Sen, OFDM radar space-time adaptive processing by exploiting spatio-temporal sparsity. IEEE Trans. Signal Process. **61**, 118–130 (2013)
22. K. Sun, H. Meng, Y. Wang, X. Wang, Direct data domain STAP using sparse representation of clutter spectrum. Signal Process. **91**, 2222–2236 (2011)
23. I.W. Selesnick, S.U. Pillai, K.Y. Li, B. Himed, Angle-Doppler processing using sparse regularization, in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2010), pp. 2750–2753
24. T.K. Sarkar, H. Wang, S. Park, R. Adve, J. Koh, K. Kim, Y. Zhang, M.C. Wicks, R.D. Brown, A deterministic least-squares approach to space-time adaptive processing (STAP). IEEE Trans. Antennas Propagat. **49**, 91–103 (2001)
25. K. Sun, H. Zhang, G. Li, H. Meng, X. Wang, A novel STAP algorithm using sparse recovery technique, in *IGARSS* (2009), pp. 336–339
26. J. Ward, *Space-Time Adaptive Processing for Airborne Radar. Technical Report* (MIT Lincoln laboratory, Lexington, MA, 1994). 1015
27. M.C. Wicks, M. Rangaswamy, R. Adve, T.B. Hale, Space-time adaptive processing: a knowledge-based perspective for airborne radar. IEEE Signal Process. Mag. **23**, 51–65 (2006)
28. Y. Wu, J. Tang, Y. Peng, On the essence of knowledge-aided clutter covariance estimate and its convergence. IEEE Trans. Aerosp. Electron. Syst. **47**, 569–585 (2011)
29. T. Xia, Y. Zheng, Q. Wan, X. Wang, Adaptive regularized FOCUSS algorithm, in *IEEE Radar Conference* (2007), pp. 282–284
30. E. Yang, R. Adve, J. Chun, J. Chun, Hybrid direct data domain sigma-delta space-time adaptive processing algorithm in non-homogeneous clutter. IET Radar Sonar Navig. **4**, 611–625 (2010)
31. Z. Yang, X. Li, H. Wang, Performance analysis of STAP algorithms based on fast sparse recovery techniques. Progr. Electromagn. Res. B **41**, 251–268 (2012)
32. Z. Yang, X. Li, H. Wang, W. Jiang, Adaptive clutter suppression based on iterative adaptive approach for airborne radar. Signal Process. **93**, 3567–3577 (2013)
33. Z. Yang, X. Li, H. Wang, W. Jiang, On clutter sparsity analysis in space-time adaptive processing airborne radar. IEEE Geosci Remote Sens. Lett. **10**, 1214–1218 (2013)
34. Z. Yang, X. Li, H. Wang, L. Nie, Sparsity-based space-time adaptive processing using complex-valued homotopy technique for airborne radar. IET Signal Process. **8**, 552–564 (2014)
35. Z. Yang, R.C. de Lamare, X. Li, L_1-regularized STAP algorithms with a generalized sidelobe canceler architecture for airborne radar. IEEE Trans. Signal Process. **60**, 674–686 (2012)
36. Z. Yang, R.C. de Lamare, X. Li, Sparsity-aware space-time adaptive processing algorithms with L_1-norm regularization for airborne radar. IET Signal Process. **6**, 413–423 (2012)
37. M. Zatman, Comments on Theory and application of covariance matrix tapers for robust adaptive beamforming. IEEE Trans. Signal Process. **48**, 1796–1800 (2000)