



Article

Random Matrix Theory-Based Reduced-Dimension Space-Time Adaptive Processing under Finite Training Samples

Di Song, Qi Feng, Shengyao Chen *, Feng Xi and Zhong Liu

School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

* Correspondence: chenshengyao@njust.com.cn

Abstract: Space-time adaptive processing (STAP) is a fundamental topic in airborne radar applications due to its clutter suppression ability. Reduced-dimension (RD)-STAP can release the requirement of the number of training samples and reduce the computational load from traditional STAP, which attracts much attention. However, under the situation that training samples are severely deficient, RD-STAP will become poor like the traditional STAP. To enhance RD-STAP performance in such cases, this paper develops a novel RD-STAP algorithm using random matrix theory (RMT), RMT-RD-STAP. By minimizing the output clutter-plus-noise power, the estimate of the inversion of clutter plus noise covariance matrix (CNCM) can be obtained through optimally manipulating its eigenvalues, thus producing the optimal STAP weight vector. Specifically, the clutter-related eigenvalues are estimated according to the clutter-related sample eigenvalues via RMT, and the noise-related eigenvalue is optimally selected from the noise-related sample eigenvalues. It is found that RMT-RD-STAP significantly outperforms the RD-STAP algorithm when the RMB rule cannot be satisfied. Theoretical analyses and numerical results demonstrate the effectiveness and the performance advantages of the proposed RMT-RD-STAP algorithm.



Citation: Song, D.; Feng, Q.; Chen, S.; Xi, F.; Liu, Z. Random Matrix Theory-Based Reduced-Dimension Space-Time Adaptive Processing under Finite Training Samples. *Remote Sens.* **2022**, *14*, 3959. <https://doi.org/10.3390/rs14163959>

Academic Editors: Jingwei Xu, Keqing Duan, Weijian Liu and Xiongpeng He

Received: 8 July 2022

Accepted: 12 August 2022

Published: 15 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

Keywords: space-time adaptive processing; reduced-dimension; finite training samples; random matrix theory

1. Introduction

Space-time adaptive processing (STAP) was first proposed almost 50 years ago [1], and since then it has been actively investigated by the radar community due to its strong capability in suppressing clutter and enhancing target [2–6]. The optimal STAP can maximize the output signal-to-interference plus noise ratio (SINR) and provide reliable target detection. However, the optimal STAP involves an ideal clutter plus noise covariance matrix (CNCM), which generally cannot be attained and must be estimated with independent identical distributed (IID) training samples, called sample covariance matrix (SCM). Once the number of training samples for SCM is insufficient, STAP would severely suffer from performance deterioration. Particularly, the famous Reed, Mallet, and Brennan (RMB) rule [4] provides an explicit theoretical guide for the requirement on the number of training samples, i.e., to limit the SINR loss to 3 dB, the number of training samples must be not less than two times the system degrees of freedom (DoFs). However, in practice, the RMB rule is commonly hard to satisfy. In addition, the computational load in optimal STAP is extremely huge owing to high-dimensional matrix inversion. Therefore, the traditional STAP needs to be improved.

With the development of STAP, a great number of methods have been developed to confront the performance deterioration stemming from finite training samples. Sparsity-recovery (SR)-STAP [7–10] can greatly improve the SCM with only a few or even one training sample, however, the computational load involved in SR is unbearable. Knowledge-aided (KA)-STAP [11–14] uses the prior knowledge to improve SCM, i.e., covariance matrix and terrain data, but its performance will severely deteriorate once the prior knowledge

deviates from the actual one. Recently, deep convolutional neural networks (CNNs) [15–17] have been used for STAP also with few training samples, but it usually needs a lot of sampling data as the training set for network learning, which is unavailable in many radar scenarios. Moreover, it is hard to obtain the ground truth in actual clutter scenarios, which is necessarily required as a label for supervised learning.

Reduced-dimension (RD)-STAP [18–23] is one of the most popular methods due to its fast convergence, simple structure, and a low requirement for training samples, which has been widely used in engineering applications. RD-STAP algorithms reduce the system DoFs through linear transformation on training samples to satisfy the RMB rule. Meanwhile, the lower system DoFs cause lower computational load, which significantly boosts real-time processing for airborne radar. In terms of different principles, several linear transformations have been constructed, such as the extended factored approach (EFA) [18] and joint domain localized (JDL) [23]. However, in a radar system, the linear transformation is often fixed in a pre-designed clutter scenario. When the scenarios deviate from the pre-designed ones, the RMB rule may not be satisfied and the RD-STAP performance becomes poor like the traditional STAP. Therefore, it is greatly meaningful to resort to other techniques to enhance RD-STAP performance.

It could be noticed that CNCM is composed of eigenvalues and eigenvectors, and a lot of works have demonstrated that the SCM can be improved by adjusting the sample eigenvalues while maintaining the sample eigenvectors [24–27]. Currently, random matrix theory (RMT), originating from quantum mechanics [28] and mainly used to study the asymptotic behavior of the empirical spectral distribution of different random matrix models as their dimensions go into infinity [29–32], has been applied to estimate the eigenvalues and eigenvectors of the SCM constructed from the finite observations [33–36]. It is found that the estimation is consistent, not only when the sample size increases without bound for a fixed observation dimension but also when the observation dimension increases to infinity at the same rate as the sample size increases [37–39]. Particularly, the spiked covariance model, defined as a low-rank perturbation of an identity matrix in RMT [40], attracts much attention since it is similar to the classical signal-plus-noise model in signal processing. It is shown in [40–43] that for such a model, consistent estimation of its eigenvalues can be obtained with small-size samples. Recently, ref. [27] applies the theory to directly estimate the inverse covariance matrix for spatial beamforming and shows superior performance under high dimension and finite training samples. Aiming at the STAP problem, some works, like [44,45], use RMT to enhance the performance under finite training samples. However, these methods are implemented in full-dimension STAP, which leads to a huge computational load and is not applicable to RD-STAP. Based on this consideration, we try to improve the performance of RD-STAP by using RMT.

Inspired by [27], this paper proposes a new RD-STAP algorithm using RMT, called RMT-RD-STAP. Like the CNCM's eigenvalues, the RD CNCM's eigenvalues consist of two parts, the noise-related part, and the clutter-related part. The noise-related part contains the large portion of the RD CNCM's eigenvalues and their values are determined by the noise power, whereas the clutter-related part takes up the small portion of the RD CNCM's eigenvalues and their values depend on the clutter power. The eigenvalues belonging to the clutter-related part are commonly much larger than those belonging to the noise-related part. Such eigenvalue distribution is similar to that of a spiked covariance model. The main difference is that the noise covariance matrix in RD CNCM is not strictly a scaled identity matrix which is required in the spiked covariance model. If the noise covariance matrix in RD CNCM can be considered a scaled identity matrix, the RMT-based technique in [27] is applicable to RD STAP with finite training samples and the performance of clutter suppression will be enhanced. In this paper, we deal with the noise covariance matrix in RD CNCM as a scaled identity matrix, and fortunately, we find that the spiked covariance model can effectively work for RD-STAP with finite training samples. We first construct a new estimate on the inverse RD CNCM by employing the eigenvectors of the RD SCM and then estimate the clutter-related eigenvalues by using the RMT. Meanwhile, we provide

the estimate on the scaled factor of the scaled identity matrix according to the eigenvalues of the RD SCM. With the proposed estimate on the inverse RD CNCM, we can directly obtain the space-time adaptive weight vector of RMT-RD-STAP in terms of the linear constraint minimum variance (LCMV) criterion. We conduct extensive numerical experiments to verify the effectiveness of the RMT-RD-STAP, and we find that the proposed RMT-RD-STAP greatly outperforms the conventional RD-STAP in the case of finite training samples.

The rest of this paper is organized as follows. Section 2 introduces the echo model for airborne radar and the fundamentals of STAP and RD-STAP. Section 3 presents the proposed RMT-RD-STAP algorithm in detail. Numerical experiments are performed in Section 4 and the conclusions of this work are drawn in Section 5.

Notation: \mathbf{I}_N denotes the $N \times N$ identity matrix. Boldface uppercase letters denote the matrices, and boldface lowercase letters denote the vectors. $j = \sqrt{-1}$. The \otimes represents the Kronecker product. The $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and Hermitian transpose, respectively. The \mathbb{C} represents the sets of complex values. The $\mathbb{E}\{\cdot\}$ denotes the statistical expectation.

2. Background

2.1. Echo Model of Airborne Radar and Optimal STAP

As shown in Figure 1, consider a side-looking airborne pulse-Doppler radar system equipped with a uniform linear array (ULA) of N array elements with the inter-element spacing $\Delta = \lambda/2$ and the wavelength λ . Let H and V denote the height and velocity of the platform, respectively, and ψ , θ , and φ denote the cone, azimuth, and elevation angle, respectively. Assume that the radar transmits K pulses at a constant pulse repetition frequency (PRF) f_r during a coherent processing interval (CPI). Then under the assumption of no range ambiguity and no internal clutter motion, the received signal $x_l \in \mathbb{C}^{NK}$ of the l -th range cell or the range cell under test (CUT) can be represented as

$$x_l = s + c_l + n, \tag{1}$$

where $s \in \mathbb{C}^{NK}$ denotes target echo, $n \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{NK})$ denotes the complex white Gaussian noise with the noise power σ_n^2 , and $c_l \in \mathbb{C}^{NK}$ denotes the clutter echo,

$$c_l = \sum_{i=1}^{N_c} \Gamma_i \mathbf{a}(f_{ci}^t, f_{ci}^s). \tag{2}$$

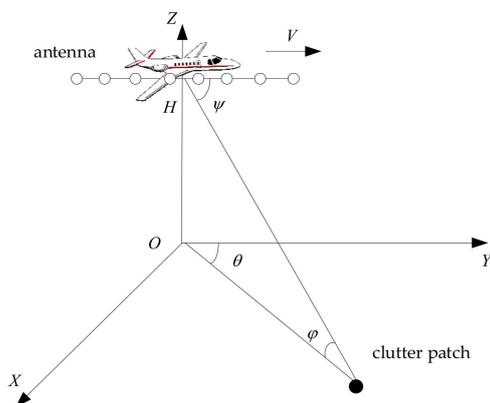


Figure 1. Side-looking airborne radar geometry relationship.

In (2), N_c is the number of clutter patches evenly divided in azimuth, Γ_i denotes the complex amplitude of the i -th clutter patch, and $\mathbf{a}(f_{ci}^t, f_{ci}^s) = \mathbf{a}_t(f_{ci}^t) \otimes \mathbf{a}_s(f_{ci}^s) \in \mathbb{C}^{NK}$ denotes the normalized spatial-temporal steering vector of the i -th clutter patch where

$\mathbf{a}_t(f_{ci}^t) \in \mathbb{C}^K$, and $\mathbf{a}_s(f_{ci}^s) \in \mathbb{C}^N$ respectively denote the temporal and spatial steering vectors as

$$\begin{aligned} \mathbf{a}_t(f_{ci}^t) &= [1, e^{j2\pi f_{ci}^t}, \dots, e^{j2\pi f_{ci}^t(K-1)}]^T \\ \mathbf{a}_s(f_{ci}^s) &= [1, e^{j2\pi f_{ci}^s}, \dots, e^{j2\pi f_{ci}^s(N-1)}]^T \end{aligned} \tag{3}$$

In (3), f_{ci}^t and f_{ci}^s denote the normalized temporal and spatial frequency, respectively,

$$\begin{aligned} f_{ci}^t &= \frac{2V}{\lambda f_r} \cos \psi_i = \frac{2V}{\lambda f_r} \cos \theta_i \cos \varphi_l \\ f_{ci}^s &= \frac{\Delta}{\lambda} \cos \psi_i = \frac{\Delta}{\lambda} \cos \theta_i \cos \varphi_l \end{aligned} \tag{4}$$

Then the CNCM \mathbf{R} of the l -th range cell can be expressed from (1) as

$$\mathbf{R} = \mathbb{E}(\mathbf{x}_l \mathbf{x}_l^H) = \sum_{i=1}^{N_c} |\Gamma_i|^2 \mathbf{a}(f_{ci}^t, f_{ci}^s) \mathbf{a}^H(f_{ci}^t, f_{ci}^s) + \sigma_n^2 \mathbf{I}_{NK}. \tag{5}$$

Under the LCMV criterion [2], the optimal STAP is defined as

$$\begin{cases} \min & P(\mathbf{w}) = \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{s.t.} & \mathbf{w}^H \mathbf{a}(f_0^t, f_0^s) = 1 \end{cases} \tag{6}$$

where $P(\mathbf{w})$ denotes the output clutter-plus-noise power, $\mathbf{a}(f_0^t, f_0^s) \in \mathbb{C}^{NK}$ denotes the spatial-temporal steering vector of the target with f_0^t and f_0^s as the normalized temporal and spatial frequency. The optimal STAP weight vector $\mathbf{w} \in \mathbb{C}^{NK}$ can be derived as

$$\mathbf{w}^{opt} = \frac{\mathbf{R}^{-1} \mathbf{a}(f_0^t, f_0^s)}{\mathbf{a}^H(f_0^t, f_0^s) \mathbf{R}^{-1} \mathbf{a}(f_0^t, f_0^s)}, \tag{7}$$

from which, the minimal clutter-plus-noise power is given by

$$P(\mathbf{w}^{opt}) = \frac{1}{\mathbf{a}^H(f_0^t, f_0^s) \mathbf{R}^{-1} \mathbf{a}(f_0^t, f_0^s)}. \tag{8}$$

2.2. Sample-Based STAP and RD-STAP

For the implementation of (7), it is fundamental to know the CNCM \mathbf{R} and its inversion. However, it is not practical to obtain the matrix \mathbf{R} . In practice, the matrix \mathbf{R} is estimated from finite echo samples or training samples,

$$\mathbf{R}_L = \sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l^H, \tag{9}$$

where \mathbf{R}_L is called the SCM with L training samples. In such case, the optimal weight vector (7) is approximated as

$$\mathbf{w}_L^{opt} = \frac{\mathbf{R}_L^{-1} \mathbf{a}(f_0^t, f_0^s)}{\mathbf{a}^H(f_0^t, f_0^s) \mathbf{R}_L^{-1} \mathbf{a}(f_0^t, f_0^s)}, \tag{10}$$

where \mathbf{w}_L^{opt} is called the adaptive weight vector of the traditional STAP. The output clutter-plus-noise power by the weight vector (10) is given by

$$P(\mathbf{w}_L^{opt}) = \frac{\mathbf{a}^H(f_0^t, f_0^s) \mathbf{R}_L^{-1} \mathbf{R} \mathbf{R}_L^{-1} \mathbf{a}(f_0^t, f_0^s)}{[\mathbf{a}^H(f_0^t, f_0^s) \mathbf{R}_L^{-1} \mathbf{a}(f_0^t, f_0^s)]^2}. \tag{11}$$

It is seen from (11) that the traditional STAP will have good performance if \mathbf{R}_L^{-1} well approximates \mathbf{R}^{-1} . However, \mathbf{R}_L^{-1} is estimated from finite training samples and deviates from \mathbf{R}^{-1} if the number of the training samples decreases. This deviation will result in a higher output clutter-plus-noise power by (11) than the theoretical one by (8). Define a constant $c \in (0, 1)$. Then from [46,47], it is known that $P(\mathbf{w}_L^{opt})/P(\mathbf{w}^{opt}) \rightarrow 1/(1-c)$ as $NK, L \rightarrow \infty$ with $c_N = NK/L \rightarrow c$, which coincides with the RMB rule [4], i.e., the sample-based STAP performance by (10) has less than 3 dB loss when the number of training samples is not less than two times the system DoFs, $L \geq 2NK$. However, in real radar scenarios, it is a common case that $L < NK$ [2]. The sample-based STAP by (10) will have much poorer performance than the optimal STAP by (7).

To reduce the requirement of (10) on the number of training samples, RD-STAP [18–23] is developed, which produces a secondary data by projecting the NK -dimensional signal vector \mathbf{x}_l into a M -dimensional one via linear transformation matrix $\mathbf{T} \in \mathbb{C}^{NK \times M}$, namely

$$\tilde{\mathbf{x}}_l = \mathbf{T}^H \mathbf{x}_l \in \mathbb{C}^M, \tag{12}$$

where M denotes the reduced-dimensional system DoFs and is much smaller than the original system DoFs, $M \ll NK$. Then, it can be predicted that the STAP with the data in (12) will have good performance when $L \geq 2M$ in terms of the RMB rule.

The RD CNCM and the corresponding target spatial-temporal steering vector are given as

$$\mathbf{R}_{rd} = \mathbb{E}(\tilde{\mathbf{x}}_l \tilde{\mathbf{x}}_l^H) = \mathbf{T}^H \mathbf{R} \mathbf{T} \in \mathbb{C}^{M \times M}, \tag{13}$$

and

$$\mathbf{a}_{rd}(f_0^t, f_0^s) = \mathbf{T}^H \mathbf{a}(f_0^t, f_0^s) \in \mathbb{C}^M. \tag{14}$$

Under the LCMV criterion, the RD-STAP is defined as

$$\begin{cases} \min & P_{rd}(\mathbf{w}_{rd}) = \mathbf{w}_{rd}^H \mathbf{R}_{rd} \mathbf{w}_{rd} \\ \text{s.t.} & \mathbf{w}_{rd}^H \mathbf{a}_{rd}(f_0^t, f_0^s) = 1 \end{cases}, \tag{15}$$

where $\mathbf{w}_{rd} \in \mathbb{C}^M$ denotes the RD-STAP weight vector. Then the optimal RD-STAP weight vector is given as

$$\mathbf{w}_{rd}^{opt} = \frac{\mathbf{R}_{rd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s)}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{R}_{rd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s)}. \tag{16}$$

When the \mathbf{R}_{rd} is generated from the finite samples, the adaptive weight vector of the sample-based RD-STAP is

$$\mathbf{w}_{Lrd}^{opt} = \frac{\mathbf{R}_{Lrd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s)}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{R}_{Lrd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s)}, \tag{17}$$

where $\mathbf{R}_{Lrd} = \mathbf{T}^H \mathbf{R}_L \mathbf{T} \in \mathbb{C}^{M \times M}$ is the RD SCM. Similar to (11), the output clutter-plus-noise power associated with the weight vector (17) is given as

$$P_{rd}(\mathbf{w}_{Lrd}^{opt}) = \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{R}_{Lrd}^{-1} \mathbf{R}_{rd} \mathbf{R}_{Lrd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{R}_{Lrd}^{-1} \mathbf{a}_{rd}(f_0^t, f_0^s) \right]^2}. \tag{18}$$

2.3. Motivation by Introducing RMT in RD-STAP

Although RD-STAP can release the requirement on the number of training samples, the RD processing structure is commonly fixed for a radar system, then the RD-STAP performance becomes poor like the traditional STAP once the number of training samples is small. Therefore, it is greatly necessary to resort to other techniques to improve the

RD-STAP performance. In this paper, we study the consistent estimate of \mathbf{R}_{rd}^{-1} from the finite training samples, especially in the case of $L \leq 2M$.

For the inverse matrix \mathbf{R}_{rd}^{-1} , let us consider the eigen-decomposition of the matrix \mathbf{R}_{rd} as

$$\mathbf{R}_{rd} = \tilde{\sigma}_n^2 \sum_{i=1}^M \gamma_i \mathbf{e}_i \mathbf{e}_i^H = \tilde{\sigma}_n^2 \sum_{i=1}^{Q_{rd}} \gamma_i \mathbf{e}_i \mathbf{e}_i^H + \tilde{\sigma}_n^2 \sum_{i=Q_{rd}+1}^M \gamma_i \mathbf{e}_i \mathbf{e}_i^H, \tag{19}$$

where Q_{rd} denotes the local clutter DoFs, which can be calculated by the method in [48], \mathbf{e}_i is the i -th eigenvector with the corresponding eigenvalue $\tilde{\sigma}_n^2 \gamma_i$ and $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_{Q_{rd}} \gg \gamma_{Q_{rd}+1} \geq \dots \geq \gamma_M$. The leading eigenvalues γ_i ($i = 1, \dots, Q_{rd}$) are related to the clutter, whereas the others are related to the noise.

Notice that Q_{rd} is generally far less than the RD system DoFs M in practice [48], and the clutter-related eigenvalues are much larger than the noise-related ones. Such eigenvalue distribution is like that of a spiked covariance model although it is not strict. In the area of RMT, this model has acquired extensive study on its statistical behaviors and the consistent estimation theory has been established on estimating the isolated eigenvalues of the spiked covariance model from finite training samples [38–41]. In particular, the spike covariance matrix estimation can be improved by managing the sample eigenvalues while keeping the sample eigenvectors [24–27]. Then, this theory can be leveraged to improve the estimation of \mathbf{R}_{rd}^{-1} and in turn enhance the RD-STAP performance.

3. RD-STAP Using RMT

In this section, we propose a novel RD-STAP algorithm by using the RMT, called RMT-RD-STAP. Firstly, the RMT-RD-STAP problem with the spiked covariance model is defined. Then, its asymptotic deterministic equivalence is given and the optimal solution is derived; ultimately an estimated adaptive weight vector is optimized.

3.1. RD-STAP Problem under the Spiked Covariance Model

It is seen from (18) that the \mathbf{R}_{rd} does not have the spiked structure as \mathbf{R} because the noise covariance matrix is not a scaled identity one [48]. Then the RMT-based technique in [27] cannot be directly applicable to the estimate of the inverse CNCM \mathbf{R}_{rd}^{-1} . Note that the optimal RD-STAP is to minimize the output clutter-plus-noise power. Then the essence of the sample-based RD-STAP is to find the estimate of \mathbf{R}_{rd}^{-1} which can reduce the output clutter-plus-noise power (17). We will find that the output clutter-plus-noise power in (17) can be reduced if the noise-related covariance matrix is replaced by a scaled identity matrix. Although there are no theoretical guarantees, a number of numerical experiments in Section 4 have confirmed the assertion. Then we make Assumption 1 as follows.

Assumption 1. *The noise component in (12) is taken to be the complex white Gaussian noise with zero mean and covariance $\sigma \mathbf{I}_M$.*

Define $\beta_i = \gamma_i - 1$. Then under Assumption 1, the RD CNCM \mathbf{R}_{rd} in (19) can be approximated as

$$\mathbf{R}_{rd} \approx \sigma \left(\sum_{i=1}^{Q_{rd}} \beta_i \mathbf{e}_i \mathbf{e}_i^H + \mathbf{I}_M \right), \tag{20}$$

where the noise power σ can be considered as the scaled factor of the identity matrix in the spiked covariance model. With the eigen-decomposition of the RD SCM \mathbf{R}_{Lrd} as

$$\mathbf{R}_{Lrd} = \sigma \sum_{i=1}^M \alpha_i \mathbf{d}_i \mathbf{d}_i^H, \tag{21}$$

in which \mathbf{d}_i is the i -th eigenvector with the corresponding eigenvalue $\sigma\alpha_i$ and $\alpha_1 \geq \dots \geq \alpha_M$, according to the spike covariance matrix estimation principle [25], an inverse RD CNCM $\bar{\mathbf{R}}_{rd}^{-1}$ can be constructed as

$$\bar{\mathbf{R}}_{rd}^{-1} = \frac{1}{\sigma} \sum_{i=1}^M \zeta_i \mathbf{d}_i \mathbf{d}_i^H, \tag{22}$$

where $1/\sigma\zeta_i$ are the eigenvalues of $\bar{\mathbf{R}}_{rd}^{-1}$ and the parameters ζ_i need to be determined. By using Assumption 1, setting $\zeta_{Q_{rd}+1} = \dots = \zeta_M = 1$, and letting $g_i = \zeta_i - 1$, (22) can be re-expressed as

$$\bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) = \frac{1}{\sigma} \left(\sum_{i=1}^{Q_{rd}} g_i \mathbf{d}_i \mathbf{d}_i^H + \mathbf{I}_M \right), \tag{23}$$

where the inverse matrix $\bar{\mathbf{R}}_{rd}^{-1}$ is implicitly expressed as a function of the vector $\mathbf{g} = [g_1 \dots g_{Q_{rd}}]^T$. Correspondingly, the adaptive weight vector $\bar{\mathbf{w}}_{rd}^{opt}$ is given as

$$\bar{\mathbf{w}}_{rd}^{opt}(\mathbf{g}) = \frac{\bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) \mathbf{a}_{rd}(f_0^t, f_0^s)}. \tag{24}$$

Then the RD-STAP problem under the spiked covariance model can be defined as

$$\mathbf{g}^* = \underset{\mathbf{g}}{\operatorname{argmin}} P_{rd}(\bar{\mathbf{w}}_{rd}^{opt}(\mathbf{g})), \tag{25}$$

where

$$\begin{aligned} P_{rd}(\bar{\mathbf{w}}_{rd}^{opt}(\mathbf{g})) &= \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) \mathbf{R}_{rd} \bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\mathbf{g}) \mathbf{a}_{rd}(f_0^t, f_0^s) \right]^2} \\ &= \sigma \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \left(\mathbf{I}_M + \sum_{q=1}^{Q_{rd}} g_q \mathbf{d}_q \mathbf{d}_q^H \right) \left(\mathbf{I}_M + \sum_{j=1}^{Q_{rd}} \beta_j \mathbf{e}_j \mathbf{e}_j^H \right) \left(\mathbf{I}_M + \sum_{i=1}^{Q_{rd}} g_i \mathbf{d}_i \mathbf{d}_i^H \right) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \left(\mathbf{I}_M + \sum_{i=1}^{Q_{rd}} g_i \mathbf{d}_i \mathbf{d}_i^H \right) \mathbf{a}_{rd}(f_0^t, f_0^s) \right]^2} \end{aligned} \tag{26}$$

The optimal RMT-RD-STAP is equivalent to finding the optimal \mathbf{g}^* .

3.2. Asymptotic Deterministic Equivalent $\tilde{P}_{rd}(\bar{\mathbf{w}}_{rd}^{opt}(\mathbf{g}))$ and the Optimal $\tilde{\mathbf{g}}^*$

Note that it is difficult to give a closed solution of (25). To address the issue, the asymptotic properties of (26) under $M, L \rightarrow \infty$ are considered. To conduct such an analysis, three assumptions are made as follows:

Assumption 2. As $M, L \rightarrow \infty$, $M/L = c_{Nrd} \rightarrow c$ for a certain $c > 0$.

Assumption 3. The number of Q_{rd} is fixed and smaller than M .

Assumption 4. $\beta_1 \geq \dots \geq \beta_{Q_{rd}} > \sqrt{c}$ and $\alpha_1 \geq \dots \geq \alpha_{Q_{cd}} > (1 + \sqrt{c})^2$.

For Assumption 2, both M and L are assumed to be reasonably large. When $c > 1/2$, the number of training samples is less than the system DoFs, i.e., $L \leq 2M$, the case considered in this paper. Assumption 3 is often taken in the airborne phased-array radar system and $Q_{rd} < M$ [2,4]. The assumption $\beta_1 \geq \dots \geq \beta_{Q_{rd}} > \sqrt{c}$ in Assumption 4 is easily satisfied in practice because the clutter power is far higher than the noise power [2,4]. For $\alpha_1 \geq \dots \geq \alpha_{Q_{cd}} > (1 + \sqrt{c})^2$, the number L of the training samples is not less than the local clutter DoFs, $L \geq Q_{rd}$. Otherwise $\alpha_1 = \dots = \alpha_{Q_{rd}} = 0 < (1 + \sqrt{c})^2$, and β_i cannot be uniquely determined from α_i [40–43], leading to the impossibility of recovering the matrix \mathbf{R}_{rd} .

The three assumptions follow [49,50], which are fundamental to deriving the asymptotic properties of (26), as discussed in [27,51]. Then it can be shown that

$$\begin{cases} \mathbf{e}_i^H \mathbf{d}_j \mathbf{d}_j^H \mathbf{e}_i - s_{rdi} \delta_{ij} \Big| \xrightarrow{a.s.} 0, j = 1, \dots, Q_{rd} \\ \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{d}_i \mathbf{d}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s) - s_{rdi} k_{rdi} \Big| \xrightarrow{a.s.} 0 \\ \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{d}_i \mathbf{d}_i^H \mathbf{e}_i \mathbf{e}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s) - s_{rdi} k_{rdi} \Big| \xrightarrow{a.s.} 0 \\ \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{d}_i \mathbf{d}_i^H \mathbf{e}_i \mathbf{e}_i^H \mathbf{d}_i \mathbf{d}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s) - s_{rdi}^2 k_{rdi} \Big| \xrightarrow{a.s.} 0 \end{cases} \quad (27)$$

where δ_{ij} is the Kronecker delta function, $s_{rdi} = (1 - c_{Nrd}/\beta_i^2)/(1 + c_{Nrd}/\beta_i)$, and $k_{rdi} = \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{e}_i \mathbf{e}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s)$. In terms of (27), $P_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g}))$ has its asymptotic deterministic equivalent $\tilde{P}_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g}))$ as

$$\tilde{P}_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g})) = \frac{\sigma \left(\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + 2 \sum_{i=1}^{Q_{rd}} g_i s_{rdi} k_{rdi} + \sum_{i=1}^{Q_{rd}} k_{rdi} \beta_i + 2 \sum_{i=1}^{Q_{rd}} g_i \beta_i s_{rdi} k_{rdi} + \sum_{i=1}^{Q_{rd}} g_i^2 s_{rdi} k_{rdi} + \sum_{i=1}^{Q_{rd}} g_i^2 \beta_i s_{rdi}^2 k_{rdi} \right)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + \sum_{i=1}^{Q_{rd}} g_i s_{rdi} k_{rdi} \right]^2}, \quad (28)$$

and

$$\left| P_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g})) - \tilde{P}_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g})) \right| \xrightarrow{a.s.} 0. \quad (29)$$

By replacing $P_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g}))$ with $\tilde{P}_{rd}(\overline{\mathbf{w}}_{rd}^{opt}(\mathbf{g}))$ in (25), we have the optimal $\tilde{\mathbf{g}}^* = [\tilde{g}_1^*, \dots, \tilde{g}_{Q_{rd}}^*]^T$ by minimizing (28) as

$$\tilde{g}_i^* = \frac{\beta_i + c_{Nrd}}{\beta_i^2 + \beta_i} \left[\frac{\sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} k_{rdq}}{\hat{\beta}_q}}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) - \sum_{q=1}^{Q_{rd}} k_{rdq} + \sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} k_{rdq}}{\hat{\beta}_q^2}} - \beta_i \right]. \quad (30)$$

3.3. Obtaining Optimal Adaptive Weight Vector $\hat{\mathbf{w}}_{rd}^{opt}$ by Estimating $\tilde{\mathbf{g}}^*$

It is seen from (30) that the computation of the optimal values \tilde{g}_i^* involves the unknown quantities β_i and k_{rdi} . Note that under Assumption 4, there exist one-to-one maps between α_i and $\beta_i, i = 1, \dots, Q_{rd}$. With the maps, the unknown quantities β_i and k_{rdi} satisfy [39–42],

$$\begin{cases} \left| \alpha_i - 1 - \beta_i - \frac{c_N(1+\beta_i)}{\beta_i} \right| \xrightarrow{a.s.} 0 \\ \left| k_{rdi} - \frac{1+c_{Nrd}/\hat{\beta}_i}{1-c_{Nrd}/\hat{\beta}_i^2} \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{d}_i \mathbf{d}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s) \right| \xrightarrow{a.s.} 0 \end{cases} \quad (31)$$

Then the consistent estimate of the two unknown quantities β_i and k_{rdi} can be given as

$$\begin{cases} \hat{\beta}_i = \frac{\alpha_i - 1 - c_{Nrd} + \sqrt{(\alpha_i - 1 - c_{Nrd})^2 - 4c_{Nrd}}}{2} \\ \hat{k}_{rdi} = \frac{1+c_{Nrd}/\hat{\beta}_i}{1-c_{Nrd}/\hat{\beta}_i^2} \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{d}_i \mathbf{d}_i^H \mathbf{a}_{rd}(f_0^t, f_0^s) \end{cases} \quad (32)$$

With (32), we have the optimal estimate \hat{g}_i^* as

$$\hat{g}_i^* = \frac{\hat{\beta}_i + c_{Nrd}}{\hat{\beta}_i^2 + \hat{\beta}_i} \left[\frac{\sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q}}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) - \sum_{q=1}^{Q_{rd}} \hat{k}_{rdq} + \sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q^2}} - \hat{\beta}_i \right], \quad (33)$$

and

$$|\hat{g}_i^* - g_i^*| \xrightarrow{a.s.} 0. \quad (34)$$

Then the estimated optimal inverse RD CNCM of $\bar{\mathbf{R}}_{rd}^{-1}$ in the asymptotic case can be expressed as

$$\hat{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) = \frac{1}{\sigma} \left(\sum_{i=1}^{Q_{rd}} \hat{g}_i^* \mathbf{d}_i \mathbf{d}_i^H + \mathbf{I}_M \right), \quad (35)$$

and the corresponding RD-STAP weight vector is given as

$$\hat{\mathbf{w}}_{rd}^{opt}(\hat{\mathbf{g}}^*) = \frac{\bar{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) \mathbf{a}_{rd}(f_0^t, f_0^s)}. \quad (36)$$

3.4. Determining the Noise Power σ

To obtain the optimal estimate of $\hat{\mathbf{R}}_{rd}^{-1}$ from (33) and (35), we need to know the noise power σ in advance. If the noise power is not determined, we would have no access to the optimal RD-STAP weight vector. In this subsection, we provide an estimation method for the noise power σ .

In terms of [39], under finite training samples, the eigenvalues distribution of the covariance matrix will be expanded, i.e., the famous M-P law. However, among all the sample eigenvalues, there must be a sample eigenvalue that can approach the actual one. For RD CNCM, we assume that the noise-related eigenvalues are all the same. Similarly, under finite training samples, there must exist a noise-related sample eigenvalue, which can minimize the output clutter-plus-noise power in (18). Therefore, we can select the optimal noise-related sample eigenvalue as the noise power σ by minimizing the output clutter-plus-noise power in (18).

With the estimated $\hat{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*)$, the output clutter-plus-noise power in (18) can be re-expressed as

$$P_{rd}(\hat{\mathbf{w}}_{rd}^{opt}(\hat{\mathbf{g}}^*)) = \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) \mathbf{R}_{rd} \bar{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \bar{\mathbf{R}}_{rd}^{-1}(\hat{\mathbf{g}}^*) \mathbf{a}_{rd}(f_0^t, f_0^s) \right]^2}, \quad (37)$$

To obtain the optimal noise power σ among all the noise-related sample eigenvalues, we need to solve the following optimization problem

$$\hat{\sigma} = \min_{\sigma} P_{rd}(\hat{\mathbf{w}}_{rd}^{opt}(\hat{\mathbf{g}}^*(\sigma))). \quad (38)$$

By solving this problem, we have derived the following result.

Proposition 1. *If $c_{Nrd} > 1$, the output clutter-plus-noise power P_{rd} in (38) decreases as σ decreases. If $c_{Nrd} < 1$, the output clutter-plus-noise power P_{rd} in (38) decreases as σ increases.*

Proof. See Appendix A. \square

According to Proposition 1, we can determine the minimum noise-related sample eigenvalue as the estimated noise power $\hat{\sigma}$ when $c_{Nrd} > 1$, and we can determine the maximum noise-related sample eigenvalue as the estimated noise power $\hat{\sigma}$ when $c_{Nrd} < 1$.

4. Numerical Results

In this section, we conduct several numerical experiments to demonstrate the reasonability and accuracy of Assumption 1 and Proposition 1 and evaluate the performance of the proposed RMT-RD-STAP algorithm. The parameters used in the experiment scenarios are listed in Table 1. In terms of these parameters, the clutter data used in experiments are

generated via the Ward model by Lincoln Laboratory at MIT [2]. For comparison, the classical EFA with three Doppler channels is taken as an example of the RD-STAP algorithm [16] with $M = 3N$. In the experiments, the noise power is $\sigma_n^2 = 1$, and the platform velocity V is set to be different values to reveal the performance of the proposed algorithms under different local clutter DoFs. STAP is taken in the main lobe of antenna arrays.

Table 1. Parameters used in the experiment scenarios.

Parameter	Value	Unit
Height	6000	m
Wavelength	0.3	m
Array number	8	/
Pulse number	8	/
PRF	2000	Hz
CNR	30	dB

4.1. Numerical Validation on Assumption 1 and Proposition 1

Firstly, we make a comparison of the output clutter-plus-noise powers generated by different STAP weight vectors (7), (17), and (36) for two platform velocities, $V = 150$ m/s and $V = 300$ m/s, corresponding to two local clutter DoFs, $Q_{Rd} = 10$ and $Q_{Rd} = 12$, respectively. Figure 2 shows the output clutter-plus-noise power versus the number of training samples with $f_0^t = 0.3$. It is obvious that for large L , both RD-STAP and RMT-RD-STAP have close output power. However, when the number L of training samples becomes small, the output clutter-plus-noise power by the RD-STAP rapidly increases, whereas the power by the RMT-RD-STAP is almost kept intact. These results illustrate that the assumption of complex white Gaussian noise in reduced dimension is reasonable from the viewpoint of reducing the output clutter-plus-noise power. Comparing Figure 2a with Figure 2b, we find that the local clutter DoFs have negative effects on the output clutter-plus-noise powers. The larger the local clutter DoFs are, the larger the optimal output powers are. But the gap between the output powers produced by RMT-RD-STAP and the optimal STAP changes slightly for these two local clutter DoFs.

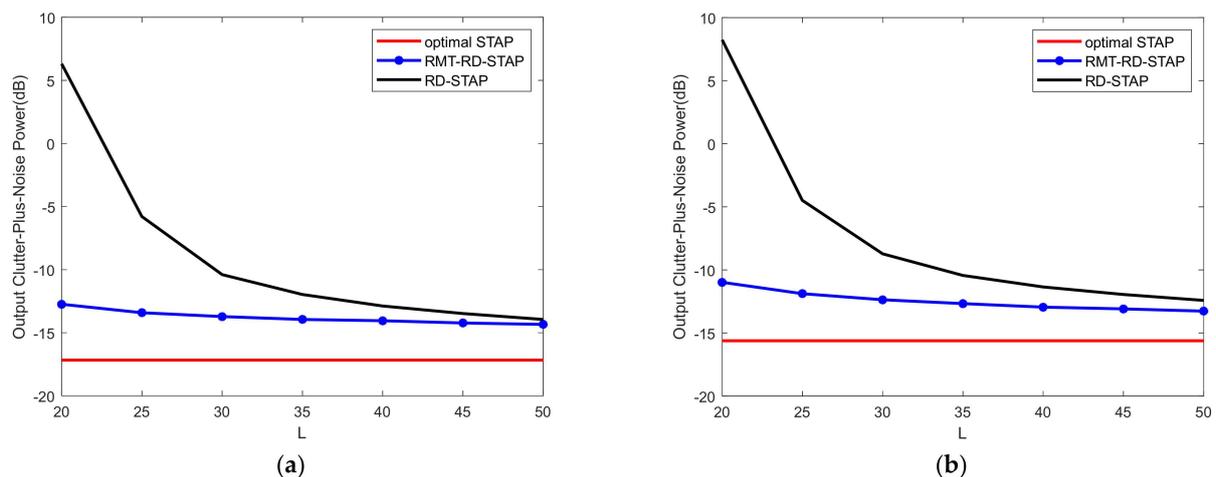


Figure 2. The output clutter-plus-noise power against the number L of the training samples. (a) $V = 150$ m/s and (b) $V = 300$ m/s.

Next, we make a comparison of the output clutter-plus-noise powers under different platform velocities for RMT-RD-STAP, $V = 150$ m/s and $V = 300$ m/s, respectively. Figure 3 shows the output clutter-plus-noise power versus the noise eigenvalues with $f_0^t = 0.3$, where $L = 20$ and $L = 30$ which correspond to two situations, $c_{Nrd} = 3N/20 > 1$ and $c_{Nrd} = 3N/30 < 1$. The noise-related eigenvalues are sorted in ascending order in the figures. It is seen that when $c_{Nrd} > 1$, the output clutter-plus-noise power decreases as the

noise power σ decreases, and when $c_{Nrd} < 1$, the output clutter-plus-noise power decreases as the noise power σ increases. These results verify the correctness of Proposition 1. By comparing the four figures, it is obvious that the above conclusion can still be maintained when the local clutter DoFs vary.

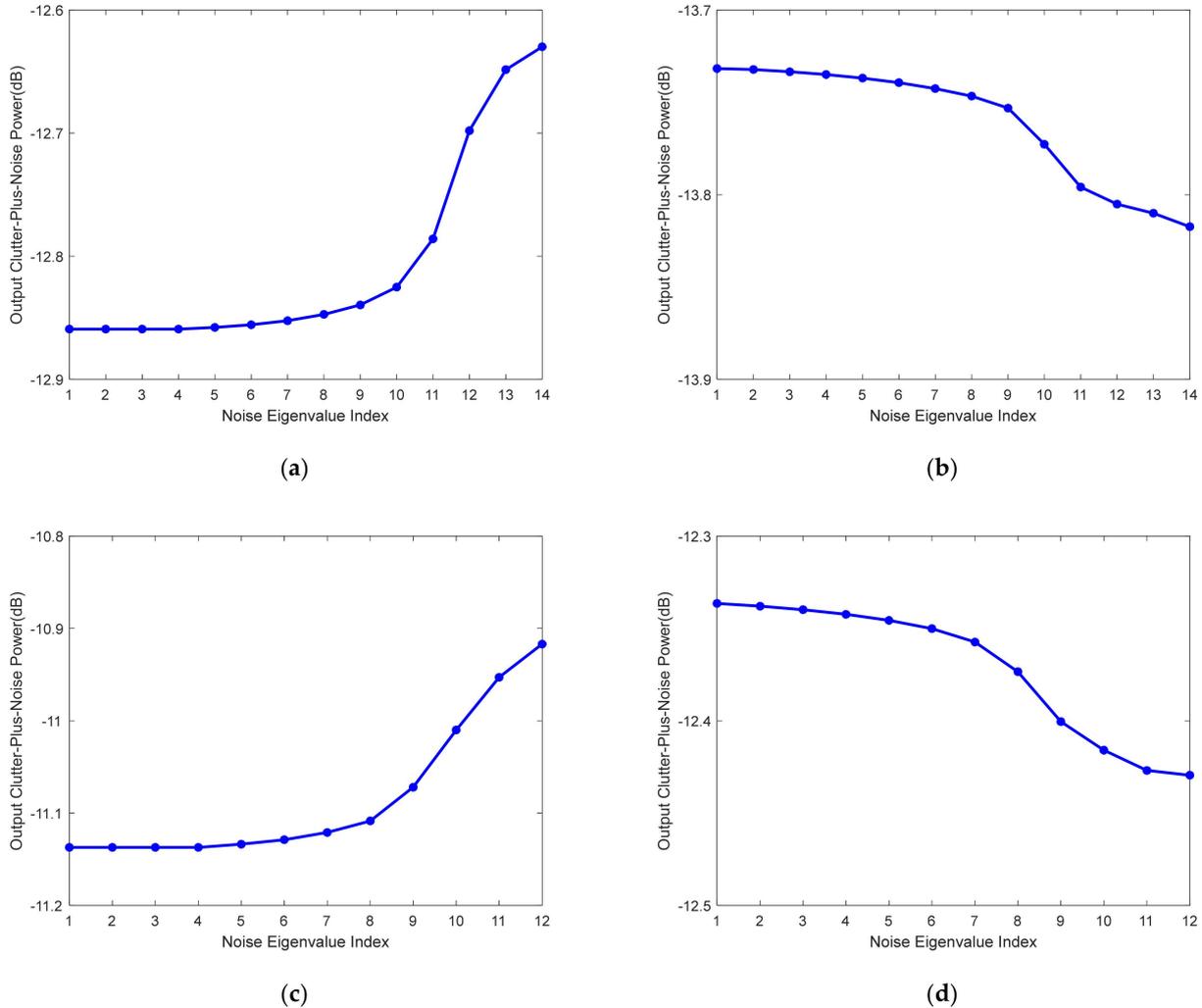


Figure 3. The output clutter-plus-noise powers versus the noise eigenvalues. (a) $V = 150$ m/s, $L = 20$; (b) $V = 150$ m/s, $L = 30$; (c) $V = 300$ m/s, $L = 20$; (d) $V = 300$ m/s, $L = 30$.

4.2. Performance Comparisons

In this subsection, output SINR loss, the ratio of the output SINR to the SNR achieved by a matched filter in a clutter-free environment [2], is taken as a performance metric to measure different STAP algorithms. For RD-STAP, the SINR loss is given as

$$SINR_{Loss}^{RD-STAP} = \frac{\sigma_n^2}{NK} \frac{\left| \mathbf{w}_{Lrd}^{optH} \mathbf{a}_{rd}(f_0^t, f_0^s) \right|^2}{\mathbf{w}_{Lrd}^{optH} \mathbf{R} \mathbf{w}_{Lrd}^{opt}}, \tag{39}$$

Similar loss functions are given by optimal STAP and RMT-RD-STAP. In the following experiments, the output SINR losses are presented by averaging 1000 independent runs.

Firstly, we present the variations of the output SINR losses as the normalized Doppler frequencies for different numbers of training samples. The results are shown in Figure 4 with $L = M$ and $L = 2M$. In such cases, we choose the maximum noise-related eigenvalue of RD SCM as the noise power. The airborne platform is assumed to move at the velocity $V = 150$ m/s. It is seen that the RMT-RD-STAP algorithm is superior to the RD-STAP

algorithm for small training samples. In Figure 4a, $L = M$, it is seen that RMT-RD-STAP is much superior to RD-STAP, and the SINR loss by RMT-RD-STAP decreases about 8 dB in comparison with RD-STAP. In Figure 4b, $L = 2M$, which means that the number of the training samples is sufficient, and the SINR losses of both two RD-STAP algorithms decrease in comparison with that in Figure 4a. Although RMT-RD-STAP and RD-STAP have almost similar performance, RMT-RD-STAP is still slightly superior to RD-STAP.

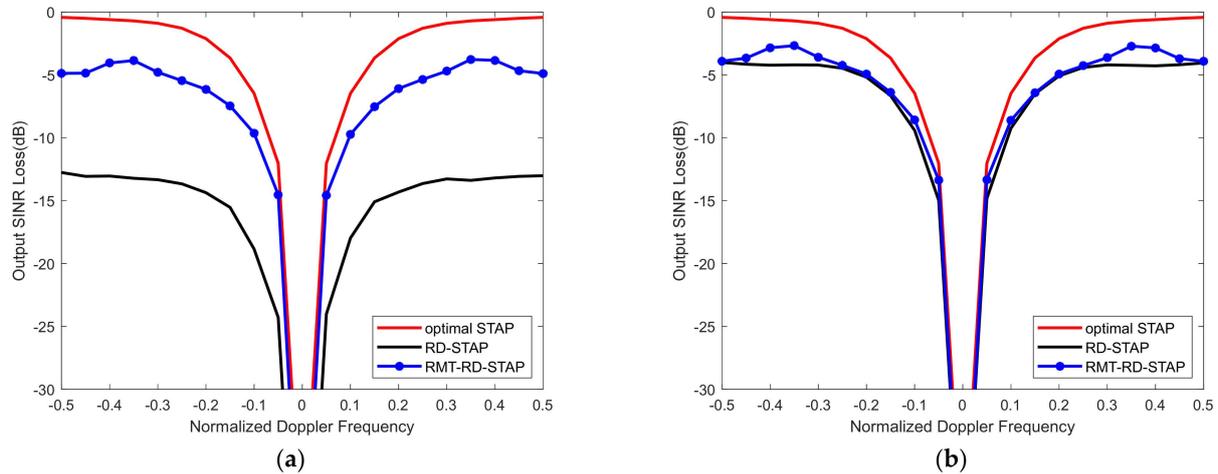


Figure 4. Output SINR losses versus normalized Doppler frequencies with $V = 150$ m/s. (a) $L = M$ training samples and (b) $L = 2M$ training samples.

Figure 5 shows the output SINR losses versus the number of training samples L with $f_0^t = 0.3$. It is seen that for large L , RMT-RD-STAP and RD-STAP achieve performance close to each other. However, as the number of training samples decreases, RMT-RD-STAP has distinct performance advantages over RD-STAP. Along with Figure 4, Figure 5 further demonstrates the advantages of the proposed RMT-RD-STAP algorithm.

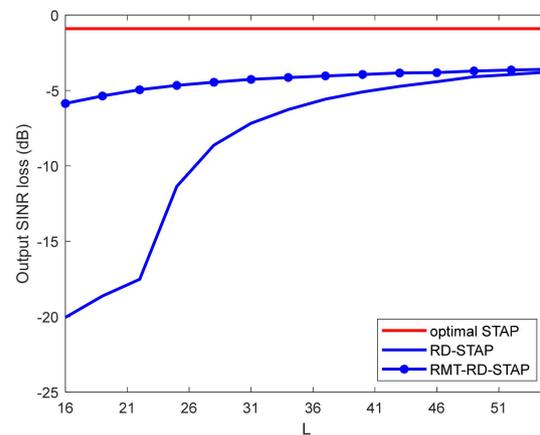


Figure 5. Output SINR losses versus the number L of the training samples.

Next, we present the variations of the output SINR losses as the function of normalized Doppler frequencies for different numbers of training samples. The results are shown in Figure 6 with $L = M$ and $L = 2M$, where the maximum noise-related sample eigenvalue is selected as the noise power. In the experiments, the platform is assumed to move at the velocity $V = 300$ m/s. Comparing Figure 6a,b, it is seen that RMT-RD-STAP is superior to RD-STAP for small-size training samples. From ref. [48], we know that the local clutter DoFs increase as the velocity increases. In comparison with Figure 4, we find that the performance of all the algorithms decreases. However, RMT-RD-STAP is still superior to

RD-STAP for small-size training samples. The SINR losses of RMT-RD-STAP in $L = M$ and $L = 2M$ decrease by about 8 dB and 1 dB in comparison with RD-STAP, respectively.

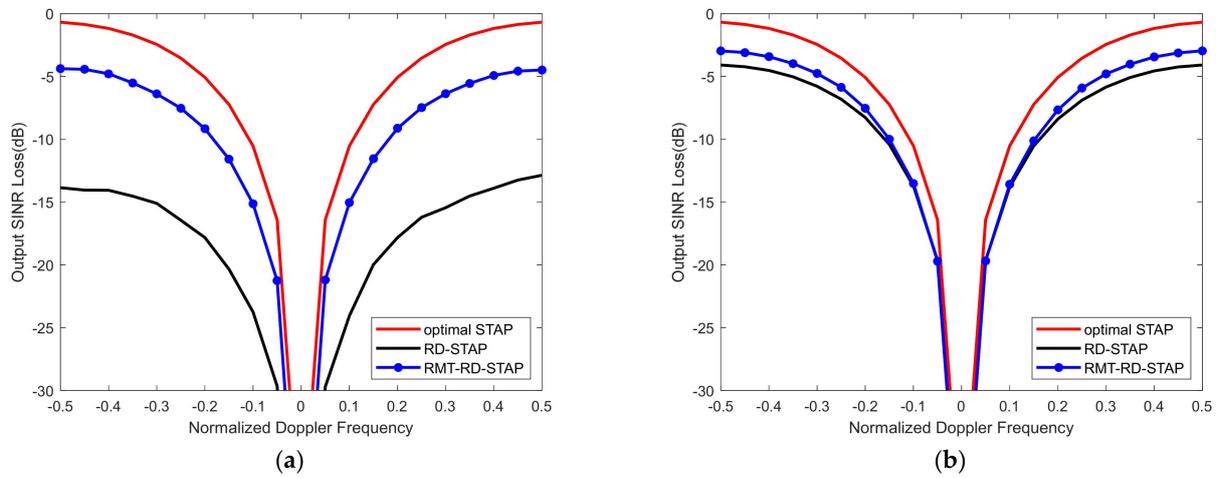


Figure 6. Output SINR losses versus normalized Doppler frequencies with $V = 300$ m/s. (a) $L = M$ training samples and (b) $L = 2M$ training samples.

In the implementation of the proposed RMT-RD-STAP algorithm, the local clutter DoFs need to be calculated according to radar parameters and platform parameters [48]. In a real environment, these parameters often deviate from the designed ones and then the calculated local clutter DoFs are different from the real local clutter DoFs. To demonstrate the robustness of the proposed algorithm, we use Figure 7 to show the output SINR losses versus the clutter of DoF errors with $f_0^t = 0.3$. In the experiments, the platform velocities are assumed to be $V = 150$ m/s and $V = 300$ m/s, and the calculated local clutter of DoFs are 10 and 12, respectively. The clutter DoF errors vary from -3 to 3 . The numbers of training samples are set to be $L = M$ and $L = 2M$, respectively. It is well-known that the conventional RD-STAP is independent of the local clutter DoFs estimation, whereas the proposed RMT-RD-STAP is related to the local clutter of DoFs estimation. As observed from Figure 7, under the predesigned clutter DoF errors scope, the performance of RD-STAP is stable for all four cases. Surprisingly, we also find that the proposed RMT-RD-STAP is not sensitive to the local clutter DoFs errors and it always provides stable output SINR losses. Furthermore, the proposed RMT-RD-STAP has superior performance to RD-STAP. Specifically, the SINR losses of RMT-RD-STAP in $L = M$ and $L = 2M$, respectively, decrease by about 8 dB and 1 dB in comparison with RD-STAP.

Under real clutter circumstance, the internal clutter motion (ICM) unavoidably exists due to wind and other factors [52], which is reflected by that the increase in clutter bandwidth in temporal domain. ICM will result in the RD-STAP performance deterioration. To illustrate the robustness of the proposed RMT-RD-STAP, we consider 2% ICM, which is a general situation. Figure 8 presents the output SINR losses versus normalized Doppler frequencies with 2% ICM. In the experiments, the platform velocities are assumed to be $V = 150$ m/s and $V = 150$ m/s, and the numbers of training samples are set to be $L = M$ and $L = 2M$, respectively. We find that, under small training samples size $L = M$, the proposed RMT-RD-STAP still has great performance advantage in comparison with RD-STAP. On the other hand, for large training samples size $L = 2M$, we find that RMT-RD-STAP and RD-STAP have almost the same performance, where RD-STAP achieves almost the optimal performance in this case according to the RMB rule.

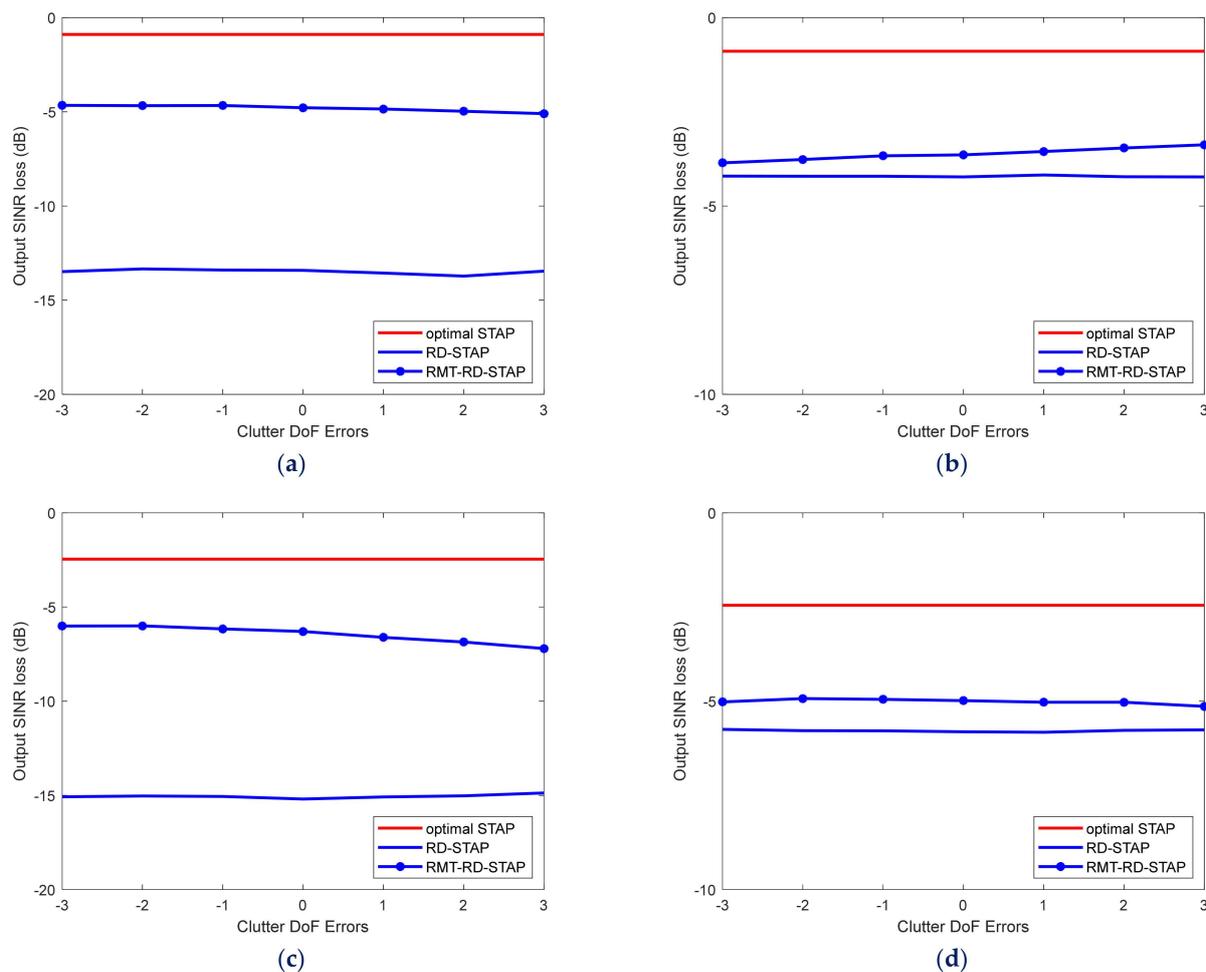


Figure 7. Output SINR losses versus clutter DoF errors. (a) $V = 150$ m/s, $L = M$; (b) $V = 150$ m/s, $L = 2M$; (c) $V = 300$ m/s, $L = M$ and (d) $V = 300$ m/s, $L = 2M$.

To demonstrate the effectiveness of the proposed RMT-RD-STAP, we finally apply it to the available real clutter data, Mountain Top data, i.e., *t38pre01v1CPI6* [53]. In this clutter data, the number K of transmitted pulses and the number N of array elements are 16 and 14, respectively. The PRF is 625 Hz and the distance resolution is 150 m. In the radar work time, there exist 403 range cells. We artificially add a target located in the 150-th range cell with a normalized Doppler frequency of 0.25 and the normalized spatial frequency 0.43 into the clutter data. In this case, the RD system DoFs is 42, and the local clutter DoFs is 16 in terms of [48]. Figure 9 presents the STAP output power verse the range cells with $L = 45$ and $L = 95$. In these two cases, four range cells adjacent to the CUT are regarded as the guard cells, and 45 and 95 range cells symmetrically located around the CUT are selected as training samples, respectively. It is seen that when $L = 95$, which satisfies the RMB rule, the proposed RMT-RD-STAP and RD-STAP have similar performance. However, when $L = 45$, which cannot satisfy the RMB rule, the RD-STAP will show a severe false alarm in the 144-th range cell, whereas the proposed RMT-RD-STAP has a distinct performance advantage and can accurately detect the target.

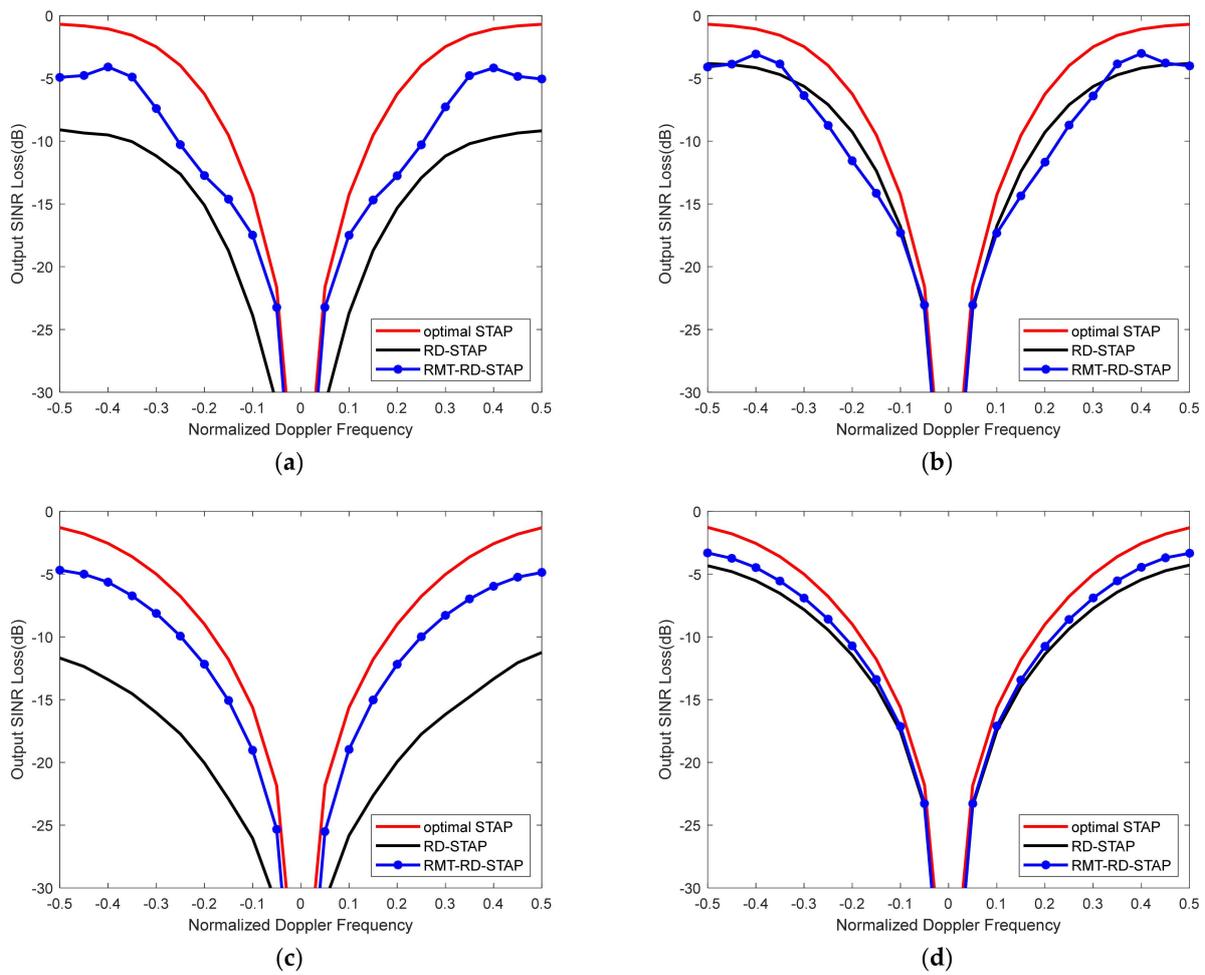


Figure 8. Output SINR losses versus normalized Doppler frequencies with 2% ICM. (a) $V = 150$ m/s, $L = M$; (b) $V = 150$ m/s, $L = 2M$; (c) $V = 300$ m/s, $L = M$ and (d) $V = 300$ m/s, $L = 2M$.

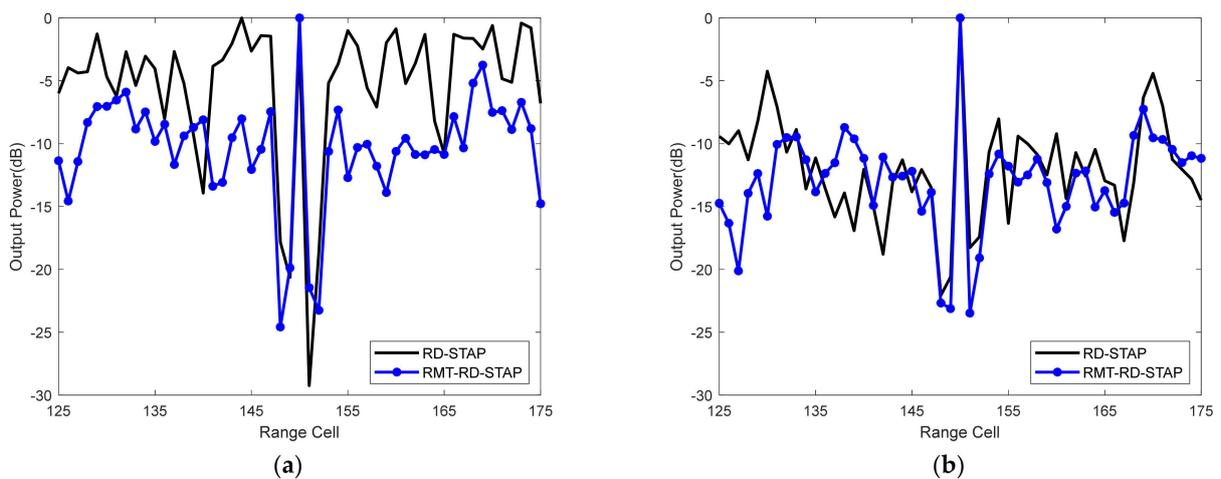


Figure 9. Output power versus range cell. (a) $L = 45$ training samples and (b) $L = 95$ training samples.

5. Conclusions

In this paper, we have proposed the RMT-based RD-STAP algorithm to enhance performance in the case of finite training samples. In terms of the spiked covariance model, the proposed algorithm estimates the inverse RD CNCM by optimally manipulating its sample eigenvalues while maintaining its sample eigenvectors. By minimizing the output

clutter-plus-noise power of RD-STAP, the clutter-related eigenvalues are estimated according to the clutter-related sample eigenvalues via RMT, and the noise-related eigenvalue is optimally selected from the noise-related sample eigenvalues. Benefitted from RMT, the proposed RMT-RD-STAP greatly outperforms the conventional RD-STAP when the RMB rule is not satisfied and has superior performance when the number of training samples satisfies the RMB rule.

Author Contributions: Conceptualization, D.S., S.C. and Z.L.; Funding acquisition, Z.L.; Methodology, D.S., S.C. and Z.L.; Software, D.S. and Q.F.; Supervision, S.C. and Z.L.; Validation, D.S. and Q.F.; Visualization, D.S. and Q.F.; Writing—original draft, D.S.; Writing—review and editing, S.C., F.X., and Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the National Natural Science Foundation of China under Grant 6217011933 and in part by Natural Science Foundation of Jiangsu Province under Grant BK20221486.

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

Appendix A

To solve the problem (37), we need to analyze the relationship between $P_{rd}(\hat{\mathbf{w}}_{rd}^{opt}(\hat{\mathbf{g}}^*))$ and σ . We consider the partial derivative

$$\frac{\partial P_{rd}}{\partial \sigma} = \frac{\partial P_{rd}}{\partial \hat{\mathbf{g}}_i^*} \frac{\partial \hat{\mathbf{g}}_i^*}{\partial \hat{\beta}_i} \frac{\partial \hat{\beta}_i}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \sigma}, \quad (\text{A1})$$

In the following, we will analyze the right four terms, respectively.

Firstly, $\sigma \alpha_i$ is the sample eigenvalue of RD SCM, and α_i will increase as σ decreases,

$$\frac{\partial \alpha_i}{\partial \sigma} < 0. \quad (\text{A2})$$

Secondly, in terms of the relationship $\hat{\beta}_i = \left[\alpha_i - 1 - c_{Nrd} + \sqrt{(\alpha_i - 1 - c_{Nrd})^2 - 4c_{Nrd}} \right] / 2$, it is obvious that $\hat{\beta}_i$ will increase as α_i increases, then

$$\frac{\partial \hat{\beta}_i}{\partial \alpha_i} > 0. \quad (\text{A3})$$

Thirdly, from (32), we can make a replacement

$$\varepsilon = \frac{\sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q}}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) - \sum_{q=1}^{Q_{rd}} \hat{k}_{rdq} + \sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q}}, \quad (\text{A4})$$

Then

$$\frac{\partial \hat{\mathbf{g}}_i^*}{\partial \hat{\beta}_i} = \frac{-(\hat{\beta}_i^2 + 2c_{Nrd} \hat{\beta}_i + c_{Nrd})}{(\hat{\beta}_i^2 + \hat{\beta}_i)^2} (\varepsilon - \hat{\beta}_i) + \frac{\hat{\beta}_i + c_{Nrd}}{\hat{\beta}_i^2 + \hat{\beta}_i} \left(\frac{\partial \varepsilon}{\partial \hat{\beta}_i} - 1 \right), \quad (\text{A5})$$

Practically, the clutter power in airborne radar system is generally greatly large, and the target is commonly not independent with clutter. Therefore, $\hat{\beta}_q$ is generally large, and $\hat{k}_{rdq} \approx 0$. Meanwhile, it is obvious that

$$\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) \gg \sum_{q=1}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q^2} - \sum_{q=1}^{Q_{rd}} \hat{k}_{rdq}, \quad (\text{A6})$$

Then

$$\varepsilon \approx 0, \tag{A7}$$

Let us make a replacement for ε

$$U_1 = \sum_{\substack{q=1 \\ q \neq i}}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q}$$

$$U_2 = \sum_{\substack{q=1 \\ q \neq i}}^{Q_{rd}} \frac{c_{Nrd} \hat{k}_{rdq}}{\hat{\beta}_q^2} + \sum_{q=1}^{Q_{rd}} \hat{k}_{rdq}$$
(A8)

Then ε can be expressed as

$$\varepsilon = \frac{\frac{c_{Nrd} \hat{k}_{rdi}}{\hat{\beta}_i} + U_1}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) - \frac{c_{Nrd} \hat{k}_{rdi}}{\hat{\beta}_i^2} + U_2}}, \tag{A9}$$

Note that $\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) \gg U_2$; therefore we have

$$\varepsilon \approx \frac{\frac{c_{Nrd} \hat{k}_{rdi}}{\hat{\beta}_i} + U_1}{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) - \frac{c_{Nrd} \hat{k}_{rdi}}{\hat{\beta}_i^2}}, \tag{A10}$$

Then, we can obtain the partial derivative of (A10) as

$$\frac{\partial \varepsilon}{\partial \hat{\beta}_i} = \frac{-\frac{\hat{k}_{rdi}}{c_{Nrd}} (\hat{\beta}_i + U_1 c_{Nrd})^2 + \hat{k}_{rdi} U_1^2 c_{Nrd} + \hat{k}_{rdi}^2}{(\frac{\hat{\beta}_i^2}{c_{Nrd}} + \hat{k}_{rdi})^2} \approx \frac{-\frac{\hat{k}_{rdi}}{c_{Nrd}} \hat{\beta}_i^2}{\frac{\hat{\beta}_i^4}{c_{Nrd}^2}} = -\frac{\hat{k}_{rdi} c_{Nrd}}{\hat{\beta}_i^2} \approx 0, \tag{A11}$$

Substituting (A7) and (A11) into (A5), we have

$$\frac{\partial \hat{\delta}_i^*}{\partial \hat{\beta}_i} \approx \frac{\hat{\beta}_i^2 + 2c_{Nrd} \hat{\beta}_i + c_{Nrd}}{\hat{\beta}_i (\hat{\beta}_i + 1)^2} - \frac{\hat{\beta}_i + c_{Nrd}}{\hat{\beta}_i^2 + \hat{\beta}_i} = \frac{c_{Nrd} - 1}{(\hat{\beta}_i + 1)^2}, \tag{A12}$$

According to (A12), notice that the relationship between $\hat{\delta}_i^*$ and $\hat{\beta}_i$ depends upon c_{Nrd} . If $c_{Nrd} > 1$, i.e., $L < M$, $\partial \hat{\delta}_i^* / \partial \hat{\beta}_i > 0$, otherwise $\partial \hat{\delta}_i^* / \partial \hat{\beta}_i < 0$.

Finally, we consider that $\partial P_{rd} / \partial \hat{\delta}_i^* \cdot P_{rd}$ can be approximated as

$$P_{rd} = \sigma \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \left(\mathbf{I}_M + \sum_{q=1}^{Q_{rd}} \hat{\delta}_q^* \mathbf{d}_q \mathbf{d}_q^H \right) \mathbf{R}_{rd} \left(\mathbf{I}_M + \sum_{i=1}^{Q_{rd}} \hat{\delta}_i^* \mathbf{d}_i \mathbf{d}_i^H \right) \mathbf{a}_{rd}(f_0^t, f_0^s)}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \left(\mathbf{I}_M + \sum_{i=1}^{Q_{rd}} \hat{\delta}_i^* \mathbf{d}_i \mathbf{d}_i^H \right) \mathbf{a}_{rd}(f_0^t, f_0^s) \right]^2}$$

$$\approx \sigma \frac{\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + \sum_{i=1}^{Q_{rd}} \hat{k}_{rdi} \beta_i + \sum_{i=1}^{Q_{rd}} s_{rdi}^2 \hat{k}_{rdi} \beta_i \left(\hat{\delta}_i^* + \frac{1}{s_{rdi}} \right)^2}{\left[\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + \sum_{i=1}^{Q_{rd}} s_{rdi} \hat{k}_{rdi} \left(\hat{\delta}_i^* + \frac{1}{s_{rdi}} \right) \right]^2}$$

$$= \sigma \frac{P_1 + s_{rdi}^2 \hat{k}_{rdi} \beta_i \left(\hat{\delta}_i^* + \frac{1}{s_{rdi}} \right)^2}{\left[P_2 + s_{rdi} \hat{k}_{rdi} \left(\hat{\delta}_i^* + \frac{1}{s_{rdi}} \right) \right]^2}$$
(A13)

where

$$P_1 = \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + \sum_{i=1}^{Q_{rd}} \hat{k}_{rdi} \beta_i + \sum_{\substack{q=1 \\ q \neq i}}^{Q_{rd}} s_{rdq}^2 \hat{k}_{rdq} \beta_q \left(\hat{g}_q^* + \frac{1}{s_{rdq}} \right)^2, \quad (\text{A14})$$

$$P_2 = \mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) + \sum_{\substack{q=1 \\ q \neq i}}^{Q_{rd}} s_{rdq} \hat{k}_{rdq} \left(\hat{g}_q^* + \frac{1}{s_{rdq}} \right)$$

Then, we can obtain the partial derivative of (A13) as

$$\frac{\partial P}{\partial \hat{g}_i^*} \approx \sigma \frac{2s_{rdi}^2 \hat{k}_{rdi} \beta_i \left(\hat{g}_i^* + \frac{1}{s_{rdi}} \right) \left[P_2 + s_{rdi} \hat{k}_{rdi} \left(\hat{g}_i^* + \frac{1}{s_{rdi}} \right) \right] - 2s_{rdi} \hat{k}_{rdi} \left[P_1 + s_{rdi}^2 \hat{k}_{rdi} \beta_i \left(\hat{g}_i^* + \frac{1}{s_{rdi}} \right)^2 \right]}{\left[P_2 + s_{rdi} \hat{k}_{rdi} \left(\hat{g}_i^* + \frac{1}{s_{rdi}} \right) \right]^3}, \quad (\text{A15})$$

As the clutter power is commonly high, we have $s_{rdi} = (1 - c_{Nrd} / \beta_i^2) / (1 + c_{Nrd} / \beta_i) \approx 1$. Substituting (32) into (A15), we have

$$\begin{aligned} \frac{\partial P}{\partial \hat{g}_i^*} &\approx \sigma \frac{2\hat{k}_{rdi} \beta_i (\hat{g}_i^* + 1) [P_2 + \hat{k}_{rdi} (\hat{g}_i^* + 1)] - 2\hat{k}_{rdi} [P_1 + \hat{k}_{rdi} \beta_i (\hat{g}_i^* + 1)^2]}{[P_2 + \hat{k}_{rdi} (\hat{g}_i^* + 1)]^3} \\ &= \sigma \frac{2\hat{k}_{rdi} [\beta_i (\hat{g}_i^* + 1) P_2 - P_1]}{[P_2 + \hat{k}_{rdi} (\hat{g}_i^* + 1)]^3} \end{aligned} \quad (\text{A16})$$

The denominator of (A16) is larger than 0 evidently, then the sign of (A16) completely depends on the numerator of (A16). Note that

$$2\hat{k}_{rdi} [\beta_i (\hat{g}_i^* + 1) P_2 - P_1] \approx -\mathbf{a}_{rd}^H(f_0^t, f_0^s) \mathbf{a}_{rd}(f_0^t, f_0^s) c_{Nrd} - \sum_{i=1}^{Q_{rd}} \hat{k}_{rdi} \beta_i. \quad (\text{A17})$$

Therefore, it is obvious that $\partial P_{rd} / \partial \hat{g}_i^* < 0$.

In summary, if $c_{Nrd} > 1$, $\partial P_{rd} / \partial \sigma > 0$, the output clutter-plus-noise power decreases as σ decreases; if $c_{Nrd} < 1$, $\partial P_{rd} / \partial \sigma < 0$, the output clutter-plus-noise power decreases as σ increases.

References

- Brennan, L.E.; Reed, L.S. Theory of adaptive radar. *IEEE Trans. Aerosp. Electron. Syst.* **1973**, *AES-9*, 237–252. [\[CrossRef\]](#)
- Ward, J. *Space-Time Adaptive Processing for Airborne Radar*; Technical Report 1015; MIT Lincoln Lab.: Lexington, MA, USA, 1994.
- Melvin, W.L. A STAP overview. *IEEE Aerosp. Electron. Syst. Mag.* **2004**, *19*, 19–35. [\[CrossRef\]](#)
- Reed, I.S.; Mallett, J.D.; Brennan, L.E. Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *AES-10*, 853–863. [\[CrossRef\]](#)
- Klemm, R. *Principles of Space-Time Adaptive Processing*, 3rd ed.; IET: London, UK, 2006.
- Guerci, J.R. *Space-Time Adaptive Processing for Radar*; Artech House: Norwood, MA, USA, 2003.
- Yang, Z.; Li, X.; Wang, H.; Jiang, W. On Clutter Sparsity Analysis in Space-Time Adaptive Processing Airborne Radar. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1214–1218. [\[CrossRef\]](#)
- Sen, S. Low-Rank Matrix Decomposition and Spatio-Temporal Sparse Recovery for STAP Radar. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 1510–1523. [\[CrossRef\]](#)
- 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**, *60*, 674–686. [\[CrossRef\]](#)
- Zhang, W.; An, R.; He, N.; He, Z.; Li, H. Reduced dimension STAP based on sparse recovery in heterogeneous clutter environments. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *56*, 785–795. [\[CrossRef\]](#)
- Guerci, J.; Baranoski, E. Knowledge-aided adaptive radar at DARPA: An overview. *IEEE Signal Process. Mag.* **2006**, *23*, 41–50. [\[CrossRef\]](#)
- Besson, O.; Tourneret, J.-Y.; Bidon, S. Knowledge-Aided Bayesian Detection in Heterogeneous Environments. *IEEE Signal Process. Lett.* **2007**, *14*, 355–358. [\[CrossRef\]](#)
- Riedl, M.; Potter, L.C. Knowledge-Aided Bayesian Space-Time Adaptive Processing. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1850–1861. [\[CrossRef\]](#)
- Riedl, M.; Potter, L.C. Multimodel Shrinkage for Knowledge-Aided Space-Time Adaptive Processing. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 2601–2610. [\[CrossRef\]](#)

15. Duan, K.; Chen, H.; Xie, W.; Wang, Y. Deep learning for high-resolution estimation of clutter angle-Doppler spectrum in STAP. *IET Radar Sonar Navig.* **2022**, *16*, 193–207. [[CrossRef](#)]
16. Zou, B.; Wang, X.; Feng, W.; Zhu, H.; Lu, F. DU-CG-STAP Method Based on Sparse Recovery and Unsupervised Learning for Airborne Radar Clutter Suppression. *Remote Sens.* **2022**, *14*, 3472. [[CrossRef](#)]
17. Zhu, H.; Feng, W.; Feng, C.; Zou, B.; Lu, F. Deep Unfolding Based Space-Time Adaptive Processing Method for Airborne Radar. *J. Radars* **2022**, *11*, 1–16.
18. 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; pp. 425–430.
19. Sarkar, T.K.; Wang, H.; Park, S.; Adve, R.; Koh, J.; Kim, K.; Zhang, Y.; Wicks, M.; Brown, R. A deterministic least-squares approach to space-time adaptive processing (STAP). *IEEE Trans. Antennas Propag.* **2001**, *49*, 91–103. [[CrossRef](#)]
20. Brown, R.; Schneible, R.; Wicks, M.; Wang, H.; Zhang, Y. STAP for clutter suppression with sum and difference beams. *IEEE Trans. Aerosp. Electron. Syst.* **2000**, *36*, 634–646. [[CrossRef](#)]
21. Zhang, W.; He, Z.; Li, J.; Liu, H.; Sun, Y. A Method for Finding Best Channels in Beam-Space Post-Doppler Reduced-Dimension STAP. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 254–264. [[CrossRef](#)]
22. Klemm, R. Adaptive airborne MTI: An auxiliary channel approach. *IEE Proc. F Commun. Radar Signal Process.* **1987**, *134*, 269–276. [[CrossRef](#)]
23. Wang, H.; Cai, L. On adaptive spatial-temporal processing for airborne surveillance radar systems. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 660–670. [[CrossRef](#)]
24. Karoui, N.E. Spectrum estimation for large dimensional covariance matrices using random matrix theory. *Ann. Statist.* **2008**, *36*, 2757–2790. [[CrossRef](#)]
25. Daniels, R.; Kass, M.J. Shrinkage estimators for covariance matrices. *Biometrics* **2001**, *57*, 1173–1184. [[CrossRef](#)] [[PubMed](#)]
26. Fan, J.; Fan, Y.; Lv, J. High dimensional covariance matrix estimation using a factor model. *J. Econ.* **2008**, *147*, 186–197. [[CrossRef](#)]
27. Yang, L.; McKay, M.R.; Couillet, R. High-Dimensional MVDR Beamforming: Optimized Solutions Based on Spiked Random Matrix Models. *IEEE Trans. Signal Process.* **2018**, *66*, 1933–1947. [[CrossRef](#)]
28. Bai, Z.; Silverstein, J.W. *Spectral Analysis of Large Dimensional Random Matrices*; Springer: New York, NY, USA, 2010; Volume 20.
29. Silverstein, J.; Choi, S. Analysis of the Limiting Spectral Distribution of Large Dimensional Random Matrices. *J. Multivar. Anal.* **1995**, *54*, 295–309. [[CrossRef](#)]
30. Schott, J.R. Asymptotics of eigenprojections of correlation matrices with some applications in principal components analysis. *Biometrika* **1997**, *84*, 327–337. [[CrossRef](#)]
31. Bai, Z.D.; Silverstein, J.W. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **1998**, *26*, 316–345. [[CrossRef](#)]
32. Silverstein, J. Strong Convergence of the Empirical Distribution of Eigenvalues of Large Dimensional Random Matrices. *J. Multivar. Anal.* **1995**, *55*, 331–339. [[CrossRef](#)]
33. Ledoit, O.; Wolf, M. Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *J. Multivar. Anal.* **2015**, *139*, 360–384. [[CrossRef](#)]
34. Ledoit, O.; Wolf, M. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Rev. Financ. Stud.* **2017**, *30*, 4349–4388. [[CrossRef](#)]
35. Girko, V. Asymptotic behavior of eigenvalues of empirical covariance matrices. *Theory Probab. Math. Stat.* **1992**, *44*, 37–44.
36. Rubio, F.; Mestre, X.; Palomar, D.P. Performance Analysis and Optimal Selection of Large Minimum Variance Portfolios Under Estimation Risk. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 337–350. [[CrossRef](#)]
37. Anderson, T.W. Asymptotic Theory for Principal Component Analysis. *Ann. Math. Stat.* **1963**, *34*, 122–148. [[CrossRef](#)]
38. Gupta, R. Asymptotic theory for principal component analysis in the complex case. *J. Indian Statist. Assoc.* **1965**, *3*, 97–106.
39. Marcenko, V.A.; Pastur, L.A. Distribution of Eigenvalues for Some Sets of Random Matrices. *Math. USSR-Sb.* **1967**, *1*, 457–483. [[CrossRef](#)]
40. Baik, J.; Arous, G.B.; Pécché, S. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **2005**, *33*, 1643–1697. [[CrossRef](#)]
41. Baik, J.; Silverstein, J.W. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivar. Anal.* **2006**, *97*, 1382–1408. [[CrossRef](#)]
42. Paul, D. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sin.* **2007**, *17*, 1617–1642.
43. Couillet, R. Robust spiked random matrices and a robust G-MUSIC estimator. *J. Multivar. Anal.* **2015**, *140*, 139–161. [[CrossRef](#)]
44. Robinson, B.D.; Malinas, R.; Hero, A.O. Space-Time Adaptive Detection at Low Sample Support. *IEEE Trans. Signal Process.* **2021**, *69*, 2939–2954. [[CrossRef](#)]
45. Combernoux, A.; Pascal, F.; Lesturgie, M.; Ginolhac, G. Performances of low rank detectors based on random matrix theory with application to STAP. In Proceedings of the 2014 International Radar Conference, Lille, France, 13–17 October 2014; pp. 1–4.
46. Pafka, S.; Kondor, I. Noisy covariance matrices and portfolio optimization II. *Phys. A Stat. Mech. Its Appl.* **2003**, *319*, 487–494. [[CrossRef](#)]
47. Pafka, S.; Kondor, I. Estimated correlation matrices and portfolio optimization. *Phys. A Stat. Mech. Table Its Appl.* **2004**, *343*, 623–634. [[CrossRef](#)]

48. Zhang, Z.; Xie, W.; Hu, W.; Yu, W. Local degrees of freedom of airborne array radar clutter for STAP. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 97–101.
49. Mestre, X. On the Asymptotic Behavior of the Sample Estimates of Eigenvalues and Eigenvectors of Covariance Matrices. *IEEE Trans. Signal Process.* **2008**, *56*, 5353–5368. [[CrossRef](#)]
50. Mestre, X. Improved Estimation of Eigenvalues and Eigenvectors of Covariance Matrices Using Their Sample Estimates. *IEEE Trans. Inf. Theory* **2008**, *54*, 5113–5129. [[CrossRef](#)]
51. Couillet, R.; Debbah, M. *Random Matrix Methods for Wireless Communications*; Cambridge Univ. Press: Cambridge, UK, 2011.
52. Melvin, W.L.; Showman, G.A. An approach to knowledge-aided covariance estimation. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 1021–1042. [[CrossRef](#)]
53. Titi, G.; Marshall, D. The ARPA/NAVY Mountaintop Program: Adaptive signal processing for airborne early warning radar. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, Atlanta, GA, USA, 9 May 1996; IEEE: Piscataway Township, NJ, USA, 2002; Volume 2, pp. 1165–1168.