

Article

Information Geometry for Covariance Estimation in Heterogeneous Clutter with Total Bregman Divergence

Xiaoqiang Hua *, Yongqiang Cheng *, Hongqiang Wang and Yuliang Qin

School of Electronic Science, National University of Defence Technology, Changsha 410073, China; oliverwhq@tom.com (H.W.); yuliang.qin@gmail.com (Y.Q.)

* Correspondence: huaxiaoqiang12@nudt.edu.cn (X.H.); yqcheng@nudt.edu.cn (Y.C.)

Received: 31 January 2018; Accepted: 6 April 2018; Published: 8 April 2018



Abstract: This paper presents a covariance matrix estimation method based on information geometry in a heterogeneous clutter. In particular, the problem of covariance estimation is reformulated as the computation of geometric median for covariance matrices estimated by the secondary data set. A new class of total Bregman divergence is presented on the Riemannian manifold of Hermitian positive-definite (HPD) matrix, which is the foundation of information geometry. On the basis of this divergence, total Bregman divergence medians are derived instead of the sample covariance matrix (SCM) of the secondary data. Unlike the SCM, resorting to the knowledge of statistical characteristics of the sample data, the geometric structure of matrix space is considered in our proposed estimators, and then the performance can be improved in a heterogeneous clutter. At the analysis stage, numerical results are given to validate the detection performance of an adaptive normalized matched filter with our estimator compared with existing alternatives.

Keywords: covariance matrix estimation; total Bregman divergence; information geometry; adaptive normalized matched filter

1. Introduction

The estimation of the disturbance covariance matrix is an important subject in the field of advanced radar signal processing. Many algorithms manipulate the covariance matrix of sample data, such as array signal processing [1,2], multichannel signal processing [3,4], adaptive radar detection [5–7], and space-time adaptive processing [8–10]. The commonly used estimator is the sample covariance matrix (SCM), which is often derived from the maximum-likelihood (ML) of K N -dimensional secondary data. A condition is assumed that these K secondary data are independent and identical distributed zero-mean complex circular Gaussian vectors. Usually, the solution of this ML exists when the amount of secondary data K is greater than the matrix dimension N ($K \geq N$). In particular, it can achieve a good performance for $K \geq 2N$ [11]. Unfortunately, in real heterogeneous clutter, the statistical distribution of the whole environment is very difficult to obtain, since the secondary data could be contaminated by power variations, clutter discretions, and/or outliers. Therefore, it is very necessary and meaningful to obtain the estimation of disturbance matrix that is not relying on the statistical characterization of the whole environment.

Many covariance estimation algorithms derived from the geometry of matrix space, not resorting to the statistical characterization of sample data, are reported in the literature. For instance, the Riemannian mean and median of covariance matrices is proposed to design the radar target detector [12–16]. In [17–20], the median is used for covariance estimation in many radar processing applications. The Riemannian mean is utilized to estimate the covariance matrix in space-time adaptive

processing [21–23]. The results have shown that the projection algorithm with Riemannian mean can yield significant performance gains. In [24,25], some geometric barycenter and medians are proposed for radar training data selection in a homogeneous environment. In [26], a geometric method is presented for covariance estimation, where each estimator is associated with a given unitary invariant norm and performs the sample covariance matrix projection into a specific set of structured covariance matrices. Recently, geometric Barycenters of symmetric positive definite matrices have been used for covariance estimators in compound-Gaussian clutter [27]. Moreover, in our previous work [28,29], we have proposed a lot of divergence means and medians of covariance matrices computed in a neighborhood of the cell under test for radar target detection. Finally, the geometric approach is used also in many other applications; for instance, the Bhattacharyya mean and median are exploited for diffusion tensor magnetic resonance (DT-MRI) image segmentation [30,31]. In these contexts, the geometric approaches have achieved good performances.

In this paper, a new class of total Bregman divergence (tBD) is presented on the Riemannian manifold. Based on the tBD, the three geometric medians, including total square loss (TSL) median, total von Neumann (TVN) median, and total LogDet (TLD) median, are derived, and are used as the estimators of disturbance matrix. As a matter of fact, the SCM of K secondary data can be seen as the arithmetic mean of K autocovariance matrices of rank one. The arithmetic mean ignores the fact that these matrices lie in a nonlinear matrix space, the Riemannian manifold, which is the foundation of information geometry. As the geometric medians are not relying on the statistical characteristics of the sample data in heterogeneous clutter, the performance of covariance estimation can be improved.

The rest of this paper is organized as follows: Section 2 formulates the problem of covariance estimation from information geometry; the total Bregman divergence-based estimators are presented in Section 3; Section 4 gives the theoretical analysis about the robustness of the proposed estimators; experimental results are presented in Section 5; Section 6 concludes our work.

Notation: Here are some notations for the description of this article. A matrix \mathbf{A} and a vector \mathbf{x} are noted as uppercase bold and lowercase bold, respectively. The conjugate transpose of matrix \mathbf{A} is denoted as \mathbf{A}^\dagger . $tr(\mathbf{A})$ is the trace of matrix \mathbf{A} . $det(\mathbf{A})$ is the determinant of matrix \mathbf{A} . \mathbf{I} denotes the identity matrix. Finally, $\mathbb{E}(\cdot)$ denotes statistical expectation.

2. Problem Formulated from Information Geometry

In this section, a heterogeneous environment is considered for covariance estimation. The secondary data $\mathbf{r}_k, k = 1, 2, \dots, K$ is a spherically invariant random vector, and can be expressed as,

$$\mathbf{r}_k = \sqrt{\tau_k} \mathbf{g}_k, k = 1, \dots, K, \quad (1)$$

where τ_k is a nonnegative scalar random variable, and \mathbf{g}_k is a N -dimensional circularly symmetric zero-mean vectors with an arbitrary joint statistical distribution and sharing the same covariance matrix,

$$\mathbb{E}[\mathbf{g}_k \mathbf{g}_k^\dagger] = \mathbf{\Sigma}, k = 1, \dots, K, \quad (2)$$

where \dagger denotes the conjugate transpose. Since the knowledge of the statistical characterization of the secondary data is not known in heterogeneous clutter, the classic approaches, e.g., ML and minimum mean-square error, cannot be applied for covariance estimation of the sample data. Thus, other covariance estimators, not dependent on the probability distribution of the whole environment, are very promising. In this paper, a covariance estimation method based on information geometry is proposed.

Recall that the SCM of K secondary data $\mathbf{r}_k, k = 1, 2, \dots, K$ can be given as

$$\hat{\Sigma} = \frac{1}{K} \sum_{k=1}^K \mathbf{r}_k \mathbf{r}_k^\dagger, \tag{3}$$

where $\mathbf{r}_k \mathbf{r}_k^\dagger$ is an autocovariance matrix of rank one, and is singular. It can be noted from Equation (3) that $\hat{\Sigma}$ is the arithmetic mean of K autocovariance matrices. In fact, these matrices lie in the nonlinear Hermitian matrix space, as the sample data is complex. It is well known that Hermitian positive-definite (HPD) matrices form a differentiable Riemannian (and also a Finsler) manifold [32,33], that is the most studied example of a manifold with nonpositive curvature [34,35]. In order to facilitate the analysis, the singular matrix $\mathbf{r}_k \mathbf{r}_k^\dagger$ is positive by adding an identity matrix, as $\mathbf{R}_k = \mathbf{r}_k \mathbf{r}_k^\dagger + \mathbf{I}$. Then, the disturbance covariance matrix can be estimated by a median related to a divergence on the Riemannian manifold of HPD matrices.

As illustrated in Figure 1, the geometric median is performed on the Riemannian manifold of HPD matrices with a non-Euclidean metric, whereas the arithmetic mean is considered in the Euclidean space. The difference implies that the different geometric structures are considered in these two estimators. Moreover, for different metrics on the Riemannian manifold, the performance of covariance estimation may be very different. These results will be found in our other reports [36]. In the next section, a new class of tBD is proposed on the Riemannian manifold. On basis of the tBD, the tBD median is derived and used for the estimator.

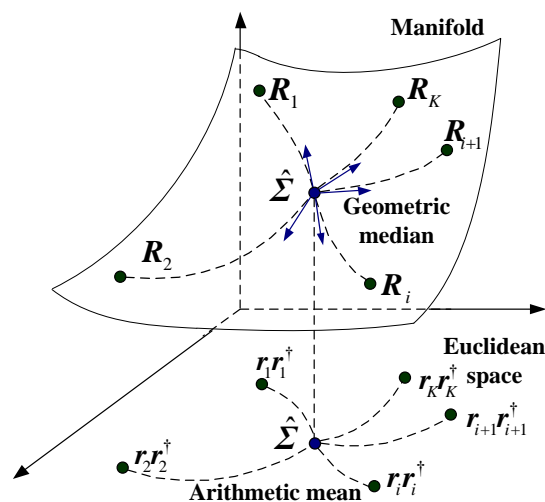


Figure 1. The arithmetic mean and geometric median.

3. Total Bregman Divergence-Based Estimators on the Manifold

In this section, a new class of tBD is proposed on the Riemannian manifold. Then, the medians associated with the tBD are derived.

3.1. The Geometry of HPD Matrices

Let $\mathbb{H}(n) = \{\mathbf{A} | \mathbf{A} = \mathbf{A}^H\}$ denote the space of $n \times n$ Hermitian matrix. For a Hermitian matrix \mathbf{A} , if the quadratic form $\mathbf{x}^H \mathbf{A} \mathbf{x} > 0, \forall \mathbf{x} \in \mathbb{C}(n)$, then \mathbf{A} is an HPD matrix, where $\mathbb{C}(n)$ is the space of n -dimensional complex vectors. All $n \times n$ HPD matrices consist of a positive-definite Hermitian matrix space $\mathbb{P}(n)$,

$$\mathbb{P}(n) = \{\mathbf{A} | \mathbf{A} > 0, \mathbf{A} \in \mathbb{H}(n)\}, \tag{4}$$

which forms a Riemannian manifold of dimension $n(n + 1)/2$ with a constant non-positive curvature [35]. For a point \mathbf{A} on the Riemannian manifold, the infinitesimal arclength between \mathbf{A} and $\mathbf{A} + d\mathbf{A}$ is given by

$$ds := (\text{tr}(\mathbf{A}^{-1}d\mathbf{A})^2)^{1/2} = \|\mathbf{A}^{-1/2}d\mathbf{A}\mathbf{A}^{-1/2}\|_F, \tag{5}$$

where ds defines a metric on the Riemannian manifold [37]. $\|\cdot\|_F$ is the Frobenius norm of a matrix. The inner product and corresponding norm on the tangent space at the point \mathbf{A} can be defined as [38]

$$\langle \mathbf{P}_1, \mathbf{P}_2 \rangle_{\mathbf{A}} = \text{tr}(\mathbf{A}^{-1}\mathbf{P}_1\mathbf{A}^{-1}\mathbf{P}_2), \|\mathbf{P}_1\|_{\mathbf{A}} = \langle \mathbf{P}_1, \mathbf{P}_1 \rangle_{\mathbf{A}}^{1/2}. \tag{6}$$

For two points \mathbf{P}_1 and \mathbf{P}_2 on the Riemannian manifold, the affine invariant (Riemannian) distance is given by [39]

$$d_R^2(\mathbf{P}_1, \mathbf{P}_2) = \|\text{logm}(\mathbf{P}_1^{-1/2}\mathbf{P}_2\mathbf{P}_1^{-1/2})\|_F^2, \tag{7}$$

where logm is the logarithmic map on the Riemannian manifold of HPD matrices.

3.2. Total Bregman Divergence

The tBD has been proposed on the space of convex functions by Baba C. Vemuri, and has been used for DT-MRI analysis [40], shape retrieval [41,42], and object tracking [43]. For $x, y \in \mathbb{R}$, f is a differentiable, strictly convex function, the total Bregman divergence δ between x, y is defined as [40]

$$\delta_f(x, y) = \frac{f(x) - f(y) - f'(y)(x - y)}{\sqrt{1 + \|f'(y)\|^2}}, \tag{8}$$

where $f'(y)$ is the derivative of $f(x)$ at y .

As illustrated in Figure 2, the tBD between x and y is defined as the orthogonal distance between the value of a convex and differentiable function f at the first argument of y and its tangent at the second argument. It can be also noted from Figure 2 that when the coordinate system rotates at an angle, the tBD between x and y does not change. This implies that the tBD is invariant to linear transformation; in other words, it is statistically robust.

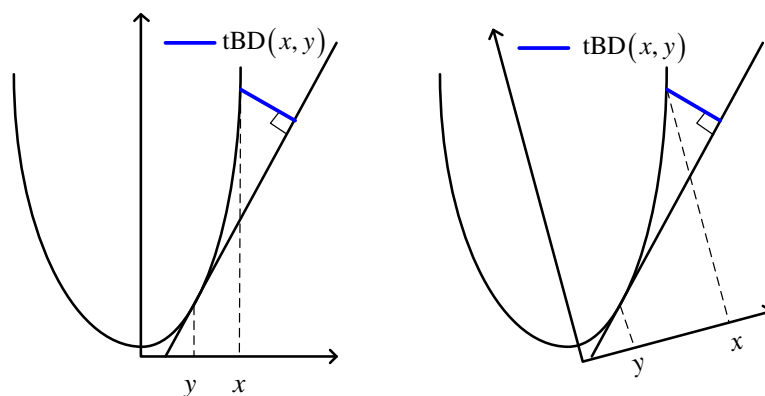


Figure 2. The geometric definition of total Bregman divergence.

In the following, we extend the definition of tBD to the Riemannian manifold of HPD matrices.

Definition 1. Let f be a strictly convex and differentiable function, for two HPD matrices \mathbf{X}, \mathbf{Y} , the tBD is defined as

$$\delta_f(\mathbf{X}, \mathbf{Y}) = \frac{\text{tr}(f(\mathbf{X})) - \text{tr}(f(\mathbf{Y})) - \text{tr}(f'(\mathbf{Y})(\mathbf{X} - \mathbf{Y}))}{\sqrt{1 + \|f'(\mathbf{Y})\|_F^2}}. \tag{9}$$

In the following, we give the definitions of the three new divergences when the function f has various forms.

Let $f(x) = x^2$, then $\text{tr}(f(\mathbf{X})) = \text{tr}(\mathbf{X}^2)$, $f'(\mathbf{Y}) = 2\mathbf{Y}$, and Equation (9) denotes the TSL

$$\delta_f(\mathbf{X}, \mathbf{Y}) = \frac{\text{tr}(\mathbf{X}^2 - \mathbf{Y}^2 - 2\mathbf{Y}(\mathbf{X} - \mathbf{Y}))}{\sqrt{1 + \|2\mathbf{Y}\|_F^2}}. \tag{10}$$

If $f(x) = x \log x - x$, then $\text{tr}(f(\mathbf{X})) = \text{tr}(\mathbf{X} \log \mathbf{X} - \mathbf{X})$, $f'(\mathbf{Y}) = \log \mathbf{Y}$, and Equation (9) yields the divergence, which is called the TVN,

$$\delta_f(\mathbf{X}, \mathbf{Y}) = \frac{\text{tr}(\mathbf{X} \log \mathbf{X} - \mathbf{X} \log \mathbf{Y} - \mathbf{X} + \mathbf{Y})}{\sqrt{1 + \|\log \mathbf{Y}\|_F^2}}. \tag{11}$$

When $f(x) = -\log x$, then $\text{tr}(f(\mathbf{X})) = -\log \det(\mathbf{X})$, $f'(\mathbf{Y}) = \mathbf{Y}^{-1}$, and we obtain the divergence referred to as the TLD or the total Stein loss,

$$\delta_f(\mathbf{X}, \mathbf{Y}) = \frac{\text{tr}(\mathbf{Y}^{-1}(\mathbf{X} - \mathbf{Y}) - \log \mathbf{X} + \log \mathbf{Y})}{\sqrt{1 + \|\mathbf{Y}^{-1}\|_F^2}}. \tag{12}$$

Based on the three divergences on the Riemannian manifold, the medians for a finite set of HPD matrices are derived in the next section.

3.3. Total Bregman Divergence Median for HPD Matrices

For a set of HPD matrices, the median related to a measure is defined as the minimizer of the sum of the distance to the given matrices.

Definition 2. Let f be a differentiable convex function; then, the median related to the tBD Equation (9), of m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m\}$ is defined as

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R}} \sum_{i=1}^m \delta_f(\mathbf{R}, \mathbf{R}_i). \tag{13}$$

It is worth pointing out that $\hat{\mathbf{R}}$ will be the arithmetic mean, if the Frobenius norm, instead of the tBD, is utilized.

Proposition 1. The median for a finite set of HPD matrices with respect to the tBD exists, and is unique.

Proof of Proposition 1. To find the median $\hat{\mathbf{R}}$, let $F(\mathbf{R}) = \sum_{i=1}^m \delta_f(\mathbf{R}, \mathbf{R}_i)$ denote the objection function

$$F(\mathbf{R}) = \sum_{i=1}^m \frac{\text{tr}(f(\mathbf{R})) - \text{tr}(f(\mathbf{R}_i)) - \text{tr}(f'(\mathbf{R}_i)(\mathbf{R} - \mathbf{R}_i))}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}}. \tag{14}$$

Setting the gradient $\nabla F(\mathbf{R})$ equal to zero,

$$\nabla F(\mathbf{R}) = \sum_{i=1}^m \frac{f'(\mathbf{R}) - f'(\mathbf{R}_i)}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} = 0. \tag{15}$$

Solving Equation (15) yields

$$f'(\hat{\mathbf{R}}) = \frac{\sum_{i=1}^m \frac{f'(\mathbf{R}_i)}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}}}{\sum_{i=1}^m \frac{1}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}}}. \tag{16}$$

The tBD $\delta_f(\mathbf{R}, \mathbf{R}_i)$ is a convex function, and $F(\mathbf{R})$, the sum of m convex functions, is also a convex function. Hence, the median $\hat{\mathbf{R}}$ indeed exists. Moreover, since f is a convex function, and f' is monotonic, the median $\hat{\mathbf{R}}$ is unique.

In the following, three concrete cases of the tBD are derived when f is given as different forms.

Proposition 2. The median related to the TSL (10), of m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m\}$ is given by

$$\hat{\mathbf{R}} = \sum_{i=1}^m \mu_i \mathbf{R}_i, \mu_i = \frac{w_i}{\sum_{i=1}^m w_i}, w_i = \frac{1}{\sqrt{1 + \|2\mathbf{R}_i\|_F^2}}. \tag{17}$$

Proposition 3. The median related to the TVN divergence (11) of m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m\}$ is given by

$$\hat{\mathbf{R}} = \exp\left(\sum_{i=1}^m \mu_i \log(\mathbf{R}_i)\right), \mu_i = \frac{w_i}{\sum_{i=1}^m w_i}, w_i = \frac{1}{\sqrt{1 + \|\log(\mathbf{R}_i)\|_F^2}}. \tag{18}$$

Proposition 4. The median related to the TLD divergence (12), of m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m\}$ is given by

$$\hat{\mathbf{R}} = \left(\sum_{i=1}^m \mu_i \mathbf{R}_i^{-1}\right)^{-1}, \mu_i = \frac{w_i}{\sum_{i=1}^m w_i}, w_i = \frac{1}{\sqrt{1 + \|\mathbf{R}_i^{-1}\|_F^2}}. \tag{19}$$

Proof of Proposition 2, 3, and 4. With the help of the formulas $\partial(\mathbf{R}^2)/\partial \mathbf{R} = 2\mathbf{R}$, $\partial(\mathbf{R} \log \mathbf{R} - \mathbf{R})/\partial \mathbf{R} = \log \mathbf{R}$, and $\partial(-\log \det(\mathbf{R}))/\partial \mathbf{R} = -\mathbf{R}^{-1}$, we then have

$$\nabla F(\mathbf{R}) = \sum_{i=1}^m \frac{2\mathbf{R} - 2\mathbf{R}_i}{\sqrt{1 + \|2\mathbf{R}_i\|_F^2}}, \tag{20}$$

$$\nabla F(\mathbf{R}) = \sum_{i=1}^m \frac{\log(\mathbf{R}) - \log(\mathbf{R}_i)}{\sqrt{1 + \|\log(\mathbf{R}_i)\|_F^2}}, \tag{21}$$

$$\nabla F(\mathbf{R}) = \sum_{i=1}^m \frac{\mathbf{R}^{-1} - \mathbf{R}_i^{-1}}{\sqrt{1 + \|\mathbf{R}_i^{-1}\|_F^2}}. \tag{22}$$

These three equations are for the three propositions above in the same order. Equating these gradients to zero and solving for \mathbf{R} yield the medians.

4. Robustness Analysis of Total Bregman Divergence Median

This section is devoted to analyzing the robustness of total Bregman divergence median and arithmetic mean via the influence function. Let $\tilde{\mathbf{R}}$ be the median, associated with total Bregman divergence, of m HPD matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$. $\tilde{\mathbf{R}}$ is the median by adding a set of n outliers $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$ with a weight ε ($\varepsilon \ll 1$) to $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$. Then, we can define $\tilde{\mathbf{R}} = \bar{\mathbf{R}} + \varepsilon H(\mathbf{Q})$, and $H(\mathbf{Q})$ denotes the influence function. In the following, four propositions are presented.

Proposition 5. *The influence function of arithmetic mean related to the Frobenius norm, of m HPD matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$ and n outliers $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$ is given by*

$$H(\mathbf{Q}) = \frac{1}{n} \sum_{i=1}^n \mathbf{Q}_i - \bar{\mathbf{R}}. \quad (23)$$

Proof of Proposition 5. Let $F(\mathbf{R})$ be the objection function

$$F(\mathbf{R}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \|\mathbf{R} - \mathbf{R}_i\|_F^2 + \varepsilon \frac{1}{n} \sum_{j=1}^n \|\mathbf{R} - \mathbf{Q}_j\|_F^2. \quad (24)$$

The derivative of objection function $F(\mathbf{R})$ is

$$\nabla F(\mathbf{R}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 2(\mathbf{R} - \mathbf{R}_i) + \varepsilon \frac{1}{n} \sum_{j=1}^n 2(\mathbf{R} - \mathbf{Q}_j). \quad (25)$$

Note that $\tilde{\mathbf{R}}$ is the median of m matrices and n outliers, and $\bar{\mathbf{R}}$ is the median of m matrices; then, we have

$$\begin{aligned} \tilde{\mathbf{R}} &= \arg \min_{\mathbf{R}} F(\mathbf{R}) \\ \Rightarrow \nabla F(\tilde{\mathbf{R}}) &= (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 2(\tilde{\mathbf{R}} - \mathbf{R}_i) + \varepsilon \frac{1}{n} \sum_{j=1}^n 2(\tilde{\mathbf{R}} - \mathbf{Q}_j) = 0 \end{aligned} \quad (26)$$

and

$$\begin{aligned} \bar{\mathbf{R}} &= \arg \min_{\mathbf{R}} G(\mathbf{R}), G(\mathbf{R}) = \frac{1}{m} \sum_{i=1}^m \|\mathbf{R} - \mathbf{R}_i\|_F^2 \\ \Rightarrow \nabla F(\bar{\mathbf{R}}) &= \frac{1}{m} \sum_{i=1}^m 2(\bar{\mathbf{R}} - \mathbf{R}_i) = 0. \end{aligned} \quad (27)$$

Substitute $\tilde{\mathbf{R}} = \bar{\mathbf{R}} + \varepsilon H(\mathbf{Q})$ into Equation (26), and we have

$$\begin{aligned} &2(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\bar{\mathbf{R}} + \varepsilon H(\mathbf{Q}) - \mathbf{R}_i) + 2\varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{R}} + \varepsilon H(\mathbf{Q}) - \mathbf{Q}_j) = 0 \\ \Rightarrow &(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\bar{\mathbf{R}} - \mathbf{R}_i) + (1 - \varepsilon)\varepsilon H(\mathbf{Q}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{R}} + \varepsilon H(\mathbf{Q}) - \mathbf{Q}_j) = 0 \\ \Rightarrow &\varepsilon H(\mathbf{Q}) - \varepsilon^2 H(\mathbf{Q}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{R}} - \mathbf{Q}_j) + \varepsilon^2 H(\mathbf{Q}) = 0 \\ \Rightarrow &H(\mathbf{Q}) = \frac{1}{n} \sum_{j=1}^n (\mathbf{Q}_j - \bar{\mathbf{R}}). \end{aligned} \quad (28)$$

Proposition 6. The influence function of TSL median related to TSL, of m HPD matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$ and n outliers $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$ is given by

$$H(\mathbf{Q}) = \frac{m}{n} \sum_{j=1}^n \frac{(\mathbf{Q}_j - \bar{\mathbf{R}})}{\sqrt{1 + \|2\mathbf{Q}_j\|_F^2}} \left(\sum_{i=1}^m \frac{1}{\sqrt{1 + \|2\mathbf{R}_i\|_F^2}} \right)^{-1}. \tag{29}$$

Proposition 7. The influence function of TVN median related to TVN, of m HPD matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$ and n outliers $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$ is given by

$$H(\mathbf{Q}) = \frac{m}{n} \sum_{j=1}^n \frac{(\log(\mathbf{Q}_j) - \log(\bar{\mathbf{R}})) \bar{\mathbf{R}}}{\sqrt{1 + \|\log(\mathbf{Q}_j)\|_F^2}} \left(\sum_{i=1}^m \frac{1}{\sqrt{1 + \|\log(\mathbf{R}_i)\|_F^2}} \right)^{-1}. \tag{30}$$

Proposition 8. The influence function of TLD median related to TLD, of m HPD matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_m\}$ and n outliers $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$ is given by

$$H(\mathbf{Q}) = \frac{m}{n} \sum_{j=1}^n \frac{(\bar{\mathbf{R}}^{-1} - \mathbf{Q}_j^{-1}) \bar{\mathbf{R}}^2}{\sqrt{1 + \|\mathbf{Q}_j^{-1}\|_F^2}} \left(\sum_{i=1}^m \frac{1}{\sqrt{1 + \|\mathbf{R}_i^{-1}\|_F^2}} \right)^{-1}. \tag{31}$$

Proof of Proposition 6, 7, and 8. According to Equation (9), let $F(\mathbf{R})$ be the objection function

$$F(\mathbf{R}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \delta_f(\mathbf{R}, \mathbf{R}_i) + \varepsilon \frac{1}{n} \sum_{j=1}^n \delta_f(\mathbf{R}, \mathbf{Q}_j). \tag{32}$$

Note that $\tilde{\mathbf{R}}$ is the median of m matrices and n outliers; then, we have,

$$\begin{aligned} \nabla F(\tilde{\mathbf{R}}) &= 0 \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \frac{f'(\tilde{\mathbf{R}}) - f'(\mathbf{R}_i)}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} + \varepsilon \frac{1}{n} \sum_{j=1}^n \frac{f'(\tilde{\mathbf{R}}) - f'(\mathbf{Q}_j)}{\sqrt{1 + \|f'(\mathbf{Q}_j)\|_F^2}} &= 0. \end{aligned} \tag{33}$$

Using Taylor expansions on $\tilde{\mathbf{R}} = \bar{\mathbf{R}} + \varepsilon H(\mathbf{Q})$ on the derivative function f' , we can obtain

$$f'(\tilde{\mathbf{R}}) = f'(\bar{\mathbf{R}}) + (\tilde{\mathbf{R}} - \bar{\mathbf{R}}) f''(\bar{\mathbf{R}}) = f'(\bar{\mathbf{R}}) + \varepsilon H(\mathbf{Q}) f''(\bar{\mathbf{R}}). \tag{34}$$

Substituting Equation (34) into Equation (33),

$$(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \frac{f'(\bar{\mathbf{R}}) + \varepsilon H(\mathbf{Q}) f''(\bar{\mathbf{R}}) - f'(\mathbf{R}_i)}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} + \varepsilon \frac{1}{n} \sum_{j=1}^n \frac{f'(\bar{\mathbf{R}}) + \varepsilon H(\mathbf{Q}) f''(\bar{\mathbf{R}}) - f'(\mathbf{Q}_j)}{\sqrt{1 + \|f'(\mathbf{Q}_j)\|_F^2}} = 0. \tag{35}$$

As $\bar{\mathbf{R}}$ is the median of m matrices, we have

$$\frac{1}{m} \sum_{i=1}^m \frac{f'(\bar{\mathbf{R}}) - f'(\mathbf{R}_i)}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} = 0. \tag{36}$$

Substituting Equation (36) into Equation (35), and ignoring the term containing ϵ^2 ,

$$\frac{1}{m} \sum_{i=1}^m \frac{\epsilon H(\mathbf{Q}) f''(\bar{\mathbf{R}})}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} + \epsilon \frac{1}{n} \sum_{j=1}^n \frac{f'(\bar{\mathbf{R}}) - f'(\mathbf{Q}_j)}{\sqrt{1 + \|f'(\mathbf{Q}_j)\|_F^2}} = 0. \tag{37}$$

Finally, the influence function of total Bregman divergence median can be derived

$$H(\mathbf{Q}) = -\frac{m}{n} \sum_{j=1}^n \frac{f'(\bar{\mathbf{R}}) - f'(\mathbf{Q}_j)}{\sqrt{1 + \|f'(\mathbf{Q}_j)\|_F^2}} \left(\sum_{i=1}^m \frac{f''(\bar{\mathbf{R}})}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} \right)^{-1}. \tag{38}$$

Let $f(\mathbf{R}) = \mathbf{R}^2$; then, $f'(\mathbf{R}) = 2\mathbf{R}$ and $f''(\mathbf{R}) = 2$. Substitute these into Equation (38), and we can obtain **Proposition 6**.

When $f(\mathbf{R}) = \mathbf{R} \log(\mathbf{R}) - \mathbf{R}$, then $f'(\mathbf{R}) = \log(\mathbf{R})$ and $f''(\mathbf{R}) = \mathbf{R}^{-1}$. Substitute these into Equation (38), and we can obtain **Proposition 7**.

If $f(\mathbf{R}) = -\log(\mathbf{R})$, then $f'(\mathbf{R}) = -\mathbf{R}^{-1}$ and $f''(\mathbf{R}) = \mathbf{R}^{-2}$. Substitute these into Equation (38), and we can obtain **Proposition 8**.

In addition, we show that the influence value of total Bregman divergence median has an upper bound when the number of outlier $n = 1$. Let

$$w(\mathbf{R}_i) = \sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}, G = \frac{1}{m} \sum_{i=1}^m \frac{f''(\bar{\mathbf{R}})}{w(\mathbf{R}_i)}. \tag{39}$$

Then, Equation (38) can be rewritten as

$$H(\mathbf{Q}) = \frac{f'(\mathbf{Q}) - f'(\bar{\mathbf{R}})}{\sqrt{1 + \|f'(\mathbf{Q})\|_F^2}} \left(\frac{1}{m} \sum_{i=1}^m \frac{f''(\bar{\mathbf{R}})}{\sqrt{1 + \|f'(\mathbf{R}_i)\|_F^2}} \right)^{-1} = \frac{f'(\mathbf{Q}) - f'(\bar{\mathbf{R}})}{w(\mathbf{Q})} G^{-1}. \tag{40}$$

Note that $w(\mathbf{Q}) = \sqrt{1 + \|f'(\mathbf{Q})\|_F^2} \geq \|f'(\mathbf{Q})\|_F$ and $w(\mathbf{Q}) \geq 1$; then, we have the following inequality:

$$\begin{aligned} \|H(\mathbf{Q})\|_F &= \frac{\|f'(\mathbf{Q}) - f'(\bar{\mathbf{R}})\|_F}{w(\mathbf{Q})} G^{-1} = \frac{\|f'(\mathbf{Q}) - f'(\bar{\mathbf{R}})\|_F}{\sqrt{1 + \|f'(\mathbf{Q})\|_F^2}} G^{-1} \\ &\leq \frac{\|f'(\mathbf{Q}) - f'(\bar{\mathbf{R}})\|_F}{\|f'(\mathbf{Q})\|_F} G^{-1} \leq \frac{\|f'(\mathbf{Q})\|_F + \|f'(\bar{\mathbf{R}})\|_F}{\|f'(\mathbf{Q})\|_F} G^{-1} \\ &= \left(1 + \frac{\|f'(\bar{\mathbf{R}})\|_F}{\|f'(\mathbf{Q})\|_F} \right) G^{-1} \leq (1 + \|f'(\bar{\mathbf{R}})\|_F) G^{-1} = c, \end{aligned} \tag{41}$$

where c is a constant. From the inequality (41), we can know that the influence function $H(\mathbf{Q})$ has an upper bound when \mathbf{Q} varies. It implies that our proposed tBD medians are robust.

5. Numerical Simulations

In this section, the performances of adaptive normalized matched filter with the proposed estimators and the normalized sample covariance matrix (NSCM) are evaluated by means of the

standard Monte Carlo techniques, as the analytical expression for the probability of detection is not available. Particularly, consider a classical target detection problem, which is formulated as follows:

$$\begin{cases} H_0 : \begin{cases} \mathbf{z} = \mathbf{r} \\ \mathbf{z}_k = \mathbf{r}_k, k = 1, \dots, K \end{cases} \\ H_1 : \begin{cases} \mathbf{z} = \beta \mathbf{s} + \mathbf{r} \\ \mathbf{z}_k = \mathbf{r}_k, k = 1, \dots, K \end{cases} \end{cases} \quad (42)$$

where \mathbf{z} and \mathbf{r} denote the received signal and the clutter data in the cell under test, respectively. β is an unknown complex quantity, accounting for the target radar cross section and the channel propagation effects. \mathbf{s} is the target steering vector,

$$\mathbf{s} = \frac{1}{\sqrt{N}} [1, \exp(j2\pi f_d), \dots, \exp(j2\pi(N-1)f_d)]^T, \quad (43)$$

where f_d is the normalized Doppler frequency. The terms \mathbf{r} and $\mathbf{r}_k, k = 1, \dots, K$ are the spherically invariant random vectors, and can be formulated as,

$$\mathbf{r} = \sqrt{\tau} \mathbf{g}, \quad \mathbf{r}_k = \sqrt{\tau_k} \mathbf{g}_k, \quad k=1, \dots, K, \quad (44)$$

where \mathbf{g} and \mathbf{g}_k are N -dimensional circularly symmetric zero-mean Gaussian vectors, and share the same covariance matrix Σ . τ and τ_k are positive and real random variables, which are also independent and identical distributed. In particular, the terms τ and τ_k are assumed to follow the inverse gamma distribution,

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), x \geq 0, \quad (45)$$

where α and β denote the shape and scale parameters, respectively. $\Gamma(\cdot)$ is the gamma function.

For the target detection problem (42), the adaptive normalized matched filtering (ANMF) is often used, which can be given by

$$\frac{\mathbf{s}^\dagger \hat{\Sigma}^{-1} \mathbf{r}}{(\mathbf{s}^\dagger \hat{\Sigma}^{-1} \mathbf{s})(\mathbf{r}^\dagger \hat{\Sigma}^{-1} \mathbf{r})} \underset{H_0}{\overset{H_1}{\geq}} \gamma, \quad (46)$$

where $\hat{\Sigma}$ is an estimator that is based on the secondary data $\mathbf{r}_k, k = 1 \dots, K$. γ denotes the threshold, which is derived by the Monte Carlo method in order to maintain the false alarm constant. In addition, the covariance matrix of \mathbf{g} and \mathbf{g}_k are given as the sum of two parts,

$$\Sigma = \Sigma_0 + \mathbf{I}, \quad (47)$$

where \mathbf{I} is accounting for the thermal noise. Σ_0 is related to the clutter, modeled as,

$$\Sigma_0(i, k) = \sigma_c^2 \rho^{|i-k|} e^{j2\pi f_{dc}(i-k)}, \quad i, k=1 \dots, N, \quad (48)$$

where ρ is the one-lag correlation coefficient. σ_c is the clutter-to-noise power ratio. f_{dc} is the clutter normalized Doppler frequency. In this simulation, we set $\rho = 0.9$, $f_{dc} = 0.1$, and $\sigma_c^2 = 25$ dB. The parameters $\alpha = 3$, and $\beta = 1$.

Performances of ANMFs with our proposed estimators are compared with the Frchet median estimator [18] and the NSCM estimator. The plots of P_d versus signal-to-clutter ratio (SCR) for different size of K , the number of secondary data, are shown in Figures 3 and 4. The P_d is estimated by the relative frequencies of 1000 simulations. The detection threshold is obtained using $100/P_{fa}$ Monte Carlo trials, for a fixed nominal P_{fa} . It can be noted from Figures 3 and 4 that the detection

performances of ANMFs with our proposed estimators and the Fréchet median estimator outperform that of ANMF with the NSCM estimator. The ANMF with the TLD estimator has the best performance, followed by the TVN estimator. The Fréchet median estimator has similar performance to the TVN estimator. However, the difference is not very clear. This implies that our proposed estimators are robust for the size of K , unlike the NSCM estimator.

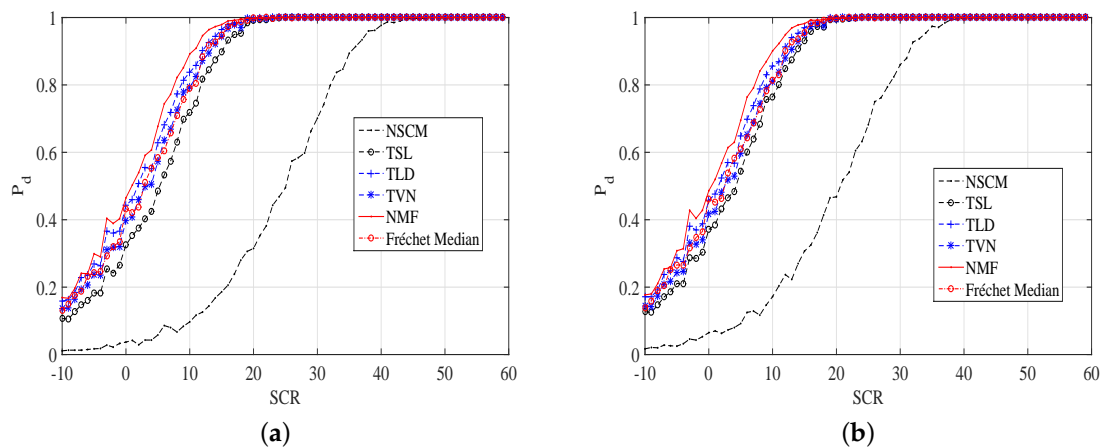


Figure 3. P_d versus SCR plots of ANMFs with the proposed estimators and the NSCM estimator, $N = 8, K = 10$. (a) $P_{fa} = 10^{-5}$; (b) $P_{fa} = 10^{-4}$.

In order to analyze the robustness of total Bregman divergence median and arithmetic mean estimators, we inject a outlier with normalized Doppler frequency $f_d = 0.1$ in 30 sample data, and compute the influence value according to (28), (29), (30), and (31). The results are shown in Figure 5. From Figure 5, we can know that the TLD median estimator has the smallest influence value, followed by the TSL median estimator. However, all of our proposed three estimators have smaller influence value than the arithmetic mean estimator, which is the NSCM estimator. These results imply that our proposed estimators are more robust than the NSCM estimator.

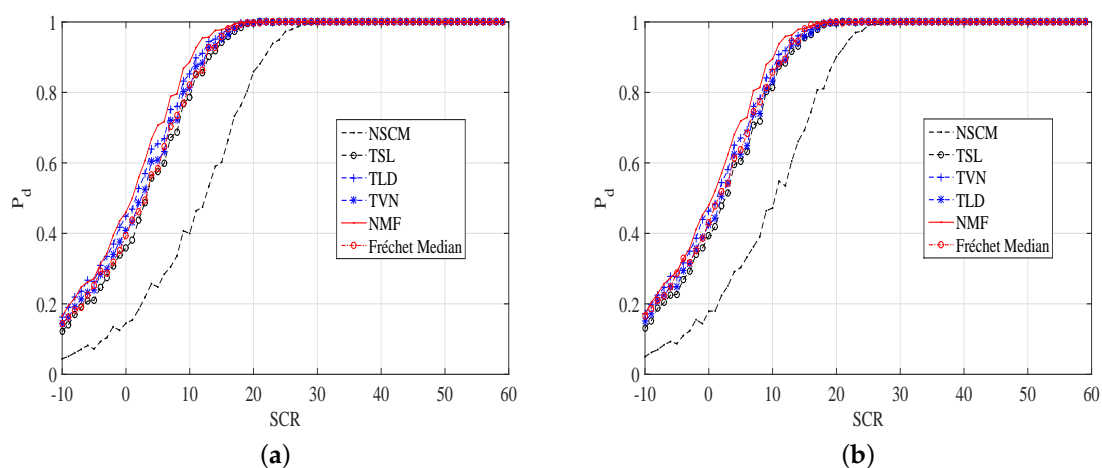


Figure 4. P_d versus SCR plots of ANMFs with the proposed estimators and the NSCM estimator, $N = 8, K = 16$. (a) $P_{fa} = 10^{-5}$; (b) $P_{fa} = 10^{-4}$.

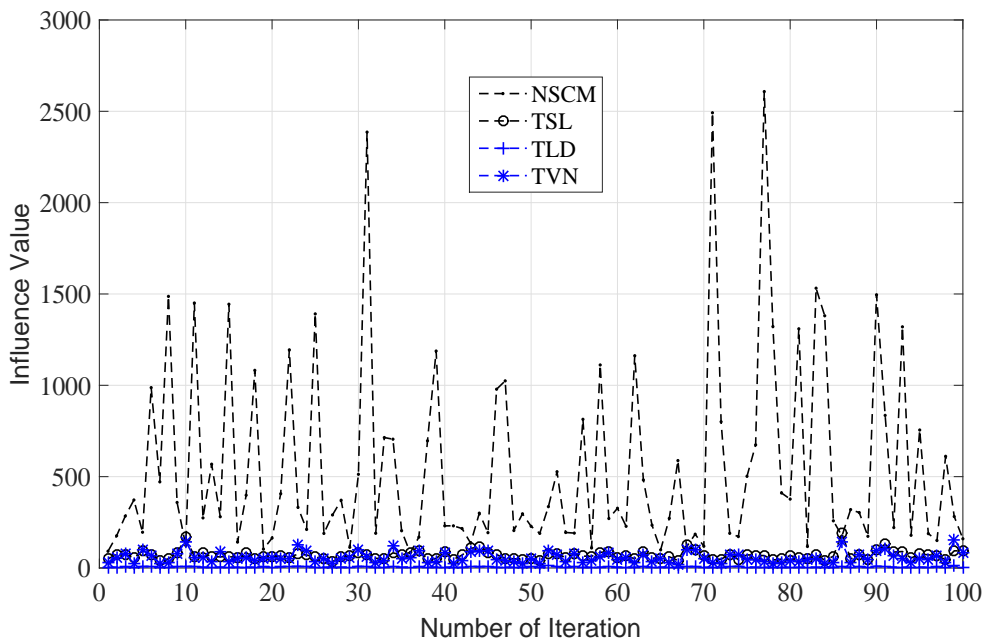


Figure 5. The influence value of arithmetic mean and total Bregman divergence median.

To enhance the persuasiveness of the results, we inject a outlier with the $f_{dc} = 0.15$, and $SCR = 20$ dB. Then, the plots of P_d versus SCR of ANMF with different estimators are given in Figures 6 and 7. It is clear from Figures 6 and 7 that the proposed TSL, TLD, and TVN estimators outperform the NSCM in a contaminated clutter. Moreover, the performance of the Fréchet median estimator is close to our proposed estimators.

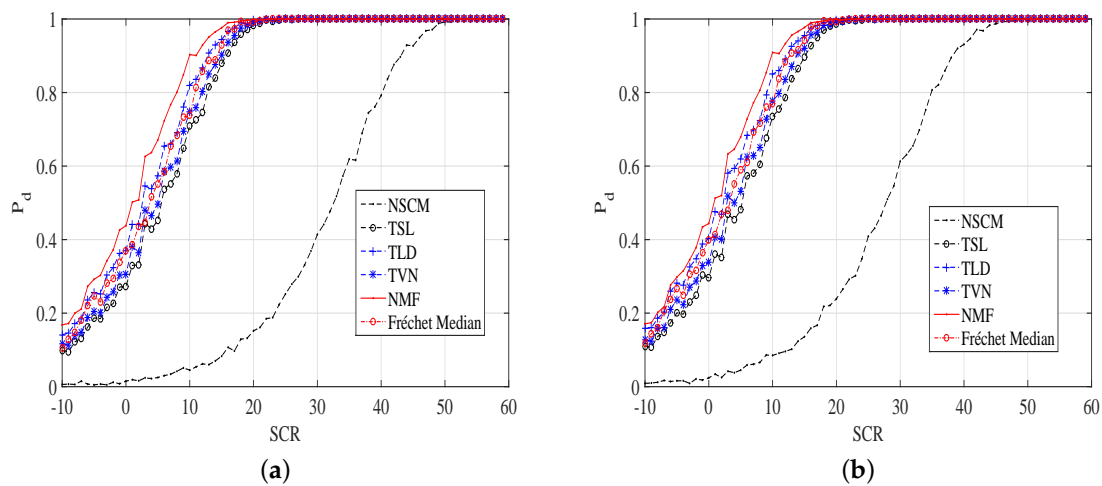


Figure 6. P_d versus SCR plots of ANMFs with the proposed estimators and the NSCM estimator, $N = 8, K = 10$. (a) $P_{fa} = 10^{-5}$; (b) $P_{fa} = 10^{-4}$.

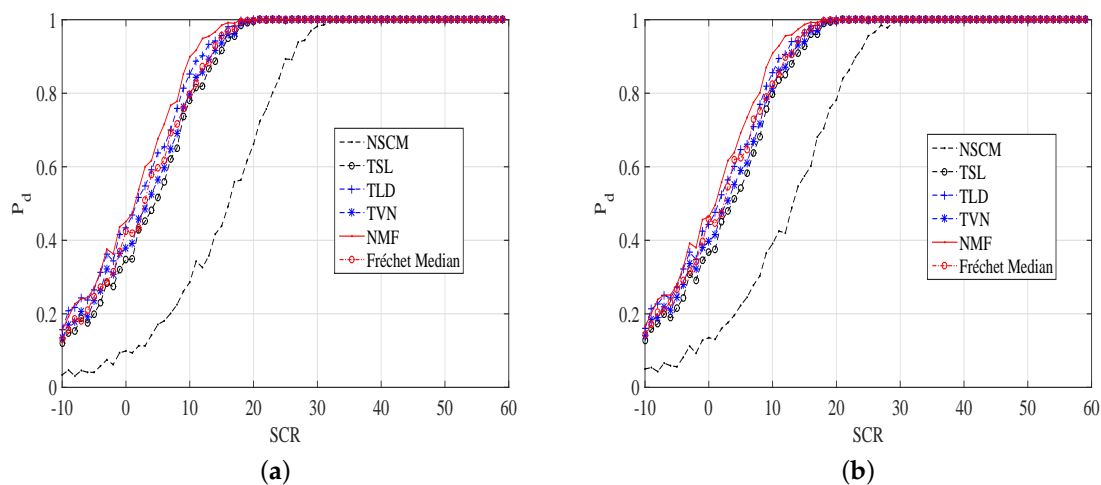


Figure 7. P_d versus SCR plots of ANMFs with the proposed estimators and the NSCM estimator, $N = 8, K = 16$. (a) $P_{fa} = 10^{-5}$; (b) $P_{fa} = 10^{-4}$.

6. Conclusions

In this work, we have proposed a covariance estimation method based on information geometry in a heterogeneous clutter. The problem of covariance estimation has been reformulated as the geometric median related to a measure. In particular, the three tBDs, including the TSL, the TLD, and the TVN, have been proposed on the Riemannian manifold. Then, we have derived the TSL, the TLD, and the TVN median estimators. At the analysis stage, the results of numerical experiments have highlighted that our proposed estimators outperform the NSCM estimator, and have similar performance to the Fréchet median estimator in heterogeneous clutter.

Acknowledgments: This work was supported by the National Natural Science Foundation of China under grant No. 61302149. The authors are grateful for the valuable comments made by the reviewers, which have assisted us with a better understanding of the underlying issues and therefore a significant improvement in the quality of the paper.

Author Contributions: Xiaoqiang Hua put forward the original ideas and performed the research. Yongqiang Cheng conceived and designed the simulations. Hongqiang Wang and Yuliang Qin reviewed the paper and provided useful comments.

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

References

1. Koichi, I.; Yu, I. On The Covariance Matrix in Array Signal Processing. *IEICE technical report. Antennas and propagation*. **2014**, *114*, 85–90.
2. Fuhrmann, D.R. Application of Toeplitz covariance estimation to adaptive beamforming and detection. *IEEE Trans. Signal Process.* **2002**, *39*, 2194–2198.
3. Samuli, V.; Hannu, O.; Visa, K. Multichannel signal processing using spatial rank covariance matrices. In Proceedings of the IEEE Eurasip Workshop on Nonlinear Signal and Image Processing, Antalya, Turkey, 20–23 June 1999; pp. 75–79.
4. Williams, D.B.; Johnson, D.H. Robust estimation of structured covariance matrices. *IEEE Trans. Signal Process.* **1993**, *41*, 2891–2906.
5. Carotenuto, V.; De Maio, A.; Orlando, D.; Pallotta, L. Adaptive Radar Detection Using Two Sets of Training Data. *IEEE Trans. Signal Process.* **2018**, *66*, 1791–1801.
6. Ciuonzo, D.; Orlando, D.; Pallotta, L. On the Maximal Invariant Statistic for Adaptive Radar Detection in Partially Homogeneous Disturbance With Persymmetric Covariance. *IEEE Signal Process. Lett.* **2016**, *23*, 1830–1834.

7. Maio, A.D.; Orlando, D. An Invariant Approach to Adaptive Radar Detection Under Covariance Persymmetry. *IEEE Trans. Signal Process.* **2015**, *63*, 1297–1309.
8. Diego, C.; Ingo, W. Joint Monostatic and Bistatic STAP for Improved SAR-GMTI Capabilities. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1834–1848.
9. Barton, T.; Smith, S. Structured Covariance Estimation for Space-Time Adaptive Processing. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, 21–24 April 1997; pp. 3493–3496.
10. Semeniaka, A.V.; Likhovitskiy, D.I.; Rachkov, D.S. Comparative analysis of Toeplitz covariance matrix estimation methods for space-time adaptive signal processing. In Proceedings of the IEEE CIE International Conference on Radar, Chengdu, China, 24–27 October 2011; pp. 696–699.
11. 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.
12. Marc, A.; Frederic, B.; Le, Y. Riemannian Medians and Means With Applications to Radar Signal Processing. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 595–604.
13. Le, Y. Medians of probability measures in Riemannian manifolds and applications to radar target detection. *Actas Urol. Esp.* **2012**, *31*, 651–659.
14. Lapuyade-Lahorgue, J.; Barbaresco, F. Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar. In Proceedings of the IEEE Radar Conference, Rome, Italy, 26–30 May 2008; pp. 1–6.
15. Hua, X.; Cheng, Y.; Li, Y.; Shi, Y.; Wang, H.; Qin, Y. Target Detection in Sea Clutter via Weighted Averaging Filter on the Riemannian Manifold. *Aerosp. Sci. Technol.* **2017**, *70*, 47–54.
16. Cheng, Y.; Hua, X.; Wang, H.; Qin, Y.; Li, X. The Geometry of Signal Detection with Applications to Radar Signal Processing. *Entropy* **2016**, *18*, 381.
17. Ruiz, M.; Barbaresco, F. Radar detection for non-stationary Doppler signal in one burst based on information geometry distance between paths. In Proceedings of the 2015 European Radar Conference (EuRAD), Paris, France, 9–11 September 2015; pp. 422–427.
18. Alexis, D.; Frederic, B. Robust Burg estimation of radar scatter matrix for autoregressive structured SIRV based on Frechet medians. *IET Radar Sonar Navig.* **2016**, *11*, 78–89.
19. Frederic, B. Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Frechet median. In *Matrix Information Geometry*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 199–255.
20. Jeuris, B.; Vandebril, R. Averaging block Toeplitz matrices with preservation of Toeplitz block structure. In Proceedings of the SIAM Conference on Applied Linear Algebra, Atlanta, GA, USA, 26–30 October 2015.
21. Balaji, B.; Barbaresco, F. Application of Riemannian mean of covariance matrices to space-time adaptive processing. In Proceedings of the 9th European IEEE Radar Conference, Amsterdam, The Netherlands, 31 October–2 November 2012; pp. 50–53.
22. Balaji, B. Riemannian mean and space-time adaptive processing using projection and inversion algorithms. *Proc. SPIE* **2013**, *8714*, 813–831.
23. Balaji, B.; Barbaresco, F.; Decurninge, A. Information geometry and estimation of Toeplitz covariance matrices. In Proceedings of the International IEEE Radar Conference, Lille, France, 13–17 October 2014; pp. 1–4.
24. Aubry, A.; Maio, A.D.; Pallotta, L.; Farina, A. Covariance matrix estimation via geometric barycenters and its application to radar training data selection. *IET Radar Sonar Navig.* **2013**, *7*, 600–614.
25. Aubry, A.; Maio, A.D.; Pallotta, L.; Farina, A. Median matrices and their application to radar training data selection. *IET Radar Sonar Navig.* **2013**, *8*, 265–274.
26. Aubry, A.; De Maio, A.; Pallotta, L. A geometric approach to covariance matrix estimation and its applications to radar problems. *IEEE Trans. Signal Process.* **2017**, *66*, 907–922.
27. Cui, G.; Li, N.; Pallotta, L.; Foglia, G.; Kong, L. Geometric Barycenters for Covariance Estimation in Compound-Gaussian Clutter. *IET Radar Sonar Navig.* **2017**, *11*, 404–409.
28. Hua, X.; Cheng, Y.; Wang, H.; Qin, Y.; Li, Y. Geometric means and medians with applications to target detection. *IET Signal Process.* **2017**, *11*, 711–720.
29. Hua, X.; Cheng, Y.; Wang, H.; Qin, Y.; Li, Y.; Zhang, W. Matrix CFAR detectors based on symmetrized Kullback Leibler and total Kullback Leibler divergences. *Digit. Signal Process.* **2017**, *69*, 106–116.

30. Malek, C.; Zeineb, C.; Maher, M.; Vemuri, B.C. Using the Bhattacharyya Mean for the Filtering and Clustering of Positive-Definite Matrices. In *Geometric Science of Information: First International Conference, GSI 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 551–558.
31. Malek, C.; Zeineb, C.; Maher, M.; Vemuri, B.C. Bhattacharyya median of symmetric positive-definite matrices and application to the denoising of diffusion-tensor fields. In Proceedings of the IEEE 10th International Symposium on Biomedical Imaging (ISBI), San Francisco, CA, USA, 7–11 April 2013; pp. 1227–1230.
32. Rhatia, R. *Positive Definite Matrices*; Princeton University Press: Princeton, NJ, USA, 2007.
33. Hiai, F.; Petz, D. Riemannian metrics on positive definite matrices related to means. *Linear Algebra Its Appl.* **2008**, *430*, 3105–3130.
34. Bridson, M.R.; Andre, H. *Metric Spaces of Nonpositive Curvature*; Springer: Berlin, Germany, 1999; pp. 52–65.
35. Suvrit, S. Positive definite matrices and the S divergence. *Proc. Am. Math. Soc.* **2011**, *144*, 1–25.
36. Hua, X.; Cheng, Y.; Wang, H.; Qin, Y. Robust covariance estimators based on information divergences and riemannian manifold. *Entropy* **2018**, *20*, 219, doi:10.3390/e20040219.
37. Maher, M. On the Averaging of Symmetric Positive Definite Tensors. *J. Elast.* **2006**, *82*, 273–296.
38. Lang, S. *Fundamentals of Differential Geometry. Graduate Texts in Mathematics*; Springer: New York, NY, USA, 1946; pp. 1757–1764.
39. Maher, M. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *Siam J. Matrix Anal. Appl.* **2008**, *26*, 735–747.
40. Vemuri, B.C.; Liu, M.; Amari, S.I.; Nielsen, F. Total bregman divergence and its applications to DTI analysis. *IEEE Trans. Med. Imaging* **2011**, *30*, 475–483.
41. Liu, M.; Vemuri, B.C.; Amari, S.I.; Nielsen, F. Total Bregman divergence and its applications to shape retrieval. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 13–18 June 2010; pp. 3463–3468.
42. Liu, M.; Vemuri, B.C.; Amari, S.I.; Nielsen, F. Shape retrieval using hierarchical total Bregman soft clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2407–2419.
43. Teran, A.R.M.Y.; Gouiffes, M.; Lacassagne, L. Total Bregman Divergence for Multiple Object Tracking. In Proceedings of the IEEE International Conference on Image Processing, Melbourne, Australia, 15–18 September 2013.



© 2018 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 (<http://creativecommons.org/licenses/by/4.0/>).