Information Geometry for Radar Target Detection with Total Jensen–Bregman Divergence

Xiaoqiang Hua [1,*] **, Haiyan Fan** [2] **, Yongqiang Cheng** [1,*] **, Hongqiang Wang** [1] **and Yuliang Qin** [1]

[1] School of Electronic Science, National University of Defence Technology, Changsha 410073, China; oliverwhq@tom.com (H.W.); yuliang.qin@gmail.com (Y.Q.)
[2] Space Engineering University, Beijing 101400, China; hyfan@yeah.net
[*] Correspondence: hxq712@yeah.net (X.H.); cyq101600@126.com (Y.C.)

Abstract: This paper proposes a radar target detection algorithm based on information geometry. In particular, the correlation of sample data is modeled as a Hermitian positive-definite (HPD) matrix. Moreover, a class of total Jensen–Bregman divergences, including the total Jensen square loss, the total Jensen log-determinant divergence, and the total Jensen von Neumann divergence, are proposed to be used as the distance-like function on the space of HPD matrices. On basis of these divergences, definitions of their corresponding median matrices are given. Finally, a decision rule of target detection is made by comparing the total Jensen-Bregman divergence between the median of reference cells and the matrix of cell under test with a given threshold. The performance analysis on both simulated and real radar data confirm the superiority of the proposed detection method over its conventional counterparts and existing ones.

1. Introduction

It is important to improve the performance of target detection in a clutter for a radar system. Due to the limitation of sample data, classical constant false alarm rate (CFAR) detector based on fast Fourier transform (FFT) [1] experiences performance degradation. This is because, the energy of the Doppler filter banks spreads and the Doppler resolution is poor, when the correlation of the sample data is captured by the FFT. To address these problems, Barbaresco investigated a geometric detection algorithm on a Hermitian positive-definite (HPD) matrix manifold, named as Riemannian distance-based geometric detector [2]. An HPD matrix $\mathbf{R}_i$ is utilized to model the information of each cell, as the sample data $\mathbf{z}$ is assumed to be following a complex circular Gaussian distribution with zero mean. Then, the detection statistic in each cell is the Riemannian distance of the matrix $\mathbf{R}_D$ of the cell under test and the mean $\bar{\mathbf{R}}$ of its surrounding cells. Finally, the detection statistic is compared with a threshold γ to determine the decision about "target absence" or "target presence". It can be referred to Figure 1.

In this geometric detector, the sample data of each cell in one coherent processing interval (CPI) is represented by an HPD matrix. According to this parameterization, the detection method is performed on the HPD matrix manifold, and the geometric characteristic is considered. Then, the metric is derived [2,3], and the existence and uniqueness of geometric mean had been proven in [4]. The geometric detector has been used to monitor the turbulence of a plane [5–7], target detection in coastal X-band and HF surface wave radars [2,3]. Experiment results show that this geometric detector outperforms the classical FFT-CFAR detector [2].
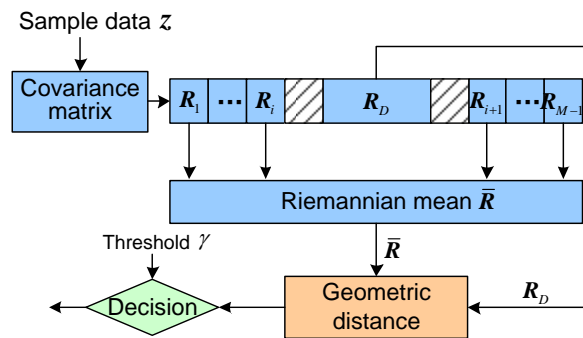
Figure 1. Riemannian mean-based geometric detector.

The classical FFT-CFAR detector has a similar scheme to the geometric detector. The main differences between them lie in the following three aspects: (1) the correlation of sample data is captured by an HPD matrix, and not the FFT coefficient; (2) the Riemannian distance is utilized in the detector, and not the Euclidean distance; and (3) the average value of HPD matrices is the geometric mean, rather than the arithmetic average. These differences illuminate that different geometry space is considered in these two detectors.

Many divergences have been used as a distance-like function on the space of HPD matrices. For instance, the Hellinger distance has been utilized to design a geometric target detector in a clutter; the Bhattacharyya divergence is used for filtering in medical imaging [8,9]; and the Log-Euclidean distance has been employed to measure the dissimilarity of two HPD matrices [10]. In these contexts, geometric measures have achieved performance improvement. In [11], the total Bregman divergence is defined on the convex function space, and many perfect properties have been analyzed in theory. The t-center of HPD matrices have been derived according to the definition of total Bregman divergence. It has been applied to diffusion tensor magnetic resonance image analysis [11], object tracking [12], and shape retrieval [13,14]. More recently, the definition of total Jensen–Bregman divergence has been presented on the space of convex functions in [15,16]. This class of divergence has not been applied in practice. In [17], a Kullback–Leibler divergence-based geometric target detection algorithm is proposed, and experimental results give a proof of performance improvement when compared to the classical FFT-CFAR detector. In [18,19], we have analyzed the anisotropy of the metric used in our geometric detection method. In particular, we have found that the detection performance of geometric detector is related to the measure used. In [20], the definition of total Bregman divergence has been extended on the Riemannian manifold, and we have validated the superiority of performance of the total Bregman divergence-based geometric detector when compared with the Riemannian distance-based geometric detector. These results inspire us to study new measures.

In this work, the definition of total Jensen–Bregman divergence is presented on the HPD matrix manifold. Specially, the total Jensen square loss, the total Jensen von Neumann divergence, and the total Jensen log-determinant divergence are proposed and defined. Based on three divergences, medians of HPD matrices are presented. As a result, a total Jensen–Bregman divergence-based radar target detection is developed. The main contribution of this paper is that a new class of total Jensen–Bregman divergence is proposed for designing the target detector on the Riemannian manifold.

The rest of this paper is organized as follows. In Section 2, we give a description about the signal model and signal manifold. In Section 3, the basic mathematical knowledge of matrix information geometry is presented. In particular, the definition of total Jensen–Bregman divergence is given. The median matrix associated with total Jensen–Bregman divergence is derived in Section 4. Then, in Section 5, we evaluate performances of total Jensen–Bregman divergences-based target detection algorithms as well as the Riemannian distance-based detector and the FFT-CFAR detector by simulated and real sea clutter data. Finally, a conclusion is provided in Section 6.

Here are some notations for the descriptions of this article. A scalar x is denoted using the math italic. 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 $\mathbf{A}^H$. $tr(\mathbf{A})$ is the trace of matrix $\mathbf{A}$. $det(\mathbf{A})$ is the determinant of matrix $\mathbf{A}$. $\mathbf{I}$ is the identity matrix. All n-dimensional vectors are noted by $\mathbb{C}(n)$. $\mathbb{H}(n)$ is the set of all $n \times n$ Hermitian matrices. $\|\mathbf{A}\|_F$ denotes the F-norm of $\mathbf{A}$. All $n \times n$ positive-definite matrices in $\mathbb{H}(n)$ consist of the space $\mathbb{P}(n)$. Finally, $\mathbb{E}(\cdot)$ denotes the statistical expectation.

2. Signal Model and Signal Manifold

The radar usually receives the echo returned from a moving target, and the phase information is included in the echo data. The correlation of the sample data $\mathbf{z} = \{z_1, z_2, \cdots, z_n\}$ can be used to capture the Doppler information of the target. Here, the data is assumed to follow a complex multivariate Gaussian distribution with zero mean, $z \sim \text{CN}(\mathbf{0}, \mathbf{R})$ [2],

$$p(z \,|\mathbf{R}) = \frac{1}{\pi^n \det(\mathbf{R})} \exp\left\{-z^H \mathbf{R}^{-1} z\right\} \tag{1}$$

where the matrix $\mathbf{R}$ is given as [2],

$$\mathbf{R} = \left[zz^H\right] = \begin{bmatrix} r_0 & \bar{r}_1 & \cdots & \bar{r}_{n-1} \\ r_1 & r_0 & \cdots & \bar{r}_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{n-1} & \cdots & r_1 & r_0 \end{bmatrix} \tag{2}$$

$$r_k = [z_i \bar{z}_{i+k}], 0 \leq k \leq n-1, 0 \leq i \leq n-1$$

where $\bar{r}_i$ denotes the conjugate of r_i, and r_k is the coefficient of the sample data. Due to the limitation of number of the sample data, according to the ergodicity, the statistical expectation $\mathbb{E}[z_i \bar{z}_{i+k}]$ can be computed by a finite time serials,

$$\hat{r}_k = \frac{1}{n-k} \sum_{j=0}^{n-1-k} z_j \bar{z}_{j+k}, 0 \leq k \leq n-1 \tag{3}$$

The sample data $\mathbf{z} = \{z_1, z_2, \cdots, z_n\}$ of each cell in one CPI is represented by a Toeplitz matrix $\mathbf{R}$. In particular, the sample data $\mathbf{z}$ lies in the linear space, while the matrix $\mathbf{R}$ is considered on the matrix manifold. Thus, the transformation can be given as,

$$\Psi : \mathbb{C}(n) \to \mathbb{P}(n), \mathbf{z} \to \mathbf{R} \in (n) \tag{4}$$

$\mathbb{P}(n)$ constitutes a differentiable non-linear manifold [21] with non-positive curvature [22,23]. Then, the target detection is performed on the Riemannian manifold. The manifold $\mathbb{P}(n)$ is a symmetric space [24], and more details are referred to [25].

3. Matrix Information Geometry

In this Section, we will introduce some basic mathematical knowledge of matrix information geometry. For instance, the geometry of Riemannian manifold of HPD matrices, and distance measures, are presented. Moreover, the definition of total Jensen–Bregman divergence is presented.

3.1. The Geometry of Riemannian Manifold

For a $n \times n$ matrix $\mathbf{A}$, $\mathbf{A}$ is a Hermitian matrix if $\mathbf{A}^H = \mathbf{A}$. $\mathbb{H}(n)$ denotes the linear space constituted of $n \times n$ Hermitian matrices. $\mathbf{A}$ is noted as a positive semi-definite if $\mathbf{x}^H \mathbf{A} \mathbf{x} \geq 0, \forall \mathbf{x} \in \mathbb{C}(n)$. In addition,

if **A** is invertible, then, we can say **A** is a positive-definite Hermitian matrix. All these HPD matrices are consisting of a convex symmetric cone [26],

$$\mathbb{P}(n) = \{\mathbf{A} \in \mathbb{H}(n), \mathbf{A} > 0\} \tag{5}$$

$\mathbb{P}(n)$ forms a non-linear manifold, and an HPD matrix is a point on the manifold. For any point **A**, its tangent space $\mathbb{T}_\mathbf{A}$ is the space $\mathbb{H}(n)$. In particular, the infinitesimal arclength at point **A** is given as [26],

$$ds := \left(tr\left(\mathbf{A}^{-1}d\mathbf{A}\right)^2 \right)^{1/2} = \left\| \mathbf{A}^{-1/2}\,d\mathbf{A}\mathbf{A}^{-1/2} \right\|_F \tag{6}$$

The metric is defined by the infinitesimal on $\mathbb{P}(n)$. For any a point **A**, the inner product can be defined as,

$$\langle \mathbf{R}_1, \mathbf{R}_2 \rangle_\mathbf{A} = tr\left(\mathbf{A}^{-1}\mathbf{R}_1\mathbf{A}^{-1}\mathbf{R}_2\right), \|\mathbf{R}_1\|_\mathbf{A} = \langle \mathbf{R}_1, \mathbf{R}_1 \rangle_\mathbf{A}^{1/2} \tag{7}$$

The distance between two points $\mathbf{R}_1$, $\mathbf{R}_2$ on the manifold is given by [26],

$$d_R^2\left(\mathbf{R}_1, \mathbf{R}_2\right) = \left\| \text{logm}\left(\mathbf{R}_1^{-1/2}\,\mathbf{R}_2\mathbf{R}_1^{-1/2}\right) \right\|_F^2 = \sum_{k=1}^n \log^2\left(\lambda_k\right) \tag{8}$$

where λ_k is the kth eigenvalue of $\mathbf{R}_1^{-1/2}\,\mathbf{R}_2\mathbf{R}_1^{-1/2}$, and $logm(\cdot)$ is a logarithmic map.

3.2. Total Jensen–Bregman Divergence on the Riemannian Manifold

The definition of total Jensen–Bregman divergence is given by Frank Nielsen in [16] on the space of convex functions, and many properties have been analyzed from the perspective of algebra and geometry. Given a strictly convex and differentiable function f, for $x, y \in \mathbb{R}$, the total Jensen–Bregman divergence δ between x, y is defined as,

$$\begin{aligned} \delta_\alpha\left(x, y\right) &= \tfrac{1}{\alpha(1-\alpha)}\rho\left(x, y\right)\left(\left(f\left(x\right)f\left(y\right)\right)_\alpha - f\left(\left(xy\right)_\alpha\right)\right) \\ \rho\left(x, y\right) &= 1/\sqrt{1 + \tfrac{\langle f(x)-f(y), f(x)-f(y) \rangle}{\langle x-y, x-y \rangle}}, \alpha \in [0, 1] \end{aligned} \tag{9}$$

where $f\left(\left(xy\right)_\alpha\right) = \alpha f\left(x\right) + \left(1 - \alpha\right)f\left(y\right)$ and $\left(xy\right)_\alpha = \alpha x + \left(1 - \alpha\right)y$. α is a skew scale parameter. In the following, the definition of total Jensen–Bregman divergence is extended from the space of convex functions to Riemannian manifold of HPD matrices. Furthermore, according to different forms of the function f, we define three new divergences, namely the total Jensen square loss, the total Jensen von Neumann divergence, and the total Jensen Log-determinant divergence.

Definition 1. *Let a strictly convex function f be differentiable, the total Jensen–Bregman divergence between two HPD matrices* **X**, **Y** *is defined as,*

$$\begin{aligned} \delta_\alpha\left(\mathbf{X}, \mathbf{Y}\right) &= \tfrac{1}{\alpha(1-\alpha)}\rho\left(\mathbf{X}, \mathbf{Y}\right)tr\left(\left(f\left(\mathbf{X}\right)f\left(\mathbf{Y}\right)\right)_\alpha - f\left(\left(\mathbf{XY}\right)_\alpha\right)\right) \\ \rho\left(\mathbf{X}, \mathbf{Y}\right) &= 1/\sqrt{1 + \tfrac{\|f(\mathbf{X})-f(\mathbf{Y})\|_F^2}{\|\mathbf{X}-\mathbf{Y}\|_F^2}}, \alpha \in [0, 1] \end{aligned} \tag{10}$$

In the following, we define three divergences when the function f has various forms.

If $f(x) = x^2$, then $tr(f(\mathbf{X})) = tr(\mathbf{X}^2)$, and (10) can be rewritten as,

$$\delta_\alpha(\mathbf{X}, \mathbf{Y})$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y}) \, tr\left(\left(\mathbf{X}^2 \mathbf{Y}^2\right)_\alpha - ((\mathbf{XY})_\alpha)^2 \right)$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y}) \, tr\left(\alpha \mathbf{X}^2 + (1-\alpha)\mathbf{Y}^2 - (\alpha \mathbf{X} + (1-\alpha)\mathbf{Y})^2 \right)$$
$$\rho(\mathbf{X}, \mathbf{Y}) = 1 \Big/ \sqrt{1 + \frac{\|\mathbf{X}^2 - \mathbf{Y}^2\|_F^2}{\|\mathbf{X} - \mathbf{Y}\|_F^2}}, \alpha \in [0, 1] \tag{11}$$

Equation (11) denotes the total Jensen square loss.

Let $f(x) = x\log x - x$, then $tr(f(\mathbf{X})) = tr(\mathbf{X}\log(\mathbf{X}) - \mathbf{X})$, and (10) yields the divergence,

$$\delta_\alpha(\mathbf{X}, \mathbf{Y}) = \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y})$$
$$\times tr\left\{ ((\mathbf{X}\log(\mathbf{X}) - \mathbf{X})(\mathbf{Y}\log(\mathbf{Y}) - \mathbf{Y}))_\alpha - (\mathbf{Z}\log(\mathbf{Z}) - \mathbf{Z}) \right\}$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y})$$
$$\times tr\left(\alpha(\mathbf{X}\log(\mathbf{X}) - \mathbf{X}) + (1-\alpha)(\mathbf{Y}\log(\mathbf{Y}) - \mathbf{Y}) - (\mathbf{Z}\log(\mathbf{Z}) - \mathbf{Z}) \right) \tag{12}$$
$$\rho(\mathbf{X}, \mathbf{Y}) = 1 \Big/ \sqrt{1 + \frac{\|(\mathbf{X}\log(\mathbf{X}) - \mathbf{X}) - (\mathbf{Y}\log(\mathbf{Y}) - \mathbf{Y})\|_F^2}{\|\mathbf{X} - \mathbf{Y}\|_F^2}}$$
$$\mathbf{Z} = \alpha\mathbf{X} + (1-\alpha)\mathbf{Y}, \alpha \in [0, 1]$$

We call Equation (12) the total Jensen von Neumann divergence.

If $f(x) = -\log x$, then $tr(f(\mathbf{X})) = -\log(\det(\mathbf{X}))$, and (10) can be rewritten as,

$$\delta_\alpha(\mathbf{X}, \mathbf{Y})$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y}) \, tr\left(((-\log(\mathbf{X}))(-\log(\mathbf{Y})))_\alpha + \log((\mathbf{XY})_\alpha) \right)$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y})$$
$$\times tr\left(\log(\alpha \mathbf{X} + (1-\alpha)\mathbf{Y}) - \alpha\log(\mathbf{X}) - (1-\alpha)\log(\mathbf{Y}) \right) \tag{13}$$
$$= \frac{1}{\alpha(1-\alpha)} \rho(\mathbf{X}, \mathbf{Y})$$
$$\times \left(\log\det(\alpha \mathbf{X} + (1-\alpha)\mathbf{Y}) - \alpha\log\det(\mathbf{X}) - (1-\alpha)\log\det(\mathbf{Y}) \right)$$
$$\rho(\mathbf{X}, \mathbf{Y}) = 1 \Big/ \sqrt{1 + \frac{\|\log(\mathbf{X}) - \log(\mathbf{Y})\|_F^2}{\|\mathbf{X} - \mathbf{Y}\|_F^2}}, \alpha \in [0, 1]$$

Equation (13) is named as the total Jensen Log-determinant divergence, or total Jensen Stein loss.

3.3. Anisotropy Analysis

In our previous work [18,19], anisotropies for six geometric measures on the Riemannian manifold are explored. An interesting phenomenon has been found that the detection performance may have some relationships with the anisotropy of a measure. However, the relationship needs further research. As a fact, the anisotropy implies the dissimilarity between metric tensors along different directions on the Riemannian manifold. It is well known that a linear space spanned by basis vectors is isotropic. However, the Riemannian manifold is a non-linear space, where different locations may have different

anisotropies. In order to illustrate the difference of anisotropy of total Jensen–Bregman divergence, we give some visual results with 3-dimensional HPD matrix. Anisotropy isosurfaces, centered at the identity matrix with radius $r = 1$, 2 and 5 respectively, are shown in Figure 2.
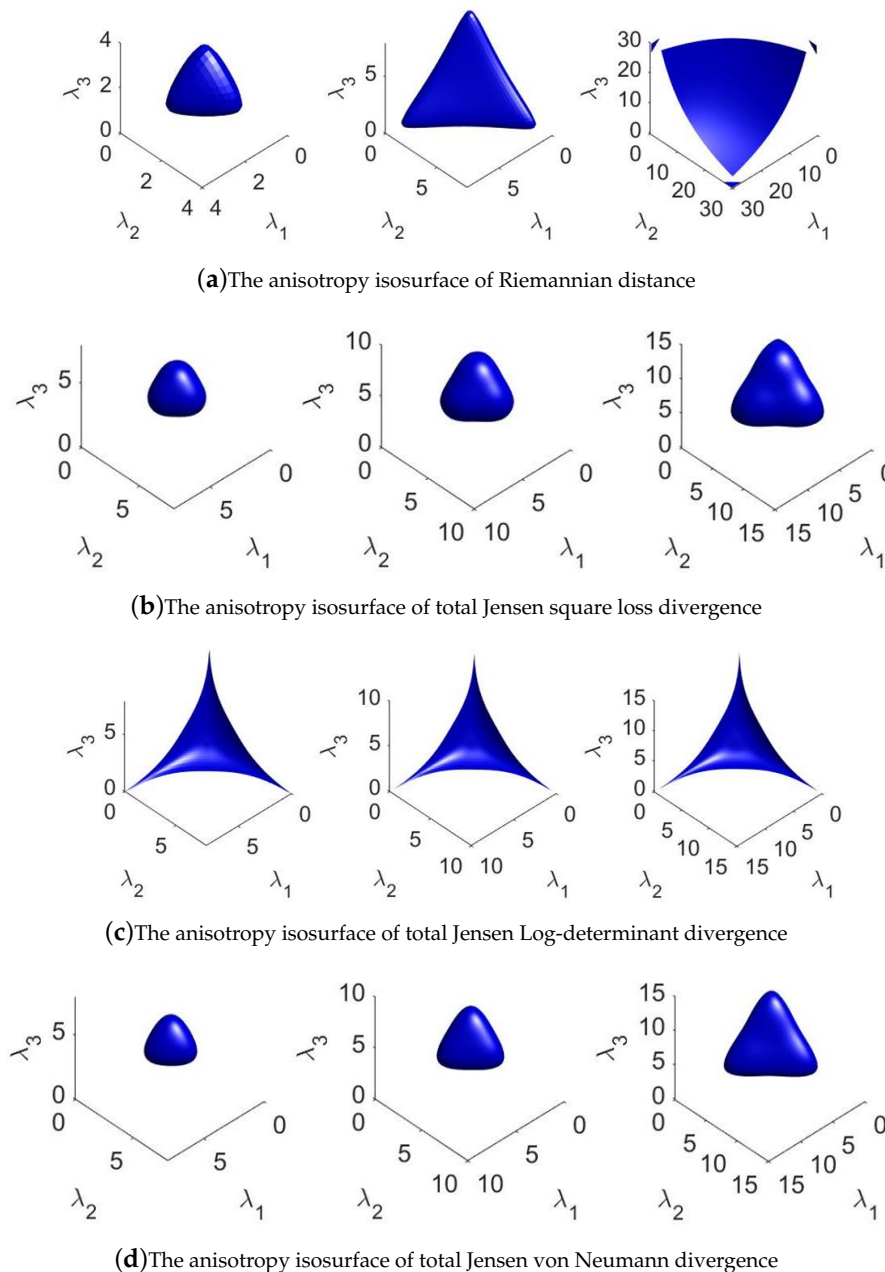

(**a**)The anisotropy isosurface of Riemannian distance

(**b**)The anisotropy isosurface of total Jensen square loss divergence

(**c**)The anisotropy isosurface of total Jensen Log-determinant divergence

(**d**)The anisotropy isosurface of total Jensen von Neumann divergence

Figure 2. The isosurfaces of total Jensen-Bregman divergence.

It is clear from Figure 2 that isosurfaces of these measures are different from each other. Different measures have different anisotropies at the point I on the manifold $\mathbb{P}(3)$. This also means that the curvatures along the different directions are not identical. In other word, the metric matrix on the Riemannian manifold is not an identity matrix. The curvature can reflect the structure of manifold.

4. Total Jensen–Bregman Divergence Median on the Riemannian Manifold of HPD Matrices

Many literatures have been reported about the median matrix of HPD matrices. For instance, in [27], the Riemannian median is defined as the minimizer of the sum of geodesic distances to given

matrices. In [9], Moakher defines the Log-Euclidean and Bhattacharyya medians for filtering the medical image. In the following, we give the definition of total Jensen–Bregman divergence median.

Definition 2. *Let a convex function f be differentiable, for m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_m\}$, the median associated with total Jensen-Bregman divergence (10) is defined as,*

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R}} \sum_{i=1}^{m} \delta_\alpha (\mathbf{R}, \mathbf{R}_i) \tag{14}$$

Note that the right-sided centroids $\hat{\mathbf{R}}'$ are obtained by minimizing the equivalent left-sided centroids for $\alpha' = 1 - \alpha$, and $\hat{\mathbf{R}}'_\alpha = \hat{\mathbf{R}}_{1-\alpha}$. Therefore, we consider the left-sided centroids in the remainder.

In order to minimize Equation (14), we proceed iteratively in two stages:

Stage 1: we consider $\hat{\mathbf{R}}^{(t)}$ given, here, $\hat{\mathbf{R}}^{(0)} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{R}_i$. This allows us to consider the following simpler minimization problem:

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R} \in (n)} \sum_{i=1}^{m} \left\{ \rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right) \frac{1}{\alpha(1-\alpha)} \times tr\left(\left(f(\mathbf{R})f(\mathbf{R}_i)\right)_\alpha - f\left((\mathbf{R}\mathbf{R}_i)_\alpha\right)\right) \right\} \tag{15}$$

Let

$$w_i^{(t)} = \frac{\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right)}{\sum_j \rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_j\right)} \tag{16}$$

be the updated renormalized weights at stage t.

Stage 2: we minimize

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R} \in (n)} \sum_{i=1}^{m} w_i^{(t)} \frac{1}{\alpha(1-\alpha)} tr\left(\left(f(\mathbf{R})f(\mathbf{R}_i)\right)_\alpha - f\left((\mathbf{R}\mathbf{R}_i)_\alpha\right)\right) \tag{17}$$

This is a convex-concave minimization procedure [28] (CCCP) that can be solved iteratively until it reaches convergence [29]. The above two iterations can be formulated as following:

$$\hat{\mathbf{R}}^{(t+1)} = (\nabla f)^{-1} \left(\sum_{i=1}^{m} w_i^{(t)} \nabla f\left(\alpha\hat{\mathbf{R}}^{(t)} + (1-\alpha)\mathbf{R}_i\right) \right) \tag{18}$$

It is an open and challenging problem to prove that the total Jensen–Bregman medians are unique whatever the chosen multivariate f, see [29].

In the following, we will give total Jensen–Bregman medians according to different forms of f.

Proposition 1. *For m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_m\}$, the median associated with total Jensen square loss (11) is given by,*

$$\hat{\mathbf{R}}^{(t+1)} = \sum_{i=1}^{m} w_i^{(t)} \left(\alpha\hat{\mathbf{R}}^{(t)} + (1-\alpha)\mathbf{R}_i\right), w_i^{(t)} = \frac{\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right)}{\sum_j \rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_j\right)}$$

$$\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right) = 1 / \sqrt{1 + \frac{\left\|\left(\hat{\mathbf{R}}^{(t)}\right)^2 - \mathbf{R}_i^2\right\|_F^2}{\left\|\hat{\mathbf{R}}^{(t)} - \mathbf{R}_i\right\|_F^2}} \tag{19}$$

Proof of Proposition 1. According to Equation (11), $f(\mathbf{R}) = \mathbf{R}^2$, then we have,

$$\nabla f(\mathbf{R}) = 2\mathbf{R}, \nabla f^{-1}(\mathbf{R}) = \mathbf{R}/2 \tag{20}$$

□

Substitution Equation (20) into Equation (18), and we can obtain the total Jensen square loss median as Equation (19).

Proposition 2. *For m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_m\}$, the median associated with total Jensen von Neumann divergence (12) is given by,*

$$\hat{\mathbf{R}}^{(t+1)} = \exp\left(\sum_{i=1}^{m} w_i^{(t)} \log\left(\alpha\hat{\mathbf{R}}^{(t)} + (1-\alpha)\mathbf{R}_i\right)\right), w_i^{(t)} = \frac{\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right)}{\sum_j \rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_j\right)},$$

$$\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right) = 1/\sqrt{1 + \frac{\left\|\left(\hat{\mathbf{R}}^{(t)}\log\left(\hat{\mathbf{R}}^{(t)}\right) - \hat{\mathbf{R}}^{(t)}\right) - \left(\mathbf{R}_i\log\left(\mathbf{R}_i\right) - \mathbf{R}_i\right)\right\|_F^2}{\left\|\hat{\mathbf{R}}^{(t)} - \mathbf{R}_i\right\|_F^2}} \tag{21}$$

Proposition 3. *The definition of total Jensen von Neumann divergence is given by choosing the convex function $f(\mathbf{R}) = \mathbf{R}\log(\mathbf{R}) - \mathbf{R}$. The first derivative of the function f and its inverse function are given as,*

$$\nabla f(\mathbf{R}) = \log(\mathbf{R}), \nabla f^{-1}(\mathbf{R}) = exp(\mathbf{R}) \tag{22}$$

Substitution Equation (22) into Equation (18), and we have the median as Equation (21).

Proposition 4. *The median associated with total Jensen log-determinant divergence (13), of m HPD matrices $\{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_m\}$ is given by,*

$$\hat{\mathbf{R}}^{(t+1)} = \left(\sum_{i=1}^{m} w_i^{(t)}\left(\alpha\hat{\mathbf{R}}^{(t)} + (1-\alpha)\mathbf{R}_i\right)^{-1}\right)^{-1}, w_i^{(t)} = \frac{\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right)}{\sum_j \rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_j\right)}$$

$$\rho\left(\hat{\mathbf{R}}^{(t)}, \mathbf{R}_i\right) = 1/\sqrt{1 + \frac{\left\|\log\left(\hat{\mathbf{R}}^{(t)}\right) - \log\left(\mathbf{R}_i\right)\right\|_F^2}{\left\|\hat{\mathbf{R}}^{(t)} - \mathbf{R}_i\right\|_F^2}} \tag{23}$$

Proof of Proposition 3. The total Jensen von Neumann divergence is obtained by choosing the convex function $f(\mathbf{R}) = -\log(\mathbf{R})$, and then we have,

$$\nabla f(\mathbf{R}) = \mathbf{R}^{-1}, \nabla f^{-1}(\mathbf{R}) = \mathbf{R}^{-1} \tag{24}$$

□

Substitution Equation (24) into Equation (18), and we have the median as Equation (23).

According to the definition of total Jensen–Bregman divergence, we can design the corresponding geometric detector illustrated in Figure 3. The HPD matrix $\mathbf{R}_i$ of each cell is estimated by the sample data $\mathbf{x}$ according to its correlation coefficient. The detection statistic is the divergence $\delta_f(\mathbf{R}_D, \hat{\mathbf{R}})$ of the matrix $\mathbf{R}_D$ of cell under test and its median matrix $\hat{\mathbf{R}}$. Thus, the statistic is compared with a given threshold γ for making the decision. The formulation of target detection can be given as,

$$\delta_f(\mathbf{R}_D, \hat{\mathbf{R}}) \underset{target\,absent}{\overset{target\,present}{\gtrless}} \gamma \tag{25}$$
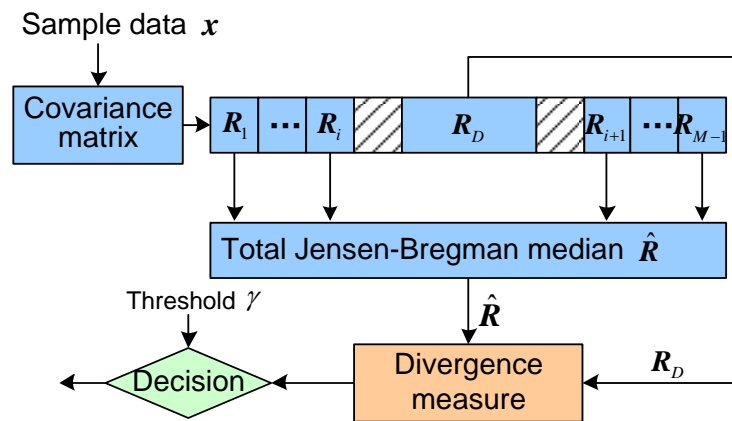
Figure 3. Total Jensen-Bregman divergence-based geometric detector.

5. Experimental Results

In this section, we analyze the performances of our proposed detection algorithms in terms of probability of detection (P_d). We also compare detection performances of our proposed geometric detectors, named as TJSL (total Jensen square loss), TJLD (total Jensen log-determinant divergence), and TJVN (total Jensen von Neumann divergence), in comparison with the classical FFT-CFAR detector [1] and the Riemannian distance-based target detection algorithm [2], abbreviated as the Riemannian detector.

5.1. Numerical Simulation

In simulation experiments, we make use of the Monte Carlo counting technique to test performances of the TJSL detector, the TJVN detector, the TJLD detector, the Riemannian detector, and the FFT-CFAR detector. We consider a radar system, the pulse repetition frequency (PRF) is 1000 Hz, and the radar central frequency f_c is 9 GHz. Seven pulses data are contained in the received echo, and the target model is given as $\tilde{a}\mathbf{p}$, where $\tilde{a}$ is a parameter relating to the backscattering and the channel propagation effects. p denotes the target steering vector,

$$\mathbf{p} = \frac{1}{\sqrt{N}}[1, \exp\left(j2\pi f_d T_r\right), ..., \exp\left(j2\pi\left(N-1\right) f_d T\right)]^T \tag{26}$$

where f_d denotes the Doppler frequency, and T_r is the pulse repetition interval. We assume the target has a velocity $v = 5$ m/s. The number of reference cells is considered for 16. In many high resolution radar scenarios, the Gaussian distribution assumption is not consistent with the actual situation, and cannot be used as clutter model. The compound Gaussian model is more suitable for describing non-Gaussian clutter. The compound Gaussian clutter can be written as the product of two independent random variables. The speckle component is a zero mean Gaussian process, and the texture component is a non-negative random process. It describes the average power level of clutter [30]. In our experiments, the non-Gaussian clutter is simulated via the *K* distribution, the probability density function of *K* distribution is given [31],

$$p\left(x\right) = \frac{\sqrt{2v/a}}{2^{v-1}\Gamma\left(v\right)}\left(\sqrt{\frac{2v}{a}}x\right)^v K_{v-1}\left(\sqrt{\frac{2v}{a}}x\right), x \geq 0 \tag{27}$$

where $K_{v-1}\left(\cdot\right)$ is the modified Bessel function of the second kind with order $v-1$, and $\Gamma\left(\cdot\right)$ is the gamma function. v is the shape parameter, and a denotes the scale parameter.

Simulation experiments are performed to validate detection performances of the TJSL detector, the TJVN detector, and the TJLD detector via the Monte Carlo technique. Results are compared with

the Riemannian and FFT-CFAR detectors. As the threshold cannot be analyzed in theory, several 10^6-trial Monte Carlo simulations in the absence of a target are used to obtain the threshold for a given false alarm probability (P_{fa}). To estimate the detection probability accurately, the P_d is computed by the relative frequency using 200 runs of Monte Carlo simulations.

Figure 4 gives plots of P_d vs SCRs for $P_{fa} = 10^{-5}$. The SCR varies from -20 to 10 dB. Inspection of Figure 4 highlights that our proposed detectors outperform the Riemannian detector. Particularly, the TJLD detector has the best performance when $\alpha = 0.1, a = 0.5, v = 1$; $\alpha = 0.1, a = 0.9, v = 1$; and $\alpha = 0.3, a = 0.5, v = 1$; $\alpha = 0.3, a = 0.9, v = 1$ and followed by the TJSL detector. In addition, the rest of the presented results show that the TJLD detector has similar performance with the TJSL detector.

The strategy of our proposed geometric detector in this paper is that the decision about "target absence" or "target presence" is made according to the dissimilarity between the matrix of cell under test and the clutter covariance matrix estimated by its surrounding cells. The dissimilarity is very different with different used metrics. In our previous work [18,19], we have explored the relationship between the anisotropy of a measure and the performance of its corresponding geometric detector. The estimation accuracy of the covariance matrix is not the factor that affects the performance of the geometric detector, but the anisotropy, which reflects the geometric structure of manifold. Further work will be reported in our other papers.

In the following, we will illustrate the robustness of covariance estimation with respect to the number of samples collected. Given a covariance matrix $\mathbf{R}_0$, several samples are generated by an N-dimension zero-mean Gaussian distribution with covariance matrix $\mathbf{R}_0$. Then, we estimate the covariance matrix $\hat{\mathbf{R}}$ according to these samples. Finally, the estimation accuracy denoted by the error value is computed as,

$$error = \frac{\left\| \mathbf{R}_0 - \hat{\mathbf{R}} \right\|_F}{\left\| \mathbf{R}_0 \right\|_F} \tag{28}$$

The number of samples varies from five to 20, and the number of pulses N is seven. A total of 100 times simulations are repeated to compute the average of the error value. Figure 5 gives the plot of the estimation accuracy of the covariance matrix under different number of samples.

It is clear from Figure 5 that the variance of estimation error is not obvious, as the number of samples varies. This means that the estimated matrix, that is the Riemannian mean, the TJSL median, the TJLD median, and the TJVN median, is robust with respect to the number of samples, especially the TJLD median. Moreover, it can be observed from Figure 5 that the TJLD median has the lowest estimation error, followed by the Riemannian mean. The TJVN median has the maximal estimation error. However, the relationship of estimation error is not consistent with the detection performance which is illustrated in Figure 4.
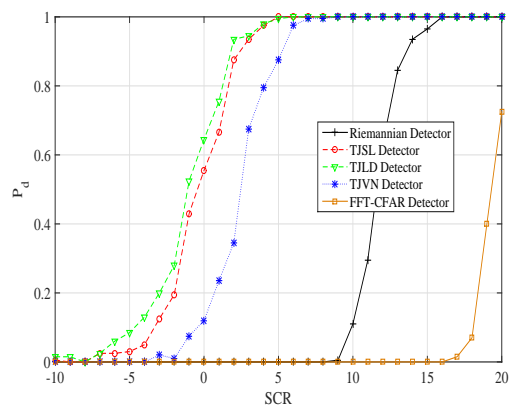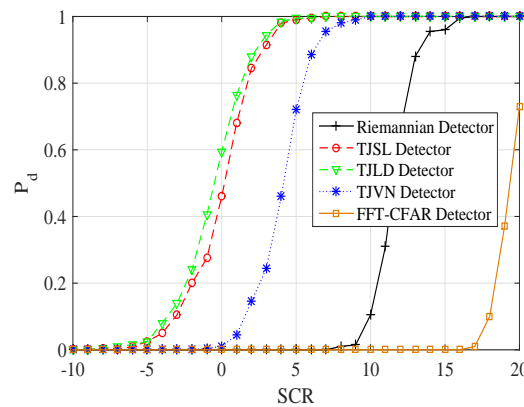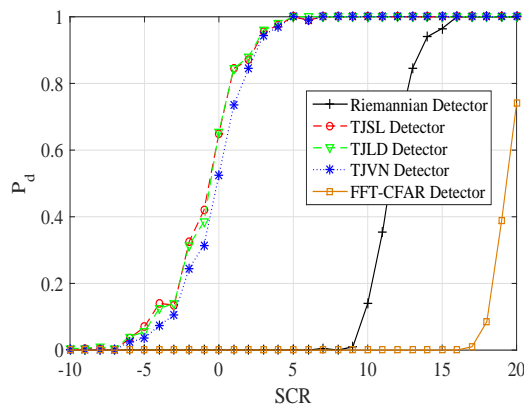
(**a**)P_d versus SCR, $a = 0.5, v = 1, \alpha = 0.1$
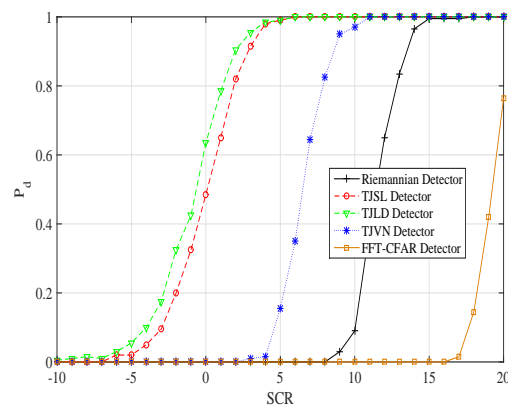
(**b**)P_d versus SCR, $a = 0.9, v = 1, \alpha = 0.1$
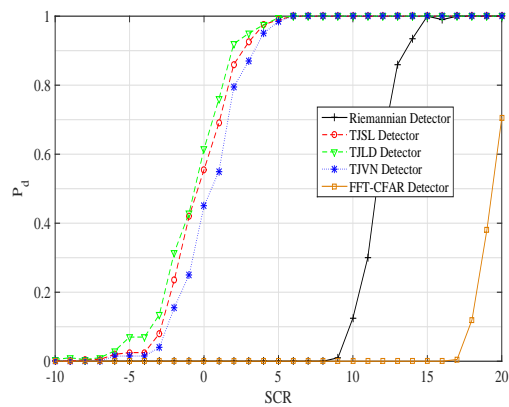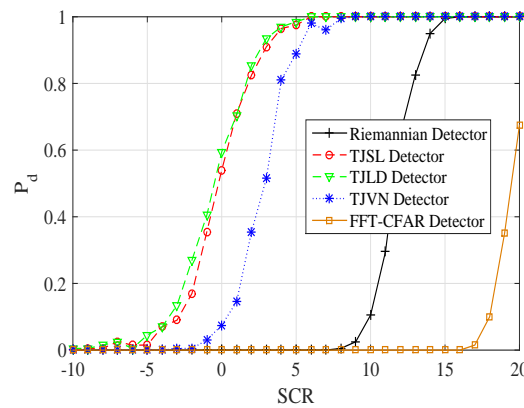
(**c**)P_d versus SCR, $a = 0.5, v = 1, \alpha = 0.3$

(**d**)P_d versus SCR, $a = 0.9, v = 1, \alpha = 0.3$

(**e**)P_d versus SCR, $a = 0.5, v = 1, \alpha = 0.5$

(**f**)P_d versus SCR, $a = 0.9, v = 1, \alpha = 0.5$

Figure 4. P_d versus SCR in K distribution, $P_{fa} = 10^{-5}$.
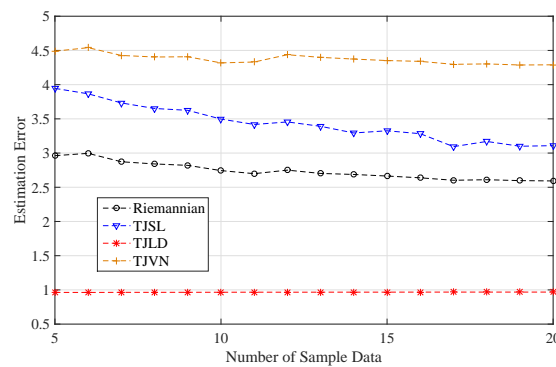
Figure 5. The estimation accuracy of the covariance matrix under different number of samples.

5.2. Real Clutter Data

To evaluate the performance of the proposed detectors in a real application situation, we use data collected by the McMaster University IPIX radar [32]. Specifically, 19980205-192053-antstep.cdf (file 1), and 19980205-185111-antstep.cdf (file 2) [33] are employed to test detection performances of our proposed detectors, Riemannian detector, and FFT-CFAR detection algorithm. For these two data files, the number of samples in the range dimension is 27, and the pulse dimension consists of 60,000 samples. The pulse repetition frequency is 1000 Hz, and the central frequency is 9.39 GHz. The horizontal polarization (HH) information of samples are considered. In particular, a synthetic target is needed to inject into one cell due to the unavailable target information, and the velocity is approximate constant, $v = 5\,\mathrm{m/s}$.

The dimension of the HPD matrix is 5, the threshold is set using the first $50,000$ pulse in each cell for $P_{fa} = 10^{-3}$. Other parameters are also as in the former subsection. The simulation is repeated 200 times to determine the Pd by means of the relative frequency.

We compare the performance of our proposed detectors using the above two real sea clutter data. Plots of detection probability for different detectors are shown in Figure 6. The P_{fa} is set to 10^{-3}, and the range of SCR is between -10 and 15 dB. From Figure 6, it is clear that our proposed detector and the Riemannian detector outperform the classical FFT-CFAR detection algorithm in real clutter environment. The detection performance of TJSL detector is the best in file1, and the performance of TJSL detector is comparable with the TJVN detector with $\alpha = 0.5$. Moreover, the results imply that our proposed detectors may have different performances in different real data files. The performance analysis of our proposed detectors under different clutter environments is needed in further work.
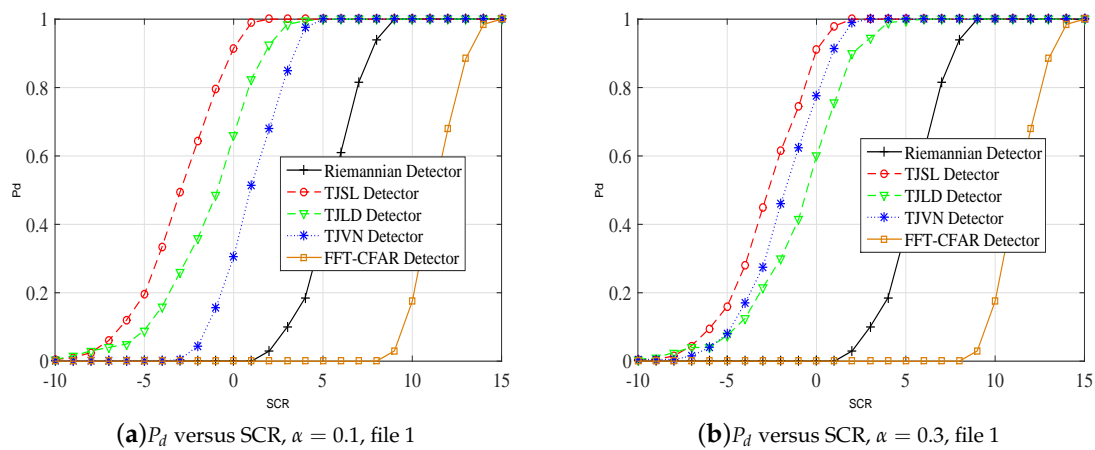

(**a**)P_d versus SCR, $\alpha = 0.1$, file 1　　　　　　　　(**b**)P_d versus SCR, $\alpha = 0.3$, file 1

Figure 6. *Cont.*

(c) P_d versus SCR, $\alpha = 0.5$, file 1

(d) P_d versus SCR, $\alpha = 0.1$, file 2

(e) P_d versus SCR, $\alpha = 0.3$, file 2
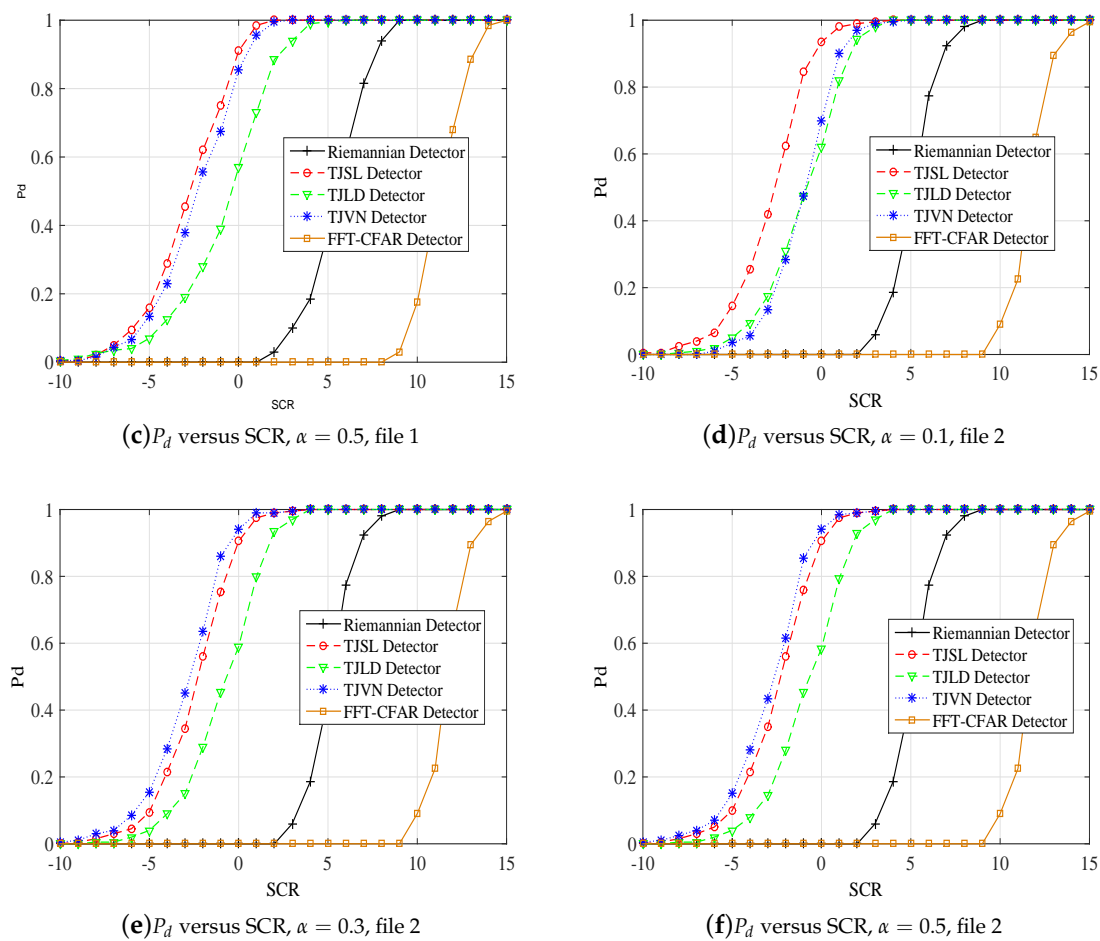
(f) P_d versus SCR, $\alpha = 0.5$, file 2

Figure 6. P_d versus SCR in real clutter environment, $P_{fa} = 10^{-3}$.

6. Conclusions

In this paper, we have developed a radar target detection algorithm using the matrix information geometry method. The total Jensen–Bregman divergence is defined and used as the distance-like function on the Riemannian manifold of Hermitian positive-definite matrix. Definitions of three divergences, including the TJSL, the TJLD, the TJVN, are given according to different forms of the convex function f. Then, medians associated with the total Jensen–Bregman divergence are derived. For the stage of the performance analysis, we compare detection performances of our proposed detectors with the Riemannian detector and the FFT-CFAR detection algorithm using the simulation data and the real sea clutter data file. The comparison results have shown a remarkable advantage with respect to the conventional FFT-CFAR detection algorithm and the Riemanian detector.

Author Contributions: Xiaoqiang Hua put forward the original ideas and performed the research. Haiyan Fan and Yongqiang Cheng conceived and designed the simulations. Hongqiang Wang and Yuliang Qin reviewed the paper and provided useful comments. All authors have read and approved the final manuscript.

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

References

1. Richards, M. *Fundamentals of Radar Signal Processing*; McGraw-Hill Professional: New York, NY, USA, 2005.
2. Lapuyade-Lahorgue, J.; Barbaresco, F. Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar. In Proceedings of the 2008 RADAR '08 IEEE Radar Conference, Rome, Italy, 26–30 May 2008; pp. 1–6.
3. Barbaresco, F. New Foundation of radar Doppler signal processing based on advanced differential geometry of symmetric spaces: Doppler matrix CFAR and radar application. In Proceedings of the 2009 RADAR '09 IEEE Radar Conference, Torino, Italy, 14–18 September 2009.
4. Arnaudon, M.; Barbaresco, F.; Yang, L. Medians and Means in Riemannian Geometry: Existence, Uniqueness and Computation. *Matrix Inf. Geom.* **2011**, *1*, 169–197.
5. Liu, Z.; Barbaresco, F. *Doppler Information Geometry for Wake Turbulence Monitoring*; Springer: Berlin/Heidelberg, Germany, 2013.
6. Barbaresco, F. Method for Radar Monitoring of Wake Turbulence. U.S. Patent 8,334,799, 18 December 2012.
7. Barbaresco, F.; Meier, U. Radar monitoring of a wake vortex: Electromagnetic reflection of wake turbulence in clear air. *Comptes Rendus Phys.* **2010**, *11*, 54–67.
8. Charfi, M.; Chebbi, Z.; Moakher, M.; Vemuri, B.C. Using the Bhattacharyya Mean for the Filtering and Clustering of Positive-Definite Matrices. In Proceedings of the GSI 2013 Geometric Science of Information: First International Conference, Paris, France, 28–30 August 2013; pp. 551–558.
9. Charfi, M.; Chebbi, Z.; Moakher, 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 International Symposium on Biomedical Imaging, San Francisco, CA, USA, 7–11 April 2013; pp. 1227–1230.
10. Kullback, S.; Leibler, R.A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.
11. 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.
12. Teran, A.R.M.Y.; Gouiffes, M.; Lacassagne, L. Total Bregman Divergence for Multiple Object Tracking. In Proceedings of the International Conference on Image Processing, Melbourne, Australia, 15–18 September 2013.
13. Liu, M.; Vemuri, B.C.; Amari, S.I.; Nielsen, F. Total Bregman divergence and its applications to shape retrieval. In Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 3463–3468.
14. 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.
15. Nielsen, F.; Nock, R. Total Jensen divergences: Definition, properties and clustering. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, South Brisbane, Australia, 19–24 April 2015; pp. 2016–2020.
16. Nielsen, F.; Nock, R. Total Jensen divergences: Definition, Properties and k-Means++ Clustering. *arXiv* **2013**, arXiv:1309.7109.
17. 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.
18. 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.
19. 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.
20. Hua, X.; Cheng, Y.; Wang, H.; Qin, Y.; Chen, D. Geometric target detection based on total Bregman divergence. *Digit. Signal Process.* **2018**, *75*, 232–241.
21. Bhatia, R. *Positive Definite Matrices*; Princeton University Press: Princeton, NJ, USA, 2007.
22. Sra, S. Positive definite matrices and the Symmetric Stein Divergence. *arXiv* **2011**, arXiv:1110.1773v2.
23. Bridson, M.R.; Haefliger, A. *Metric Spaces of Non-Positive Curvature*; Springer: Berlin/Heidelberg, Germany, 1999.
24. Terras, A. *Harmonic Analysis on Symmetric Spaces and Applications II*; Springer-Verlag: Berlin, Germany, 2014.
25. Wolkowicz, H.; Saigal, R.; Vandenberghe, L. *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*; Springer: New York, NY, USA, 2000.
26. Lang, S. *Fundamentals of Differential Geometry*; Springer: Berlin/Heidelberg, Germany, 1999.

27. Fletcher, P.T.; Venkatasubramanian, S.; Joshi, S. Robust statistics on Riemannian manifolds via the geometric median. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.

28. Yuille, A.L.; Rangarajan, A. The Concave-Convex Procedure. *Neural Comput.* **2003**, *15*, 915–936.

29. Nielsen, F.; Boltz, S. The Burbea-Rao and Bhattacharyya Centroids. *IEEE Trans. Inf. Theory* **2011**, *57*, 5455–5466.

30. Roy, L.P.; Raja Kumar, R.V. A GLRT detector in partially correlated texture based compound-Gaussian clutter. In Proceedings of the 2010 National Conference on, Communications (NCC), Chennai, India, 29–31 January 2010; pp. 1–5.

31. Watts, S. Radar Detection Prediction in K-Distributed Sea Clutter and Thermal Noise. *IEEE Trans. Aerosp. Electron. Syst.* **1987**, *23*, 40–45.

32. Haykin, S.; Krasnor, C.; Nohara, T.J.; Currie, B.W. A coherent dual-polarized radar for studying the ocean environment. *IEEE Trans. Geosci. Remote Sens.* **1991**, *29*, 189–191.

33. IPIX Radar File. IPIX Radar Dataset Files in Dartmouth. 1993. Available online: http://www.ipixbpo.com/ (accessed on 6 April 2018).