

Article

# Robust Covariance Estimators Based on Information Divergences and Riemannian Manifold

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: hxq712@yeah.net (X.H.); cyq101600@126.com (Y.C.)

Received: 25 February 2018; Accepted: 12 March 2018; Published: 23 March 2018



**Abstract:** This paper proposes a class of covariance estimators based on information divergences in heterogeneous environments. In particular, the problem of covariance estimation is reformulated on the Riemannian manifold of Hermitian positive-definite (HPD) matrices. The means associated with information divergences are derived and used as the estimators. Without resorting to the complete knowledge of the probability distribution of the sample data, the geometry of the Riemannian manifold of HPD matrices is considered in mean estimators. Moreover, the robustness of mean estimators is analyzed using the influence function. Simulation results indicate the robustness and superiority of an adaptive normalized matched filter with our proposed estimators compared with the existing alternatives.

**Keywords:** information divergence; Riemannian manifold; covariance estimation; mean estimator; heterogeneous clutter

## 1. Introduction

Covariance estimation plays an important role in adaptive signal processing, such as multichannel signal processing [1,2], space-time adaptive processing (STAP) [3,4], and radar target detection. Conventional covariance estimation methods, derived from the maximum-likelihood (ML) of the clutter data, are based on the assumption that the clutter remains stationary and homogeneous during the adaptation process. However, in real heterogeneous clutter, the number of the sample data where the clutter is homogeneous is very limited, and the estimation performance is seriously degraded. Therefore, it is necessary and important to improve the performance of estimated covariance in heterogeneous clutter.

A commonly used strategy to ameliorate the performance of estimated covariance is to exploit some a priori information about the clutter environment. For instance, the geographic information is used for covariance estimation, and the performance of target detection with the estimator is significantly improved [5]. In [6], the authors employ the Bayesian method to perform the target detection in interference environment, where the unknown covariance matrix is assumed to follow a suitable probability model. In [7,8], a Bayesian framework is used together with the structural information about the estimated covariance. In [9], a condition number upper-bound constraint is imposed on the problem of covariance estimation to achieve a signal-to-noise ratio improvement. Moreover, a symmetrically structured spectral density is constrained on the covariance estimation, and these results show the superiority of the estimator [10]. These mentioned methods rely on the knowledge of statistics characterization of the clutter data. However, the probability distribution of the whole environment is difficult to obtain, and a mismatched distribution results in a remarkable degradation of the estimation performance in heterogeneous clutter.

Many covariance estimation algorithms derived from the geometry of matrix space, not resorting to the statistical characterization of the sample data, are reported in the literature. For instance, the Riemannian mean is used for monitoring the wake turbulence [11,12] and target detection in HF and X-band radar [13]. In [14–16], Riemannian mean and median are employed for covariance estimation in STAP, and results have shown that the projection algorithm with Riemannian mean can yield significant performance gains. In [17,18], some geometric barycenter and medians are proposed for radar training data selection in homogeneous environment. In recent times, we have explored information divergence means and medians for target detection in non-Gaussian clutter [19–21]. Moreover, in image processing applications, Bhattacharyya mean and median are used for filtering and clustering of diffusion tensor magnetic resonance image [22,23]. In [24], the Log-Euclidean mean, together with the reproducing kernel Hilbert mapping, is used for texture recognition. These geometric approaches have achieved good performances.

In this paper, a class of covariance estimators based on information divergences is proposed in heterogeneous clutter. In particular, six means related to geometric measures are derived on the Riemannian manifold of Hermitian positive-definite (HPD) matrices. These means do not rely on the knowledge of statistics characterization of sample data, and the geometry of the matrix space is considered. Moreover, the robustness of means is analyzed with injected outliers via the influence function. Simulation results are given to validate the superiority of proposed estimators.

The rest of this paper is organized as follows. In Section 2, we reformulate the problem of covariance estimation on the Riemannian manifold. In Section 3, the geometry of Riemannian manifold of HPD matrices is presented, in particular, six distance measures are given on the Riemannian manifold, and means associated with these measures are derived. The robustness of means is analyzed via the influence function in Section 4. Then, we evaluate performances of an adaptive normalized matched filter with geometric means as well as the normalized sample covariance matrix in Section 5. Finally, conclusion is provided in Section 6.

### Notation

Here is some notation for the descriptions 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  $\mathbf{A}^H$ .  $\text{tr}(\mathbf{A})$  is the trace of matrix  $\mathbf{A}$ .  $|\mathbf{A}|$  is the determinant of matrix  $\mathbf{A}$ .  $\mathbf{I}$  denotes the identity matrix. Finally,  $\mathbb{E}(\cdot)$  denotes the statistical expectation.

## 2. Problem Reformulated on the Riemannian Manifold

A heterogeneous environment is considered for covariance estimation. For a set of  $K$  secondary data  $\{\mathbf{r}_1, \dots, \mathbf{r}_K\}$ , the normalized sample covariance matrix (NSCM)  $\hat{\mathbf{R}}$  based on the ML of probability distribution of the sample data is estimated as,

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^K \frac{\mathbf{r}_k \mathbf{r}_k^H}{\frac{1}{N} \mathbf{r}_k^H \mathbf{r}_k} = \frac{1}{K} \sum_{k=1}^K \mathbf{R}_k, \quad (1)$$

$$\mathbf{R}_k = \frac{\mathbf{r}_k \mathbf{r}_k^H}{\frac{1}{N} \mathbf{r}_k^H \mathbf{r}_k}$$

where  $N$  is the dimension of the vector  $\mathbf{r}_k$ .  $\mathbf{r}_k, k = 1, \dots, K$  are modeled as a compound-Gaussian random vector, and can be expressed as,

$$\mathbf{r}_k = \sqrt{\tau_k} \mathbf{g}_k, k = 1, \dots, K \quad (2)$$

where  $\tau_k$  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,

$$\mathbb{E}[\mathbf{g}_k \mathbf{g}_k^H] = \mathbf{\Sigma}, k = 1, \dots, K \tag{3}$$

It is clear from Equation (1) that the NSCM is the arithmetic mean of  $K$  auto-covariance matrices  $\mathbf{R}_k$  of rank one. Since the knowledge of probability distribution of the whole environment is difficult to obtain in heterogeneous clutter, the performance of NSCM is severely degraded. Actually, these  $K$  auto-covariance matrices lie in a non-linear Hermitian matrix space, as the sample data is complex. It is well known that HPD matrices form a differentiable Riemannian manifold [25], that is the most studied example of a manifold with non-positive curvature [26]. In order to facilitate the analysis, the matrix  $\mathbf{R}_k$  is transformed to the positive-definite matrix  $\mathbf{P}_k$  using the following three ways:

- (1)  $\mathbf{P}_k$  is obtained by adding an identity matrix  $\mathbf{I}$ , as  $\mathbf{P}_k = \mathbf{R}_k + \mathbf{I}$ ;
- (2) A Toeplitz HPD matrix  $\mathbf{T}_k$  is utilized. As in [13],  $\mathbf{T}_k$  can be expressed as,

$$\mathbf{T}_k = [\mathbf{r}_k \mathbf{r}_k^H] = \begin{bmatrix} c_0 & \bar{c}_1 & \cdots & \bar{c}_{N-1} \\ c_1 & c_0 & \cdots & \bar{c}_{N-2} \\ \vdots & \ddots & \ddots & \vdots \\ c_{N-1} & \cdots & c_1 & c_0 \end{bmatrix}, c_k = [r_i \bar{r}_{i+k}], 0 \leq k \leq N - 1, 0 \leq i \leq N - 1 \tag{4}$$

where  $c_k$  denotes the correlation coefficient of sample data, and  $\bar{c}_i$  is the complex conjugate of  $c_i$ .  $c_k$  can be computed as,

$$c_k = \frac{1}{N - k} \sum_{j=1}^{N-1-k} r_j \bar{r}_{j+k}, 0 \leq k \leq N - 1 \tag{5}$$

- (3) The HPD matrix  $\mathbf{S}_k$  is the solution of the optimization problem as follows [17],

$$\begin{cases} \min_{\mathbf{S}_k} \|\mathbf{r}_k \mathbf{r}_k^H / \mathbf{r}_k^H \mathbf{r}_k - \mathbf{S}_k\|^2 \\ \mathbf{S}_k \geq \mathbf{I} \\ \frac{\lambda_{\max}(\mathbf{S}_k)}{\lambda_{\min}(\mathbf{S}_k)} \leq \kappa_M \end{cases} \tag{6}$$

The optimal solution can be given by [17],

$$\begin{aligned} \mathbf{S}_k &= \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H \\ \mathbf{\Lambda}_k &= \text{diag}([\kappa_M \lambda_k, \lambda_k, \dots, \lambda_k]) \\ \lambda_k &= \max\left(1, \frac{\|\mathbf{r}_k\|^2 \kappa_M}{\kappa_M^2 + N - 1}\right) \end{aligned} \tag{7}$$

where  $\mathbf{U}_k$  is a unitary matrix of the eigenvectors of  $\mathbf{r}_k \mathbf{r}_k^H$  with the first eigenvector corresponding to the eigenvalue  $\|\mathbf{r}_k\|^2$ . The  $\kappa_M$  is the condition number.

According to above transformations, we can obtain the HPD matrix. Then, the problem of covariance estimation can be reformulated on the Riemannian manifold. In general, for a set of  $m$

scalar numbers  $x_1, \dots, x_m$ , the arithmetic mean is defined as the minimum of sum of squared distances to the given point  $x$ ,

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i = \arg \min_{x>0} \sum_{i=1}^m |x - x_i|^2 \tag{8}$$

From Equation (8), we can understand the arithmetic mean from a geometric viewpoint. Similar to Equation (8), for  $K$  HPD matrices  $\mathbf{P}_1, \dots, \mathbf{P}_K$ , the mean related to a measure can be defined as,

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P}} \frac{1}{K} \sum_{k=1}^K d^2(\mathbf{P}, \mathbf{P}_k) \tag{9}$$

where  $d(.,.)$  denotes the measure. It is worth pointing out that the arithmetic mean Equation (1) is obtained, when  $d$  is the Frobenius norm and  $\mathbf{P}_k$  is replaced by  $\mathbf{R}_k$ . The difference between the arithmetic mean and the geometric mean is shown in Figure 1. 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.

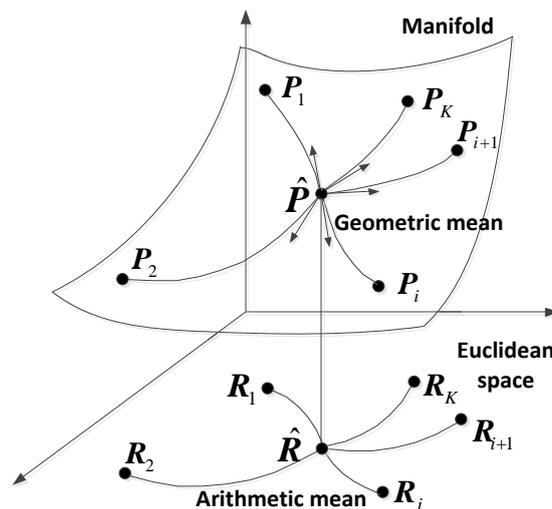


Figure 1. The geometric mean and the arithmetic mean.

### 3. The Geometry of Riemannian Manifold of HPD Matrices

In this Section, the fundamental mathematic knowledge related to this paper is presented. Firstly, the Riemannian manifold of HPD matrices is introduced. Then, six distance measures are presented. Finally, the means related to measures are derived.

#### 3.1. The Riemannian Manifold of HPD Matrices

Let  $\mathbb{H}(n) = \{\mathbf{A} | \mathbf{A} = \mathbf{A}^H\}$  denotes 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 vector. All  $n \times n$  HPD matrices consist of a positive-definite Hermitian matrix space  $\mathbb{P}(n)$ ,

$$\mathbb{P}(n) = \{\mathbf{A} | \mathbf{A} > 0, \mathbf{A} \in \mathbb{H}(n)\} \tag{10}$$

forms a Riemannian manifold of dimension  $n(n+1)/2$  with a constant non-positive curvature [26]. For a point  $\mathbf{A}$  on the Riemannian manifold, the infinitesimal arclength between  $\mathbf{A}$  and  $\mathbf{A} + d\mathbf{A}$  is given by [27],

$$ds := (\text{tr}(\mathbf{A}^{-1}d\mathbf{A})^2)^{1/2} = \|\mathbf{A}^{-1/2}d\mathbf{A}\mathbf{A}^{-1/2}\|_F \quad (11)$$

where  $ds$  defines a metric on the Riemannian manifold [27].  $\|\cdot\|_F$  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 [28],

$$\langle \mathbf{P}_1, \mathbf{P}_2 \rangle_{\mathbf{A}} = \text{tr}(\mathbf{A}^{-1}\mathbf{P}_1\mathbf{A}^{-1}\mathbf{P}_2), \|\mathbf{P}_1\|_{\mathbf{A}} = \langle \mathbf{P}_1, \mathbf{P}_1 \rangle_{\mathbf{A}}^{1/2} \quad (12)$$

For two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$  on the Riemannian manifold, the affine invariant (Riemannian) distance is given by [29],

$$d_{\mathbb{R}}^2(\mathbf{P}_1, \mathbf{P}_2) = \|\text{logm}(\mathbf{P}_1^{-1/2}\mathbf{P}_2\mathbf{P}_1^{-1/2})\|_F^2 \quad (13)$$

where  $\text{logm}$  is the logarithmic map on the Riemannian manifold of HPD matrices.

### 3.2. The Geometric Measure on the Riemannian Manifold

In addition to the Riemannian distance, a lot of distance or information divergences can be used as the measurement on the Riemannian manifold. Here, five geometric measures are presented in the following.

#### (1) Log-Euclidean distance

The Log-Euclidean distance is also a geodesic distance. It is defined on the tangent space at a point on the Riemannian manifold, which is isomorphic and diffeomorphic to the tangent space identified with a Hermitian matrix space. For two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the Log-Euclidean distance  $d_{LE}(\mathbf{P}_1, \mathbf{P}_2)$  is given by [30],

$$d_{LE}(\mathbf{P}_1, \mathbf{P}_2) = \|\text{logm}(\mathbf{P}_1) - \text{logm}(\mathbf{P}_2)\|_F \quad (14)$$

#### (2) Hellinger distance

The Hellinger distance is a special case of the  $\alpha$ -divergence with  $\alpha = 0$ . Given two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the Hellinger distance  $d_H(\mathbf{P}_1, \mathbf{P}_2)$  is [31],

$$d_H(\mathbf{P}_1, \mathbf{P}_2) = \sqrt{2 - 2 \frac{|\mathbf{P}_1|^{1/4} |\mathbf{P}_2|^{1/4}}{|\frac{1}{2}(\mathbf{P}_1 + \mathbf{P}_2)|^{1/2}}} \quad (15)$$

#### (3) Kullback-Leibler divergence

It is well known that the Kullback-Leibler (KL) divergence is the most widely used measure on the Riemannian manifold. The KL divergence is also a special case of the  $\alpha$ -divergence with  $\alpha = \pm 1$ . In addition, the KL divergence is called the Stein loss or the log-determinant divergence. The KL divergence  $d_{KL}(\mathbf{P}_1, \mathbf{P}_2)$  between two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$  can be given by [32],

$$d_{KL}(\mathbf{P}_1, \mathbf{P}_2) = \text{tr}(\mathbf{P}_2^{-1}\mathbf{P}_1 - \mathbf{I}) - \log |\mathbf{P}_2^{-1}\mathbf{P}_1| \quad (16)$$

(4) Bhattacharyya distance

The Bhattacharyya distance is a common used measure, and has been used in medical image segmentation [22]. In particular, the Bhattacharyya distance is a Jensen version of the KL divergence. For two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the Bhattacharyya distance  $d_B(\mathbf{P}_1, \mathbf{P}_2)$  can be given by,

$$d_B(\mathbf{P}_1, \mathbf{P}_2) = 2\sqrt{\log \frac{|\frac{\mathbf{P}_1 + \mathbf{P}_2}{2}|}{\sqrt{|\mathbf{P}_1| |\mathbf{P}_2|}}} \tag{17}$$

(5) Symmetrized Kullback-Leibler divergence

The symmetrized Kullback-Leibler (SKL) divergence is a Jeffreys divergence [33]. It behaves as the square of a distance; however, it is not a distance, as the triangle inequality does not hold. Given two points  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the SKL  $d_{SKL}(\mathbf{P}_1, \mathbf{P}_2)$  between them is,

$$d_{SKL}(\mathbf{P}_1, \mathbf{P}_2) = \frac{1}{2}\{d_{KL}(\mathbf{P}_1, \mathbf{P}_2) + d_{KL}(\mathbf{P}_2, \mathbf{P}_1)\} = \frac{1}{2}tr(\mathbf{P}_2^{-1}\mathbf{P}_1 + \mathbf{P}_1^{-1}\mathbf{P}_2 - 2\mathbf{I}) \tag{18}$$

3.3. The Geometric Mean for A Set of HPD Matrices

In this Section, the means related to the above mentioned six measures are derived using the fix-point algorithm. This work has been done in our previous article [19]. In the following, six means are presented in Table 1.

**Table 1.** Geometric means related to different measures.

Geometric Measure	Mean
Riemannian	$\bar{\mathbf{P}}_{t+1} = \bar{\mathbf{P}}_t^{1/2} \exp \left\{ \varepsilon_t \left( \sum_{k=1}^K \log \left( \bar{\mathbf{P}}_t^{-1/2} \mathbf{P}_k \bar{\mathbf{P}}_t^{-1/2} \right) \right) \right\} \bar{\mathbf{P}}_t^{1/2}$
Log-Euclidean	$\bar{\mathbf{P}} = \exp \left( \frac{1}{K} \sum_{i=1}^K \log(\mathbf{P}_i) \right)$
Hellinger	$\bar{\mathbf{P}}_{t+1} = \left( \frac{\sum_{i=1}^K  \bar{\mathbf{P}}_t ^{1/4}  \mathbf{P}_i ^{1/4}  \frac{1}{2}(\bar{\mathbf{P}}_t + \mathbf{P}_i) ^{-1/2} (\frac{1}{2}(\bar{\mathbf{P}}_t + \mathbf{P}_i))^{-1}}{\left( \sum_{i=1}^K  \bar{\mathbf{P}}_t ^{1/4}  \mathbf{P}_i ^{1/4}  \frac{1}{2}(\bar{\mathbf{P}}_t + \mathbf{P}_i) ^{-1/2} \right)} \right)^{-1}$
KL	$\bar{\mathbf{P}} = \left( \frac{1}{K} \sum_{i=1}^K \mathbf{P}_i^{-1} \right)^{-1}$
Bhattacharyya	$\bar{\mathbf{P}}_{t+1} = \left( \frac{1}{K} \sum_{i=1}^K \left( \frac{\bar{\mathbf{P}}_t + \mathbf{P}_i}{2} \right)^{-1} \right)^{-1}$
SKL	$\bar{\mathbf{P}} = \left( \left( \frac{1}{K} \sum_{i=1}^K \mathbf{P}_i \right) \left( \frac{1}{K} \sum_{k=1}^K \mathbf{P}_k^{-1} \right)^{-1} \right)^{1/2}$

Where  $t$  is the number of iteration, and  $\varepsilon_t$  is the step size of iteration.

**4. Robustness Analysis of Geometric Means**

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

**Proposition 1.** The influence function of arithmetic mean related to the Frobenius norm, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,

$$H(\mathbf{Q}) = \frac{1}{n} \sum_{i=1}^n \mathbf{Q}_i - \bar{\mathbf{P}} \tag{19}$$

**Proof of Proposition 1.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \|\mathbf{P} - \mathbf{P}_i\|_F^2 + \varepsilon \frac{1}{n} \sum_{j=1}^n \|\mathbf{P} - \mathbf{Q}_j\|_F^2 \tag{20}$$

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

$$\nabla F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 2(\mathbf{P} - \mathbf{P}_i) + \varepsilon \frac{1}{n} \sum_{j=1}^n 2(\mathbf{P} - \mathbf{Q}_j) \tag{21}$$

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

$$\begin{aligned} \tilde{\mathbf{P}} &= \arg \min_{\mathbf{P}} F(\mathbf{P}) \\ \Rightarrow \nabla F(\tilde{\mathbf{P}}) &= (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 2(\tilde{\mathbf{P}} - \mathbf{P}_i) + \varepsilon \frac{1}{n} \sum_{j=1}^n 2(\tilde{\mathbf{P}} - \mathbf{Q}_j) = 0 \end{aligned} \tag{22}$$

and

$$\begin{aligned} \bar{\mathbf{P}} &= \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m \|\mathbf{P} - \mathbf{P}_i\|_F^2 \\ \Rightarrow \nabla F(\bar{\mathbf{P}}) &= \frac{1}{m} \sum_{i=1}^m 2(\bar{\mathbf{P}} - \mathbf{P}_i) = 0 \end{aligned} \tag{23}$$

Substitute  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$  into Equation (22), and we have

$$\begin{aligned} &2(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\bar{\mathbf{P}} + \varepsilon H(\mathbf{Q}) - \mathbf{P}_i) + 2\varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{P}} + \varepsilon H(\mathbf{Q}) - \mathbf{Q}_j) = 0 \\ \Rightarrow &(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\bar{\mathbf{P}} - \mathbf{P}_i) + (1 - \varepsilon)\varepsilon H(\mathbf{Q}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{P}} + \varepsilon H(\mathbf{Q}) - \mathbf{Q}_j) = 0 \\ \Rightarrow &\varepsilon H(\mathbf{Q}) - \varepsilon^2 H(\mathbf{Q}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{P}} - \mathbf{Q}_j) + \varepsilon^2 H(\mathbf{Q}) = 0 \\ \Rightarrow &H(\mathbf{Q}) = \frac{1}{n} \sum_{j=1}^n (\mathbf{Q}_j - \bar{\mathbf{P}}) \end{aligned} \tag{24}$$

□

**Proposition 2.** The influence function of Riemannian mean related to the Riemannian distance, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,

$$H(\mathbf{Q}) = -\frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) \bar{\mathbf{P}} \tag{25}$$

**Proof of Proposition 2.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \|\log(\mathbf{P}_i^{-1} \mathbf{P})\|_F^2 + \varepsilon \frac{1}{n} \sum_{j=1}^n \|\log(\mathbf{Q}_j^{-1} \mathbf{P})\|_F^2 \tag{26}$$

As  $\tilde{\mathbf{P}}$  is the mean of  $m$  matrices and  $n$  outliers, and  $\bar{\mathbf{P}}$  is the mean of  $m$  matrices, then, we have,

$$\begin{aligned} \tilde{\mathbf{P}} &= \arg \min_{\mathbf{P}} F(\mathbf{P}) \\ \Rightarrow \nabla F(\tilde{\mathbf{P}}) &= 2(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \log(\mathbf{P}_i^{-1} \tilde{\mathbf{P}}) \tilde{\mathbf{P}}^{-1} + 2\varepsilon \frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \tilde{\mathbf{P}}) \tilde{\mathbf{P}}^{-1} = 0 \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \log(\mathbf{P}_i^{-1} \tilde{\mathbf{P}}) &+ \varepsilon \frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \tilde{\mathbf{P}}) = 0 \end{aligned} \tag{27}$$

and

$$\begin{aligned} \bar{\mathbf{P}} &= \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m \|\log(\mathbf{P}_i^{-1} \mathbf{P})\|_F^2 \\ \Rightarrow \nabla G(\bar{\mathbf{P}}) &= \frac{1}{m} \sum_{i=1}^m 2 \log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) \bar{\mathbf{P}}^{-1} = 0 \Rightarrow \frac{1}{m} \sum_{i=1}^m \log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) = 0 \end{aligned} \tag{28}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have

$$\begin{aligned} \log(\mathbf{P}_i^{-1} \tilde{\mathbf{P}}) &= \log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \mathbf{P}_i^{-1} (\mathbf{P}_i^{-1} \bar{\mathbf{P}})^{-1} = \log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} \\ \log(\mathbf{Q}_j^{-1} \tilde{\mathbf{P}}) &= \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \mathbf{Q}_j^{-1} (\mathbf{Q}_j^{-1} \bar{\mathbf{P}})^{-1} = \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} \end{aligned} \tag{29}$$

Substitute Equations (28) and (29) into Equation (27), and we have

$$\begin{aligned} (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1}) &+ \varepsilon \frac{1}{n} \sum_{j=1}^n (\log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1}) = 0 \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \log(\mathbf{P}_i^{-1} \bar{\mathbf{P}}) &+ (1 - \varepsilon) \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} + \varepsilon \frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) + \varepsilon^2 H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} = 0 \\ \Rightarrow (1 - \varepsilon) \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} &+ \varepsilon \frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) + \varepsilon^2 H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} = 0 \end{aligned} \tag{30}$$

Ignore the terms contain  $\varepsilon^2$  for the constant  $\varepsilon \ll 1$ , and we can obtain,

$$\begin{aligned} \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-1} + \varepsilon \frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) &= 0 \\ \Rightarrow H(\mathbf{Q}) &= -\frac{1}{n} \sum_{j=1}^n \log(\mathbf{Q}_j^{-1} \bar{\mathbf{P}}) \bar{\mathbf{P}} \end{aligned} \tag{31}$$

□

**Proposition 3.** The influence function of Log-Euclidean mean related to the Log-Euclidean distance, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,

$$H(\mathbf{Q}) = \frac{1}{n} \sum_{j=1}^n (\log(\mathbf{Q}_j) - \log(\bar{\mathbf{P}})) \bar{\mathbf{P}} \tag{32}$$

**Proof of Proposition 3.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \|\log(\mathbf{P}) - \log(\mathbf{P}_i)\|_F^2 + \varepsilon \frac{1}{n} \sum_{j=1}^n \|\log(\mathbf{P}) - \log(\mathbf{Q}_j)\|_F^2 \tag{33}$$

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

$$\begin{aligned} \tilde{\mathbf{P}} &= \arg \min_{\mathbf{P}} F(\mathbf{P}) \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 2(\log(\tilde{\mathbf{P}}) - \log(\mathbf{P}_i))\tilde{\mathbf{P}}^{-1} + \varepsilon \frac{1}{n} \sum_{j=1}^n 2(\log(\tilde{\mathbf{P}}) - \log(\mathbf{Q}_j))\tilde{\mathbf{P}}^{-1} &= 0 \tag{34} \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\log(\tilde{\mathbf{P}}) - \log(\mathbf{P}_i)) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\log(\tilde{\mathbf{P}}) - \log(\mathbf{Q}_j)) &= 0 \end{aligned}$$

and

$$\begin{aligned} \bar{\mathbf{P}} &= \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m \|\log(\bar{\mathbf{P}}) - \log(\mathbf{P}_i)\|_F^2 \\ \Rightarrow \nabla G(\mathbf{P}) &= \frac{1}{m} \sum_{i=1}^m 2(\log(\bar{\mathbf{P}}) - \log(\mathbf{P}_i))\bar{\mathbf{P}}^{-1} = 0 \tag{35} \\ \Rightarrow \nabla G(\mathbf{P}) &= \frac{1}{m} \sum_{i=1}^m (\log(\bar{\mathbf{P}}) - \log(\mathbf{P}_i)) = 0 \end{aligned}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have

$$\log(\tilde{\mathbf{P}}) = \log(\bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q})\bar{\mathbf{P}}^{-1} \tag{36}$$

Substitute Equations (35) and (36) into Equation (34), and ignore the terms contain  $\varepsilon^2$ ,

$$\begin{aligned} (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\log(\bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q})\bar{\mathbf{P}}^{-1} - \log(\mathbf{P}_i)) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\log(\bar{\mathbf{P}}) + \varepsilon H(\mathbf{Q})\bar{\mathbf{P}}^{-1} - \log(\mathbf{Q}_j)) &= 0 \\ \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\log(\bar{\mathbf{P}}) - \log(\mathbf{P}_i)) + (1 - \varepsilon)\varepsilon H(\mathbf{Q})\bar{\mathbf{P}}^{-1} + \varepsilon \frac{1}{n} \sum_{j=1}^n (\log(\bar{\mathbf{P}}) - \log(\mathbf{Q}_j)) + \varepsilon^2 H(\mathbf{Q})\bar{\mathbf{P}}^{-1} &= 0 \tag{37} \\ \Rightarrow \varepsilon H(\mathbf{Q})\bar{\mathbf{P}}^{-1} + \varepsilon \frac{1}{n} \sum_{j=1}^n (\log(\bar{\mathbf{P}}) - \log(\mathbf{Q}_j)) &= 0 \\ \Rightarrow H(\mathbf{Q}) &= \frac{1}{n} \sum_{j=1}^n (\log(\mathbf{Q}_j) - \log(\bar{\mathbf{P}}))\bar{\mathbf{P}} \end{aligned}$$

□

**Proposition 4.** The influence function of Hellinger mean related to the Hellinger distance, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,

$$\begin{aligned} H(\mathbf{Q}) &= \left\{ \frac{1}{n} \sum_{j=1}^n |\mathbf{Q}_j|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \bar{\mathbf{P}}^{-1} \right) \right\} \\ &\quad \times \left\{ \frac{1}{m} \sum_{i=1}^m |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \bar{\mathbf{P}}^{-1} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right)^2 \right\}^{-1} \end{aligned} \tag{38}$$

**Proof of Proposition 4.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( 1 - |\mathbf{P}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\mathbf{P} + \mathbf{P}_i}{2} \right|^{-1/2} \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( 1 - |\mathbf{P}|^{1/4} |\mathbf{Q}_j|^{1/4} \left| \frac{\mathbf{P} + \mathbf{Q}_j}{2} \right|^{-1/2} \right) \quad (39)$$

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

$$\begin{aligned} \tilde{\mathbf{P}} &= \arg \min_{\mathbf{P}} F(\mathbf{P}), \nabla F(\tilde{\mathbf{P}}) = 0 \\ \Rightarrow &-(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{4} |\tilde{\mathbf{P}}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \tilde{\mathbf{P}}^{-1} - \frac{1}{4} |\tilde{\mathbf{P}}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right) \\ &- \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{4} |\tilde{\mathbf{P}}|^{1/4} |\mathbf{Q}_j|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \tilde{\mathbf{P}}^{-1} - \frac{1}{4} |\tilde{\mathbf{P}}|^{1/4} |\mathbf{Q}_j|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \left( \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} \right) = 0 \quad (40) \\ \Rightarrow &(1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( |\mathbf{P}_i|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \tilde{\mathbf{P}}^{-1} - |\mathbf{P}_i|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right) \\ &+ \varepsilon \frac{1}{n} \sum_{j=1}^n \left( |\mathbf{Q}_j|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \tilde{\mathbf{P}}^{-1} - |\mathbf{Q}_j|^{1/4} \left| \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \left( \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} \right) = 0 \end{aligned}$$

and

$$\begin{aligned} \bar{\mathbf{P}} &= \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m \left( 1 - |\mathbf{P}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\mathbf{P} + \mathbf{P}_i}{2} \right|^{-1/2} \right) \\ \Rightarrow \nabla G(\bar{\mathbf{P}}) &= \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{4} |\bar{\mathbf{P}}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \bar{\mathbf{P}}^{-1} - \frac{1}{4} |\bar{\mathbf{P}}|^{1/4} |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right) = 0 \quad (41) \\ \Rightarrow &\frac{1}{m} \sum_{i=1}^m \left( |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \bar{\mathbf{P}}^{-1} - |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right) = 0 \end{aligned}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have

$$\begin{aligned} \left| \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} &= \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} - \frac{1}{4} \varepsilon H(\mathbf{Q}) \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \\ \tilde{\mathbf{P}}^{-1} &= \bar{\mathbf{P}}^{-1} - \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} \quad (42) \\ \left( \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} &= \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} - \frac{1}{2} \varepsilon H(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \end{aligned}$$

Substitute Equations (41) and (42) into Equation (40), and ignore the terms contain  $\varepsilon^2$ ,

$$\begin{aligned}
 & (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (|\mathbf{P}_i|^{1/4} (|\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{P}_i)|^{-1/2} - \frac{1}{4}\varepsilon H(\mathbf{Q}) |\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{P}_i)|^{-1/2} (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{P}_i))^{-1}) \\
 & \times (\bar{\mathbf{P}}^{-1} - \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} - (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{P}_i))^{-1} + \frac{1}{2}\varepsilon H(\mathbf{Q}) (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{P}_i))^{-2}) \\
 & + \varepsilon \frac{1}{n} \sum_{j=1}^n (|\mathbf{Q}_j|^{1/4} (|\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{Q}_j)|^{-1/2} - \frac{1}{4}\varepsilon H(\mathbf{Q}) |\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{Q}_j)|^{-1/2} (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{Q}_j))^{-1}) \\
 & \times (\bar{\mathbf{P}}^{-1} - \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} - (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{Q}_j))^{-1} + \frac{1}{2}\varepsilon H(\mathbf{Q}) (\frac{1}{2}(\bar{\mathbf{P}} + \mathbf{Q}_j))^{-2}) = 0 \\
 & \Rightarrow \varepsilon H(\mathbf{Q}) \frac{1}{m} \sum_{i=1}^m |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \bar{\mathbf{P}}^{-1} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right)^2 \\
 & = \varepsilon \frac{1}{n} \sum_{j=1}^n |\mathbf{Q}_j|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \bar{\mathbf{P}}^{-1} \right) \\
 & \Rightarrow H(\mathbf{Q}) = \left\{ \frac{1}{n} \sum_{j=1}^n |\mathbf{Q}_j|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right|^{-1/2} \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \bar{\mathbf{P}}^{-1} \right) \right\} \\
 & \times \left\{ \frac{1}{m} \sum_{i=1}^n |\mathbf{P}_i|^{1/4} \left| \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right|^{-1/2} \left( \bar{\mathbf{P}}^{-1} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} \right)^2 \right\}
 \end{aligned} \tag{43}$$

□

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

$$H(\mathbf{Q}) = \bar{\mathbf{P}} - \frac{1}{n} \sum_{j=1}^n \mathbf{Q}_j^{-1} \bar{\mathbf{P}}^2 \tag{44}$$

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

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \text{tr}(\mathbf{P}_i^{-1} \mathbf{P} - \log(\mathbf{P}_i^{-1} \mathbf{P}) - \mathbf{I}) + \varepsilon \frac{1}{n} \sum_{j=1}^n \text{tr}(\mathbf{Q}_j^{-1} \mathbf{P} - \log(\mathbf{Q}_j^{-1} \mathbf{P}) - \mathbf{I}) \tag{45}$$

As  $\tilde{\mathbf{P}}$  is the mean of  $m$  matrices and  $n$  outliers, and  $\bar{\mathbf{P}}$  is the mean of  $m$  matrices, then, we have,

$$\begin{aligned}
 & \tilde{\mathbf{P}} = \arg \min_{\mathbf{P}} F(\mathbf{P}), \nabla F(\tilde{\mathbf{P}}) = 0 \\
 & \Rightarrow (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\mathbf{P}_i^{-1} - \tilde{\mathbf{P}}^{-1}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\mathbf{Q}_j^{-1} - \tilde{\mathbf{P}}^{-1}) = 0
 \end{aligned} \tag{46}$$

and

$$\begin{aligned}
 & \bar{\mathbf{P}} = \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m \text{tr}(\mathbf{P}_i^{-1} \mathbf{P} - \log(\mathbf{P}_i^{-1} \mathbf{P}) - \mathbf{I}) \\
 & \Rightarrow \nabla G(\bar{\mathbf{P}}) = 0 \Rightarrow \frac{1}{m} \sum_{i=1}^m (\mathbf{P}_i^{-1} - \bar{\mathbf{P}}^{-1}) = 0
 \end{aligned} \tag{47}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have,

$$\tilde{\mathbf{P}}^{-1} = \bar{\mathbf{P}}^{-1} - \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} \tag{48}$$

Substitute Equations (47) and (48) into Equation (46), and ignore the terms contain  $\varepsilon^2$ ,

$$\begin{aligned}
 & (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m (\mathbf{P}_i^{-1} - \bar{\mathbf{P}}^{-1} + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2}) + \varepsilon \frac{1}{n} \sum_{j=1}^n (\mathbf{Q}_j^{-1} - \bar{\mathbf{P}}^{-1} + \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2}) = 0 \\
 \Rightarrow & (1 - \varepsilon) \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} + \varepsilon \frac{1}{n} \sum_{j=1}^n (\mathbf{Q}_j^{-1} - \bar{\mathbf{P}}^{-1}) = 0 \\
 \Rightarrow & H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} = \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{P}}^{-1} - \mathbf{Q}_j^{-1}) \\
 \Rightarrow & H(\mathbf{Q}) = \bar{\mathbf{P}} - \frac{1}{n} \sum_{j=1}^n \mathbf{Q}_j^{-1} \bar{\mathbf{P}}^2
 \end{aligned} \tag{49}$$

□

**Proposition 6.** *The influence function of Bhattacharyya mean related to the Bhattacharyya divergence, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,*

$$H(\mathbf{Q}) = \left\{ \frac{1}{n} \sum_{j=1}^n \left( \bar{\mathbf{P}}^{-1} - \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} \right) \right\} \left\{ \frac{1}{m} \sum_{i=1}^m \left( \bar{\mathbf{P}}^{-2} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \right) \right\}^{-1} \tag{50}$$

**Proof of Proposition 6.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 4 \left( \log \left| \frac{\mathbf{P} + \mathbf{P}_i}{2} \right| - \frac{1}{2} \log |\mathbf{P}| |\mathbf{P}_i| \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n 4 \left( \log \left| \frac{\mathbf{P} + \mathbf{Q}_j}{2} \right| - \frac{1}{2} \log |\mathbf{P}| |\mathbf{Q}_j| \right) \tag{51}$$

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

$$\begin{aligned}
 & \tilde{\mathbf{P}} = \arg \min_{\mathbf{P}} F(\mathbf{P}), \nabla F(\tilde{\mathbf{P}}) = 0 \\
 \Rightarrow & (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m 4 \left( \left( \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} - \tilde{\mathbf{P}}^{-1} \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n 4 \left( \left( \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \tilde{\mathbf{P}}^{-1} \right) = 0
 \end{aligned} \tag{52}$$

and

$$\begin{aligned}
 & \bar{\mathbf{P}} = \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m 4 \left( \log \left| \frac{\mathbf{P} + \mathbf{P}_i}{2} \right| - \frac{1}{2} \log |\mathbf{P}| |\mathbf{P}_i| \right) \\
 \Rightarrow & \frac{1}{m} \sum_{i=1}^m 4 \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} - \bar{\mathbf{P}}^{-1} \right) = 0
 \end{aligned} \tag{53}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have,

$$\begin{aligned}
 \left( \frac{\tilde{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} &= \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} - \frac{1}{2} \varepsilon H(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \\
 \left( \frac{\tilde{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} &= \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \frac{1}{2} \varepsilon H(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-2} \\
 \tilde{\mathbf{P}}^{-1} &= \bar{\mathbf{P}}^{-1} - \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2}
 \end{aligned} \tag{54}$$

Substitute Equations (53) and (54) into Equation (52), and ignore the terms contain  $\varepsilon^2$ ,

$$\begin{aligned}
 & (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-1} - \frac{1}{2} Hz(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} - \bar{\mathbf{P}}^{-1} + Hz(\mathbf{Q}) \bar{\mathbf{P}}^{-2} \right) \\
 & + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \frac{1}{2} Hz(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-2} - \bar{\mathbf{P}}^{-1} + Hz(\mathbf{Q}) \bar{\mathbf{P}}^{-2} \right) = 0 \\
 \Rightarrow & \frac{1}{m} \sum_{i=1}^m \left( \varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-2} - \frac{1}{2} \varepsilon H(\mathbf{Q}) \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} - \bar{\mathbf{P}}^{-1} \right) = 0 \tag{55} \\
 \Rightarrow & H(\mathbf{Q}) \frac{1}{m} \sum_{i=1}^m \left( \bar{\mathbf{P}}^{-2} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \right) = \frac{1}{n} \sum_{j=1}^n \left( \bar{\mathbf{P}}^{-1} - \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} \right) \\
 \Rightarrow & H(\mathbf{Q}) = \left\{ \frac{1}{n} \sum_{j=1}^n \left( \bar{\mathbf{P}}^{-1} - \left( \frac{\bar{\mathbf{P}} + \mathbf{Q}_j}{2} \right)^{-1} \right) \right\} \left\{ \frac{1}{m} \sum_{i=1}^m \left( \bar{\mathbf{P}}^{-2} - \frac{1}{2} \left( \frac{\bar{\mathbf{P}} + \mathbf{P}_i}{2} \right)^{-2} \right) \right\}^{-1}
 \end{aligned}$$

□

**Proposition 7.** The influence function of SKL mean related to the SKL divergence, of  $m$  HPD matrices  $\{\mathbf{P}_1, \dots, \mathbf{P}_m\}$  and  $n$  outliers  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  is given by,

$$H(\mathbf{Q}) = \left( \frac{2}{m} \sum_{i=1}^m \mathbf{P}_i \right)^{-1} \left( \frac{1}{n} \sum_{j=1}^n (\bar{\mathbf{P}} \mathbf{Q}_j - \mathbf{Q}_j^{-1} \bar{\mathbf{P}}^3) \right) \tag{56}$$

**Proof of Proposition 7.** Let  $F(\mathbf{P})$  be the objection function,

$$F(\mathbf{P}) = (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m tr \left( \mathbf{P}_i^{-1} \mathbf{P} + \mathbf{P}^{-1} \mathbf{P}_i - 2\mathbf{I} \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n tr \left( \mathbf{Q}_j^{-1} \mathbf{P} + \mathbf{P}^{-1} \mathbf{Q}_j - 2\mathbf{I} \right) \tag{57}$$

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

$$\begin{aligned}
 \tilde{\mathbf{P}} &= \arg \min_{\mathbf{P}} F(\mathbf{P}), \nabla F(\tilde{\mathbf{P}}) = 0 \\
 \Rightarrow & (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( \mathbf{P}_i^{-1} - \tilde{\mathbf{P}}^{-2} \mathbf{P}_i \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \mathbf{Q}_j^{-1} - \tilde{\mathbf{P}}^{-2} \mathbf{Q}_j \right) = 0 \tag{58}
 \end{aligned}$$

and

$$\begin{aligned}
 \bar{\mathbf{P}} &= \arg \min_{\mathbf{P}} G(\mathbf{P}), G(\mathbf{P}) = \frac{1}{m} \sum_{i=1}^m tr \left( \mathbf{P}_i^{-1} \mathbf{P} + \mathbf{P}^{-1} \mathbf{P}_i - 2\mathbf{I} \right) \\
 \Rightarrow & \frac{1}{m} \sum_{i=1}^m \left( \mathbf{P}_i^{-1} - \bar{\mathbf{P}}^{-2} \mathbf{P}_i \right) = 0 \tag{59}
 \end{aligned}$$

Using the Taylor expansion on  $\tilde{\mathbf{P}} = \bar{\mathbf{P}} + \varepsilon H(\mathbf{Q})$ , and we have,

$$\tilde{\mathbf{P}}^{-2} = \bar{\mathbf{P}}^{-2} - 2\varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-3} \tag{60}$$

Substitute Equations (59) and (60) into Equation (58), and ignore the terms contain  $\varepsilon^2$ ,

$$\begin{aligned} (1 - \varepsilon) \frac{1}{m} \sum_{i=1}^m \left( \mathbf{P}_i^{-1} - \left( \bar{\mathbf{P}}^{-2} - 2\varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-3} \right) \mathbf{P}_i \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \mathbf{Q}_j^{-1} - \left( \bar{\mathbf{P}}^{-2} - 2\varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-3} \right) \mathbf{Q}_j \right) &= 0 \\ \Rightarrow \frac{1}{m} \sum_{i=1}^m \left( 2\varepsilon H(\mathbf{Q}) \bar{\mathbf{P}}^{-3} \mathbf{P}_i \right) + \varepsilon \frac{1}{n} \sum_{j=1}^n \left( \mathbf{Q}_j^{-1} - \bar{\mathbf{P}}^{-2} \mathbf{Q}_j \right) &= 0 \\ \Rightarrow H(\mathbf{Q}) = \left( \frac{2}{m} \sum_{i=1}^m \mathbf{P}_i \right)^{-1} \left( \frac{1}{n} \sum_{j=1}^n \left( \bar{\mathbf{P}} \mathbf{Q}_j - \mathbf{Q}_j^{-1} \bar{\mathbf{P}}^3 \right) \right) \end{aligned} \tag{61}$$

□

### 5. Numerical Simulations

In order to gain a better understanding of the superiority of proposed estimators, simulation results of the performance of an ANMF with the proposed estimator in heterogeneous clutter are presented. As there is not an analytical expression for the detection threshold, the standard Monte Carlo technique [34] is utilized. A similar approach was recently used to solve several problems from different areas, such as physics [35], decision theory [36], engineering [37], computational geometry [38], finance [39], etc. The rule of adaptive normalized matched filter (ANMF) is given as [40],

$$\frac{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{r}}{(\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s})(\mathbf{r}^H \hat{\Sigma}^{-1} \mathbf{r})} \stackrel{H_1}{\underset{H_0}{\gtrless}} \gamma \tag{62}$$

where  $\hat{\Sigma}$  is the clutter covariance estimation.  $\mathbf{r}$  is the sample data in the cell under test.  $\gamma$  denotes the threshold, which is derived by Monte Carlo method in order to maintain the false alarm constant.  $\mathbf{s}$  is the target steering vector, and is given by,

$$\mathbf{s} = \frac{1}{\sqrt{N}} [1, \exp(j2\pi f_d), \dots, \exp(j2\pi(N-1)f_d)]^T \tag{63}$$

where  $f_d$  is the normalized Doppler frequency. According to Equation (2), the terms  $\mathbf{r}_k, k = 1, \dots, K$  are compound-Gaussian random vectors, and sharing the same covariance matrix  $\Sigma$ ,

$$\Sigma = \Sigma_0 + \mathbf{I} \tag{64}$$

where  $\mathbf{I}$  is accounting for the thermal noise.  $\Sigma_0$  is related to the clutter, modeled as,

$$\Sigma_0(i, k) = \sigma_c^2 \rho^{|i-k|} e^{j2\pi f_{dc}(i-k)}, i, k = 1, \dots, N \tag{65}$$

where  $\rho$  is the one-lag correlation coefficient.  $\sigma_c$  is the clutter-to-noise power ratio.  $f_{dc}$  is the clutter normalized Doppler frequency.

In addition,  $\tau$  and  $\tau_k$  are positive and real independent and identical distributed random variables, and are assumed to follow the inverse gamma distribution,

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), x \geq 0 \tag{66}$$

where  $\alpha$  and  $\beta$  denote the shape and scale parameters, respectively.  $\Gamma(\cdot)$  is the gamma function. In the simulation, we set  $\rho = 0.9, f_{dc} = 0.1$ , and  $\sigma_c^2 = 25$  dB. The parameters  $\alpha = 3$ , and  $\beta = 1$ .

In the following, we analyze the performance of an ANMF with the proposed estimators, in terms of detection probability ( $P_d$ ), also in comparison with the optimum detector, which assumes the perfect knowledge of the disturbance covariance matrix, NMF. In particular, the positive-definite

matrix obtained using (1), (2), and (3) is noted as the SPDF, THPD, and SPD matrix, respectively. Simulation results are shown in Figure 2 with  $N = 8$ ,  $P_{fa} = 10^{-4}$ , it is clear that the proposed estimators have different performances, and all the proposed estimators have better performances than the NSCM estimator when  $K = 10$  and  $K = 16$ . In particular, our proposed estimators have different performance when different positive-definite matrix is utilized. For the SPDF positive-definite matrix, the Bhat estimator, the Hel estimator, and the KL estimator have comparable performances, and outperform others. The SKL estimator has the worst performance, while the KL estimator has the best performance. However, this relationship is different on the condition of the THPD positive-definite matrix. Particularly, performances of the KL estimator and the SKL estimator are poor. Performances of the other proposed estimators are close to the optimum. For the SPD positive-definite matrix, relationships of proposed estimators are similar to the case of SPDF positive-definite matrix. The KL estimator has the best performance, while the performance of SKL estimator is the worst. In addition, the performance of Hel estimator is poor. These results imply that performances of proposed estimators are related to the used positive-definite matrix.

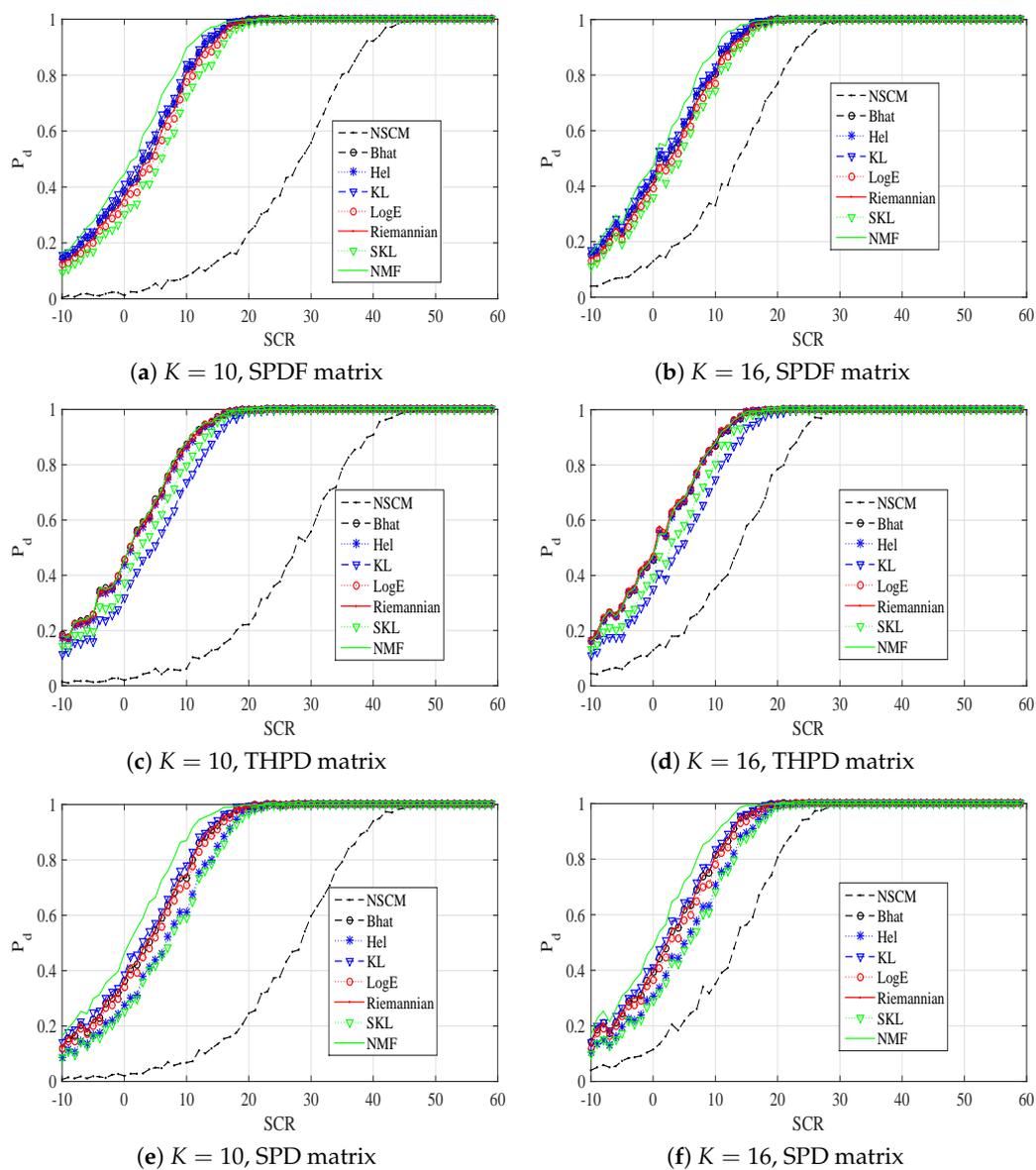


Figure 2.  $P_d$  versus SCR plots of ANMFs with proposed estimators, the NSCM estimator, and NMF.

In order to show the influence function of robustness of proposed estimators, 17 positive-definite matrices with an injected outlier are considered. The value of influence function is computed as Propositions 1–7, and 100 times simulations are repeated. A total of 100 simulation results and the average of values of influence function are shown in Figure 3. From Figure 3 we can know that our proposed estimators are more robust than the NSCM estimator. In particular, the robustness of SKL estimator is poor, while the KL estimator has the best robustness when the SPDF or THPD positive-definite matrix is utilized. For the SPD positive-definite matrix, all proposed estimators have comparable robustness. It can be concluded that the robustness of proposed estimators is related to the used positive-definite matrix. It is worth pointing out that the three HPD matrices, namely the SPDF, THPD, and SPD matrix, have different structures. Both the SPDF and the SPD matrices have a largest eigenvalue and  $N - 1$  equal eigenvalues. Their differences lie in the multiple between the maximum eigenvalue and the minimum eigenvalue. In particular, this multiple of the SPD matrix is the constant number  $\kappa_M$ , while the SPDF matrix has a varied multiple. For the THPD matrix, all eigenvalues are different. Riemannian manifolds composed of different positive-definite matrices have different geometric structures. Thus, the estimators associated with different metrics on Riemannian manifold may have different behaviors.

Figure 4 plots the  $P_d$  of the ANMF with our proposed estimators, the NSCM estimator, and the NMF detector in a contaminated clutter. An outlier is injected in one reference cell, the number of reference cell  $K$  is set to 10 and 16, respectively. The dimension of the sample data is 8. It can be noted from Figure 4 that performances of our proposed estimators have not been significantly reduced, while there is a degradation in the performance of the NSCM estimator. Relationships of performances are similar to curves of Figure 2. These results prove the advantage of our proposed estimators sufficiently.

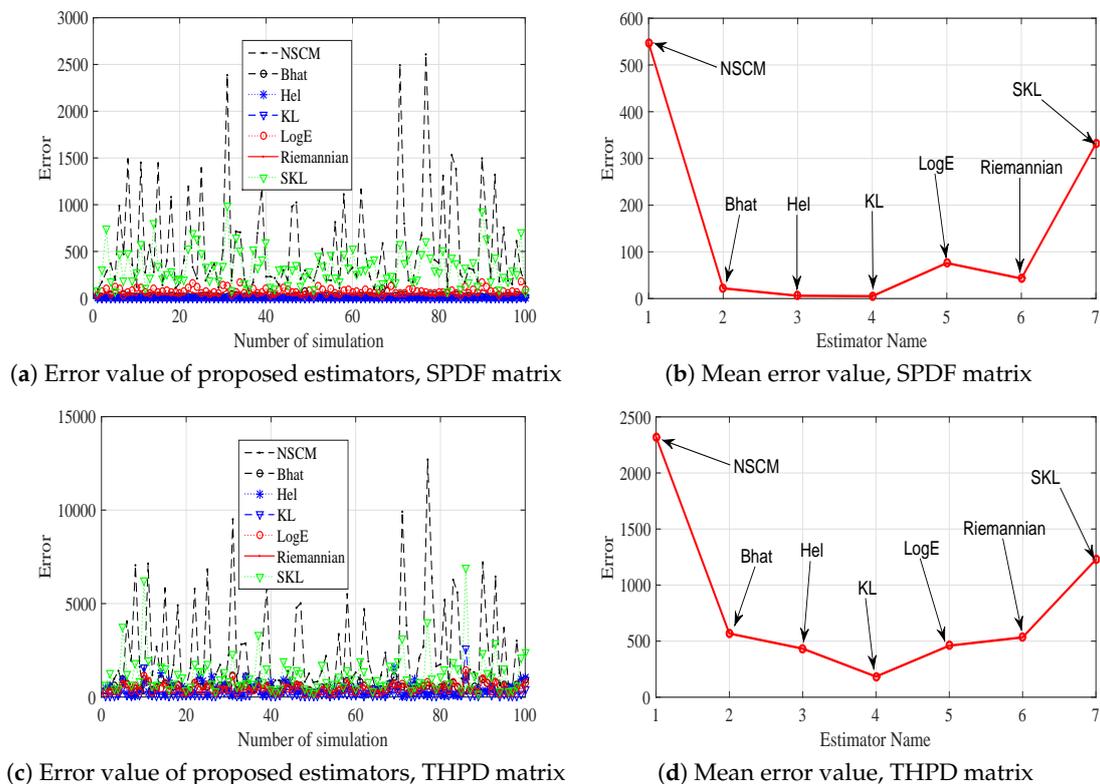


Figure 3. Cont.

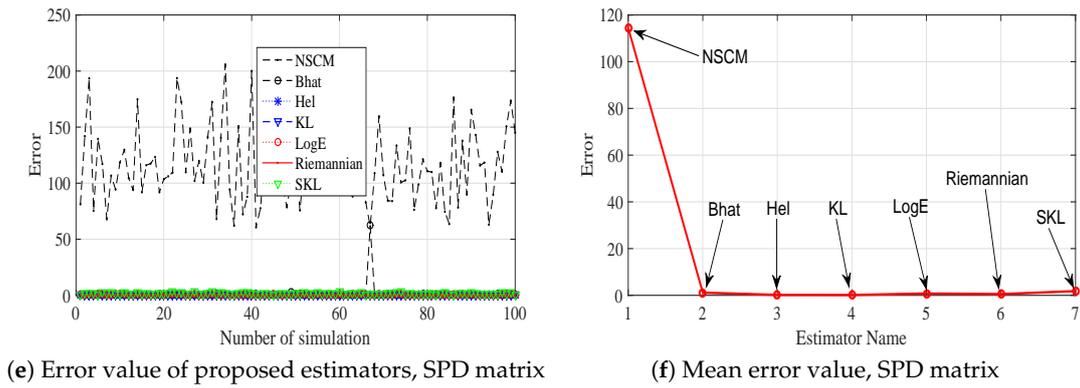


Figure 3. The error value of proposed estimators and their corresponding mean vlaue.

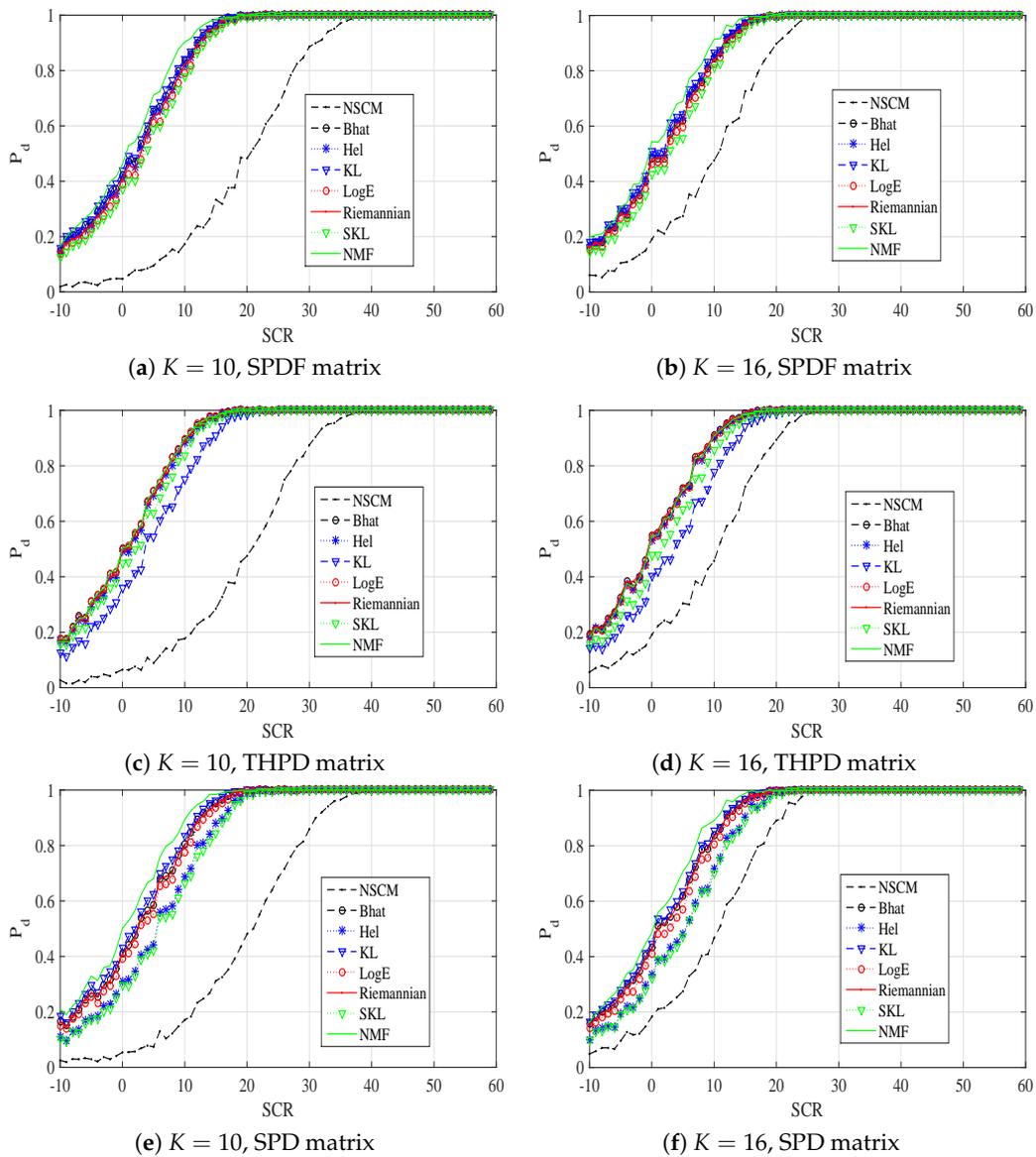


Figure 4.  $P_d$  versus SCR plots of ANMFs with proposed estimators, the NSCM estimator, and NMF in a contaminated environment.

## 6. Conclusions

In this paper, a class of covariance estimators based on information divergences is proposed in heterogeneous clutter. Particularly, the problem of disturbance covariance estimation is reformulated as obtaining the geometric mean on the Riemannian manifold. Six mean estimators related to information measures are derived. Moreover, the robustness of proposed estimators are analyzed via the influence function, and the analytic expression of influence function is deduced. At the analysis stage, the performance advantage and robustness of our proposed estimators are verified by means of simulation results in heterogeneous environment.

**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. All authors have read and approved the final manuscript.

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

## References

1. Visuri, S.; Oja, H.; Koivunen, V. 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.
2. Williams, D.B.; Johnson, D.H. Robust estimation of structured covariance matrices. *IEEE Trans. Signal Process.* **1993**, *41*, 2891–2906.
3. 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.
4. 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.
5. Maio, A.D.; Farina, A.; Foglia, G. Design and experimental validation of knowledge-based constant false alarm rate detectors. *IET Radar Sonar Navig.* **2007**, *1*, 308–316.
6. Maio, A.D.; Farina, A.; Foglia, G. Knowledge-Aided Bayesian Radar Detectors Their Application to Live Data. *IEEE Trans. Aerosp. Electron. Syst.* **2010**, *46*, 170–183.
7. Wang, P.; Li, H.; Himed, B. Knowledge-Aided Parametric Tests for Multichannel Adaptive Signal Detection. *IEEE Trans. Signal Process.* **2011**, *59*, 5970–5982.
8. Li, H.; Michels, J.H. Parametric adaptive signal detection for hyperspectral imaging. *IEEE Trans. Signal Process.* **2006**, *54*, 2704–2715.
9. Aubry, A.; Maio, A.D.; Pallotta, L.; Farina, A. Maximum Likelihood Estimation of a Structured Covariance Matrix with a Condition Number Constraint. *IEEE Trans. Signal Process.* **2012**, *60*, 3004–3021.
10. Maio, A.D.; Orlando, D.; Foglia, G.; Foglia, G. Adaptive Detection of Point-Like Targets in Spectrally Symmetric Interference. *IEEE Trans. Signal Process.* **2016**, *64*, 3207–3220.
11. Liu, Z.; Barbaresco, F. *Doppler Information Geometry for Wake Turbulence Monitoring*; Springer: Berlin/Heidelberg, Germany, 2013.
12. Barbaresco, F.; Meier, U. Radar monitoring of a wake vortex: Electromagnetic reflection of wake turbulence in clear air. *C. R. Phys.* **2010**, *11*, 54–67.
13. 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.

14. Balaji, B.; Barbaresco, F. Application of Riemannian mean of covariance matrices to space-time adaptive processing. In Proceedings of the IEEE Radar Conference, Amsterdam, The Netherlands, 31 October–2 November 2012; pp. 50–53.
15. Balaji, B. Riemannian mean and space-time adaptive processing using projection and inversion algorithms. *Proc. SPIE* **2013**, *8714*, 813–831.
16. Barbaresco, F. Robust statistical Radar Processing in Frechet metric space: OS-HDR-CFAR and OS-STAP Processing in Siegel homogeneous bounded domains. In Proceedings of the International Radar Symposium (IRS), Leipzig, Germany, 7–9 September 2011; pp. 639–644.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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*, doi:10.1016/j.ast.2017.07.042.
22. 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 Geometric Science of Information: First International Conference, GSI 2013, Paris, France, 28–30 August 2013; pp. 551–558.
23. 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.
24. Jayasumana, S.; Hartley, R.; Salzmann, M.; Li, H.; Harandi, M. Kernel Methods on the Riemannian Manifold of Symmetric Positive Definite Matrices. *Comput. Vis. Pattern Recog.* **2013**, 73–80.
25. Hiai, F.; Petz, D. Riemannian metrics on positive definite matrices related to means. *Linear Algebra Appl.* **2008**, *430*, 3105–3130.
26. Sra, S. Positive definite matrices and the S divergence. *Proc. Am. Math. Soc.* **2011**, *144*, 1–25.
27. Moakher, M. On the Averaging of Symmetric Positive Definite Tensors. *J. Elast.* **2006**, *82*, 273–296.
28. Lang, S. Fundamentals of differential geometry. In *Graduate Texts in Mathematics*; Springer: New York, NY, USA, 1946; pp. 1757–1764.
29. Moakher, M. A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices. *Siam J. Matrix Anal. Appl.* **2008**, *26*, 735–747.
30. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **2006**, *56*, 411–421.
31. Lindsay, B.G. Efficiency Versus Robustness: The Case for Minimum Hellinger Distance and Related Methods. *Ann. Stat.* **1994**, *22*, 1081–1114.
32. Csiszar, I. Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems. *Ann. Stat.* **1991**, *19*, 2032–2066.
33. Menendez, M.L.; Pardo, J.A.; Pardo, L.; Pardo, M.C. The Jensen-Shannon divergence. *J. Frankl. Inst.* **1997**, *334*, 307–318.
34. Fishman, G.S. Monte Carlo. Concepts, algorithms, and applications. *Technometrics* **1996**, *39*, 338.
35. Klein, S.R.; Nystrand, J.; Seger, J.; Gorbunov, Y.; Butterworth, J. STARlight: A Monte Carlo simulation program for ultra peripheral collisions of relativistic ions. *Comput. Phys. Commun.* **2016**, *212*, 258–268.
36. Pask, F.; Lake, P.; Yang, A.; Tokos, H.; Sadhukhan, J. Sustainability indicators for industrial ovens and assessment using Fuzzy set theory and Monte Carlo simulation. *J. Clean. Prod.* **2016**, *140*, 1217–1225.
37. Ma, Y.; Chen, X.; Biegler, L.T. Monte Carlo simulation-based optimization for copolymerization processes with embedded chemical composition distribution. *Comput. Chem. Eng.* **2018**, *109*, 261–275.
38. Vilcu, A.D.; Vilcu, G.E. An algorithm to estimate the vertices of a tetrahedron from uniform random points inside. *Ann. Mat. Pura Appl.* **2016**, *4*, 1–14.

39. Bormetti, G.; Callegaro, G.; Livieri, G.; Pallavicini, A. A backward Monte Carlo approach to exotic option pricing. *Eur. J. Appl. Math.* **2018**, *29*, 146–187.
40. Conte, E.; Lops, M.; Ricci, G. Adaptive detection schemes in compound-Gaussian clutter. *IEEE Trans. Aerosp. Electron. Syst.* **1998**, *34*, 1058–1069.



© 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/>).