# Kernel Density Estimation on the Siegel Space with an Application to Radar Processing [†]

**Emmanuel Chevallier [1,\*], Thibault Forget [2,3], Frédéric Barbaresco [2] and Jesus Angulo [3]**

[1]   Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel
[2]   Thales Air Systems, Surface Radar Business Line, Advanced Radar Concepts Business Unit, Voie Pierre-Gilles de Gennes, Limours 91470, France; thibault.forget@mines-paristech.fr (T.F.); frederic.barbaresco@thalesgroup.com (F.B.)
[3]   CMM-Centre de Morphologie Mathématique, MINES ParisTech, PSL-Research University, Paris 75006, France; jesus.angulo@mines-paristech.fr
**\***   Correspondence: emmanuelchevallier1@gmail.com; Tel.: +972-58-693-7744
†   This paper is an extended version of our paper published in the 2nd conference on Geometric Science of Information, Paris, France, 28–30 October 2015.

**Abstract:** This paper studies probability density estimation on the Siegel space. The Siegel space is a generalization of the hyperbolic space. Its Riemannian metric provides an interesting structure to the Toeplitz block Toeplitz matrices that appear in the covariance estimation of radar signals. The main techniques of probability density estimation on Riemannian manifolds are reviewed. For computational reasons, we chose to focus on the kernel density estimation. The main result of the paper is the expression of Pelletier's kernel density estimator. The computation of the kernels is made possible by the symmetric structure of the Siegel space. The method is applied to density estimation of reflection coefficients from radar observations.

**Keywords:** kernel density estimation; Siegel space; symmetric spaces; radar signals

## 1. Introduction

Various techniques can be used to estimate the density of probability measure in the Euclidean spaces, such as histograms, kernel methods, or orthogonal series. These methods can sometimes be adapted to densities in Riemannian manifolds. The computational cost of the density estimation depends on the isometry group of the manifold. In this paper, we study the special case of the Siegel space. The Siegel space is a generalization of the hyperbolic space. It has a structure of symmetric Riemannian manifold, which enables the adaptation of different density estimation methods at a reasonable cost. Convergence rates of the density estimation using kernels and orthogonal series were gradually generalized to Riemannian manifolds (see [1–3]).

The Siegel space appears in radar processing in the study of Toeplitz block Toeplitz matrices, whose blocks represent covariance matrices of a radar signal (see [4–6]). The Siegel also appears in statistical mechanics, see [7] and was recently used in image processing (see [8]). Information geometry is now a standard framework in radar processing (see [4–6,9–13]). The information geometry on positive definite Teoplitz block Teoplitz matrices is directly related to the metric on the Siegel space (see [14]). Indeed, Toeplitz block Toeplitz matrices can be represented by a symmetric positive definite matrix and a point laying in a product of Siegel disks. The metric considered on Toeplitz block Toeplitz matrices is induced by the product metric between a metric on the symmetric positive definite matrices and the Siegel disks metrics (see [4–6,9,14]).

One already encounters the problem of density estimation in the hyperbolic space for electrical impedance [15], networks [16] and radar signals [17]. In [18], a generalization of the Gaussian law on the hyperbolic space was proposed. Apart from [19], where authors propose a generalization of the Gaussian law, probability density estimation on the Siegel space has not yet been addressed.

The contributions of the paper are the following. We review the main non parametric density estimation techniques on the Siegel disk. We provide some rather simple explicit expressions of the kernels defined by Pelletier in [1]. These expressions make the kernel density estimation the most adapted method. We present visual results of estimated densities in the simple case where the Siegel disk reduces to the Poincaré disk.

The paper begins with an introduction to the Siegel space in Section 2. Section 3 reviews the main non-parametric density estimation techniques on the Siegel space. Section 3.3 contains the original results of the paper. Section 4 presents an application to radar data estimation.

## 2. The Siegel Space

This section presents facts about the Siegel space. The interested reader can find more details in [20,21]. The necessary background on Lie groups and symmetric space can be found in [22].

### 2.1. The Siegel Upper Half Space

The Siegel upper half space is a generalization of the Poincaré upper half space (see [23]) for a description of the hyperbolic space. Let $Sym(n)$ be the space of real symmetric matrices of size $n \times n$ and $Sym_+(n)$ the set of real symmetric positive definite matrices of size $n \times n$. The Siegel upper half space is defined by

$$\mathcal{H}_n = \{Z = X + iY | X \in Sym(n), Y \in Sym_+(n)\}.$$

$\mathcal{H}_n$ is equipped with the following metric:

$$ds = 2tr(Y^{-1}dZY^{-1}d\overline{Z}).$$

The set of real symplectic matrices $Sp(n, \mathbb{R})$ is defined by

$$g \in Sp(n, \mathbb{R}) \Leftrightarrow g^t J g = J,$$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

and $I_n$ is the $n \times n$ identity matrix. $Sp(n, \mathbb{R})$ is a subgroup of $SL_{2n}(\mathbb{R})$, the set of $2n \times 2n$ invertible matrices of determinant 1. Let $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, \mathbb{R})$. The metric $ds$ is invariant under the following action of $Sp(n, \mathbb{R})$,

$$g.Z = (AZ + B)(CZ + D)^{-1}.$$

This action is transitive, i.e.,

$$\forall Z \in \mathcal{H}_n, \exists g \in Sp(n, \mathbb{R}), g.iI = Z.$$

The stabilizer $K$ of $iI$ is the set of elements $g$ of $Sp(n, \mathbb{R})$ whose action leaves $iI$ fixed. $K$ is a subgroup of $Sp(n, \mathbb{R})$ called the isotropy group. We can verify that

$$K = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix}, A + iB \in SU(n) \right\}.$$

A symmetric space is a Riemannian manifold, where the reversal of the geodesics is well defined and is an isometry. Formally, $exp_p(u) \mapsto exp_p(-u)$ is an isometry for each $p$ on the manifold, where $u$ is a vector in the tangent space at $p$, and $exp_p$ the Riemannian exponential application at $p$. In other words, the symmetry around each point is an isometry. $\mathcal{H}_n$ is a symmetric space (see [20]). The structure of a symmetric space can be studied through its isometry group and the Lie algebra of its isometry group. The present work will make use of the Cartan and Iwasawa decompositions of the Lie algebra of $Sp(n, \mathbb{R})$ (see [22]). Let $\mathfrak{sp}(n, \mathbb{R})$ be the Lie algebra of $Sp(n, \mathbb{R})$. Given $A$, $B$ and $C$ three real $n \times n$ matrices, let denote $\begin{pmatrix} A & B \\ C & -A^t \end{pmatrix} = (A, B, C)$. We have

$$\mathfrak{sp}(n, \mathbb{R}) = \{(A, B, C) | B \text{ and } C \text{ symmetric}\}.$$

The Cartan decomposition of $\mathfrak{sp}(n, \mathbb{R})$ is given by

$$\mathfrak{sp}(n, \mathbb{R}) = \mathfrak{t} \oplus \mathfrak{p},$$

where

$$\mathfrak{t} = \{(A, B, -B) | B \text{ symmetric and } A \text{ skew-symmetric}\},$$

$$\mathfrak{p} = \{(A, B, B) | A, B, \text{ symmetric}\}. \tag{1}$$

The Iwasawa decomposition is given by

$$\mathfrak{sp}(n, \mathbb{R}) = \mathfrak{t} \oplus \mathfrak{a} \oplus \mathfrak{n},$$

where

$$\mathfrak{a} = \{(H, 0, 0) | H \text{ diagonal}\},$$

$$\mathfrak{n} = \{(A, B, 0) | A \text{ upper triangular with } 0 \text{ on the diagonal}, B \text{ symmetric}\}.$$

It can be shown that

$$\mathfrak{p} = \cup_{k \in K} Ad_k(\mathfrak{a}), \tag{2}$$

where $Ad$ is the adjoint representation of $Sp(n, \mathbb{R})$.

*2.2. The Siegel Disk*

The Siegel disk $\mathcal{D}_n$ is the set of complex matrices $\{Z | I - Z^*Z \geq 0\}$, where $\geq$ stands for the Loewner order (see [24] for details on the Loewner order). Recall that for $A$ and $B$ two Hermitian matrices, $A \geq B$ with respect to the Loewner order means that $A - B$ is positive definite. The transformation

$$Z \in \mathcal{H}_n \mapsto (Z - iI)(Z + iI)^{-1} \in \mathcal{D}_n$$

is an isometry between the Siegel upper half space and the Siegel disk. Let $C = \begin{pmatrix} I & -iI \\ I & iI \end{pmatrix}$. The application $g \in Sp(n, \mathbb{R}) \mapsto CgC^{-1}$ identifies the set of isometries of $\mathcal{H}_n$ and of $\mathcal{D}_n$. Thus, it can be shown that a matrix $g = \begin{pmatrix} A & B \\ \overline{A} & \overline{B} \end{pmatrix} \in Sp(n, \mathbb{C})$ acts isometrically on $\mathcal{D}_n$ by

$$g.Z = (AZ + B)(\overline{A}Z + \overline{B})^{-1},$$

where $\overline{A}$ stands for the conjugate of $A$. The point $iI$ in $\mathcal{H}_n$ is identified with the null matrix noted 0 in $\mathcal{D}_n$. Let $Z \in \mathcal{D}_n$. There exists $P$ a diagonal matrix with decreasing positive real entries and $U$ a unitary matrix such that $Z = UPU^t$. Let $\tau_1 \geq ... \geq \tau_n$ be such that

$$P = \begin{pmatrix} th(\tau_1) & & \\ & \ddots & \\ & & th(\tau_n) \end{pmatrix}.$$

Let

$$A_0 = \begin{pmatrix} ch(\tau_1) & & \\ & \ddots & \\ & & ch(\tau_n) \end{pmatrix}, B_0 = \begin{pmatrix} sh(\tau_1) & & \\ & \ddots & \\ & & sh(\tau_n) \end{pmatrix}$$

and

$$g_Z = \begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix} \cdot \begin{pmatrix} A_0 & B_0 \\ A_0 & B_0 \end{pmatrix}.$$

It can be shown that

$$g_Z \in Sp(n, \mathbb{C}) \text{ and } g_Z.0 = Z. \tag{3}$$

We provide now a correspondence between the elements of $\mathcal{D}_n$ and the elements of $\mathfrak{p}$ defined in Equation (1). Let

$$H_Z = \begin{pmatrix} \tau_1 & & & & & \\ & \ddots & & & & \\ & & \tau_n & & & \\ & & & -\tau_1 & & \\ & & & & \ddots & \\ & & & & & -\tau_n \end{pmatrix} \in \mathfrak{a}, \tag{4}$$

and

$$a_Z = \begin{pmatrix} e^{\tau_1} & & & & & \\ & \ddots & & & & \\ & & e^{\tau_n} & & & \\ & & & e^{-\tau_1} & & \\ & & & & \ddots & \\ & & & & & e^{-\tau_n} \end{pmatrix} \in A = exp(\mathfrak{a}).$$

It can be shown that there exists $k \in K$ such that

$$C exp(Ad_k(H_Z)) C^{-1}.0 = Z,$$

or equivalently

$$C k a_Z k C^{-1}.0 = Z.$$

Recall that Equation (2) gives $Ad_k(H) \in \mathfrak{p}$ and $kak \in exp(\mathfrak{p})$. The distance between $Z$ and 0 in $\mathcal{D}_n$ is given by

$$d(0, Z) = \left(2 \sum \tau_i^2\right)^{1/2} \tag{5}$$

### 3. Non Parametric Density Estimation on the Siegel Space

Let $\Omega$ be a space, endowed with a $\sigma$-algebra and a probability measure $p$. Let $X$ be a random variable $\Omega \to \mathcal{D}_n$. The Riemannian measure of $\mathcal{D}_n$ is called *vol* and the measure on $\mathcal{D}_n$ induced by $X$ is noted $\mu_X$. We assume that $\mu_X$ has a density, noted $f$, with respect to *vol*, and that the support of $X$ is a compact set noted *Supp*. Let $(x_1, ..., x_k) \in \mathcal{D}_n^k$ be a set of draws of $X$.

The Dirac measure at a point $a \in \mathcal{D}_n$ is denoted $\delta_a$. Let $\mu_k = \frac{1}{k} \sum_{i=1}^k \delta_{x_i}$ denotes the empirical measure of the set of draws. This section presents four non-parametric techniques of estimation of the density $f$ from the set of draws $(x_1, ..., x_k)$. The estimated density at $x$ in $\mathcal{D}_n$ is noted $\hat{f}_k(x) = \hat{f}(x, x_1, ..., x_k)$. The relevance of a density estimation technique depends on several aspects. When the space allows it, the estimation technique should equally consider each direction and location. This leads to an isotropy and a homogeneity condition. In the kernel method, for instance, a kernel density function $K_{x_i}$ is placed at each observation $x_i$. Firstly, in order to treat directions equally, the function $K_{x_i}$ should be invariant under the isotropy group of $x_i$; Secondly, for another observation $x_j$, functions $K_{x_i}$ and $K_{x_j}$ should be similar up to the isometries that send $x_i$ on $x_j$. These considerations strongly depend on the geometry of the space: if the space is not homogeneous and the isotropy group is empty, these indifference principles have no meaning. Since the Siegel space is symmetric, it is homogeneous and has a non empty isotropy group. Thus, the density estimation technique should be chosen accordingly.

The convergence of the different estimation techniques is widely studied. Results were first obtained in the Euclidean case, and are gradually extended to the probability densities on manifold (see [1,2,15,25]).

The last relevant aspect is computational. Each estimation technique has its own computational framework that presents pros and cons given the different applications. For instance, the estimation by orthogonal series needs an initial pre-processing, but provides a fast evaluation of the estimated density in compact manifolds.

#### 3.1. Histograms

The histogram is the simplest density estimation method. Given a partition of the space $\mathcal{D}_n = \cup_i A_i$, the estimated density is given by

$$\hat{f}(x \in A_i) = \frac{1}{vol(A_i)} \sum_{j=1}^k 1_{A_i}(x_j),$$

where $1_{A_i}$ stands for the indicator function of $A_i$. Following the considerations of the previous sections, the elements of the partition should firstly be as isotropic as possible, and secondly as similar as possible to each other. Regarding the problem of histograms, the case of the Siegel space is similar to the case of the hyperbolic space. There exist various uniform polygonal tilings on the Siegel space that could be used to compute histograms. However, there are ratio $\lambda \in \mathbb{R}$ for which there is no homothety. Thus, it is not always possible to adapt the size of the bins to a given set of draws of the random variable. Modifying the size of the bins can require a change of the structure of the tiling. This is why the study of histograms has not been deepened.

#### 3.2. Orthogonal Series

The estimation of the density $f$ can be made out of the estimation of the scalar product between $f$ and a set of "orthonormal" functions $\{e_j\}$. The most standard choice for $\{e_j\}$ is the eigenfunctions of the Laplacian. When the variable $X$ takes its values in $\mathbb{R}^n$, this estimation technique becomes the characteristic function method. When the underlying space is compact, the spectrum of the Laplacian operator is countable, while when the space is non-compact, the spectrum is uncountable. In the first case, the estimation of the density $f$ is made through the estimation of a sum, while in the second case is made through the estimation of an integral. In practice, the second situation presents a larger

computational complexity. Unfortunately, eigenfunctions of the Laplacian operator are known on $\mathcal{D}_n$ but not on compact sub-domains. This is why the study of this method has not been deepened.

*3.3. Kernels*

Let $\mathcal{K} : \mathbb{R}_+ \to \mathbb{R}_+$ be a map which verifies the following properties:

(i)    $\int_{\mathbb{R}^d} \mathcal{K}(||x||)dx = 1$;
(ii)   $\int_{\mathbb{R}^d} x\mathcal{K}(||x||)dx = 0$;
(iii)  $\mathcal{K}(x > 1) = 0$;
(iv)  $sup(\mathcal{K}(x)) = \mathcal{K}(0)$.

Let $p \in \mathcal{D}_n$. Generally, given a point $p$ on a Riemannian manifold, $exp_p$ defines an injective application only on a neighborhood of 0. On the Siegel space, $exp_p$ is injective on the whole space. When the tangent space $T_p\mathcal{D}_n$ is endowed with the local scalar product,

$$||u|| = d(p, exp_p(u)),$$

where $||.||$ is the Euclidean distance associated with the local scalar product and $d(.,.)$ is the Riemannian distance. The corresponding Lebesgue measure on $T_p\mathcal{D}_n$ is noted $Leb_p$. Let $exp_p^*(Leb_p)$ denote the push-forward measure of $Lep_p$ by $exp_p$. The function $\theta_p$ defined by:

$$\theta_p : q \mapsto \theta_p(q) = \frac{dvol}{dexp_p^*(Leb_p)}(q) \tag{6}$$

is the density of the Riemannian measure on $\mathcal{D}_n$ with respect to the Lebesgue measure $Leb_p$ after the identification of $\mathcal{D}_n$ and $T_p\mathcal{D}_n$ induced by $exp_p$ (see Figure 1).
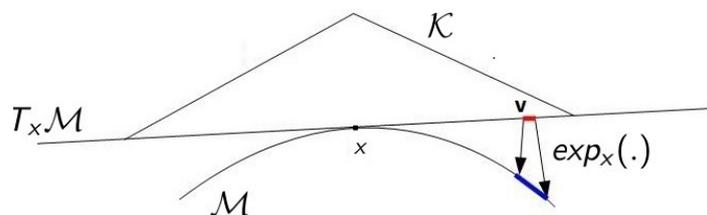


**Figure 1.** $\mathcal{M}$ is a Riemannian manifold, and $T_x\mathcal{M}$ is its tangent space at $x$. The exponential application induces a volume change $\theta_x$ between $T_x\mathcal{M}$ and $\mathcal{M}$.

Given $\mathcal{K}$ and a positive radius $r$, the estimator of $f$ proposed by [1] is defined by:

$$\hat{f}_k = \frac{1}{k} \sum_i \frac{1}{r^n} \frac{1}{\theta_{x_i}(x)} \mathcal{K}\left(\frac{d(x, x_i)}{r}\right). \tag{7}$$

The corrective factor $\theta_{x_i}(x)^{-1}$ is necessary since the kernel $\mathcal{K}$ originally integrates to one with respect to the Lebesgue measure and not with respect to the Riemannian measure. It can be noticed that this estimator is the usual kernel estimator in the case of Euclidean space. When the curvature of the space is negative, which is the case of the Siegel space, the distribution placed over each sample $x_i$ has $x_i$ as intrinsic mean. The following theorem provides convergence rate of the estimator. It is a direct adaptation of Theorem 3.1 of [1].

**Theorem 1.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold of dimension n and $\mu$ its Riemannian volume measure. Let X be a random variable taking its values in a compact subset C of $(\mathcal{M}, g)$. Let $0 < r \leq r_{inj}$, where $r_{inj}$ is the infimum of the injectivity radius on C. Assume the law of X has a twice differentiable density f with respect to*

the Riemannian volume measure. Let $\hat{f}_k$ be the estimator defined in Equation (7). There exists a constant $C_f$ such that

$$\int_{x \in \mathcal{M}} E_{x_1,...,x_k}[(f(x) - \hat{f}_k(x))^2]d\mu \le C_f(\frac{1}{kr^n} + r^4). \tag{8}$$

If $r \sim k^{\frac{-1}{n+4}}$,

$$\int_{x \in \mathcal{M}} E_{x_1,...,x_k}[(f(x) - \hat{f}_k(x))^2]d\mu = O(k^{-\frac{4}{n+4}}). \tag{9}$$

**Proof.** See Appendix A. □

It can be checked that on the Siegel space $r_{inj} = +\infty$ and that, for an isometry $\alpha$, we have:

$$\hat{f}_k(x, x_1, ..., x_k) = \hat{f}_k(\alpha(x), \alpha(x_1), ..., \alpha(x_k)).$$

Each location and direction are processed as similarly as possible. This density estimator can be used for data classification on Riemannian manifolds, see [26].

In order to obtain the explicit expression of the estimator, one must have the explicit expression of the Riemannian exponential, of its inverse, and of the function $\theta_p$ (see Equations (6) and (7)). These expressions are difficult and sometimes impossible to obtain for general Riemannian manifolds. In the case of the Siegel space, the symmetric structure makes the computation possible. Since the space is homogeneous, the computation can be made at the origin $iI \in \mathcal{H}_n$ or $0 \in \mathcal{D}_n$ and transported to the whole space. In the present work, the random variable lays in $\mathcal{D}_n$. However, in the literature, the Cartan and Iwasawa decompositions are usually given for the isometry group of $\mathcal{H}_n$. Thus, our computation starts in $\mathcal{H}_n$ before moving to $\mathcal{D}_n$.

The Killing form on the Lie algebra $\mathfrak{sp}(n, \mathbb{R})$ of the isometry group of $\mathcal{H}_n$ induces a scalar product on $\mathfrak{p}$. This scalar product can be transported on $exp(\mathfrak{p})$ by left multiplication. This operation gives $exp(\mathfrak{p})$ a Riemannian structure. It can be shown that on this Riemannian manifold, the Riemannian exponential at the identity coincides with the group exponential. Furthermore,

$$\begin{array}{rcl} \phi & : & exp(\mathfrak{p}) & \to & \mathcal{H}_n \\ & & g & \mapsto & g.iI \end{array} \tag{10}$$

is a bijective isometry, up to a scaling factor. Since the volume change $\theta_p$ is invariant under rescaling of the metric, this scaling factor has no impact. Thus, $\mathcal{H}_n$ can be identified with $exp(\mathfrak{p})$ and $exp_{iI \in \mathcal{H}_n}$ with $exp_{|\mathfrak{p}}$. The expression of the Riemannian exponential is difficult to obtain in general; however, it boils down to the group exponential in the case of symmetric spaces. This is the main element of the computation of $\theta_p$. The Riemannian volume measure on $exp(\mathfrak{p})$ is noted $vol'$. Let

$$\begin{array}{rcl} \psi & : & K \times \mathfrak{a} & \to & \mathfrak{p} \\ & & (k, H) & \mapsto & Ad_k(H). \end{array}$$

Let $\mathfrak{a}^+$ be the diagonal matrices with strictly decreasing positive eigenvalues. Let $\Lambda^+$ be the set of positive roots of $\mathfrak{sp}(n, \mathbb{R})$ as described in p. 282 in [20] ,

$$\Lambda^+ = \{e_i + e_j, i \le j\} \cup \{e_i - e_j, i < j\},$$

where $e_i(H)$ is the $i$-th diagonal term of the diagonal matrix $H$. Let $C_c(E)$ be the set of continuous compactly supported functions on the space $E$. In [27], at page 73, it is given that for all $t \in C_c(\mathfrak{p})$, there exists $c_1 > 0$ such that

$$\int_{\mathfrak{p}} t(Y)dY = c_1 \int_K \int_{\mathfrak{a}^+} t(\psi(k, H)) \prod_{\lambda \in \Lambda^+} \lambda(H)dkdH, \tag{11}$$

where $dY$ is a Lebesgue measure on the coefficients of $Y$. Let $\tilde{\mathfrak{p}} = \psi(K \times \mathfrak{a}^+)$. $\lambda \in \Lambda^+$ never vanishes on $\mathfrak{a}^+$ and $\mathfrak{p} \setminus \tilde{\mathfrak{p}}$ has a null measure. Thus,

$$\int_{\tilde{\mathfrak{p}}} t(Y) \prod_{\lambda \in \Lambda^+} \frac{1}{\lambda(H_Y)} dY = c_1 \int_{K'} \int_{\mathfrak{a}^+} t(Ad_k(H)) dk dH, \tag{12}$$

where $H_Y$ is the point in $\mathfrak{a}^+$ such that there exists $k$ in $K$ such that $\psi(k, H_Y) = Y$. Calculation in p. 73 in [27] also gives that for all $t \in C_c(\mathfrak{p})$, there exists $c_2 > 0$, such that

$$\int_{Sp(n,\mathbb{R})} t(g) dg = c_2 \int_K \int_{\mathfrak{a}^+} \int_K t(k_2.exp(Ad_{k_1}(H))) J(H) dk_1 dH dk_2, \tag{13}$$

where $dg$ is the Haar measure on $Sp(n, \mathbb{R})$ and

$$\begin{aligned} J(H) &= \prod_{\lambda \in \Lambda^+} e^{\lambda(H)} - e^{-\lambda(H)} \\ &= 2^{|\Lambda^+|} \prod_{\lambda \in \Lambda^+} \sinh(\lambda(H)). \end{aligned}$$

Thus, for all $t \in C_c(Sp(n, \mathbb{R})/K)$,

$$\int_{Sp(n,\mathbb{R})/K} t(x) dx = c_2 \int_K \int_{\mathfrak{a}^+} t(exp(Ad_k(H))) J(H) dk dH, \tag{14}$$

where $dx$ is the invariant measure on $Sp(n, \mathbb{R})/K$. After identifying $Sp(n, \mathbb{R})/K$ and $exp(\mathfrak{p})$, the Riemannian measure on $exp(\mathfrak{p})$ coincides with the invariant measure on $Sp(n, \mathbb{R})/K$. Thus, for all $t \in C_c(exp(\mathfrak{p}))$,

$$\int_{exp(\mathfrak{p})} t(x) dvol' = c_2 \int_K \int_{\mathfrak{a}^+} t(exp(Ad_k(H))) J(H) dk dH. \tag{15}$$

Using the notation $H_Y$ of Equation (12),

$$\int_{\tilde{\mathfrak{p}}} t(exp(Y)) J(H_Y) \prod_{\lambda \in \Lambda^+} \frac{1}{\lambda(H_Y)} dY = c_1 \int_K \int_{\mathfrak{a}^+} t(exp(Ad_k(H))) J(H) dk dH. \tag{16}$$

Combining Equations (15) and (16), we obtain that there exists $c_3$ such that

$$\int_{\tilde{\mathfrak{p}}} t(exp(Y)) \prod_{\lambda \in \Lambda^+} \frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)} dY = c_3 \int_{exp(\mathfrak{p})} t(x) dvol'. \tag{17}$$

The term $\frac{\sinh(\lambda(H))}{\lambda(H)}$ can be extended by continuity on $\mathfrak{a}$; thus,

$$\int_{\mathfrak{p}} t(exp(Y)) \prod_{\lambda \in \Lambda^+} \frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)} dY = c_3 \int_{exp(\mathfrak{p})} t(x) dvol'. \tag{18}$$

Let $dY$ be the Lebesgue measure corresponding to the metric. Then, the exponential application does not introduce a volume change at $0 \in \mathfrak{p}$. Since $H_0 = 0$ and $\frac{\sinh(\lambda(H))}{\lambda(H)} \xrightarrow[H \to 0]{} 1$, we have $c_3 = 1$. Let $log$ denote the inverse of the exponential application. We have

$$\frac{dlog^*(vol')}{dY} = \prod_{\lambda \in \Lambda^+} \frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)}.$$

Since $\phi$ from Equation (10) is an isometry up to a scaling factor, if $Y \in \mathfrak{p}$ and $C\phi(exp(Y))C^{-1} = exp_0(u \in T_0\mathcal{D}_n)$, then

$$\frac{dlog^*(vol)}{dLeb_0}(u) = \frac{dlog^*(vol')}{dY}(Y),$$

where $Leb_0$ refers to the Lebesgue measure on the tangent space $T_0\mathcal{D}_n$ as in Equation (6). Given $Z \in \mathcal{D}_n$, $H_Z$ from Equation (4) verifies $C\phi(exp(Ad_k(H_Z)))C^{-1} = Z$ for some $k$ in $K$. Thus,

$$\theta_0(Z) = \frac{dlog^*(vol')}{dY}(Ad_k(H_Z)) = \prod_{\lambda \in \Lambda^+} \frac{\sinh(\lambda(H_Z))}{\lambda(H_Z)}.$$

We have then

$$\theta_0(Z) = \prod_{i<j} \frac{\sinh(\tau_i - \tau_j)}{\tau_i - \tau_j} \prod_{i\leq j} \frac{\sinh(\tau_i + \tau_j)}{\tau_i + \tau_j},$$

where the $(\tau_i)$ are described in Section 2.2. Given $Z_1, Z_2 \in \mathcal{D}_n$,

$$\theta_{Z_1}(Z_2) = \theta_0(g_{Z_1}^{-1}.Z_2),$$

where $g_{Z_1}^{-1}$ is defined in Equation (3). It is thus possible to use the density estimator defined in Equation (7). Indeed,

$$\frac{1}{\theta_{Z_1}(Z_2)} \mathcal{K}\left(\frac{d(Z_1, Z_2)}{r}\right) = \prod_{i<j} \frac{\tau_i - \tau_j}{\sinh(\tau_i - \tau_j)} \prod_{i\leq j} \frac{\tau_i + \tau_j}{\sinh(\tau_i + \tau_j)} \mathcal{K}\left(\frac{(2\sum \tau_i^2)^{1/2}}{r}\right), \quad (19)$$

where the $(\tau_i)$ are the diagonal elements of $H_{g_{Z_1}^{-1}.Z_2}$. Recall that when $n = 1$, the Siegel disk corresponds to the Poincaré disk. Thus, we retrieve the expression of the kernel for the hyperbolic space,

$$\frac{1}{\theta_{Z_1}(Z_2)} \mathcal{K}\left(\frac{d(Z_1, Z_2)}{r}\right) = \frac{2\tau}{\sinh(2\tau)} \mathcal{K}\left(\frac{(2\tau^2)^{1/2}}{r}\right). \quad (20)$$

## 4. Application to Radar Processing

### 4.1. Radar Data

In space time adaptive radar processing (STAP), the signal is formed by a succession of matrices $X$ representing the realization of a temporal and spatial process. Let $\mathcal{B}_{n,m}^+$ be the set of positive definite block Teoplitz matrices composed of $n \times n$ blocks of $m \times m$ matrices (PD BT). For a stationary signal, the autocorrelation matrix $R$ is PD BT (see [5,6,14]). Authors of [5,6,14] proposed a generalization of Verblunsky coefficients and defined a parametrization of PD BT matrices,

$$\begin{array}{ccc} \mathcal{B}_{n,m}^+ & \to & Sym^+ \times \mathbb{D}_n^{m-1} \\ \mathcal{R} & \mapsto & (P_0, \Omega_1, ..., \Omega_{m-1}), \end{array} \quad (21)$$

in which the metric induced by the Kähler potential is the product metric of an affine invariant metric on $Sym^+$ and $m-1$ times the metric of the Siegel disk, up to a scaling factor. When the signal is not Gaussian, reflection/Verblunsky coefficients in Poincaré or Siegel Disks should be normalized as described in [28] by a normalized Burg algorithm. Among other references, positive definite block Teoplitz matrices have been studied in the context of STAP-radar processing in [4–6].

### 4.2. Marginal Densities of Reflection Coefficients

In this section, we show density estimation results of the marginal parameters $\Omega_k$. For the sake of visualization, only the Siegel disk $\mathbb{D}_1$ is considered. Recall that $\mathbb{D}_1$ coincides with the Poincaré disk.

The results are partly extracted from the conference paper [17]. Data used in the experimental tests are radar observations from THALES X-band Radar, recorded during 2014 field trials campaign at Toulouse Blagnac Airport for European FP7 UFO study (Ultra-Fast wind sensOrs for wake-vortex hazards mitigation) (see [29,30]). Data are representative of Turbulent atmosphere monitored by radar. Figure 2 illustrates the density estimation of six coefficients on the Poincaré unit disk under a rainy environment. The densities are individually re-scaled for visualization purposes. For each environment, the dataset is composed of 120 draws. The densities of the coefficients $\Omega_k$ are representative of the background. This information on the background is expected to ease the detection of interesting targets.
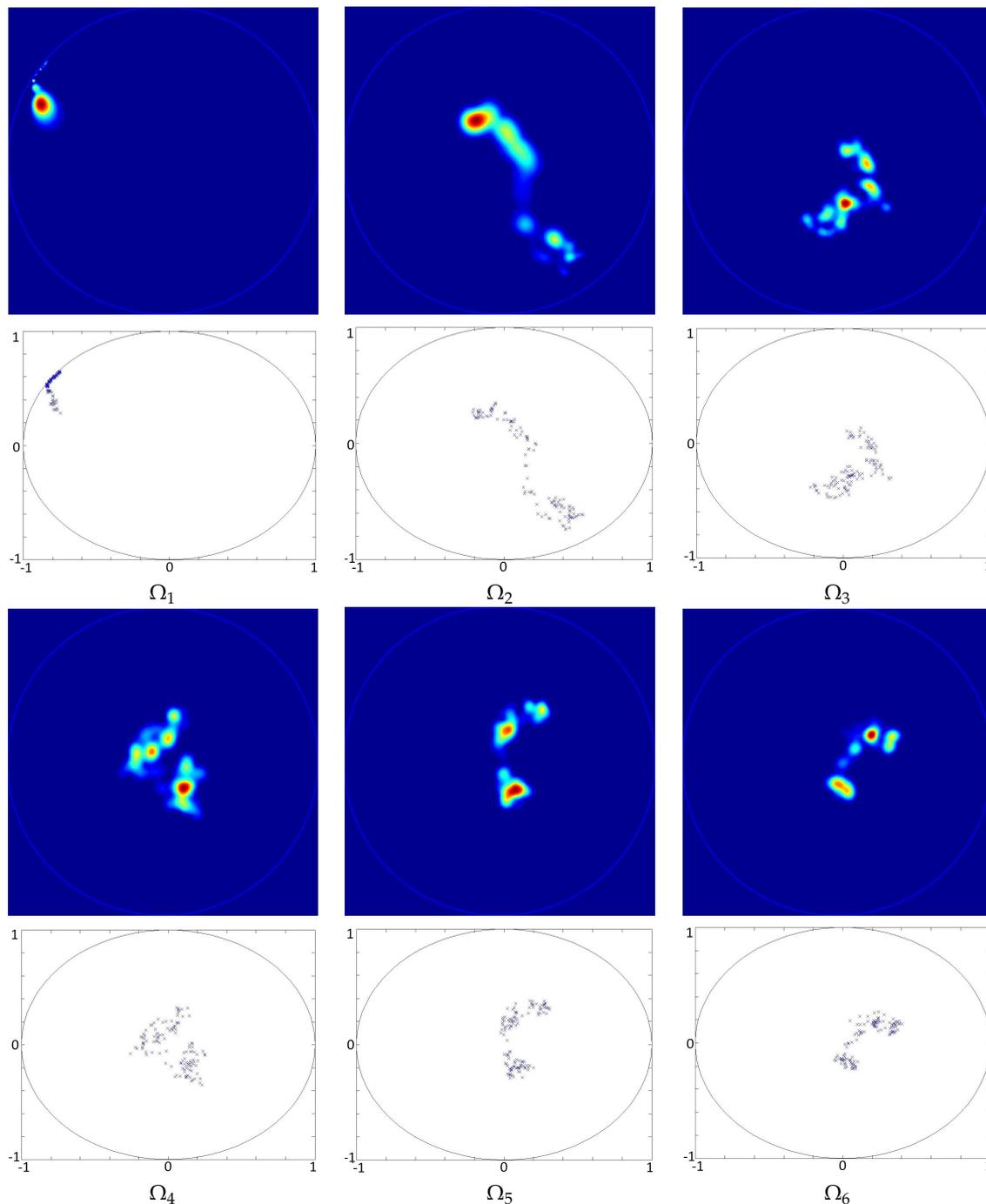
**Figure 2.** Estimation of the density of six coefficients $\Omega_k$ under rainy conditions. The expression of the used kernel is $K(x) = \frac{3}{\pi}(1 - x^2)^2 \mathbf{1}_{x<1}$. Densities are rescaled for visual purposes.

### 4.3. Radar Clutter Segmentation

Clutter refers to background Doppler signal related to meteorological conditions (e.g., wind in wooded areas, currents and breaking waves on water), which hinders detection of small and slow targets. At each range, a set of reflection coefficients are computed from the Doppler spectrum (see [31]). This set of coefficients is a point in the Poincaré poly-disk. From this set of points in the poly-disk, it is possible to estimate the underlying density. Segmenting clutter, i.e., determining zones of homogeneous Doppler characteristics (see Figure 3), enables the improvement of detection algorithms on each zone. The mean-shift algorithm enables segmentation of the space according to the kernel density estimation of a set of points. It was introduced by Fukunaga and Hostetler in 1975 (see [32]). It corresponds to a gradient ascent of the density estimator (see [33]) for a study of the statistical consistency of the gradient lines estimation. Each data point moves to a local mode of the density estimator, which yields as many clusters as modes. This algorithm has been generalized on manifolds in [34], and applied to radar images in [35]. It can thus be used to segment the set of points in the Poincaré poly-disk. Unfortunately, the mean-shift algorithm requires working with a kernel depending only on the distance to its barycenter, which is not the case of the kernel defined in Equation (19). Thus, the computations are performed without the use of the corrective term $\theta_p$. It is possible to solve this problem by replacing the corrective term by its average at a given radius, which leads to a kernel depending only on the distance to its barycenter. Our future work will focus on the computation of these averages. Let

$$\hat{f}_r^{\mathcal{K}}(x) = \frac{c_d}{k} \sum_{i=1}^{k} \frac{1}{r^n} \mathcal{K}\left(\frac{d(x_i, x)^2}{r^2}\right),$$

where $c_n$ is a normalization constant. Let $g = -k'$.
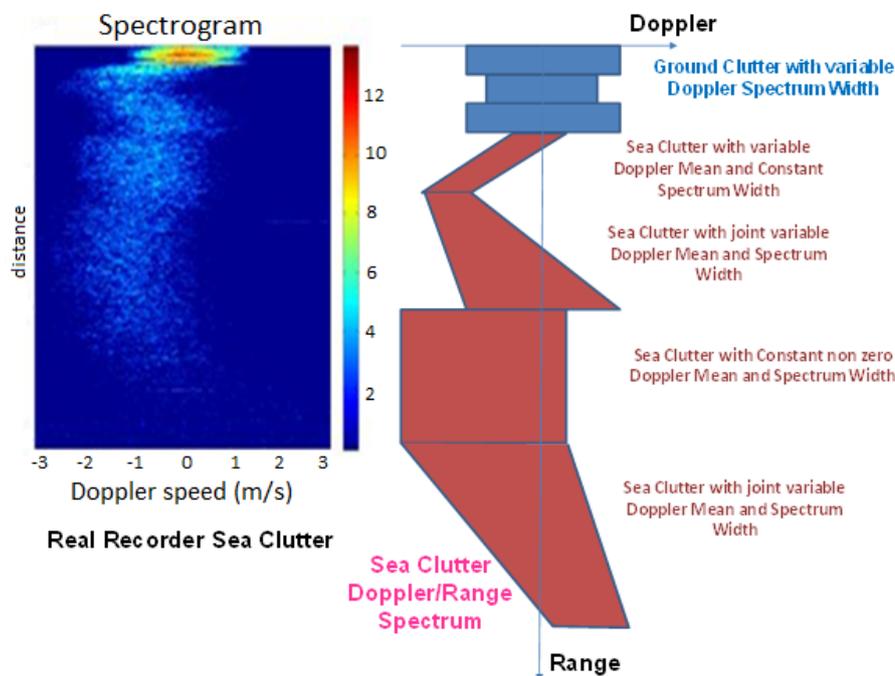


**Figure 3.** Mean and width variability of sea clutter Doppler spectrum.

The mean-shift is defined by

$$m(x) = \sum_{i=1}^{k} \frac{\frac{1}{r^{n+2}} g\left(\frac{d(x,x_i)^2}{r^2}\right)}{\sum_{i=1}^{k} \frac{1}{r^{n+2}} g\left(\frac{d(x,x_i)^2}{r^2}\right)} log_x(x_i) \propto \frac{\nabla f_r^{\mathcal{K}}}{f_r^g},$$

where $m(x)$ is in the tangent space at $x$. The algorithm moves from $x$ to $exp_x(m(x))$ until convergence to a local maximum. The points of the space are segmented according to the local maxima to which they converge.

In order to assess the quality of unsupervised classification, we use the notion of Silhouette, see [36], which computes for each point a proximity criterion with respect to other points of the same cluster and other points of different clusters (see Figure 4). Let $x$ be in the cluster $A$. We respectively define $a(x) = min_{y \in A} d(x,y)$ and $b(x) = min_{y \neq A} d(x,y)$, the minimum distance to points of the same (resp. other) class(es). The Silhouette of $x$ is

$$\frac{a(x) - b(x)}{max\{a(x), b(x)\}},$$

which takes values between $-1$ and 1, respectively, when the data point is considered "badly" and "well" clustered. The average of all the silhouettes provides an indication of the relevance of the classification. One can represent graphically the silhouette profile by plotting for each class horizontal segments of the length of the silhouette value (see Figure 5).
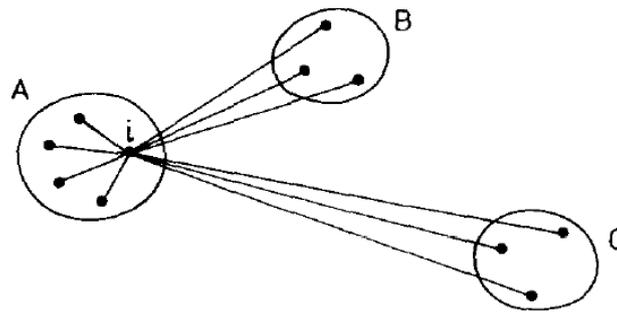


**Figure 4.** Intra and inter cluster distances.
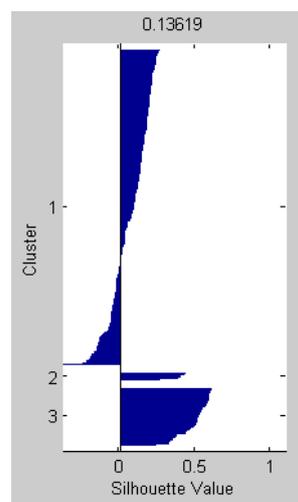


**Figure 5.** Example of silhouette.

In order to test the Riemannian Mean Shift performance, we generate simple synthetic radar clutter data. Given 250 range cells, we generate 125 cells of ground clutter (wind) centered at 0 m·s$^{-1}$, of spectral width 5 m·s$^{-1}$, to which we add 125 cells of rain clutter, centered at 5 m·s$^{-1}$, of spectral width 10 m·s$^{-1}$. This clutter is sampled 10 times and the segmentation is performed on each simulation (see Figures 6–8).
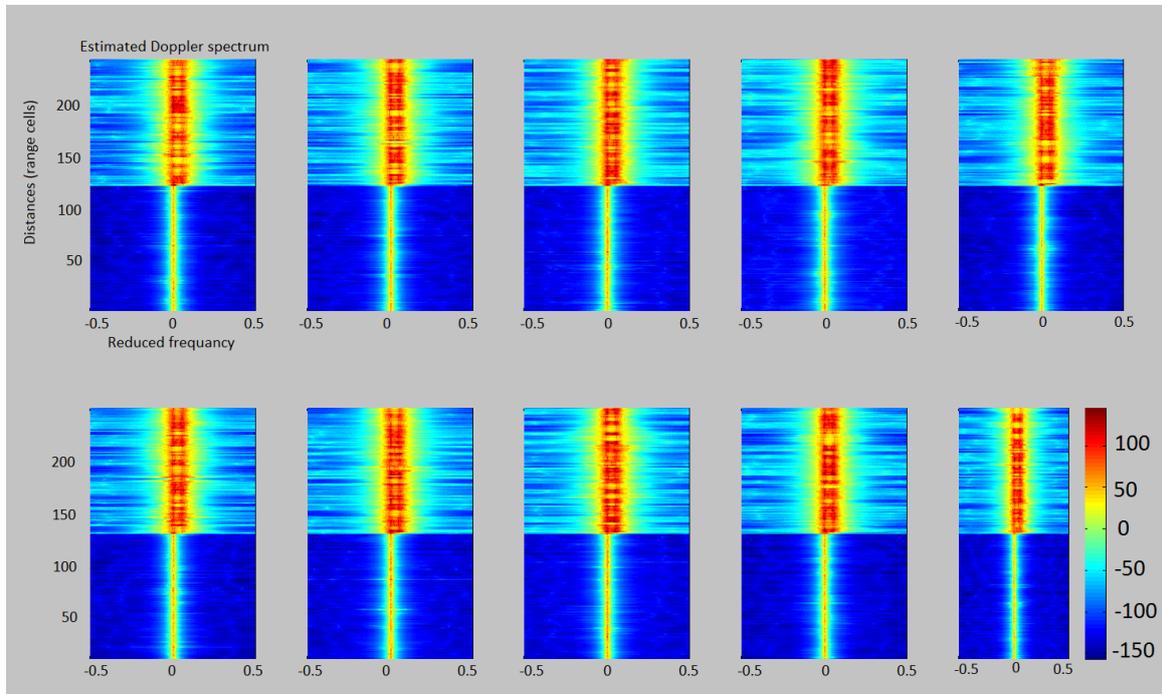


**Figure 6.** Autoregressive spectra.
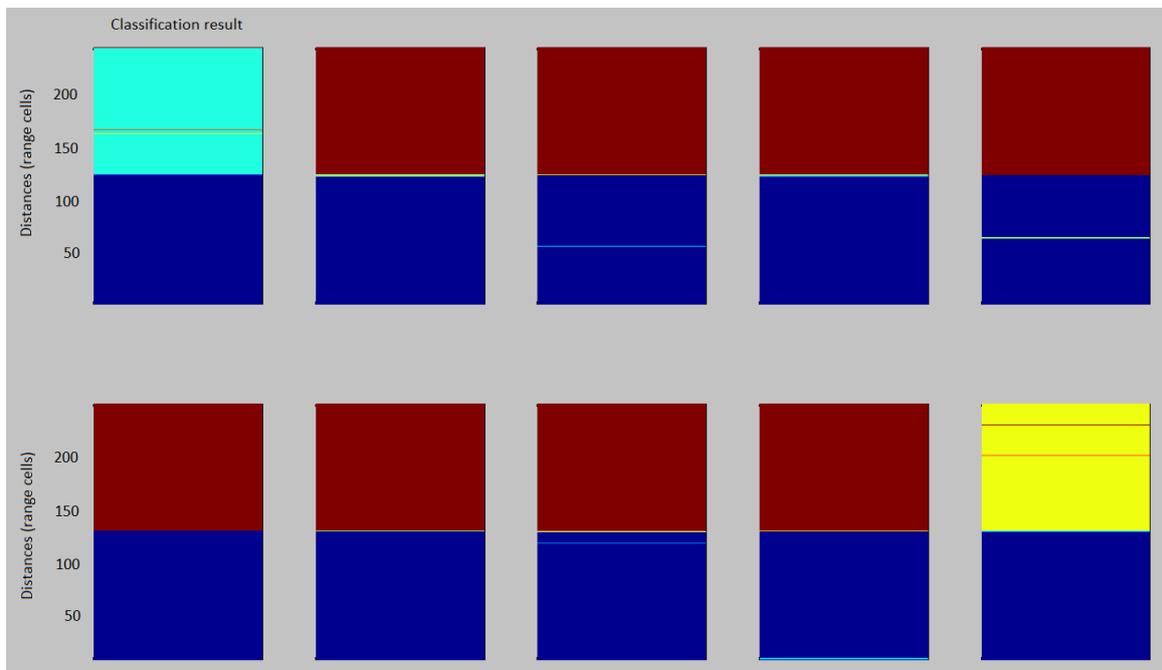


**Figure 7.** Classification results (one color per cluster).
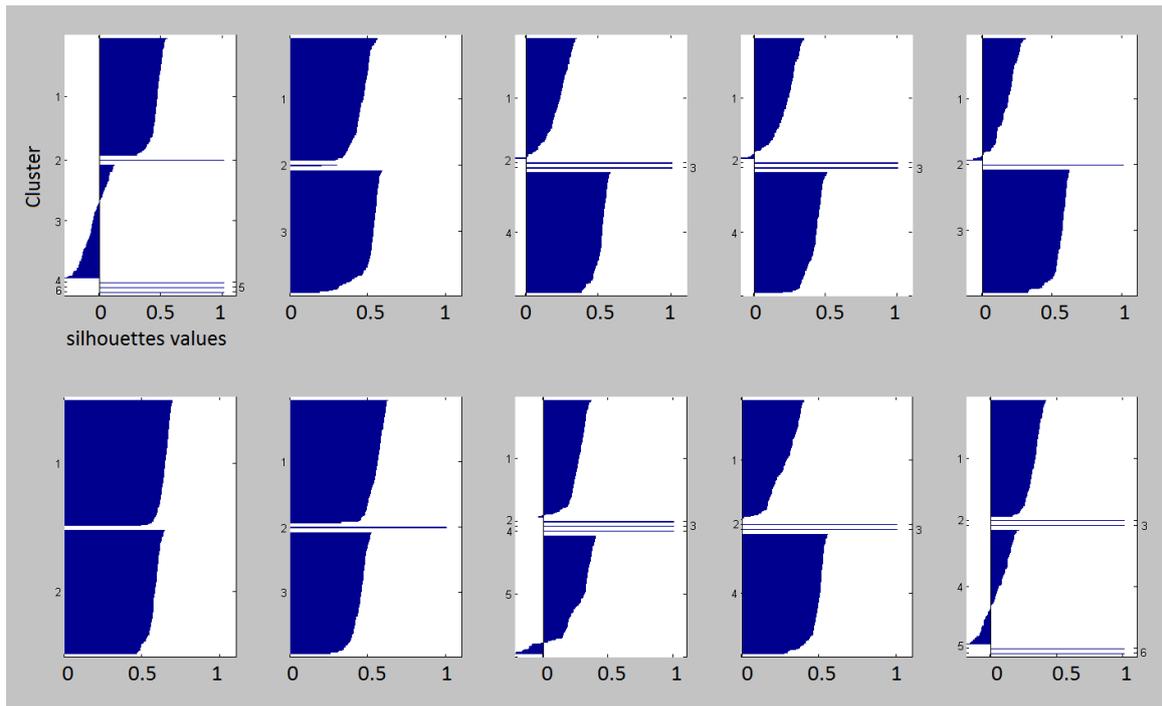
**Figure 8.** Silhouettes.

It can be seen that, apart from a few outliers, the two clutters are well classified and that the algorithm was able to distinguish between two zones of different Doppler characteristics.

We then test our algorithm on real sea clutter data (see Figures 9–11).
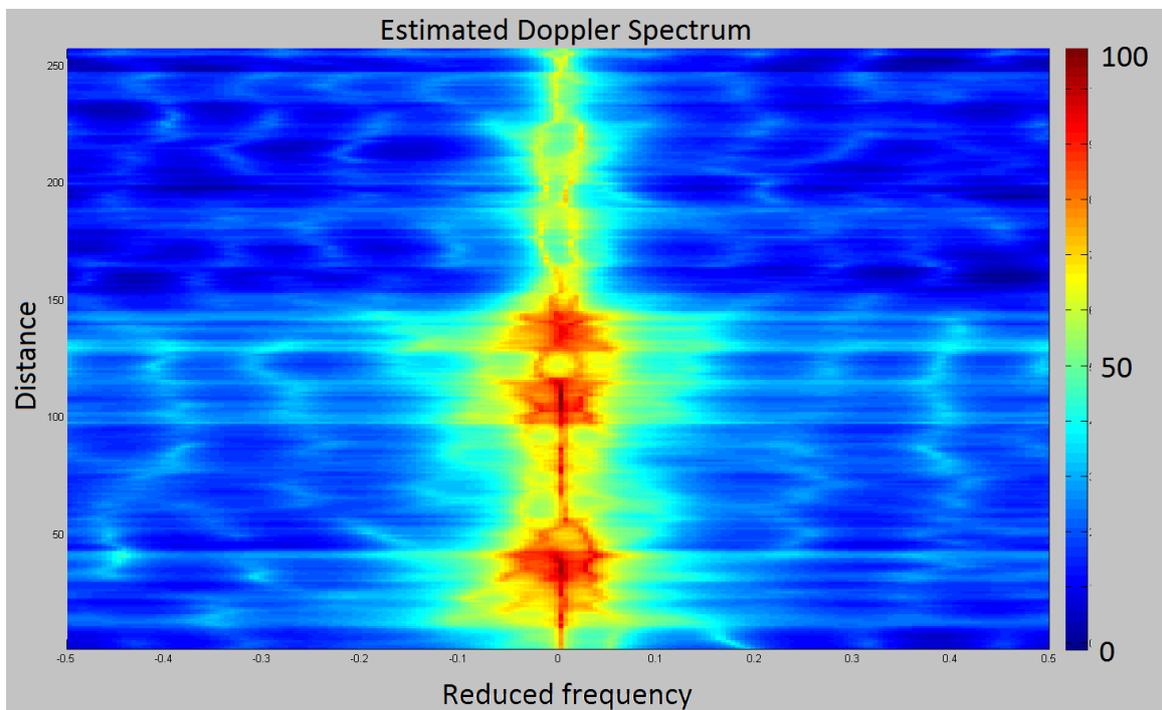


**Figure 9.** Autoregressive spectrum.
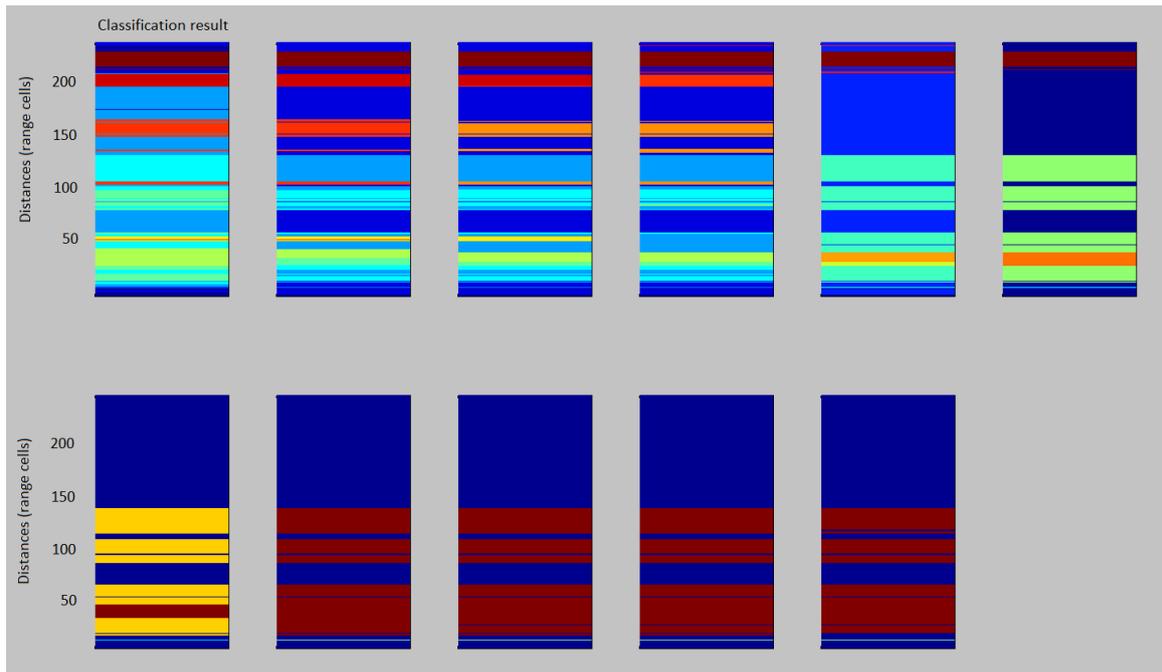
**Figure 10.** Classification results for varying radii size in the density estimator (10 to 20 closest neighbours).
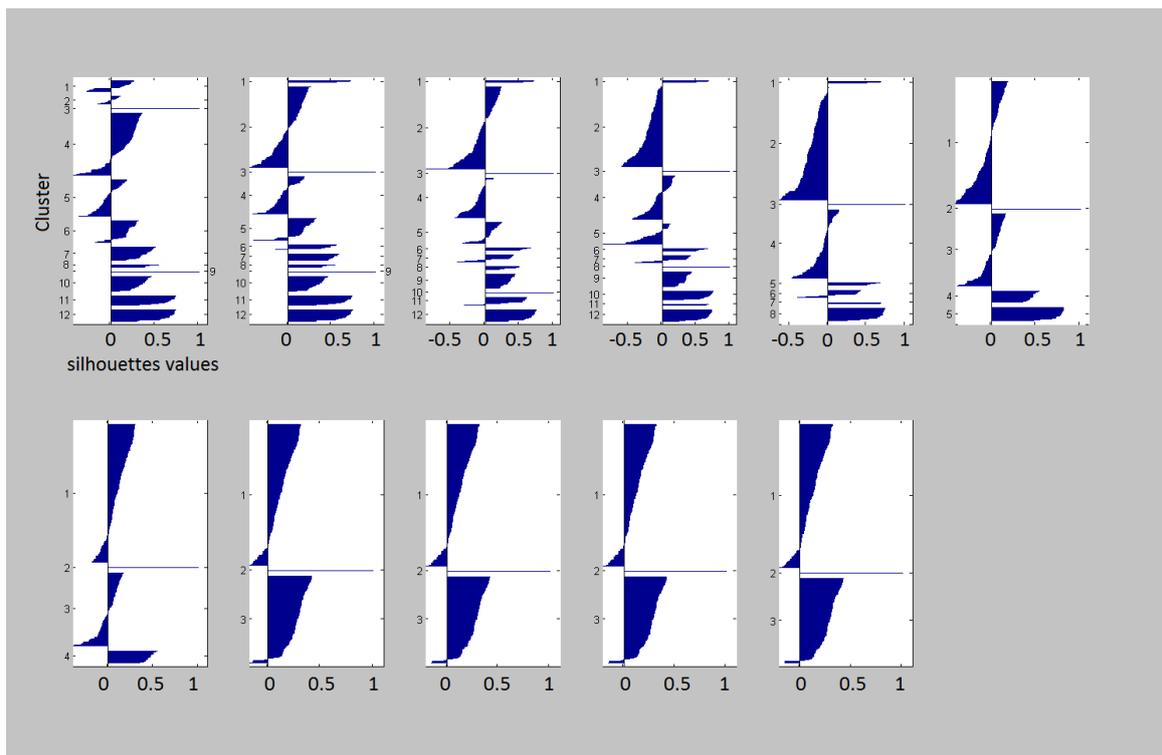


**Figure 11.** Silhouettes.

The results are more difficult to interpret in that case. The Doppler spectra are varying quite a lot along the range axis. Even though it looks over-segmented, the first classification (kernel size defined by the distance to the 10th closest neighbor point) displays the highest average silhouette value.

## 5. Conclusions

Three non parametric density estimation techniques have been considered. The main advantage of histograms in the Euclidean context is their simplicity of use. This makes histograms an interesting tool despite the fact that they do not present optimal convergence rates. On the Siegel space, histograms lose their simplicity advantage. They were thus not deeply studied. The orthogonal series density estimation also presents technical disadvantages on the Siegel space. Indeed, the series become integrals, which make the numerical computation of the estimator more difficult than in the Euclidean case. On the other hand, the use of the kernel density estimator does not present major differences with the Euclidean case. The convergence rate obtained in [1] can be extended to compactly supported random variables on non compact Riemannian manifolds. Furthermore, the corrective term whose computation is required to use Euclidean kernels on Riemannian manifolds turns out to have a reasonably simple expression. Our future efforts will concentrate on the use of kernel density estimation on the Siegel space in radar signal processing. As the experimental section suggests, we strongly believe that the estimation of the densities of the $\Omega_k$ will provide an interesting description of the different backgrounds. This non-parametric method of density estimation should be compared with parametric ones, as "Maximum Entropy Density" (Gibbs density) on homogenesous manifold as proposed in [37] based on the works of Jean-Marie Souriau. As proposed in [38], a median-shift approach might also be investigated.

**Author Contributions:** Emmanuel Chevallier carried out the mathematical development. Thibault Forget has set up the Radar clutter segmentation. Frédéric Barbaresco has introduced Poincaré/Siegel Half space and Poincaré/Siegel Disk parameterization for Radar Doppler and Space-Time Adaptive Processing based on Metric spaces deduced from Information Geometry. This parameterization has been re-used in this paper. Jesus Angulo was the Ph.D. supervisor of Emmanuel Chevallier and participates in the supervision of master thesis of Thibault Forget, both thesis are at the origin of this study. All authors have read and approved the final manuscript.

## Appendix A. Demonstration of Theorem 1

**Lemma A1.** *Let $(M, g)$ be a Riemannian manifold, let $C$ be a compact subset of $M$ and let $U$ be a relatively compact open subset of $M$ containing $C$. Then, there is a compact Riemannian manifold $(M', g')$ such that $U$ is an open subset of $M'$, the inclusion $i : U \hookrightarrow M'$ is a diffeomorphism onto its image and $g' = g$ on $\overline{U}$.*

**Proof.** We can assume that $M$ is not compact. Let $f : M \to \mathbb{R}$ be a smooth function on $M$ which tends to $+\infty$ at infinity. Since $\overline{U}$ is compact, $f^{-1}(]-\infty, a[)$ contains $\overline{U}$ for $a$ large enough. By Sard Theorem, there exists a value $a \in \mathbb{R}$ such that $f^{-1}(a)$ contains no critical point of $f$ and such that $f^{-1}(]-\infty, a[)$ contains $\overline{U}$. It follows that $N = f^{-1}(]-\infty, a])$ is a submanifold with boundary of $M$. Since $f$ tends to $+\infty$ at infinity, $N$ is compact as well as its boundary $\partial N = f^{-1}(\{a\})$.

Call $M'$ the double of $N$. It is a compact manifold which contains $N$ such that the inclusion $i : N \hookrightarrow M'$ is a diffeomorphism onto its image (see [39], Theorem 5.9 and Definition 5.10 ). Choose any metric $g_0$ on $M'$. Consider two open subsets $W_1$ and $W_2$ in $M'$ and two smooth functions $f_1, f_2 : M' \to [0, 1]$ such that

$$\overline{U} \subset W_1 \subset \overline{W}_1 \subset W_2 \subset \overline{W}_2 \subset \text{int } N,$$

the interior of $N$,

$$f_1(x) = 1$$

on $\overline{W}_1$, vanishes outside of $W_2$, and

$$f_2(x) = 1$$

outside $W_1$, and vanishes in $\overline{U}$. Define $g'$ on $M'$ by

$$g' = f_1 g + f_2 g_0$$

on $N$ and

$$g' = f_2 g_0$$

outside of $N$. Since $f_1 + f_2 > 0$, $g'$ is positive definite everywhere on $M'$. Since $f_1$ vanishes outside of $W_2$, $g'$ is smooth on $M'$. Finally, since $f_1 = 1$ and $f_2 = 0$ on $\overline{U}$, $g' = g$ on $\overline{U}$. $\square$

We can now prove Theorem 1. Let $X$ be a random variable as in Theorem 1. Following the notations of the theorem and the lemma, let $U = \{x \in \mathcal{M}, d(x, C) < r_{inj}\}$. $U$ is open, relatively compact and contains $C$. Let $(\mathcal{M}', g')$ be as in the lemma. Let $\hat{f}$ and $\hat{f}'$ be the kernel density estimators defined on $M$ and $M'$, respectively. Theorem 3.1 of [1] provides the desired results for $\hat{f}'$. For $r \leq r_{inj}$, the support and the values on the support of $\hat{f}'$ and $\hat{f}$ coincide. Thus, the desired result also holds for $\hat{f}$.

## References

1. Pelletier, B. Kernel density estimation on Riemannian manifolds. *Stat. Probab. Lett.* **2005**, *73*, 297–304.
2. Hendriks, H. Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *Ann. Stat.* **1990**, *18*, 832–849.
3. Asta, D.M. Kernel Density Estimation on Symmetric Spaces. In *Geometric Science of Information*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9389, pp. 779–787.
4. Barbaresco, F. Robust statistical radar processing in Fréchet metric space: OS-HDR-CFAR and OS-STAP processing in siegel homogeneous bounded domains. In Proceedings of the 2011 12th International Radar Symposium (IRS), Leipzig, Germany, 7–9 Septerber 2011.
5. Barbaresco, F. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median. In *Matrix Information Geometry*; Bhatia, R., Nielsen, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 199–256.
6. Barbaresco, F. Information geometry manifold of Toeplitz Hermitian positive definite covariance matrices: Mostow/Berger fibration and Berezin quantization of Cartan-Siegel domains. *Int. J. Emerg. Trends Signal Process.* **2013**, *1*, 1–87.
7. Berezin, F.A. Quantization in complex symmetric spaces. *Izv. Math.* **1975**, *9*, 341–379.
8. Lenz, R. Siegel Descriptors for Image Processing. *IEEE Signal Process. Lett.* **2016**, *25*, 625–628.
9. Barbaresco, F. Robust Median-Based STAP in Inhomogeneous Secondary Data: Frechet Information Geometry of Covariance Matrices. In Proceedings of the 2nd French-Singaporian SONDRA Workshop on EM Modeling, New Concepts and Signal Processing For Radar Detection and Remote Sensing, Cargese, France, 25–28 May 2010.
10. Degurse, J.F.; Savy, L.; Molinie, J.P.; Marcos, S. A Riemannian Approach for Training Data Selection in Space-Time Adaptive Processing Applications. In Proceedings of the 2013 14th International Radar Symposium (IRS), Dresden, Germany, 19–21 June 2013; Volume 1, pp. 319–324.
11. Degurse, J.F.; Savy, L.; Marcos, S. Information Geometry for radar detection in heterogeneous environments. In Proceedings of the 33rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Amboise, France, 21–26 September 2014.
12. Barbaresco, F. Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics. *Entropy* **2014**, *16*, 4521–4565.
13. Barbaresco, F. New Generation of Statistical Radar Processing based on Geometric Science of Information: Information Geometry, Metric Spaces and Lie Groups Models of Radar Signal Manifolds. In Proceedings of the 4th French-Singaporian Radar Workshop SONDRA, Lacanau, France, 23 May 2016.
14. Jeuris, B.; Vandebril, R. The Kahler mean of Block-Toeplitz matrices with Toeplitz structured block. *SIAM J. Matrix Anal. Appl.* **2015**, *37*, 1151–1175.
15. Huckemann, S.; Kim, P.; Koo, J.; Munk, A. Mobius deconvolution on the hyperbolic plan with application to impedance density estimation. *Ann. Stat.* **2010**, *38*, 2465–2498.
16. Asta, D.; Shalizi, C. Geometric network comparison. 2014, arXiv:1411.1350.

17. Chevallier, E.; Barbaresco, F.; Angulo, J. Probability density estimation on the hyperbolic space applied to radar processing. In *Geometric Science of Information*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 753–761.

18. Said, S.; Bombrun, L.; Berthoumieu, Y. New Riemannian Priors on the Univariate Normal Model. *Entropy* **2014**, *16*, 4015–4031.

19. Said, S.; Hatem, H.; Bombrun, L.; Baba, C.; Vemuri, B.C. Gaussian distributions on Riemannian symmetric spaces: Statistical learning with structured covariance matrices. 2016, arXiv:1607.06929.

20. Terras, A. *Harmonic Analysis on Symmetric Spaces and Applications II*; Springer: Berlin/Heidelberg, Germany, 2012.

21. Siegel, C.L. Symplectic geometry. *Am. J. Math.* **1943**, *65*, doi:10.2307/2371774.

22. Helgason, S. *Differential Geometry, Lie Groups, and Symmetric Spaces*; Academic Press: Cambridge, MA, USA, 1979.

23. Cannon, J.W.; Floyd, W.J.; Kenyon, R.; Parry, W.R. Hyperbolic geometry. In *Flavors of Geometry*; Cambridge University Press: Cambridge, UK, 1997; Volume 31, pp. 59–115.

24. Bhatia, R. Matrix Analysis. In *Graduate Texts in Mathematics-169*; Springer: Berlin/Heidelberg, Germany, 1997.

25. Kim, P.; Richards, D. Deconvolution density estimation on the space of positive definite symmetric matrices. In *Nonparametric Statistics and Mixture Models: A Festschrift in Honor of Thomas P. Hettmansperger*; World Scientific Publishing: Singapore, 2008; pp. 147–168.

26. Loubes, J.-M.; Pelletier, B. A kernel-based classifier on a Riemannian manifold. *Stat. Decis.* **2016**, *26*, 35–51.

27. Gangolli, R.; Varadarajan, V.S. *Harmonic Analysis of Spherical Functions on Real Reductive Groups*; Springer: Berlin/Heidelberg, Germany, 1988.

28. Decurninge, A.; Barbaresco, F. Robust Burg Estimation of Radar Scatter Matrix for Mixtures of Gaussian Stationary Autoregressive Vectors. 2016, arxiv:1601.02804.

29. Barbaresco, F. Eddy Dissipation Rate (EDR) retrieval with ultra-fast high range resolution electronic-scanning X-band airport radar: Results of European FP7 UFO Toulouse Airport trials. In Proceedings of the 2015 16th International Radar Symposium, Dresden, Germany, 24–26 June 2015.

30. Oude Nijhuis, A.C.P.; Thobois, L.P.; Barbaresco, F. *Monitoring of Wind Hazards and Turbulence at Airports with Lidar and Radar Sensors and Mode-S Downlinks: The UFO Project*; Bulletin of the American Meteorological Society, 2016, submitted for publication.

31. Barbaresco, F.; Forget, T.; Chevallier, E.; Angulo, J. Doppler spectrum segmentation of radar sea clutter by mean-shift and information geometry metric. In Proceedings of the 17th International Radar Symposium (IRS), Krakow, Poland, 10–12 May 2016; pp. 1–6.

32. Fukunaga, K.; Hostetler, L.D. The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition. *Proc. IEEE Trans. Inf. Theory* **1975**, *21*, 32–40.

33. Arias-Castro, E.; Mason, D.; Pelletier, B. On the estimation of the gradient lines of a density and the consistency of the mean-shift algorithm. *J. Mach. Learn. Res.* **2000**, *17*, 1–28.

34. Subbarao, R.; Meer, P. Nonlinear Mean Shift over Riemannian Manifolds. *Int. J. Comput. Vis.* **2009**, *84*, doi:10.1007/s11263-008-0195-8.

35. Wang, Y.H.; Han, C.Z. PolSAR Image Segmentation by Mean Shift Clustering in the Tensor Space. *Acta Autom. Sin.* **2010**, *36*, 798–806.

36. Rousseeuw, P. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *1*, 53–65.

37. Barbaresco, F. Geometric Theory of Heat from Souriau Lie Groups Thermodynamics and Koszul Hessian Geometry: Applications in Information Geometry for Exponential Families. *Entropy* **2016**, *18*, 386.

38. Wanga, Y.; Huang, X.; Wua, L. Clustering via geometric median shift over Riemannian manifolds. *Inf. Sci.* **2013**, *220*, 292–305.

39. Munkres, J.R. Elementary Differential Topology. In *Annals of Mathematics Studies-54*; Princeton University Press: Princeton, NJ, USA, 1967.