Innovative Tools for Radar Signal Processing Based on Cartan's Geometry of SPD Matrices & Information Geometry

F. Barbaresco, *Senior Member SEE (SEE Ampere medal 2007)*
THALES AIR SYSTEMS, Surface Radar Business Line, Strategy Technology & Innovation Dept.
Hameau de Roussigny, F-91470, Limours, France
phone: +33.(0)1.64.91.99.24, fax: +33.(0)1.64.91.67.66, email: frederic.barbaresco@thalesgroup.com

***Abstract*— New operational requirements for stealth targets detection in dense & inhomogeneous clutter are emerging (littoral warfare, low altitude asymmetric threats, battlefield in urban area…). Classical Radar approaches for Doppler & Array signal processing have reached their limits. We propose new improvements based on advanced Mathematical studies on Geometry of SPD matrix (Symmetric Positive Definite matrix) and Information Geometry, using that Radar data Covariance matrices include all information of the sensor signal.**

First, Information Geometry allows to take into account statistics of Radar covariance matrix (by mean of Fisher information matrix used in Cramer-Rao bound) to built a robust distance, called Jensen, Siegel or Bruhat-Tits metric. Geometry on "Symmetric cones", developed in frameworks of Lie Group and Jordan Algebra, provides new algorithms to compute Matrix Geometric Means that could be used for "matrix CFAR". This innovative approach avoids classical drawbacks of Doppler processing by filter banks or FFT in case of bursts with very few pulses.

***Index Terms*— Information Geometry, Bruhat-Tits metric, Cartan-Hadamard Manifold.**

I. Introduction

New Operational Requests are emerging :

- Detection of new low altitude threats (small and/or stealth, agile, asymmetric…)
- Increase Reaction Time against "Ultra Critical threats"

But, classical CFARs have many drawbacks for detection in dense and inhomogeneous clutters. More particularly, their probability of detection is sub-optimal close to clutter transitions because they are not edge-preserving & poorly take into account Clutter statistic. This bad behavior is critical as these Clutter transitions correspond to the most threatening areas: crest-line & unmasking areas (threats: furtive helicopter pop-up with missile shooting, low altitude cruise missile & UAV, asymmetric threats, rockets/batteries…). In this context, definition of an Edge-preserving CFAR based on clutter statistic will increase Survivability against Lethal Threat (with Limited Time Exposure behind Crest-Lines) by improving system Reaction Time.

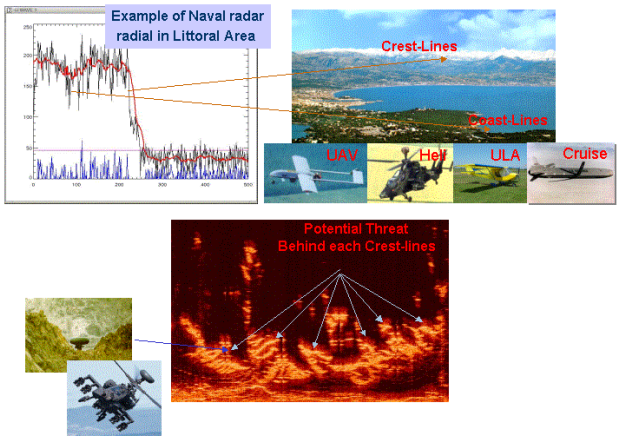


Figure 1 - Critical threatening areas (coast-lines, crest-lines, …) with emerging of new low altitude threats

This challenge of new Radar missions will increase pressure for Doppler processing improvement. This constraint is bal-anced by an antagonistic one that pushes to use short Doppler bursts with very few pulses in order to relax time budget for Multi-function/Multi-mission radars. For short waveforms, Classical FFT or Doppler Filter Banks are not efficient, and suffer of the following drawbacks:

- Poor Doppler Resolution
- If Target Doppler is between two Doppler filters, detec-tion is sub-optimal
- High intensity of Ground Clutter is not limited to zero-Doppler filter but pollution is spread over all filters due to poor Filter-Banks Resolution & Doppler Filter side lobes.

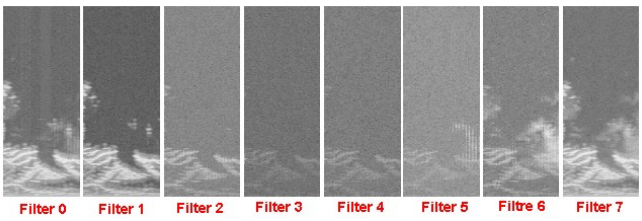


Figure 2 - Pollution of all Doppler Filters by "strong" ground clutter" in case of short Doppler Burst

Then, we have to define a robust and efficient detector on

Funded by French MoD (DGA/MRIS)

short bursts in case of dense and inhomogeneous clutters and to find an alternative to classical use of Doppler filter banks (or FFT) and CFAR strategies.

We propose to apply a new "matrix CFAR" that will act directly on Radar data covariance matrix [8]. For this purpose, we need to establish the definition of:

1. **"Statistical" distance between covariance matrices** based on Information Geometry

2. **Geometric means of a set of N covariance matrices** based on Geometry of SPD matrix (called symmetric cones in Mathematical literature) and its natural extension to "median matrix"

We illustrate this new concept of "matrix CFAR" in the following figure:

a Geometric mean or Median matrix is computed to estimate the ambience of the "cell under test" neighborhood, and robust detector is obtained by a "matrix distance" with specific invariance properties.

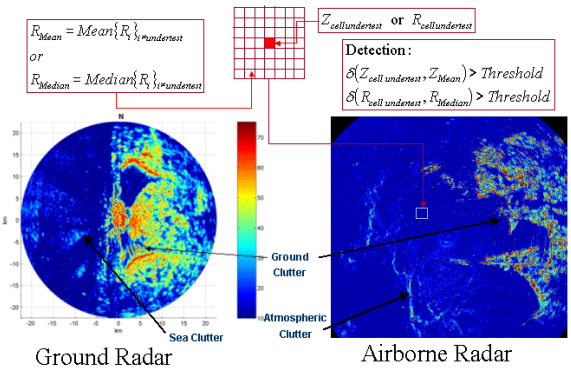


Figure 3 - New robust " matrix" CFAR based on distance & geometric mean/median of radar covariance SPD matrices

We will introduce this new "matrix CFAR" tool in the general framework of different mathematical theories:

1. **Information Geometry**, that has been introduced by C.R.Rao [2], and axiomatized by N. Chentsov [13], (with same roots that the well-known Cramer-Rao bound), allows to build a distance between statistical distributions that is invariant to non-singular parameterization transformations.

2. **Symplectic Geometry**, used by C.L. Siegel [1][11] to define distance between complex symmetric matrix whose the imaginary part is Positive Definite. This is an extension for matrix of the upper half space of Poincare. The associate metric and distance is invariant under generalized Möbius transform.

3. **Geometry on Symmetric Cones** [17][19], where Symmetric Hermitian Space is equivalent to a tube domain, where we can define a "geodesic between matrices" by mean of Lie Group Theory [14] and Jordan Algebra [15][16]. In case of Symmetric matrix, this space is called Bruhat-Tits [9] space or Cartan-Hadamard Manifold.

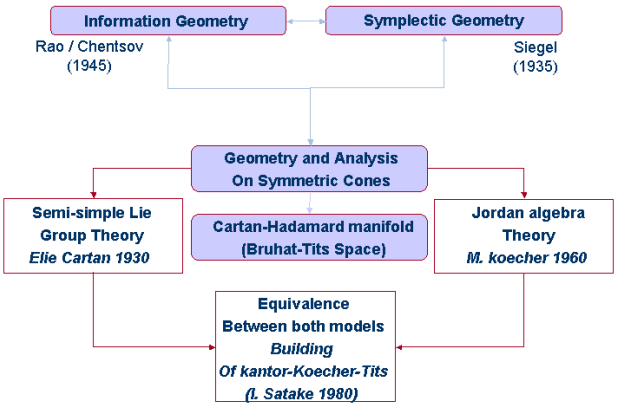


Figure 4 - Three Geometries used for processing on SPD matrix & their application for Radar covariance matrix

II. INFORMATION GEOMETRY

In his seminal paper of 1945, C.R.Rao [2] has introduced two concepts: Cramer-Rao bound and Information Geometry. The tool of Cramer-Rao bound is largely used in Signal Processing and Radar communities, but Information Geometry is less popular. The Cramer-Rao bound is given by the inverse of the Fisher Information Matrix $I(\theta)$:

$$E\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^T\right] \geq I(\theta)^{-1} \tag{1}$$

A. *Rao & Chentsov's Information Geometry*

Chentsov [13] was the first to introduce the Fisher information matrix as a Riemannian metric on the parameter space, considered as a differentiable manifold. Chentsov was led by decision theory when he considered a category whose objects are probability spaces and whose morphisms are Markov Kernels. Chentsov's great achievement was that up to a constant factor the Fisher information yields the only monotone family of Riemannian metrics on the class of finite probability simplexes. In parallel, Burbea & Rao [2] have introduced a family of distance measures, based on the so-called α-order entropy metric, generalizing the Fisher Information metric that corresponds to the Shannon entropy. Such a choice of the matrix for the quadratic differential metric was shown to have attractive properties through the concepts of discrimination and divergence measures between probability distribution. As is well known from differential geometry, the Fisher information matrix is a covariant symmetric tensor of the second order, and hence, the associate metric is invariant under the admissible transformations of the parameters. The information geometry considers probability distributions as differentiable manifolds, while the random variables and their expectation appear as vectors and inner products in tangent spaces to these manifolds.

Chentsov [13] has introduced a distance between parametric families of probability distributions $G_\Theta = \{p(./\theta) : \theta \in \Theta\}$ with Θ the space of parameters, by

considering, to the first order, the difference between the log-density functions. Its variance defines a positive definite quadratic differential form based on the elements of the Fisher matrix. If we note traditional Fisher information matrix:

$$I(\theta) = E\left[s(x,\theta).s(x,\theta)^{+}\right] \text{ where } s(x,\theta) = \frac{\partial \ln p(x/\theta)}{\partial \theta} \quad (2)$$

$$d\ln p(x/\theta) = \sum_i \frac{\partial \ln p(x/\theta)}{\partial \theta_i}.d\theta_i \quad (3)$$

Its variance defines a positive definite quadratic differential form based on the elements of the Fisher matrix and a Taylor expansion to the 2nd order of the Kullback divergence gives a Riemannian metric:

$$K\left[p(./\theta), p(./\theta + d\theta)\right] = \frac{1}{2!}\sum_{i,j} g_{ij}(\theta).d\theta_i.d\theta_j^{*} + O\left(\left|d\theta\right|^3\right) \quad (4)$$

where $I(\theta) = \left[g_{ij}(\theta)\right]$ and $g_{ij}(\theta) = E\left[\frac{\partial \ln p(x/\theta)}{\partial \theta_i}.\frac{\partial \ln p(x/\theta)}{\partial \theta_j^{*}}\right]$

B. Information Geometry for Radar Complex Circular Multivariate Gaussian Distribution

Classically, a complex circular multivariate Gaussian distribution of zero mean models Radar data and is given by :

$$p(X_n/R_n) = (\pi)^{-n}.\left|R_n\right|^{-1}.e^{-Tr\left[\hat{R}_n.R_n^{-1}\right]} \quad (5)$$

with $\hat{R}_n = (X_n - m_n).(X_n - m_n)^{+}$ and $E\left[\hat{R}_n\right] = R_n$

Fisher matrix elements are provided by derivatives of first moments:

$$g_{ij}(\theta) = -E\left[\frac{\partial \ln p(X_n/\theta_n)}{\partial \theta_i.\partial \theta_j}\right] = -Tr\left[\partial_i R_n.\partial_j R_n^{-1}\right] + \partial_i m_n^{+}.R_n^{-1}.\partial_j m_n \quad (6)$$

If we assume zero-mean process $m_n = 0$, we deduce from

$$R_n.R_n^{-1} = I_n \Rightarrow \partial R_n = -R_n.\partial R_n^{-1}.R_n \quad (7)$$

that : $g_{ij}(\theta) = Tr\left[\left(R_n.\partial_i R_n^{-1}\right)\left(R_n.\partial_j R_n^{-1}\right)\right]$ (8)

Then, we obtain :

$$ds^2(\theta) = Tr\left[R_n.\left(\sum_i \partial_i R_n^{-1}.d\theta_i\right).R_n.\left(\sum_j \partial_j R_n^{-1}.d\theta_j\right)\right] \quad (9)$$

But as $dR_n^{-1} = \sum_i \partial_i R_n^{-1} d\theta_i$

We conclude that $ds^2 = Tr\left[\left(R_n dR_n^{-1}\right)^2\right] = Tr\left[\left(d\ln R_n\right)^2\right]$ (10)

This metric is a *GL(g,R)*-invariant Riemannian metric and its Laplacian is given by : $\Delta = Tr\left(\left(R_n \frac{\partial}{\partial R_n}\right)^2\right)$

$$ds^2 = \left\|R_n^{-1/2} dR_n R_n^{-1/2}\right\|^2 \quad (11)$$

with $\|A\|^2 = \langle A, A\rangle$ and $\langle A, B\rangle = Tr\left(AB^T\right)$

By integration, the distance between 2 Radar Hermitian SPD matrices, R_1 and R_2, is given by mean of their extended eigen-values $\{\lambda_k\}_{k=1}^{n}$ [19]:

$$d^2(R_1, R_2) = \left\|\log\left(R_1^{-1/2}.R_2.R_1^{-1/2}\right)\right\|^2 = \sum_{k=1}^{n} \log^2(\lambda_k) \quad (12)$$

with $\det\left(R_1^{-1/2}.R_2.R_1^{-1/2} - \lambda.I\right) = \det\left(R_2 - \lambda R_1\right) = 0$

C. Dual Differential Information Geometry

The geometry of a family of probability distribution is also characterized by a dual differential geometry determined by a couple of affine connections and a divergence associated with a couple of dual potential functions [3]. This kind of dual geometry was studied first by Eugenio Calabi. If a Riemannian manifold *(M,g)* is flat with respect to a pair of torsion-free dual affine connections, then there exists a pair of dual coordinate systems which is associated with dual potential functions via a Legendre transformations. If we consider a multivariate Gaussian distribution in exponential form:

$$p(X/m, R) = e^{-\frac{1}{2}X^T R^{-1} X + X^T R^{-1} m - \psi(m,R)} \quad (13)$$

Dual coordinates and dual potential functions are related by the following Legendre transformation:

$$\tilde{\Phi} \equiv \left\langle \tilde{\Theta}, \tilde{H}\right\rangle - \tilde{\Psi} \text{ where } \left\langle \tilde{\Theta}, \tilde{H}\right\rangle = Tr\left(\theta\eta^T + \Theta H^T\right) \quad (14)$$

$$\begin{cases} \tilde{\Theta} = (\theta, \Theta) = \left(R^{-1}m, (2R)^{-1}\right) \\ \tilde{H} = (\eta, H) = \left(m, -R + mm^T\right) \end{cases} \text{ and } \begin{cases} \dfrac{\partial \tilde{\Psi}}{\partial \theta} = \eta \\ \dfrac{\partial \tilde{\Psi}}{\partial \Theta} = H \end{cases} \text{ and } \begin{cases} \dfrac{\partial \tilde{\Phi}}{\partial \eta} = \theta \\ \dfrac{\partial \tilde{\Phi}}{\partial H} = \Theta \end{cases}$$

$$\Rightarrow \begin{cases} \tilde{\Psi}\left(\tilde{\Theta}\right) = 2^{-2} Tr\left(\Theta^{-1}\theta\theta^T\right) - 2^{-1}\log\left(\det\Theta\right) + 2^{-1} n \log(\pi) \\ \tilde{\Phi}\left(\tilde{H}\right) = -2^{-1}\log\left(1 + \eta^T H^{-1}\eta\right) - 2^{-1}\log\left(\det(-H)\right) - 2^{-1} n \log(2\pi e) \end{cases}$$

We can observe that one of the potential function is the Entropy of the process X : $\tilde{\Phi}\left(\tilde{H}\right) = E\left[\log p\right]$. These dual potential functions and coordinates are related to Kullback divergence:

$$\int p(X/m_1, R_1)\log\frac{p(X/m_1, R_1)}{p(X/m_2, R_2)}dX \quad (15)$$
$$= \tilde{\Psi}\left(\tilde{\Theta}_2\right) + \tilde{\Phi}\left(\tilde{H}_1\right) - \left\langle \tilde{\Theta}_2, \tilde{H}_1\right\rangle = \tilde{\Psi}\left(\tilde{\Theta}_1\right) + \tilde{\Phi}\left(\tilde{H}_2\right) - \left\langle \tilde{\Theta}_1, \tilde{H}_2\right\rangle$$

As these potential function are convex, their Hessian $g_{ij} = \frac{\partial^2 \tilde{\Psi}}{\partial \Theta_i \partial \Theta_j}$ and $g_{ij}^{*} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j}$ with $g_{ij}(\Theta) g_{jk}^{*}(H) = \delta_k^i$ define Riemannian metrics, as previously explained for Information Geometry:

$$ds^2 = \frac{1}{2} g_{ij} d\Theta_i d\Theta_j + O\left(\left|d\Theta_i\right|^3\right) = \frac{1}{2} g^{ij} dH_i dH_j + O\left(\left|dH_i\right|^3\right) \quad (16)$$

III. Information Geometry with Radar Complex Autoregressive Model

If we apply Information Geometry on Radar data modeled by a complex autoregressive process [6], we can define the metric as the Hessian of Entropy given by reflection coefficients (or PARCOR coefficients) $\{\mu_k / |\mu_k| < 1\}_{k=1}^{n-1}$:

$$g_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j} \text{ with } \tilde{\Phi}(R) = -\log\left(\det R\right) - n\log(\pi e) \quad (17)$$

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1}(n-k).\ln\left[1 - \left|\mu_k\right|^2\right] + n.\ln\left[\pi.e.P_0\right]$$

A seminal paper of Erich Kähler [5] has introduced natural extension of Riemannian geometry to Complex Manifold during 30th 's of last century. We can easily apply this geometric framework for information metric definition. Let a complex Manifold M^n of dimension *n*, we can associate a Kählerian metric, which can be locally defined by its

definite positive Riemannian form: $ds^2 = 2\sum_{i,j=1}^{n} g_{i\bar{j}}.dz^i dz^{\bar{j}}$.

Kähler assumption sets that we can define a Kähler potential Φ , such that $g_{i\bar{j}} = \frac{\partial^2 \Phi}{\partial z^i \partial z^{\bar{j}}}$. Fundamental relation, given by Erich Kähler, is that Ricci tensor can be expressed by :

$$R_{i\bar{j}} = -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j} \quad (18)$$

with the associated scalar curvature $R = \sum_{k,l=1}^{n} g^{k\bar{l}} R_{k\bar{l}}$.

In case of Complex Auto-Regressive (CAR) models, if we choose as Kähler potential Φ the Entropy of the process expressed according to reflection coefficients in the unit Polydisk $\{z / |z_k| < 1 \ \forall k = 1,...n\}$, Kähler potential is given by:

$$\Phi = \sum_{k=1}^{n-1} \rho_k \ln\left[1 - |z_k|^2\right] = \ln K_D(z,z) \quad (19)$$

and Bergman kernel $K_D(z,z) = \prod_{k=1}^{n-1} \left(1 - |z_k|^2\right)^{\rho_k}$ (20)

Very surprisingly, this case was the first example of potential studied by Erich Kähler in his seminal paper, named by Erich Kähler "*Hyper-Abelian case*" [5].

If we compute the Hessian of Entropy with parametrization [7]:

$$\theta^{(n)} = \begin{bmatrix} P_0 & \mu_1 & \cdots & \mu_{n-1} \end{bmatrix}^T = \begin{bmatrix} \theta_1^{(n)} & \cdots & \theta_n^{(n)} \end{bmatrix}^T \quad (21)$$

, we can deduce expression of Information metric :

$$g_{11} = n\alpha_0^2 = nP_0^{-2} \quad \text{and} \quad g_{ij} = \frac{(n-i).\delta_{ij}}{\left(1 - |\mu_i|^2\right)^2} \quad (22)$$

Then, Ricci tensor is obtain by Kähler formula: $R_{k\bar{l}} = -2\delta_{kl}\left(1 - |\mu_k|^2\right)^{-2}$ for $k = 2,...,n-1$ and $R_{11} = -2P_0^{-2}$

Identically, we can deduce scalar curvature:

$$R = -2.\left[n^{-1} + \sum_{j=1}^{n-1}(n-j)^{-1}\right] = -2.\left[\sum_{j=0}^{n-1}(n-j)^{-1}\right] \quad (23)$$

This curvature diverges when n tends to infinity. We can observe that we have not a Kähler-Einstein metric ($R_{i\bar{j}} = k_0 g_{k\bar{l}}$) but a more general relation defined by :

$$\left[R_{ij}\right] = B^{(n)}\left[g_{ij}\right] \text{ with } R = Tr\left[B^{(n)}\right] \quad (24)$$

where $B^{(n)} = -2diag\left\{.., (n-i)^{-1}, ..\right\}$

There are also some links with Calabi Geometric Flow that drives evolution of potential ϕ , that represents Entropy for AR model:

$$\frac{\partial \phi}{\partial t} = R - \bar{R} \quad \text{with} \quad R \text{ scalar curvature} \quad (25)$$

that minimizes the functional : $S(g) = \int_M R^2(g) dV(g)$ (26)

Recently, Bakas has identified this flow as heat flow on Kähler potential:

$$ds^2 = 2e^{\Phi(z,z^*,t)} dz.dz^* \Rightarrow \frac{\partial \Phi}{\partial t} = -\Delta\Delta\Phi \text{ with } \Delta. = e^{-\Phi} \frac{\partial^2 .}{\partial z \partial z^*} \quad (27)$$

Kähler geometry can be used to define Frechet's mean on Kähler potential function by its geodesics.

IV. SYMPLECTIC GEOMETRY AND SIEGEL UPPER HALF PLANE

Symplectic Group $Sp_{2n}R$ is one possible generalization of the group $SL_2R = Sp_2R$ (group of invertible matrices with determinant 1) to higher dimensions. This generalization goes further, since they act on a symmetric homogeneous space, the Siegel upper half plane, and this action has quite a few similarities with the action of SL_2R on the hyperbolic plane. Carl Ludwig Siegel did a study of this action in 1935 in his book "Symplectic Geometry" [1].

Let F be either the real or the complex field, the symplectic Group is the group of all matrices $M \in GL_{2n}F$ satisfying :

$$Sp(n,F) \equiv \left\{M \in GL(2n,F) / M^T JM = J\right\}, J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \in SL(2n,R)$$

$$\text{or } M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n,F) \Leftrightarrow A^T C \text{ et } B^T D \text{ symmetric and } A^T D - C^T D = I_n$$

The Siegel upper half plane is the set of all complex symmetric (n x n) matrices with positive definite imaginary part. We denote it by :

$$SH_n = \left\{Z = X + iY \in Sym(n,C) / \operatorname{Im}(Z) = Y > 0\right\} \quad (28)$$

The action of the Symplectic Group on the Siegel upper half plane is transitive. The group $PSp(n,R) \equiv Sp(n,R)/\{\pm I_{2n}\}$ is group of SH_n biholomorphisms via generalized Möbius transformations: $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ + B)(CZ + D)^{-1}$

$PSp(n,R)$ acts as a sub-group of isometries. Siegel has proved that Symplectic transformations are isometries for the Siegel metric in SH_n. It can be defined on SH_n using the distance element at the point $Z = X + iY$, as defined by [12][18] :

$$ds^2_{Siegel} = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ with } Z = X + iY \quad (29)$$

This metric is a *Sp(n,F)*-invariant Kähler metric on SH_n and Maass has proved that its Laplacian is given by [11] :

$$\Delta = 4.Tr\left(Y^t\left(Y\frac{\partial}{\partial \bar{Z}}\right)\frac{\partial}{\partial Z}\right)$$

we can observed that is a more general case that previous one studied in framework of Information Geometry. If we set $X = 0$ and $Y = R$, such that $Z = iR$, then we recover previous metric : $ds^2 = Tr\left(\left(R^{-1}(dR)\right)^2\right)$.

The distance induced in SH_n by this Riemannian metric can be found in Siegel work and is given by [19] :

$$d_{Siegel}(Z_1, Z_2) = \left(\sum_{k=1}^{n} \log^2\left(\frac{1+\sqrt{r_k}}{1-\sqrt{r_k}}\right)\right)^{1/2} \text{ with } Z_1, Z_2 \in SH_n \quad (30)$$

where the r_k's are the eigenvalues of the cross-ratio :

$$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1} \quad (31)$$

For our application case, $Z = iR$, and the distance is given by the eigenvalues of the cross-ratio :

$$R(Z_1, Z_2) = (R_1 - R_2)(R_1 + R_2)^{-1}(R_1 - R_2)(R_1 + R_2)^{-1} \quad (32)$$

As we previously observed for AR, this metric is Kählerian [12] :

$$\Omega = \frac{1}{2i} Tr\left(Y^{-1} dZ \wedge Y^{-1} d\bar{Z}\right) = Tr\left(dX \wedge d\left(Y^{-1}\right)\right) \Rightarrow d\Omega = 0 \tag{33}$$

V. GEOMETRY OF SYMMETRIC CONES

Geometry of Symmetric Cones has been studied in the framework of Semi-Lie Group theory by Elie Cartan in 1930 [14] an in the framework of Euclidean Algebra theory by M. Koecher in 1960 [15]. Let Ω be an open convex cone in a Euclidean vector space V of dimension n, T_Ω is the tube domain : $T_\Omega = V + i\Omega = \{z = x + iy / x \in V, y \in \Omega\}$. It is a Hermitian symmetric space isomorphic to *G/K* where *G* is the group of holomorphic automorphisms of T_Ω and *K* is the stabilizer of *i.e* in *G*. In the case of $V = Sym(n, R)$, the tube domain is the Siegel upper half plane $T_{\Omega Sym}$, the group *G* is the Symplectic group $Sp_n R$ and *K* is the unitary group $U(n)$.

VI. BRUHAT-TITS SPACE & CARTAN-HADAMARD MANIFOLD

With respect to the distance metric δ (arising from the previously defined Riemannian trace metric) on space of symmetric positive definite matrices, the space is a Bruhat-Tits space [9][16], and the unique midpoint of any two points is given by their geometric mean $A \circ B$.

$$\begin{aligned} &\delta(A,B) = \left(\sum_{i=1}^{n} \log^2 \lambda_i\right) \text{ with } \det\left(AB^{-1} - \lambda I\right) = 0 \\ &\delta(A,B) = 2\delta(A, A \circ B) = 2\delta(B, A \circ B) \\ &\delta\left(C^T AC, C^T BC\right) = \delta(A,B) = \delta(A^{-1}, B^{-1}) \end{aligned} \tag{34}$$

A Bruhat-Tits space is a space with complete metric that satisfies the semi-parallelogram law:

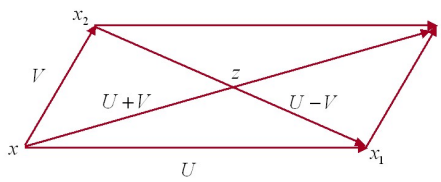
$$\begin{aligned} &\forall x_1, x_2 \in \mathrm{X} \ \ \exists z \in X \text{ such that:} \\ &\delta(x_1,x_2)^2 + 4\delta(x,z)^2 \le 2\delta(x,x_1)^2 + 2\delta(x,x_2)^2 \ \ \forall x \in X \end{aligned} \tag{35}$$

This inequality is motivated by the parallelogram law, in planar geometry which states that, using the Euclidean distance d^2, the sum of the squared lengths of the diagonals equals the sum of the squared lengths of the sides of the parallelogram :

$$d_2(x_1,x_2)^2 + d_2(x,x_3)^2 = 2d_2(x,x_1)^2 + 2d_2(x,x_2)^2 \tag{36}$$

Let z the midpoint between x_1 and x_2. Since $2.d_2(x,z) = d_2(x,x_3)$, substitution in the parallelogram law yields:

$$d_2(x_1,x_2)^2 + 4d_2(x,z)^2 = 2d_2(x,x_1)^2 + 2d_2(x,x_2)^2 \tag{37}$$



$$\|U-V\|^2 + \|U+V\|^2 = 2\|U\|^2 + 2\|V\|^2$$

Figure 5 – Parallelogram Law

Generalizing this to Bruhat-Tits space (X, δ) and allowing for a weak inequality, (X, δ) is said to satisfy the semi-parallelogram law given by (35). It is a theorem of P. Jordan and J. von Neumann that each normed space satisfying parallelogram law is Euclidean.

Bruhat-Tits spaces arise from a much larger class of Riemannian manifolds, the Cartan-Hadamard manifolds, which are complete simply connected Riemannian manifold with semi-negative curvature [10]. A Cartan-Hadamard Manifold is contractible (it has the homotopy type of a single point) and, ***between two points, there is a unique geodesic segment***.

The midpoint property of the geometric mean can be established independently of the theory of Bruhat-Tits spaces. Geometric mean can be defined as solution of the Riccati Equation: $XA^{-1}X = B$ or by Frechet definition [22] :

$$\arg\underset{R}{Min} H(R) \text{ with } H(R) = \left\|\log\left(A^{-1/2}.R..A^{-1/2}\right)\right\|^2 + \left\|\log\left(B^{-1/2}.R..B^{-1/2}\right)\right\|^2$$

$$\Rightarrow \ \nabla H(R) = R\left[\log\left(A^{-1/2}.R..A^{-1/2}\right) + \log\left(B^{-1/2}.R..B^{-1/2}\right)\right] = 0$$

The Geometric mean is then given by:

$$A \circ B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{1/2} A^{1/2} \tag{38}$$

with $A^{1/2} = e^{\frac{1}{2}\log A}$ and $e^W = \sum_{n=0}^{\infty} \frac{A^n}{n!}$

This definition satisfy the following properties:

$$\begin{aligned} &A \circ A = A \text{ Idempotence}, \left(A \circ B\right)^{-1} = A^{-1} \circ B^{-1} \text{ Inversion} \\ &A \circ B = B \circ A \text{ Commutativity} \\ &\Gamma_C(A) \circ \Gamma_C(B) = \Gamma_C(A \circ B) \text{ with } \Gamma_C(A) = C^T AC \\ &2\left(A^{-1} + B^{-1}\right)^{-1} \le A \circ B \le \left(A + B\right)/2 \end{aligned} \tag{39}$$

We can also defined the unique geodesic in this space joining the two matrices A and B. If $t \to \gamma(t)$ is the geodesic between A and B , where $t \in [0,1]$ is such that $\delta(A, \gamma(t)) = t.\delta(A,B)$, then the mean of A and B is the matrix $A \circ B = \gamma(1/2)$. The geodesic parameterized by the length as previously is given by::

$$\begin{aligned} &\gamma(t) = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^t A^{1/2} \text{ with } 0 \le t \le 1 \\ &\gamma(0) = A \ , \ \gamma(1) = B \text{ and } \gamma(1/2) = A \circ B \end{aligned} \tag{40}$$

We have seen that for SPD matrices, we have to consider the geometric mean. To convince reader by intuition , that $(R_1 + R_2)/2$ is not a well adapted mean of two SPD matrices R_1 and R_2, consider Multivariate Gaussian model case. We have seen that variance of R is proportional to R, that means to take into account variance of $(R_1 + R_2)/2$, we should consider the inverse matrix that depends on $\det[(R_1 + R_2)/2]^{-1}$ that could have a bad behavior. Inversely, metric proposed by Rao $ds^2 = Tr\left(\left(R^{-1}(dR)\right)^2\right)$ takes into account variance of R to build a robust metric invariant by non singular parameters transformation : $w = \Theta(\theta) \Rightarrow ds_w^2 = ds_\theta^2$

VII. SQUARE ROOT MATRIX

We have seen that we have to compute the geometric mean $A \circ B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{1/2} A^{1/2}$. Then, we have to compute square root of positive definite matrix.

A first approach could be to use the Denman-Beabers iteration for the square root of a matrix A with no negative

eigen-values :

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} 0 & Y_k \\ Z_k & 0 \end{bmatrix} + \begin{bmatrix} 0 & Z_k^{-1} \\ Y_k^{-1} & 0 \end{bmatrix}\right) \text{with } X_0 = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}$$

The iteration has the properties that : $\underset{k\to\infty}{Lim}\, X_k = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}$

To avoid matrix inversion, Schulz iteration can be used :

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2} X_k (3I - X_k^2) \tag{41}$$

We can also used Schur Method for Matrix Square root introduced by Björk [20]. Based on Schur decomposition of A, $Q^+ AQ = T$ (T upper triangular), uses a recurrence to obtain an upper triangular U such that $U^2 = T$. A square root of A is given by : $X = QUQ^+$ (42)

VIII. Geometric means of N SPD matrices

Different authors have tried to extend this notion of Geometric means for a set of N Symmetric Positive definite matrices.

For 3 matrices, D. Petz has proposed a symmetrization procedures by a Cauchy sequences with respect to the geodesic distance δ. The space is complete with respect to this metric and the three sequences (of interleaved triangles) have a common limit points :

$$\begin{aligned} & A \circ B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{1/2} A^{1/2} \\ & \begin{cases} A_1 = A \ , \ B_1 = B \ , \ C_1 = C \\ A_{n+1} = A_n \circ B_n \ , \ B_{n+1} = A_n \circ C_n \ , \ C_{n+1} = B_n \circ C_n \end{cases} \\ & G(A,B,C) = \underset{n\to\infty}{Lim}\, A_n = \underset{n\to\infty}{Lim}\, B_n = \underset{n\to\infty}{Lim}\, C_n \end{aligned} \tag{43}$$

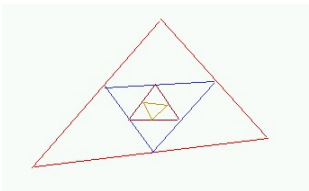


Figure 6 - symmetrization procedures with Cauchy sequences

For the extension to N matrices of geometric means, we can observe that naïve extension cannot be applied due to the following limitations:

1) $(A_1 A_2 ... A_N)^{1/N}$ is not positive definite if A_i don't commute

2) $e^{(\log A_1 + \log A_2)/2}$ is not monotone

because $A \geq B \Rightarrow e^A \geq e^B$ is not true

T. Ando [4] has proposed a definition for geometric mean of N positive definite matrices by mean of an iterative sequence that converges to same limit for all components :

$$G(A_1, A_2) = A_1^{1/2}\left(A_1^{-1/2} A_2 A_1^{-1/2}\right)^{1/2} A_1^{1/2} \tag{44}$$

for k = 3,..., N do

$$\left| \begin{aligned} & \{G((A_i)_{i\neq l})\}_{l=1}^{k+1} \text{ have been iteratively defined} \\ & \begin{cases} A^{(1)} = (A_1, ..., A_{k+1}) \\ A^{(r+1)} = T(A^{(r)}) = \left(G((A_i^{(r)})_{i\neq 1}), ..., G((A_i^{(r)})_{i\neq k+1})\right) \ , \ r = 1,2,... \end{cases} \\ & G(A_1, ..., A_k, A_{k+1}) = \tilde{A} \ \text{ with } \ \underset{r\to\infty}{Lim}\, A^{(r)} = (\tilde{A}, ..., \tilde{A}) \end{aligned} \right.$$

IX. Conclusion

We have introduced a new strategy for radar detection optimization by using "matrix CFAR" acting on SPD covariance matrices of data. For this purpose, we have introduced the Siegel metric and distance by Information Geometry and Geometry of symmetric cones. Notion of "geodesic" and "Geometric mean" between 2 matrices have allowed to extend the tool for the Geometric mean definition of N SPD matrices that will be used for new "matrix CFAR" acting on Radar covariance matrices to replace poor resolution FFT or Doppler Filters approaches. Results on HFSWR radar data are presented in another paper presented in the conference [21]. Approach is validated on real Doppler Radar Data from an HFSWR radar and an X-band Coastal Radar.

References

[1] C.L. Siegel, "Symplectic Geometry", Academic Press, New York, 1964

[2] C.R. Rao, "Information and Accuracy attainable in the estimation of statistical parameters", Bull. Calcutta Math. Soc., n°37, pp.81-91, 1945

[3] S. Yoshizawa & K. Tanabe, "Dual Differential Geometry associated with the Kullback-Leibler Information on the Gaussian Distributions and its 2-parameter Deformations", SUT Journal of Mathematics, vol.35,n°1, pp.113-137, 1999

[4] T.Ando & R. Mathias, "Geometric Means", Linear Algebra Appl., vol.385, pp.305-334, 2004

[5] "Kähler Erich, Mathematical Works", Berlin, Walter de Gruyter, ix, 2003

[6] I. Bakas, "The Algebraic Structure of Geometric Flows in Two Dimensions", Inst. of Physics, SISSA, October 2005

[7] F. Barbaresco, "Information Intrinsic Geometric Flows", MaxEnt'06 Conference, Paris, June 2006, published in American Institute of Physics, n° 872, 2006

[8] F. Barbaresco, "Les 2 sources des traitements critiques radar basés sur la métrique de C.L. Siegel", THALES-SMAI Workshop on ISTAR, http://smai.emath.fr/breve.php3?id_breve=49 ,Oct. 2007

[9] F. Bruhat & J. Tits, « Groupes réductifs sur un corps local », IHES, n°41, pp.5-251, 1972

[10] M. Gromov, "Hyperbolic Groups", Essays in Group Theory, Math. Sci. Res. Inst. Publ. 8, New york, pp.75-263, 1987

[11] H. Maass, "Siegel Modular Forms and Derichlet Series", Lecture Notes in Math., n° 216, Springer-Verlag, Berlin, 1971

[12] H. Cartan, "Ouverts fondamentaux pour le groupe modulaire",Séminaire Henri Cartan, tome n°10, n°1, exp.n°3, p.1-12, 1957

[13] N.N. Chentsov, "Statistical Decision Rules and Optimal Inferences", Trans. of Math. Monog., n°53, Amer. Math. Society, Providence, 1982

[14] E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes", Abh. Math. Semin. hamb. Univ., n°11, pp.116-162, 1935

[15] M. Koecher, "Jordan Algebras and their Applications", Lect. Notes, Univ. of Minnesota, Minneapolis, 1962

[16] I. Satake, "Algebraic Structures of Symmetric Domains", Kano memorial Lectures, n°4, Princeton University Press, 1980

[17] J. Faraut & A. Koranyi, "Analysis on Symmetric Cones", Oxford University Press, 1994

[18] P. Bougerol, "Kalman Filtering with Random Coefficients and Contractions", SIAM J. Control and Optimization, vol.31, n°4 pp.942-959, Juillet 1993

[19] K. Koufany, "Analyse et Géométrie des domaines bornés symétriques", HDR, Institut de Mathematiques Elie Cartan, Nancy, Nov. 2006

[20] A. Björk & S. Hammarling, "A Schur method for the square root of a matrix", Linear Algebra and Appl., n°52/53,pp.127-140, 1983

[21] J.Lapuyade-Lahorgue & F. Barbaresco, "Innovative CFAR detection using Siegel Metric in the case of Gaussian random distributions", IEEE Radar Conference, Rome, May 2008

[22] H. Karcher, "Riemannian center of mass and mollifier smoothing", Comm. Pure Applied Math., °30,pp.509-541,1977