

Une nouvelle approche pour la détection de cibles dans les images radar basée sur des distances et moyennes dans des espaces de matrices de covariance

A new approach for target detection in radar images based on geometric properties of covariance matrices' spaces

Nicolas Charon¹, Frédéric Barbaresco²

¹CMLA, Ecole Normale Supérieure de Cachan, 61 avenue du président Wilson, 94235 Cachan
Nicolas.Charon@cmla.ens-cachan.fr

²Thales Air Systems, Département Stratégie Technologie et Innovation, Hameau de Roussigny, 91470 Limours
frederic.barbaresco@thalesgroup.com

Manuscrit reçu le 25 novembre 2008



Résumé et mots clés

La présente étude vise à introduire une toute nouvelle classe d'algorithmes en vue de l'application à la détection automatique des cibles sur les images radar dans lesquelles on dispose d'un faible nombre d'échantillons par salve d'émission. Contrairement aux méthodes déjà mises en œuvre (filtrage TFAC, etc...) qui agissent directement sur les données radar, celle exposée ici consiste à calculer des matrices de covariance à partir de ces données puis de chercher en quels points une matrice de covariance diffère de la matrice « moyenne » calculée dans un voisinage. On espère ainsi obtenir des pics de détection plus prononcés au niveau des cibles qu'avec les algorithmes standards. Ceci impose dans un premier temps l'exploitation de résultats déjà mis au point en géométrie différentielle pour définir distances et moyennes dans de tels espaces (cf. [5], [11], [16], [18]) ce que nous récapitulerons dans les sections 2 et 3. L'algorithme à proprement dit ainsi que quelques résultats de détection par cette nouvelle approche sont ensuite présentés dans la section 4 et comparés aux algorithmes classiques. Enfin, dans la section 5, on verra une amélioration possible des méthodes basées sur les matrices de covariance notamment en terme de temps de calcul, grâce à l'utilisation des processus auto-régressifs ([7], [10]).

Images radar, détection, géométrie de l'information, espaces riemanniens, moyennes géométriques.

Abstract and key words

The following paper aims at presenting new theoretical and algorithmic developments to the problem of target detection in radar imaging. This concept has been explored very recently by the authors as a possible alternative to the standard Fourier transform-based algorithm called TFAC. Indeed, TFAC proves inaccurate in cases where the number of impulses

per emission salvo in the radar image becomes small. To avoid lack of resolution of the Fourier transform, we propose an entirely different approach that involves covariance structures of the data instead of raw data themselves. Working with covariance matrices, as defined by (1), of the complex vectors that represents the radar data has already appeared useful in some other issues related to radar image processing, which motivated their use in this particular problem. The basic idea of the algorithm is identical to other detection schemes: we want to compare a given location on the radar image with its neighbourhood and detect a target as a point where radar data is highly different from the average of its closest neighbours. The difference in our setting comes from the fact that we try to evaluate it through these covariance matrices defined at each point. Assuming that we can derive them from the data, the problem becomes one of finding and computing adapted distances and averages of such objects. This involves recent results in differential geometric studies of hermitian positive definite matrices' space and is developed in section 2 and 3. The metric on $HPD_n(\mathbb{C})$ needs to satisfy some invariances due to the special structure of covariance matrices, in particular invariance by basis changes which imposes condition (2). This excludes all the usual distances we know like for instance the Frobenius norm. A pleasant framework to derive a good distance is the one of Riemannian geometry. Considering the manifold $HPD_n(\mathbb{C})$, we define a metric on the identity tangent space which we transport to the whole space according to the invariances we want (4). Taking the most simple metric on the identity tangent space, ie the usual euclidean metric, we have defined a riemannian metric satisfying the conditions. Another good point comes from the fact that we have an explicit parametrisation of the geodesics (9) and therefore, integrating along them, the expression of the distance as in (7) and (8). It is very remarkable that the resulting distance is actually the Siegel distance in $HPD_n(\mathbb{C})$, which is also Rao's distance measure between zero-mean Gaussian distributions in information geometry.

Once the metric is defined, there is a canonical way to express the mean of several covariance matrices so that we preserve the same invariances. It is given by equation (10). Finding the mean then requires the resolution of a minimization problem to which we cannot express the solution in the general case. This is solved by a gradient descent scheme made possible both by the expression of the geodesics and the gradient. Computation steps are given by the recursive relations (11). Eventually, the defined mean on $HPD_n(\mathbb{C})$ has many interesting properties summarized in theorem 3.1, which makes it more appropriate to our applications than other definitions that can be found in literature. In section 4, we give a first set of results on real images provided by a coastal radar. Starting with raw data, we need to estimate covariance matrices on each radar location. This is done classically by considering auto-regressive models. Since it is not the point of the article, we simply refer to the papers cited about the subject ([7], [10]). Once calculated, we can apply all the previous framework on the covariance matrices. For detection puposes, we compute at each location the mean of the adjacent matrices and compare it to the current covariance matrix through the Siegel distance. In figure 1, this distance is represented function to the point's distance to the radar. Several peaks corresponding to special objects appear on the graph and can be compared to the equivalent peaks obtained by the usual TFAC scheme (figure 2), clearly showing a far better detection efficiency. This result is highlighted when artificial targets are inserted on the initial image (cf. figures 3 and 4).

Finally, in section 5, we describe an improved version which was considered in order to reduce computational cost of the algorithm. The strategy consists in the use of a more compact representation of the data given directly by reflexion coefficients of the auto-regressive model. Again, a metric and a mean are defined on the new space following the ideas of information geometry, leading to equations (13) and (14). As shown on figures 6 and 7, detection quality remains similar while computation time is notably improved by a factor 4 within this new scheme. Although, as can be imagined, results are still perfectible, we have presented several new ideas that, in our sense, could pave the way for an interesting approach to radar image processing in general.

Radar imaging, detection, information geometry, riemannian spaces, geometric means.

Notations

Au fil du texte, on utilisera les notations suivantes :

- D : disque unité ouvert de \mathbb{C} .
- $M_n(\mathbb{K})$: ensemble des matrices de taille (n, n) à coefficients dans le corps \mathbb{K} .
- M^T : transposée de la matrice M .
- M^* : transconjuguée de la matrice complexe M .
- $Sym_n(\mathbb{R})$: ensemble des matrices symétriques réelles.
- $PD_n(\mathbb{R})$: ensemble des matrices symétriques définies positives.
- $Herm_n(\mathbb{C})$: ensemble des matrices hermitiennes complexes.
- $HPD_n(\mathbb{C})$: ensemble des matrices hermitiennes définies positives.
- \exp : exponentielle matricielle définie sur $M_n(\mathbb{C})$; c'est en particulier un C^∞ -difféomorphisme de $Herm_n(\mathbb{C})$ sur $HPD_n(\mathbb{C})$.
- \log : logarithme matriciel principal, défini comme la réciproque de l'exponentielle pour toute matrice n'ayant aucune valeur propre sur l'axe réel négatif \mathbb{R}_- (c'est en particulier le cas pour $M \in HPD_n(\mathbb{C})$).

1. Introduction

Avant de rentrer véritablement dans le vif du sujet, présentons brièvement la nature des données radar dont il sera question dans les sections 4 et 5. Celles-ci sont représentées par une matrice de nombres complexes dont chaque pixel (aussi appelé case radar) correspond à une distance et un azimuth précis par rapport à l'antenne émettrice. Le radar, en tournant sur lui-même, émet des salves de n impulsions. Pour chaque case radar, on obtient ainsi un vecteur de nombres complexes $Z = (z_1, \dots, z_n)^T$ où n est le nombre d'impulsions émises dans une salve. Le module de chacun de ces nombres complexes donne une information sur la réflectivité de l'objet alors que les déphasages sont directement liés à sa vitesse radiale.

L'objectif est maintenant de pouvoir détecter de façon algorithmique une « cible » sur l'image. Une cible étant bien sûr un objet dont le comportement notamment en vitesse diffère de son environnement, l'approche standard (baptisée TFAC) consiste à travailler sur les transformées de Fourier des vecteurs Z et de là chercher à quels endroits la valeur en une case est éloignée de la valeur moyenne calculée sur les cases adjacentes. Cette méthode donne d'assez bons résultats. Cependant, lorsque le nombre d'impulsions par salve est assez faible (de l'ordre de 8 pour les données utilisées dans cet article), la transformée de Fourier de Z souffre d'un manque de résolution évident. C'est ce défaut qu'on aimerait en partie corriger pour espérer aboutir à de meilleurs résultats. L'idée sur laquelle se focalisera l'essentiel de cet article consiste à travailler sur une structure a priori plus complexe que les données « brutes ». Plutôt que de considérer en chaque case radar le vecteur Z , on va s'intéresser à sa matrice de covariance Σ_n définie par :

$$\Sigma_n = E(ZZ^*) = E \begin{pmatrix} |z_1|^2 & z_1 \bar{z}_2 & \dots & z_1 \bar{z}_n \\ z_2 \bar{z}_1 & |z_2|^2 & \dots & z_2 \bar{z}_n \\ \vdots & \ddots & \ddots & \vdots \\ z_n \bar{z}_1 & \dots & z_n \bar{z}_{n-1} & |z_n|^2 \end{pmatrix} \quad (1)$$

Sans hypothèse supplémentaire, il est assez facile de voir que Σ_n est une matrice hermitienne définie positive. On a donc changé d'espace d'étude de l'espace classique \mathbb{C}^n à $HPD_n(\mathbb{C})$. En laissant pour le moment le problème de l'estimation de la matrice Σ_n directement à partir des données, qu'on évoquera plus en détail dans la section 4, on va en fait adopter une démarche très semblable à celle qu'on vient d'évoquer, à savoir détecter des cibles comme des points où la matrice de covariance Σ_n diffère grandement de la matrice « moyenne » sur les cases adjacentes. Ceci suppose avant tout deux choses : savoir définir la distance sur $HPD_n(\mathbb{C})$ et la moyenne. Ces problèmes ont donné lieu à de multiples travaux dans des domaines aussi variés que la géométrie différentielle, les probabilités et la géométrie de l'information. Parmi les différentes méthodes à disposition, nous exposons dans les deux sections qui suivent celles que nous avons retenues en vue de la détection de cibles.

2. Définition d'une métrique appropriée sur l'espace des matrices hermitiennes définies positives

L'objet de la section qui suit est d'exhiber une distance sur l'espace d'étude qui sera désormais le nôtre, à savoir l'espace des matrices hermitiennes définies positives noté $HPD_n(\mathbb{C})$. Cette distance doit cependant vérifier un certain nombre de propriétés d'invariance liées à la structure des matrices de covariance. En effet, un premier élément crucial est le fait que la distance entre deux matrices de covariances Σ_1 et Σ_2 des vecteurs complexes Z_1 et Z_2 ne doit bien sûr pas dépendre de la base choisie pour représenter ces deux vecteurs. Ceci se traduit par la relation :

$$\forall P \in GL_n(\mathbb{C}), d(P\Sigma_1P^*, P\Sigma_2P^*) = d(\Sigma_1, \Sigma_2) \quad (2)$$

Considérons par exemple la métrique la plus simple à laquelle on puisse penser : la distance de Frobenius qu'on restreint à $HPD_n(\mathbb{C})$. Rappelons que la norme de Frobenius dans $M_n(\mathbb{C})$ est définie par $\|A\|_F = \sqrt{\text{Tr}(A^*A)}$. On peut vérifier facilement que cette distance jouit d'un certain nombre d'invariances, en particulier :

$$\forall U \in U_n(\mathbb{C}), d(U\Sigma_1U^*, U\Sigma_2U^*) = d(\Sigma_1, \Sigma_2)$$

On n'obtient pas toutefois la totalité des invariants donnés par l'équation (2), ce qui justifie l'utilisation de métriques plus complexes dans le problème en question. Outre (2), on veut imposer à la métrique d'être en plus invariante par inversion matricielle pour des raisons qui apparaîtront plus clairement dans la section suivante, ce qui s'écrit :

$$d(\Sigma_1^{-1}, \Sigma_2^{-1}) = d(\Sigma_1, \Sigma_2) \quad (3)$$

Pour répondre à ces deux critères, on cherche une métrique du type métrique riemannienne sur $HPD_n(\mathbb{C})$. Rappelons qu'en effet, $HPD_n(\mathbb{C})$ étant un ouvert de l'espace vectoriel $Herm_n(\mathbb{C})$, il est naturellement muni d'une structure de variété de dimension $n(n+1)/2$. Se donner une métrique riemannienne sur cette variété revient ainsi à déterminer en chaque point p de $HPD_n(\mathbb{C})$ un produit scalaire noté $\langle \cdot, \cdot \rangle_p$ sur l'espace tangent T_p .

On peut commencer par simplifier le problème en ramenant l'étude de $HPD_n(\mathbb{C})$ à celle de l'espace réel $PD_n(\mathbb{R})$, ceci grâce au fait qu'il existe un morphisme d'algèbre injectif de $GL_n(\mathbb{C})$ dans $GL_{2n}(\mathbb{R})$ donné par :

$$\Psi : GL_n(\mathbb{C}) \rightarrow GL_{2n}(\mathbb{R})$$

$$A + iB \rightarrow \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$



Ce morphisme a notamment la bonne propriété d'envoyer une matrice hermitienne définie positive vers une matrice symétrique définie positive ce qui permet de transposer chacun des résultats ci-dessous d'un espace vers l'autre. On est donc ramené maintenant à la recherche d'une métrique sur $PD_n(\mathbb{R})$ tel que le groupe G engendré par les transformations $\Sigma \rightarrow \Sigma^{-1}$ et $\Sigma \rightarrow P^T \Sigma P$ soit inclus dans le groupe des isométries pour cette métrique.

Plusieurs approches tout à fait différentes permettent en fait de répondre à la question. Une première est présentée notamment dans les articles [5] et [11] auxquels nous invitons le lecteur à se référer. Cette méthode générale permet de définir, à partir d'une distance entre des distributions de probabilités paramétrées, une métrique sur l'espace des paramètres via la matrice de Fisher. Appliquée au cas de distributions de probabilité gaussiennes à moyenne fixée, on obtient une métrique riemannienne sur $PD_n(\mathbb{R})$ dont on peut déterminer ensuite l'expression de la distance entre deux points et vérifier à posteriori les propriétés d'invariance voulues.

Dans une tout autre optique, on peut retrouver la même métrique lorsqu'on restreint la métrique de Siegel de la géométrie symplectique à l'espace $PD_n(\mathbb{R})$. À ce titre, on pourra se reporter à l'article historique de Siegel [21]. Il est tout à fait remarquable que les deux approches donnent en définitive le même résultat.

Dans ce qui suit, nous introduisons toutefois une troisième méthode pour déterminer la métrique en question qui présente l'avantage d'être plus constructive et plus directe vis à vis du problème qui nous concerne. On a vu que $PD_n(\mathbb{R})$ est une variété dont l'espace tangent en chaque point est isomorphe à

$Sym_n(\mathbb{R})$. Définissons comme précédemment le groupe G comme le groupe engendré par $\Theta : \Sigma \rightarrow \Sigma^{-1}$ et l'ensemble des transformations $\Phi_P : \Sigma \rightarrow P \Sigma P^T$ pour $P \in GL_n(\mathbb{R})$. G agit sur $PD_n(\mathbb{R})$ et l'action est de plus transitive car pour tout $\Sigma \in PD_n(\mathbb{R})$ s'écrit $\Sigma = \Sigma^{1/2} (\Sigma^{1/2})^T = \Phi_{\Sigma^{1/2}}(Id)$. En raisonnant par conditions nécessaires, si tous les éléments de G sont supposés être des isométries pour la métrique riemannienne alors la donnée du produit scalaire en l'identité suffit à déterminer la métrique sur l'espace entier. En effet, le fait que $\Phi_{\Sigma^{1/2}}$ soit une isométrie se traduit par l'équation :

$$\forall A, B \in T_\Sigma, \quad \langle A, B \rangle_\Sigma = \langle D\phi_\Sigma(A), D\phi_\Sigma(B) \rangle_{Id}$$

soit encore :

$$\langle A, B \rangle_\Sigma = \langle \Sigma^{-1/2} A \Sigma^{-1/2}, \Sigma^{-1/2} B \Sigma^{-1/2} \rangle_{Id} \quad (4)$$

Il faut cependant faire attention au fait que, pour avoir un sens bien défini, l'expression précédente ne doit pas dépendre de la racine carrée de Σ qu'on considère. En effet, il n'y a pas unicité de la matrice $P \in GL_n(\mathbb{R})$ telle que $P P^T = \Sigma$. Il faut donc que tous les membres de droite de l'équation (4) coïncident pour les diverses racines carrées. Or on sait que toutes les matrices P satisfaisant la relation précédente se déduisent les unes des autres par multiplication à droite par une matrice orthogonale, c'est-à-dire

$$P P^T = Q Q^T = \Sigma \Rightarrow \exists U \in O_n(\mathbb{R}), \quad Q = P U$$

Ceci impose la condition suivante sur le produit scalaire en l'identité :

$$\forall U \in O_n(\mathbb{R}), \forall A, B \in Sym_n(\mathbb{R}),$$

$$\langle U A U^T, U B U^T \rangle_{Id} = \langle A, B \rangle_{Id} \quad (5)$$

Prenons maintenant un produit scalaire en l'identité qui satisfasse la condition (5) (condition d'invariance par changement de base orthonormée). On peut étendre ce produit scalaire sur $PD_n(\mathbb{R})$ sans ambiguïté par la relation (4). La métrique riemannienne qu'on obtient admet alors toutes les applications Φ_P comme isométries, ce qu'on peut vérifier aisément à partir de ce qui a été dit. Mais il y a mieux : on obtient de surcroît l'invariance par inversion matricielle. En effet, si $\Sigma \in PD_n(\mathbb{R})$ alors on a :

$$\langle D\Theta_{\Sigma^{-1}}(A), D\Theta_\Sigma(B) \rangle_\Sigma = \langle \Sigma^{-1} A \Sigma^{-1}, \Sigma^{-1} B \Sigma^{-1} \rangle_{\Sigma^{-1}}$$

$$= \langle \Sigma^{-1/2} A \Sigma^{-1/2}, \Sigma^{-1/2} B \Sigma^{-1/2} \rangle_{Id}$$

$$= \langle A, B \rangle_\Sigma$$

Ceci implique que ψ est bien une isométrie. Il en résulte que tous les éléments du groupe G sont des isométries pour la métrique riemannienne sur $PD_n(\mathbb{R})$ qu'on a construite. Au final, on peut construire une telle métrique simplement en se donnant un produit scalaire en l'identité qui vérifie la condition (5). C'est en particulier le cas si l'on prend le produit scalaire euclidien habituel : $\langle A, B \rangle_{Id} = Tr(AB^T)$. Le produit scalaire en Σ

qu'on obtient avec (4) devient alors :

$$\langle A, B \rangle_{\Sigma} = \text{Tr}(\Sigma^{-1/2} A \Sigma^{-1} B \Sigma^{-1/2}) = \text{Tr}(\Sigma^{-1} A \Sigma^{-1} B) \quad (6)$$

C'est précisément la métrique de Siegel. Celle-ci peut s'interpréter comme la métrique euclidienne en l'identité qu'on projette à l'espace $HPD_n(\mathbb{R})$ par les isométries qu'on a évoquées. Ce que cette analyse montre en supplément, c'est qu'on peut en fait définir d'autres métriques invariantes par les mêmes transformations. La seule condition à vérifier porte sur le produit scalaire en l'identité et se résume aux relations (5). La métrique de Siegel est toutefois celle qui présente la structure la plus simple, ce qui explique pourquoi on la privilégie.

À partir de la métrique locale donnée par (6), on peut déterminer tous les outils habituels de la géométrie riemannienne : carte exponentielle, géodésiques et expression de la distance entre deux éléments. Nous renvoyons à [5] pour de plus amples détails. Transposés au cas des matrices hermitiennes définies positives grâce à l'application Ψ introduite plus haut, on peut obtenir l'expression suivante de la distance entre Σ_1 et Σ_2 :

$$d(\Sigma_1, \Sigma_2) = \left\| \log((\Sigma_1^{-1/2})^* \Sigma_2 \Sigma_1^{-1/2}) \right\|_F \quad (7)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius et la notation $\Sigma^{1/2}$ une matrice vérifiant la relation $A^* A = \Sigma$. On notera que l'expression précédente ne dépend pas des racines carrées choisies. Une autre expression plus habituelle de d est donnée par :

$$d(\Sigma_1, \Sigma_2) = \left(\sum_{k=1}^n \log^2 \lambda_k \right)^{1/2} \quad (8)$$

les λ_k représentant les valeurs propres généralisées de Σ_1 et Σ_2 , ie les solutions de l'équation $\det(\lambda \Sigma_1 - \Sigma_2) = 0$. Avec les mêmes conventions, on dispose de l'équation suivante du segment de géodésique joignant Σ_1 et Σ_2 :

$$\Sigma(t) = (\Sigma_1^{1/2})^* ((\Sigma_1^{-1/2})^* \Sigma_2 \Sigma_1^{-1/2})^t \Sigma_1^{1/2}, \quad t \in [0, 1]. \quad (9)$$

La métrique ainsi définie, outre le fait qu'elle respecte les invariances requises, donne finalement une expression simple et facilement implémentable de la distance et des géodésiques.

3. Moyenne géométrique de matrices hermitiennes définies positives

Maintenant que l'on sait mesurer l'écart entre différentes matrices de covariance par la distance de Siegel, l'objectif est de pouvoir définir une moyenne entre de tels objets qui une fois

encore respecte leur structure. C'est un sujet qui a été plusieurs fois abordé. Dans [2] par exemple, la moyenne de N matrices de $HPD_n(\mathbb{C})$ est définie de façon récursive. La moyenne obtenue présente un grand nombre d'invariances mais s'avère trop coûteuse pour les applications en terme de temps de calcul lorsque N est trop grand. Dans [4], est définie la moyenne dite « log-exponentielle » très rapidement calculable mais n'ayant pas toutes les propriétés d'invariance voulues. L'approche que nous présentons ici est l'analogie de la moyenne de réels positifs : on pourra consulter en particulier [18] pour davantage de détails. Dans le cas des réels, on peut en effet obtenir différentes moyennes de N points $(x_i)_{i=1 \dots N} \in \mathbb{R}_+^*$ en recherchant :

$$\inf_{x \in \mathbb{R}_+^*} \sum_{i=1}^N d(x, x_i)^2$$

Lorsque d désigne la distance euclidienne classique, l'infimum est atteint lorsque x est égal à la moyenne arithmétique des (x_i) . Si d est la distance hyperbolique sur \mathbb{R}_+^* , on tombe sur la moyenne géométrique. C'est à partir de ce constat qu'on définit la moyenne dite « géométrique » de N matrices de $HPD_n(\mathbb{C})$ par :

$$\text{Moy}(\Sigma_1, \dots, \Sigma_N) = \arg \inf_{\Sigma \in HPD_n(\mathbb{C})} \sum_{i=1}^N d(\Sigma, \Sigma_i)^2 \quad (10)$$

où d désigne cette fois-ci la distance de Siegel qu'on a définie dans la section précédente. Se pose alors un premier problème délicat d'existence et d'unicité d'une telle moyenne. La réponse est donné grâce au fait que l'espace $HPD_n(\mathbb{C})$ muni de la métrique de Siegel est une variété de Cartan-Hadamard (cf. [19]) dans laquelle on a donc unicité de la géodésique de longueur minimale joignant deux points quelconques (c'est la courbe donnée par l'équation (9)). On trouvera alors dans [16] une justification d'existence et d'unicité de la moyenne géométrique.

En vue d'applications numériques, il est également indispensable de pouvoir calculer la moyenne. Contrairement au cas réel, on ne dispose pas d'une expression explicite générale. On doit se résoudre à une méthode de descente de gradient dans la variété riemannienne $HPD_n(\mathbb{C})$. Pour les applications ultérieures, il nous est apparu qu'une méthode de gradient à pas constant $\varepsilon > 0$ s'avérait suffisante. Notons $H(P) = \sum_{i=1}^N d(P, \Sigma_i)^2$. La méthode itérative consiste à calculer le gradient de H au point courant au sens de la variété $HPD_n(\mathbb{C})$ et à se déplacer le long de la géodésique dirigée par ce gradient. Après quelques calculs, on obtient l'expression suivante :

$$\nabla H(P) = P \sum_{i=1}^N \log(\Sigma_i^{-1} P)$$

Fixant un point de départ quelconque $S_0 \in HPD_n(\mathbb{C})$, et grâce à l'expression des géodésiques, on obtient les relations de récurrence qui suivent pour la descente de gradient :

$$\begin{aligned}
 S_{n+1} &= (S_n^{1/2})^* \exp\left(-\varepsilon \sum_{i=1}^N \log(S_n^{1/2} \Sigma_i^{-1} (S_n^{1/2})^*)\right) S_n^{1/2} \\
 &= (S_n^{1/2})^* \exp\left(\varepsilon \sum_{i=1}^N \log((S_n^{-1/2})^* \Sigma_i S_n^{-1/2})\right) S_n^{1/2}
 \end{aligned} \tag{11}$$

où la notation $S_n^{1/2}$ désigne, comme dans la section précédente, une matrice quelconque vérifiant $AA^* = S_n$ qu'on prend généralement égale à l'unique racine carrée hermitienne définie positive de S_n pour simplifier les expressions. On arrête ensuite l'algorithme lorsque la distance entre deux itérés successifs passe en deçà d'un certain seuil de précision.

La moyenne précédente vérifie t'elle les propriétés qu'on est en droit d'attendre d'elle ? La réponse est donnée dans le théorème suivant :

Théorème 3.1 *La moyenne géométrique donnée par (10) vérifie les six propriétés suivantes :*

P1 : si $\Sigma_1, \dots, \Sigma_N$ commutent deux à deux alors :

$$\text{Moy}(\Sigma_1, \dots, \Sigma_N) = (\Sigma_1 \dots \Sigma_N)^{1/N}$$

P2 :

$$\text{Moy}(\lambda_1 \Sigma_1, \dots, \lambda_N \Sigma_N) = (\lambda_1 \dots \lambda_N)^{1/N} \text{Moy}(\Sigma_1, \dots, \Sigma_N),$$

$$\forall \lambda_1, \dots, \lambda_N > 0$$

P3 : pour toute permutation σ de $1, \dots, N$, on a :

$$\text{Moy}(\Sigma_{\sigma(1)}, \dots, \Sigma_{\sigma(N)}) = \text{Moy}(\Sigma_1, \dots, \Sigma_N)$$

P4 : pour toute matrice $A \in GL_n(\mathbb{C})$,

$$\text{Moy}(A \Sigma_1 A^*, \dots, A \Sigma_N A^*) = A \text{Moy}(\Sigma_1, \dots, \Sigma_N) A^*$$

P5 : $\text{Moy}(\Sigma_1^{-1}, \dots, \Sigma_N^{-1}) = \text{Moy}(\Sigma_1, \dots, \Sigma_N)^{-1}$

P6 : $\det(\text{Moy}(\Sigma_1, \dots, \Sigma_N)) = (\det(\Sigma_1) \dots \det(\Sigma_N))^{1/N}$

Preuve : Du fait que, dans le calcul de la moyenne, on minimise une fonction symétrique en les Σ_i , il résulte immédiatement la propriété P3. Par ailleurs, l'invariance de la distance d par les applications $\Sigma \rightarrow \Sigma^{-1}$ et $\Sigma \rightarrow A \Sigma A^*$ pour $A \in GL_n(\mathbb{C})$ donne directement les propriétés P4 et P5. Pour les trois autres propriétés, on utilise la caractérisation suivante de la moyenne notée M des $(\Sigma_i)_{i=1 \dots N}$:

$$\sum_{i=1}^N \log(\Sigma_i^{-1} M) = 0 \tag{12}$$

Cette dernière équation exprime simplement le fait qu'en M , le gradient de H est nul. On vérifie alors aisément P2 à partir de cette caractérisation. En prenant la trace de (12) et en se rappelant que $\text{Tr}(\log(X)) = \log(\det(X))$:

$$\begin{aligned}
 0 &= \sum_{i=1}^N \log(\det(\Sigma_i^{-1} M)) \\
 &= \sum_{i=1}^N \log(\det(\Sigma_i)^{-1} \det(M)) \\
 &= \log\left(\left(\prod_{i=1}^N \det(\Sigma_i)\right)^{-1} \det(M)^N\right)
 \end{aligned}$$

d'où $\det(M) = (\det(\Sigma_1) \dots \det(\Sigma_N))^{1/N}$, ce qui donne P6. Enfin, pour la propriété P1, on utilise le fait que l'algorithme de gradient à pas constant converge vers la moyenne quelle que soit l'initialisation $S_0 \in \text{HPD}_n(\mathbb{C})$ (dont on trouvera la preuve à nouveau dans [16]). En particulier, en supposant que les Σ_i commutent deux à deux, si $S_0 = \Sigma_1$, on obtient facilement, grâce à la relation de récurrence, que S_n commute avec Σ_i pour tout n . En passant à la limite et par continuité du produit matriciel, on en déduit que M commute avec chacune des Σ_i et (12) devient alors :

$$\log\left(\left(\prod_{i=1}^N \Sigma_i\right)^{-1} P^N\right) = 0$$

puis en prenant l'exponentielle matricielle $P = (\Sigma_1 \dots \Sigma_N)^{1/N}$. Qu'expriment ces diverses propriétés ? P1 et P2 justifient le nom de moyenne « géométrique » puisque, d'une certaine manière, on généralise la moyenne géométrique dans \mathbb{R}_+^* . Avec P4 et P5, on retrouve l'invariance de la moyenne par le groupe des isométries de la métrique de Siegel qu'on a introduit dans la section précédente. Ces deux propriétés se transfèrent naturellement de la distance à la moyenne compte tenu de la manière dont on a défini celle-ci via (10). Enfin, la propriété P6 est également digne d'intérêt dans la mesure où elle montre que le déterminant de la moyenne reste confiné entre le minimum et le maximum des déterminants des Σ_i . C'est en particulier ce qui fait défaut à la moyenne arithmétique dans $\text{HPD}_n(\mathbb{C})$.

Quant au calcul numérique de la moyenne, bien qu'on ne dispose pas d'une expression explicite, la méthode de descente de gradient à pas constant converge en un nombre relativement peu élevé d'itérations (expérimentalement, on peut constater qu'il y a convergence en moins d'une dizaine d'itérations lorsqu'on calcule des moyennes de 8 éléments avec un seuil d'arrêt à 10^{-3}). Ceci reste donc satisfaisant en vue des applications qui feront l'objet de la suite de cet article.

4. Application aux données radar

Ayant maintenant à disposition une distance entre matrices de covariance ainsi qu'un algorithme pour construire la moyenne de tels objets, nous présentons plus avant le schéma algorithmique mis en place dans le cas de l'imagerie radar que nous avons évoqué en introduction. Afin d'utiliser les outils des sections 2 et 3, il faut pouvoir estimer de façon robuste les matrices de covariance à partir des données brutes. À ce titre, on introduit généralement des hypothèses supplémentaires a priori sur les données en supposant que le vecteur de complexes pour chaque case radar peut se décrire par un modèle auto-régressif. Nous renvoyons aux articles [7] et [10] pour les définitions précises de ce type de modèles. Le point important, en ce qui nous

concerne, tient au fait que l'on peut estimer de manière fiable les coefficients auto-régressifs à partir des données. C'est l'objet des algorithmes récurrents de Levinson et Burg qu'on trouvera exposés en détail dans les articles précités. Ces mêmes coefficients auto-régressifs permettent à leur tour d'estimer en chaque case radar la matrice de covariance recherchée. Il est à noter que, sous l'hypothèse auto-régressive, cette matrice est non seulement hermitienne définie positive mais est également une matrice de Toeplitz.

En combinant les algorithmes de Levinson et Burg avec le calcul des distances et moyennes de matrices de covariance, on peut calculer, en chaque case radar, la distance entre la matrice de covariance courante et la moyenne des matrices provenant des cases adjacentes. Ceci traduit la différence de comportement entre un objet et son environnement proche. En figure 1, on a représenté ce qu'on obtient pour un échantillon de données provenant d'un radar côtier et constitué de cases allant jusqu'à 150 km de distance avec 8 impulsions par salve. La moyenne glissante est effectuée sur cinq cases adjacentes. On peut observer l'existence de pics de détection s'élevant au dessus du bruit, correspondant à des « cibles » (par exemple navires, ...). À titre de comparaison, nous donnons en figure 2 le même type de graphique, obtenu lorsqu'on utilise la méthode standard de détection dite « TFAC ». Celle-ci repose toujours sur l'idée de moyenne et de distance entre cases radar mais en travaillant plus directement sur les transformées de Fourier des données. Bien que plus rapide, cette méthode s'avère peu efficace lorsqu'on dispose d'un faible nombre d'échantillons par salves. C'est ce qu'on peut constater sur la figure. Les pics de détection se confondent avec le bruit.

Afin de mieux mettre cela en évidence, on a rajouté, dans l'image radar précédente, une cible artificielle à une distance de 110 km. En figure 3, correspondant à l'algorithme par matrices de covariance, on voit apparaître un pic très marqué à l'emplacement de la cible. Dans la méthode TFAC (figure 4), le pic en question se distingue à nouveau très peu du bruit. En définitive, travailler avec les matrices de covariance permet d'augmenter

significativement le rapport signal sur bruit. Ceci est dû à l'utilisation de distances et moyennes adaptées à la structure des matrices de covariance. À titre de comparaison, on peut voir en figure 5 le même schéma algorithmique appliqué dans le cas précédent mais en représentant cette fois-ci la distance de Frobenius au lieu de celle de Siegel. La réponse obtenue au niveau de la cible synthétique est peu marquée, compte tenu du fait que les distances de Frobenius prennent des valeurs très variables et ont tendance à exploser en certains points, rendant de surcroît la définition d'un seuil de détection particulièrement délicate.

Malgré une efficacité de détection tout à fait satisfaisante, l'ensemble du schéma algorithmique en question (calcul des matrices de covariance, moyenne glissante et distance de Siegel) présente l'inconvénient de nécessiter un temps de calcul beaucoup plus important que TFAC. À ce stade, il paraît vraisemblable que ceci puisse être amélioré si l'on en vient à exploiter davantage la structure autorégressive des données et pas seulement la structure hermitienne définie positive des matrices de covariance, ce dont il est question dans la suite.

S

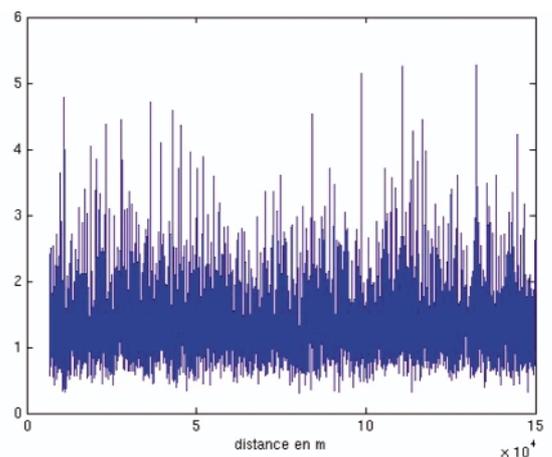


Figure 2. Résultats de l'algorithme de détection de cibles par FFT.

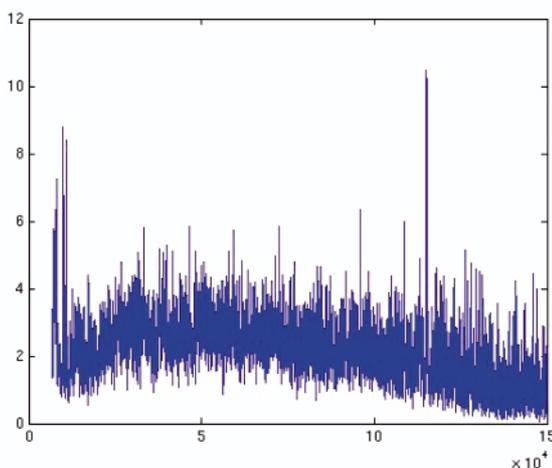


Figure 1. Résultats de l'algorithme de détection de cibles via les matrices de covariance.

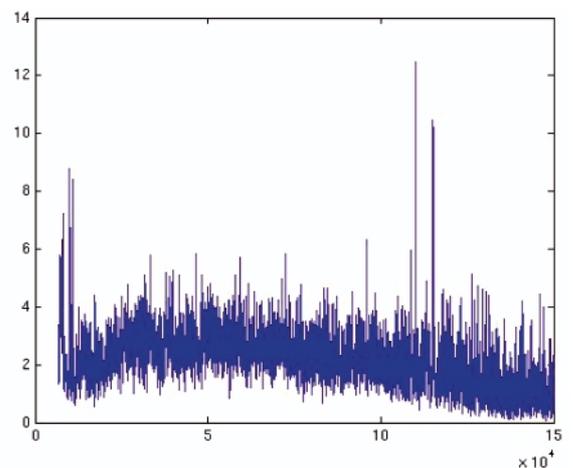


Figure 3. Ajout d'une cible synthétique dans les données initiales : détection par la nouvelle méthode.

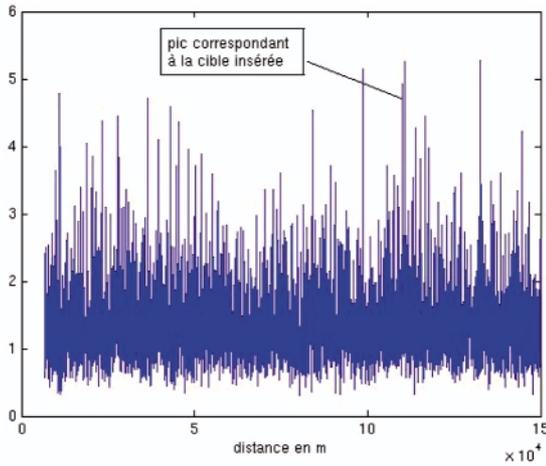


Figure 4. Ajout d'une cible synthétique dans les données initiales : détection par FFT.

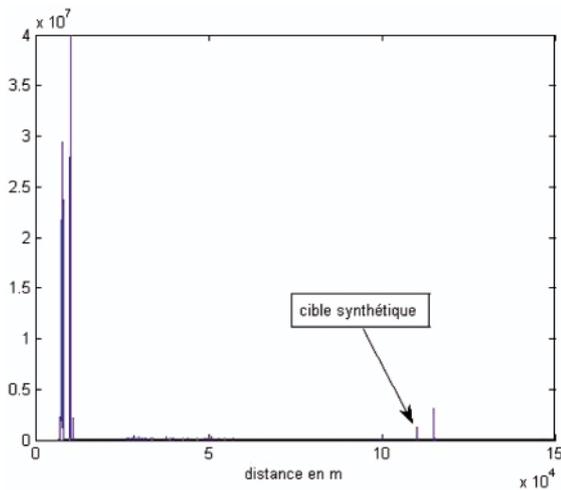


Figure 5. Ajout d'une cible synthétique dans les données initiales : détection via la norme de Frobenius.

5. Amélioration des résultats : moyennes et distances dans l'espace des coefficients de réflexion

On a vu que le modèle auto-régressif impose aux matrices de covariance d'être Toeplitz hermitiennes définies positives. Elles possèdent ainsi une structure beaucoup plus compacte que le simple fait d'appartenir à l'espace $HPD_n(\mathbb{C})$, puisqu'on peut représenter une telle matrice simplement par son premier vecteur colonne. C'est l'idée qu'on aimerait exploiter à priori dans

le but d'accélérer l'algorithme précédent. Malheureusement, il est facile de constater que le calcul de moyenne dans $HPD_n(\mathbb{C})$ introduit dans la section 3 ne préserve pas la structure Toeplitz des matrices, ce qui rend impossible son utilisation directe. Pour remédier à ce problème, on peut songer plutôt à moyenniser les coefficients auto-régressifs du modèle. Ceci revient à considérer en chaque case radar un vecteur de paramètres $\theta = (P_0, \mu_1, \dots, \mu_{n-1})$ où $P_0 \in \mathbb{R}_+^*$ est le coefficient de puissance du modèle auto-régressif et $(\mu_1, \dots, \mu_{n-1}) \in D^{n-1}$ les coefficients de réflexion. Nous renvoyons une fois encore à [7] et [10] pour les définitions de ces divers coefficients ainsi que pour les relations entre ceux-ci et la matrice de covariance des données. Remplacer la matrice de covariance par le vecteur θ présente le double avantage de travailler sur seulement n coefficients au lieu d'une matrice (n, n) tout en préservant la structure auto-régressive des données radar. Ceci nécessite cependant de définir une distance et une moyenne adéquate dans le nouvel espace $\mathbb{R}_+^* \times D^{n-1}$. Dans cette optique, il s'agit de reprendre la démarche de la géométrie de l'information dont nous avons déjà fait mention au début de la section 2. Appliquée au vecteur de paramètres θ , on trouve, après calculs, l'expression suivante pour la métrique de Fisher :

$$ds^2 = n \left(\frac{dP_0}{P_0} \right)^2 + \sum_{i=1}^n (n-i) \frac{|d\mu_i|^2}{(1-|\mu_i|^2)^2} \quad (13)$$

On obtient de cette façon une métrique riemannienne qui a la particularité intéressante que la matrice du produit scalaire en chaque point est diagonale. L'expression (13) montre par ailleurs que la métrique se réduit en fait à une combinaison

linéaire de métriques bien connues, à savoir $ds_1^2 = \left(\frac{dP_0}{P_0} \right)^2$ qui donne la distance logarithmique dans \mathbb{R}_+^* et

$ds_2^2 = \frac{|d\mu|^2}{(1-|\mu|^2)^2}$ qui n'est autre que la métrique hyperbo-

lique dans le disque de Poincaré D . Cette dernière jouit de nombreuses propriétés intéressantes (on connaît notamment l'expression explicite de la distance ainsi que des géodésiques) : on pourra se reporter à ce sujet aux premiers chapitres de [1]. Ces propriétés de l'expression (13) permettent de déterminer facilement les géodésiques pour la métrique : ce sont, comme on pourra le vérifier aisément, les produits des géodésiques de chaque espace \mathbb{R}_+^* muni de ds_1 et D muni de ds_2 . En intégrant le long de ces géodésiques, on peut expliciter la distance riemannienne entre deux vecteurs $\theta^{(1)}$ et $\theta^{(2)}$ de $\mathbb{R}_+^* \times D^{n-1}$:

$$d(\theta^{(1)}, \theta^{(2)}) = \sqrt{nd_0(P_0^{(1)}, P_0^{(2)})^2 + \sum_{k=1}^{n-1} (n-k) d_H(\mu_k^{(1)}, \mu_k^{(2)})^2} \quad (14)$$

avec $d_0(P_0^{(1)}, P_0^{(2)}) = \left| \ln \left(\frac{P_0^{(2)}}{P_0^{(1)}} \right) \right|$ (distance donnée par ds_1)

et d_H la distance hyperbolique dans le disque de Poincaré ayant pour expression (cf. [1]) :

$$d_H(\mu^{(1)}, \mu^{(2)}) = \frac{1}{4} \operatorname{arccosh} \left(1 + 2 \frac{|\mu^{(2)} - \mu^{(1)}|^2}{(1-|\mu^{(1)}|^2)(1-|\mu^{(2)}|^2)} \right)$$

Ceci donne au final une distance complètement explicite dans $\mathbb{R}_+^* \times D^{n-1}$.

Suivant le même cheminement que dans la section 3, on est naturellement amené à définir la moyenne de $\theta_1, \dots, \theta_N \in \mathbb{R}_+^* \times D^{n-1}$ comme le minimum de la fonctionnelle $H(\theta) = \sum_{i=1}^N d(\theta, \theta_i)^2$. Il est alors remarquable, au vu de (14), que ceci conduise en définitive à calculer la moyenne séparément pour chaque coordonnée dans l'espace qui lui correspond. Ainsi, pour les termes de puissance $P_0^{(i)}$, on calcule la moyenne géométrique habituelle dans \mathbb{R}_+^* . Pour ce qui est des coefficients de réflexion, on calcule chacune des moyennes par descente de gradient comme dans la section 3 de façon à minimiser $\sum_{i=1}^N d_H(\mu, \mu_i)^2$. Ceci est rendu algorithmiquement possible par le fait que l'on connaisse l'expression de la distance et des géodésiques dans le disque de Poincaré. C'est ce schéma numérique que nous avons implémenté et testé, à titre de comparaison, sur les mêmes données que la section précédente. La figure 5 montre ce que l'on obtient en appliquant l'algorithme sur les données initiales. On voit apparaître des pics de distance aux mêmes emplacements qu'avec l'algorithme de détection basée sur les matrices de covariance. En figure 6, on observe le résultat obtenu lorsqu'on a préalablement ajouté une cible synthétique. On notera à nouveau un pic de détection très marqué qu'on pourra comparer aux faibles résultats qu'on avait obtenus avec la méthode TFAC.

En définitive, travailler à partir des coefficients de réflexion plutôt que les matrices de covariance n'altère en rien la qualité de détection des cibles. En revanche, comme on pouvait s'y attendre, on diminue à la fois la durée des calculs et l'espace requis pour le stockage des variables. Ainsi, en Matlab, on a pu constater le gain d'un facteur 4 en terme de vitesse d'exécution des programmes.

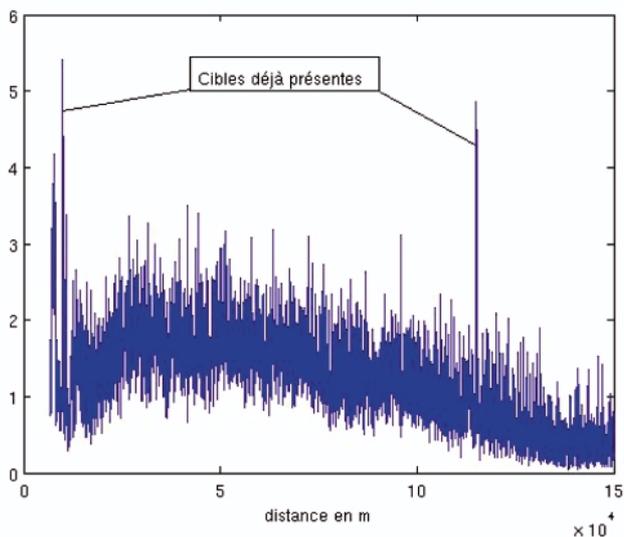


Figure 6. Détection des cibles sur les données initiales.

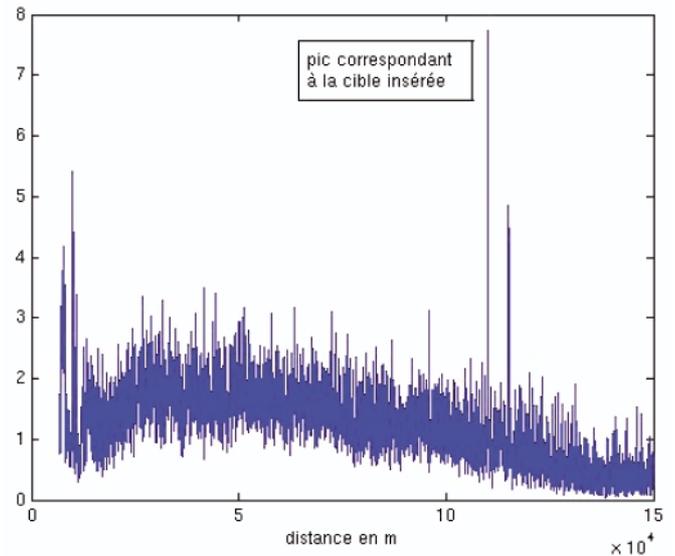


Figure 7. Ajout d'une cible synthétique.

6. Conclusion

Nous avons donc vu en quoi le passage par une structure mathématique a priori plus complexe que les données brutes (à savoir des matrices de covariance) permettait d'améliorer considérablement la qualité de détection lors d'une acquisition radar avec un faible nombre d'échantillons par salve d'émission. Ceci constitue dans ces cas précis une bonne alternative aux méthodes de type TFAC bien que le temps de calcul en soit sensiblement augmenté. Dans un souci d'économie, la section 5 présente une approche parallèle explorée par les auteurs utilisant pleinement les avantages des modèles auto-régressifs, qui constituent une représentation efficace et condensée des données. Sans pour autant prétendre avoir développé un algorithme satisfaisant à tout point de vue, on peut observer là une démarche générique consistant à déterminer sans cesse des représentations plus appropriées des données (et à chaque fois des métriques sur les nouveaux espaces rencontrés) qui, à notre sens, pourrait conduire à terme à des algorithmes encore plus performants. Citons à ce titre les récents travaux présentés dans [8], où une approche tout à fait similaire est employée pour construire une notion de médiane dans les espaces riemanniens dont il a été question ici.

Références

- [1] ANDERSON J., *Hyperbolic geometry*, Springer, 1964.
- [2] ANDO T., LI C-K., *Geometric means*, Linear Algebra Appl., 2004.
- [3] ARNAUDON M., XUE-MEI L., *Barycenters of measures transported by stochastic flow*, The annals of probability, 2005.
- [4] ARSIGNY V., FILLARD P., PENNEC X., *Geometric means in a novel vector space structure on symmetric positive-definite matrices*, Society for industrial and applied mathematics.

- [5] ATKINSON C. et MITCHELL A., *Rao's distance measure*, Shankhya, 1981.
- [6] BARBARESCO F., *Interactions between Symmetric Cone and Information Geometries*, ETVIC'08, Springer Lecture Notes in Computer Science, vol. 5416, pp. 124-163, 2009.
- [7] BARBARESCO F., *Procédé et dispositif de détermination du spectre de fréquence d'un signal*, Fascicule de brevet européen, 1996.
- [8] BARBARESCO F., BOUYT G., *Espace Riemannien symétrique et géométrie des espaces de matrices de covariance: équations de diffusion et calculs de médianes*, Colloque GRETSI'09, Dijon, Septembre 2009.
- [9] BARBARESCO F., *New Foundation of Radar Doppler Signal Processing based on Advanced Differential Geometry of Symmetric Spaces: Doppler Matrix CFAR and Radar Applications*, IEEE Radar'09 Conference, Bordeaux, France, Octobre 2009.
- [10] BULTHEEL A., VAN BAREL M., *Linear prediction: mathematics and engineering*, Belg. Math. Soc, 1994.
- [11] BURBEA J., *Informative geometry of probability spaces*, Expo. Math, 1986.
- [12] CHARON N., *Une nouvelle approche pour la détection de cibles dans les images radar*, Rapport de stage, 2008.
- [13] DO CARMO M., *Riemannian geometry*, Birkhäuser, 1992.
- [14] FARAUT J., *Analysis on symmetric cones*, Clarendon Press, 1994.
- [15] FREITAS P.J., *On the action of the symplectic group on the Siegel upper half plane*, 1997.
- [16] HUILING L., *Estimation of Riemannian barycentres*, London Mathematical society, 2004.
- [17] KREYSZIG E., *Differential geometry*, Dover books, 1991.
- [18] MOAKHER M., *A differential geometric approach to the geometric mean of symmetric positive-definite matrices*, 2002.
- [19] NEEB K.H., *Lie groups and matrix groups*, 2005.
- [20] OLLER J. et CALVO M., *A distance between multivariate normal distributions based in an embedding into the Siegel group*, Journal of multivariate analysis, 1990.
- [21] SIEGEL C-L., *Symplectic geometry*, Academic press, 1964.
- [22] STURM K-T., *Probability measures on metric spaces of nonpositive curvature*, Contemporary mathematics, 1991.



Nicolas **Charon**

Nicolas Charon est élève à l'Ecole Normale Supérieure de Cachan depuis 2006. Reçu à l'agrégation de mathématiques en 2008, il est désormais doctorant au CMLA et poursuit des recherches dans le domaine de la géométrie appliquée au traitement du signal.



Frédéric **Barbaresco**

Frédéric Barbaresco est « senior scientist and advanced studies manager » au département Stratégie Technologie et Innovation de la BL Surface Radar au sein de THALES AIR SYSTEMS. Ses domaines de recherche concernent la géométrie de l'information pour leurs applications dans le domaine radar. Auteur de plus de 100 articles et de 20 brevets, il a reçu la médaille Ampère de la SEE en 2007.