

# Comparaisons de méthodes robustes pour l'estimation de la covariance

Alexis DECURNINGE<sup>1,2</sup>, Frédéric BARBARESCO<sup>1</sup>

<sup>1</sup>Thales Air Systems  
Voie Pierre Gilles de Gennes 91470 Limours, France

<sup>2</sup>Laboratoire de Statistique Théorique et Appliquée  
4, place Jussieu, 75252 Paris cedex 05, France  
alexis.decurninge@thalesgroup.com  
frederic.barbaresco@thalesgroup.com

**Résumé** – Cet article traite de l'estimation de matrices de covariance dans le cadre de signaux stationnaires provenant d'un environnement non Gaussien. Nous introduisons la matrice de covariance du médian spatial d'un vecteur aléatoire réel que nous utilisons pour proposer un nouveau détecteur pour le problème de la détection de cibles. Nous comparons cette approche avec des estimateurs robustes de la covariance du signal complexe en terme de performance de détection de cibles.

**Abstract** – This article deals with the estimation of covariance matrices for stationary signals coming from a non-Gaussian environment. We introduce the covariance matrix of the spatial median of a real random vector that we use in order to propose a new detector for the problem of the target detection. We compare this approach to robust estimators of the complex signal covariance with respect to their performance in target detection.

## 1 Introduction

Cet article traite de la modélisation et de l'estimation de la distribution statistique du fouillis dans un environnement non Gaussien. Il est bien connu que la puissance des échos radars renvoyés par des fouillis de mer ou des fouillis de sol ont une distribution dite « à queue lourde » plus piquée. Ces pics apparaissent comme des valeurs aberrantes si l'on se place sous l'hypothèse gaussienne d'où la nécessité de construire des procédures « robustes » pour diminuer l'impact des erreurs de modélisation ou d'hypothèse.

Pour modéliser ces situations, l'hypothèse naturelle de circularité des signaux conduit à considérer les distributions elliptiques symétriques complexes (cf Ollila et al. [16]). Nous nous concentrerons dans la suite sur une sous-famille importante : les modèles CG [10] [12] [16] (pour « Compound Gaussian » aussi appelés SIRV pour « Spherically Invariant Random Vectors »). Ces modèles sont des généralisations des modèles Gaussiens avec un terme à queue lourde appelé « texture », la composante gaussienne étant appelée « speckle ».

Une approche robuste consiste à considérer une notion de médian spatial multivarié d'un vecteur aléatoire : dans un article peu connu [7], Rao introduisit une matrice d'association adaptée aux distributions à queue lourde dans le cadre de la régression. Nous reprenons cette matrice et les tests d'hypothèses associés basés pour construire un détecteur de cibles non paramétrique dans le cadre des signaux radars. Nous comparons cette

approche avec des estimateurs robustes proposés dans le cadre paramétrique des vecteurs autorégressifs Gaussiens et dans le cadre semi-paramétrique des vecteurs SIRV.

Dans la suite de l'article, la transposée d'un vecteur  $a$  sera notée  $a'$  et sa transposée conjuguée  $a^+$ .

## 2 Modélisation statistique non gaussienne

### 2.1 Contexte

Avant de présenter la modélisation du fouillis, il est utile de considérer le problème de la détection dans son intégralité. Soit un vecteur  $z \in \mathbb{C}^d$  représentant une rafale de  $d$  impulsions (nous considérerons dans la suite le cas  $d > 1$ ). Nous disposons d'une rafale pour chaque case d'espace. Le test de détection est construit à partir des  $N$  échantillons les plus proches spatialement de la case sous test. Les échantillons sont supposés venir en majorité de la même distribution  $H_0$  alors que les cibles recherchées viennent d'une distribution  $H_1$  :

$$\begin{cases} H_0 : \mathbf{z} = \mathbf{c} \\ H_1 : \mathbf{z} = \mathbf{c} + \mathbf{s} \end{cases} \quad (1)$$

$\mathbf{c}$  est la réponse du fouillis,  $\mathbf{s}$  la réponse de la cible. La fausse alarme étant due à  $H_0$ , nous considérons dans le cadre de cet article la modélisation de l'hypothèse nulle.

## 2.2 Les Vecteurs Aléatoires Sphériquement Invariants (SIRV)

Un vecteur SIRV  $\mathbf{c}$  résulte de la multiplication de deux composantes  $\mathbf{c} = \tau \mathbf{g}$  (voir [3]) avec :

- $\tau$  un scalaire aléatoire positif de densité  $h$
- $\mathbf{g}$  un vecteur aléatoire complexe Gaussien de moyenne nulle et de covariance  $\Sigma$

La densité de  $\mathbf{c}$  est alors donnée par :

$$f_{\mathbf{c}}(\mathbf{c}) = \frac{1}{\pi^d \det(\Sigma)} \int_0^{+\infty} \frac{1}{t^{2d}} \exp\left(\frac{-\mathbf{c}^+ \Sigma^{-1} \mathbf{c}}{2t^2}\right) h(t) dt \quad (2)$$

## 3 L'approche Haute Résolution Doppler basée sur la Géométrie de l'Information

Tout d'abord, nous présentons une approche robuste basée sur une modélisation gaussienne particulière : les vecteurs gaussiens autorégressifs. Grâce à l'hypothèse de stationnarité, la matrice de covariance du vecteur gaussien est Toeplitz Hermitienne définie positive. Nous reparamétrisons la matrice de covariance  $\Sigma$  avec le vecteur  $[r_0, \mu_1, \dots, \mu_{d-1}]$  (voir les théorèmes de Trench et Verblunsky [1]-[2] pour les détails sur la reparamétrisation) :

$$\Sigma \leftrightarrow [r_0, \mu_1, \dots, \mu_{d-1}] \quad (3)$$

avec  $r_0 = \mathbb{E}[|\mathbf{c}|^2]$  and  $\mu_i$  le  $i$ -ème paramètre de réflexion de Verblunsky.

Supposons que nous observons un échantillon  $\mathbf{c}_1, \dots, \mathbf{c}_N$ . Pour chaque  $\mathbf{c}_i$ , nous pouvons faire une estimation régularisée de  $(r_0, \mu_1, \dots, \mu_{d-1})$  (cf [8]). Nous pouvons alors estimer la matrice de covariance des  $N$  échantillons en prenant la médiane (critère  $L_1$ ) des estimées précédentes. Pour ce faire, nous voyons chaque matrice estimée dans l'espace des matrices Hermitiennes Toeplitz définies positives dotée d'une métrique venant de la géométrie de l'information (voir [14] et la thèse de Yang [13] récompensée par un Thales PhD Award).

Cette méthode a l'avantage d'être très géométrique et, malgré l'hypothèse d'une modélisation gaussienne stationnaire, l'utilisation du médian rend la procédure robuste aux valeurs aberrantes.

## 4 L'approche de Huber

Huber a utilisé [4] les M-estimateurs comme généralisation robuste et souple des estimateurs du maximum de vraisemblance. Mahot [15] considère l'estimation de la matrice de covariance dans les modèles SIRV avec ces M-estimateurs. Ces estimations sont les solutions d'une équation impliquant une fonction  $u$  laissée comme degré de liberté :

$$\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^N u(\mathbf{c}_i^+ \hat{\Sigma}^{-1} \mathbf{c}_i) \mathbf{c}_i \mathbf{c}_i^+ \quad (4)$$

Si  $u(x) = \frac{1}{x}$ , nous retrouvons l'estimateur du point fixe proposé par Conte [10] et Pascal [11].

Cependant, la structure Toeplitz de l'estimateur ci-dessus n'est pas conservée. Nous proposons donc un estimateur adapté de l'algorithme SEM (Stochastic Expectation Maximization) pour pouvoir calculer le « maximum de vraisemblance » d'un échantillon observé  $\mathbf{c}_1, \dots, \mathbf{c}_N$  pour le modèle SIRV :

- Initialisation des paramètres :  $\Sigma = \Sigma_0$  et  $h = h_0$
- Avant l'étape  $t + 1$ , nous avons les estimateurs  $\Sigma = \Sigma_t$  et  $h = h_t$ 
  - étape SE : pour chaque  $1 \leq i \leq N$  nous simulons les échantillons  $\tau_{1i}, \dots, \tau_{Pi}$  (avec  $P$  fixé) selon la distribution de  $\tau | \mathbf{c} = \mathbf{c}_i$  qui dépend de  $h_t$  et  $\Sigma_t$
  - étape M : nous effectuons les estimations spécifiques de  $h_{t+1}$  avec  $(\tau_{11}, \dots, \tau_{PN})$  et de  $\Sigma_{t+1}$  avec  $(\mathbf{x}_{11} = \frac{\mathbf{c}_1}{\tau_{11}}, \dots, \mathbf{x}_{PN} = \frac{\mathbf{c}_N}{\tau_{PN}})$
- Nous répétons l'opération jusqu'à une faible variation des paramètres

Cet algorithme permet de décorrélérer l'estimation de la distribution à queue lourde  $h$  de l'estimation de la covariance  $\Sigma$  du vecteur Gaussien. Malheureusement, la phase SE est très longue à calculer numériquement. Cette méthode reste donc théorique et nécessite des approximations simplificatrices pour être implémentable.

## 5 L'approche de Rao

Dans [7], Rao fait une revue des méthodes robustes basées sur la norme  $L_1$  présentes dans la littérature statistique. Dans le cas univarié, le médian (estimateur minimisant la norme  $L_1$  aux échantillons) devient plus efficace que la moyenne (sa contrepartie  $L_2$ ) si la distribution sous-jacente est à queue lourde. Huber donne plusieurs propriétés intéressantes du médian dans [5] :

- Il minimise le biais asymptotique maximum causé par des perturbations asymétriques
- Il est peu perturbé par la contamination des données : le « breakdown point » du médian est élevé
- C'est une limite de M-estimateurs

Rao tente de définir une notion de médian multivarié conservant ces propriétés. Pour une variable aléatoire réelle  $\mathbf{x} = (x_1, \dots, x_d)$ , soit  $\boldsymbol{\rho} = (\rho_1, \dots, \rho_d)$  le vecteur des médians marginaux,  $f_i$  la densité de  $x_i$  en  $\rho_i$  (abusivement  $f_i = \mathbb{P}(x_i = \rho_i)$ ) et  $\gamma_{ij} = \mathbb{P}(x_i \leq \rho_i, x_j \leq \rho_j) - \frac{1}{4}$ . Nous définissons la matrice d'association de  $\mathbf{x}$  :

$$\Gamma = \begin{pmatrix} \frac{\gamma_{11}}{f_1^2} & \cdots & \frac{\gamma_{1d}}{f_1 f_d} \\ \vdots & \ddots & \vdots \\ \frac{\gamma_{d1}}{f_d f_1} & \cdots & \frac{\gamma_{dd}}{f_d^2} \end{pmatrix} \quad (5)$$

Babu et Rao [6] ont montré la convergence en distribution du vecteur des médians marginaux  $\mathbf{m} = (m_1, \dots, m_d)$  :

$$\sqrt{N}(\mathbf{m} - \boldsymbol{\rho}) \xrightarrow{d} N_d(\mathbf{0}, \Gamma) \quad (6)$$

Rao propose un test d'hypothèse  $H_0 : \rho = \rho_0$  contre  $H_1 : \rho \neq \rho_0$  qui est l'analogue du  $T^2$  d'Hotelling. Soit  $\mathbf{x}_1, \dots, \mathbf{x}_N$  les  $N$  échantillons multivariés et  $\mathbf{m}$  les médians marginaux associés. La statistique de test est donnée par :

$$T^2 = N(\mathbf{m} - \rho_0)' \hat{\Gamma}^{-1} (\mathbf{m} - \rho_0) \quad (7)$$

Pour l'estimateur  $\hat{\Gamma}$ , il propose :

$$\gamma_{ij} = \frac{1}{n} |\{1 \leq r \leq N \text{ s.t. } e_{ir} \leq 0, e_{jr} \leq 0\}| - \frac{1}{4} \quad (8)$$

avec  $\forall 1 \leq r \leq N$ ,  $(e_{1r}, \dots, e_{dr}) = \mathbf{x}_r - \mathbf{m}$ .  $f_i$  étant une densité, nous utilisons l'estimateur à noyau proposé dans Babu et Rao [6] :

$$\hat{f}_i^{-1} = N^\delta \int_0^{\log(N)} \left[ \hat{F}_N^{-1} \left( \frac{1}{2} + vN^{-\delta} \right) - \hat{F}_N^{-1} \left( \frac{1}{2} \right) \right] k(v) dv \quad (9)$$

avec  $\delta = \frac{1}{2d-1}$ ,  $\hat{F}_N^{-1}$  le quantile empirique et le noyau  $k(v) = (-0.5v^2 + v + 1)e^{-v}$ .

Afin d'appliquer ce test au problème de la détection de cibles, nous devons spécifier un vecteur réel  $\mathbf{x}$  : nous proposons de prendre  $\mathbf{x} = \left( \sum_{j=0}^{d-1} z_k e^{-\frac{2i\pi jk}{d}} \right)_{1 \leq k \leq d}$  et, comme statistique de test :

$$T^2 = N(\mathbf{m} - \mathbf{x})' \hat{\Gamma}^{-1} (\mathbf{m} - \mathbf{x}) \quad (10)$$

## 6 Résultats expérimentaux

L'approche de Rao n'effectuant pas d'estimation de la covariance du signal, nous ne pouvons comparer les performances des algorithmes directement sur la qualité de l'estimation. Nous illustrons donc celles-ci à travers leur capacité de détection.

Ainsi, nous simulons le fouillis  $\mathbf{c}$  par le modèle SIRV de texture suivant une distribution Weibull ou une distribution Gamma. Le paramètre d'échelle de la texture représente la puissance du fouillis (fixée à 30 dB). Plus son paramètre de forme est petit, plus la distribution est piquée.

Nous simulons  $N = 64$  cellules environnantes selon la distribution SIRV avec une matrice de covariance paramétrée par  $\rho = 0,9 : \Sigma = (\rho^{|i-j|})_{1 \leq i \leq d, 1 \leq j \leq d}$ .

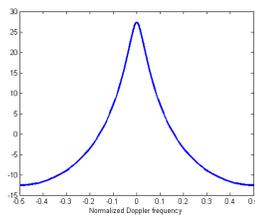


FIGURE 1 – Spectre du fouillis

La cible simulée est impulsionnelle

$$\mathbf{s}(f_d) = \alpha [1, e^{2i\pi f_d}, \dots, e^{2i\pi f_d(d-1)}] \quad (11)$$

avec un paramètre de puissance  $\alpha$  et un paramètre de fréquence Doppler normalisé  $f_d$  (choisi égal à 0, 1).

Nous utilisons une statistique de test commune pour les algorithmes estimant la matrice de covariance  $\Sigma$  : l'ANMF (Adaptive Normalized Matched Filter) donnée par

$$\max_{f_d} \frac{|\mathbf{s}(f_d)' \hat{\Sigma}^{-1} \mathbf{z}|^2}{[\mathbf{s}(f_d)' \hat{\Sigma}^{-1} \mathbf{s}(f_d)] [\mathbf{z}' \hat{\Sigma}^{-1} \mathbf{z}]} \quad (12)$$

L'erreur d'estimation pour  $\Sigma$  doit être normalisée pour avoir un sens. Nous définissons une erreur moyenne pour  $N$  estimations :

$$\frac{1}{N} \sum_{i=1}^N \left\| \frac{\hat{\Sigma}_i}{\|\hat{\Sigma}_i\|} - \frac{\Sigma}{\|\Sigma\|} \right\| \quad (13)$$

Nous présentons les performances pour la détection de cibles pour un taux de fausse alarme fixe ainsi que les erreurs d'estimation pour la matrice de covariance  $\Sigma$ . Les acronymes utilisées pour les courbes de résultat sont :

- AR : estimation de la matrice de covariance présentée à la section 3
- FP : algorithme du point fixe proposé par Pascal [11] spécifié à la section 4
- « empirical » : la matrice de covariance empirique comme estimateur de  $\Sigma$  ( $\hat{\Sigma} = \sum \mathbf{z}_i \mathbf{z}_i'$ )
- « theoretical » : la matrice de covariance  $\Sigma$  est supposée connue
- Rao-FFT : le test et l'estimation spécifiés à la section 5

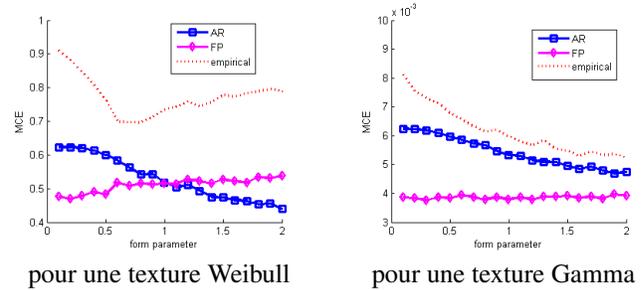
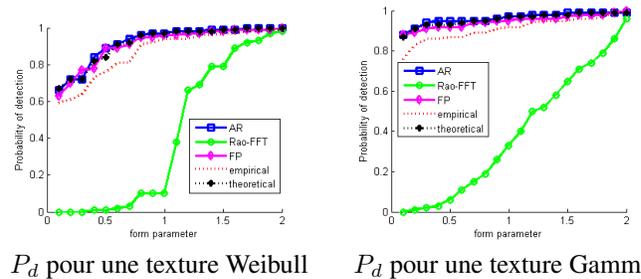
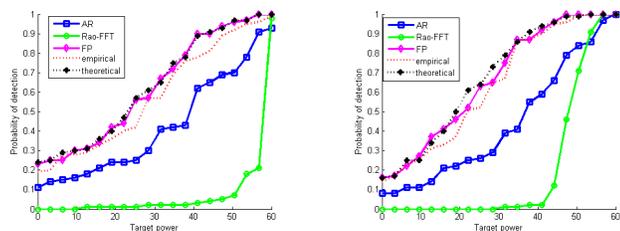


FIGURE 2 – Erreur d'estimation de la matrice de covariance  $\Sigma$  en fonction du paramètre de forme de la texture



$P_d$  pour une texture Weibull  $P_d$  pour une texture Gamma

FIGURE 3 – Influence du paramètre de forme pour une cible de puissance 30 dB



$P_d$  pour une texture Weibull  $P_d$  pour une texture Gamma

FIGURE 4 – Influence de la puissance de la cible pour un paramètre de forme égal à 0, 1

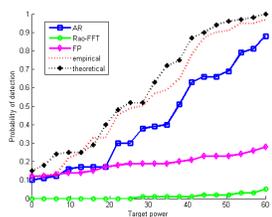


FIGURE 5 – Cas multi-cibles : nous polluons l'échantillon des 64 cellules environnantes par 30 cellules correspondant à l'hypothèse  $H_1$

## 7 Conclusion

Nous avons présenté plusieurs algorithmes pour l'estimation de la matrice de covariance du signal complexe et comparé leurs performances en simulant un modèle SIRV. Pour des textures à queue lourde, le gain des méthodes robustes par rapport à l'estimation naïve empirique est significatif. Cependant, on ne retrouve pas ce gain dans la performance en terme de détection de cibles car le test utilisé (ANMF) est adapté au modèle simulé. De même, les performances de l'approche de Rao qui est totalement non paramétrique ne sont pas compétitives sous ce modèle.

## 8 Remerciements

Cette étude est financée par le ministère de la Défense française (DGA/MRIS) et Thales Air Systems.

## Références

[1] S. Verblunsky. *On positive harmonic functions*. Proc. London Math. Soc., 38, pp 125-157, 1935 ; Proc. London Math. Soc., 40, pp 290-320, 1936

[2] W.F. Trench, *An algorithm for the inversion of finite Toeplitz matrices*, J. Soc. Indust. Appl. Math., 12, pp 515-522, 1964

[3] K. Yao, *A representation theorem and its application to spherically-invariant random process*, IEEE Trans., September 1973

[4] P. J. Huber, *Robust statistics*, John Wiley & Sons, 1981

[5] P. J. Huber, *The place of  $L_1$ -norm in Robust estimation*, Computational statistics & data Analysis, no. 5, pp 255-262, 1987

[6] G. J. Babu, C. R. Rao, *Joint asymptotic distribution of marginal quantile functions in samples from a multivariate population*, Journal of Multivariate Analysis, vol. 27, pp. 15-23, 1988

[7] C. R. Rao, *Methodology based on the  $L_1$ -Norm in statistical inference*, Sankhyā, vol. 50, no. 3, pp. 289-313, Oct. 1988

[8] F. Barbaresco, *Super Resolution Spectrum Analysis Regularization : Burg, Capon and Aco-antagonistic Algorithms*, EUSPICO-96, pp. 2005-2008, Trieste, Sept. 1996

[9] F. Gini, *Sub-optimum Coherent Radar Detection in a Mixture of  $K$ -distributed and Gaussian Clutter*, IEE Proc. Radar, Sonar and Navigation, vol. 144, no. 1, pp. 39-48, Feb. 1997

[10] E. Conte, A. De Maio, G. Ricci, *Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection*, IEEE Trans.-SP, vol. 50, no. 8, Aug. 2002

[11] F. Pascal, P. Forster, J.P. Ovarlez, P. Larzabal, *Performance Analysis of Covariance Matrix Estimates in Impulsive Noise*, IEEE Transactions on signal processing, vol. 56, no. 6, pp. 2206-2217, June 2008

[12] J. Carretero-Moya, J. Gismero-Menoyo, A. Asensio-Lopez, A. Blanco Del Campo. *Small-Target Detection in High-Resolution Heterogeneous Sea-Clutter : An Empirical Analysis*, IEEE Transactions on Aerospace and Electronic Systems, July 2011

[13] L. Yang, *Médianes de mesures de probabilité dans les variétés riemanniennes et applications à la détection de cibles radar*, Thèse de l'Université de Poitiers, tel-00664188, 2011, Thales PhD Award 2012

[14] F. Barbaresco, *Information Geometry of Covariance Matrix : Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger fibration and Frechet median*, in R. Bhatia & F. Nielsen Ed., "Matrix Information Geometry", Springer Lecture Notes in Mathematics, 2012

[15] M. Mahot, F. Pascal, P. Forster, J-P. Ovarlez. *Robust ANMF test using Huber's  $M$ -estimator*. IEEE Sensor Array and Multichannel Signal Processing Workshop, June 2012

[16] E. Ollila, D. Tyler, V. Koivunen, V. Poor *Complex Elliptically Symmetric Distributions : Survey, New Results and Applications*, IEEE Transactions on signal processing, vol. 60, no. 11, 2012