

# Sur des tests d’hypothèses composites pour des paramètres parcimonieux

Silvia PARIS, David MARY, André FERRARI

UMR 6525 H. Fizeau, Université de Nice-Sophia Antipolis  
CNRS, Observatoire de la Côte d’Azur, Campus Valrose, F-06108 Nice cedex, FRANCE

Silvia.Paris@unice.fr, David.Mary@unice.fr  
Andre.Ferrari@unice.fr

**Résumé** – Nous proposons une analyse des liens existant entre des fonctions seuillantes utilisées en estimation et certains tests de détection. Nous nous intéressons à quatre tests de détection: le Rapport de vraisemblance généralisé (Generalized Likelihood Ratio, GLR), le Facteur de Bayes et deux tests basés sur l’estimation au sens du Maximum *A Posteriori* (MAP). Dans le cas scalaire, nous montrons que les tests considérés sont équivalents au GLR. Dans le cas de vecteurs de paramètres parcimonieux, les deux tests basés sur le MAP présentent de meilleures performances que les deux autres grâce à des statistiques de test qui tendent à se focaliser sur les composantes non nulles du vecteur de paramètres, en lien direct avec la fonction de seuillage opérée par le MAP dans le cas scalaire. Nous proposons une analyse asymptotique (en la taille du vecteur) et une analyse à taille finie des performances des tests basés sur le MAP. Les résultats théoriques sont illustrés par des exemples numériques.

**Abstract** – We propose an analysis of some connections existing between sparse estimation and detection tests. We compare four detection tests: in addition to the Generalized Likelihood Ratio (GLR) and to the Bayes Factor, we consider two tests based on Maximum *A Posteriori* (MAP) estimates. In the case of a scalar parameter, we show that all considered tests are equivalent to the GLR. In the case of a sparse vector of parameters, the two tests based on MAP estimates outperform the GLR and the Bayes Factor thanks to thresholded test statistics. We propose an asymptotic analysis (in the vector length) and an analysis at finite length of the performances of the two MAP-based tests. Theoretical results are illustrated by numerical exemples.

## 1 Introduction

La littérature sur les problèmes inverses témoigne depuis quelques décennies d’un intérêt particulier pour différentes méthodes d’estimation et de détection prenant en compte la parcimonie des signaux à traiter [3]. Dans le cadre de l’estimation, ces méthodes sont souvent basées sur l’utilisation de fonctions de seuillage, dont certaines peuvent être interprétées dans le cadre du Maximum *A Posteriori* (MAP), avec des lois *a priori* appropriées [1]. Par construction, de telles méthodes d’estimation tendent à localiser des espaces de dimensions réduites où serait contenue l’information d’intérêt. Les vecteurs de paramètres considérés ici sont déterministes, et parcimonieux (peu de coefficients sont non nuls). Ils peuvent correspondre par exemple aux coefficients d’analyse obtenus en projetant des images à l’aide de dictionnaires d’ondelettes ou d’autres transformations. Le problème concernant l’estimation d’un vecteur de paramètres parcimonieux est donc lié à celui de la détection d’un support, ou d’un sous-espace. Malgré cette connexion évidente entre estimation et détection parcimonieuse on ne trouve pas, à notre connaissance, d’étude formelle de cette interaction dans la littérature, exception faite, récemment, de [5].

Le travail proposé ici vise à clarifier les connexions entre l’estimation parcimonieuse, reposant sur l’utilisation de fonc-

tions de seuillage, et des tests de détection qui prennent en compte la parcimonie en tentant de n’inclure dans le test que les composantes actives du vecteur de paramètres inconnus [4]. Dans la suite nous présentons en Section 2 les tests et considérons le cas scalaire. Ce cas est une étude préliminaire pour le cas vectoriel, qui est bien sûr le cas qui nous intéresse. Le cas vectoriel est traité en Section 3. La Section 4 illustre les résultats obtenus par quelques exemple numériques.

## 2 Cas scalaire

On s’intéresse au test d’hypothèses suivant :

$$\begin{cases} \mathcal{H}_0 : x = \epsilon, & \epsilon \sim \mathcal{N}(0, \sigma^2) \\ \mathcal{H}_1 : x = \theta + \epsilon \end{cases} \quad (1)$$

où  $\theta$  est un paramètre scalaire déterministe inconnu et  $\epsilon$  est Gaussien. Nous considérons ci-dessous quatre tests de détection dont l’intérêt sera justifié dans le cas vectoriel.

### 2.1 Rapport de vraisemblance généralisé (GLR)

Pour une vraisemblance Gaussienne  $p$  et les hypothèses considérées ci-dessus, le rapport de vraisemblance généralisé s’écrit :

$$GLR(x) = \frac{\max_{\theta \neq 0} p(x | \theta)}{p(x | 0)}. \quad (2)$$

Le GLR fait intervenir l'estimation par Maximum de Vraisemblance du paramètre  $\theta$ ,  $\hat{\theta}_{MV} = x$ . En comparant  $GLR(x)$  à un certain seuil  $\gamma$  on obtient l'expression de la statistique de test  $T_{GLR}$  :

$$GLR(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma \Leftrightarrow T_{GLR}(x) = \frac{x^2}{\sigma^2} \underset{H_0}{\overset{H_1}{\geq}} 2 \ln \gamma = \gamma_{GLR}. \quad (3)$$

Pour un seuil donné, la probabilité de fausse alarme  $P_{FA_{GLR}}$  est

$$P_{FA_{GLR}} = P \{T_{GLR}(x) > \gamma_{GLR} \mid \mathcal{H}_0\}.$$

## 2.2 Rapport du maximum des probabilités a posteriori (PDR)

Le PDR [2] est défini comme

$$PDR(x) = \frac{\max_{\theta \neq 0} p(\theta \mid x)}{p(0 \mid x)} = \frac{p(x \mid \hat{\theta}_{MAP}) \pi(\hat{\theta}_{MAP})}{p(x \mid 0) \pi(0)} \quad (4)$$

où  $\pi(\theta)$  représente l'*a priori*. En estimation parcimonieuse,  $\pi(\theta)$  est choisi de façon que l'estimation MAP de  $x$  soit une fonction seuillante de  $x$ . C'est en particulier le cas pour un *a priori* Laplacien  $\pi(\theta) = \frac{1}{2\lambda} e^{-\frac{|\theta|}{\lambda}}$ , pour lequel  $\hat{\theta}_{MAP}$  correspond à la fonction de seuillage doux [6] :

$$\hat{\theta}_{MAP} = \begin{cases} 0, & \text{si } -\frac{\sigma}{\lambda} < \frac{x}{\sigma} < \frac{\sigma}{\lambda} \\ \frac{x}{\sigma} - \text{sgn}(x) \frac{\sigma}{\lambda}, & \text{sinon.} \end{cases} \quad (5)$$

En comparant  $PDR(x)$  à un seuil  $\gamma$  on obtient le test :

$$PDR(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma \Leftrightarrow T_{PDR}(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma_{PDR},$$

où  $\gamma_{PDR} = 2 \ln \gamma$  et la statistique de test  $T_{PDR}(x)$  est avec (5) :

$$T_{PDR}(x) = \begin{cases} 0, & \text{si } \frac{|x|}{\sigma} < \frac{\sigma}{\lambda} \\ \left(\frac{|x|}{\sigma} - \frac{\sigma}{\lambda}\right)^2, & \text{si } \frac{|x|}{\sigma} \geq \frac{\sigma}{\lambda} \end{cases} \quad (6)$$

## 2.3 Rapport de vraisemblance avec estimation MAP (LRMAP)

Nous proposons ici le LRMAP, un test inspiré du GLR qui à différence de celui-ci utilise l'estimation MAP de façon à corréler les données avec une estimation pertinente dans le cas de signaux parcimonieux. Le LRMAP est défini comme :

$$LRMAP(x) = \frac{p(x \mid \hat{\theta}_{MAP})}{p(x \mid 0)}. \quad (7)$$

Pour une vraisemblance Gaussienne on obtient :

$$LRMAP(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma \Leftrightarrow T_{LRMAP}(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma_{LRMAP},$$

où la statistique de test  $T_{LRMAP}(x)$  s'écrit :

$$T_{LRMAP}(x) = \begin{cases} 0, & \text{si } \frac{|x|}{\sigma} < \frac{\sigma}{\lambda} \\ \frac{x^2}{\sigma^2} - \frac{\sigma^2}{\lambda^2}, & \text{si } \frac{|x|}{\sigma} \geq \frac{\sigma}{\lambda} \end{cases},$$

en utilisant l'eq.(5) et où  $\gamma_{LRMAP} = 2 \ln \gamma$ .

## 2.4 Facteur de Bayes (BF)

Le BF est défini par :

$$BF(x) = \frac{\int_{\mathbb{R}} p(x \mid \theta) \pi(\theta) d\theta}{p(x \mid 0)} \underset{H_0}{\overset{H_1}{\geq}} \gamma. \quad (8)$$

En considérant une vraisemblance Gaussienne et un *a priori* Laplacien, la statistique de test du BF s'écrit :

$$T_{BF}(x) = e^{\frac{x^2}{2\sigma^2}} \left\{ \left[ 2 \cosh\left(\frac{x}{\lambda}\right) - e^{-\frac{x}{\lambda}} \Phi\left(-\frac{x}{\sigma} + \frac{\sigma}{\lambda}\right) - e^{\frac{x}{\lambda}} \Phi\left(\frac{x}{\sigma} + \frac{\sigma}{\lambda}\right) \right] \right\}$$

où  $\Phi(\cdot)$  est la fonction de répartition d'une loi normale centrée réduite.

## 2.5 Discussion

On peut vérifier que les statistiques de test obtenues pour ces quatre tests sont des fonctions paires, strictement croissantes pour  $|x| \geq x_0$ . Or, nous avons montré en [7] que toute statistique  $T(x)$  possédant ces propriétés est équivalente au GLR. Par conséquent, les courbes CORE (Caractéristique Opérationnelle de Réception) de ces tests coïncident avec celle du GLR et la  $P_{FA}$  associée à chaque statistique de test  $T$  est

$$P_{FA_{test}}(\gamma_{test}) = 1 - \Phi\left(\frac{1}{\sigma^2} [T^{-1}(\gamma_{test})]^2\right). \quad (9)$$

Dans le cas scalaire, les courbes CORE des tests présentés coïncident donc avec celle du GLR. Pour illustrer cette équivalence, prenons l'exemple du PDR. Selon l'expression (6),  $T_{PDR}$  est une fonction paire et strictement croissante pour  $\frac{|x|}{\sigma} \geq x_0 = \frac{\sigma}{\lambda}$ . Par conséquent, il y aura détection si  $x^2$  est suffisamment élevé, ce qui est essentiellement le test du GLR. Précisément, on montre que le test  $T_{PDR}(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma_{PDR}$  est le même test que

$T_{GLR}(x) \underset{H_0}{\overset{H_1}{\geq}} \gamma_{GLR}$  en prenant  $\gamma_{GLR} = [\sqrt{\gamma_{PDR}} + \frac{\sigma}{\lambda}]^2$ . Notons encore que dans les cas du PDR et du LRMAP la détection ne peut se faire que si  $\frac{|x|}{\sigma} \geq \frac{\sigma}{\lambda}$  : la quantité  $\frac{\sigma}{\lambda} = \eta$  fixe donc l'extension maximale en fausse alarme de ces courbes. Les courbes CORE ne décrivent par conséquent qu'une partie de la courbe du GLR.

L'équivalence de ces quatre tests se généralise pour des *a priori* Gaussiens généralisés centrés avec  $0 < p < 1$  [7]. La preuve repose sur le fait que ces lois sont paires, et produisent des estimateurs  $\hat{\theta}_{MAP}(x)$  qui sont impairs, croissants en  $x$  et dont le biais est positif pour  $x > 0$  [1].

## 3 Cas vectoriel

Considérons à présent les hypothèses :

$$\begin{cases} \mathcal{H}_0 : \underline{x} = \underline{\epsilon}, & \underline{\epsilon} \sim \mathcal{N}(\underline{0}, \Sigma) \\ \mathcal{H}_1 : \underline{x} = \underline{\theta} + \underline{\epsilon} \end{cases} \quad (10)$$

où  $\underline{x}$ ,  $\underline{\theta}$  et  $\underline{\epsilon}$  sont des vecteurs à  $N$  composantes,  $\underline{\theta}$  est parcimonieux et  $\underline{\epsilon}$  est Gaussien avec  $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_N^2)$ . En estimation, on sait que le maximum de vraisemblance et l'estimateur d'Erreur Quadratique Moyenne Minimale ne produisent

par des solutions parcimonieuses, par conséquent le GLR et le BF ne semblent pas appropriés pour la détection de tels vecteurs. En revanche, l'estimateur du MAP produit de telles solutions pour des lois *a priori* strictement décroissantes en  $\theta = 0$ , telles que celles considérées en Sec. 2. L' injection de l'estimation MAP avec un *a priori* approprié dans un test de détection promet la parcimonie et rend donc le PDR et le LRMAP très pertinents pour la détection de signaux parcimonieux.

Reprenons les calculs des statistiques des quatre tests dans le cas vectoriel en considérant un *a priori* Laplacien  $\pi(\theta) = \prod_i \frac{1}{2\lambda_i} e^{-\frac{|\theta_i|}{\lambda_i}}$ . Pour le GLR et le BF l'extension du cas scalaire est immédiate en calculant les statistiques de tests à partir du produit sur les  $N$  composantes. Pour  $T_{PDR}(\underline{x})$  et  $T_{LRMAP}(\underline{x})$  on obtient, en notant  $u_i = x_i/\sigma_i$  :

$$T_{PDR}(\underline{x}) = \sum_{i=1}^N (|u_i| - \frac{\sigma_i}{\lambda_i})^2 I(|u_i| > \frac{\sigma_i}{\lambda_i}), \quad (11)$$

$$T_{LRMAP}(\underline{x}) = \sum_{i=1}^N (u_i^2 - \frac{\sigma_i^2}{\lambda_i^2}) I(|u_i| > \frac{\sigma_i}{\lambda_i})$$

où  $I(\cdot)$  est la fonction indicatrice provenant des seuillages scalaires (5) et (6) sur chaque composante.

Le taux de fausse alarme pour une composante pondérée  $u_i = \frac{x_i}{\sigma_i}$  est  $\Pr(|u_i| > \frac{\sigma_i}{\lambda_i} | \mathcal{H}_0)$  : il paraît donc naturel de choisir les hyperparamètres  $\lambda_i$  de telle sorte que le seuil par composante  $\sigma_i/\lambda_i$  soit constant. Appelons  $\eta = \sigma_i/\lambda_i$  et reconsidérons les tests (11) pour cette valeur.

La première statistique permet de retrouver la procédure de seuillage doux, proposée en détection par Fan [4], avec l'objectif d'éliminer des composantes de la statistique de test. On voit ici que ce test correspond en fait à faire le test de détection du PDR (ou *type 2 ML test*) discuté par Basu (hors d'un contexte de parcimonie) dans [2].

Les deux tests PDR et LRMAP dépendent du couple de paramètres  $(\eta, \gamma)$ . Différents couples de valeurs  $(\eta, \gamma)$  peuvent représenter le même taux de fausse alarme, comme illustré en Fig. 1. Les deux premières figures de la colonne de gauche montrent comment deux différents couples  $(\eta = 1, \gamma_1 : \text{rouge})$  et  $(\eta = 2, \gamma_2 : \text{vert})$  peuvent représenter la même  $P_{FA}$  (aire à droite des seuils  $\gamma_1$  et  $\gamma_2$ ). Or, le calcul de la  $P_{FA} = \Pr(T(\eta) > \gamma)$  n'est pas possible dans le cas général. En effet, les distributions de  $T_{PDR}$  et  $T_{LRMAP}$  sont très complexes. Pour le LRMAP par exemple, la distribution de  $T_{LRMAP}$  est la convolution de  $N \chi^2$  tronqués sur  $[\eta; +\infty[$ , décalés en 0 et avec une masse en 0 ( $\Pr(0 \leq \chi^2 < \eta) = [2\Phi(\sqrt{\eta}) - 1]^N$ ). Une façon de surmonter cette difficulté consiste à poser  $\gamma = 0$ . Dans ce cas les deux tests sont les mêmes et ne dépendent que du paramètre  $\eta$ . Ceci permet de calculer facilement le taux de fausse alarme maximal. En effet, comme  $T_{PDR}$  et  $T_{LRMAP}$  sont positifs dans (11) :

$$\begin{aligned} P_{FA_0} = \Pr(T > 0 | \mathcal{H}_0) &= 1 - \Pr(|u_i| < \eta \forall i) \\ &= 1 - (2\Phi(\eta) - 1)^N. \end{aligned} \quad (12)$$

De la même manière on obtient la probabilité de détection pour

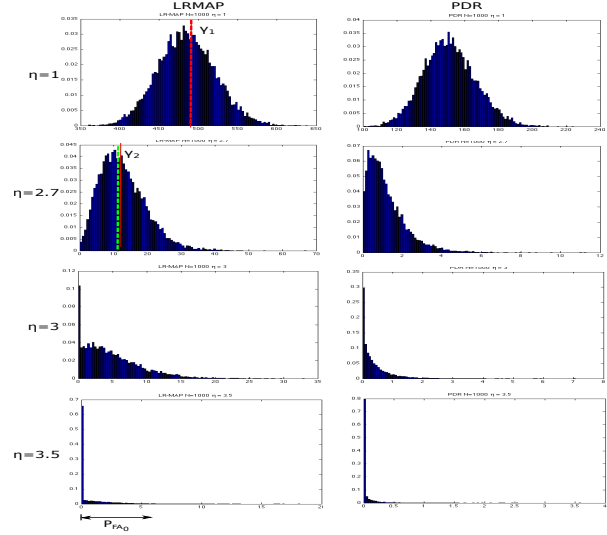


FIGURE 1 – Cas vectoriel : Distributions empiriques des statistiques  $T_{LRMAP}$  (colonne de gauche) et  $T_{PDR}$  (colonne de droite) pour un vecteur de paramètres à  $N = 1000$  composantes et différentes valeurs du paramètre  $\eta$ . Les deux premières figures dans la colonne à gauche montrent comment deux différents valeurs du couple  $(\eta, \gamma)$  peuvent donner lieu au même taux de fausse alarme.

$\gamma = 0$  :

$$\begin{aligned} P_{DET_0} &= \Pr(T > 0 | \mathcal{H}_1) \\ &= 1 - \prod_{i=1}^N [\Phi(\eta - \theta_i) + \Phi(\eta + \theta_i) - 1]. \end{aligned} \quad (13)$$

Les distributions complexes de  $T_{PDR}$  et  $T_{LRMAP}$  rendent la comparaison analytique de leurs performances difficile. Cependant on peut montrer que pour  $\eta \rightarrow 0$  les deux tests sont équivalents. En comparant les deux expressions en (11) nous avons en effet

$$T_{LRMAP} = T_{PDR} + 2(|u_i|\eta - \eta^2) I(|u_i| > \eta), \quad (14)$$

ce qui montre que  $T_{LRMAP}$  converge en loi vers  $T_{PDR}$  quand  $\eta = \frac{\sigma_i}{\lambda_i} \rightarrow 0$ . Ceci se vérifie dans les deux cas suivants :

$$\begin{cases} \sigma_i \rightarrow 0, & \text{cas de fort rapport signal à bruit (SNR)} \\ \lambda_i \rightarrow \infty, & \text{cas d'a priori uniforme.} \end{cases} \quad (15)$$

Pour finir, nous considérons ci-dessous une approximation asymptotique de  $T_{LRMAP}(\underline{x})$  quand  $N \rightarrow \infty$ , analogue à celle de [4] pour  $T_{PDR}(\underline{x})$ . Le maximum de  $N$  variables indépendantes  $\mathcal{N}(0, 1)$  converge vers  $\sqrt{2 \log N}$ . Afin de ne pas accumuler des composantes "bruit seul" dans les tests (11) il est donc nécessaire de choisir un seuil  $\eta$  de la forme  $\eta = \sqrt{2 \log N} a_N$  où  $a_N$  (par exemple  $a_N = 1/\log(N)^2$ ) permettra de conserver les composantes "signal". On peut alors démontrer la normalité asymptotique en  $N$  des deux statistiques de test (11) sous  $\mathcal{H}_0$  par un théorème centrale limite en vérifiant la condition de Lyapounov.

De façon similaire à [4], les moments d'ordre  $k$  de  $T_{LRMAP}$  font intervenir  $E[(u^2 - \eta^2) I(|u| > \eta)]^k$  et peuvent être calculés

par intégrations par parties successives de  $\int_{\eta}^{\infty} u^{2k} e^{-\frac{u^2}{2}} du$  suivies d'une approximation pour  $\eta \rightarrow \infty$ . Pour une probabilité de fausse alarme cible  $P_{FA}$  nous obtenons dans ce cas le test : Rejeter  $\mathcal{H}_0$  si

$$\sigma_{LRMAP}^{-1}(T_{LRMAP}(\underline{x}) - \mu_{LRMAP}) > \Phi^{-1}(1 - P_{FA}),$$

$$\text{avec } \begin{cases} \mu_{LRMAP} = \sqrt{\frac{2}{\pi}} e^{-\eta^2/2} \left( \frac{2}{\eta} - \frac{4}{\eta^3} + O(\eta^{-5}) \right) \\ \sigma_{LRMAP}^2 = \sqrt{\frac{2}{\pi}} e^{-\eta^2/2} \left( \frac{8}{\eta} - \frac{24}{\eta^3} + O(\eta^{-5}) \right) \end{cases}.$$

On observe cependant une convergence très lente en pratique, comme observé dans [4] à propos du PDR.

## 4 Simulations et conclusions

La Fig. 2 montre les courbes CORE du GLR, du BF, du LRMAP et du PDR dans le cas scalaire, pour  $\theta = 0.8$  et  $\sigma = \lambda = 1$ . Comme discuté en Sec.2, les quatre courbes correspondent à celle du GLR. Le test de Neyman-Pearson, qui maximise la probabilité de détection pour un certain taux de fausse alarme, en corrélant les données aux vraies valeurs du paramètre  $\theta$  sous l'hypothèse  $\mathcal{H}_1$ , est indiqué en cyan pour référence. Pour le LRMAP et le PDR, seule une partie de la courbe est accessible, et  $\max(P_{FA}) = P_{FA_0}$  est fixé par  $\eta = \sigma/\lambda$ . Dans le cas considéré et d'après (12) pour  $N=1$ ,  $P_{FA_0} \approx 0.32$  (ligne verticale en pointillé). Des résultats pour le cas vectoriel sont

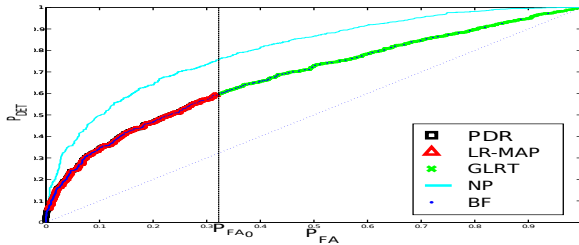


FIGURE 2 – Cas scalaire : Courbes CORE pour les tests du GLR, BF, LRMAP, PDR et de NP.

montrés en Fig.3. Ici, les quatre tests (PDR : noir, LRMAP : rouge, BF : bleu, GLR : vert) sont comparés pour un vecteur de paramètres où  $N = 400$ . Les écarts types  $\sigma_i$  du bruit augmentent comme la racine des index des composantes de  $\sigma_1 = 1$  à  $\sigma_{400} = 4, 5$ . Seules 2 composantes sont non nulles, d'amplitudes 3 et 5. Les hyperparamètres sont réglés de telle sorte que  $\eta = 2$ . Pour le PDR et le LRMAP,  $P_{FA_0} \approx P_{DET_0} \approx 1$  (eq.(12) et (13)) dans ce cas. Cette simulation montre l'intérêt des tests PDR et LRMAP. Soulignons que suivant les valeurs des paramètres, les performances relatives de PDR et LRMAP peuvent être inversées, comme le montrent les figures du centre et de droite en Fig.4. Ici, les courbes CORE du LRMAP (en bleu) et du PDR (en rouge) sont comparées pour  $\eta = 0.05$ ,  $\eta = 1$  et  $\eta = 3.3$ .  $\underline{\theta}$  est un vecteur à  $N = 400$  composantes dont 10 sont non nulles. Dans la première figure à gauche, correspondant à  $\eta = 0.05$ , on remarque que LRMAP et PDR sont proches à  $\eta \rightarrow 0$ , ce qui confirme les résultats obtenus en (14) et (15). Pour deux autres valeurs de  $\eta$  (centre et droite), les performances des deux tests sont inversées. Comme mentionnée

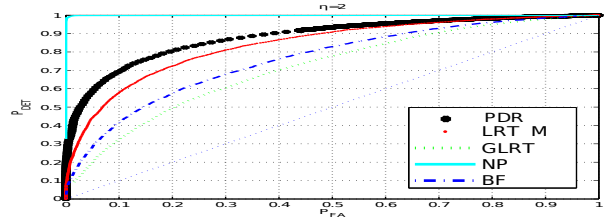


FIGURE 3 – Cas vectoriel :  $\eta = 2$ . Courbes CORE pour les tests du GLR, BF, LRMAP, PDR et de NP.

précédemment, l'étude théorique comparée de ces tests est à poursuivre, mais elle est analytiquement difficile. Des applications de ces travaux à la détection de données hyperspectrales sont présentées en [8].

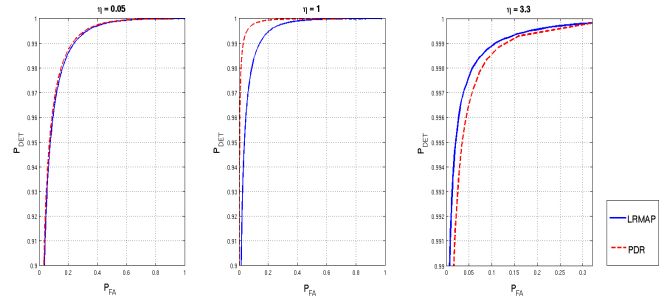


FIGURE 4 – Cas vectoriel : Courbes CORE pour les tests du LRMAP et du PDR pour 3 différentes valeurs du paramètre  $\eta$ . Pour  $\eta = 0.05$  (à gauche), les deux tests coïncident ; pour  $\eta = 1$  (au centre), le PDR est supérieur au LRMAP tandis que pour  $\eta = 3.3$  (à droite) c'est l'inverse.

## Références

- [1] A. Antoniadis and J. Fan. Regularization of wavelet approximations. *Journal of the American Statistical Association*, 96(455) :939–955, 2001.
- [2] S. Basu. Bayesian hypotheses testing using posterior density function. *Statistics and Probability Letters*, (30), 1996.
- [3] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems equations to sparse modeling of signal images. *SIAM Review*, 51(1) :34–81, 2009.
- [4] J. Fan. Test of significance based on wavelet thresholding and Neyman's truncation. *Journal of the American Statistical Association*, 91(434) :674–688, 1996.
- [5] J. J. Fuchs. The generalized likelihood ratio test and the sparse representation approach. *ICISP 2010, LNCS 6134*, 2010.
- [6] S. Mallat. *A wavelet tour of signal processing : the sparse way*. Academic Press, 2008.
- [7] S. Paris, D. Mary, and A. Ferrari. Composite hypothesis tests for sparse parameters. *IEEE SSP 2011*.
- [8] S. Paris, D. Mary, and A. Ferrari. Sparsity-based composite detection tests. Application to astrophysical hyperspectral data. *IEEE EUSIPCO 2011*.