

---

# K-produits, un critère de classification pour le traitement du signal

Michel Terré<sup>1</sup>, Luc Féty<sup>1</sup>, Nicolas Paul<sup>2</sup>

1. CNAM – Laboratoire CEDRIC  
Équipe LAETITIA  
292 rue Saint Martin, 75143, Paris cedex 03  
michel.terre@cnam.fr, luc.fety@cnam.fr
2. EDF R&D – Département STEP  
6 Quai Watier, 78400 Chatou cédex  
nicolas.paul@edf.fr

---

*RÉSUMÉ.* Cet article passe en revue plusieurs domaines d'applications du critère de classification « K-produits » introduit dans [10]. Ce critère est appliqué à l'estimation de mélange de lois, à l'extraction de lignes droites dans des images binaires et à l'estimation aveugle de canal en radiocommunications. Plusieurs algorithmes d'optimisation du critère sont présentés et des comparaisons avec d'autres approches sont conduites.

*ABSTRACT.* This article presents several applications of the clustering criterion « K-products » introduced in [10]. This criterion is applied to the unsupervised mixture estimation, the straight lines extraction in binary images and the blind channel estimation in radiocommunications. Some optimization algorithms are proposed and comparisons with other known methods are presented.

*MOTS-CLÉS :* classification, détection de lignes droites, égalisation aveugle.

*KEYWORDS:* clustering, straight lines detection, blind equalization.

---

DOI:10.3166/TS.27.221-239 © 2010 Lavoisier, Paris

## Extended Abstract

The *K-products* algorithm is a clustering algorithm that could be compared to the well known *K-means* algorithm. Given a set of univariate observations  $\{x_i\}_{i \in [1, N]}$  of a *K*-components mixture, the *K-products* (KP) algorithm estimates the mixture component expectations  $\mathbf{m} = (m_1, \dots, m_K)$ . The algorithm is based on the minimization of the KP criterion  $J_{KP}(u)$  defined by:

$$J_{KP}(u) = \sum_{i=1}^N \prod_{k=1}^K |x_i - u_k|^2 \quad (1)$$

The first intuitive motivation for defining this criterion is its behaviour in the limit case, when the variances of the components are null. In this particular case, all the observations are equal to one of the  $m_k$  so  $J_{KP}(\mathbf{m}) = 0$  and  $J_{KP}(\mathbf{u})$  is minimal in  $\mathbf{u} = \mathbf{m}$ . The second motivation to define the KP criterion is that, in the general case, it does not have any local non-global minima. The global minimum of  $J_{KP}$  can be reached for instance by the simple relaxation algorithm or by a non-iterative algorithm. It provides a biased but useful estimation of the mixture component expectation which can be used to cluster the data observation (nearest neighbour classification) and separate the *K* mixture components.

The *K-products* algorithm is analysed in this paper and its extensions to different other application fields are presented.

1) We introduce a new algorithm for extracting several straight lines in a binary image (*LK-products*). The modified criterion is presented hereafter:

$$J_{LKP}(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^N \prod_{k=1}^K \frac{(y_i - a_k x_i - b_k)^2}{(1 + a_k^2)} \quad (2)$$

where the observation is a set of couples  $\{x_n, y_n\}_{n \in [1, N]}$  which indicate the coordinates of the black pixels in the binary image. Among this set of (potentially noisy) observations we are looking for *K* Cartesian straight lines equations defined by parameters  $\{a_k, b_k\}_{k \in [1, K]}$ :  $y = a_k x + b_k$ .

We compare this algorithm with the well-known Hough Transform and we show that, in many simulation cases, performances obtained are greater than those obtained with the Hough Transform. The main advantage of this new algorithm is the absence of threshold process or clustering stages. The only parameter of the algorithm is a number of iterations for a relaxation process. The straight lines extraction in binary images is involved in many algorithms and the new proposed algorithm offers finally an interesting non-parametric approach to this problem.

2) We introduce a blind channel estimation algorithm (*HK-products*), based on a constrained clustering approach. The proposed algorithm is devoted to the case of an OFDMA uplink with high order *K*-ary modulations. The aim of the study is to reduce as much as possible the number of pilots on the uplink of an OFDM access radio network. We show that dealing with a highly reduced number of received symbols (<50), the proposed algorithm is able to give a good channel estimation.

The modified criterion is presented hereafter:

$$J_{HKP}(\hat{H}) = \sum_{i=1}^N \prod_{k=1}^K |x_i - \hat{H}s_k|^2 \quad (3)$$

where  $x_i = Hd_i + n_i$

$n_i$  is stationary, zero-mean, complex additive noise.

$\{d_i\}$  is the zero-mean, iid, unknown symbol sequence belonging to a  $K$ -ary QAM set.

$H$  represents the frequency response of the channel for the considered sub-carrier.

In conclusion, the paper proposes extensions of the  $K$ -products approach to different application fields. Moreover, concerning  $K$ -products, some analytical results are demonstrated in the paper:

- Unique solution for the bidimensionnal case (complex data) independently of the number of components of the mixture.
- Unique solution for any dimension for a mixture restricted to two components.
- Equivalence between the  $HK$ -products criterion, the maximum Likelihood criterion optimized through the Viterbi-Viterbi algorithm.

## 1. Introduction

Devoir classer un ensemble de  $N$  observations  $\{x_i\}_{i \in [1, N]}$  dans  $K$  classes est un problème classique [1] (*clustering*) qui se rencontre dans de nombreux domaines d'applications du traitement du signal ou du traitement d'images.

Ce problème de classification est aussi souvent présenté comme l'estimation de mélanges de lois : Les mélanges de lois [2][3] constituent un outil général utilisé dans les problèmes de modélisation et d'apprentissage statistique.

Les méthodes classiques d'estimation de mélanges de lois, telles que l'algorithme EM [4] ou l'algorithme  $K$ -means [5][6], proposent de rechercher un ensemble de paramètres (espérances et variances des classes par exemple) qui optimisent des critères tels que la log-vraisemblance, éventuellement contrainte [7], ou la somme des carrés des distances par classe ( $K$ -means).

Les optima globaux de ces critères sont en général non biaisés, ils tendent vers l'ensemble des paramètres du mélange lorsque le nombre d'observations tend vers l'infini, mais ils sont difficiles à atteindre. En effet ces critères possèdent de nombreux points stationnaires différents de leur optimum global. En pratique, même dans le cas le plus simple où les supports des densités de probabilités des composantes du mélange sont disjoints, l'algorithme EM ou l'algorithme  $K$ -means doivent être relancés [8][9] plusieurs fois pour obtenir une estimation correcte des paramètres du mélange.

Pour résoudre ce problème, nous avons proposé d'introduire un nouveau critère, appelé *K-produits* [10]. Pour des observations de dimension 1 ou 2, il a été montré que *K-produits* n'admet aucun minimum local non global. Le critère a en fait  $K!$  minimum globaux qui sont les  $K!$  permutations d'un vecteur (à  $K$  composantes). Le critère n'a donc aucun minimum local non-global.

Il a aussi été montré que le minimum global de ce critère est accessible analytiquement [3].

Dans cet article nous commencerons par rappeler la définition et les propriétés de ce critère *K-produits*, puis nous exposerons trois domaines d'applications : l'estimation des mélanges de lois, l'extraction de lignes droites en traitement d'images et l'estimation aveugle de canal en radiocommunications.

## 2. K-produits : critère et minimisation

### 2.1. Observations réelles ou complexes

Partant d'un ensemble d'observations réelles ou complexes  $\{x_i\}_{i \in [1, N]}$  d'un mélange de  $K$  composantes, l'algorithme proposé estime les espérances  $\mathbf{m} = (m_1, \dots, m_K)$  des  $K$  lois. L'algorithme est basé sur la minimisation du critère  $J_{KP}(u) = J_{KP}(u_1, \dots, u_K)$  défini par :

$$J_{KP}(u) = \sum_{i=1}^N \prod_{k=1}^K |x_i - u_k|^2 \tag{4}$$

Une justification intuitive de ce critère provient du fait que, dans le cas limite où les variances des composantes sont nulles, alors toutes les observations sont égales à un des  $m_k$ . Dès lors  $J_{KP}(\mathbf{m}) = 0$  et  $J_{KP}(\mathbf{u})$ , qui est un critère positif, est ainsi minimal pour  $\mathbf{u} = \mathbf{m}$ .

En développant l'équation (4), on obtient [10] :

$$J_{KP}(u) = \sum_{i=1}^N \left| x_i^K - c_1 x_i^{K-1} - c_2 x_i^{K-2} - \dots - c_K \right|^2 \tag{5}$$

Avec :  $c_p = (-1)^{p+1} \sum_{i_1=1}^K \sum_{i_2=i_1+1}^K \dots \sum_{i_p=i_{p-1}+1}^K u_{i_1} u_{i_2} \dots u_{i_p}$

La minimisation de  $J_{KP}(\mathbf{u})$  est donc équivalente à la prédiction linéaire de  $x_i^K$  à partir de la base de prédiction :  $\{x_i^{K-1}, x_i^{K-2}, \dots, x_i, 1\}$ .

On introduit alors les moments estimés des observations :

$$m_{p,q} = \frac{1}{N} \sum_{i=1}^N x_i^p \bar{x}_i^q \tag{6}$$

La solution de la prédiction linéaire conduit à :

$$\begin{pmatrix} m_{0,0} & m_{0,1} & m_{0,2} & \dots & m_{0,K-1} \\ m_{1,0} & m_{1,1} & & & \\ m_{2,0} & & m_{2,2} & & \\ \vdots & & & \ddots & \\ m_{K-1,0} & & & & m_{K-1,K-1} \end{pmatrix} \begin{pmatrix} c_K \\ c_{K-1} \\ c_{K-2} \\ \vdots \\ c_1 \end{pmatrix} = \begin{pmatrix} m_{0,K} \\ m_{1,K} \\ m_{2,K} \\ \vdots \\ m_{K-1,K} \end{pmatrix} \quad (7)$$

La résolution de ce système linéaire donne alors directement les termes  $\{c_i\}_{i \in [1,K]}$ . Le minimum de  $J_{KP}(u)$ , représenté par  $(\widehat{m}_k)$ , est le vecteur des racines d'un polynôme de coefficients  $\{c_i\}_{i \in [1,K]}$  :

$$(\widehat{m}_k) \equiv \text{Racines} (z^K - c_1 z^{K-1} - c_2 z^{K-2} - \dots - c_K) \quad (8)$$

*Application au mélange de lois*

Les figures 1 et 2 montrent les minima du critère *K-produits* dans le cas de l'observation d'un mélange de 5 lois Gaussiennes bidimensionnelles avec différentes matrices de covariances.

On observe un biais, relativement faible qui dépend de la covariance des composantes du mélange : lorsque le nombre d'observations tend vers l'infini, l'ensemble de points qui minimise *K-produits* ne tend pas vers l'ensemble des espérances du mélange. Ce biais a été démontré théoriquement [3] en dimension 1 et en dimension 2 pour  $K = 2$ .

Pour minimiser ce biais, on peut recourir à une dernière étape d'allocation des observations aux modes, au sens des plus proches voisins. Il suffit alors de recalculer les moyennes des classes pour obtenir une estimation nettement moins biaisée des espérances du mélange.

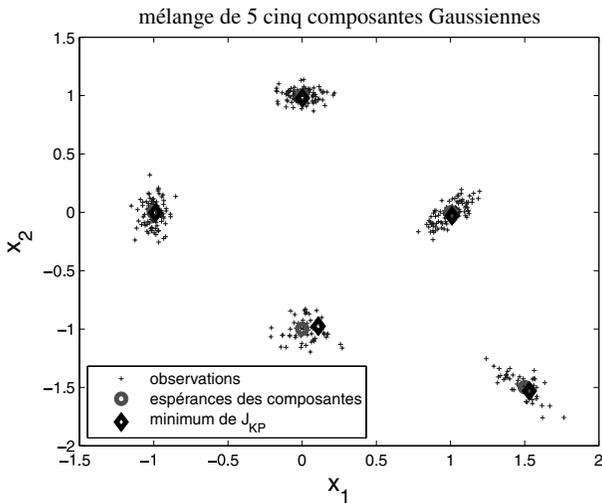


Figure 1. Observations issues de 5 lois gaussiennes complexes (axes réels et imaginaires) d'écart type 0,1

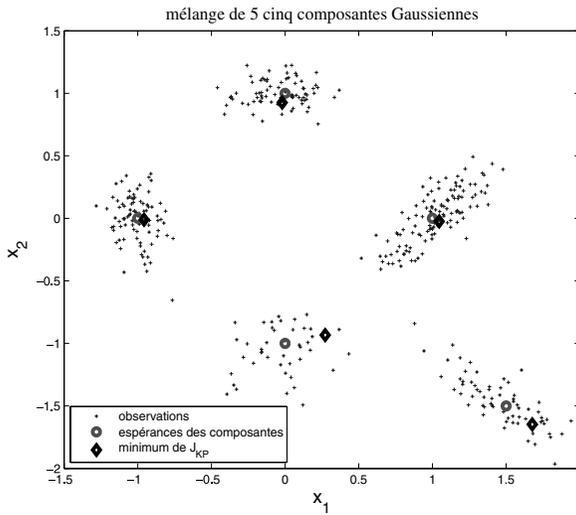


Figure 2. Observations issues de 5 lois gaussiennes complexes (axes réels et imaginaires) d'écart type 0,2

Concernant le biais il a aussi été montré [3] que dans le cas d'observations bidimensionnelles, pour  $K=2$  et en considérant un mélange de deux lois gaussiennes complexes équiprobables, de moyennes et variances respectives  $(a, \sigma_1^2 I_{2 \times 2})$  et  $(-a, \sigma_2^2 I_{2 \times 2})$ , alors la solution analytique de  $K$ -produits conduisait au deux estimations suivantes :  $-a \left( \sqrt{1 + \rho^2} - \rho \right), a \left( \sqrt{1 + \rho^2} + \rho \right)$  (cf. annexe 2) avec

$\rho = \frac{\sigma_2^2 - \sigma_1^2}{\|a\|^2 + \sigma_1^2 + \sigma_2^2}$ . Dans ce cas particulier, on constate que les moyennes sont estimées sans biais si  $\sigma_1^2 = \sigma_2^2$  ou si les variances des composantes sont négligeables devant l'écart entre les deux modes.

**2.2. Observations de dimensions supérieures à 2**

La solution exposée précédemment fonctionne pour des observations unidimensionnelles (réelles) ou bidimensionnelles (complexes). Par contre elle ne se généralise pas directement aux dimensions supérieures. Dans de tels cas, on peut utiliser un algorithme de relaxation. Pour cela on écrit le critère de la manière suivante :

$$J_{KP}(u) = \sum_{i=1}^N C_{k,i} \|x_i - u_k\|^2 \tag{9}$$

avec  $C_{k,i} = \prod_{\substack{k'=1 \\ k' \neq k}}^K \|x_i - u_{k'}\|^2$

On propose alors d'annuler successivement les dérivées partielles du critère par rapport aux différents termes  $u_k$  :  $\frac{\partial J_{KP}(\mathbf{u})}{\partial u_k} = 0$ . À la  $n^{ième}$  itération, l'annulation de cette dérivée partielle conduit ainsi à :

$$u_k^n = \sum_{i=1}^N \left( \frac{C_{k,i}^{n-1}}{\sum_{m=1}^N C_{k,m}^{n-1}} \right) x_i \tag{10}$$

$C_{k,i}^{n-1}$  représente le terme calculé avec des estimations obtenues à l'issue de l'itération  $n - 1$ , i.e avec des termes du type  $u_k^{n-1}$ . L'algorithme va alors mettre à jour successivement, et plusieurs fois, l'ensemble des termes  $u_k$ . Cette approche itérative ne peut se stabiliser que sur des minima du critère. La question de l'existence de minima secondaires est alors cruciale pour pouvoir affirmer que cet algorithme itératif converge.

Or l'unicité du minimum du critère a été prouvée pour les dimensions 1 et 2 [3]. Pour les dimensions supérieures, on peut montrer (voir annexe 1) l'unicité du minimum dans le cas bimodal ( $K = 2$ ) dans le cas limite où les matrices de covariance des composantes sont nulles. La preuve de cette unicité reste à faire dans le cas où il y a plus de deux modes ( $K > 2$ ) et pour des matrices de covariance quelconque.

On notera que, jusqu'à maintenant, les simulations effectuées pour ces dimensions supérieures, n'ont jamais fait apparaître de minima secondaires.

### 2.3. Modèle paramétrique

Le critère *K-produits* peut être étendu dans le cas de fonctions paramétriques.

$$J_{KP}(\mathbf{u}) = \sum_{i=1}^N \prod_{k=1}^K f_k(z_i) \tag{11}$$

La fonction  $f_k(\cdot)$  fait intervenir un vecteur de paramètres  $u_k$ . L'observation, notée ici  $z_i$ , est éventuellement multidimensionnelle. On peut illustrer ce cas avec le problème d'extraction de lignes droites à partir d'une image binaire.

On dispose alors de couples d'observations, qui sont les coordonnées :  $z_i = (x_i, y_i)$ , de points appartenant à des droites. On cherche un ensemble de  $K$  équations de droites qui « expliquent » au mieux cet ensemble de coordonnées. Le problème revient donc à identifier les  $K$  couples  $u_k = \{a_k, b_k\}_{k \in [1, K]}$  tels que  $y_i = a_k x_i + b_k$ . Issu de *K-produits*, nous proposons donc le critère suivant :

$$J_{LKP}(a, b) = \sum_{i=1}^N \prod_{k=1}^K \frac{(y_i - a_k x_i - b_k)^2}{(1 + a_k^2)} \tag{12}$$

Ce nouveau critère (appelé *LK-produits*) représente les produits des distances des observations aux  $K$  équations de droites. L'optimisation peut être effectuée au moyen

d'un algorithme de relaxation qui cherche à déterminer successivement les paramètres des  $K$  droites. Il procède par annulations successives des dérivées partielles du critère par rapport aux différents paramètres des droites à identifier.

Il a été montré [11] que cette stratégie offre une alternative intéressante à la Transformée de Hough [12] étendue au cas de plusieurs droites. Des exemples d'applications sont donnés sur les figures 3-6.

La figure 3 correspond à un cas assez facile avec trois droites et peu de bruit. On constate que les deux algorithmes donnent des estimations précises des coefficients des trois droites.

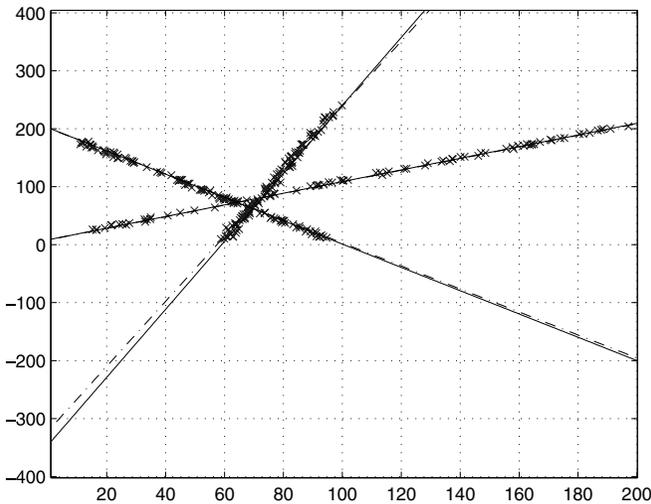


Figure 3. 3 lignes droites, variance du bruit ajouté sur les coordonnées des pixels  $\sigma^2 = 1$ , Transformée de Hough en trait plein, LK-produits en trait pointillé, observations représentées par des croix

Tableau 1. Estimation des paramètres des droites par LK-produits (LKP) et Hough (HT). 3 lignes,  $\sigma^2 = 1$

$(a; b)$	6; -350	1; 10	-2; 200
HT	6; -351	1; 10	-2; 199
LKP	5.7; -326	1; 12	-2; 199

La figure 4 correspond à un cas un peu plus difficile avec trois droites et un niveau de bruit plus important.

En effectuant les deux algorithmes sur plusieurs réalisations du bruit, on constate une moindre dispersion pour les solutions identifiées par l'approche *LK-produits* que pour celles identifiées au moyen de la Transformée de Hough. Ceci est illustré par la

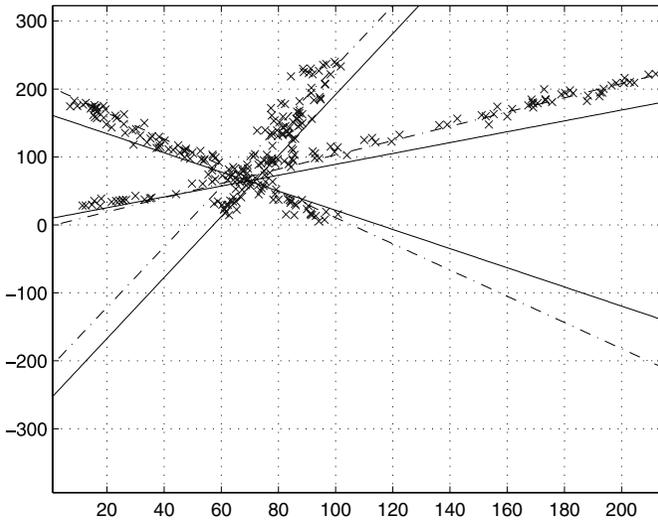


Figure 4. 3 lignes droites, variance du bruit ajouté sur les coordonnées des pixels  $\sigma^2 = 4$ , Transformée de Hough en trait plein, LK-produits en trait pointillé, observations représentées par des croix

figure 5. La figure 6 correspond à un cas difficile avec cinq droites et un niveau de bruit important.

On constate sur l'exemple de la figure 6 que la Transformée de Hough converge parfois vers des solutions erronées et ceci se traduit, en définitive, par des équations de lignes droites qui se croisent, afin de correspondre au mieux aux observations.

Tableau 2. Estimation des paramètres des droites par LK- produits (LKP) et Hough (HT). 3 lignes,  $\sigma^2 = 4$

<b>(a ; b)</b>	6; -350	1; 10	-2; 200
<b>HT</b>	4.5; -252	0.8; 10	-1.4; 161
<b>LKP</b>	4.4; -205	1; -1	-1.9; 201

Tableau 3. Estimation des paramètres des droites par LK-produits (LKP) et Hough (HT). 5 lignes,  $\sigma^2 = 4$

<b>(a ; b)</b>	6; -350	1; 10	-2; 200	-3; 350	1; 50
<b>HT</b>	3.9; -223	1.4; 0	-1.2; 161	-0.3; 62	-0.21; 168
<b>LKP</b>	5.4; -308	1; 7	-1.3; 191	-3.2; 357	1; 62

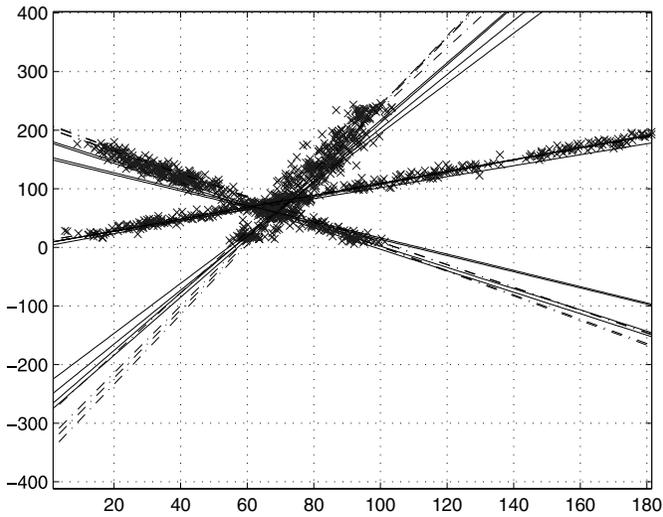


Figure 5. 4 réalisations différentes de bruit pour une simulation avec 3 lignes droites, variance du bruit ajouté sur les coordonnées des pixels  $\sigma^2 = 4$ , Transformée de Hough en trait plein, LK-produits en trait pointillé, observations représentées par des croix

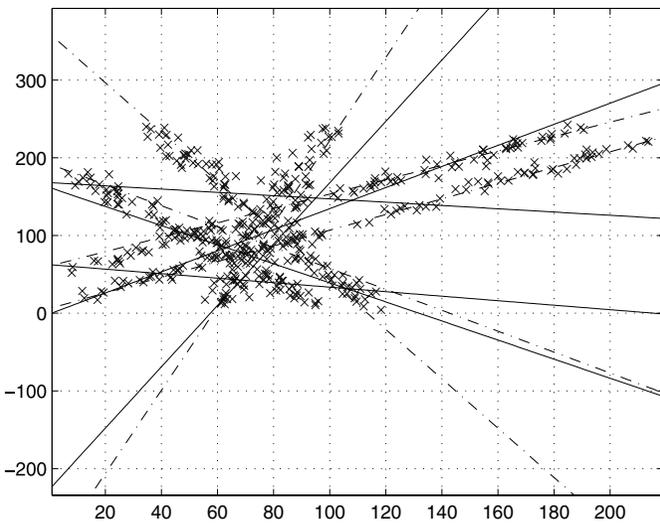


Figure 6. 5 lignes droites, variance du bruit ajouté sur les coordonnées des pixels  $\sigma^2 = 4$ , Transformée de Hough en trait plein, LK-produits en traits pointillé, observations représentées par des croix

### 3. Application à l'estimation aveugle de canal

On considère ici la réception de symboles de communications appartenant à une modulation d'amplitude et de phase *MDAQ* (*QAM*). Il s'agit d'identifier, sans séquence d'apprentissage (« en aveugle »), la réponse complexe  $H$  du canal de propagation. On reçoit ainsi des symboles  $x_i$  :

$$x_i = Hs_i + w_i \quad (13)$$

où  $w_i$  représente un bruit complexe additif et  $H$  la réponse du canal. Ce modèle simple correspond idéalement à l'observation faite sur une porteuse, à l'issue de la FFT de réception, dans le cas typique d'une transmission en OFDMA. On propose alors d'identifier la transformation qu'il faut appliquer à la constellation de départ pour rendre compte au mieux des observations.

Si on note  $(s_k)_{k \in [1, K]}$  les  $K$  symboles de cette constellation, alors on peut proposer une modification [13] du critère *K-produits* :

$$J_{HKP}(\hat{H}) = \sum_{i=1}^N \prod_{k=1}^K |x_i - \hat{H}s_k|^2 \quad (14)$$

$N$  représente toujours le nombre d'observations mais  $K$  représente maintenant le nombre de symboles de la modulation utilisée (4, 16, 64, ...). On appellera dorénavant *HK-produits* ce critère modifié.

Une approche relativement similaire a été proposée dans [14], pour le cas d'une réception sur plusieurs antennes. Dans cette référence les auteurs cherchent à agir sur les observations et non sur la constellation comme nous le faisons ici. De plus, ils restreignent leurs travaux au cas des modulations dont tous les symboles sont de même puissance : constellations circulaires de type modulation de phase *MDP* (*PSK*). Ils proposent des équivalences avec le critère du maximum de vraisemblance.

Pour minimiser, dans le cas général des modulations *MDAQ*, le critère proposé à l'équation (14), des propriétés de symétrie de ces modulations peuvent être utilisées. Nous savons que  $\{s_k, js_k, -s_k, -js_k\}$  appartiennent à la constellation. On utilise alors la propriété suivante :

$$(x_i - \hat{H}s_k)(x_i - j\hat{H}s_k)(x_i + \hat{H}s_k)(x_i + j\hat{H}s_k) = x_i^4 - \hat{H}^4 s_k^4 \quad (15)$$

Plusieurs équivalences entre le critère *HK-produits*, le critère du maximum de vraisemblance et l'approche proposée par l'algorithme de Viterbi-Viterbi [16] peuvent être identifiées (cf. annexe 3).

On introduit ensuite la nouvelle observation :  $y_i = x_i^4$ , le nouveau pseudo symbole  $v_k = s_k^4$ , et la variable complexe  $z = \hat{H}^4$ . Dès lors le critère de l'équation (14) devient, en introduisant  $M = K/4$  :

$$J_{HKP}(z) = \sum_{i=1}^N \prod_{k=1}^M |y_i - zv_k|^2 \quad (16)$$

On introduit alors :

$$c_0 = 1, c_1 = - \sum_{k=1}^M v_k, c_2 = \sum_{i=1}^M \sum_{j=i+1}^M v_i v_j, \dots$$

$$c_p = (-1)^p \sum_{i_1=1}^M \sum_{i_2=i_1+1}^M \dots \sum_{i_p=i_{p-1}+1}^M v_{i_1} v_{i_2} \dots v_{i_p}, c_M = \prod_{k=1}^M v_k,$$

$$m_{p,q} = \frac{1}{N} \sum_{i=1}^N y_i^p \bar{y}_i^q \text{ et } a_{n,m} = c_n \bar{c}_m m_{M-n, M-m}.$$

Le critère peut alors être ré-exprimé sous la forme :

$$J_{HKP}(z) = \sum_{n=0}^M \sum_{m=0}^M a_{n,m} z^n \bar{z}^m \tag{17}$$

À cette étape il est judicieux d'utiliser une représentation polaire :  $z = \rho e^{j\theta}$ . On obtient alors :

$$J_{HKP}(\rho, \theta) = \sum_{n=0}^M \sum_{m=0}^M a_{n,m} \rho^{n+m} e^{j(n-m)\theta} \tag{18}$$

La minimisation de  $J_{HKP}(\rho, \theta)$  suivant les deux paramètres est alors une étape difficile. Une approche par relaxation entre  $\rho$  et  $\theta$  peut alors être proposée. Elle conduit, après quelques changements de variables, aux solutions simples suivantes :

cas  $\rho$  connu

$$J_{HKP}(\theta) = \text{Re} \left\{ \sum_{m=0}^M a_{m,m} \rho^{2m} + 2 \sum_{t=1}^M \sum_{m=0}^{M-t} a_{m+t,m} \rho^{t+2m} e^{jt\theta} \right\} \tag{19}$$

$$\hat{\theta} = \underset{\theta}{\text{Argmin}} \text{Re} \left\{ \text{IFFT} \left\{ \frac{\alpha_0}{2}, \alpha_1, \alpha_2, \dots, \alpha_M \right\} \right\} \tag{20}$$

avec :

$$\alpha_t = 2 \sum_{m=0}^{M-t} a_{m+t,m} \rho^{t+2m} \tag{21}$$

La phase estimée  $\hat{\theta}$  est donc simplement donnée par la minimisation de la partie réelle d'une Transformée de Fourier Discrète Inverse (IFFT) :

cas  $\theta$  connue

$$J(\rho) = \sum_{t=0}^{2M} \sum_{m=t-M}^t a_{t-m,m} \rho^t e^{j(t-2m)\theta} \tag{21}$$

$$\hat{\rho} = \underset{\rho \in \{\text{racines réelles positives de } J(\rho)\}}{\text{Argmin}} \{J(\rho)\} \quad (23)$$

On montre alors que, même avec très peu d'observations, cette approche permet d'identifier la réponse du canal avec précision. La figure 7 illustre la variance de l'estimation du canal en fonction de la taille de l'observation ( $N$ ). Les performances sont comparées à celles d'un algorithme de référence [15] pour l'estimation aveugle de la phase du canal.

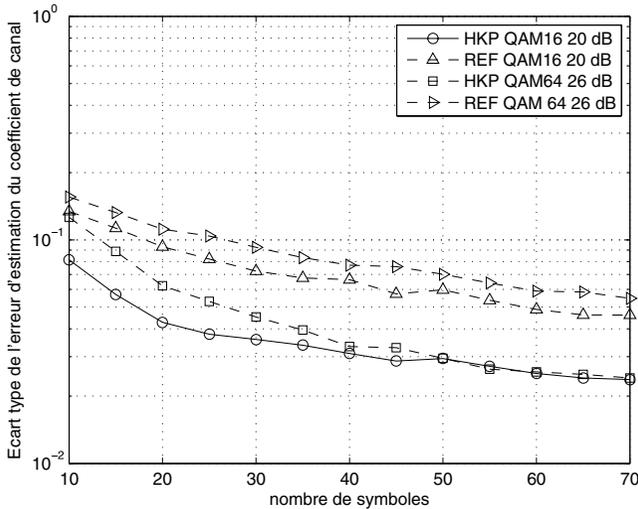


Figure 7. Comparaisons des performances obtenues avec l'approche HK-produits (HKP) pour l'estimation du canal par rapport à une méthode de référence (REF). Les courbes sont tracées en fonction du nombre de symboles observés

#### 4. Conclusion

Nous avons passé en revue un ensemble d'applications du critère  $K$ -produits pour la classification de données. Des résultats prometteurs ont été présentés pour l'application de ce critère dans des domaines aussi différents que l'estimation non supervisée des mélanges de lois, l'extraction de plusieurs droites dans des images ou l'estimation aveugle d'un canal en radiocommunications.

L'unicité de la solution du critère a été démontrée pour des données bi-dimensionnelles, indépendamment du nombre de modes et pour des données de dimensions quelconques en l'absence de bruit dans le cas restreint de deux modes. Parallèlement l'algorithme d'optimisation itérative n'a jamais identifié de minimums secondaires du critère en dimensions quelconques et quel que soit le nombre de modes. On peut espérer que la démonstration de l'unicité (sans considérer les permutations) de la solution doit être atteignable. C'est un axe de travail en cours.

## Bibliographie

- [1] Xu .R, Wunsch D. (2005). « Survey of clustering algorithms, » *IEEE trans. on Neural Networks*, vol 16, pp. 645-678.
- [2] O. Cappé (2000). Modèles de mélange et modèles de Markov cachés pour le traitement automatique de la parole, notes de cours, ENST, Paris.
- [3] N. Paul (2008). Une nouvelle approche pour l'estimation des mélanges de lois, application à la radio cognitive, thèse de doctorat, CNAM, Paris, soutenue le 14 mai 2008.
- [4] G. McLachlan, T. Krishnan (1997). *The EM algorithm and extensions*, John Wiley and Sons, New York.
- [5] J. Hartigan, M. Wong (1979). « A K-means algorithm, » *Journal of Applied Statistics*, vol 28, pp. 100-108.
- [6] H. H. Bock (2007). « Clustering methods: A history of the K-means algorithm, » *in Selected contributions in data Analysis and Classification*, pp. 161-172, Springer Verlag, Berlin.
- [7] D. Ormoneit, V. Tresp (1998). « Averaging maximum penalized likelihood and Bayesian estimation for improving Gaussain mixture probability density estimate, » *IEEE trans. on Neural Network*, vol 9, pp. 639-649.
- [8] C. Biernacki, G. Celeux, G. Govaert (2003). « Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models, » *Computational Statistics and Data Analysis*, vol 41, pp. 561-575.
- [9] P.S. Bradley, U.M. Fayyad (1998). « Refining initial points for K-means clustering, » *in proc. of the 15<sup>th</sup> Int. Conf. on Machine Learning*, pp. 91-99, San Francisco.
- [10] N. Paul, M. Terré, L. Féty (2007). « A global algorithm to estimate the expectations of the components of an observed univariate mixture, » *Advances in Data Analysis and Classification*, Volume 1, pp. 201-219.
- [11] C. Bidot, N. Paul, L. Féty, M. Terré (2008). « A new optimisation criterion for extracting several straight lines in a binary image, » *in proc EUSIPCO*, Lausanne, August.
- [12] Henri Maitre (1986). « Contribution to the prediction of performances of the Hough transform, » *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 8, Issue 5, pp. 669-674, September.
- [13] M. Terré, L. Féty, N. Paul (2009). « Optimisation of a clustering-based criterion for blind channel estimation, » *in proc. SPAWC 2009*, Perugia, June.
- [14] O. Grellier, P. Comon (1998). « Blind separation of Discrete Sources, » *IEEE signal processing letters*, vol 5, n°8, pp. 212-214 August.
- [15] E. Serpedin, P. Ciblat, G. B. Giannakis, and P. Loubaton (2001). « Performance Analysis of Blind Carrier Phase Estimators for General QAM Constellations, » *IEEE trans. on Signal Processing*, vol. 49, no. 8, August.
- [16] A.J Viterbi, A.M Viterbi (1983). « Nonlinear Estimation of PSK-modulated Carrier Phase with Application to Burst Digital transmissions, » *IEEE trans. on Inform. Theory*, vol IYT 29, pp. 543-551, July.

### Annexe 1. Preuve de l'absence de minimums secondaires en dimension quelconque pour deux composantes ( $K = 2$ ) de matrice de covariance nulle

On considère le cas avec deux modes  $\mathbf{m}_1$   $\mathbf{m}_2$  distincts, de dimension  $D$  quelconque, et en absence de bruit. Les  $N$  observations ne peuvent donc être égales qu'à  $\mathbf{m}_1$ , ce qui se produit dans  $N_1$  cas ou égales à  $\mathbf{m}_2$ , ce qui se produit dans  $N_2$  cas ( $N = N_1 + N_2$ ).

Sans perte de généralité on suppose que  $\mathbf{m}_1$  est le vecteur nul (si ce n'est pas le cas il suffit de faire un changement de repère pour se retrouver dans cette configuration) et on pose  $\mathbf{m}_2 = \mathbf{m}$  ( $\mathbf{m}$  différent du vecteur nul). Le critère s'écrit alors :

$$J(\mathbf{u}_1, \mathbf{u}_2) = N_1 \|\mathbf{u}_1\|^2 \|\mathbf{u}_2\|^2 + N_2 \|\mathbf{u}_1 - \mathbf{m}\|^2 \|\mathbf{u}_2 - \mathbf{m}\|^2 \quad (24)$$

Les couples  $(\mathbf{u}_1, \mathbf{u}_2)$  qui annulent le gradient de  $J$  vérifient :

$$\mathbf{0} = N_1 \mathbf{u}_1 \|\mathbf{u}_2\|^2 + N_2 (\mathbf{u}_1 - \mathbf{m}) \|\mathbf{u}_2 - \mathbf{m}\|^2 \quad (25)$$

$$\mathbf{0} = N_1 \mathbf{u}_2 \|\mathbf{u}_1\|^2 + N_2 (\mathbf{u}_2 - \mathbf{m}) \|\mathbf{u}_1 - \mathbf{m}\|^2 \quad (26)$$

D'après (25) et (26)  $\mathbf{u}_1$  et  $\mathbf{u}_2$  sont colinéaires à  $\mathbf{m}$ . Il existe donc deux réels positifs  $a$  et  $b$  tels que :

$$\mathbf{u}_1 = a \cdot \mathbf{m} \quad (27)$$

$$\mathbf{u}_2 = b \cdot \mathbf{m} \quad (28)$$

En remplaçant (27) et (28) dans (25) et (26), on obtient :

$$\mathbf{0} = N_1 a \mathbf{m} b^2 \|\mathbf{m}\|^2 + N_2 (a - 1) \mathbf{m} (b - 1)^2 \|\mathbf{m}\|^2 \quad (29)$$

$$\mathbf{0} = N_1 b \mathbf{m} a^2 \|\mathbf{m}\|^2 + N_2 (b - 1) \mathbf{m} (a - 1)^2 \|\mathbf{m}\|^2 \quad (30)$$

Le vecteur  $\mathbf{m}$  étant non nul :

$$0 = N_1 a b^2 + N_2 (a - 1)(b - 1)^2 \quad (31)$$

$$0 = N_1 a^2 b + N_2 (a - 1)^2 (b - 1) \quad (32)$$

En calculant a.(31) - b.(32), on obtient :

$$0 = N_2 (a - 1)(b - 1)(b - a) \quad (33)$$

Les seuls points stationnaires sont donc :

$a = 1$  (alors  $b = 0$  d'après (31)), il s'agit d'un minimum global qui annule le critère.

$b = 1$  (alors  $a = 0$  d'après (32)), il s'agit d'un minimum global qui annule le critère. Il s'agit en fait de la permutation de la première solution.

$a = b$ , dans ce cas on a donc  $\mathbf{u}_1 = \mathbf{u}_2$ , et l'équation (31) conduit à :

$$N_1 a^3 = N_2 (1 - a)^3 \quad (34)$$

$$a = \frac{N_2^{1/3}}{N_1^{1/3} + N_2^{1/3}} \quad (35)$$

Il est alors possible d'évaluer le critère à partir de l'équation (24) et de montrer qu'il s'agit d'un maximum du critère. Dans le cas particulier d'équiprobabilité des deux modes, c'est à dire pour  $N_1 = N_2$ , on obtient  $a = 0,5$ , d'où  $\mathbf{u}_1 = \mathbf{u}_2 = 0,5\mathbf{m}$ .

## Annexe 2. Preuve de l'absence de biais en dimension deux dans le cas de deux composantes ( $K = 2$ ) équiprobables gaussiennes de même variance

En dimension 2 ( $x_i \in \mathbb{R}^2$ ) on représente l'observation sous la forme d'un nombre complexe  $z_i = x_{i1} + jx_{i2}$ . La densité de probabilité de la variable aléatoire  $Z$  s'écrit donc :

$$f_Z(z) = \frac{0.5}{2\pi} \left( \frac{1}{\sigma_1^2} \exp\left(-\frac{\|z + a\|^2}{2\sigma_1^2}\right) + \frac{1}{\sigma_2^2} \exp\left(-\frac{\|z - a\|^2}{2\sigma_2^2}\right) \right) \quad (36)$$

On peut alors calculer

$$E_Z [Z] = 0 \quad (37)$$

$$E_Z [\bar{Z}] = 0 \quad (38)$$

$$E_Z [\bar{Z}Z] = \|a\|^2 + \sigma_1^2 + \sigma_2^2 \quad (39)$$

$$E_Z [Z^2] = a^2 \quad (40)$$

$$E_Z [Z^2 \bar{Z}] = 2a(\sigma_2^2 - \sigma_1^2) \quad (41)$$

L'équation (7) devient alors

$$\begin{pmatrix} c_2 \\ c_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \|a\|^2 + \sigma_1^2 + \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} a^2 \\ 2a(\sigma_2^2 - \sigma_1^2) \end{pmatrix} \quad (42)$$

Ce qui conduit à :

$$\begin{pmatrix} c_2 \\ c_1 \end{pmatrix} = \begin{pmatrix} a^2 \\ 2a\rho \end{pmatrix} \quad (43)$$

$$\text{avec } \rho = \frac{\sigma_2^2 - \sigma_1^2}{\|a\|^2 + \sigma_1^2 + \sigma_2^2}$$

Le minimum du critère est alors donné par les racines du polynôme  $x^2 - 2a\rho x - a^2$ . Celles ci sont égales à :  $-a\left(\sqrt{1+\rho^2} - \rho\right)$ ,  $a\left(\sqrt{1+\rho^2} + \rho\right)$ .

On constate donc que la minimisation du critère conduit à une estimation biaisée des espérances des deux modes. Cependant dans le cas où les deux composantes ont la même variance ( $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ) alors  $\rho = 0$  et **l'estimation est non biaisée**.

Dans le cas d'observations unidimensionnelles d'un mélange de deux gaussiennes d'espérance (réelle)  $-a$  et  $+a$ , de même probabilité et de variance  $\sigma_1^2$  et  $\sigma_2^2$ , le même raisonnement permet de trouver le minimum du critère lorsque le nombre d'observations tends vers l'infini.

Les deux racines sont égales à :  $-a\left(\sqrt{\rho_1^2 + \rho_2} - \rho_1\right)$ ,  $a\left(\sqrt{\rho_1^2 + \rho_2} + \rho_1\right)$

$$\text{avec } \rho_1 = \frac{3(\sigma_2^2 - \sigma_1^2)}{2a^2 + \sigma_2^2 + \sigma_1^2} \text{ et } \rho_2 = 1 + \frac{\sigma_2^2 + \sigma_1^2}{2a^2}$$

Contrairement au cas bidimensionnel, l'estimation reste biaisée même lorsque les deux composantes ont des variances identiques. Dans un tel cas on a :

$$\sigma_1^2 = \sigma_2^2 = \sigma^2 \Rightarrow -a\left(\sqrt{1 + \frac{\sigma^2}{a^2}}\right), a\left(\sqrt{1 + \frac{\sigma^2}{a^2}}\right)$$

### Annexe 3. Critère *HK-produits*, algorithme de Viterbi-Viterbi et critère du maximum de vraisemblance

Si on se restreint au cas de la modulation MDAQ 4 (QPSK), les phases des symboles de la constellations sont telles que :  $\forall k, s_k^4 = a$ , où  $a$  représente une constante réelle qui est fonction du module des symboles de la constellation. Le critère *HK-produits* s'écrit alors :

$$J_{HKP}(\hat{H}) = \sum_{i=1}^N |x_i^4 - a\hat{H}^4|^2 \quad (44)$$

La solution est alors donnée par :

$$\hat{H}^4 = \frac{1}{aN} \sum_{i=1}^N x_i^4 \quad (45)$$

Si on s'intéresse juste à l'estimation de la phase d'un canal de module unité :  $H = e^{j\theta}$ , on obtient alors (pour  $a > 0$ ) :

$$\hat{\theta} = \frac{1}{4} \left( \text{angle} \left\{ \sum_{i=1}^N x_i^4 \right\} \pm 2m\pi \right), \quad m \in N \quad (46)$$

On reconnaît alors la solution donnée par l'algorithme de Viterbi-Viterbi [16].

On peut aussi remarquer, que dans le cas d'un canal de module unité, le critère de l'équation (44) devient :

$$J_{HKP}(\hat{\theta}) = \sum_{i=1}^N \left| x_i^4 - ae^{j4\hat{\theta}} \right|^2 \tag{47}$$

Ce critère peut alors être vu comme la maximisation de la probabilité conditionnelle  $p \left( \{x_i^4\}_{i=1, \dots, N} \mid \theta \right)$

En effet, les observations  $x_i$  s'écrivent de la manière suivante :

$$x_i = e^{j\theta} d_i + n_i \tag{48}$$

$d_i$  représente le symbole émis à l'instant « i » et  $n_i$  représente un terme de bruit additif complexe de moyenne nulle et de variance  $\sigma_n^2$ .

On a alors :

$$x_i^4 = e^{j4\theta} d_i^4 + 4e^{j3\theta} d_i^3 n_i + 6e^{j2\theta} d_i^2 n_i^2 + 4e^{j\theta} d_i n_i^3 + n_i^4 \tag{49}$$

ou encore :

$$x_i^4 = ae^{j4\theta} + u_i \tag{50}$$

avec :

$$u_i = 4e^{j3\theta} d_i^3 n_i + 6e^{j2\theta} d_i^2 n_i^2 + 4e^{j\theta} d_i n_i^3 + n_i^4 \tag{51}$$

Comme  $n_i$  est un bruit circulaire complexe gaussien, avec  $E \{n_i^p n_i^{m*}\} = 0 \forall m \neq p$ , on en déduit que  $u_i$  est un bruit gaussien de moyenne nulle dont la variance  $\sigma_u^2$  dépend uniquement du module des symboles MDAQ et de la variance  $\sigma_n^2$  du processus  $n_i$ .

On peut alors exprimer la probabilité conditionnelle  $p(x_i^4 \mid \theta)$  :

$$p(x_i^4 \mid \theta) = \frac{1}{\sqrt{2\pi}\sigma_u} \exp \left\{ -\frac{|x_i^4 - ae^{j4\theta}|^2}{2\sigma_u^2} \right\} \tag{52}$$

Comme les termes de bruit  $n_i$  sont indépendants, on en déduit que les termes  $u_i$  sont aussi indépendants, la probabilité conditionnelle du  $N$ -uplet observé s'écrit donc :

$$p \left( \{x_i^4\}_{i=1, \dots, N} \mid \theta \right) = \prod_{i=1}^N p(x_i^4 \mid \theta) \tag{53}$$

$$p\left(\{x_i^4\}_{i=1,\dots,N} \mid \theta\right) = \frac{1}{(\sqrt{2\pi}\sigma_u)} \exp\left\{-\frac{\sum_{i=1}^N |x_i^4 - ae^{4j\theta}|^2}{2\sigma_u^2}\right\} \quad (54)$$

Comme  $\sigma_u^2$  ne dépend pas de  $\theta$  la maximisation de  $p\left(\{x_i^4\}_{i=1,\dots,N} \mid \theta\right)$  est donc équivalente à la minimisation de l'équation (47). Ainsi, dans le cas de l'estimation de la phase d'un canal de module connu, le critère du maximum de vraisemblance est équivalent au critère *HK-produits*.

Article reçu le : 1/10/2009

Accepté le : 15/05/2010



**Michel Terré** est né en 1964, il est ingénieur de Télécom SudParis (INT 87), docteur du Conservatoire National des Arts et Métiers (1995) et habilité à diriger des recherches (HDR Paris XIII 2004). Il a passé 10 années chez TRT, Thalès et Alcatel avant de rejoindre le Conservatoire National des Arts et Métiers en 1998 où il est actuellement professeur des Universités. Michel Terré est membre émérite de la SEE et a été organisateur de plusieurs journées d'études SEE sur les applications du traitement du signal dans le domaine des communications numériques. Ses thèmes de recherche concernent le traitement du signal appliqué aux télécommunications.



**Luc Féty** est né à Chateau-Thierry en 1958. Il est diplômé de l'Université d'Orsay (82) et Docteur en sciences de Télécom Paris (88). Jusqu'en 1991, il a mené des activités de recherche à TRT puis Thomson TRT Défense en traitement d'antenne et en traitement d'image. Depuis 1991, il est maître de conférences au Conservatoire National des Arts et Métiers. Il poursuit des activités de recherche en traitement du signal : traitement d'antenne en radiocommunications, séparation de sources, transmissions multiporteuses.



**Nicolas Paul**. Diplômé de Supélec (1999) et docteur en traitement du signal (CNAM, 2008), Nicolas Paul a commencé sa carrière chez Alcatel, comme ingénieur de recherche en radiocommunication. Il travaille actuellement dans le département STEP (Simulation et Traitement de l'Information pour l'Exploitation des systèmes de Production) d'EDF R&D (Chatou), où ses activités principales sont la restauration de séquences vidéos et le contrôle non destructif.

