

# Algorithme QML approchant la borne de Cramer Rao pour l'estimation de phase et de fréquence

Valérie Cueff<sup>1</sup>, Michel Terré<sup>2</sup>

<sup>1</sup>Alcatel CIT  
5 rue Noël Pons, 92734 Nanterre Cedex

<sup>2</sup>CNAM  
292 rue Saint Martin, 75141 Paris Cedex 03

valerie.cueff@alcatel.fr, terre@cnam.fr

**Résumé** - L'article décrit un estimateur de phase et de fréquence approchant la borne de Cramer Rao. Cet algorithme, appelé QML, particulièrement adapté aux modulations de phase en mode paquet, consiste à tester des couples de phase et de fréquence. Une fonction de corrélation calculée à partir des symboles de référence et des symboles d'information permet alors d'estimer le couple phase-fréquence le plus probable. Les performances de l'algorithme peuvent être améliorées par un choix judicieux du nombre et de la position des symboles de référence.

**Abstract** - This paper describes a phase and frequency recovery algorithm nearing the Cramer Rao bound. This algorithm, called QML, is particularly convenient for burst-mode PSK signals, and consists in testing phase and frequency hypothesis. A correlation function, using both reference and information symbols allows the estimation of the more likely phase-frequency couple. The performances of the algorithm are increased by a convenient number and position of the reference symbols.

## 1. Introduction

Transmettre sur une voie radio un trafic de données de nature sporadique, tel que peut l'être celui d'une source internet ou d'une source multimedia, conduit à utiliser des transmissions par paquets. Cette tendance est visible dans les évolutions des différents projets de radiocommunications cellulaires terrestre ou satellitaires, pour lesquels des paquets assez courts, véhiculant entre 100 et 500 bits sont assez fréquemment rencontrés. Enfin, des avancées récentes dans le domaine des turbo codes laissent entrevoir l'utilisation de solutions fonctionnant à un rapport  $E_b/N_0$  transmis proche de 0 dB.

Ces solutions techniques imposent des contraintes de performances très sévères aux algorithmes d'estimation de phase et de fréquence qui constituent dans les modems l'étage précédant le décodeur.

Estimer un décalage de phase et de fréquence sur un paquet de symboles de communication peut être assez simple, si le nombre de symboles de référence est suffisant. Cependant, pour limiter la perte d'insertion générée, il est légitime chercher à réduire leur nombre. Une classification sépare ainsi les algorithmes de récupération de phase/fréquence qui utilisent de tels symboles, dits de type Data Aided (DA) des algorithmes qui n'en utilisent pas, dits Non Data Aided (NDA).

L'article rappelle d'abord les performances atteintes par les algorithmes classiques dans ces différents cas et les compare à la borne théorique de Cramer Rao [1]. Puis un nouvel algorithme d'estimation, appelé Quantified Maximum Likelihood (QML), est présenté. Ses performances sont présentées et une analyse théorique de la position des symboles de référence, en fonction des informations a priori sur l'écart de fréquence recherché, est menée.

## 2. Modèle de signal

L'article utilise les conventions suivantes : la modulation considérée est une MPSK, le bruit du canal est supposé blanc, additif et gaussien. Le signal à synchroniser en fréquence est un paquet de  $N$  symboles. Soient  $e_k$  et  $s_k$  les  $k^{\text{ièmes}}$  symboles émis et reçu,  $T_s$  le temps symbole,  $f_d$  et  $\varphi_0$  les erreurs de fréquence et de phase et  $n_k$  le bruit sur le  $k^{\text{ième}}$  symbole du paquet. Le signal reçu est modélisé de la manière suivante :

$$\begin{aligned} s_k &= e_k e^{j(2\pi f_d k T_s + \varphi_0)} + n_k \\ &= e^{j\left(\frac{\pi(2l+1)}{M}\right)} e^{j(2\pi f_d k T_s + \varphi_0)} + n_k \quad (\text{lentier}) \end{aligned} \quad (1.)$$

En DA, les symboles connus en réception sont corrigés de la valeur émise, et en NDA, la modulation est supprimée en multipliant par  $M$  la phase et en appliquant une non linéarité au module du signal reçu :

$$\begin{aligned} s_k^{DA} &= s_k e_k^* = e^{j(2\pi f_d k T_s + \varphi_0)} + n_k e_k^* \\ s_k^{NDA} &= \rho(s_k) e^{j(M \text{Arg}(s_k))} \end{aligned} \quad (2.)$$

## 3. Estimation de fréquence en MPSK : algorithme Rife et Boorstyn

Rife et Boorstyn ont montré dans [2] qu'estimer la fréquence  $f_d$  au sens du maximum de vraisemblance en DA revient à trouver la fréquence entre 0 et  $\frac{1}{T_s}$  qui maximise le module de  $Z(f)$ .

$$Z(f) = \frac{1}{N} \sum_{k=0}^{k=N-1} s_k^{DA} e^{-j2\pi k f T_s} \quad (3.)$$

La phase estimée est alors donnée par :

$$\hat{\phi}_0 = \text{Arg} \left\{ \max_f (|Z(f)|) \right\}$$

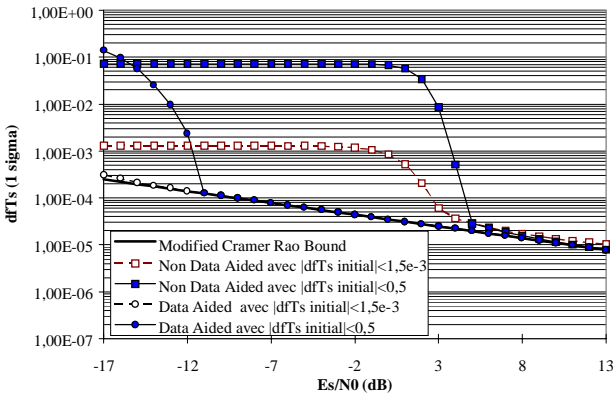
Les algorithmes classiques NDA nécessitent une incertitude initiale de fréquence inférieure à  $\frac{1}{4T_s}$ . Ils suppriment la

modulation et appliquent une non linéarité [3] au module comme indiqué dans l'équation 2.

Dans les cas DA et NDA, le signal temporel est complété par  $2^P - N$  zéros. Une FFT est réalisée, dont la résolution peut être affinée par interpolation parabolique sur le module de la FFT autour de son maximum. Les résultats de cette estimation peuvent être comparés avec la version modifiée de la borne de Cramer-Rao dont l'expression est donnée dans [1] :

$$\sigma_{f, MCRB}^2 = \frac{3T_s}{2\pi^2 (NT_s)^3 \frac{E_s}{N_0}} \quad (4.)$$

Les performances de l'algorithme R&B sont représentées sur la Figure 1 dans le cas DA pour des paquets de 500 symboles de référence et dans le cas NDA pour des paquets de 500 symboles d'information pour différentes incertitudes initiales de fréquence. Si cette incertitude est limitée, du fait de la précision des oscillateurs locaux et de la connaissance des dopplers, il n'est plus utile de tester toutes les fréquences ce qui réduit la complexité et améliore les performances de l'estimation.



**Figure 1 - Performances de l'algorithme R&B sur un signal QPSK de 500 symboles, abscisse : rapport  $E_s/N_0$ , ordonnée : écart-type de l'erreur résiduelle de fréquence multipliée par le temps symbole -**

## 4. Algorithme QML : Quantified Maximum Likelihood

Les résultats de la partie 3 montrent des performances dégradées d'environ 16 dB dans le cas NDA par rapport au cas DA car la suppression de la modulation augmente la variance du bruit à faible  $E_b/N_0$ , et le rend non gaussien.

L'algorithme QML consiste à tester des hypothèses de phase et des hypothèses de fréquences dans la plage d'incertitude de fréquence initiale sans lever la modulation.

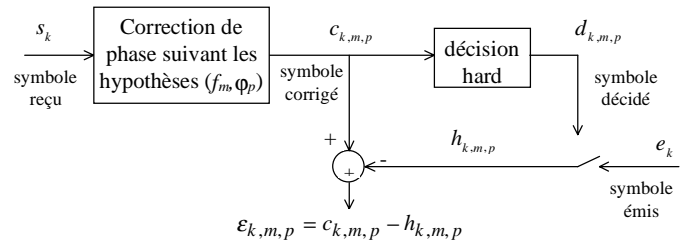
### 4.1 Principe général de l'algorithme QML

Pour chaque couple fréquence-phase testé  $(f_m, \phi_p)$ , les symboles  $s_k$  du paquet sont corrigés en phase, et sont notés :

$$c_{k,m,p} = s_k e^{-j(2\pi f_m k T_s + \phi_p)}$$

Une décision "hard"  $d_{k,m,p}$  est prise pour chaque symbole d'information. Soit  $h_{k,m,p}$  le symbole égal à  $e_k$  si le  $k^{\text{ième}}$  symbole est connu, ou égal à  $d_{k,m,p}$  dans le cas contraire, et soit  $\epsilon_{k,m,p}$  le signal d'erreur calculé entre  $c_{k,m,p}$  et  $h_{k,m,p}$  comme indiqué sur la Figure 2.

Le couple fréquence-phase qui minimise la fonction de coût obtenue en sommant tous les signaux d'erreurs quadratiques du paquet est choisi comme étant le plus probable. L'estimation peut être affinée par interpolation autour du couple le plus probable.



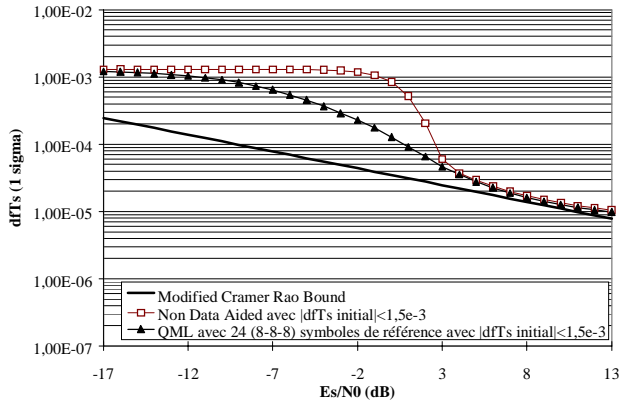
**Figure 2 - Calcul d'un signal d'erreur -**

### 4.2 Fonction de corrélation

On peut montrer que minimiser la fonction de coût équivaut à maximiser la fonction  $\text{Corrélation}(f)$  :

$$\begin{aligned} \text{Corrélation}(f_m) &= F(f_m, \phi(f_m)) \\ &= \text{Max}_{\phi_p} \left( \sum_{k=1}^N \text{Re} \{ c_{k,m,p} h_{k,m,p}^* \} \right) \end{aligned} \quad (5.)$$

Cette fonction présente des lobes secondaires qui dépendent du nombre de symboles de référence et de leur position. L'article explique comment choisir la position de ces symboles pour rejeter les ambiguïtés de fréquence hors de la plage de recherche et ainsi améliorer les performances de l'estimation.



**Figure 3 - Performances du QML pour un choix adapté de la position des symboles de référence, abscisse : rapport Es/N0, ordonnée : écart-type de l'erreur résiduelle de fréquence multipliée par le temps symbole -**

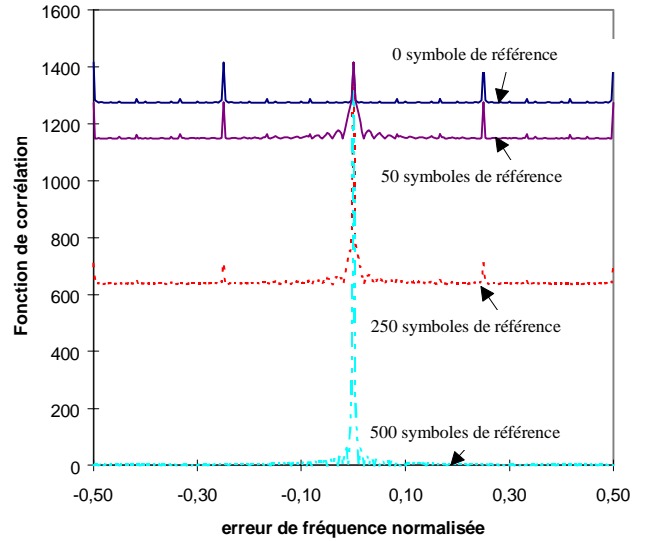
### 4.3 Influence du nombre de symboles de référence

Sur la Figure 4 ont été représentées les fonctions de corrélation correspondant à des paquets de 500 symboles QPSK dont les symboles de référence ont été placés en début de paquet.

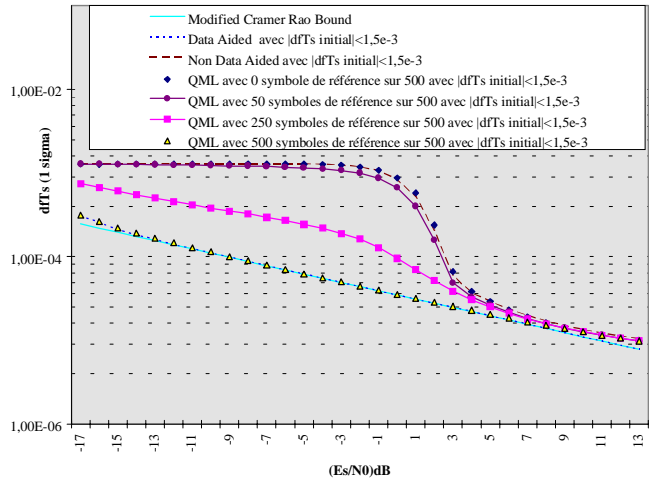
Dans le cas où il n'y a que des symboles de référence, la fonction de corrélation en l'absence de bruit est égale au module de la Transformée de Fourier du paquet ( cf courbe inférieure Figure 4). En effet :

$$\begin{aligned}
 \text{Corrélation}(f_m) &= \text{Max}_{\varphi_p} \left( \sum_k \text{Re} \left\{ c_{k,m,p} h_{k,m,p}^* \right\} \right) \\
 &= \text{Max}_{\varphi_p} \left( \text{Re} \left\{ \sum_k \left( s_k e^{-j(2\pi f_m k T_s + \varphi_p)} e_k^* \right) \right\} \right) \\
 &= \text{Max}_{\varphi_p} \left( \text{Re} \left\{ e^{-j\varphi_p} \sum_k \left( s_k e_k^* \right) e^{-j(2\pi f_m k T_s)} \right\} \right) \\
 &= \left| \sum_k \left( s_k e_k^* \right) e^{-j(2\pi f_m k T_s)} \right|
 \end{aligned} \quad (6.)$$

Dans le cas où il n'y a pas de symboles de référence, la fonction de corrélation a l'allure représentée sur la Figure 4 (cf courbe supérieure). Dans le cas où il y a des symboles de référence et des symboles connus, la fonction de coût est intermédiaire entre les deux fonctions précédentes, et dépend de la position des symboles de référence. Plus il y a de symboles de référence, et plus la bonne hypothèse de fréquence se distingue des mauvaises hypothèses. Les performances associées sont représentées sur la Figure 5, et sont effectivement meilleures quand le nombre de symboles de référence augmente.



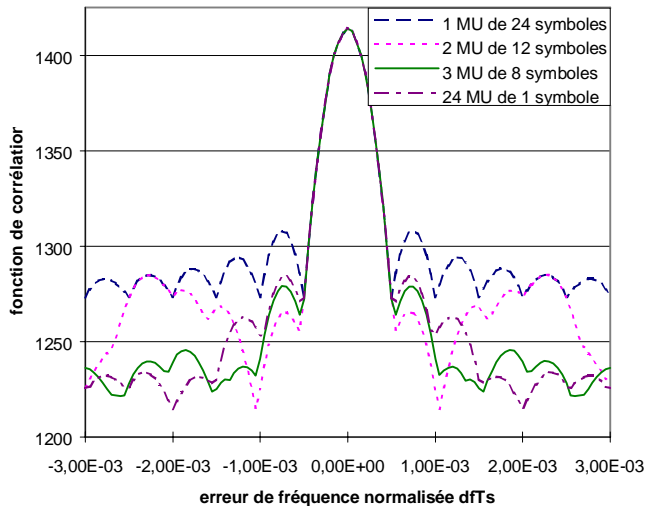
**Figure 4 - Fonction de corrélation calculée pour un paquet de 500 symboles QPSK en fonction du nombre de symboles de référence -**



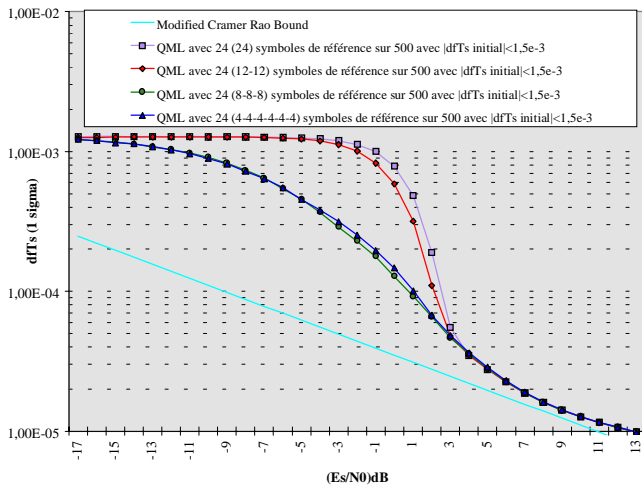
**Figure 5 - Performances du QML en fonction du nombre de symboles de référence -**

### 4.4 Influence de la position des symboles de référence

Sur la Figure 6, la fonction de corrélation a été calculée pour 24 symboles de référence sur 500 symboles QPSK. En fonction de la position des symboles de référence, l'allure de la fonction de coût pour de faibles erreurs de fréquence n'a pas la même allure et est plus avantageuse lorsque les symboles de référence sont répartis sur tout le paquet (au moins 3 mots uniques), en ce sens que la bonne hypothèse de fréquence se dégage mieux des mauvaises hypothèses. Les performances associées sont représentées sur la Figure 7.



**Figure 6 - Fonction de corrélation pour un paquet de 500 symboles de QPSK dont 24 symboles de référence répartis en un ou plusieurs mots uniques -**



**Figure 7 - Performances du QML en fonction de la position des symboles de référence -**

Des simulations non présentées ici ont montré que pour une incertitude initiale de fréquence donnée, des ambiguïtés de fréquence apparaissent pour certaines configurations des symboles de référence. En particulier, dans le cas où les symboles de référence sont séparés en deux mots uniques (début et fin de paquet), des ambiguïtés de fréquences apparaissent pour les valeurs  $df_{ambiguïté} \approx \frac{k}{NT_s} (k = 1, 2, \dots)$ .

Dans le cas où les symboles de référence sont séparés en trois mots uniques (début, milieu et fin de paquet), des ambiguïtés de fréquences apparaissent pour les valeurs

$$df_{ambiguïté} \approx \frac{2k}{NT_s} (k = 1, 2, \dots).$$

D'une manière générale, les ambiguïtés apparaissent pour les valeurs de fréquences telles que les symboles de référence semblent n'avoir pas tourné (ils font en fait un nombre entier de tours).

Par exemple, dans le cas où les symboles de référence sont répartis sur  $p$  mots uniques équirépartis sur le paquet (premier mot unique en début de paquet, dernier mot unique en fin de paquet), les ambiguïtés de fréquences apparaissent

$$df_{ambiguïté} \approx \frac{(p-1)k}{NT_s} k = 1, 2, \dots$$

Pour un nombre donné de symboles de référence, plus le nombre de mots uniques répartis sur le paquet est important et plus les ambiguïtés de fréquence se produisent à des fréquences loin de la valeur à estimer. Cependant, lorsque le nombre de mots uniques augmente, les performances de l'estimation de fréquence se dégradent. La Figure 7 illustre ce phénomène, puisque les performances du QML sont meilleures dans le cas de 3 mots uniques que dans les autres cas présentés.

## 5. Conclusion

L'article proposé aborde le problème de l'estimation de phase et de fréquence pour les transmissions radio utilisant des paquets de données courts à faible rapport signal sur bruit. Un nouvel algorithme est proposé et une analyse de performances est menée. Associé à un turbo code produit cet algorithme a permis de répondre à une spécification de modem TDMA extrêmement sévère.

## Références

- [1] A. N. D'Andrea, U. Mengali, R. Reggiannini : The modified Cramer-Rao bound and its applications to synchronization problems. « IEEE Trans. Commun. », Vol. COM-42, Feb./March/April 1994, P.1391-1399
- [2] D. C. Rife, R. R. Boorstyn : Single-tone parameter estimation from discrete-time observations. « IEEE Trans. Inform. Theory », Vol IT-20, Sept. 1974, p.591-598
- [3] A. J. Viterbi, A. M. Viterbi : Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission. « IEEE Trans. Inform. Theory », Vol. IT-29, July 1983, p543-550