



LA DETECTION SONAR ADAPTATIVE D'UN SIGNAL CONNU EN PRESENCE  
D'UN BRUIT DE TYPE MELANGE GAUSSIEN-GAUSSIEN  
PAR L'ALGORITHME "EXPECTATION MAXIMIZATION"

Mohamed SIALA et Ghassan KAWAS KALEH

Ecole Nationale Supérieure des Télécommunications,  
46, rue Barrault, 75 634 Paris cedex 13.

RÉSUMÉ

ABSTRACT

Nous présentons un détecteur sonar d'un signal connu en présence d'un bruit de type mélange gaussien-gaussien dont les paramètres (facteur de mélange et variances des composantes) sont inconnus. Le détecteur est adaptatif. Il est capable d'effectuer conjointement une estimation des paramètres et une détection. Des re-estimations itératives basées sur l'algorithme "*Expectation-Maximization*" convergent vers une estimée au sens du maximum de vraisemblance. Les résultats des simulations montrent la supériorité du détecteur proposé par rapport au détecteur robuste.

We present a sonar detector of known signal in the presence of a Gaussian-Gaussian mixture noise whose parameters (mixture factor and components variance) are unknown. The detector is adaptive. It is able to make joint parameters estimation and detection. Iterative estimates, based on the Expectation-Maximization algorithm, converge to a maximum likelihood estimate. Simulation results show the superiority of the proposed detector compared with the robust detector.

1. INTRODUCTION

Nous étudions ici la détection sonar en présence d'un bruit sous-marin de type mélange gaussien-gaussien (MGG). L'estimation des paramètres du bruit peut conduire à une amélioration significative des performances de la détection lorsque le rapport signal à bruit est faible. Ce concept a été étudié par Bouvet et Schwartz [2,3]. La conclusion à laquelle ils sont arrivés est qu'un détecteur adaptatif possède de meilleures performances qu'un détecteur utilisant un filtre adapté [3] et, pour que cette affirmation reste vraie, il est nécessaire que les paramètres de la loi de probabilité du bruit utilisés par le détecteur soient correctes. Cependant, la seule mesure préalable de ces paramètres n'est pas suffisante à cause de la nature non stationnaire du bruit et de sa dépendance du lieu et du temps. Une question importante qui

se pose alors est de savoir si une détection et une estimation des paramètres peuvent se faire conjointement.

Nous présentons ici une méthode adaptative de détection (d'un signal connu en présence d'un bruit de type MGG) et d'estimation (de paramètres) conjointes. Elle est basée sur la modélisation du signal reçu par une chaîne de Markov et l'utilisation de l'algorithme EM (*Expectation-Maximization*) [4, 5]. C'est un algorithme itératif d'estimation selon le critère du maximum de vraisemblance. Au lieu de maximiser directement la fonction de vraisemblance, on maximise une fonction intermédiaire qui est la fonction de mesure de l'information de Kullback-Leibler. La vraisemblance *a posteriori* des paramètres à estimer augmente à chaque itération. L'algorithme conduit à une estimation fiable des variances du bruit, de leur facteur de mélange ainsi que du



rapport de vraisemblance. Après un nombre fini d'itérations, une décision optimale sur la présence du signal maximisant la probabilité de détection pour une probabilité de fausse alarme donnée peut alors être obtenue.

L'article est organisé de la façon suivante. Le modèle du système est défini au § 2. Nous déduisons au §3 le détecteur recherché après une présentation sommaire de l'algorithme EM. Les résultats des simulations sont donnés au §4.

## 2. MODELISATION

On considère le système de sonar sous-marin qui transmet ou non un signal vectoriel  $s_1, s_2, \dots, s_T$  connu. Le récepteur qui dispose de l'observation vectorielle  $Y = (y_1, y_2, \dots, y_T)$  du signal reçu doit fournir une décision sur l'une des hypothèses suivantes [1]:

Hypothèse  $H_0$ :  $y_t = b_t, t = 1, 2, \dots, T$ ;

Hypothèse  $H_1$ :  $y_t = s_t + b_t, t = 1, 2, \dots, T$ .

où  $b_t$  est une réalisation d'un processus de bruit modélisé par un mélange gaussien-gaussien, de densité de probabilité conditionnelle donnée par

$$p(b|Z=z) = (1-z)g(b;\sigma_1) + z g(b;\sigma_2) \quad (1)$$

où  $Z$  est une variable aléatoire de Bernoulli prenant les valeurs 0 ou 1. La densité de probabilité  $g(b;\sigma)$  est celle d'une variable aléatoire gaussienne centrée et de variance  $\sigma^2$ . On soulage cette formulation quelque peu encombrante en posant:

Hypothèse  $H_\delta$ :  $y_t = \delta s_t + b_t, t = 1, 2, \dots, T$ , avec  $\delta = 0, 1$

Ce système peut être représenté par une chaîne de Markov de premier ordre à quatre états déterminés par les réalisations du couple  $(\delta, Z)$ . A chaque instant discret de temps, le processus est supposé être dans un des quatre états, et une observation est générée par la fonction aléatoire correspondante à l'état courant. Le diagramme de treillis de Figure 1 est une représentation utile.

La chaîne de Markov sous-jacente change d'état suivant une matrice de transition. L'observateur n'observe que la réalisation de la fonction aléatoire associée à chaque état et ne peut donc pas observer directement l'état de la chaîne de Markov sous-jacente. On associe à la chaîne de Markov un vecteur probabilité des états initiaux  $\pi = (\pi_0, \pi_1, \pi_2, \pi_3)$  et une matrice de transition entre états  $A[a_{ij}]$ ,  $0 \leq i, j \leq 3$ . Ici,  $\pi_i$  est la probabilité de l'état  $i$  au temps initial  $t = 1$  et  $a_{ij}$  est la probabilité de transition vers l'état  $j$ , conditionnellement à l'état courant  $i$ . On suppose que la transition entre les états

correspondant à  $\delta = 0$  (resp. 1) et les états correspondant à  $\delta = 1$  (resp. 0) est de probabilité nulle. Le processus aléatoire associé aux quatre états  $i$  peut être représenté par la famille  $f = \{f_0^t, f_1^t, f_2^t, f_3^t; t = 1, 2, \dots, T\}$ , où  $f_i^t(y)$ ,  $i = 0, 1, 2, 3$  est la probabilité d'observer au temps  $t$  la valeur  $y$  de l'échantillon du signal reçu étant donné l'état  $i$ . On a alors:

$$f_0^t(y) = g(y;\sigma_1), f_1^t(y) = g(y;\sigma_2)$$

$$f_2^t(y) = g(y - s_t;\sigma_1), f_3^t(y) = g(y - s_t;\sigma_2)$$

La chaîne de Markov peut alors être identifiée par l'ensemble des paramètres .

$$\theta \triangleq (\pi, A, \sigma_1, \sigma_2)$$

Etant donné l'observation  $Y$ , on se propose dans la suite de mettre en oeuvre l'algorithme EM pour l'estimation itérative selon le critère du maximum de vraisemblance des paramètres  $\theta$ .

## 3. ESTIMATION ET DETECTION

On note par  $p(Y; \theta)$  la densité de probabilité de l'observation  $Y$  indexé par le vecteur des paramètres  $\theta$ . On définit la mesure suivante:

$$L(Y, V=v; \theta) \triangleq P[V=v] p(Y; \theta / V=v) \quad (2)$$

où  $V$  est une variable aléatoire (scalaire ou vectorielle) de probabilité *a priori*  $P[V=v]$  et  $p(Y; \theta / .)$  est une densité de probabilité conditionnelle. Parfois, l'expression précédente sera simplifiée en  $L(Y, v; \theta)$ .

Soit  $X_t$  l'état à l'instant  $t$ . Il prend ses valeurs dans  $\{0, 1, 2, 3\}$ . Soit  $X = (x_1, x_2, \dots, x_{T+1})$  une séquence de  $T+1$  états où  $x_t$  est une réalisation de  $X_t$ .  $X$  prend ses valeurs dans l'ensemble des trajets possibles, c.à.d. dans un sous-ensemble de  $\{0, 1, 2, 3\}^{T+1}$ . L'estimation optimale au sens du maximum de vraisemblance de  $\theta$  est la valeur de  $\theta$  qui maximise la vraisemblance

$$L(Y; \theta) = p(Y; \theta) = \sum_X L(Y, X; \theta)$$

C'est un problème difficile. Cependant la solution peut être trouvée par l'algorithme itératif EM qui augmente la vraisemblance à posteriori de ses paramètres estimés à chaque itération. A cette fin on introduit la fonction de mesure de l'information de Kullback-Leibler [ 5]

$$Q(\theta, \theta^*) = \sum_X L(Y, X; \theta) \log L(Y, X; \theta^*)$$

L'algorithme commence par une estimation arbitraire  $\theta^{(0)}$  de  $\theta$ . La valeur estimée  $\theta^{(p)}$  obtenue à la p-ème itération est utilisée pour calculer une meilleure estimation  $\theta^{(p+1)}$  à la (p+1)-ème itération en faisant les deux étapes suivantes:

- l'étape d'estimation: calcul de  $Q(\theta^{(p)}, \theta)$ .
- l'étape de maximisation: recherche de  $\theta^{(p+1)}$  qui maximise  $Q(\theta^{(p)}, \theta)$  par rapport à  $\theta$ .

Sous certaines conditions, l'algorithme converge vers un estimateur de maximum de vraisemblance.

On montre que les estimations des paramètres de la chaîne de Markov sont données par les équations récurrentes suivantes:

$$\pi_i^{(p+1)} = \frac{L(Y, X_i=i; \theta^{(p)})}{L(Y; \theta^{(p)})}; i=0, 1, 2 \text{ et } 3 \quad (3)$$

$$a_{ij}^{(p+1)} = \frac{\sum_{t=1}^{T-1} L(Y, X_t=i, X_{t+1}=j; \theta^{(p)})}{\sum_{t=1}^{T-1} L(Y, X_t=i; \theta^{(p)})}; i, j=0, 1, 2 \text{ et } 3 \quad (4)$$

$$(\sigma_1^{(p+1)})^2 = \frac{\sum_{t=1}^T L(Y, X_t=0; \theta^{(p)}) y_t^2 + \sum_{t=1}^T L(Y, X_t=2; \theta^{(p)}) (y_t - s_t)^2}{\sum_{t=1}^T L(Y, X_t=0; \theta^{(p)}) + \sum_{t=1}^T L(Y, X_t=2; \theta^{(p)})} \quad (5)$$

$$(\sigma_2^{(p+1)})^2 = \quad (6)$$

$$\frac{\sum_{t=1}^T L(Y, X_t=1; \theta^{(p)}) y_t^2 + \sum_{t=1}^T L(Y, X_t=3; \theta^{(p)}) (y_t - s_t)^2}{\sum_{t=1}^T L(Y, X_t=1; \theta^{(p)}) + \sum_{t=1}^T L(Y, X_t=3; \theta^{(p)})}$$

où  $L(Y, X_i=i; \theta)$  est définie comme en (1).

Les quantités  $L(Y, X_i=i, X_{t+1}=j; \theta^{(p)})$  et  $L(Y, X_i=i; \theta^{(p)})$  qui apparaissent dans les formules précédentes se calculent de la façon suivante ( $\theta$  ne sera pas écrit):

$$L(Y, X_i=i, X_{t+1}=j) = \alpha_t(i) \beta_{t+1}(j) a_{ij} \quad (7)$$

$$L(Y, X_i=i) = \sum_{j=0}^3 L(Y, X_i=i, X_{t+1}=j) \quad (8)$$

où

$$\alpha_t(i) = P[X_t=i] p(y_1, y_2, \dots, y_t / X_t=i) \quad (9)$$

$$\beta_{t+1}(j) = p(y_{t+1}, y_{t+2}, \dots, y_T / X_{t+1}=j) \quad (10)$$

Les  $\alpha_t(i)$  et  $\beta_{t+1}(j)$  s'obtiennent par les récurrences:

$$\alpha_{t+1}(j) = [\sum_{i=0}^3 \alpha_t(i) a_{ij}] f_j^{t+1}(y_{t+1}) \quad (11)$$

$$\beta_t(i) = \sum_{j=0}^3 a_{ij}^{t+1}(y_{t+1}) \beta_{t+1}(j) \quad (12)$$

$$\alpha_i(i) = \pi_i f_i^t(y_i); \beta_i(i) = 1; i=0, 1, 2, 3 \quad (13)$$

L'algorithme se résume de la façon suivante:

- 1) Calculer  $\alpha_i(i)$  et  $\beta_i(i)$  en utilisant (11-13)
- 2) Calculer  $L(\cdot)$  en utilisant (7-8)
- 3) Re-estimer les paramètres en utilisant (3-6)
- 4) Si  $L(Y; \theta) <$  un seuil aller à 1, sinon fin.

Après l'estimation des paramètres  $\pi$ , A et f on procède à la détection par comparaison du rapport de vraisemblance

$$\Lambda(y_1^T) = \frac{L(Y, X_i=2; \theta) + L(Y, X_i=3; \theta)}{L(Y, X_i=0; \theta) + L(Y, X_i=1; \theta)}$$

à un seuil  $\Lambda_0$  fixé à l'avance. Si  $\Lambda(y_1^T) \geq \Lambda_0$ , l'Hypothèse  $H_1$  est retenue;  $H_0$  est retenue dans le cas contraire.

#### 4. RESULTATS DE LA SIMULATION

Les courbes COR (*Receiver Operating Characteristic*) du détecteur proposé, du détecteur utilisant un filtre adapté ainsi que le détecteur de [2,3], qui suppose une connaissance parfaite des paramètres, ont été estimées par une méthode de type Monte-Carlo en prenant 500 blocs de bruit et une longueur de bloc  $T=100$  échantillons. Les paramètres choisis sont:  $P[Z=1] = 0.25$ ,  $\sigma_1 = 0.65760$  et  $\sigma_2 = 1.64399$ , voir (1).

Les résultats sont présentés dans Figures 2: a, b et c qui correspondent respectivement à un rapport signal-sur-bruit de -10, -15 et -20 dB. Les estimations de  $P[Z=1]$ ,  $\sigma_1$  et  $\sigma_2$  prises au démarrage de l'algorithme sont telles que

$$\sigma_2^{(0)} / \sigma_1^{(0)} = 5.0, P[Z=1]^{(0)} = 0.5 \text{ et } (\sigma_1^{(0)})^2 + (\sigma_2^{(0)})^2 = 1.$$

Quatre courbes COR, obtenues respectivement pour le détecteur utilisant un filtre adapté, celui de [2,3], la première et la dixième itération de l'algorithme proposé, sont comparées dans chacune de ces trois figures. Elles montrent qu'après un certain nombre d'itération, la performance dépasse celle du récepteur à filtre adapté et qu'elle est proche du récepteur qui dispose d'une connaissance parfaite des paramètres.

#### REFERENCES

- [1] H.L. Van Trees, "Detection, estimation, and modulation theory", Tome I, J. Wiley & Sons, p.19-46, p.246-271.
- [2] M. Bouvet et S.C. Schwartz, "Sur l'intérêt de l'adaptativité par rapport à la robustesse en détection sonar". GRETSI 1989, pp. 111-114.
- [3] M. Bouvet et S.C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise", IEEE Trans. on ASSP, vol. 37, n° 5, May 1989, p.621-626.



- [4] A.P. Dempster, N.M. Laird et D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", J. Roy. Stat. Soc., Ser. 39, 1977, p.1-38.
- [5] L.E. Baum, T. Petrie, G. Soules et N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains", Ann. Math. Stat., vol.41, n° 1, 1970, p.164-171.

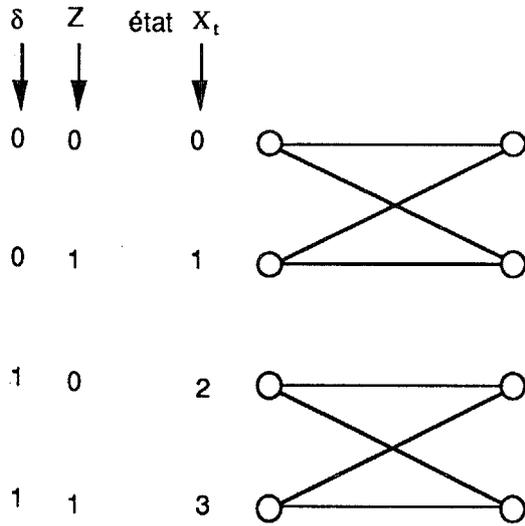
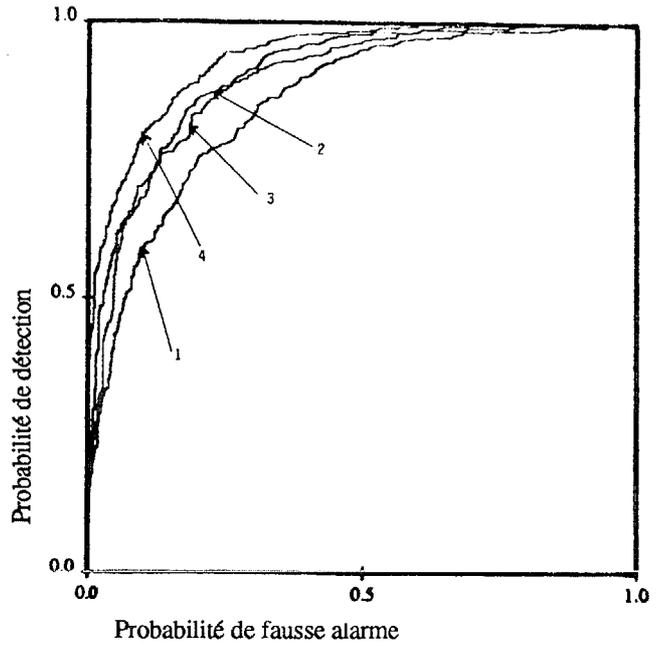
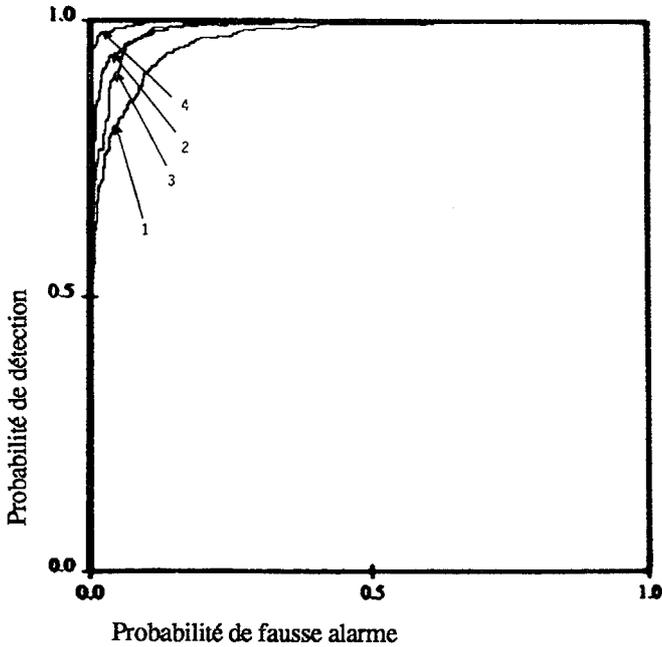


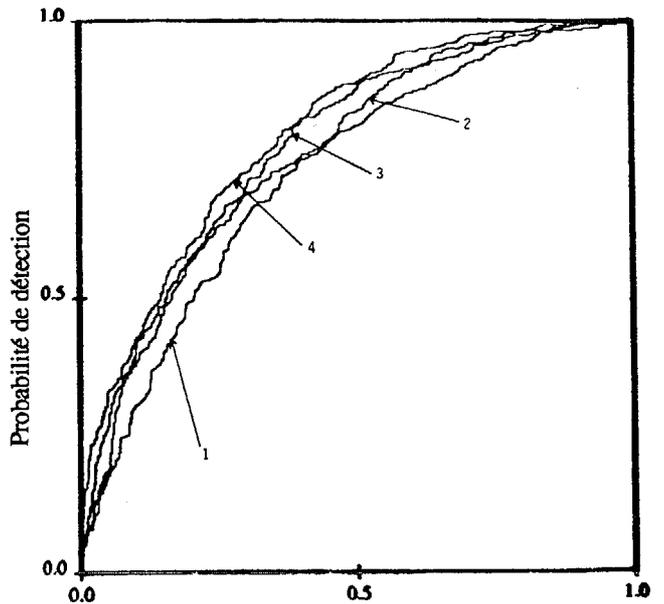
Figure 1- Le treillis.



(b) Rapport signal sur bruit= -15 dB



(a) Rapport signal sur bruit= -10 dB



(c) Rapport signal sur bruit= -20 dB

Figure 2 - Performance du détecteur proposé (Courbe 1: après une itération, Courbe 2: après 10 itérations) en comparaison avec celle du filtre adapté (Courbe 3) et celle du détecteur de [3,4] qui suppose une connaissance parfaite des paramètres (Courbe 4).