

ELEMENTS DE COMPARAISON ENTRE DIFFERENTES METHODES DE SOUSTRACTION DE BRUIT

Christine SERVIERE(*) - Daniel BAUDOIS (*)
 André SILVENT (**) - Pierre LORENZINO (*)

(*) CEPHAG - Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble
 B.P. 46 - 38402 SAINT MARTIN D'HERES (FRANCE)

(**) Attaché aux services techniques des Armées

Nous envisageons différentes méthodes de soustraction de bruit perturbant un signal, lorsque nous disposons d'une "référence-bruit seul". Ces différentes approches sont comparées en tenant compte de tous les paramètres intervenant : nombre d'échantillons utilisés, temps de calcul L'accent a été mis sur une approche utilisant un algorithme des moindres carrés fréquentiels. Des résultats théoriques concernant la caractérisation de cette méthode ont été établis.

This note concerns different methods of noise cancelling when we have a noise reference. We compare the algorithms with respect of all parameters : number of datas which are used, computational requirements.

We particularly insist on the frequency domain LMS algorithm. Theoretical results concerning this method are presented.

I. INTRODUCTION

Un des problèmes majeurs du traiteur de signaux est l'estimation d'un signal, lorsqu'il est additivement perturbé. Différentes approches de ce problème ont été envisagées par divers auteurs. Les hypothèses nécessaires à la mise en place de ces méthodes sont fortes : la technique développée par WIENER (I) exige la connaissance des moments du second ordre du signal, du bruit et du signal bruité, l'approche de KALMAN-BUCY (II) la modélisation du signal et du bruit à l'aide d'une représentation d'état...

Ces méthodes, devenues classiques ont en commun :

- le choix du critère de minimisation du risque quadratique qui n'est autre qu'une optimisation au sens de la norme classique [critère E.Q.M.],

- l'utilisation d'un estimateur linéaire.

Dans la suite, nous nous intéressons particulièrement au cas où, en plus de l'observation bruitée $z(t) = s(t) + b(t)$, l'expérimentateur a accès à un signal $x(t)$, corrélé avec le bruit perturbateur $b(t)$ et indépendant du signal utile $s(t)$. [Pour cette raison, il est convenu d'appeler $x(t)$ "Référence-bruit seul" - en abrégé R.B.S.]

La connaissance de cette référence permet dans une première étape, l'estimation du bruit $b(t)$, puis par soustraction à l'observation $z(t)$, l'estimation du signal $s(t)$ [fig. 1].

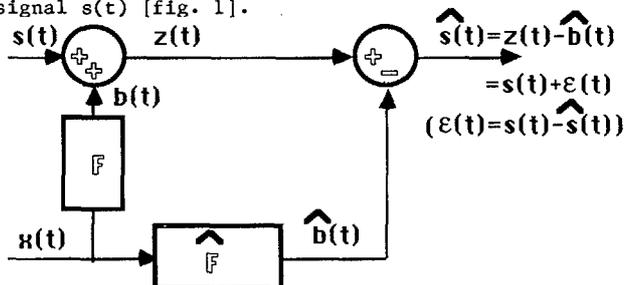


Fig. 1 - Schéma de principe du soustracteur de bruit

Il s'agit en réalité d'identifier le filtre F qui engendre $b(t)$ à partir de $x(t)$ (dans la bande de $x(t)$), puis de filtrer $x(t)$ à l'aide d'un filtre \hat{F} identique à F . Cette méthode a été développée par WIDROW [III] qui a défini un algorithme récursif sur lequel nous reviendrons plus loin. On peut classer les méthodes de soustraction de bruit en deux catégories : les méthodes directes et les méthodes bouclées. Cette communication a pour but de rappeler rapidement les particularités de chaque approche et de définir des critères objectifs permettant de les comparer.

II. LES DIVERSES METHODES DE SOUSTRACTION DE BRUIT

II.1. Méthodes non bouclées

Lorsque $x(t)$ est un signal aléatoire à spectre continu, l'approche la plus classique consiste à décomposer le filtre \hat{F} en deux filtres en cascade. Cette méthodologie décrite par H. MERMOZ [IV] a été largement étudiée au CEPHAG [V, VI, VII]. Il s'agit d'abord de blanchir la "référence-bruit-seul". Cette référence blanchie est alors filtrée par un filtre dont la réponse impulsionnelle est égale à l'intercorrélacion entre le signal bruité et cette référence blanchie. La structure complète a été appelée corrélofiltre.

Lorsque $x(t)$ se compose de raies, la méthode se transpose aisément [VIII]. Le filtre banchisseur est alors remplacé par un système égaliseur [IX,X].

Cet estimateur peut être construit en travaillant dans la représentation fréquentielle. Le filtre \hat{F} est alors caractérisé par son gain complexe qui est estimé en calculant le rapport de la densité spectrale de puissance croisée entre le signal bruité et la référence bruit à la densité spectrale de puissance moyenne de la référence. Cette division réalise le blanchissement ; l'ensemble est appelé spectrofiltre [XI].

II.2. Algorithme L.M.S. temps

La réduction de bruit peut être obtenue à l'aide d'algorithmes récursifs. L'utilisation de l'algorithme du Gradient par WIDROW [III, XII] conduit l'auteur à



définir un système sous-optimal en substituant au gradient réel de la puissance de l'erreur résiduelle un gradient estimé. Cette méthodologie, parce que très simple à appliquer, a connu un vif succès. Elle nous servira de structure-test pour notre comparaison, même si des structures plus performantes - moins carrés récurrents - Kalman rapide... [XIII] sont maintenant plus utilisées. L'étude de la convergence de l'algorithme de WIDROW a fait l'objet de nombreux papiers [XIV, XV].

II.3. Algorithme L.M.S. fréquence

Il est possible de définir un algorithme ayant pour objectif d'estimer le gain complexe du filtre \hat{F} de manière récursive. La démarche est analogue à celle proposée par B. PICINBONO dans [XIV].

Le filtre \hat{F} est caractérisé par le vecteur \hat{W} , constitué de $N+1$ composantes $W(0), W(1), \dots, W(N)$ représentant $N+1$ valeurs de son gain complexe $W(v)$.

$$\text{Ayant } \begin{cases} \hat{B}(v) = W(v) \cdot \hat{X}(v) \\ \hat{\varepsilon}(v) = B(v) - \hat{B}(v), \end{cases} \quad v = 0, \dots, N$$

sous réserve d'indépendance de $S(v)$ et de $X(v)$, la minimisation du risque quadratique

$$E \left\{ \left| \hat{\varepsilon}(v) \right|^2 \right\} \quad v = 0, \dots, N, \text{ donc de } E \left\{ \left| S(v) \right|^2 \right\}$$

entraîne

$$(1) \quad W(v) = \frac{E \{ Z(v) \cdot X^*(v) \}}{E \{ X(v) \cdot X^*(v) \}}, \quad v = 0, \dots, N$$

qui est l'expression classique du gain du filtre optimal obtenue par la méthode dite de Wiener-Hopf.

Il est alors possible d'obtenir une formule récursive adaptative à partir de la formule (1).

Les signaux $z(t)$ et $x(t)$ étant découpés en blocs adjacents composés chacun de N échantillons temporels, l'estimation des $N+1$ coefficients $W(v)$ après traitement des k premiers blocs de $z(t)$ et $x(t)$ peut être obtenue par application de la formule (2) :

$$(2) \quad W_{k+1}(v) = \frac{\sum_{j=0}^k Z(j) \cdot X^*(j)}{\sum_{j=0}^k X(j) \cdot X^*(j)} \quad v = 0, \dots, N$$

en remplaçant les espérances mathématiques par des moyennes arithmétiques.

Les coefficients $W_{k+1}(v)$ sont utilisés pour estimer les valeurs de \hat{B} correspondant au bloc de rang $k+1$ de X .

Un développement analogue à celui présenté en [XIV] permet d'obtenir un algorithme récursif :

$$(3) \quad W_{k+1}(v) = W_k(v) + M_k(v) \cdot X_k^*(v) \cdot \hat{S}_k(v) \quad v=0, \dots, N$$

$$M_k(v) = \frac{1}{\sum_{j=0}^k X_j(v) \cdot X_j^*(v)}$$

$X_j(v)$ et $\hat{S}_j(v)$ représentant les transformées de Fourier des échantillons des blocs temporels de rang j de $x(t)$ et de $\hat{s}(t)$.

L'expression (3) peut être simplifiée en évitant le calcul de $M_k(v)$ à chaque itération, soit en introduisant une suite a priori, soit en remplaçant $M_k(v)$ par un terme $\mu(v)$ qui ne dépend plus de k . On obtient ainsi l'algorithme L.M.S. fréquence [XVI-XVIII].

Certains auteurs ont fait état d'algorithmes utilisant des formules de renouvellement voisines : C. OGUE & al [XVIII] proposent :

$$M_k(v) = \frac{-1}{[M_0(v) + \sum_{l=1}^k X_l^*(v) X_l(v)]^{-1}}$$

Ils indiquent des risques d'instabilité due au choix de la condition initiale et proposent d'introduire une mémoire exponentielle pour le calcul de

$$\sum_{l=1}^k X_l^*(v) X_l(v).$$

En Juin 1986, BERSHAD & al [XIX] proposent un algorithme voisin. La formule itérative est empirique. Elle assure une bonne stabilité de l'algorithme. Le coefficient $M_k(v)$ est inversement proportionnel à l'estimée de la puissance de la référence bruit.

A condition de supposer $X_k(v)$ et $X_{k-1}(v)$ décorrélés, on peut montrer que l'algorithme

$$W_{k+1}(v) = W_k(v) + \mu(v) X_k^*(v) \hat{S}_k(v)$$

converge en valeur moyenne vers la solution optimale de Wiener si

$$(4) \quad \mu(v) < \frac{2}{E \{ X_k(v) X_k^*(v) \}}$$

Cette borne supérieure, calculée pour chaque canal de fréquence traité est en fait indépendante du rang de la tranche traitée lorsque $x(t)$ est stationnaire au second ordre.

En effet :

$$X_k(v) = \int_{kT}^{(k+1)T} x(t) e^{-2\pi jvt} dt \quad (T=(N-1)T_{ech})$$

et

$$E \{ X_k(v) X_k^*(v) \} = \int_{kT}^{(k+1)T} \int_{kT}^{(k+1)T} E \{ x(t_1) x(t_2) \} e^{-2\pi jv(t_1-t_2)} dt_1 dt_2$$

$$= \int_{-\frac{T}{2}}^{+\frac{T}{2}} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \Gamma_{xx}(u_1-u_2) e^{-2\pi jv(u_1-u_2)} du_1 du_2$$

$$(5) \quad = \int_{-T}^{+T} [T - |u|] \Gamma_{xx}(u) e^{-2\pi jvu} du,$$

expression dans laquelle $\Gamma_{xx}(u)$ est la fonction d'autocorrélation de la référence bruit $x(t)$.

On remarque en particulier que si $x(t)$ est blanc (ie : $\Gamma_{xx}(u) = N_0 \delta(u)$), $\mu(v) < \frac{2}{N_0 T}$;

ce qui signifie que cette borne est indépendante du canal de fréquence étudié, mais varie avec la taille des blocs traités.

III COMPARAISON DES QUATRE ALGORITHMES DE BASE

III.1. Mesure des paramètres caractéristiques

Il est classique de caractériser le comportement des systèmes adaptatifs par la courbe d'apprentissage qui indique l'évolution de la puissance de l'erreur d'estimation en fonction du temps. Dans le cas d'études en simulation, nous pouvons avoir accès à cette quantité $P_{\varepsilon}(t) = [s(t) - \hat{s}(t)]^2$

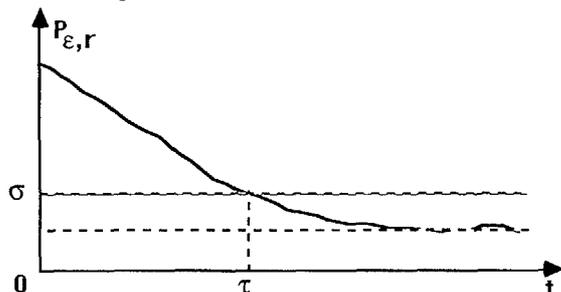


Fig. 2

Cette courbe, dont l'allure générale est indiquée fig.2, possède une valeur initiale $P_{\varepsilon}(0)$ qui dépend des conditions d'expérimentation. Elle converge en un temps τ vers une valeur asymptotique $P_{\varepsilon,r}$; le couple de valeurs $[P_{\varepsilon,r}, \tau]$ mesurées dans une situation

donnée nous a semblé convenablement approprié pour définir l'efficacité d'une méthode de réduction de bruit face à cette situation.

Pour mesurer correctement l'erreur résiduelle $P_{\epsilon,r}$, la durée de chaque expérience doit être suffisamment longue. Ainsi le palier de convergence de la courbe est atteint, puis maintenu sur un nombre d'échantillons assez grand pour permettre le calcul d'une moyenne significative. Dans les exemples présentés, nous avons vérifié que la pente de la droite moyenne couvrant les 2000 derniers points est voisine de zéro. Lorsque cette condition est remplie, nous choisissons de prendre pour valeur de $P_{\epsilon,r}$ la moyenne des 2000 derniers points de la courbe d'apprentissage. Dans le cas contraire, l'expérience n'est pas prise en compte.

Le temps de convergence τ est le temps au bout duquel la courbe d'apprentissage est très proche de son asymptote.

Après avoir convenablement lissé cette courbe [XX], nous avons déterminé la valeur de τ en tant qu'abscisse de l'intersection de la courbe lissée et de la droite d'ordonnée

$$\tau = P_{\epsilon,r} + \frac{P_{\epsilon}(0) - P_{\epsilon,r}}{10} \quad [\text{fig.2}].$$

Chaque algorithme sera caractérisé par un réseau de courbes tracé dans le plan $[P_{\epsilon,r}, \tau]$ pour différentes valeurs du paramètre $N+1$ (nombre de coefficients du filtre F).

III.2. Précautions concernant les algorithmes fréquentiels

Lorsqu'on examine la courbe d'apprentissage d'une méthode fréquentielle (par exemple - spectrofiltre, fig.3),

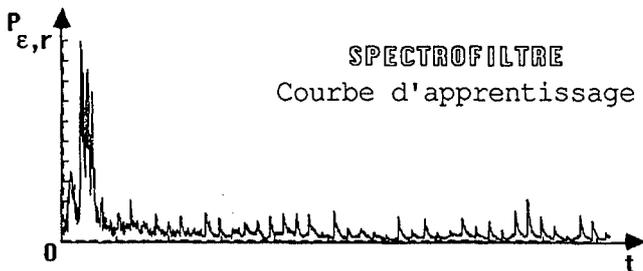


Fig.3

On se rend compte de la présence de transitoires apparaissant à chaque début de bloc traité. En effet le traitement du bloc de rang k est obtenu par application de la formule du filtrage fréquentiel

$$(6) \quad \hat{B}_k(v) = W_k(v) \cdot X_k(v) \quad v = 0, \dots, N$$

Le calcul de $X_k(v)$, transformée de Fourier des échantillons du bloc de rang k de $x(t)$ est obtenu sans tenir compte des échantillons des blocs des rangs précédents $k-1, \dots$

En conséquence, apparaît lors de l'application de la formule (6) un transitoire dont la longueur est égale au support de la réponse impulsionnelle du filtre F .

Deux méthodes ont été mises en oeuvre pour éliminer ce transitoire.

La première consiste à ne retenir que les estimations significatives (donc situées hors du transitoire) en travaillant sur des blocs se recouvrant partiellement. Deux blocs de rangs successifs ont en commun $N-L$ échantillons. A la suite du traitement d'un bloc, ne seront conservées que les estimations correspondant aux L derniers échantillons temporels de ce bloc. Si $N-L$ est suffisamment grand (en tout cas, plus grand que le support de la réponse du filtre F) les estimations retenues seront hors du transitoire, donc correctes. [XXI].

La seconde technique, dite par addition de blocs consiste à calculer par (6) $2N$ échantillons temporels à partir de N échantillons d'entrée (complétés par

N zéros). Ainsi les transitoires seront éliminés deux à deux par addition terme à terme des blocs contenant les échantillons temporels de $B(t)$. [XXI]

Les courbes de la figure 4 ont été établies pour le spectrofiltre. $x(t)$ est un bruit blanc normé et le filtre F est du premier ordre. On constate que les résultats obtenus avec la première méthode sont meilleurs. Ceci provient du fait que le nombre d'accumulations équivalent est plus élevé que dans la seconde méthode (mais demande un plus grand nombre d'opérations).

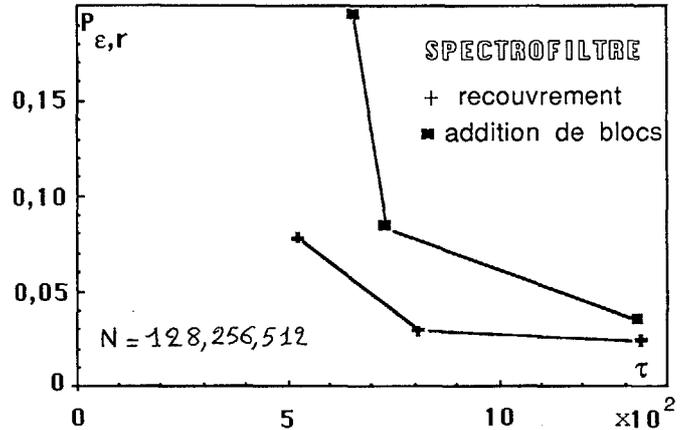


Fig. 4

On a constaté que pour le spectrofiltre, les résultats s'améliorent lorsque le recouvrement $(N-L)$ augmente. Au delà d'une certaine valeur, le gain en performances stagne alors que le coût de calcul croît rapidement.

Dans le cas du LMS-fréquence, l'augmentation de $N-L$ entraîne une augmentation de $P_{\epsilon,r}$, voire une divergence de l'algorithme. Ceci est dû au fait que les conditions de convergence de l'algorithme ne sont plus remplies (cf.II.3).

La figure 5 représente dans le cas d'une RBS blanche et normée, pour diverses valeurs de μ indépendantes de la fréquence, l'évolution de $P_{\epsilon,r}$ en fonction de $(N-L)$; les zones de divergence apparaissent (μ trop grand, $N-L$ trop grand pour un μ donné). Ce réseau a été tracé pour des tailles de blocs de 128 points.

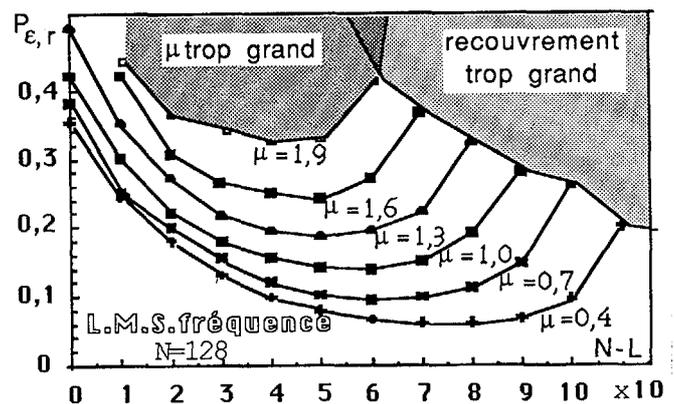


Fig. 5

Dans le paragraphe précédent, nous avons indiqué la technique permettant d'éliminer le transitoire lié à chaque bloc traité. Ce transitoire apparaît également dans le terme correcteur $\hat{S}(v)$. Nous avons donc corrigé $\hat{S}(t)$ en ne retenant que les échantillons correctement estimés, puis calculé $\hat{S}(v)$ par transformation de Fourier. Dans ces conditions, les résultats sont devenus comparables à ceux obtenus par les autres algorithmes. Les courbes de la figure 6 représentent les résultats de différentes versions de l'algorithme LMS-fréquence.

L'utilisation de l'algorithme (3) peut conduire à des instabilités, lorsque le dénominateur de $M_k(v)$ est proche de zéro. Ces instabilités diminuent en introduisant une moyenne exponentielle :



$$(7) M_k(v) = \frac{1}{\sum_{j=0}^k \lambda^{k-j} x_j(v) x_j^*(v)}$$

Enfin, l'introduction de la cohérence $C(v)$ entre $z(t)$ et $x(t)$ donne de meilleurs résultats, mais au prix de calculs plus importants.

$$(8) W_{k+1}(v) = W_k(v) + \mu C(v) \hat{S}(v) X^*(v)$$

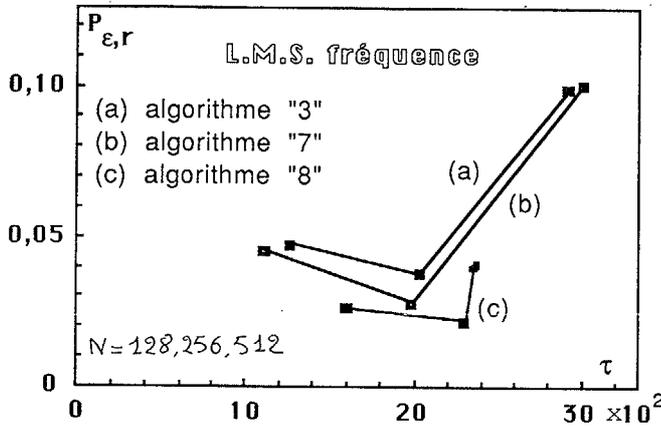


Fig. 6

La figure 7 représente pour trois tailles de blocs (128, 256 et 512) les performances des 4 algorithmes.

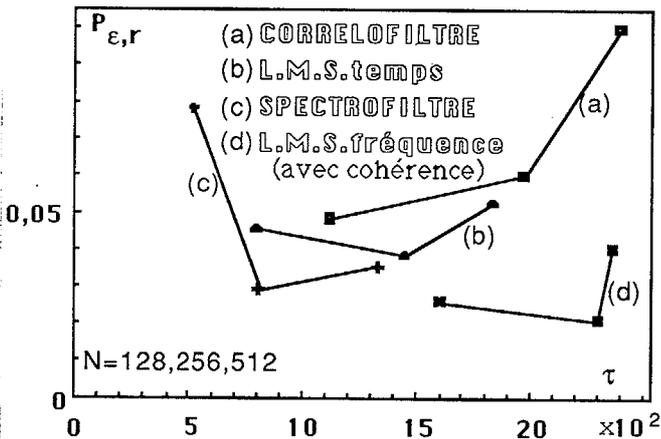


Fig. 7

On constate que par le choix des paramètres on peut rendre proches ces performances.

CÔÛT DE CALCUL	
LMS temps	$4N + 2$
Corrélofiltre	$4N + 2$
Spectrofiltre sans recouvrement	$\frac{3}{2} \log_2 N + 7$
avec recouvrement	$\left[\frac{3}{2} \log_2 N + 7 \right] \frac{N}{L}$
LMS fréquence μ constant	$\frac{N}{L} \left[\frac{3}{2} \log_2 N + 4 \right]$
théorique	$\frac{N}{L} \left[\frac{3}{2} \log_2 N + 6 \right]$
avec la cohérence	$\frac{N}{L} \left[\frac{3}{2} \log_2 N + 12 \right]$

Le tableau ci-dessus indique pour chacun des algorithmes étudiés le nombre d'opérations élémentaires nécessaires pour obtenir un échantillon de sortie.

En pratique, il est clair que chaque utilisateur sera amené à faire un choix en fonction des performances qu'il souhaite obtenir mais également en fonction du volume de calcul (choix du calculateur utilisé, rapidité ...).

- [I] N.WIENER : Extrapolation, interpolation and smoothing of stationary time series, with engineering applications. New York, Wiley, 1949.
- [II] R.KALMAN : A new approach to linear filtering and prediction problems. J. Basic eng. Vol.2, mars 1962.
- [III] B.WIDROW & al. : Adaptive noise cancelling. Principles and applications. Proc IEEE, Vol 63, Dec. 1975.
- [IV] H.MERMOZ : "Modularité" du traitement adaptatif d'antenne. Ann. des Télécomm. 1974, 29, N° 1-2 .
- [V] D.BAUDOIS - W.KOFMAN - A. SILVENT : Etude théorique et expérimentale du corrélofiltre - 8ème colloque GRETSI 1981.
- [VI] G.BOUTHEMY - W.KOFMAN - A.SILVENT : Theoretical and computer simulation study of the "corrélofiltre" adaptive system for non stationary processes - ICASSP 1982.
- [VII] G.BOUTHEMY : Etude d'un corrélofiltre autoadaptatif. Thèse de docteur ingénieur - Grenoble Novembre 1982.
- [VIII] D.BAUDOIS - A.SILVENT : Application du corrélofiltre à l'élimination de raies en présence de bruit stationnaire - 7ème colloque GRETSI 1979.
- [IX] D.BAUDOIS - E.MOISAN - A.SILVENT - J.C.BALLUET : Etude d'un égaliseur de raies - 9ème colloque GRETSI 1983
- [X] P.JOUVE : Etude de classes de filtres isolant une sinusoïde d'un signal composé de fréquences pures bruitées. Application à l'égalisation de raies. Thèse de 3ème cycle, Grenoble 1985.
- [XI] P.COMON : Traitement de signaux magnétiques multivariés. Thèse de doctorat de l'INPG, Décembre 1985.
- [XII] B.WIDROW & al. : Stationary and non stationary learning characteristics of the LMS adaptive filter. Proc. IEEE, vol. 64. Août 1976.
- [XIII] M.BELLANGER : Les algorithmes des moindres carrés rapides et leurs applications au filtrage numérique adaptatif. Rapport LSS MB 85 026, mai 85.
- [XIV] B.PICINBONO : Adaptive signal processing for detection and communication. NATO, 1978 pp 639, 660.
- [XV] O.MACCHI : Estimation linéaire adaptative - Application aux transmissions de données. Colloque national sur le traitement du signal & ses applications. Avril 1977, Nice.
- [XVI] M.DENTINO - J.Mc COOL - B. WIDROW : Adaptive filtering in the frequency domain. Proc. IEEE, Dec. 78, p. 1658 - 1667.
- [XVII] N.J.BERSHAD - P.L.FEINTUCH : Analysis of the frequency domain adaptive filter. Proc. IEEE, Dec. 79, p. 1658-1659.
- [XVIII] C.OGUE - T.SAITO - Y.HOSHIKO : A fast convergence frequency domain adaptive filter. ASSP Oct. 1983.
- [XIX] N.J.BERSHAD - P.L. FEINTUCH : A normalized frequency domain LMS adaptive algorithm. Prod. IEEE, Juin 86.
- [XX] D.BAUDOIS - P.LORENZINO - A.SILVENT : Sous-tructeurs de bruits. Etude d'une méthode de comparaison. Rapport CEPHAG n° 19-86
- [XXI] C.SERVIERE : Soustracteurs de bruit. Rapport DEA Sept.86.