

REDUCTION DE LA COMPLEXITE DES ALGORITHMES DES MOINDRES CARRÉS RAPIDES

Thierry PETILLON, André GILLOIRE
Centre National d'Etudes des Télécommunications
CNET LAA/TSS/CMC BP 40 22301 LANNION CEDEX FRANCE
et Groupement de Recherche CNRS Traitement du Signal et Image

RÉSUMÉ

Dans certaines applications comme l'annulation d'écho acoustique dans les mobiles, les algorithmes des moindres carrés rapides (MCR) sont arithmétiquement trop complexes pour une identification en temps réel des réponses impulsionnelles de grande taille. Une nouvelle classe d'algorithmes [1] est basée sur la réduction de l'ordre des prédicteurs. La complexité de ces algorithmes est alors fortement réduite si le signal d'excitation est prédictible par un modèle court. Nous présentons un algorithme de cette classe adapté à la parole comme signal d'excitation du filtre à identifier. Nous montrons expérimentalement que les performances obtenues sont très proches de celles des MCR pour une complexité très inférieure.

1 INTRODUCTION

L'identification adaptative des réponses impulsionnelles longues pose actuellement des problèmes de complexité et de performances. Deux approches sont possibles pour traiter ce problème. L'utilisation d'algorithmes de type gradient stochastique, traitant l'information par blocs [2], ou dans le domaine fréquence [3] permet principalement de réduire la complexité, et peut apporter simultanément un gain de performance. Une autre approche, qui privilégie la vitesse de convergence pour pallier la perte de performance due à la longueur du filtre, s'appuie sur l'utilisation des algorithmes de type moindres carrés rapides (MCR). Mais la complexité de la partie prédiction atteignant $6L$ multiplications par nouvelle observation, l'implantation d'une telle solution pour un filtre de longueur L supérieure à une centaine de points semble difficile sur un seul DSP flottant de la génération actuelle travaillant à une fréquence d'échantillonnage de 8 ou 16 kHz. Dans les MCR, l'ordre de prédiction du signal d'excitation dépend de la taille L du système à identifier, et non de la statistique de ce même signal. Une nouvelle classe d'algorithmes [1], appelés FNTF (Fast Newton Transversal Filter), introduit un degré de liberté supplémentaire quant au choix de l'ordre de prédiction N . La complexité de la partie prédiction passe de $O(L)$ à $O(N)$ et permet d'atteindre une complexité proche de celle du gradient stochastique, en conservant tout ou partie des qualités des MCR. Après avoir introduit ces nouveaux algorithmes, nous présentons une méthode d'initialisation, et un exemple de mise en oeuvre à partir du FTF [4]. Nous adaptons ensuite les techniques de réinitialisation nécessaire pour le traitement de la parole. Enfin, nous les comparons avec les MCR stabilisés numériquement (MCRS) classiques et le gradient stochastique normalisé

ABSTRACT

In some applications, like acoustic echo cancellation in mobiles, the computational cost of fast transversal recursive least squares algorithms is too high for real time identification of long impulse responses. Prediction order reduction has been recently introduced in a new class of algorithms [1]. Thus, the computational complexity can be highly reduced if the input signal is predictable with a short model. We present a member of this class which has been designed to deal with speech as input. Then, we demonstrate experimentally that the performance of this algorithm is almost equivalent to the usual FTF while its complexity is much smaller.

(NLMS), en termes de complexité et de performance sur des simulations d'annulation d'écho où les signaux d'excitation sont, soit des processus stationnaires, soit de la parole.

2 UNE NOUVELLE CLASSE D'ALGORITHMES

2.1 Les moindres carrés rapides stabilisés numériquement

Les MCR sous forme transverse, et en particulier le FTF [4], sont composés de deux parties indépendantes: la partie prédiction fournit le gain de Kalman Dual $\tilde{C}_{L,t}$ et la variable de vraisemblance $\gamma_{L,t}$ à la partie filtrage.

$$\bar{\epsilon}_{L,t} = y_t - H_{L,t-1}^T X_{L,t} \quad (\text{filtrage}) \quad (1)$$

$$H_{L,t} = H_{L,t-1} - \tilde{C}_{L,t} \gamma_{L,t} \bar{\epsilon}_{L,t} \quad (\text{adaptation}) \quad (2)$$

où $H_{L,t}^T = [H_{L,t}^1 \dots H_{L,t}^L]$ modélise la réponse impulsionnelle tronquée à L points, y_t est l'observation bruitée en sortie, $X_{L,t}$ le vecteur des L derniers échantillons du signal d'excitation et enfin $\bar{\epsilon}_{L,t}$ représente l'erreur de filtrage a priori. La partie prédiction du FTF stabilisé permet une mise à jour de $\tilde{C}_{L,t} = -1/\lambda R_{L,t-1}^{-1} X_{L,t}$ (3) et de $\gamma_{L,t} = 1/(1 - \tilde{C}_{L,t}^T X_{L,t})$ (4) en $6L$ multiplications au lieu des $O(L^2)$ du RLS. Rappelons pour la suite que les deux relations à la base de l'obtention des formes rapides par une approche algébrique sont,

$$R_{L+1,t}^{-1} = \begin{bmatrix} R_{L,t}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -b_{L,t} \\ 1 \end{bmatrix} \begin{bmatrix} -b_{L,t} & 1 \end{bmatrix} \frac{1}{\beta_{L,t}} \quad (5a)$$

$$R_{L+1,t}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & R_{L,t}^{-1} \end{bmatrix} + \begin{bmatrix} 1 \\ -a_{L,t} \end{bmatrix} \begin{bmatrix} 1 & -a_{L,t} \end{bmatrix} \frac{1}{\alpha_{L,t}} \quad (5b)$$



selon les partitionnements haut et bas de la matrice d'autocorrélation de taille $L+1$ ($R_{L+1,t}$) où les couples $(a_{L,t}, \alpha_{L,t})$ et $(b_{L,t}, \beta_{L,t})$ sont respectivement les prédicteurs et variances direct et rétrograde. Les algorithmes des moindres carrés rapides sont théoriquement équivalents aux moindres carrés récursifs (RLS). Cependant, l'accumulation des erreurs d'arrondi due à la précision finie des machines rend les MCR instables. L'analyse de la propagation des erreurs numériques a conduit à des méthodes de stabilisation sans dégradation des performances [5][6]. Le FTF8L utilisé dans les simulations est un algorithme stabilisé par contre réaction des erreurs numériques sur les variables instables du FTF. Cette version [6] nécessite le calcul des erreurs de prédiction retour par convolution et donc incrémente de L le nombre total de multiplications par échantillon (voir équations E1/E2/E3 en annexe).

2.2 Extrapolation de la matrice d'autocorrélation

La simplification obtenue dans les FNTF provient d'une estimation de $R_{L,t}$ par une méthode d'extrapolation à partir de $R_{N,t}$. Notons que ce paragraphe n'est pas nécessaire pour expliciter le nouvel algorithme mais justifie la troncature des prédicteurs. Dans le cas stationnaire gaussien, il est bien connu que l'extrapolation des N premiers coefficients de corrélation d'un processus peut s'effectuer grâce à une modélisation AR(N), équivalente, dans ce cas, à une extrapolation selon le principe du Maximum d'Entropie, ou encore à une maximisation de la variance de prédiction d'ordre N [7]. Reprenons l'analyse dans le cas non-stationnaire: supposons que les couples $(a_{N-1,t}, \alpha_{N-1,t})$ et $(b_{N-1,t}, \beta_{N-1,t})$ soient connus grâce à l'estimation de $R_{N,t}$ sur un processus et que nous voulions trouver $R_{N+1,t}$ à partir de ces seules données. Nous supposons en plus que ce processus est autorégressif d'ordre $N-1$. Nous imposons alors la troncature des prédicteurs d'ordre N tel que:

$$a_{N,t} = \begin{bmatrix} a_{N-1,t} \\ 0 \end{bmatrix} \quad \alpha_{N,t} = \alpha_{N-1,t} \quad (6)$$

$$b_{N,t} = \begin{bmatrix} 0 \\ b_{N-1,t-1} \end{bmatrix} \quad \beta_{N,t} = \beta_{N-1,t-1} \quad (7)$$

Ecrivons ensuite les équations de Yule-Walker d'ordre $N+1$ pour les deux prédicteurs d'ordre N où seuls $\hat{r}_{N,t}^a$ et $\hat{r}_{N,t}^r$ sont inconnus et représentent respectivement l'extrapolation de la matrice d'autocorrélation selon le prédicteur direct et selon le prédicteur rétrograde.

$$\begin{bmatrix} \Gamma_{0,t} & \Gamma_{N-1,t}^{aT} & \hat{r}_{N,t}^a \\ \Gamma_{N-1,t}^a & R_{N-1,t-1} & \Gamma_{N-1,t-1}^r \\ \hat{r}_{N,t}^r & \Gamma_{N-1,t-1}^{rT} & \Gamma_{0,t-N} \end{bmatrix} \begin{bmatrix} 1 \\ -a_{N-1,t} \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_{N-1,t} \\ 0_{N-1} \\ 0 \end{bmatrix} \quad (9a)$$

$$\begin{bmatrix} \Gamma_{0,t} & \Gamma_{N-1,t}^{aT} & \hat{r}_{N,t}^a \\ \Gamma_{N-1,t}^a & R_{N-1,t-1} & \Gamma_{N-1,t-1}^r \\ \hat{r}_{N,t}^r & \Gamma_{N-1,t-1}^{rT} & \Gamma_{0,t-N} \end{bmatrix} \begin{bmatrix} 0 \\ -b_{N-1,t-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0_{N-1} \\ \beta_{N-1,t-1} \end{bmatrix} \quad (9b)$$

avec les vecteurs de corrélation $\Gamma_{N,t}^{aT} = [r_{1,t} \ r_{2,t} \ \dots \ \hat{r}_{N,t}^a]$, $\Gamma_{N,t}^{rT} = [\hat{r}_{N,t}^r \ r_{N-1,t-1} \ \dots \ r_{1,t-N+1}]$ où $r_{i,t} = \lambda r_{i,t-1} + x_t x_{t-i}$ pour $i=1, \dots, N-1$. Les prédicteurs optimaux sont définis comme: $a_{N-1,t} = R_{N-1,t-1}^{-1} \Gamma_{N-1,t}^a$ et $b_{N-1,t-1} = R_{N-1,t-1}^{-1} \Gamma_{N-1,t-1}^r$. En prenant la dernière ligne de (9a) et la première de (9b), nous constatons que les deux approches aboutissent au même coefficient extrapolé, soit: $\hat{r}_{N,t}^a = \Gamma_{N-1,t}^{aT} R_{N-1,t-1}^{-1} \Gamma_{N-1,t-1}^r$ (10). Ceci nous enseigne que le

fait de tronquer les prédicteurs direct et rétrograde à un ordre $N-1$, permet implicitement d'extrapoler d'une manière unique la matrice d'autocorrélation d'ordre N . Il est prouvé [1] que cette matrice extrapolée est définie positive; elle peut donc représenter une matrice d'autocorrélation. De plus, il est possible d'itérer l'extrapolation de $R_{N,t}$ à $R_{L,t}$ par des formules récursives. Mais l'intérêt de la méthode d'extrapolation réside dans l'application des équations (5a) (5b) à l'inverse de cette matrice extrapolée grâce aux prédicteurs tronqués.

2.3 Un algorithme simplifié

Supposons maintenant que le processus observé soit autorégressif d'ordre N , alors la matrice d'autocorrélation d'ordre $L+1$ construite à partir d'une estimation à l'ordre N , possède des prédicteurs associés qui s'écrivent comme:

$$a_{L,t} = \begin{bmatrix} a_{N,t} \\ 0_{L-N} \end{bmatrix} \quad \alpha_{L,t} = \alpha_{N,t} \quad (11)$$

$$b_{L,t} = \begin{bmatrix} 0_{L-N} \\ b_{N,t-L+N} \end{bmatrix} \quad \beta_{L,t} = \beta_{N,t-L+N} \quad (12)$$

en itérant les propriétés (6) et (7). Appliquons maintenant (5a) et (5b) à la matrice de corrélation extrapolée:

$$R_{L+1,t}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & R_{L,t-1}^{-1} \end{bmatrix} + \begin{bmatrix} 1 \\ -a_{N,t} \\ 0_{L-N} \end{bmatrix} [1 \ -a_{N,t} \ 0_{L-N}] \frac{1}{\alpha_{N,t}} \quad (13)$$

en posant $t_D = t - L + N$, on obtient d'une manière similaire,

$$R_{L+1,t}^{-1} = \begin{bmatrix} R_{L,t}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0_{L-N} \\ -b_{N,t_D} \\ 1 \end{bmatrix} [0_{L-N} \ -b_{N,t_D} \ 1] \frac{1}{\beta_{N,t_D}} \quad (14)$$

Ces équations à l'instant $t-1$, multipliées à droite par $-\lambda^{-1} X_{L+1,t}$ selon (3) permettent la mise à jour du gain de Kalman Dual:

$$\tilde{C}_{L+1,t} = \begin{bmatrix} 0 \\ \tilde{C}_{L,t-1} \end{bmatrix} - \begin{bmatrix} 1 \\ -a_{N,t-1} \\ 0_{L-N} \end{bmatrix} \frac{\bar{e}_{N,t}}{\lambda \alpha_{N,t-1}} \quad (15)$$

ou,

$$\tilde{C}_{L+1,t} = \begin{bmatrix} \tilde{C}_{L,t} \\ 0 \end{bmatrix} - \begin{bmatrix} 0_{L-N} \\ -b_{N,t_D-1} \\ 1 \end{bmatrix} \frac{\bar{r}_{N,t_D}}{\lambda \beta_{N,t_D-1}} \quad (16)$$

De plus, ces dernières équations multipliées à droite par $X_{L,t}$ selon

(4), nous donnent une mise à jour de $\delta_{L,t} = \tilde{C}_{L,t}^T X_{L,t}$:

$$\delta_{L+1,t} = \delta_{L,t-1} - \frac{\bar{e}_{N,t}^2}{\lambda \alpha_{N,t-1}} \quad (17) \quad \delta_{L+1,t} = \delta_{L,t} - \frac{\bar{r}_{L,t}^2}{\lambda \beta_{L,t-1}} \quad (18)$$

En posant

$$\tilde{S}_{N+1,t} = \begin{bmatrix} 1 \\ -a_{N,t-1} \end{bmatrix} \frac{\bar{e}_{N,t}}{\lambda \alpha_{N,t-1}} \quad \text{et} \quad \tilde{U}_{N+1,t} = \begin{bmatrix} -b_{N,t-1} \\ 1 \end{bmatrix} \frac{\bar{r}_{N,t-1}}{\lambda \beta_{N,t-1}}$$

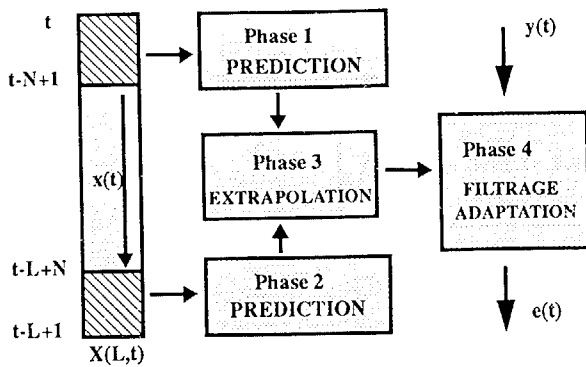
nous résumons ci-dessous la technique de mise à jour du gain de Kalman Dual d'ordre L ainsi que celle de la variable de vraisemblance.

$$\tilde{C}_{L+1,t} = \begin{bmatrix} 0 \\ \tilde{C}_{L,t-1} \end{bmatrix} - \begin{bmatrix} \tilde{S}_{N+1,t} \\ 0_{L-N} \end{bmatrix} \quad (19) \quad \begin{bmatrix} \tilde{C}_{L,t} \\ 0 \end{bmatrix} = \tilde{C}_{L+1,t} - \begin{bmatrix} 0_{L-N} \\ \tilde{U}_{N+1,t} \end{bmatrix} \quad (20)$$

$$\delta_{L,t} = \delta_{L,t-1} + \tilde{S}_{N,t}^1 \bar{e}_{N,t} + \tilde{U}_{N,t_D}^{N+1} \bar{r}_{N,t_D} \quad (21) \quad \gamma_{L,t} = \frac{1}{\delta_{L,t}} \quad (22)$$



Les équations précédentes permettent de constituer un algorithme comportant quatre phases représentées sur le synoptique suivant:



Phase 1: La partie prédiction d'ordre N d'un algorithme rapide (ex: FTF), traitant l'information à l'instant t permettant de calculer $\hat{s}_{L,t}$ et l'erreur de prédiction directe a priori $\bar{e}_{N,t}$. **Phase 2:** Cette même partie prédiction estime $\hat{U}_{N,t}$ et l'erreur de prédiction rétrograde a priori $\bar{r}_{N,t}$ à l'instant $t-L+N$. **Phase 3:** Mise à jour gain de Kalman Dual et de la variable de vraisemblance d'ordre L (Partie Extrapolation). **Phase 4:** Filtrage et adaptation

2.4 Initialisation, stabilisation et implantations

La présence de deux prédicteurs de même ordre N agissant sur des données décalées dans le temps de $L-N$ périodes permet d'envisager a priori trois types d'implantation. **Type a:** un recalcul complet du vecteur $\hat{U}_{N,t}$ et de $\bar{r}_{N,t}$ à l'instant t_D . **Type b:** stockage du vecteur et des deux scalaires pendant $L-N$ itérations. **Type c:** un recalcul partiel associé à un stockage. La seconde solution qui nécessite $(L-N) \times (N+2)$ mémoires est rédhibitoire sur les microprocesseurs de traitement de signal actuels. Nous avons choisi la solution de recalculer l'ensemble des variables par un second prédicteur retardé (type a). L'initialisation et la complexité, détaillées ci-après, concernent implicitement ce type d'implantation. Dans un contexte de préfenêtrage des signaux, seul $x_{-L} = \sqrt{e_0}$ est non nul avant le début de l'exécution, les variances des prédicteurs aux instant $t=0$ et $t=-L+N$ s'écrivent:

$\beta_{N,0} = \lambda^{L-N} e_0$	$\alpha_{N,0} = \lambda^L e_0$	$\beta_{N,-L+N} = e_0$	$\alpha_{N,-L+N} = \lambda^N e_0$
-----------------------------------	--------------------------------	------------------------	-----------------------------------

Quant aux problèmes de stabilisation intrinsèques à la partie prédiction des MCR, ils seront résolus directement par l'utilisation d'algorithmes stabilisés dans les phases (1) et (2) de l'algorithme complet car la partie extrapolation (phase 3) n'engendre pas d'instabilité numérique.

2.5 Complexité

Le degré de liberté supplémentaire apporté par le choix de N permet de trouver un bon compromis entre les performances nécessaires et la complexité imposée par les contraintes temps réel. Nous avons estimé dans le tableau ci-après, la complexité de l'algorithme simplifié, appelé FTF12N2L, en nombre de multiplications, divisions et mémoires, et comparé avec deux autres algorithmes classiques (FTF8L, NLMS) sur un exemple classique d'annulation d'écho acoustique dans les mobiles. Le temps de réverbération de l'habitacle du véhicule étant faible dans ce contexte, il est possible de tronquer la réponse impulsionnelle estimée à $L = 256$ points. Le choix de la taille du prédicteur correspond à celui de la modélisation des formants de la parole ($8 \leq N \leq 16$). Nous supposons de plus que la complexité d'une

division est équivalente à celle de 16 multiplications (mult.). Les rapports de complexité avec le NLMS sont indiqués dans les colonnes /NLMS.

$L=256$	Mult.	/NLMS	Mémoire	/NLMS
FTF8L	2096	4.2	1280	2.5
FTF12N2L(N=16)	704	1.4	864	1.7
FTF12N2L(N=8)	608	1.2	816	1.6
NLMS	512	1	512	1

Il est intéressant de noter que les simplifications dans la partie prédiction permettent de réduire le nombre de multiplications par 3 par rapport aux moindres carrés rapides classiques et d'atteindre une complexité inférieure au double de celle du gradient stochastique. De plus, la réduction du nombre de multiplications s'accompagne d'une réduction de la taille mémoire utilisée. Notons que plus L devient grand devant N , plus la partie prédiction devient négligeable en complexité devant la partie filtrage.

3 SIMULATIONS

3.1 Comparaison sur un processus stationnaire

Pour tester l'algorithme à complexité réduite, nous avons d'abord simulé une identification de réponse sur un processus AR stationnaire:

$-x(t)$ est un signal stationnaire dont le spectre est formé de trois formants d'une voyelle (U en l'occurrence) dont le modèle provient d'une analyse LPC.

$-y(t)$ est le résultat de la convolution de $x(t)$ par une réponse impulsionnelle de voiture tronquée à 256 points. La comparaison des algorithmes est faite sur l'évaluation temporelle de l'énergie moyenne de l'écho résiduel (EQMR) calculée par blocs de 256 échantillons successifs.

remarques: Nous avons choisi, a priori, un facteur d'oubli homogène à la taille du filtre et non à la taille du prédicteur pour la comparaison avec les MCR classiques ($\lambda = 0.9987$). Le signal de sortie y , n'étant pas bruité et L étant exact, l'EQMR asymptotique représente le bruit de quantification.

Nous observons sur la figure 2 des performances du même ordre pour $N=L$, $N=100$, $N=16$, $N=8$, une forte dégradation pour $N=4$, et une similitude entre $N=0$ et le NLMS. Une interprétation qualitative en terme de pouvoir de décorrélation ou de préblanchiment du signal d'excitation par la partie prédiction explique ces résultats. En effet, la forte dégradation pour $N=4$ vient de la modélisation incomplète des trois formants, alors que pour $N \geq 8$ les résultats sont proches de ceux des MCR. Remarquons aussi les performances prévisibles pour $N=0$ et pour le NLMS où il n'y pas de prédiction. Nous n'avons pas prouvé que les algorithmes simplifiés convergent vers la solution optimale, mais nous vérifions que l'EQMR asymptotique est égale à celle des MCR classiques, du moins pour N supérieur ou égal à 8. Notons que lorsque $L \gg N$ (ex: $L=1024$), la performance du FTF12N2L est sensiblement dégradée par rapport à celle FTF8L. Une explication plausible de cette dégradation est que la matrice d'autocorrélation implicitement extrapolée ne représente alors plus le processus observé.

3.2 Comparaison sur de la parole

Pour comparer les différents algorithmes sur des signaux naturels, nous utilisons de la parole. La technique de réinitialisation appliquée sur la partie prédiction, nécessaire sur ces signaux [8], est aussi utilisée sur les deux prédicteurs décalés dans le temps. Cette technique est basée sur la scrutation de $\gamma_{N,t}$ qui chute vers 0 lorsque l'algorithme se bloque à cause des non-stationnarités de



la parole. Dans ce cas, nous réinitialisons la partie prédiction, tout en conservant le filtre $H_{N,t}$ déjà estimé. Une valeur classique pour le seuil de détection γ_{\min} est 0.1. Nous constatons sur la figure 3 une confirmation des résultats obtenus dans le cas stationnaire. Le FTF12N2L converge beaucoup plus rapidement que le NLMS et atteint rapidement le régime asymptotique du FTF8L.

4 CONCLUSIONS

Les nouveaux algorithmes MCR à complexité réduite permettent d'augmenter d'une manière significative la taille maximale des filtres identifiables en temps réel par une technique de moindres carrés sur un processeur donné. Le degré de liberté apporté par le choix de la taille du prédicteur donne une grande souplesse d'utilisation, et peut résoudre le compromis entre complexité et performances. De plus, ils héritent directement de tous les résultats acquis sur les MCR classiques. Quant à leur capacité de poursuite, des simulations ont montré des performances équivalentes à celles des MCR à taux d'oubli identique.

Remerciements: Nous remercions Sergios Theodoridis (Université de Patras-Grèce) pour sa collaboration à la partie théorique de ce travail.

ANNEXE

Partie Prédiction du FTF8N(phases 1 et 2)

$\bar{e}_{N,t} = x_t - a_{N,t-1}^T X_{N,t-1}$ $e_{N,t} = \gamma_{N,t-1} \bar{e}_{N,t}$ $\alpha_{N,t} = \lambda \alpha_{N,t-1} + \bar{e}_{N,t} e_{N,t}$ $\gamma_{N+1,t} = \frac{\lambda \alpha_{N,t-1}}{\alpha_{N,t}} \gamma_{N,t-1}$ $\tilde{S}_{N+1,t} = \frac{\bar{e}_{N,t}}{\lambda \alpha_{N,t-1}} \begin{bmatrix} 1 \\ -a_{N,t-1} \end{bmatrix}$ $\tilde{C}_{N+1,t} = \begin{bmatrix} 0 \\ \tilde{C}_{N,t-1} \end{bmatrix} - \tilde{S}_{N+1,t}$ $a_{N,t} = a_{N,t-1} - e_{N,t} \tilde{C}_{N,t-1}$ $\tilde{r}_{N,t}^{conv} = x_{t-N} - b_{N,t-1}^T X_{N,t-1} \quad (E1)$	$\eta_{N,t} = \tilde{r}_{N,t}^{conv} + \lambda \beta_{N,t-1} \tilde{C}_{N+1,t} \quad (E2)$ $\tilde{r}_{L,t} = \tilde{r}_{L,t}^{conv} + \eta_{L,t} \quad (E3)$ $\gamma_{L,t} = \frac{\gamma_{L+1,t}}{1 + \gamma_{L+1,t} \tilde{r}_{L,t}^T \tilde{C}_{L+1,t}^L}$ $r_{N,t} = \gamma_{N,t} \tilde{r}_{N,t}$ $\tilde{U}_{N+1,t} = \tilde{C}_{N+1,t} \begin{bmatrix} -b_{N,t-1} \\ 1 \end{bmatrix}$ $\begin{bmatrix} \tilde{C}_{N,t} \\ 0 \end{bmatrix} = \tilde{C}_{N+1,t} - \tilde{U}_{N+1,t}$ $b_{N,t} = b_{N,t-1} - r_{N,t} \tilde{C}_{N,t}$ $\beta_{N,t} = \lambda \beta_{N,t-1} + \tilde{r}_{N,t} r_{N,t}$
--	--

Partie Extrapolation (Phase 3)

$\tilde{C}_{L+1,t} = \begin{bmatrix} 0 \\ \tilde{C}_{L,t-1} \end{bmatrix} - \begin{bmatrix} \tilde{S}_{N+1,t} \\ 0_{L-N} \end{bmatrix} \quad \begin{bmatrix} \tilde{C}_{L,t} \\ 0 \end{bmatrix} = \tilde{C}_{L+1,t} - \begin{bmatrix} 0_{L-N} \\ \tilde{U}_{N+1,t,D} \end{bmatrix}$ $\delta_{L,t} = \delta_{L,t-1} + \tilde{S}_{N,t}^T \bar{e}_{N,t} + \tilde{U}_{N+1,t,D}^T \tilde{r}_{N,t,D} \quad \gamma_{L,t} = \frac{1}{1 - \delta_{L,t}}$

Partie Filtrage (Phase 4)

$\bar{e}_{L,t} = y_t - H_{L,t-1}^T X_{L,t} \quad H_{L,t} = H_{L,t-1} - \tilde{C}_{L,t} \gamma_{L,t} \bar{e}_{L,t}$
--

Conditions initiales

<p>Prédiction non retardée: (Phase 1):</p> $a_{N,0} = 0_N \quad b_{N,0} = 0_N \quad \gamma_{N,0} = 1 \quad C_{N,0} = 0_N \quad \alpha_{N,0} = \lambda^L e_0 \quad \beta_{N,0} = \lambda^{L-N} e_0$ <p>Prédiction retardée: (phase 2):</p> $a_{N,0} = 0_N \quad b_{N,0} = 0_N \quad \gamma_{N,0} = 1 \quad C_{N,0} = 0_N \quad \alpha_{N,0} = \lambda^N e_0 \quad \beta_{N,0} = e_0$ <p>$e_0 \approx \frac{\sigma^2}{1-\lambda}$ σ^2 variance du signal $x(t)$</p> <p>Extrapolation et filtrage:</p> $\tilde{C}_{L,0} = 0_L \quad \delta_{L,0} = 0 \quad H_{L,0} = 0_L$
--

FIGURE 2 Performances du FTF à complexité réduite sur signal stationnaire à spectre de voyelle

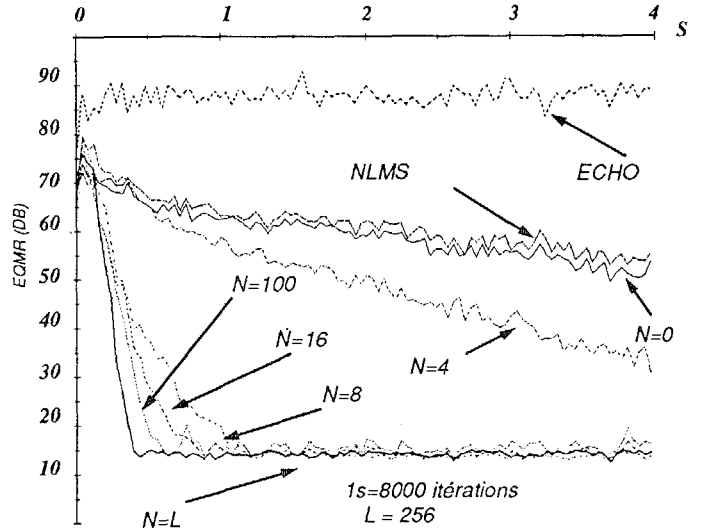
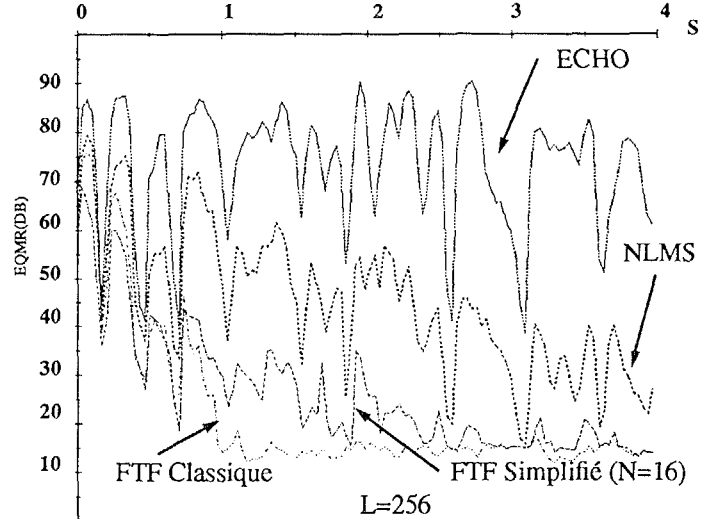


FIGURE 3 Performances du FTF à complexité réduite sur signal de parole



REFERENCES

[1] G.V MOUSTAKIDES S.THEODORIDIS
Fast Newton Transversal filters
A New class of Adaptive Estimation
submitted to IEEE ASSP

[2] J. BENESTY P.DUHAMEL
A Fast Exact Least Mean Square Adaptive Algorithm
submitted to IEEE ASSP

[3] M. XU Y. GRENIER
Time Frequency Domain Adaptive Filters
ICASSP 89

[4] J.CIOFFI T.KAILATH
Fast RLS transversal filters for adaptive filtering
IEEE ASSP vol 32 pp 304-337 April 1984

[5] D.T.M SLOCK et T.KAILATH
Numerically Stable fast RLS Transversal filters
IEEE ICASSP 88 New-York

[6] A.BENALLAL A.GILLOIRE
A New method to stabilize Fast RLS Algorithms Based
on a first-order model of propagation of numerical errors
ICASSP 1988 New York

[7] B. PICINBONO M.BARRET
Nouvelle Présentation de la Méthode du Maximum d'Entropie
Traitement du Signal Vol 7 N°2 1990

[8] A.BENALLAL A.GILLOIRE
Instabilité et stabilité des MCR transversaux
rapides excités par la parole.
GRETSI 1989