

PROPRIETES ET OPTIMISATION DU TAUX DE CONVERGENCE
DES ALGORITHMES DE FILTRAGE EN TREILLIS/GRADIENT

J. BOUDY - G. FAVIER

CNRS - LASSY- 41 Bd Napoléon III- 06041 NICE CEDEX
et
GRECO "Systèmes Adaptatifs"

Résumé : Cet article est consacré à l'étude théorique des propriétés de convergence en moyenne et en moyenne quadratique du filtre en treillis/gradient adaptatif. D'autre part une méthode de correction du taux de convergence est proposée en vue d'améliorer la capacité d'adaptation du filtre vis à vis de non stationnarités du signal à filtrer. Des résultats de simulation sont présentés pour illustrer le comportement du filtre en treillis/gradient sans correction et avec correction du taux de convergence.

Summary : This paper is concerned with the theoretical study of convergence properties in the mean and the mean squares senses of the gradient adaptive lattice filter. Moreover a correction method of the convergence rate is proposed in order to improve the filter adaptation capability with respect to non stationarities of the signal to be filtered. Simulation results are presented to illustrate the behavior of the gradient adaptive filter without correction and with correction of the convergence rate.

I. INTRODUCTION

Les filtres en treillis ont fait l'objet de nombreuses publications ces dernières années tant sur le plan théorique qu'au niveau des applications (analyse et synthèse de la parole, estimation spectrale, égalisation de données, déconvolution de signaux sismiques, traitement d'antenne, reconnaissance et détection de fouillis). Ces filtres sont caractérisés par des paramètres, appelés coefficients de corrélation partielle (PARCOR) direct et rétrograde.

Dans la plupart des applications de traitement du signal, l'entrée du filtre ne pouvant être supposée stationnaire, il est fait appel à des filtres en treillis adaptatifs, c'est à dire pour lesquels les coefficients PARCOR sont réajustés à chaque instant. Deux types de méthodes récursives peuvent-être utilisés : des méthodes optimales basées sur l'utilisation des Moindres Carrés (treillis/Moindres Carrés) et des méthodes sous-optimales utilisant l'algorithme du gradient (treillis/gradient) ou l'algorithme du signe.

Cet article est consacré à l'étude théorique des propriétés de convergence des algorithmes en treillis/gradient introduits par [Griffiths -1977,1978] et [Griffiths, Medaugh - 1979], qui présentent l'intérêt d'une plus grande simplicité de mise en oeuvre que les algorithmes en treillis/Moindres Carrés.

En revanche il est assez difficile d'effectuer le réglage du taux de convergence, paramètre caractéristique de la méthode du gradient, pour chaque cellule du filtre en treillis. Griffiths et Medaugh ont proposé une normalisation de ce taux de convergence afin d'obtenir une marge de fonctionnement assez large tout en respectant le caractère d'indépendance de chaque étage du filtre en treillis. Cette normalisation peut être rapprochée de celle développée par [Frost-1972], [Bienvenu-1973] et [Widrow et al.-1976] dans le cas d'un filtre transverse, et elle est sous-tendue par une évaluation des bornes théoriques de convergence en moyenne et en moyenne quadratique de l'algorithme LMS.

En s'inspirant de ces travaux nous montrons comment peuvent-être obtenues les conditions

de convergence théoriques relatives à l'algorithme en treillis/gradient de Griffiths et Medaugh. D'autre part, nous proposons une méthode de correction du taux de convergence dans le cas d'un estimateur biaisé de la covariance des résidus avance et rétrograde présents dans le filtre. Des résultats de simulation sont présentés pour illustrer le comportement des filtres en treillis/gradient (version normalisée, avec correction et sans correction) vis-à-vis de non stationnarités du signal à filtrer.

II. FILTRE EN TREILLIS/GRADIENT

La structure standard d'un filtre en treillis d'ordre N est décrite à l'aide des équations suivantes :

$$\epsilon_n(t) = \epsilon_n(t) - K_n(t) r_{n-1}(t-1) \quad (1)$$

$$r_n(t) = r_{n-1}(t-1) - K_n(t) \epsilon_n(t) \quad (2)$$

Les coefficients K_n et K_n , paramètres caractéristiques de la cellule (n), sont appelés

coefficients de corrélation partielle (partial correlation : PARCOR) direct et rétrograde respectivement. D'autre part les signaux ϵ_n et r_n sont appelés résidus direct et rétrograde respectivement. Le filtre en treillis/gradient à deux coefficients [Favier-1986] est obtenu en choisissant K_n^E et K_n^R de manière à minimiser les énergies moyennes des résidus rétrograde et direct respectivement, ce qui donne :

$$\epsilon_n(t+1) = K_n(t) + 2\mu \epsilon_n(t) r_{n-1}(t) \quad (3)$$

$$K_n(t+1) = K_n(t) + 2\mu r_{n-1}(t-1) \epsilon_n(t) \quad (4)$$

Les solutions optimales correspondantes des coefficients PARCOR sont :

$$K_n^* = E[\epsilon_n(t) r_{n-1}(t-1)] / E[\epsilon_n(t)] \quad (5)$$

$$K_n^* = E[r_{n-1}(t-1) \epsilon_n(t)] / E[r_{n-1}(t-1)] \quad (6)$$



III. PROPRIETES DE CONVERGENCE

En raison de la non linéarité des équations (3)-(4) vis-à-vis des signaux résiduels, il est nécessaire de faire des hypothèses simplificatrices afin de pouvoir déterminer les propriétés de convergence du filtre en treillis/gradient. Ces hypothèses sont identiques à celles utilisées par [Griffiths, Medaugh-1979] :

- a) Tous les signaux intervenant dans le filtre sont supposés gaussiens et centrés.
- b) Les coefficients PARCOR calculés récursivement à l'aide de (3)-(4) sont supposés être indépendants des signaux intervenant dans le filtre.

Convergence en moyenne du filtre :

En définissant les écarts ΔK_n , ΔK_n^r par rapport à la solution optimale (K_n) , (K_n^r)

définie à l'aide de (5) - (6) :

$$\Delta K_n(t) = K_n(t) - (K_n) \quad (7)$$

$$\Delta K_n^r(t) = K_n^r(t) - (K_n^r) \quad (8)$$

et en utilisant les relations (2)-(3), nous obtenons :

$$K_n(t+1) = K_n(t) + 2\mu_n [\epsilon(t)r(t-1) - K_n(t)\epsilon(t)] \quad (9)$$

Par suite, compte tenu de l'hypothèse b) et de l'expression (5), nous déduisons :

$$E[\Delta K_n(t+1)] = (1 - 2\mu_n E[\epsilon(t)^2]) E[\Delta K_n(t)] \quad (10)$$

D'où nous pouvons tirer la condition de convergence en moyenne du coefficient K_n , que

doit satisfaire le taux de convergence μ_n :

$$0 < \mu_n < 1/E[\epsilon(t)^2] \quad (11)$$

En procédant de la même façon pour le coefficient K_n^r nous obtenons la condition de convergence en moyenne suivante pour le taux

de convergence μ_n^r :

$$0 < \mu_n^r < 1/E[r(t-1)^2] \quad (12)$$

Convergence en moyenne quadratique du filtre: En supposant que les conditions de convergence en moyenne (11) et (12) soient satisfaites, nous allons déterminer maintenant les conditions de convergence en moyenne quadratique du filtre en treillis/gradient, c'est à dire les conditions pour que la variance des écarts $\Delta K_n^e(t)$ et $\Delta K_n^r(t)$ reste bornée. Les calculs étant identiques pour chacun des écarts nous développons ci-après les équations relatives à $\Delta K_n^e(t)$. A partir de (7) et (9) nous avons :

$$\Delta K_n^e(t+1) = A(t) \Delta K_n^e(t) + 2\mu_n r(t) \epsilon(t) \quad (13)$$

avec

$$A(t) = 1 - 2\mu_n E[\epsilon(t)^2] \quad (14)$$

$$r(t) = r(t-1) - (K_n^r) \epsilon(t) \quad (15)$$

où $r(t)$ représente le résidu rétrograde op-

timal, obtenu en remplaçant $K_n(t)$ par (K_n) dans (2).

Par suite la variance de l'écart $\Delta K_n(t+1)$

est égale à :

$$E\{[\Delta K_n(t+1)]^2\} = E\{A(t)^2 [\Delta K_n(t)]^2\} + 4\mu_n^2 E\{A(t)r(t)\epsilon(t)\Delta K_n(t)\} + 4\mu_n^2 E\{r(t)\epsilon(t)\} \quad (16)$$

En prenant en compte les hypothèses a) et b) nous avons :

$$E\{A(t)^2 [\Delta K_n(t)]^2\} = E\{A(t)^2\} E\{[\Delta K_n(t)]^2\} \quad (17)$$

avec

$$E\{A(t)^2\} = 1 - 4\mu_n^2 E[\epsilon(t)^2] + 12\mu_n^2 E[\epsilon(t)] E[\epsilon(t)] \quad (18)$$

D'autre part

$$E\{r(t)\epsilon(t)\} = \{E[r(t-1)] - (K_n^r) E[\epsilon(t)]\} E[\epsilon(t)] \quad (19)$$

où le terme entre accolades représente la variance du résidu rétrograde optimal $r(t)$, soit :

$$E\{r(t)\epsilon(t)\} = E\{r(t)\} E[\epsilon(t)] \quad (20)$$

Enfin en supposant que la condition de convergence (11) soit satisfaite, le deuxième terme du membre de droite de (16) est nul.

Par suite, la variance de $\Delta K_n(t+1)$ s'écrit

$$E[\Delta K_n(t+1)]^2 = E\{A(t)^2\} E[\Delta K_n(t)]^2 + 4\mu_n^2 E\{r(t)\epsilon(t)\} E[\epsilon(t)] \quad (21)$$

D'où il est facile de tirer la condition de convergence en moyenne quadratique

$$E\{A(t)^2\} < 1 \quad (22)$$

ou encore à partir de l'expression (18)

$$0 < \mu_n < 1/3 E[\epsilon(t)^2] \quad (23)$$

En procédant de la même façon pour le coefficient K_n^r nous obtenons la condition de

convergence en moyenne quadratique suivante

$$0 < \mu_n^r < 1/3 E[r(t-1)^2] \quad (24)$$

Comme il était prévisible les conditions de convergence (23)-(24) sont plus restrictives que les conditions (11)-(12).

IV. METHODE DE CORRECTION DES TAUX DE CONVERGENCE

L'introduction d'une méthode de correction des taux de convergence a pour but d'améliorer les propriétés de convergence du filtre en treillis/gradient de manière à aboutir à un compromis entre d'une part la capacité d'adaptation du filtre vis à vis de non stationnarités du signal à filtrer et d'autre part la sensibilité au bruit en régime permanent.

Dans ce qui suit nous considérons la version normalisée à deux coefficients du filtre en treillis/gradient. La normalisation est réalisée par rapport à l'énergie des résidus direct et rétrograde se trouvant en entrée de chaque cellule c'est-à-dire :

$$\mu = \alpha / E[\epsilon(t)] \quad (25) \quad \mu = \alpha / E[r(t-1)] \quad (26)$$

Il est à noter que pour satisfaire les conditions de convergence en moyenne (11)-(12) et en moyenne quadratique (23)-(24), nous devons avoir :

$$\alpha < 1/3 \quad (27)$$

En pratique la variance des résidus direct et rétrograde est estimée à l'aide de l'estimateur biaisé suivant [Friedlander-1982], [Honig, Messerschmitt-1981] :

$$R_n(t) = \lambda R_n(t-1) + \epsilon_n(t) \quad (28)$$

$$R_n(t) = \lambda R_n(t-1) + r_n(t-1) \quad (29)$$

avec $0 < \lambda < 1$, ce qui conduit aux valeurs estimées suivantes des taux de convergence

$$\mu(t) = \alpha(t)/R_n(t) \quad (30) \quad \mu(t) = \alpha(t)/R_n(t) \quad (31)$$

Pour obtenir les formules de correction relatives à μ et μ nous déterminons $\alpha(t)$

et $\alpha(t)$ de manière que (30) et (31) soient des estimateurs non biaisés de (25) et (26) respectivement. Cela nous donne :

$$\alpha(t) = \alpha E[R_n(t)] / E[\epsilon(t)] \quad (32)$$

$$\alpha(t) = \alpha E[r(t)] / E[r(t-1)] \quad (33)$$

En remarquant que (28) peut s'écrire aussi :

$$R_n(t) = \sum_{k=t_0+1}^t \lambda^{t-k} \epsilon_n(k) + \lambda^{t-t_0} R_n(t_0) \quad (34)$$

où t_0 représente l'instant initial de la fenêtre de calcul considérée, et en notant que pour tout $n \in \{0, N\}$:

$$E[\epsilon_n(t_0+1)] \geq E[\epsilon_n(t_0+2)] \geq \dots \geq E[\epsilon_n(t)] \quad (35) \text{ et}$$

$$E[\epsilon_n(t)] \leq R_0 \quad \text{où} \quad R_0 = E[y_n(t)] \quad (36)$$

nous pouvons déduire l'inégalité suivante

$$E[R_n(t)]/E[\epsilon_n(t)] \geq \frac{1-\lambda^{t-t_0}}{1-\lambda} + \lambda^{t-t_0} R_n(t_0)/R_0 \quad (37)$$

et par suite la méthode de correction que nous proposons consiste à prendre

$$\alpha_n(t) = \alpha \frac{1 - \lambda^{t-t_0}}{1 - \lambda} + \lambda^{t-t_0} \frac{R_n(t_0)}{R_0} \quad (38)$$

où α est choisi de manière à réaliser un bon compromis entre la sensibilité au bruit et la capacité d'adaptation de l'algorithme.

De la même façon nous choisissons :

$$\alpha_n(t) = \alpha \frac{1 - \lambda^{t-t_0}}{1 - \lambda} + \lambda^{t-t_0} \frac{R_n(t_0)}{R_0} \quad (39)$$

V. SIMULATIONS

Les simulations que nous présentons ci-après se rapportent d'une part à l'étude des propriétés de convergence du filtre en treillis gradient, d'autre part à la méthode de correction développée au § IV. Pour l'ensemble des essais nous avons utilisé une formule de correction de la même forme que (38) ou (39), en fixant :

$$\alpha_n(t) = \alpha_n(t) = \alpha_n(t), \text{ avec } \alpha = k \quad \forall n \in N$$

$$R_n(t_0) = R_n(t_0) = R_n(t_0) = 1 \text{ et } R_0 = 1$$

Ceci se traduit par l'utilisation d'une formule unique de correction pour le calcul de $\alpha(t)$.

Les figures 1-12 ont été obtenues à partir du traitement d'un processus AR d'ordre 2 obtenu par passage d'un bruit blanc de variance unité à travers la fonction de transfert $1/A_1(z)$ avec $A_1(z) = 1 - 1,6z^{-1} + 0,95z^{-2}$. Ces figures représentent, en fonction du nombre d'itérations traitées t , le premier coefficient estimé $a_1(t)$ obtenu à partir des valeurs estimées de K_1 et K_2 (Fig 1, 4, 7, 10) et une valeur estimée, à l'aide d'une moyenne temporelle, des quantités $E[\Delta K_1]$ (Fig 2, 5, 8, 11) et $E[(\Delta K_1)^2]$ (Fig 3, 6, 9, 12). L'ensemble de ces figures permet de vérifier les conditions théoriques de convergence en moyenne et en moyenne quadratique établies au §III. A noter cependant que pour $\alpha_n = k \geq 1/3$ le paramètre a_1 n'est plus identifié correctement. Le deuxième ensemble de simulations (Fig 13-21) est consacré à l'étude de la méthode de correction lorsque le signal d'entrée du filtre présente une non stationnarité (rupture de modèle à la 500ème itération : $A_1(z) \rightarrow A_2(z) = 1 - 1,8z^{-1} + 1,6z^{-2} - 0,7z^{-3}$; voir fig 13). A l'appui de ces simulations, nous pouvons formuler un certain nombre de remarques intéressantes :

-Pour une même vitesse de convergence à l'initialisation, il y a augmentation de la capacité d'adaptation de l'algorithme au niveau de la rupture du modèle lorsqu'on utilise la méthode de correction (comparer fig. 16 et 17), au détriment de l'apparition d'oscillations plus importantes durant le transitoire qui suit la rupture.

-Inversement, pour une même capacité d'adaptation à la rupture, le régime transitoire présente de plus fortes oscillations dans le cas sans correction (Fig 20) que dans le cas avec correction (fig 18).

-Les figures 14 et 15 ainsi que 19 et 20 mettent en évidence, en régime permanent, l'intérêt du lissage des paramètres estimés à l'aide d'un filtre du 1er ordre.

-La variance de l'erreur d'estimation des paramètres AR augmente proportionnellement à la valeur du taux de convergence $\alpha = k$. (voir fig 15, 18).

-Dans le cas sans correction, nous pouvons ajouter que, pour une même capacité d'adaptation à la rupture et pour une erreur d'estimation comparable en régime permanent, les transitoires à l'initialisation sont plus rapides lorsque le taux de convergence est normalisé à l'aide d'un estimateur biaisé de la variance des résidus (équations (28)-(31) ; fig 14) que dans le cas où l'on utilise un estimateur non biaisé (fig 21).

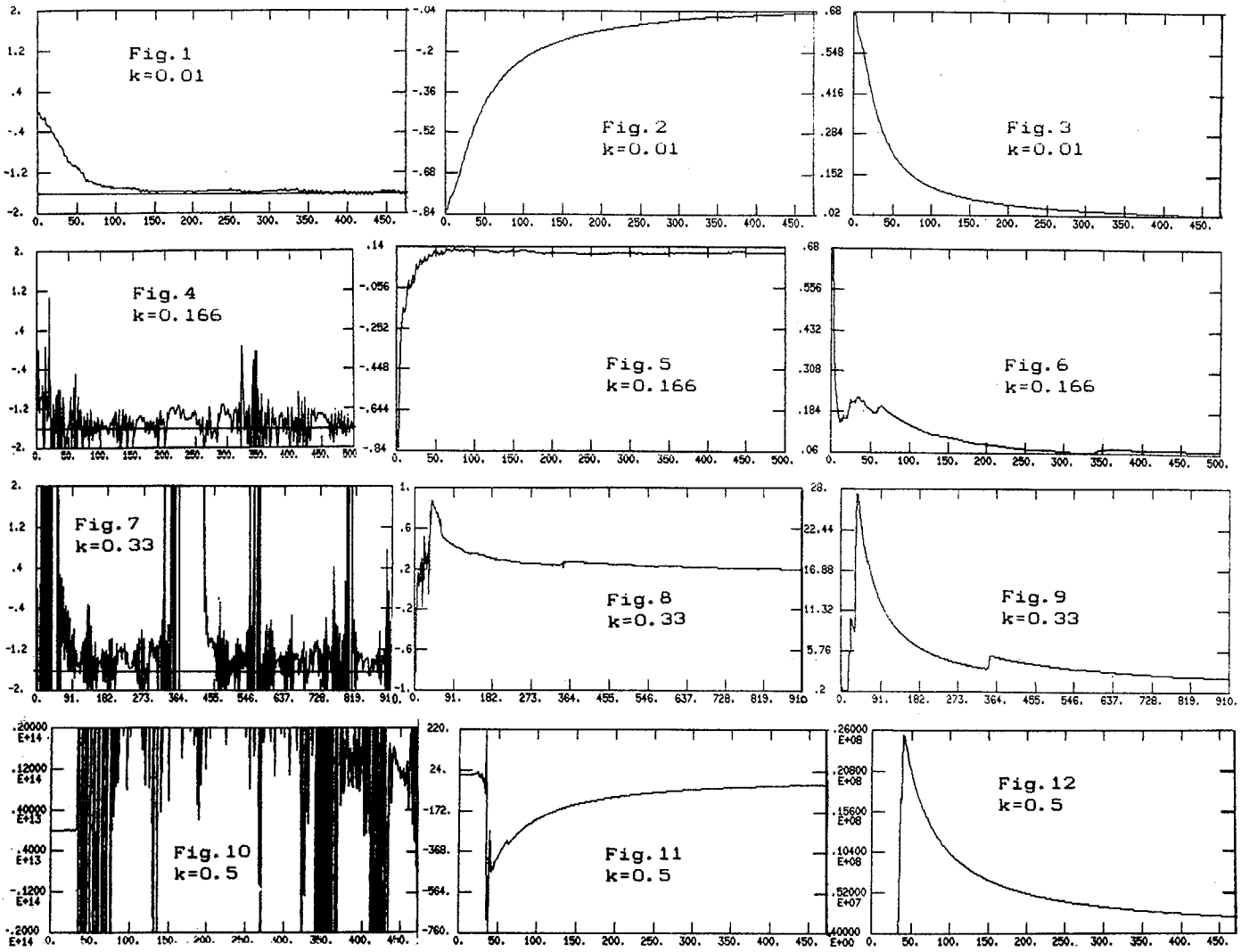
References :

[Bienvenu-1973]: Thèse de D. I. Ursay
 [Favier-1986]: Ann. Télécom. tome 41, no 5-6 pp. 305-321
 [Frost-1972]: Proc. IEEE vol. 60 pp. 926-935
 [Griffiths-1977, 1978]: Proc. ICASSP 77 pp. 683-686 et ICASSP 78 pp. 87-90
 [Griffiths, Medaugh-1979]: Proc. IEEE-UDC 79 pp. 1357-1361
 [Friedlander-1982]: Proc. IEEE no 8 pp. 829-867
 [Honig, Messerschmitt-1981]: IEEE ASSP no 9 pp. 642-653
 [Widrow et al.-1976]: Proc. IEEE vol. 64 no 8 pp. 1151-1162



1ere Partie des simulations : PROPRIETES DE CONVERGENCE

$\lambda = 0.98$



2eme Partie des simulations : METHODE DE CORRECTION en régime non stationnaire $\lambda = 0.98$

