

## Identification des paramètres d'un modèle *MA* au moyen de l'algorithme de l'anti-treillis

L. Féty<sup>1</sup>, M. Terré<sup>2</sup>

<sup>1</sup> Conservatoire National des Arts et Métiers, 292 rue St-Martin, 75141 Paris Cedex 03, France.

<sup>2</sup> Thomson CSF / CNI, 160 Boulevard de Valmy, Parc d'activité Kléber, 92704 Colombes Cedex, France.

### RÉSUMÉ

L'algorithme de l'anti-treillis est rappelé, puis ses capacités d'identification des modèles *MA* (Moving Average) sont mises en évidence. On montre notamment que la base de prédiction "arrière" utilisée dans l'anti-treillis conduit à une structure très particulière qui rappelle en partie celle d'un filtre à réponse impulsionnelle infinie. Cette particularité est mise à profit pour dériver un nouvel algorithme spécialisé dans l'identification des modèles *MA*. Les performances de cet algorithme sont ensuite évaluées en termes de convergence, de biais et de coût de calcul.

### ABSTRACT

The anti-lattice algorithm is presented, then its capabilities in identifying *MA* (Moving Average) models are highlighted. We show more specifically that the "backward" prediction base used in the anti-lattice algorithm leads to a very particular structure evoking partly that of an infinite impulse response filter. This particularity is used to derive a new algorithm specialized in the identification of *MA* models. The performances of this algorithm are then evaluated in terms of convergence, bias and calculation cost.

## 1 Introduction

Un algorithme nommé "anti-treillis" et destiné à l'adaptation des filtres à réponse impulsionnelle finie a été présenté récemment [1]. Par opposition aux algorithmes rapides en treillis [2] et *QR* [3] qui entretiennent une base de prédiction "arrière", l'anti-treillis repose, lui, sur une base de prédiction "avant". Cela lui confère une structure très particulière constituée de trois branches, dont une partie, en gras sur la figure 1, rappelle la structure d'un filtre à réponse impulsionnelle infinie. C'est cette particularité qu'on se propose de mettre à profit dans le cadre de l'identification *MA* (Moving Average). Après avoir rappelé l'algorithme, on mettra en évidence le rôle des différentes composantes de l'anti-treillis et on proposera de façon informelle un nouvel algorithme obtenu en supprimant la branche supérieure de l'anti-treillis et mieux adapté à l'identification des paramètres des modèles *MA*.

## 2 Algorithme de l'Anti-Treillis

La structure de filtrage dans l'anti-treillis comporte trois branches (figure 1). La principale est la branche intermédiaire qui contient la base de prédiction  $\{e_1(t-1) \dots e_p(t-1)\}$ . Cette base est une version orthogonalisée du signal  $\{y(t-p) \dots y(t-1)\}$  présent dans le filtre. L'orthogonalisation mise en jeu est l'inverse de celle rencontrée dans les treillis classiques. Elle est organisée du passé vers l'avenir. La première composante  $e_1(t-1)$  est constituée de

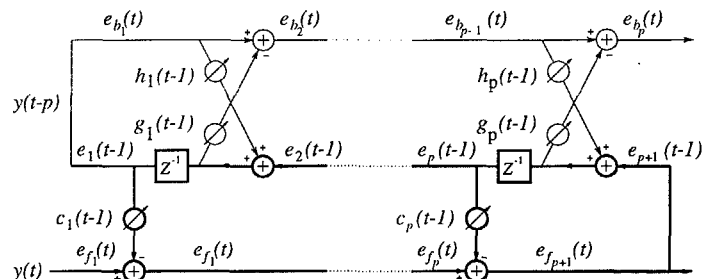


FIG. 1 - *Anti-treillis d'ordre p*

l'échantillon de signal le plus ancien. Les composantes suivantes sont les innovations des échantillons suivants par rapport à ceux plus anciens. Jusqu'à la dernière composante qui représente l'innovation de  $y(t-1)$  par rapport à  $\{y(t-p) \dots y(t-2)\}$ . L'entrée d'un nouvel échantillon de signal dans le filtre est relativement simple puisqu'il suffit de calculer son innovation par rapport à la base en place, puis de l'injecter dans celle-ci. Cette tâche est assurée par la branche inférieure. Par contre, la sortie d'un échantillon de signal du filtre est beaucoup plus délicate puisqu'il faut en quelque sorte, recorréler les composantes de la base avec celle qui sort. Cette "recorrélation" est assurée d'une façon progressive et orientée du passé vers l'avenir, par la branche supérieure de l'anti-treillis.

## 3 Identification d'un modèle *MA*

Le plus souvent, les méthodes d'identification des paramètres d'un modèle *MA* passent par l'identification d'un



$$\begin{aligned}
 e_{f_1}(t) &= y(t) \\
 \gamma_{f_1}(t) &= 1 \\
 e_{b_1}(t) &= e_1(t-1) \\
 \text{pour } m \text{ allant de } 1 \text{ à } p \\
 e_{f_{m+1}}(t) &= e_{f_m}(t) - c_m^T(t-1).e_m(t-1) \\
 \gamma_{f_{m+1}}(t) &= \gamma_{f_m}(t) - \frac{\gamma_m^T(t-1).e_m(t-1).e_m^T(t-1)}{\alpha_m(t-1)} \\
 c_m(t) &= c_m(t-1) + \frac{\gamma_m(t-1).e_m(t-1).e_{f_{m+1}}^T(t)}{\alpha_m(t-1)} \\
 e_m(t) &= e_{m+1}(t-1) + h_m(t-1).e_{b_m}(t) \\
 \text{avec } e_{p+1}(t-1) &= e_{f_{p+1}}(t) \\
 e_{b_{m+1}}(t) &= e_{b_m}(t) - g_m^T(t-1).e_m(t) \\
 \gamma_{b_m}(t) &= \gamma_{b_{m-1}}(t) - \frac{\gamma_{m-1}^2(t).e_{m-1}(t).e_{m-1}^T(t)}{\alpha_{m-1}(t)} \\
 \text{avec } \gamma_{b_1}(t) &= \gamma_1(t-1) \\
 \alpha_{b_m}(t) &= \lambda.\alpha_{b_m}(t-1) + \gamma_{b_m}(t).e_{b_m}(t).e_{b_m}^T(t) \\
 \gamma_m(t) &= \gamma_{m+1}(t-1) + \frac{\gamma_{b_m}^2(t).e_{b_m}(t).e_{b_m}^T(t)}{\alpha_{b_m}(t)} \\
 \text{avec } \gamma_{p+1}(t-1) &= \gamma_{f_{p+1}}(t) \\
 \alpha_m(t) &= \lambda.\alpha_m(t-1) + \gamma_m(t).c_m(t).c_m^T(t) \\
 h_m(t) &= h_m(t-1) + \frac{\gamma_{b_m}(t).e_{m+1}(t-1).e_{b_m}^T(t)}{\alpha_{b_m}(t)} \\
 g_m(t) &= g_m(t-1) + \frac{\gamma_m(t).e_m(t).e_{f_{m+1}}^T(t)}{\alpha_m(t)}
 \end{aligned}$$

TAB. 1 - *Algorithme de l'Anti-Treillis*

modèle *AR* long<sup>1</sup> puis par une phase d'inversion. Ce n'est pas le cas de l'anti-treillis qui permet l'identification **directe** des paramètres.

Soit un signal *MA* d'ordre *M* :

$$y(t) = x(t) + \sum_{k=1}^M \tilde{a}_k.x(t-k) \quad (1)$$

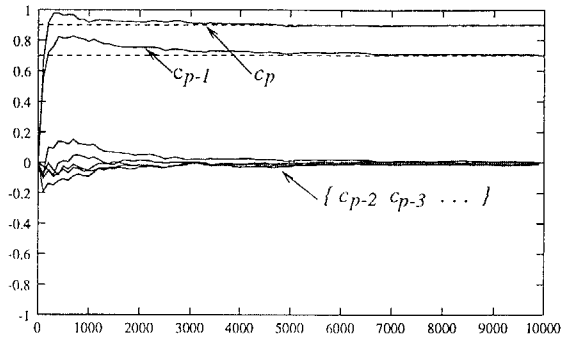
où  $x(t)$  est une excitation blanche. Lorsque qu'un anti-treillis d'ordre *p* effectue la prédiction linéaire de  $y(t)$ , l'erreur avant représente l'innovation de  $y(t)$  par rapport à son passé :  $\{y(t-p) \ y(t-p+1) \ \dots \ y(t-1)\}$ . Lorsque l'ordre *p* grandit et que le modèle est à **phase minimale**, cette erreur avant tend vers l'innovation de  $y(t)$  par rapport à :  $\{\dots \ x(t-2) \ x(t-1)\}$  et devient donc égale à  $x(t)$ . **L'inversion du modèle a donc eu lieu.**

De même, l'erreur arrière  $e_{b_m}(t)$  circulant sur la branche supérieure de l'anti-treillis représente l'innovation de  $y(t-p)$  par rapport à son futur :  $\{y(t-p+1) \ \dots \ y(t-p+i)\}$  et tend, aux ordres grands, à devenir décorrélée des valeurs récentes de  $y(t)$  et donc a fortiori des innovations récentes de  $y(t)$  qui circulent sur la branche intermédiaire. Les coefficients de prédiction  $h_m(t)$  et  $g_m(t)$  tendent alors vers 0 et la branche supérieure n'interagit plus avec la branche intermédiaire. Cette dernière contient alors, aux ordres grands, les précédentes innovations de  $y(t)$  qui demeurent les précédentes valeurs de  $x(t)$ , et les derniers coefficients de prédiction  $\{c_{p-M+1} \ \dots \ c_p\}$  tendent vers les paramètres du modèle *MA* à identifier tandis que les autres tendent vers 0. La figure 2 met en évidence ce phénomène. Les coefficients  $c_p$  et  $c_{p-1}$  ( $p = 20$ ) convergent respectivement vers

les paramètres du modèle *MA* adopté pour la simulation :

$$y(t) = x(t) + 0.9.x(t-1) + 0.7.x(t-2) \quad (2)$$

tandis que les autres convergent vers 0.

FIG. 2 - *Evolution des coefficients  $c_m$* 

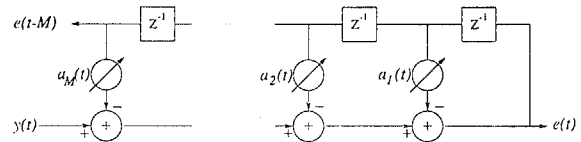
L'algorithme de l'anti-treillis représente donc un **moyen direct et adaptatif** d'estimation des paramètres d'un modèle *MA* à phase minimale.

Cependant, le fait de devoir choisir un ordre *p* grand pour l'anti-treillis ralentit la convergence et accroît le coût de calcul. Et ce, alors qu'en fait, seuls les *M* derniers coefficients  $c_m$  sont utiles. Il est légitime dans ces conditions de chercher à dériver un nouvel algorithme à partir de l'anti-treillis en amputant ce dernier de sa branche supérieure et de la majeure partie de sa branche inférieure.

## 4 Vers un modèle *MA* adaptatif

La branche inférieure de l'anti-treillis, prise séparément, est analogue à un **filtre à réponse impulsionnelle infinie**. Et c'est une contre-réaction de la branche supérieure qui dans l'anti-treillis, tronque cette réponse impulsionnelle infinie et permet d'obtenir en définitive un filtre à réponse impulsionnelle finie. C'est cette particularité que nous allons exploiter en cherchant à faire fonctionner la branche inférieure de l'anti-treillis en absence de la branche supérieure de façon à obtenir un filtre à réponse impulsionnelle infinie capable d'effectuer la prédiction linéaire adaptative d'un signal. Autrement dit, un **modèle *MA* adaptatif**.

La figure 3 présente la nouvelle structure de filtrage recherchée : seuls les derniers coefficients de la branche in-

FIG. 3 - *Modèle *MA* adaptatif*

férieure  $\{c_{p-M+1} \ \dots \ c_p\}$  ont été conservés et renommés  $\{a_M \ \dots \ a_1\}$ . L'erreur avant  $e_f(t)$  a été renommé  $e(t)$ . La modélisation *MA* adaptative consiste à minimiser cette grandeur au sens d'un critère qui peut avantageusement être celui des moindres carrés :

$$\sum_{i=0}^t \lambda^{t-i}.e_i^2(i) \text{ où } e_t(i) = \sum_{k=1}^M a_k(t).e_t(i-k) \quad (3)$$

1. La réponse impulsionnelle finie du filtre doit être aussi longue que la partie significative de celle du modèle à identifier.

Il s'agit là d'un problème d'optimisation non-linéaire auquel de nombreux chercheurs se sont intéressés mais, à ce jour, aucune réponse réellement satisfaisante n'a été apportée.

#### 4.1 La branche supérieure de l'anti-treillis

Au niveau algorithmique, la suppression de la branche supérieure peut être obtenue moyennant la suppression du rebouclage de l'erreur arrière, qui apparaît sur la figure 1 et dans la table 1 ( $e_{b_i}(t) = e_1(t - 1)$ ). De cette façon, l'erreur arrière qui circulait dans la branche supérieure est contrainte à 0 et indirectement également, les coefficients de prédiction  $h_m(t)$  et  $g_m(t)$ . La figure 4 présente une comparaison des coefficients  $c_m$  avec et sans le rebouclage. Les courbes font apparaître des écarts relativement faibles

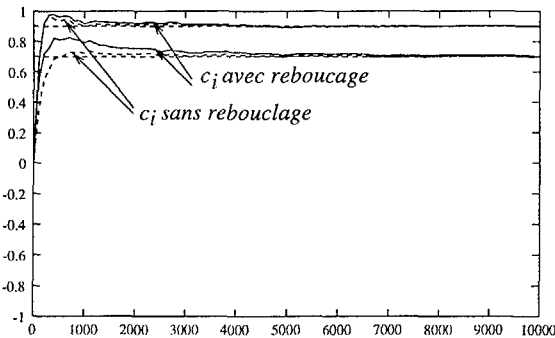


FIG. 4 -  $c_m$  avec et sans rebouclage

qui s'estompent avec la convergence. Les  $c_m$  sans rebouclage semblent converger plus rapidement vers les paramètres du modèle; phénomène qui peut s'expliquer par le fait que contraindre les coefficients  $h_m(t)$  et  $g_m(t)$  à 0 pour les ordres élevés, revient à injecter de l'information a priori. Il apparaît ainsi que la suppression de la branche supérieure est tout à fait envisageable.

#### 4.2 Troncature du filtre

Rappelons qu'à l'origine, l'identification qu'un modèle  $MA$  avec l'anti-treillis nécessitait un ordre élevé. Mais après la suppression de la branche supérieure qui était responsable de la troncature de la réponse impulsionnelle, l'ordre élevé ne paraît plus justifié. Cependant, la démarche correcte consisterait à conserver les ordres élevés du filtre tout en s'affranchissant des ordres faibles pour lesquels, les coefficients de prédiction  $c_m$  convergent vers 0. Supprimer ces coefficients revient à les contraindre à 0 et peut être considéré là encore, comme l'injection dans l'algorithme d'une information a priori. Des simulations ont confirmé que cette contrainte était plutôt bénéfique au niveau des performances.

Reste la variable  $\gamma_{f_m}(t)$  qui circule depuis les ordres inférieurs vers les ordres supérieurs. Il apparaît que la contraindre à 1 (sa valeur d'initialisation à l'ordre 1) en entrée des ordres supérieurs, provoque une chute spectaculaire de la vitesse de convergence sur certains signaux<sup>2</sup>. La figure 5 illustre ce phénomène en présentant la convergence des  $c_m$  avec et sans la contrainte sur  $\gamma_{f_m}(t)$ . Les écarts entre les

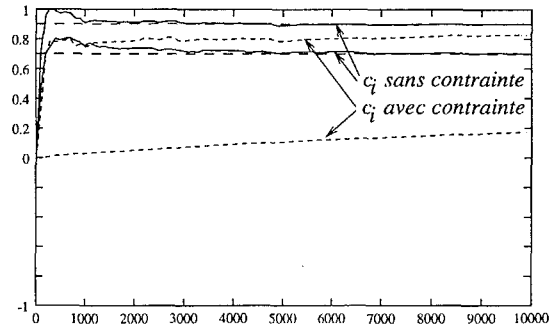


FIG. 5 -  $c_m(t)$  avec et sans contrainte sur  $\gamma_{f_m}(t)$

courbes sont importants et largement en défaveur de la contrainte sur  $\gamma_{f_m}(t)$ .

Cette variable  $\gamma_{f_m}(t)$ , quelquefois appelée pseudo-vraisemblance, est courante dans les algorithmes des moindres carrés. Elle est définie comme le rapport entre les erreurs a priori et celles a posteriori. Elle est faible en début de convergence car la correction du filtre amène une diminution importante de l'erreur de prédiction, et tend avec le temps vers 1. Elle quantifie en quelque sorte la confiance qu'il faut accorder aux prédictions délivrées par le filtre. Dans le cas d'un modèle  $MA$  adaptatif, les prédictions en question sont réinjectées dans le filtre pour servir à prédire le signal à venir. Il est alors crucial d'accompagner ces prédictions de leur fiabilité de façon qu'il en soit tenu compte. S'agissant d'un filtre récursif, il est évident en particulier que les premières prédictions seront fausses. Contraindre leur fiabilité à 1 revient à considérer le signal réinjecté dans le filtre comme un signal extérieur sans tenir compte de la récursivité du système. La figure 6 présente une évolution typique de cette variable  $\gamma_{f_m}(t)$ . La simulation a été me-

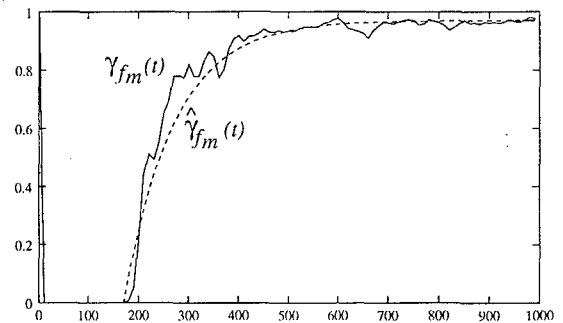


FIG. 6 - Evolution de la variable  $\gamma_{f_m}(t)$

née avec un ordre 100, ce qui explique le palier des 200 premières itérations environ, pendant lesquelles cette variable reste nulle. Après ce palier  $\gamma_{f_m}(t)$  suit une évolution exponentielle et tend vers 1. Les écarts erratiques autour de cette évolution moyenne sont dus à la réalisation particulière du signal. Et on peut penser que ce ne sont pas tant ces écarts qui sont importants, mais l'évolution exponentielle, et remplacer cette variable par un modèle :

$$\hat{\gamma}_{f_m}(t) = \lambda_\gamma \cdot \hat{\gamma}_{f_m}(t - 1) + (1 - \lambda_\gamma) \quad (4)$$

C'est la solution que nous proposons dans l'algorithme final donné dans la table 2.  $\lambda_\gamma$  est une quantité qui définit la rapidité avec laquelle  $\hat{\gamma}_{f_m}(t)$  tend vers 1. Elle doit être

2. Ces mêmes signaux font diverger l'anti-treillis lorsque la précision numérique est faible ou que l'ordre est grand.



réglée en accord avec la longueur  $l$  de la partie significative de la réponse impulsionnelle du modèle à identifier comme suit :  $\lambda_\gamma = \frac{l-1}{l}$ . L'algorithme présenté table 2 pré-

$$\begin{aligned} \alpha_m(0) &= \sigma^2, \quad c_m(0) = 0, \quad e_m(0) = 0, \quad \forall m \\ e_{f_1}(t) &= y(t) \\ \gamma_{f_1}(t) &= \lambda_\gamma \cdot \gamma_{f_1}(t-1) + (1 - \lambda_\gamma) \text{ avec } \gamma_{f_1}(t) = 0 \\ &\text{pour } m \text{ allant de } 1 \text{ à } p \\ e_{f_{m+1}}(t) &= e_{f_m}(t) - c_m^T(t-1) \cdot e_m(t-1) \\ \gamma_{f_{m+1}}(t) &= \gamma_{f_m}(t) - \frac{\gamma_m^2(t-1) \cdot e_m(t-1) \cdot e_m^T(t-1)}{\alpha_m(t-1)} \\ c_m(t) &= c_m(t-1) + \frac{\gamma_m(t-1) \cdot e_m(t-1) \cdot e_{f_{m+1}}^T(t)}{\alpha_m(t-1)} \\ e_m(t) &= e_{m+1}(t-1) \text{ avec } e_{p+1}(t-1) = e_{fp+1}(t) \\ \gamma_m(t) &= \gamma_{m+1}(t-1) \text{ avec } \gamma_{p+1}(t-1) = \gamma_{fp+1}(t) \\ \alpha_m(t) &= \lambda \cdot \alpha_m(t-1) + \gamma_m(t) \cdot e_m(t) \cdot e_m^T(t) \end{aligned}$$

TAB. 2 - *Modèle MA adaptatif*

sente un coût de calcul en  $O(p)$  extrêmement réduit. Il est comparable et même inférieur à ceux engendrés par certains algorithmes rapides d'adaptation de filtres à réponse impulsionnelle finie.

## 5 Performances

Parmi les algorithmes d'adaptation des filtres à réponse impulsionnelle infinie [4], ce sont ceux dits "à erreur de sortie", par opposition à ceux dits "à équation d'erreur", qui sont les plus satisfaisants. L'algorithme proposé ici se classe dans cette famille en ce sens qu'il optimise directement l'erreur en sortie du filtre récursif sans passer par l'emploi d'un signal désiré indisponible dans le cas du modèle MA adaptatif. Dans cette famille, le meilleur algorithme en terme de performances sur une fenêtre temporelle donnée est celui qui consiste à optimiser jusqu'à la convergence, le critère des moindres carrés :

$$J(n) = \sum_{i=0}^n \lambda^{t-i} \cdot \left( y(t) - \sum_{k=1}^M a_k(n) \cdot e_{n-1}(i-k) \right)^2 \quad (5)$$

où  $e_{n-1}(i)$  est l'erreur obtenue à l'instant  $i$  avec le filtre  $a(n-1)$ . Cet algorithme présente évidemment un coût de calcul très important puisqu'il nécessite à chaque itération, de refiltrer entièrement le signal. De plus, il est à noter que des précautions doivent être prise pour garantir la stabilité du filtre obtenu à chaque itération.

La confrontation, grâce à des simulations, de cet algorithme avec celui présenté auparavant a permis de mettre en évidence l'**extrême similitude des résultats**. Ce n'est qu'en tout début de convergence qu'apparaissent de faibles écarts qui peuvent probablement être mis sur le compte des conditions initiales. L'algorithme présente donc de très bonnes facultés de convergence puisqu'il est capable d'exploiter un signal en une passe aussi bien qu'un algorithme extrêmement coûteux procédant par itération. L'algorithme proposé apparaît donc très intéressant en termes de performances. Cependant, les algorithmes dits "à erreur de sortie" ne sont pas exempts de problèmes. Ainsi, il ne suffit pas que le modèle soit à phase minimale pour qu'il puisse être identifié correctement. Il faut en plus

qu'une condition sur la positivité du modèle [4] soit vérifiée :  $Re(\tilde{A}(f)) > 0, \forall f$ .  $\tilde{A}(f)$  représente la réponse en fréquence du modèle à identifier (1). La zone de positivité est représentée sur la figure 7 dans le cas d'un modèle d'ordre 2. Elle apparaît à l'intérieur du triangle de stabilité. Les segments qui apparaissent dans le triangle repré-

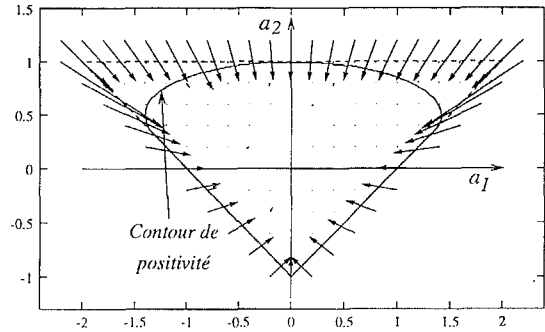


FIG. 7 - *Biais d'identification*

sentent les directions et les amplitudes des biais observés lors de l'identification par l'algorithme des paramètres du modèle :  $(a_1 - \tilde{a}_1)$  et  $(a_2 - \tilde{a}_2)$ . Les biais les plus importants sont en plus orientés. Ces derniers surviennent essentiellement lorsque les paramètres du modèle sont à l'extérieur du triangle de stabilité. L'identification parfaite est alors impossible et l'algorithme converge vers une solution située à l'intérieur du triangle de stabilité. Cette observation met en évidence la capacité de l'algorithme à fournir une solution stable. Les biais observés lorsque le modèle est stable sont extrêmement faibles. Ce n'est que dans les coins supérieurs du triangle, lorsque la condition de positivité n'est plus remplie, qu'ils grandissent quelque peu.

## 6 Conclusion

Au terme de cette étude, il apparaît que l'algorithme de l'anti-treillis permet d'identifier de façon directe et adaptative un modèle MA à phase minimale. Il faut cependant pour cela adopter un ordre grand, ce qui engendre un coût de calcul important. L'alternative proposée consiste en une version tronquée du filtre mis en oeuvre dans l'anti-treillis. Le nouvel algorithme ainsi obtenu présente d'excellentes performances aussi bien en termes de convergence que de stabilité ainsi qu'un coût de calcul extrêmement réduit.

## Références

- [1] L. Féty, M. Terré and M. Bellanger, *Fast Anti-Lattice Algorithm*, Proc EUSIPCO, Edimbourg Sept 1994 1257-1260.
- [2] M. Morf, A. Vieira, and D.T. Lee, *Ladder forms for identification and speech processing*, Proc IEEE Conf. Decision Contr. Dec 1977 1074-1078.
- [3] J.G. McWhirter, *Recursive least-squares minimization using systolic array*, SPIE vol 431, Real Time Signal Processing VI 1983.
- [4] J. J. Shynk, *Adaptive IIR Filtering*, IEEE ASSP Magazine April 1989 4-21.