Sur les Points Stationnaires des Algorithmes en Filtrage Récursif Adaptatif

Mamadou MBOUP
Université René Descartes
UFR Mathématiques et Informatique
45, rue des Saints Pères
75270 Paris cedex 06.

Phillip A. REGALIA
Institut National des Télécommunications
Département Electronique et Communications
9, rue Charles Fourier
91011 Evry cedex.

RÉSUMÉ

Des conditions nécessaires et suffisantes sur la fonction d'erreur correspondant aux points stationnaires de trois algorithmes de filtrage adaptatif RII en sous modélisation sont données. Ces conditions sont obtenues dans l'espace des fonctions de transfert, grâce à une application judicieuse du théorème de Beurling. Notre approche permet aussi de montrer des connexions intimes avec la théorie des opérateurs de Hankel qui joue un rôle essentiel dans des problèmes de réduction de modèle.

ABSTRACT

Necessary and sufficient conditions on the error function corresponding to the stationary points of three common adaptive IIR filtering algorithms are given. These constrains are obtained in the transfert functions space, by the judicious application of Beurling's theorem. Our approach also allows to show some deep connections with the Hankel operator theory which is essential for many model reduction problems.

1 Introduction

Le filtrage récursif adaptatif est considéré dans de nombreuses applications où le filtre optimal nécessite une réponse impulsionnelle très longue, comme par exemple en annulation d'écho acoustique. Les avantages de modélisation qu'offrent ces filtres sont cependant ternis par l'absence de descriptions analytiques des points stationnaires des algorithmes adaptatifs en régime sous-modélisé.

Ce papier résume les résultats présentés dans [1], donnant des conditions nécessaires et suffisantes sur la fonction d'erreur correspondant aux points stationnaires des algorithmes du gradient récursif, Steiglitz-McBride et SHARF, en sous modélisation.

On cherche à identifier une fonction de transfert inconnue, causale et stable, $H(z) = \sum_{k=0}^{\infty} h_k\, z^{-k}$ par un modèle ajustable $G(z)$ causal et stable, de degré M, de la forme

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \ldots + b_M z^{-M}}{1 + a_1 z^{-1} + \ldots + a_M z^{-M}}.$$

Si $\Theta(n) = [b_0 \cdots b_M a_1 \cdots a_M]^t$ dénote le vecteur des paramètres de G à l'instant n, on peut lui associer un algorithme d'adaptation selon

$$\Theta(n+1) = \Theta(n) + \mu\, e(n) \Psi(n)$$

où $e(n)$ est un signal d'erreur qui dépend de $H(z) - G(z)$, et $\Psi(n)$ est un vecteur de régression filtré. Les formes exactes de $e(n)$ et $\Psi(n)$ varient selon l'algorithme adaptatif utilisé [2].

Pour un algorithme adaptatif donné, les points stationnaires sont obtenus en trouvant les valeurs de Θ qui annulent le terme de correction :

$$\mathrm{E}[e(n)\, \Psi(n)] = 0.$$

Cet article a été préparé au L.S.S., CNRS - ESE

L'essentiel des résultats de convergence existants dans la littérature sur l'identification [2] suppose que l'ordre du filtre optimal est connu, ce qui est, bien entendu, qu'un idéal [3]. Il est donc plus réaliste de considérer qu'une modélisation exacte n'est pas réalisable ; le filtre adaptatif fonctionne alors en régime sous-modélisé. Dans ces conditions, le mieux que l'on puisse espérer et que le(s) point(s) stationnaire(s) corresponde(ent) à ce que $G(z)$ soit une bonne approximation de $H(z)$.

Le problème des points stationnaires en sous modélisation a toujours été posé dans l'espace des paramètres. Or, ceci conduit à des équations non linéaires (en les paramètres $\{a_k\}$) qui n'admettent pas de solutions analytiques, d'où, l'absence de résultats analytiques malgré l'importance de la question. Notre approche repose le problème sous sa forme la plus naturelle *i.e.* dans l'espace des fonctions de transfert. Cette approche permet, grâce à une judicieuse application du théorème de Beurling, de trouver des conditions nécessaires et suffisantes pour qu'une fonction de transfert corresponde à un point stationnaire. Elle permet aussi de montrer des connexions intimes avec la théorie des opérateurs de Hankel, qui joue un rôle essentiel dans des problèmes de réduction de modèle.

2 Préliminaires

On considère des fonctions de $L_2(-\pi, \pi)$, $S(z) = \sum_{k=-\infty}^{\infty} s_k z^{-k}$, *i.e.* que

$$\langle S(z), S(z) \rangle \triangleq \frac{1}{j2\pi} \oint_C S(z) S(z^{-1}) \frac{dz}{z} < \infty.$$

Pour simplifier, on suppose que l'entrée $u(n)$ du filtre $G(z)$ est un bruit blanc ; le cas d'une entrée colorée est traité dans [1], en utilisant la même méthode. On a le modèle d'état suivant :

$$\left[\begin{array}{c} \mathbf{x}(n+1) \\ y(n) \end{array}\right] = \left[\begin{array}{cc} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^t & d \end{array}\right]\left[\begin{array}{c} \mathbf{x}(n) \\ u(n) \end{array}\right], \qquad (1)$$

avec $G(z) = \mathbf{c}^t(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d$, où $\mathbf{A}$ a toutes ses valeurs propres à l'intérieur du cercle unité, $(\mathbf{A}, \mathbf{b})$ est complètement commandable et $(\mathbf{A}, \mathbf{c})$ complètement observable. On s'intéresse, ici, à la caractérisation de l'ensemble $\mathcal{N}$ des fonctions analytiques appartenant à l'orthogonal de l'opérateur de commandabilité $\mathcal{C}(z)$ associé à (1) :

$$\mathcal{C}(z) = z(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{b} z^{-k}, \ |z| \geq 1. \qquad (2)$$

On vérifie aisément, pour toute fonction analytique $S(z)$, la relation : $\langle \mathcal{C}(z), z^{-1}S(z)\rangle = \mathbf{A}\langle \mathcal{C}(z), S(z)\rangle$. Ainsi, si $S(z) \in \mathcal{N}$ *i.e.* $\langle \mathcal{C}(z), S(z)\rangle = \mathbf{0}$ alors on a $z^{-1}S(z) \in \mathcal{N}$. $\mathcal{N}$ définit donc un sous espace de Hardy invariant par décalage [4]. Le théorème de Beurling [5] montre que tout sous espace de Hardy invariant par décalage admet une représentation unique $V(z)\mathcal{H}_2$, avec $V(z)$ une fonction causale passe-tout ($V(z)V(z^{-1}) \equiv 1$) et $\mathcal{H}_2$, l'ensemble des fonctions causales à carré sommable. Pour notre problème, on montre [1] que

$$V(z) = \frac{z^{-M}A(z^{-1})}{A(z)} = \frac{\hat{A}(z)}{A(z)} \qquad (3)$$

est la fonction passe-tout propre à $\mathcal{C}(z)$; elle est construite à partir des pôles de $G(z)$. En résumé, nous avons :

Proposition 1 *$S(z) \in \mathcal{N} \iff S(z) = V(z)R(z)$ avec $R(z) \in \mathcal{H}_2$.*

3 Surface d'erreur réduite

Dans cette section, nous traitons la première partie du problème des points stationnaires, à savoir, le problème de la surface d'erreur réduite. Il s'agit de trouver ($A(z)$ étant fixé) le numérateur $B(z)$ qui minimise la norme $||H(z) - G(z)||_2^2 = \langle H(z) - G(z), H(z) - G(z)\rangle$. C'est un problème linéaire quadratique dont la solution est obtenue en annulant les dérivées. On obtient le système d'équations suivant :

$$\langle \mathcal{C}(z), H(z) - G(z)\rangle = \mathbf{0}_M \qquad (4)$$

$$\langle z^{-M}/A(z), H(z) - G(z)\rangle = 0 \qquad (5)$$

Proposition 2 *Soit $G(z) = B(z)/A(z)$ une fonction causale. Pour $A(z)$ fixé, la norme $||H(z) - G(z)||_2$ est minimisée par rapport à $B(z)$ ssi*

$$H(z) - G(z) = V(z)\sum_{k=1}^{\infty} r_k z^{-k} = V(z)R(z) \qquad (6)$$

où $V(z)$ est la fonction passe-tout définie par (3).

Le fait que la fonction d'erreur $H(z) - G(z)$ soit causalement divisible par $V(z)$ découle du système d'équations (4) et de la proposition 1 tandis qu'une manipulation simple de (5) montre que $R(z)$ est strictement causale. En multipliant l'équation (6) par $V(z^{-1})$ et en prenant les projections causale et anti-causale, on obtient :

$$[V(z^{-1})H(z)]_+ = \sum_{k=1}^{\infty} r_k z^{-k} = R(z), \qquad (7)$$

$$[V(z^{-1})H(z)]_- = V(z^{-1})G(z), \qquad (8)$$

où $[\cdot]_+$ et $[\cdot]_-$ désignent respectivement les projections causale et anti-causale. En écrivant $V(z) = \sum_{k=0}^{\infty} v_k z^{-k}$, on montre

Corollaire 1 *La relation $R(z) = [V(z^{-1})H(z)]_+$ est équivalente à $\mathbf{r} = \mathbf{\Gamma}_H \mathbf{v}$, avec $\mathbf{\Gamma}_H \triangleq \mathrm{Hankel}[h_1, h_2, \ldots]$ et où $\mathbf{r}$ et $\mathbf{v}$ désignent respectivement les vecteurs formés par les coefficients de $R(z)$ et $V(z)$. Ainsi, la surface d'erreur réduite devient*

$$\min_{B(z)} ||H(z) - G(z)||_2^2 = \mathbf{v}^t \mathbf{\Gamma}_H^2 \mathbf{v}.$$

Remarque : Noter l'apparence explicite de l'opérateur de Hankel $\mathbf{\Gamma}_H$.

4 Points stationnaires

4.1 Gradient Récursif

Les points stationnaires sont obtenus en annulant le gradient, ce qui, par rapport à $B(z)$, conduit à l'équation (6). Par rapport à $A(z)$, on obtient en utilisant (6) :

$$\begin{aligned} \mathbf{0}_M &= \langle z^{-1}\mathcal{C}(z)G(z), H(z) - G(z)\rangle \\ &= \langle \mathcal{C}(z), G(z^{-1})V(z)[zR(z)]\rangle \\ &= \langle \mathcal{C}(z), [z^{-M})B(z^{-1})/A(z)][zR(z)]\rangle. \end{aligned} \qquad (9)$$

En vertu de la proposition 1, $[z^{-M})B(z^{-1})/A(z)][zR(z)]$ doit être causalement divisible par $V(z)$. Ceci implique que c'est $R(z)$ qui contient un facteur $V(z)$ car autrement, il y aurait une compensation pôle-zéro pour $G(z)$ et il en resulterait un point en selle [6]. On donne alors :

Proposition 3 *Une fonction rationnelle causale $G(z)$, de degré M correspond à un point stationnaire de l'algorithme adaptatif du gradient récursif ssi*

$$H(z) - G(z) = V^2(z)R'(z), \qquad (10)$$

où $V(z)$ est construite comme en (3) et $R'(z)$ une fonction strictement causale.

On a d'après (10), $||H(z) - G(z)||_2 = ||R'(z)||_2$. Une question ouverte est de savoir si la structure analytique de $R'(z)$ pourrait permettre de distinguer les points minima de l'algorithme du gradient récursif des points en selle.
Si $H(z) = D(z)/C(z)$ est rationnelle de degré M_H, alors on montre [1] :

Proposition 4 *Si $\deg G(z) = M = M_H$, alors $G(z) = H(z)$ est le seul point critique de l'algorithme du gradient récursif.*

Revenons au cas sous-modélisé. La contrainte sur $R(z)$ s'écrit :

$$R(z) = V(z)R'(z) \qquad (11)$$

dans le domaine fréquentiel et dans le domaine temporel :

$$\mathbf{r} = V(\mathcal{T})\mathbf{r}' \qquad (12)$$

où $V(\mathcal{T}) = \sum_{k=0}^{\infty} v_k \mathcal{T}^k$ est un opérateur de Toeplitz triangulaire et $\mathcal{T}$ est l'opérateur de décalage, avec des uns sur la sous diagonale et des zéros partout. Avec cette notation, l'équation (7) devient, en utilisant la contrainte sur $R(z)$ et le corollaire 1

$$\mathbf{\Gamma}_H \mathbf{v} = V(\mathcal{T})\mathbf{r}', \qquad (13)$$

dans le domaine temporel et dans le domaine frquentiel, $[V(z^{-1})H(z)]_+ = V(z)R'(z)$. Cette dernière forme montre que $V(z)$ se comporte comme une fonction propre de l'opérateur de Hankel $\mathbf{\Gamma}_H[V(z)] \triangleq [V(z^{-1})H(z)]_+$.

4.2 Steiglitz-McBride

Pour une entrée blanche, les points stationnaires de l'algorithme Steiglitz-McBride sont les solutions [7] du système d'équations (4 - 5) (pour la minimisation par rapport à $B(z)$) et du système (concernant $A(z)$) :

$$\langle z^{-1}\mathcal{C}(z)H(z), H(z) - G(z)\rangle = \mathbf{0}_M. \qquad (14)$$

Ainsi, pour cet algorithme aussi, les points stationnaires sont situés sur la surface d'erreur réduite, bien que ne correspondant pas, en général, à un minimum de cette surface. En tenant compte de cette propriété (équation (6)), le système (14) devient :

$$\begin{aligned} \mathbf{0}_M = & \langle z^{-1}\mathcal{C}(z)H(z), V(z)R(z)\rangle \\ = & \langle z^{-1}\mathcal{C}(z)V(z^{-1}), H(z^{-1})R(z)\rangle. \end{aligned} \qquad (15)$$

On montre [1] que $\mathcal{C}'(z) \triangleq z^{-1}\mathcal{C}(z)V(z^{-1})$ définit un opérateur de commandabilité anti-causale, dont la fonction passe-tout propre est $V(z^{-1})$. Donc, en utilisant la projection anti-causale de la proposition 1, on obtient $[H(z^{-1})R(z)]_- = V(z^{-1})L(z^{-1})$ où $L(z^{-1})$ est une fonction anti-causale. Un changement de variable (z^{-1} en z) montre que la projection causale de $H(z)R(z^{-1})$ doit contenir un facteur $V(z)$. Si $P(z) = \sum_{k=0}^{\infty} p_k z^{-k}$ désigne cette projection, alors on montre en utilisant la relation $\mathbf{r} = \mathbf{\Gamma}_H\mathbf{v}$

$$\mathbf{p} = [p_0, p_1, \ldots]^t = \mathbf{\Gamma}_H\mathbf{r} = \mathbf{\Gamma}_H^2\mathbf{v}. \qquad (16)$$

La contrainte $P(z) = V(z)P'(z)$, où $P'(z)$ est une fonction causale, combinée avec (16) donne :

Proposition 5 *$V(z)$ correspond à un point stationnaire de l'algorithme Steiglitz-McBride ssi*

$$\mathbf{\Gamma}_H^2\mathbf{v} = V(\mathcal{T})\mathbf{p}', \ \mathbf{p}' = [p'_0, p'_1, \ldots]^t. \qquad (17)$$

L'erreur résultante est $\|H(z) - G(z)\|_2 = \sqrt{p'_0}$.

L'équation (17) ressemble beaucoup à l'équation (13) obtenue pour le gradient récursif. La seule différence est que la matrice de Hankel $\mathbf{\Gamma}_H$ est remplacée par $\mathbf{\Gamma}_H^2$. Ceci donne le Lemme suivant dont une preuve est donnée dans [1] :

Lemme 1 *La fonction $P'(z) - \frac{1}{2}p'_0$ est réelle positive :*

$$\mathrm{Re}[P'(e^{j\omega}) - \frac{1}{2}p'_0] \geq 0, \ \forall\omega.$$

Si $H(z)$ est une fonction rationnelle, alors on peut montrer que les pôles de $P'(z)$ coïncident avec ceux $H(z)$ et que les zéros de $P'(z)$ contiennent, en facteur, les pôles de $G(z)$. D'après cette propriété, on peut conclure que pour toute fonction $G(z)$ correspondant à un point stationnaire de l'algorithme de Steiglitz-McBride, ses pôles sont localisés dans un domaine réel positif des pôles de $H(z)$.

4.3 SHARF

Avec un bruit blanc en entrée, les points stationnaires de l'algorithme SHARF vérifient [7] :

$$\begin{aligned} \langle z^{-i}, F(z)[H(z) - G(z)]\rangle &= 0, \quad i = 0, \ldots, M \\ \langle z^{-1}\mathcal{C}(z)B(z), F(z)[H(z) - G(z)]\rangle &= \mathbf{0}. \end{aligned} \qquad (18)$$

où $F(z)$ est un filtre polynômial, à minimum de phase, choisi pour satisfaire une condition de réalité strictement positive [8]. Les $M+1$ premières équations de ce système impliquent l'existence d'un zéro à l'infini, d'ordre $M+1$, pour la fonction $F(z)[H(z) - G(z)]$. Or le polynôme $F(z)$ est à minimum de phase, le zéro d'ordre $M+1$ provient de $H(z) - G(z)$. Autrement dit, $H(z) - G(z)$ s'exprime par :

$$H(z) - G(z) = \sum_{k=M+1}^{\infty} r_k z^{-k}, \qquad (19)$$

pour une certaine séquence r_k. Les $M+1$ premiers coefficients des réponses impulsionnelles de $G(z)$ et de $H(z)$ coïncident. Il apparaît donc, que $G(z)$ est une approximation de *Padé* d'ordre $M+1$, de $H(z)$.
La deuxième partie du système d'équations (18) se réarrange comme

$$\langle \mathcal{C}(z), zB(z^{-1})F(z)[H(z) - G(z)]\rangle = \mathbf{0}_M. \qquad (20)$$

On montre [1], [3] que $zB(z^{-1})F(z)[H(z) - G(z)]$ est une fonction causale, d'où, d'après la proposition 1, elle contient un facteur $V(z)$. Les zéros de $V(z)$ ne peuvent pas diviser $B(z^{-1})$ car il en résulterait une compensation pôle-zéro, ce qui ne peut pas correspondre à un point stationnaire. De même, $F(z)$ étant à minimum de phase, ne peut contenir un facteur passe-tout. Finalement, on obtient

Proposition 6 (Points stationnaires de SHARF) *Soit* $\deg G(z) = M$. *Le système (18) est satisfait ssi*

$$H(z) - G(z) = V(z) \sum_{k=M+1}^{\infty} r_k z^{-k} = V(z)R(z) \qquad (21)$$

Cette proposition montre que les points stationnaires de l'algorithme SHARF sont aussi localisés sur la surface d'erreur réduite comme pour les deux algorithmes précédents. Par contre, SHARF est le seul pour lequel, la contrainte placée sur $R(z)$ ne dépend pas de $V(z)$. Trouver les points stationnaires de SHARF revient à trouver une fonction passe-tout $V(z)$ satisfaisant :

$$\langle H(z), z^{-k}V(z)\rangle = 0, \ k = 1, \ldots, M, \qquad (22)$$

dans le domaine fréquentiel, ou bien

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & h_4 & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ h_M & h_{M+1} & h_{M+2} & \cdots \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \end{bmatrix} \qquad (23)$$

dans le domaine temporel.

5 Lien avec les opérateurs de Hankel

On reprend l'algorithme du gradient récursif. En utilisant la factorisation [4] de toute fonction non nulle de L_1, $S(z)$ selon

$$S(z) = U(z)S_0(z), \qquad (24)$$

où $U(z)$ est une fonction passe-tout et $S_0(z)$ une fonction à minimum de phase, on peut réécrire l'équation (11) par

$$R(z) = V(z)U(z)R_0(z) \qquad (25)$$

avec $U(z)$, une fonction passe-tout et $R_0(z)$, une fonction à minimum de phase. Avec cette factorisation, le vecteur r de l'équation (12) devient $\mathbf{r} = \mathbf{v}(\mathcal{T})\mathbf{u}(\mathcal{T})\mathbf{r}_0$ et l'équation (13) s'exprime par :

$$\mathbf{\Gamma}_H[\mathbf{r}_0(\mathcal{T})]^{-1}\mathbf{r}_0(\mathcal{T})\mathbf{v} = \mathbf{v}(\mathcal{T})\mathbf{u}(\mathcal{T})\mathbf{r}_0, \qquad (26)$$

où l'inversibilité de $\mathbf{r}_0(\mathcal{T})$ est assurée par le fait que $R_0(z)$ soit à minimum de phase. En posant

$$\mathbf{\Gamma} \triangleq \mathbf{\Gamma}_H[\mathbf{r}_0(\mathcal{T})]^{-1}\,; \quad \boldsymbol{\zeta} \triangleq \mathbf{r}_0(\mathcal{T})\mathbf{v}\,; \quad \boldsymbol{\eta} \triangleq \mathbf{v}(\mathcal{T})\mathbf{u}(\mathcal{T})\mathbf{r}_0, \qquad (27)$$

on obtient une matrice de Hankel : $\mathbf{\Gamma}$ et une paire de *Schmidt* associée : $\boldsymbol{\zeta}$ et $\boldsymbol{\eta}$. En effet, étant le résultat d'un produit Hankel–Toeplitz, $\mathbf{\Gamma}$ est une matrice de Hankel. Par ailleurs, à partir des définitions de $\mathbf{\Gamma}$, $\boldsymbol{\zeta}$ et $\boldsymbol{\eta}$, on lit directement sur l'équation (26) : $\mathbf{\Gamma}\boldsymbol{\zeta} = \boldsymbol{\eta}$. Enfin, on vérifie [3] $\mathbf{\Gamma}\boldsymbol{\eta} = \boldsymbol{\zeta}$. La paire $\{\boldsymbol{\zeta}, \boldsymbol{\eta}\}$ est donc une paire de Schmidt de la matrice de Hankel $\mathbf{\Gamma}$, associée à la valeur singulière $\sigma = 1$. Dans la suite, les valeurs singulières, σ_k, de $\mathbf{\Gamma}$ sont ordonnées par convention selon : $\sigma_1 = \|\mathbf{\Gamma}\|_H \geq \sigma_2 \geq \cdots \geq \sigma_k \geq \cdots \geq 0$. Considérons maintenant le problème de Hankel suivant :

"Etant donnée une fonction rationnelle causale et stable $H'(z)$, dont la matrice de Hankel correpondante est la matrice $\mathbf{\Gamma}$ donnée ci-dessus, trouver une fonction rationnelle causale et stable, $G'(z)$ de degré κ, de telle sorte la norme $\|H'(z) - G'(z)\|_H$ soit minimale."

Ce problème est largement évoqué dans la littérature, sous différentes formes [9]. Il admet une solution analytique unique, donnée par *Adamjan et al.* par :

$$G'(z) = H'(z) - \sigma_\kappa \left[\frac{\eta_\kappa^-(z)}{\zeta_\kappa^+(z)}\right]_+, \qquad (28)$$

où σ_κ est la $\kappa^{\text{ième}}$ valeur singulière de $\mathbf{\Gamma}$, associée à la paire de Schmidt $\{\boldsymbol{\zeta}_\kappa, \boldsymbol{\eta}_\kappa\}$. Cette paire détermine la fonction passe–tout dont la projection causale apparaît dans la solution optimale (28), où :

$$\zeta_\kappa^+(z) = \sum_{j=0}^{\infty} \zeta_{\kappa_j} z^{-j} \quad \text{et} \quad \eta_\kappa^-(z) = z\sum_{i=0}^{\infty} \eta_{\kappa_i} z^i,$$

les coefficients ζ_{κ_j}, η_{κ_i} étant les composantes respectives des vecteurs $\boldsymbol{\zeta}_\kappa$ et $\boldsymbol{\eta}_\kappa$. La norme minimale de l'erreur d'approximation est alors donnée par σ_κ [9]. On déduit de la régularité de $\mathbf{r}_0(\mathcal{T})$ que les matrices $\mathbf{\Gamma}$ et $\mathbf{\Gamma}_H$ sont de même rang. Les fonctions de transfert $H(z)$ et $H'(z)$ sont alors de même degré de *McMillan N*. La correspondance entre ces deux fonctions est encore plus explicite au travers du lemme suivant démontré dans [3].

Lemme 2 *Soient les matrices de Hankel $\mathbf{\Gamma}_H$ et $\mathbf{\Gamma}$ definies ci-dessus. Les fonctions de transfert associées $H(z)$ et $H'(z)$ ont exactement les mêmes pôles.*

Pour montrer que la fonction rationnelle $G(z)$ qui optimise (localement) notre problème d'approximation est aussi la solution du problème utilisant le critère de Hankel, il suffit de montrer que le rang de la valeur singulière $\sigma = 1$, associée à la paire de Schmidt $\{\boldsymbol{\zeta}, \boldsymbol{\eta}\}$, est égal à M, le degré de $G(z)$. Pour cela, nous utiliserons le théorème suivant, démontré par *Adamjan et al.* [9].

Théorème 1 *Soit $\mathbf{\Gamma}$ une matrice de Hankel bornée, de dimension infinie et soit σ une valeur singulière de $\mathbf{\Gamma}$. Alors, les vecteurs $\boldsymbol{\zeta}$ et $\boldsymbol{\eta}$ formant une paire de Schmidt de $\mathbf{\Gamma}$ associée à σ peuvent être factorisés sous les formes*

$$\boldsymbol{\zeta} = \mathbf{u}'(\mathcal{T})\mathbf{v}(\mathcal{T})\mathbf{r}_0 \qquad \boldsymbol{\eta} = \mathbf{u}(\mathcal{T})\mathbf{v}(\mathcal{T})\mathbf{r}_0, \qquad (29)$$

où $R_0(z)$ est une fonction à minimum de phase, $U(z)$ et $U'(z)$ sont des fonctions passe–tout dont la somme des degrés est strictement inférieure à l'ordre de multiplicité de la valeur propre σ^2 de $\mathbf{\Gamma}^2$ et où $V(z)$ est aussi une fonction passe–tout, indépendante du choix de la paire de Schmidt associée à σ. De plus, on a :

(i) si $\sigma = \sigma_1 = \|\mathbf{\Gamma}\|_H$ alors $V(z) \equiv 1$ ($\mathbf{v}(\mathcal{T}) \equiv \mathbf{I}$)
(ii) si $\sigma_1 \geq \sigma_{\kappa-1} > \sigma = \sigma_\kappa \geq \sigma_{\kappa+1}$ alors le degré de $V(z)$ est égal κ.

La paire de Schmidt définie en (27) est déjà factorisée sous la forme (29) donnée dans ce théorème, avec $U'(z) \equiv 1$. L'application directe du théorème 1 permet de conclure que le rang de la valeur singulière $\sigma = 1$ est égal à M, c'est à dire, le degré de $V(z)$ ou encore, le degré de $G(z)$. En d'autres termes, la fonction $G'(z) = G(z)$ est la solution optimale (unique) du problème de Hankel avec $\kappa = M$. Notons que si la valeur singulière $\sigma = 1$ est multiple d'ordre m (degré de $U(z)$ égal $m-1$), la solution du problème de Hankel reste unique et est toujours donnée par l'équation (28). La fonction passe–tout apparaissant dans cette équation est indépendante de la paire de Schmidt choisie, correspondant à σ [9]. Réciproquement, il est évident que si $G(z)$ est la solution optimale du problème de Hankel, alors $G(z)$ est aussi optimale pour notre problème d'approximation dès lors que la matrice $\mathbf{\Gamma}$ est construite à partir de cette même fonction.

Cette démarche s'applique aussi aux algorithmes adaptatifs Steiglitz-McBride et SHARF [3].

6 Conclusion

Nos résultats donnent pour la première fois des conditions nécessaires et suffisantes sur la fonction d'erreur correspondant aux points stationnaires des algorithmes du gradient récursif, de Steiglitz-McBride et de SHARF. La connexion établie avec la théorie des opérateurs de Hankel devrait permettre à terme, de pouvoir borner, pour chacun des algorithmes adaptatif étudiés, la fonction d'erreur en termes des valeurs singulières de la matrice de Hankel $\mathbf{\Gamma}_H$.

Références

[1] P. A. Regalia, M. Mboup, Rapport interne, L.S.S 1993.

[2] L. Ljung, T. Söderström, MIT Press, Cambridge MA, 1983.

[3] M. Mboup, Thèse Doctorat, Univ. Paris Sud, 92.

[4] K. Hoffman, Prentice–Hill, Inc., Englewood Cliffs, N. J., 1962.

[5] A. Beurling, *Acta Math.*, vol. 81, 1949.

[6] N. Nayeri, W.K. Jenkins, *IEEE Trans. on Circ. and Syst.*, vol. 36, n° 4, pp. 485-496, Apr. 1989.

[7] H. Fan, *IEEE Trans. Info. Theor*, vol. 34, July 88.

[8] M. G. Larimore, J. R. Treichler, C. R. Johnson, *IEEE Trans. on ASSP*, vol. 28, Aug. 80.

[9] V. M. Adamjan, D. Z. Arov, M. G. Kreĭn, *Math. U.S.S.R. Sbornik*, vol. 15, n° 1, pp. 31-73, 1971.