



NON STATIONNARITES MARKOVIENNES RAPIDES ET FILTRAGE ADAPTATIF

M. TURKI (1), O. MACCHI (2), M. JAIDANE-SAIDANE (1)

(1) Laboratoire des Systèmes de Télécommunications, ENIT, BP37, Tunis, Tunisie
 (2) Laboratoire des Signaux et Systèmes, CNRS-ESE, Plateau du Moulon, Gif sur Yvette, France

RESUME

Les algorithmes d'identification adaptative dans un contexte non stationnaire de type markovien sont étudiés ici à partir d'un formalisme mathématique qui stationnarise le problème. La décroissance exponentielle de la composante transitoire est montrée pour différents algorithmes. L'étude de la capacité de poursuite en régime permanent permet de comparer les non stationnarités de type promenade aléatoire et de type markovien. Les non stationnarités rapides non traitées dans la littérature, reçoivent ici une solution.

I- PRESENTATION DU PROBLEME

Nous étudions les performances de l'identification adaptative dans un contexte non stationnaire. Le signal de référence $y(k)$, résulte du filtrage d'une suite stationnaire $x(k)$, par un filtre linéaire de longueur N , variable et non mesurable, de vecteur paramètre $F(k)$:

$$y(k) = s(k) + n(k), \quad s(k) = F(k)^T X(k), \quad (I-1)$$

où $n(k)$ est un bruit de sortie non mesurable stationnaire, suite i.i.d., centrée, de puissance P_n , indépendante de la suite $x(k)$, et où $X(k) = (x(k), \dots, x(k - N + 1))^T$ est le vecteur observation ; $s(k)$ est appelé "signal". Il est inconnu.

Le vecteur $F(k)$ est supposé évoluer au cours du temps selon un modèle markovien du premier ordre défini par :

$$F(k) = (I - \underline{\eta})F(k - 1) + \Omega(k), \quad k \geq 0, \quad (I-2)$$

où $\underline{\eta} = \text{diag}(1 - \eta_i)$ ($0 < \eta_i < 1$) est une matrice diagonale d'ordre N et I est la matrice identité. Chacune des composantes $\omega_i(k)$ de $\Omega(k)$ est une suite i.i.d., centrée, de puissance P_Ω , indépendante des suites $x(k)$ et $n(k)$.

Pour différents algorithmes adaptatifs de type erreur a priori, comme le LMS, le LMS normalisé (NLMS), le LMS avec signe sur les $x(k)$ (SLMS) et le RLS avec oubli exponentiel, nous prouvons la décroissance exponentielle de la composante transitoire de l'erreur paramétrique.

Contrairement aux hypothèses utilisées dans [1] pour le LMS ou dans [2] pour le RLS, ceci est fait sans l'hypothèse irréaliste d'indépendance des vecteurs observation $X(k)$. L'hypothèse (réaliste) est celle de la M-indépendance des vecteurs $X(k)$. A partir du formalisme de découplage introduit dans [3] puis [4], cette contribution développe des résultats de convergence exponentielle et de poursuite pour des variations de type général markovien.

ABSTRACT

Adaptive identification algorithms in non stationary markovian context are studied using a mathematical formulation that makes the problem stationary. Exponential decay of the transient component is proved for various algorithms. This framework allows to compare the tracking performances in non stationary contexts of random walk type and markovian type. Fast non stationarities, that are not dealt with in literature, are investigated here.

En régime permanent, on établit des résultats analytiques sur la poursuite des non stationnarités aussi bien rapides que lentes alors que les travaux antérieurs [5][6] ne traitent que les variations lentes. L'ensemble des résultats est obtenu grâce à un formalisme mathématique qui stationnarise le problème.

II- STATIONNARISATION DU PROBLEME

II-1 Formalisme des équations

Le filtre adaptatif $H(k)$ estimé selon divers algorithmes avec erreur a priori peut être décrit par l'équation générale :

$$H(k + 1) = H(k) + G(k)e(k), \quad (II-1)$$

où $G(k)$ est un gain vectoriel qui dépend de l'algorithme (par exemple pour le LMS, $G(k) = \mu X(k)$, où $\mu > 0$ est le pas d'adaptation de l'algorithme, et pour le RLS avec facteur d'oubli λ , $G(k) = R(k)^{-1} X(k)$, où la quantité $(1 - \lambda)R(k)$ est l'estimation de la matrice d'autocorrélation R de l'entrée $X(k)$, définie récursivement $R(k) = \lambda R(k - 1) + X(k)X(k)^T$ où $0 < \lambda < 1$). (II-2)

A partir de $H(k)$, on obtient le signal estimé

$$\hat{s}(k) = H(k)^T X(k). \quad (II-3)$$

A l'aide du formalisme mathématique de [4], le vecteur

$$W(k) = (V(k)^T, F(k - 1)^T)^T \text{ évolue selon}$$

$$W(k + 1) = A(k)W(k) + B(k)n(k) + C(k)\Omega(k) \quad (II-4)$$

avec

$$A(k) = \begin{bmatrix} I - G(k)X(k)^T & \underline{\eta}(I - \underline{\eta}) \\ 0 & (I - \underline{\eta}) \end{bmatrix}, \quad (II-5)$$

$$B(k) = \begin{bmatrix} G(k) \\ 0 \end{bmatrix}, \quad C(k) = \begin{bmatrix} -(I - \underline{\eta}) - G(k)X(k)^T \\ I \end{bmatrix}, \quad (II-6)$$



et où $V(k) = H(k) - (I - \underline{\eta})F(k-1)$ mesure la déviation entre le filtre adaptatif $H(k)$ et le vecteur paramètre optimal qui peut être atteint par un algorithme a priori. D'après (II-4), $W(k)$ a trois composantes, $W^t(k)$ liée au régime transitoire, $W^n(k)$ causée par le bruit d'observation $n(k)$ et $W^\Omega(k)$ se rapportant à la non stationnarité du filtre vrai ($\Omega(k)$). Ainsi,

$$W(k) = W^t(k) + W^n(k) + W^\Omega(k), \quad (\text{II-7})$$

$$W^t(k+1) = A(k)W^t(k), \quad (\text{II-8-a})$$

$$W^n(k+1) = A(k)W^n(k) + B(k)n(k), \quad (\text{II-8-b})$$

$$W^\Omega(k+1) = A(k)W^\Omega(k) + C(k)\Omega(k), \quad (\text{II-8-c})$$

$$\text{avec } W^t(0) = W(0), W^n(0) = W^\Omega(0) = 0. \quad (\text{II-9})$$

II-2 Découplage des erreurs

Les performances de l'algorithme sont mesurées par l'erreur quadratique moyenne résiduelle (EQMR)

$$m(k) = E[(s(k) - \hat{s}(k))^2] / P_n. \quad (\text{II-10})$$

Cette erreur est normalisée par le bruit. Il découle de (II-7) et des hypothèses d'indépendance, la formule de découplage

$$m(k) = m^t(k) + m^n(k) + m^\Omega(k) + \delta, \quad (\text{II-11})$$

où

$$\delta = E[\|\Omega(k)^T X(k)\|^2] / P_n \quad (\text{II-12})$$

est appelé **degré de non stationnarité** et où les composantes d'erreur

$$m^a(k) = E[\|V^a(k)^T X(k)\|^2] / P_n; \quad a = t, n, \Omega \quad (\text{II-13})$$

correspondent aux trois déviations différentes de (II-8).

III DECROISSANCE EXPONENTIELLE DU TRANSITOIRE.

La vitesse de convergence est caractérisée par la décroissance vers 0 du transitoire $W^t(k)$. Celle-ci entraîne que les moments de $W^n(k)$ et de $W^\Omega(k)$ sont bornés et tendent vers une limite finie quand k tend vers l'infini. En effet, d'après (II-8 a,b,c)

$$W^t(k) = E_{k-1,-1} W^t(0), \quad (\text{III-1-a})$$

$$W^n(k) = \sum_{j=0}^{k-1} E_{k-1,j} B(j)n(j), \quad (\text{III-1-b})$$

$$W^\Omega(k) = \sum_{j=0}^{k-1} E_{k-1,j} C(j)\Omega(j), \quad (\text{III-1-c})$$

où

$$E_{k-1,j}^\Delta = A(k-1)A(k-2)\dots A(j+1). \quad (\text{III-2})$$

On a

$$E_{k-1,j} = \begin{bmatrix} U_{k-1,j} & \sum_{l=1}^{k-j-1} U_{k-l-1,j} \underline{\eta}(I - \underline{\eta})^l \\ O & (I - \underline{\eta})^{k-j-1} \end{bmatrix} \quad (\text{III-3})$$

avec

$$U_{k-1,j} = T(k-1)T(k-2)\dots T(j+1), \quad (\text{III-4})$$

$$T(k) = (I - G(k)X(k)^T). \quad (\text{III-5})$$

Pour chaque algorithme, la matrice U correspond au cas stationnaire. Sa décroissance a été prouvée dans [3][7][8][9], respectivement pour le LMS, NLMS, SLMS et le RLS (pour des vecteurs $X(k)$ corrélés). D'après la structure particulière (III-3), la matrice E a donc elle aussi une décroissance exponentielle. Il en est de même pour ses moments. Il découle alors de (III-1-a) que $W^t(k)$ décroît exponentiellement ainsi que ses moments. Par conséquent, en régime permanent, la composante de l'erreur paramétrique transitoire est nulle.

IV- NON STATIONNARITES RAPIDES

La capacité de poursuite d'un algorithme en régime permanent est mesurée par son EQMR. Elle dépend du taux d'oubli ν défini par exemple par $V_{LMS} = \mu P_x$ et $V_{RLS} = 1 - \lambda$; on note m_ν la limite de $m(k)$. D'après ce qui précède on a

$$m_\nu = m_\nu^n + m_\nu^\Omega + \delta. \quad (\text{IV-1})$$

D'une manière générale, la composante de bruit m_ν^n est croissante en fonction de ν . Inversement la composante non stationnaire m_ν^Ω est décroissante par rapport à ν . Il peut en résulter une valeur de compromis V_{opt} minimisant m_ν , ainsi qu'il sera détaillé plus bas.

IV-1 Composantes de m_ν en régime permanent

Les composantes de m_ν sont calculées ici théoriquement pour une suite $x(k)$ i.i.d. centrée. Le calcul est valable pour toute valeur de $\eta_i, i=1, \dots, N$, donc **que la non stationnarité soit lente ou rapide**.

Comme exemple nous prenons le LMS avec l'hypothèse $|X(k)|^2 = cte = NP_x$. Sous l'hypothèse gaussienne les démonstrations sont similaires et les résultats qualitatifs sont identiques. Le calcul donne [10]

$$m_\nu^n = \frac{\nu}{2 - \nu} \quad (\text{IV-2})$$

$$m_\nu^\Omega = \frac{\delta}{(2 - \nu)} [\nu - 2(1 - \eta)] + \frac{2\delta(1 - \eta)}{(2 - \nu)(2 - \eta)} \frac{[N(1 + \eta - \eta^2) - \nu\eta(2 - \eta)]}{[\eta N + \nu(1 - \eta)]} \quad (\text{IV-3})$$



L'EQMR est alors donnée par :

$$m_v = \frac{1}{(2-v)} \left[v + \frac{2\delta N}{(2-\eta)[v(1-\eta) + \eta N]} \right]. \quad (IV-4)$$

Il est important de noter que, à la différence des travaux antérieurs, nous ne faisons pas l'hypothèse que l'algorithme est lent. Ainsi, **le taux d'oubli peut être grand.**

Pour le RLS, on trouve la même expression de m_v à condition que V soit très petit, de sorte que les fluctuations de la matrice R_k^{-1} autour de sa limite moyenne vR^{-1} soient faibles.

IV-2 De la promenade aléatoire au modèle markovien.

Lorsque η tend vers 0, la non stationnarité de $F(k)$ devient une promenade aléatoire. L'étude de cette non stationnarité a été faite dans [3]. Le passage à une non stationnarité markovienne ($\eta \neq 0$) modifie le comportement du LMS de manière significative :

1- On constate que pour $v = 0$, l'EQMR est finie :

$$m_0 = \delta / [\eta(2-\eta)] \quad (IV-5)$$

Le modèle de non stationnarité markovienne possède donc une meilleure signification physique que la promenade aléatoire pour laquelle l'EQMR est infinie pour $v=0$.

2- m_v est croissante pour $v > 0$ ($v_{opt} < 0$) lorsque

$$\eta > \eta_0 = 2/(N+2). \quad (IV-6)$$

Ceci montre que la poursuite est nuisible lorsque η est très grand. Naturellement une valeur non nulle de v est nécessaire à l'initialisation. Mais il faut ensuite bloquer l'adaptation ($v = 0$). Elle ne fait que dégrader les performances. La borne (IV-6) se comprend aisément car alors le filtre $F(k)$ est complètement renouvelé avant même que l'échantillon $x(k)$ ait parcouru tout le filtre.

3- Si η n'est pas trop grand ($\eta < \eta_0$), la situation est semblable (m_v croissante) lorsque :

$$\delta < \delta_0 = \frac{N\eta^2(2-\eta)}{2(1-\eta) - N\eta}. \quad (IV-7)$$

Si δ est trop faible par rapport à η , il faut là encore bloquer la poursuite après l'acquisition. Pour $\eta \ll 2/(N+2)$, ceci arrive lorsque la puissance P_B du bruit de sortie, mesuré dans la bande η qui caractérise les variations du filtre $F(k)$, vérifie $P_B (= N\eta P_n) > 2P_s$, où P_s est la puissance du signal à la sortie du filtre $F(k)$. Dans ce cas, l'observation de $y(k)$ est trop bruitée pour permettre la poursuite. Là encore, le modèle markovien possède une meilleure interprétation physique que la promenade aléatoire, pour laquelle, plus δ est faible (P_n élevé) plus l'EQMR minimale est faible (elle vaut δ).

4- On peut trouver des couples (η, δ) pour lesquels $m_{min} > 1$. Ainsi $y(k)$ est plus proche de $s(k)$ que $\hat{s}(k)$. Il vaut mieux ne pas estimer $s(k)$.

IV-3 Simulations

Les résultats théoriques trouvés dans le cas de l'algorithme LMS sont en parfait accord avec les simulations dans le cas d'une entrée $x(k)$ i.i.d., que le taux d'oubli soit élevé ou non et que la non stationnarité soit lente ou rapide.

On implémente un filtre d'ordre 3. La figure 1 représente m_v pour différentes valeurs de η ($\delta=0.01$). En particulier les cas $\eta=0.5$ et $\eta=0.1$ illustrent deux situations où la poursuite est néfaste ; le cas $\eta=0.01$ montre la présence d'un minimum d'EQMR avec une valeur v_{opt} voisine de 0.05. On remarque que pour δ et v donnés, m_v est une fonction décroissante de η . On montre facilement en effet, que $E(|F(k)|^2)$ tend vers $NP_\Omega / (\eta(2-\eta))$. Donc les fluctuations de $F(k)$ s'atténuent lorsque η augmente, ce qui explique le résultat observé.

Des simulations analogues sont effectuées pour l'algorithme RLS. La figure 1 montre que $m_v^{LMS} = m_v^{RLS}$ lorsque v est faible, comme prévu au paragraphe IV-1. Mais quand v augmente ($v > 0.2$), l'algorithme RLS surclasse l'algorithme LMS.

La figure 2 représente m_v^{LMS} à $\eta=0.1$ fixé, pour différentes valeurs de δ . Sa valeur minimale peut s'approcher de 1, rendant inutile ou néfaste l'adaptation.

Dans le cas d'un signal corrélé $x(k) = s \sin(kT) + b(k)$ où S/B est le rapport signal sur bruit, l'algorithme LMS peut être supérieur ou inférieur au RLS selon le cas. Il suffit de faire varier la période T , le rapport S/B et la corrélation normalisée $\rho_{i,j} = E(\omega_i(k) \omega_j(k)) / E(\omega_i^2(k))$ entre les composantes du vecteur $\Omega(k)$ qui génère $F(k)$. Les simulations des figures 3 et 4 illustrent respectivement les cas où $m_{min}^{LMS} < m_{min}^{RLS}$ et le cas inverse. Pour une promenade aléatoire et pour δ petit, un tel résultat avait déjà été prouvé théoriquement [11].

V- CONCLUSION

Dans un contexte non stationnaire de type markovien, à partir d'un formalisme mathématique qui stationnarise le problème d'identification, il est possible de montrer pour différents types d'algorithmes a priori, la décroissance exponentielle de la composante transitoire de l'erreur paramétrique. Ce résultat permet de retrouver, dans le cas non stationnaire, des propriétés connues dans le cas stationnaire. Ce formalisme permet aussi, en régime permanent, de traiter la poursuite de non stationnarités rapides. Dans le cas d'une entrée $x(k)$ i.i.d. les calculs analytiques ont pu être menés jusqu'à leur terme pour l'algorithme LMS. Ils montrent la supériorité du modèle markovien sur la promenade aléatoire pour traduire la réalité physique des phénomènes non stationnaires.



Références

- [1] B. WIDROW, J.M. MC COOL, "Stationary and Non Stationary Characteristics of the LMS Adaptive Filter", IEEE Trans. ASSP, Vol 64, n°8, pp1151-1162, 1976.
- [2] E. ELEFThERIOU, D. FALCONER, "Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithm", IEEE Trans. ASSP, Vol 34, pp1097-1109, 1986.
- [3] O. MACCHI, "Optimization of adaptive identification for time varying filters", IEEE Trans. AC, Vol AC31, n°3 pp283-287, 1986.
- [4] O. MACCHI, M. TURKI, "The Non Stationary Degree : Can an Adaptive Filter be Worse than No Processing", Proc. IFAC Int. Symp. on ACASP, pp 743-747, 1992.
- [5] L. LJUNG, P. PRIOURET, "A Result on Mean Square Error Obtained Using General Algorithms", Int. J. on ACASP, Vol 5, pp231-250, 1991.
- [6] A. BENVENISTE, "Design of Adaptive Algorithms for the Tracking of Time Varying Systems ", Int. J. on ACASP, Vol 1, pp3-21, 1987.
- [7] BITMEAD, ANDERSON, "Performances of adaptive estimation algorithms in dependant random environment", IEEE Trans. AC25, p788-793, 1980.
- [8] O. MACCHI et al, "Compared speed and accuracy of the RLS and LMS algorithm with constant forgetting factor", RAIRO-APII, vol 22, n°3, 1988.
- [9] W.A. STETHARES, C.R. JOHNSON, "A comparison of two quantized state adaptive algorithms", IEEE Trans. ASSP, pp 138-143, Janvier 1989.
- [10] M. TURKI, O. MACCHI, M. JAIDANE-SAIDANE, "Calcul du régime permanent de l'algorithme LMS en poursuite de non stationnarités markoviennes", Rapport LSS, Juillet 1993.
- [11] F. MICHAUT, "Méthodes adaptatives pour le signal", Hermès, 1992.

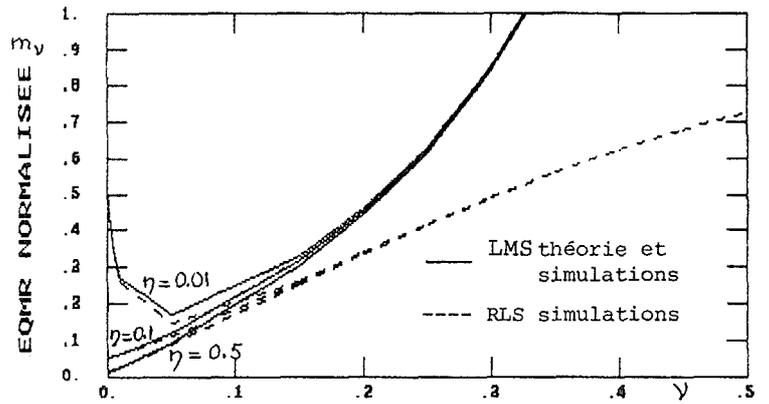


Figure 1 : Entrée $x(k)$ i.i.d., $\delta = 10^{-2}$

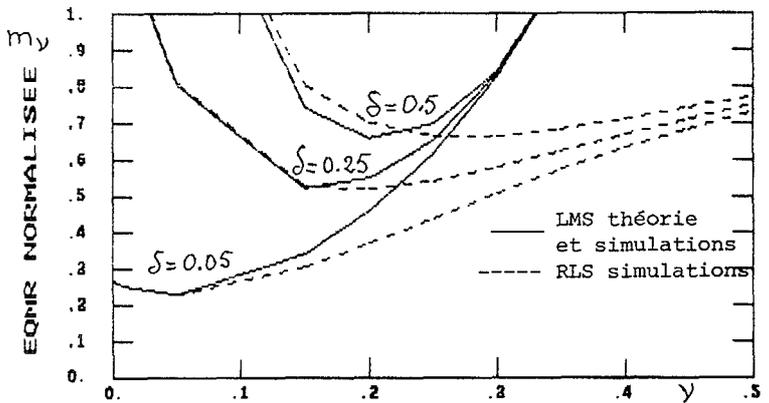


Figure 2 : Entrée $x(k)$ i.i.d., $\eta = 10^{-1}$

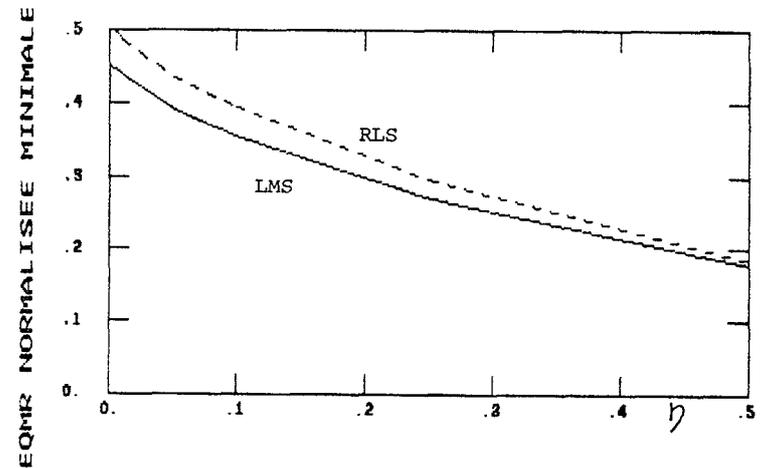


Figure 3 : Entrée $x(k)$ corrélée, supériorité du LMS
($T = \pi/4$, $S/B = 2.4$, $\rho_{i,j} = 0.5$, $\delta = 0.1$)

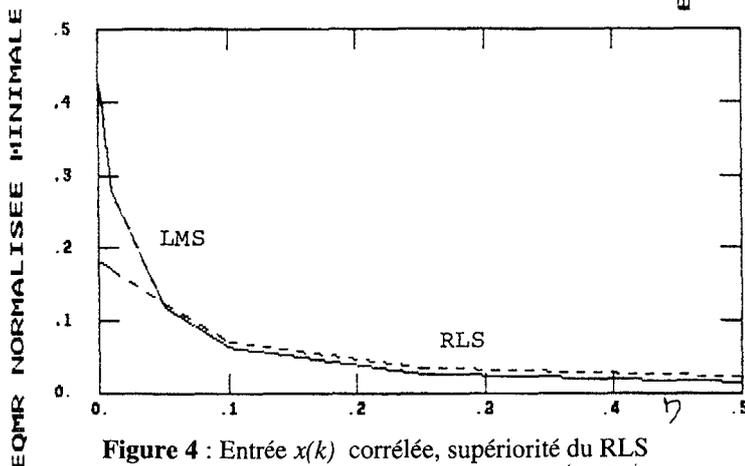


Figure 4 : Entrée $x(k)$ corrélée, supériorité du RLS
($T = 3\pi/4$, $S/B = 23.3$, $\rho_{i,j} = 0.97$, $\delta = 0.1$)