

Analyse de performance de l'algorithme du gradient stochastique en régime permanent par la méthode de développement par perturbation

A. Perrier^{*}, B. Delyon^{**} et E. Moulines^{*}

^{*} ENST, Dépt. Signal, 46 rue Barrault, 75634 Paris Cedex 13, France. ^{**} IRISA/INRIA, Campus de Beaulieu, Rennes

E-mail: perrier@sig.enst.fr, Fax: 1 45887935.

RÉSUMÉ

L'analyse de performance de l'algorithme du gradient stochastique fait l'objet de nombreux travaux. Dans la littérature orientée 'approximation stochastique', des résultats sont établis pour des pas d'adaptation très petits par des techniques de convergence faible. Ces résultats, importants d'un point de vue théorique, ne sont pas utilisables en pratique. Les hypothèses d'indépendance, malgré le fait qu'elle ne sont pratiquement jamais vérifiées, connaissent une grande popularité de part leur simplicité analytique et l'exactitude des résultats obtenus dans certains cas. Dans cette contribution, nous présentons une méthode originale qui permet d'obtenir le développement explicite de la matrice de covariance du vecteur écart par rapport au pas d'adaptation en régime permanent. La connaissance des termes du développement permet d'expliquer de nombreux aspects du comportement de l'algorithme. La validité des hypothèses d'indépendance est examinée et le contexte où la dimension du filtre inconnu est sous-modélisé est étudié.

ABSTRACT

The performance analysis of the LMS algorithm has been the subject of many contributions. In stochastic approximation literature, much emphasis has been put on infinitesimal stepsize analysis based on weak convergence results; although theoretically important, these results lack practical use. As a way to overcome these limitations, the independence assumptions, despite their obvious theoretical shortcomings, have gained much popularity in the engineering community due to their analytic convenience and their moderate agreement with experimental results. In this contribution, an original method to carry out the steady-state performance analysis of stochastic gradient type algorithms is introduced. This method allows the explicit computation of the expansion of the weight-error covariance matrix with respect to the step size for the LMS algorithm. The inspection of the expansion terms enlightens many observed behaviours which up to now remained unexplained. The validity of the independence assumptions are examined and the context of sub-modelization of the unknown parameter vector is investigated.

1 Introduction

L'algorithme du gradient stochastique [1] est un des algorithmes les plus attrayants pour l'identification de filtre à réponse impulsionnelle finie (RIF) de part sa simplicité et sa faible complexité numérique. Cet algorithme est défini par les récursions suivantes

$$\begin{cases} \epsilon_{t+1} = d_{t+1} - \mathbf{x}_t^T \hat{\mathbf{h}}_t \\ \hat{\mathbf{h}}_{t+1} = \hat{\mathbf{h}}_t + \mu \mathbf{x}_t \epsilon_{t+1} \end{cases} \quad (1)$$

où $\mathbf{x}_t = [x(t), \dots, x(t-L+1)]^T$ est le vecteur de dimension L du signal d'entrée $x(t)$, d_t est le signal de sortie ou signal désiré, μ est le pas d'adaptation scalaire et $\hat{\mathbf{h}}_t$ est le vecteur de dimension L des coefficients du filtre adaptatif à l'instant t . ϵ_t est l'erreur a priori définie comme la différence entre le signal désiré et le signal reconstruit.

L'analyse théorique des performances des algorithmes d'identification de type gradient stochastique a fait l'objet de nombreux travaux. Dans la littérature orientée 'Traitement du Signal', la plupart des analyses de performance de ces algorithmes ont été menées sous les hypothèses d'indépendance, introduites par Widrow et al [1] puis popularisées par Haykin [2]. Toutefois, ces hypothèses ne sont pratiquement jamais vérifiées. On observe néanmoins dans certaines situations, une assez bonne cohérence entre les observations et les prédictions théoriques, ce qui semble plaider en faveur de la 'validité' pratique de ces hypothèses (la validité expérimentale est souvent mise en exergue par

les défenseurs des hypothèses d'indépendance).

Dans la littérature orientée 'Approximation Stochastique' [3], les performances de ce type d'algorithmes ont été étudiées à la limite des pas d'adaptation infinitésimalement petits et des conditions très générales sur les signaux d'entrée (α -mélangeants). Bien que ces résultats aient un intérêt théorique certain, les valeurs de pas considérées sont beaucoup plus faibles que celles que l'on utilise couramment en pratique.

Les travaux de Macchi et Eweda [4] sur l'algorithme du gradient stochastique ont montrés que, pour des signaux m -dépendants, la Matrice de Covariance du Vecteur Ecart (MCVE) est bornée proportionnellement au pas d'adaptation (voir aussi [5]). Des résultats récents sur la stabilité des algorithmes d'identification stochastiques ont été développés par Guo [6]. Dans [7], le développement de la MCVE est effectué jusqu'au premier ordre par rapport au pas d'adaptation et la convergence presque sûre de l'algorithme est établie.

Notre but est d'obtenir une expression explicite du développement de la MCVE en fonction du pas d'adaptation μ pour des signaux m -dépendants. Nous utilisons pour cela une méthode dite de développement par perturbation inspirée des travaux de Solo [7]; sa forme générale est due à Priouret (on se reportera à la thèse de Mazroui [8]). Le vecteur écart est décomposé en une série de sous-processus de telle façon que (i) les variances de ces sous-



processus correspondent à un ordre donné d'approximation en μ et (ii) ces variances sont solutions d'équations de Lyapounov.

Nous présentons brièvement les bases théorique de la méthode dans une première partie, puis nous l'appliquons à l'algorithme du gradient stochastique dans une deuxième partie. Enfin des simulations présentées dans une troisième partie démontrent la validité des résultats obtenus.

notations La norme euclidienne $\|\cdot\|$ est utilisée pour les vecteurs. La norme vectorielle p aléatoire est définie par $\|\mathbf{v}\|_p = (E\{\|\mathbf{v}\|^p\})^{\frac{1}{p}}$. La norme matricielle spectrale est utilisé $\|X\| = \{\lambda_{\max}(XX^T)\}^{\frac{1}{2}}$ où $\lambda_{\max}(X)$ dénote la valeur propre maximum de la matrice X . La norme p aléatoire matricielle est définie par $\|X\|_p = [E\{\|X\|^p\}]^{\frac{1}{p}}$. L'opérateur de transposition est noté par un exposant T et la matrice identité par I .

2 Bases de la méthode

Le modèle de base est celui de la régression linéaire, le signal de sortie est donné par

$$d_t = \mathbf{x}_t^T \mathbf{h}_t + \nu_t \quad t \geq 0 \quad (2)$$

où ν_t est un bruit scalaire et \mathbf{h}_t est le vecteur de dimension L des coefficients du filtre à identifier. Nous supposons ici, que ce filtre est constant, $\mathbf{h}_t = \mathbf{h}$. L'étude du comportement en poursuite fera l'objet d'un autre article (voir [9] pour des résultats préliminaires). La récursion du vecteur écart $\mathbf{\Lambda}_t = \hat{\mathbf{h}}_t - \mathbf{h}$ définie comme la différence entre le filtre adaptatif et le vrai filtre est alors

$$\mathbf{\Lambda}_{t+1} = (I - \mu \mathbf{x}_t \mathbf{x}_t^T) \mathbf{\Lambda}_t + \mu \xi_{t+1} \quad (3)$$

où $\xi_{t+1} = \mathbf{x}_t \nu_t$. Le vecteur écart vérifie donc une équation linéaire du premier ordre de matrice de transition $(I - \mu \mathbf{x}_t \mathbf{x}_t^T)$ excité par une séquence de vecteur aléatoire ξ_t . Nous considérons les hypothèses suivantes sur les signaux

H1) Le signal d'entrée $x(t)$ est strictement stationnaire, centré, m -dépendant et ses moments à tous les ordres sont bornés, c.à.d. $\forall p \geq 2 \quad |x(t)|_p < \infty$

H2) Le bruit $\nu(t)$ est strictement stationnaire, centré, m -dépendant et ses moments à tous les ordres sont bornés.

La condition de m -dépendance est en fait une condition technique. Des résultats similaires sont obtenus pour des signaux vérifiant des conditions de mélanges moins restrictives mais l'aspect théorique de la méthode est plus difficile à établir. Dans la plupart des travaux sur les algorithmes stochastiques d'identification, l'hypothèse est faite que le bruit est indépendant du signal d'entrée. Ce n'est pas le cas ici, ce qui nous permettra d'étudier le cas où la dimension du filtre à identifier est sous-estimée. L'hypothèse **H1** implique que la valeur propre minimum de la matrice d'autocorrélation du signal d'entrée $R = E\{\mathbf{x}_t \mathbf{x}_t^T\}$ est strictement positive $\rho = \lambda_{\min}(R) > 0$. La stabilité exponentielle de l'algorithme est démontrée sous les hypothèses **H1-H2** dans [9].

2.1 Décomposition de $\mathbf{\Lambda}_t$

L'idée sous-jacente à la décomposition de $\mathbf{\Lambda}_t$ est de s'affranchir du caractère aléatoire de la matrice $(I - \mu \mathbf{x}_t \mathbf{x}_t^T)$ en la remplaçant par son espérance $(I - \mu R)$. La décomposition de $\mathbf{\Lambda}_t$ est basée sur la décomposition de la matrice de transition: $(I - \mu \mathbf{x}_t \mathbf{x}_t^T) = (I - \mu R) + \mu Z_t$ où $Z_t = R - \mathbf{x}_t \mathbf{x}_t^T$. Le vecteur écart se décompose alors en n sous-processus de la façon suivante

$$\mathbf{\Lambda}_t = \mathbf{J}_t^0 + \mathbf{J}_t^1 + \dots + \mathbf{J}_t^{n-1} + \bar{\mathbf{J}}_t^n \quad (4)$$

et

$$\begin{aligned} \mathbf{J}_{t+1}^0 &= (I - \mu R) \mathbf{J}_t^0 + \mu \xi_{t+1} \\ \mathbf{J}_{t+1}^k &= (I - \mu R) \mathbf{J}_t^k + \mu Z_t \mathbf{J}_t^{k-1} \quad 1 \leq k \leq n \\ \bar{\mathbf{J}}_{t+1}^n &= (I - \mu \mathbf{x}_t \mathbf{x}_t^T) \bar{\mathbf{J}}_t^n + \mu Z_t \mathbf{J}_t^{n-1} \quad k \geq 1 \end{aligned} \quad (5)$$

Le théorème suivant montre l'intérêt de cette décomposition.

Théorème 1 Sous les hypothèses **H1-H2**, $\forall p \geq 2$, $\exists \mu_p > 0$, $C_{0,p}^k < \infty$, $C_{1,p}^k < \infty$, $D_{0,p}^k < \infty$, $D_{1,p}^k < \infty$, tel que $\forall 0 < \mu < \mu_p$

$$\begin{aligned} \|\mathbf{J}_t^k\|_p &\leq C_{0,p}^k \mu^{\frac{k}{2}} |E\{\xi_t\}|_{2p} + C_{1,p}^k \mu^{\frac{k+1}{2}} \quad k \geq 0 \\ \|\bar{\mathbf{J}}_t^k\|_p &\leq D_{0,p}^k \mu^{\frac{k}{2}} |E\{\xi_t\}|_{2p} + D_{1,p}^k \mu^{\frac{k+1}{2}} \quad k \geq 1 \end{aligned} \quad (6)$$

Ce théorème est démontré dans [9]. Il permet d'obtenir le développement de la MCVE à n'importe quel ordre en μ grâce aux inégalités suivantes

$$\begin{aligned} \|E\{\mathbf{J}_t^i (\mathbf{J}_t^k)^T\}\| &\leq C_{1,2}^i C_{1,2}^k \mu^{\frac{i+k}{2}} \\ \|E\{\bar{\mathbf{J}}_t^{2q+1} (\mathbf{J}_t^k)^T\}\| &\leq D_{1,2}^{2q-1} C_{1,2}^k \mu^{q+\frac{k}{2}} \end{aligned} \quad (7)$$

Par conséquent, pour obtenir un développement de $\Pi^\Lambda = E\{\mathbf{\Lambda}_t \mathbf{\Lambda}_t^T\}$ à un ordre donné q en μ , il suffit de décomposer $\mathbf{\Lambda}_t$ en $n = (2q + 1)$ sous-processus ($n = (2q - 1)$ si $E\{\xi_t\} = 0$) et de calculer les variances des sous-processus d'ordre inférieur à μ^q . Notons $\Pi^{i,k} = E\{\mathbf{J}_t^i (\mathbf{J}_t^k)^T\}$, le développement à l'ordre 2 de la MCVE, est donné par les relations suivantes

- si $E\{\xi_t\} = 0$, $\Pi^\Lambda = \sum_{0 \leq i+k \leq 2} \Pi^{i,k} + O(\mu^3)$ et le premier terme du développement est d'ordre μ .
- Si $E\{\xi_t\} \neq 0$, $\Pi^\Lambda = \sum_{0 \leq i+k \leq 4} \Pi^{i,k} + O(\mu^3)$ et le premier terme du développement est constant par rapport à μ .

Dans la suite nous présentons les deux premiers termes du développement par rapport à μ de Π^Λ dans le cas classique où le bruit est indépendant du signal d'entrée; ξ_t est alors centré; puis dans le contexte de sous-modélisation du filtre inconnu. Le bruit et le signal d'entrée sont alors corrélés et ξ_t n'est pas centré. Les calculs des différents termes présents dans les dérivations sont trop longs pour être présentés dans cette contribution. Le lecteur peut se reporter à [9].

3 Bruit et signal d'entrée indépendants

Nous supposons ici que

H3 le bruit est indépendant du signal d'entrée

D'après le théorème 1, le terme d'ordre μ du développement de Π^Λ est obtenu en ne considérant que la covariance du premier sous-processus \mathbf{J}_t^0 . Cette covariance vérifie l'équation suivante obtenue en multipliant l'Eq. (5) par sa transposée et en prenant l'espérance

$$R\Pi^{0,0} + \Pi^{0,0}R - \mu R\Pi^{0,0}R = \mu R\sigma_\nu^2 + (I - \mu R)E\{\mathbf{J}_t^0 \xi_{t+1}^T\} + E\{\xi_{t+1}(\mathbf{J}_t^0)^T\}(I - \mu R) \quad (8)$$

De la même façon, le développement au second ordre est obtenu en résolvant les équations vérifiées par les covariances des sous-processus $(\Pi^{i,k})_{\{0 \leq i+k \leq 2\}}$. Le résultat final est présenté dans la proposition suivante

Proposition 1 *Les hypothèses H1-H3 sont vérifiées, Le développement de Π^Λ au second ordre est*

$$\Pi^\Lambda = \mu \Xi^0 + \mu^2(\Xi^0 + \Xi^1 + (\Xi^1)^T + \Xi^2) + O(\mu^3) \quad (9)$$

où Ξ^0 , Ξ^1 and Ξ^2 sont les uniques solutions des équations de Lyapounov

$$\begin{aligned} R\Xi^0 + \Xi^0R &= \Gamma_{x,\nu} \\ R\Xi^1 + \Xi^1R &= -R\Delta_{x,\nu} + \Delta_{Z,x,\nu} + \Gamma_{Z,x,\nu} + \Delta_Z\Xi^0 \\ R\Xi^2 + \Xi^2R &= \Gamma_{Z,x,\nu,Z} \end{aligned} \quad (10)$$

et

$$\begin{aligned} \Gamma_{x,\nu} &= \sum_{s=-m+1}^{m-1} R(s)r_\nu(s) \\ \Delta_{x,\nu} &= \sum_{s=2}^{m-1} (s-1)R(-s)r_\nu(-s) \\ \Delta_{Z,x,\nu} &= \sum_{s=0}^{m-1} \sum_{\tau=s+1}^{m-1} E\{Z_0\mathbf{x}_{s-\tau}\mathbf{x}_s^T\}r_\nu(-\tau) \\ \Gamma_{Z,x,\nu} &= \sum_{s=1}^{m-1} (E\{Z_s\mathbf{x}_0\mathbf{x}_0^T\}\sigma_\nu^2 + \sum_{\tau=1}^{m-1} E\{Z_s(\mathbf{x}_\tau\mathbf{x}_0^T + \mathbf{x}_0\mathbf{x}_\tau^T)\}r_\nu(\tau)) \\ \Delta_Z &= \sum_{s=1}^{m-1} E\{Z_sZ_0\} \\ \Gamma_{Z,x,\nu,Z} &= \sum_{s=-m+1}^{m-1} E\{Z_s\Xi^0Z_0\} \\ R(s) &= E\{\mathbf{x}_{t+s}\mathbf{x}_t^T\} \quad r_\nu(s) = E\{\nu_{t+s}\nu_t\} \end{aligned} \quad (11)$$

Ce résultat permet de prédire le comportement de l'algorithme de façon bien plus précise qu'auparavant. La connaissance du terme quadratique permet de considérer des valeurs de pas utilisées en pratique. Il est intéressant de comparer le résultat obtenu dans le cas particulier d'un bruit i.i.d. avec celui obtenu sous les hypothèses d'indépendance. Si le bruit est i.i.d., les termes de l'Eq. (9) se simplifient. On obtient comme résultat final

$$\Pi^\Lambda = \frac{\mu}{2}\sigma_\nu^2I + \frac{\mu^2}{2}\sigma_\nu^2\Xi^3 + O(\mu^3) \quad (12)$$

où Ξ^3 est solution de l'équation de Lyapounov

$$R\Xi^3 + \Xi^3R = E\{x_t x_t^T x_t x_t^T\} \quad (13)$$

Considérons maintenant la récursion vérifiée par la MCVE sous les hypothèses d'indépendance. Le filtre adaptatif est le signal d'entrée étant indépendant, on peut écrire que

$$\Pi^\Lambda = E\{(I - \mu x_t x_t^T)\Pi^\Lambda(I - \mu x_t x_t^T)\} + \mu^2 R\sigma_\nu^2 \quad (14)$$

En développant cette équation et après simplification, on obtient

$$R\Pi^\Lambda + \Pi^\Lambda R = \mu E\{x_t x_t^T \Pi^\Lambda x_t x_t^T\} + \mu R\sigma_\nu^2 \quad (15)$$

Cette équation se résout en posant $\Pi^\Lambda = \mu A + \mu^2 B + O(\mu^3)$, et l'on obtient par conséquent le développement à l'ordre 2 de la MCVE sous les hypothèses d'indépendance

$$\Pi^\Lambda = \frac{\mu}{2}\sigma_\nu^2I + \frac{\mu^2}{2}\sigma_\nu^2\Xi^3 + O(\mu^3) \quad (16)$$

où Ξ^3 est solution de l'Eq. (13). Le développement de Π^Λ au second ordre en μ est donc identique sous les hypothèses d'indépendance et par la méthode de développement par perturbation qui est établie d'un point de vue théorique. L'apparente validité des résultats issue des hypothèses d'indépendance trouve ainsi une justification théorique. Les limites de ces hypothèses sont aussi soulignées. En effet, si le bruit est corrélé, ces hypothèses ne rendent absolument pas compte de l'influence de cette corrélation. Nous considérons maintenant le contexte de la sous-modélisation.

4 Sous modélisation

La généralité des hypothèses relatives au bruit permet d'étudier le cas où la dimension du filtre à identifier est sous-modélisée. Nous supposons donc que

H4 La taille du filtre adaptatif L est inférieure à la taille du filtre à identifier L' .

Notons $N = L' - L$ et écrivons le filtre à identifier sous la forme

$$\mathbf{h} = [\mathbf{h}^T, \tilde{\mathbf{h}}^T]^T \quad (17)$$

où $\mathbf{h} = [h(0), \dots, h(L-1)]^T$ est le vecteur de dimension L qui correspond à la partie identifiée du filtre et $\tilde{\mathbf{h}} = [h(L), \dots, h(L'-1)]^T$ le vecteur de taille N qui n'est pas pris en compte par le filtre adaptatif. Le vecteur du signal d'entrée se réécrit sous une forme similaire

$$\mathbf{x}_t = [\mathbf{x}_t^T, \tilde{\mathbf{x}}_t^T]^T \quad (18)$$

où $\mathbf{x}_t = [x(t), \dots, x(t-L+1)]^T$ est la partie du vecteur \mathbf{x}_t contenue dans les équations de l'algorithme et $\tilde{\mathbf{x}}_t = [x(t-L), \dots, x(t-L'+1)]^T$. $\tilde{\mathbf{x}}_t$ correspond aux N premiers composants du vecteur \mathbf{x}_{t-L} (retardé de L itérations), on peut donc écrire

$$\tilde{\mathbf{x}}_t = I_{N \times L'} \mathbf{x}_{t-L} \quad (19)$$

où $I_{N \times L'}$ est la matrice (identité) de taille $(N \times L')$ composée de 1 sur la diagonale principale et de 0 ailleurs. Le signal désiré est alors donné par

$$d_t = \mathbf{x}_t^T \mathbf{h} + \alpha_{t+1} = \mathbf{x}_t^T \mathbf{h} + \mathbf{x}_{t-L}^T I_{L' \times N} \tilde{\mathbf{h}} + \alpha_{t+1} \quad (20)$$

On suppose que

H5 α_t est un bruit centré, strictement stationnaire, m -dépendant et indépendant du signal d'entrée

Le signal désiré peut donc s'écrire

$$d_t = \mathbf{x}_t^T \mathbf{h} + \nu_{t+1} \quad (21)$$

où $\nu_{t+1} = \mathbf{x}_{t-L}^T I_{L' \times N} \tilde{\mathbf{h}} + \alpha_t$ est un bruit de mesure général qui, lui, est dépendant du signal d'entrée. Ce bruit vérifie l'hypothèse **H2** et on peut donc appliquer les résultats obtenus dans le cas où le bruit et le signal d'entrée sont dépendants [9]. Le résultat final est donné dans la proposition suivante:



Proposition 2 Les Hypothèses **H1-H2** et **H4-H5** étant vérifiées,

$$\Pi^\Lambda = R^{-1} \tilde{R}(L) D \tilde{R}(L)^T R^{-1} + \mu (X_1 + X_1^T) + \mu \Xi^5 + O(\mu^2) \quad (22)$$

où

$$\begin{aligned} X_1 &= \sum_{s=1}^{m-1} E \left\{ \mathbf{x}_s \mathbf{x}_s^T \left(\mathbf{x}_0 \mathbf{x}_0^T R^{-1} \tilde{R}(L) - I_{L \times L'} \mathbf{x}_0 \mathbf{x}_0^T \right) \right\} D \tilde{R}(L)^T \\ D &= I_{L' \times N} \tilde{\mathbf{h}} \tilde{\mathbf{h}}^T I_{N \times L'} \\ \tilde{R}(L) &= E \{ \mathbf{x}_{t+L} \mathbf{x}_t \} \end{aligned} \quad (23)$$

et Ξ^4 est solution de l'équation

$$\begin{aligned} R \Xi^4 + \Xi^4 R &= \sum_{s=-m+1}^{m-1} E \left\{ \left(\mathbf{x}_s \mathbf{x}_{s-L}^T - \mathbf{x}_s \mathbf{x}_s^T R^{-1} \tilde{R}(L) \right) \right. \\ &\quad \left. D \left(\mathbf{x}_{-L} \mathbf{x}_0^T - \tilde{R}(L)^T R^{-1} \mathbf{x}_0 \mathbf{x}_0^T \right) \right\} \\ &\quad + R_s r_\alpha(s) \end{aligned} \quad (24)$$

La sous-modélisation a une grande influence sur le comportement de l'algorithme. Le premier terme étant constant par rapport à μ , l'algorithme ne peut pas converger en dessous d'un certain seuil d'identification.

5 simulations

Afin de démontrer l'exactitude des expressions obtenues par la méthode de développement par perturbation, nous présentons deux simulations. La performance de l'algorithme est évaluée par la trace (en dB) de la matrice de covariance du vecteur écart moyenné sur 64 échantillons. Les réponses impulsionnelles sont de taille $L = 64$. La performance de l'algorithme est en trait plein, le premier terme du développement est en trait espacé pointillé (---) et le second terme en trait espacé (—). On présente d'abord les cas d'un bruit corrélé, indépendant du signal d'entrée (fig 1). Dans cette simulation, le signal d'entrée et le bruit sont des processus MA d'ordre 60. Le rapport signal à bruit est de -20 dB. Contrairement à ce que l'on aurait pu croire au vu de l'expression de la proposition (1), l'influence de la corrélation du bruit n'est pas très importante. Le contexte de la sous-modélisation est ensuite présenté (fig 2). Le signal d'entrée est un processus AR(1) de coefficient 0.9. Le filtre adaptatif est de taille $L = 30$ tandis que le vrai filtre est de taille $L' = 64$, et le bruit α_t est nul. Le signal d'entrée a une séquence de coefficients de corrélation infinie et ne vérifie donc pas l'hypothèse **H1**. Néanmoins, la simulation montre bien que les résultats obtenus restent valables pour ces signaux.

6 Conclusion

Nous avons présenté une méthode de développement par perturbation qui permet d'obtenir le développement explicite de la matrice de covariance du vecteur écart par rapport au pas d'adaptation. Cette méthode s'applique dans des contextes peu étudiés auparavant. La validité des hypothèses d'indépendance est examinée. Le contexte où la dimension du filtre est sous-modélisé est étudié. Des simulations montrent l'exactitude des expressions obtenues.

References

- [1] B. Widrow, J. M. M. Cool, M. G. Larimore, and R. C. Johnson, "Stationary and non stationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151-1162, august 1976.

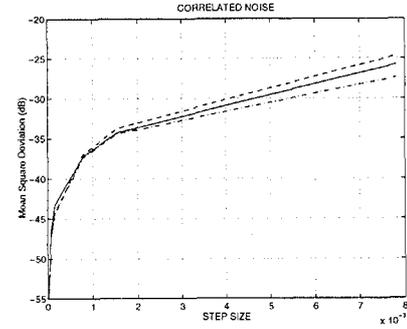


Figure 1: Erreur quadratique d'écart pour différents pas d'adaptation en présence de bruit corrélés (MA(60))

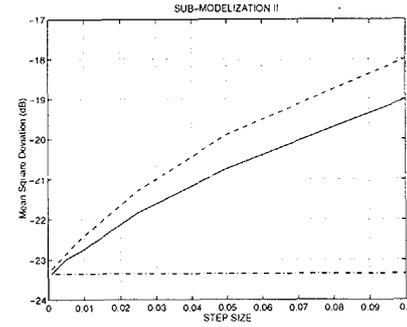


Figure 2: Erreur quadratique d'écart pour différents pas d'adaptation dans le contexte de sous-modélisation. signal d'entrée AR(0.9)

- [2] S. Haykin, *Adaptive filter theory*. Prentice Hall Info. and Syst. Science series, 1991.
- [3] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, 1990.
- [4] E. Eweda and O. Macchi, "Tracking error bounds of adaptive nonstationary filtering," *Automatica*, vol. 21, no. 3, pp. 293-302, 1985.
- [5] L. Ljung and P. Priouret, "A result on the mean square error obtained using general tracking algorithms," *Int J. Adapt. Control Signal Process.*, vol. 5, pp. 231-250, 1991.
- [6] L. Guo, "Stability of recursive stochastic tracking algorithms," *SIAM J. Control and Optimization*, vol. 32, pp. 1195-1225, Sept. 1994.
- [7] V. Solo, "The error variance of LMS with time-varying weights," *IEEE trans. on Signal Processing*, vol. 40, pp. 803-813, Apr. 1992.
- [8] A. Mazroui, *Poursuite et controle pour certains modèles de régressions*. PhD thesis, Dépt. de Probabilité, Université Pierre et Marie Curie, Paris, France, 1994.
- [9] A. Perrier, E. Moulines, and B. Delyon, "A perturbation expansion method for the performance analysis of stochastic gradient algorithms." To be submitted, 1995.