ONZIEME COLLOQUE GRETSI - NICE DU 1er AU 5 JUIN 1987

ANALYSE DES EFFETS DE QUANTIFICATION DANS L'IMPLANTATION DE L'ALGORITHME LMS

H. DEDIEU* - F. CASTANIE*

GAPSE/ENSEEIHT 2 rue Camichel 31071 TOULOUSE Cedex FRANCE

RESUME

Une étude exhaustive de l'ensemble des effets de quantification dans l'algorithme du gradient stochastique (LMS) est présentée.

Après avoir décrit les erreurs affectant l'algorithme lors de son implémentation, on cherche à exprimer les équations de récurrence qui régissent l'évolution des moments d'ordre un et deux, de l'écart du vecteur courant des coefficients au vecteur de Wiener.

Cette étude permet alors de poser aisement les conditions de non-blocage de l'algorithme, mais aussi de déduire quantitativement le comportement de la moyenne quadratique de l'erreur de sortie.

SUMMARY

This communication introduces an exhaustive study of the quantization effects in the Least Mean Square Algorithm (LMS).

After describing the errors present in the algorithm while implementing it, we try to write the recursion equations that rule the evolution of the first and the second moment of the distance between the actual coefficient vector and the Wiener vector.

This study allows us in one hand to define easily the conditions for which the adaptation ceases in the algorithm, and in the other hand to evaluate the behavior of the mean square output error.

I INTRODUCTION

Bien que l'algorithme LMS ait donné lieu à une abondante littérature, la plupart des travaux ont été effectués en supposant une précision infinie. Les effets non linéaires de l'adaptation des coefficients ont été étudiés par les auteurs dans le cas de quantification de type signe sur les données [1] ou sur l'erreur [2] et dans le cas plus général toutes les opérations non linéaires [3] de la forme :

$$\underline{W}_{n+1} = \underline{W}_n + 2\mu\, g(e_n)\, f(\underline{X})$$

On peut déjà noter que l'inclusion de ces effets non linéaires complique de façon très significative l'étude du premier et second moment du vecteur coefficient.

Si l'on considère le cas d'une opération de quantification de la forme :

$$\underline{W}_{n+1} = \underline{W}_n + Q\left[2\mu\left\{Q[e_n]\right\}Q[\underline{X}]\right]$$

dans lequel Q représente un quantificateur uniforme, la difficulté du problème apparait facilement et il n'est pas étonnant que les travaux soient peu nombreux. Ce manque théorique n'occulte pourtant pas un problème de l'implantation virgule fixe dont le praticien sait bien que la solution n'est pas toujours triviale.

Dans l'étude des effets de quantification des filtres digitaux à coefficients fixes, on suppose généralement que chaque quantificateur peut être modélisé par une source de bruit donnant naissance à une séquence d'une variable aléatoire uniformément distribuée et indépendante des autres variables issues des autres quantificateurs.

Cette façon de procéder ne peut être simplement étendue aux filtres adaptatifs, du fait de la récursivité en temps de l'algorithme, puisque tout biais de quantification à l'étape n influe sur le biais qui existera à l'étape n + 1 (report d'erreur cumulatif).

Ainsi il est intuitif que le quantificateur ne fonctionne pas comme une simple non linéarité, mais plutôt comme une non linéarité à mémoire.

Les premiers travaux sur le problème posé par l'effet des quantificateurs concluaient que toute adaptation cessait dès que la valeur moyenne des composantes du vecteur d'actualisation $2\mu e.\underline{X}$ devenait inférieur au demi-pas de quantification. On pouvait alors déduire un choix pour le compromis (facteur de gain, précision requise pour les accumulateurs), dans le cas de l'algorithme LMS appliqué à l'égaliseur adaptatif [4].

Ce choix était dicté à l'aide de modèles validés par des simulations. Cependant on n'avait pas encore analysé l'effet du quantificateur présent dans la "boucle d'asservissement des coefficients".

Dans un travail récent [5], une analyse des effets de quantification a été effectuée, ce travail sous certaines hypothèses permet de prendre en compte au niveau de l'erreur quadratique de sortie :

a les effets dus à la quantification des données d'entrée

b les erreurs dues aux erreurs d'arrondi dans le calcul de la sortie du filtre

c les erreurs dues à la différence qui existe entre les valeurs des coefficients en précision finie et en précision infinie.

On peut aussi noter les travaux [6] et [7], dans lesquels est décrite une procédure de quantification aléatoire des coefficients permettant d'éviter le phénomène de blocage de l'algorithme et d'assurer la convergence en moyenne vers le vecteur de Wiener.

On trouve aussi dans [6] et [7] le calcul de l'estimation de la dégradation de la puissance d'erreur de sortie, ainsi que le calcul du bruit sur les coefficients dans l'hypothèse de la seule quantification des coefficients. Citons aussi les

* Ce travail a été accompli dans le cadre du GRECO 69 (SARTA) et fait partie de la thèse de doctorat de H. Dedieu.

travaux [8], qui développent un procédé de linéarisation des quantificateurs semblable à celui présenté dans [6] et [7], conduisant à des constatations qualitatives similaires.

Avec des hypothèses semblables à celles adoptées en [5], nous développons dans la suite un méthodologie différente, afin d'estimer la dégradation appportée par la quantification, non seulement au niveau du bruit de sortie de l'algorithme, mais aussi au niveau du bruit existant sur le vecteur des coefficients, ceci dans le but de modéliser plus facilement les régimes bloqué ou non bloqué de l'algorithme.

II METHODOLOGIE DE L'ETUDE

Considérons l'algorithme donné par (1) :

$$\underline{W}_Q(n+1) = \underline{W}_Q(n) + Q\left[2\mu . Q(e(n))\, Q[\underline{X}(n)]\right] \quad (1)$$

dans lequel Q symbolise un quantificateur uniforme.

On peut montrer que d'une façon équivalente (1) , sous une hypothèse de non débordement peut être présenté sous la forme (2) :

$$\underline{W}(n+1) = \underline{W}_Q(n) + 2\mu\, Q[e(n)] . Q[\underline{X}(n)]$$
$$\underline{W}_Q(n+1) = Q[\underline{W}(n+1)] \quad (2)$$

Si nous notons $\tilde{\underline{W}}(n)$ le vecteur coefficient produit par l'algorithme qui travaillerait en précision infinie , la récurrence sur $\tilde{\underline{W}}(n)$ s'écrit :

$$\tilde{\underline{W}}(n+1) = \tilde{\underline{W}}(n) + 2\mu\, \tilde{e}(n)\, \underline{X}(n) \quad (3)$$

où $\tilde{e}(n)$ est l'erreur en précision infinie.

On peut mettre en évidence au niveau du vecteur de filtrage deux types d'erreurs différentes :

$$\underline{W}_Q(n+1) - \underline{W}_{opt} =$$
$$\left[\tilde{\underline{W}}(n+1) - \underline{W}_{opt}\right] + \left[\underline{W}_Q(n+1) - \underline{W}(n+1)\right] + \left[\underline{W}(n+1) - \tilde{\underline{W}}(n+1)\right] \quad (4)$$

L'écart qui existe à une étape de calcul donnée entre le vecteur coefficient quantifié et le vecteur de Wiener, est donc la somme de l'écart du vecteur coefficient non quantifié, obtenu à la même étape de calcul par le même algorithme fonctionnant sur les mêmes données (non quantifiées), avec le vecteur de Wiener, et de deux termes d'erreurs :

- le premier type d'erreur $[\underline{W}_Q(n+1) - \underline{W}(n+1)]$ est un simple terme d'erreur d'arrondi $Q[\underline{W}_{n+1}] - \underline{W}_{n+1}$

- le deuxième type d'erreur de la forme $[\underline{W}_{n+1} - \tilde{\underline{W}}_{n+1}]$ est la différence qui existe entre les deux vecteurs $\underline{W}(n+1)$ et $\tilde{\underline{W}}(n+1)$:

$\underline{W}(n+1)$ étant généralement un vecteur quantifié beaucoup plus finement que $\underline{W}_Q(n+1)$.

$\tilde{\underline{W}}(n+1)$ étant le vecteur obtenu par le même algorithme fonctionnant avec les mêmes données, mais en précision infinie.

L'objet de l'étude qui suit est de faire l'analyse de ces deux effets, on s'interessera au calcul des deux premiers moments du vecteur $[\underline{W}_Q(n+1) - \underline{W}_{opt}]$ De cette analyse on déduira le comportement de l'erreur quadratique moyenne .

III ETUDE COMPLETE DU REPORT DE L'ENSEMBLE DES BRUITS DE QUANTIFICATION

III-1) LES HYPOTHESES

Hypothèse 1 : On suppose que les seules multiplications introduisent des erreurs, on suppose qu'il n'y a pas dépassement et que les additions n'introduisent pas d'erreur.

Hypothèse 2 : On suppose que tous les bruits de quantification sont indépendants entre eux, sont indépendants avec toutes les autres variables de l'algorithme, sont indépendants composante à composante.

Hypothèse 3 : Les données x(n) et d(n) sont supposées stationnaires, de plus les composantes $x(n+1-j) \triangleq x_j(n)$ du vecteur $\underline{X}(n)$ sont indépendantes entre elles.

Hypothèse 4 : Le vecteur $d\underline{W}(n) = \underline{W}(n) - \underline{W}_{opt}$ est indépendant de $\underline{X}(n)$. Cette hypothèse semble raisonnable, elle est en tous cas nécessaire pour conduire les calculs, on trouvera une justification théorique dans [9].

Dans la suite, nous convenons de désigner par α les bruits ayant pour origine une quantification des coefficients (quantificateur Qc), par β les bruits ayant pour origine une quantification de données ou d (quantificateur Qd), par γ un bruit de quantification de l'erreur (quantificateur Qe).

Soit l'algorithme :

$$\underline{W}_{Qc}(n+1) = \underline{W}_{Qc}(n) + Q_c\left\{2\mu\, Q_d[\underline{X}(n)]\, Q_e\left[d_{Qd}(n) - Q_e\left[\underline{W}^T_{Qc}(n) . Q_d[\underline{X}(n)]\right]\right]\right\} \quad (5)$$

avec

Qc désignant le quantificateur agissant sur les coefficients [Bc bits]
Qd désignant le quantificateut agissant sur les données [Bd bits]
Qe désignant le quantificateur agissant sur l'erreur [Bc ou Bd bits]

Nous supposons par ailleurs $B_c \gg B_d$, cette hypothèse est naturelle dans la mesure où c'est la quantification au niveau des coefficients qui apporte la majeure dégradation à la puissance de l'erreur de sortie.

D'autre part la plupart des microprocesseurs de signal actuels travaillent au moins sur des registres seize bits, alors que généralement des convertisseurs douze bits sont adoptés pour discrétiser les signaux d'entrée.

Analysons tout d'abord l'effet du quantificateur Qe, l'erreur quantifiée s'écrit dans (5) :

$$e_Q(n) = Q_e\left[d_{Qd}(n) - Q_e\left[\underline{W}^T_{Qc}(n) . Q_d[\underline{X}(n)]\right]\right] \quad (6)$$

Comme $B_c \gg B_d$

$$e_Q(n) = d_{Qd}(n) - Q_e\left[\underline{W}^T_{Qc}(n) . Q_d[\underline{X}(n)]\right] \quad (7)$$

Au niveau de (7), il nous faut faire des hypothèses sur la structure de calcul du produit scalaire

Hypothèse dite de double précision :

Si l'on suppose que tous les produits de $\underline{W}^T_{Qc}(n) . Q_d[\underline{X}(n)]$ sont effectués en "double précision", la somme étant quantifiée sur Be = Bc bits ou Be = Bd bits , (7) s'écrit alors :

$$e_Q(n) = d_{Qd}(n) - \underline{W}^T_{Qc}(n) . Q_d[\underline{X}(n)] + \gamma(n) \quad (8)$$

soit

$$e_Q(n) = d(n) + \beta_2(n) - \underline{W}^T_{Qc}(n) . [\underline{X}(n) + \underline{\beta}_1(n)] + \gamma(n) \quad (9)$$

donc

$$e_Q(n) = \left[d(n) - \underline{W}^T_{Qc}(n)\, \underline{X}(n)\right] - \underline{W}^T_{Qc}(n)\, \underline{\beta}_1(n) + \beta_2(n) + \gamma(n) \quad (10)$$

Avec les définitions suivantes :

$\underline{\beta}_1(n)$ est le vecteur des arrondis d'erreur sur le vecteur $\underline{X}(n)$ quantifié sur B_d bits.

$\beta_2(n)$ est l'erreur d'arrondi sur les données d(n) quantifiées sur B_d bits.

$\gamma(n)$ est l'erreur d'arrondi sur la somme résultant du produit scalaire $\underline{W}^T_{Q}(n).\underline{X}(n)$ quantifiée sur B_e bits.

Hypothèse dite de simple précision :

Si l'on suppose que tous les produits de $W^T_{Qc}(n).Q_d[\underline{X}(n)]$ sont quantifiés sur Bc bits, la somme étant quantifiée sur Be = Bc bits ou Be = Bd bits, (7) devient :

$$e_Q(n) = d_{Qd}(n) - \underline{W}^T_{Qc}(n).Q_d[\underline{X}(n)] + \sum_{i=1}^{N} \gamma_i(n) \qquad (11)$$

soit

$$e_Q(n) = d(n) + \beta_2(n) - \underline{W}^T_{Qc}(n)\left[\underline{X}(n) + \underline{\beta}_1(n)\right] + \sum_{i=1}^{N} \gamma_i(n) \qquad (12)$$

donc

$$e_Q(n) = \left[d(n) - W^T_{Qc}(n)\underline{X}(n)\right] - \underline{W}^T_{Qc}(n)\underline{\beta}_1(n) + \beta_2(n) + \sum_{i=1}^{N}\gamma_i(n) \qquad (13)$$

Avec les définitions suivantes :

$\underline{\beta}_1(n)$ est le vecteur des arrondis d'erreur sur le vecteur _(n) quantifié sur Bd bits.

$\beta_2(n)$ est l'erreur d'arrondi sur les données d(n) quantifiées sur B_d bits.

$\gamma_i(n)$ est l'erreur d'arrondi qui entâche chacun des N produits constituant le produit scalaire, chacun étant quantifié sur B_e bits.

III-2) CALCUL DES EQUATIONS DE RECURRENCE SUR W

Posons :

$$\underline{u}_n = Q_c\left[2\mu\, Q_d[\underline{X}(n)]\, e_Q(n)\right] \qquad (14)$$

Sous les hypothèses habituelles faites sur les quantificateurs il vient :

$$\underline{u}_n = 2\mu.\left[\underline{X}(n) + \underline{\beta}_1(n)\right].e_Q(n) + \underline{\alpha}(n) \qquad (15)$$

$\underline{\alpha}(n)$ étant le bruit d'arrondi sur les coefficients.

Si nous plaçons sous l'hypothèse double précision, il vient en reportant (10) dans (15) :

$$\begin{aligned} \underline{u}_n = {} & 2\mu[\underline{X}(n)]\left\{e_{opt}(n) - d\underline{W}^T_{Qc}(n).\underline{X}(n)\right\} \\ & + 2\mu[\underline{X}(n)].\left\{\gamma(n) - d\underline{W}^T_{Qc}(n).\underline{\beta}_1(n) - \underline{W}^T_{opt}\underline{\beta}_1(n) + \beta_2(n)\right\} \\ & + 2\mu\,\underline{\beta}_1(n)\left\{e_{opt}(n) - d\underline{W}^T_{Qc}(n)\underline{X}(n) + \gamma(n) - d\underline{W}^T_{Qc}(n)\underline{\beta}_1(n) \right. \\ & \left. - \underline{W}^T_{opt}.\underline{\beta}_1(n) + \beta_2(n)\right\} + \underline{\alpha}(n) \end{aligned} \qquad (16)$$

Dans l'hypothèse simple précision, il suffirait de remplacer dans (16), $\gamma(n)$ par $\sum_{i=1}^{N}\gamma_i(n)$.

Par un jeu d'écriture relativement simple mais fastidieux, dont on trouvera le détail dans [10], on peut exprimer le vecteur sous la forme récurrente :

$$d\underline{W}_{Qc}(n+1) = A\, d\underline{W}_Q(n) + B + C \qquad (17)$$

avec :

$$A = \mathbb{I} - 2\mu\left\{\underline{X}(n)\underline{X}^T(n) + \underline{X}(n)\underline{\beta}_1^T(n) + \underline{\beta}_1(n)\underline{X}^T(n) + \underline{\beta}_1(n)\underline{\beta}_1^T(n)\right\} \qquad (18)$$

$$B = 2\mu\, e_{opt}(n).\underline{X}(n) \qquad (19)$$

$$\begin{aligned} C = {} & 2\mu\,\underline{X}(n)\left[\gamma'(n) - \underline{W}^T_{opt}.\underline{\beta}_1(n) + \beta_2(n)\right] \\ & + 2\mu\,\underline{\beta}_1(n)\left\{e_{opt}(n) + \gamma'(n) - \underline{W}^T_{opt}\underline{\beta}_1(n) + \beta_2(n)\right\} + \underline{\alpha}(n) \end{aligned} \qquad (20)$$

Avec dans (20), si l'on se place dans l'hypothèse double précision :

$$\gamma'(n) = \gamma(n) \qquad (21a)$$

et si l'on se place dans l'hypothèse simple précision

$$\gamma'(n) = \sum_{i=1}^{N}\gamma_i(n) \qquad (21b)$$

III-2) ETUDE DU MOMENT D'ORDRE 1 DE dWq

A partir de l'équation récurrente (17), il est aisé [10], compte tenu des hypothèses 1,2,3 et 4 d'obtenir la récurrence sur le vecteur $E[d\underline{W}_{Qc}(n)]$, soit :

$$E[d\underline{W}_{Qc}(n+1)] = \mathbb{I} - 2\mu\left\{R + \sigma_\beta^2\mathbb{I}\right\}E[d\underline{W}_{Qc}(n)] - 2\mu\sigma_\beta^2\,\mathbb{I}\,\underline{W}_{opt} \qquad (22)$$

Sous la condition $0<\mu<1/\lambda_{max}$, l'équation (22) converge vers la quantité (23), qui exprime le biais existant sur le régime limite .

$$E[d\underline{W}_{Qc}(\infty)] = -\sigma_\beta^2\left\{R + \sigma_\beta^2\mathbb{I}\right\}^{-1}\underline{W}_{opt} \qquad (23)$$

Dans l'expression (22), σ_β^2 mesure la puissance du bruit de quantification sur les données X(n) et d(n) quantifiées sur B_d bits.

III-3) ETUDE DU MOMENT D'ORDRE 2 DE dWq

Posons :

$$M_{W_Q}(n+1) = d\underline{W}_{Qc}(n+1).d\underline{W}^T_{Qc}(n+1) \qquad (24)$$

A partir de (17), (24) s'écrit :

$$\begin{aligned} M_{W_Q}(n+1) = {} & A\,M_{W_Q}(n)A^T + \underline{B}\,dW^T_Q(n)A^T \\ & + \underline{C}\,d\underline{W}^T_{Qc}(n)A^T + A\,d\underline{W}_{Qc}(n)\underline{B}^T \\ & + BB^T + CB^T + A\,d\underline{W}_{Qc}(n)C^T + BC^T + CC^T \end{aligned} \qquad (25)$$

III-3-1) *Etude de la quantité* $A\,M_{W_Q}(n)A^T$ *dans (25)*

D'après (18) :

$$\begin{aligned} A\,M_{W_Q}(n)A^T = {} & M_{W_Q}(n) \\ & - 2\mu\left[H\,M_{W_Q}(n) + M_{W_Q}(n)H^T\right] \\ & + 4\mu^2\, H\,M_{W_Q}(n)H^T \end{aligned} \qquad (26)$$

avec

$$H = H^T \triangleq \left[X(n)X^T(n) + X(n)\beta_1^T(n) + \beta_1(n)X^T(n) + \beta_1(n)\beta_1^T(n)\right] \qquad (27)$$

On trouvera dans [10], les calculs relativement lourds qui permettent de déterminer le terme général de la matrice (26), soit :

si $i \neq j$

$$E\left[A\,M_{W_Q}(n)A^T\right](i,j) = \left\{1 - 4\mu(\sigma_x^2+\sigma_\beta^2) + 8\mu^2(\sigma_x^4 + 2\sigma_x^2\sigma_\beta^2 + \sigma_\beta^4)\right\}E[M_{W_Q}(i,j)] \qquad (28)$$

et pour les termes diagonaux :

$$\begin{aligned} E\left[A\,M_{W_Q}(n)A^T\right](i,i) = {} & \left\{1 - 4\mu(\sigma_x^2+\sigma_\beta^2) + 4\mu^2\left[(N-1)(\sigma_x^4+\sigma_\beta^4) + (2N+4)\sigma_x^2\sigma_\beta^2\right.\right. \\ & \left.\left. + \gamma_{4x} + \gamma_{4\beta}\right]\right\}.\,E\left[M_{W_Q}(n)\right](i,i) \end{aligned} \qquad (29)$$

Les quantités γ_{4x} et $\gamma_{4\beta}$ représentant respectivement les moments d'ordre 4 des données x et des bruits de quantification de ces mêmes données.

III-3-2) *Etude de la quantité* $B\,dW^T_Q(n)A^T$ *dans (25)*

D'après (18) et (19), et en supposant l'orthogonalité entre l'erreur de Wiener et les données X(n), on a :

$$E\left[\underline{B}.d\underline{W}^T_Q(n)A^T\right] = E\left[A\,dW_Q(n)\underline{B}^T\right] = 0 \qquad (30)$$

III-3-3) *Etude de la quantité* $C\,dW^T_Q(n)A^T$ *dans (25)*

D'après (18) et (20), on peut montrer que les termes non nuls de $E[\underline{C}\,dW^T_Q(n)A^T](i,j)$ sont proportionnels à $\sigma_\beta^2 E[dW^T_Q(n)(i)].W_{opt}(j)$, soit d'après (22) à σ_β^4, on donc :

$$E\left[\underline{C}\,d\underline{W}^T_Q(n)A^T\right] = E\left[A\,d\underline{W}_Q(n)\underline{C}^T\right] = 0 \qquad (31)$$

III-3-4) *Etude de la quantité* BB^T *dans (25)*

D'après (19) :

$$E\left[BB^T\right] = 4\mu^2\,\xi_{min}.R \qquad (32)$$

III-3-5) Etude de la quantité CB^T dans (25)

D'après (19)et (20) :

$$[CB^T] = 4\mu^2 \underline{X}\,\underline{X}^T e_{opt}(n)\left\{\gamma'(n) - \underline{W}_{opt}^T \cdot B_1(n) + B_2(n)\right\}$$
$$+ 4\mu^2 \underline{B}_1 \underline{X}^T e_{opt}(n)\left\{e_{opt}(n) + \gamma'(n) - \underline{W}_{opt}^T B_1(n) + B_2(n)\right\} \quad (33)$$
$$+ 2\mu\,\underline{\alpha}(n)\,X^T e_{opt}(n)$$

Du principe d'orthogonalité de l'erreur de Wiener avec $\underline{X}(n)$ et des différentes hypothèses d'indépendance résultent l'égalité :

$$E[CB^T] = E[BC^T] = 0 \quad (34)$$

III-3-6) Etude de la quantité $\underline{C}\,\underline{C}^T$ dans (25)

D'après (20) :

$$E[\underline{C}\,\underline{C}^T] = 4\mu^2 E\left[\underline{X}\,\underline{X}^T\left\{\gamma'(n) - \underline{W}_{opt}^T B_1(n) + B_2(n)\right\}\right]$$
$$+ 4\mu^2 E\left\{[\underline{X}\underline{B}_1^T + \underline{B}_1\underline{X}^T]\left(e_{opt}(n) + \gamma'(n) - \underline{W}_{opt}^T B_1(n)\right)\left(\gamma'(n) - W_{opt}B_1(n) + B_2(n)\right)\right\}$$
$$+ E[\underline{\alpha}(n)\underline{\alpha}^T(n)] \quad (35)$$

soit

$$E[\underline{C}C^T] = 4\mu^2\sigma_{\gamma'}^2 R + 4\mu^2|W_{opt}|^2\sigma_\beta^2 R$$
$$+ 4\mu^2\sigma_\beta^2 R + \sigma_\alpha^2 \mathbb{I}$$
$$+ 4\mu^2\left\{\xi_{min} + \sigma_{\gamma'}^2 + \sigma_\beta^2\right\}\sigma_\beta^2 \mathbb{I} \quad (36)$$

III-3-5) *Equation de récurrence sur les termes diagonaux de la matrice (24)* :

Dans la suite, seuls nous interesserons les régimes limites des termes diagonaux de (24), que nous pourrons exprimer en tenant compte des résultats intermédiaires (29), (30), (31), (34) et (36) par :

$$E[M_{W_Q}(\infty)](i,i) = \mu \cdot \frac{\left\{\xi_{min}(\sigma_x^2+\sigma_\beta^2) + \sigma_x^2\left(\sigma_{\gamma'}^2 + |W_{opt}|^2\sigma_\beta^2 + \sigma_\beta^4\right) + \sigma_\alpha^2/4\mu^2\right\}}{(\sigma_x^2+\sigma_\beta^2) - \mu\left\{(N-1)(\sigma_x^4+\sigma_\beta^4) + (2N+4)\sigma_x^2\sigma_\beta^2 + \gamma_{4x} + \gamma_{4\beta}\right\}} \quad (36)$$

III-4) REPORT SUR L'ERREUR DE SORTIE

Ecrivons l'erreur de sortie de l'algorithme :

$$e(n) = d(n) - Q_e\left[W_{Qc}^T(n) X_d(n)\right] \quad (37)$$

soit encore :

$$e(n) = d(n) - \underline{W}_{Qc}^T(n)\left[\underline{X}_d(n)\right] \quad (38)$$

$$e(n) = d(n) - \underline{W}_{Qc}^T(n)\left[\underline{X}(n) + B_1(n)\right] + \gamma'(n) \quad (39)$$

avec $\gamma'(n)$ donné par (21a) ou (21b) suivant l'hypothèse double ou simple précision dans laquelle on s'est placé.

(39) s'écrit après quelques manipulations :

$$e(n) = e_{opt}(n) - \underline{dW}_{Qc}^T(n)\left[\underline{X}(n) + \underline{B}_1(n)\right] - \underline{W}_{opt}\underline{B}_1(n) + \gamma'(n) \quad (40)$$

L'erreur quadratique moyenne de sortie de l'algorithme est alors donnée par :

$$E[e^2(n)] = \xi_{min} + E\left[\underline{X}^T(n) M_{W_Q}(n)\underline{X}(n)\right]$$
$$+ E\left[\underline{B}_1^T(n) M_{W_Q}(n)\underline{B}_1(n)\right] + |\underline{W}_{opt}|^2\sigma_\beta^2 + \sigma_{\gamma'}^2 \quad (41)$$

En développant (41) il vient :

$$E[e^2(n)] = \xi_{min} + \left(\sum_{i=1}^{N} M_{W_Q}(n)(i,i)\right)\cdot\left(\sigma_x^2 + \sigma_\beta^2\right)$$
$$+ |W_{opt}|^2\sigma_\beta^2 + \sigma_{\gamma'}^2 \quad (42)$$

En reportant dans (42), l'équation (36) on peut écrire le régime limite de l'erreur quadratique soit

$$\xi = \xi_{min} + \frac{N\mu\sigma_x^2\,\xi_{min}}{1 - D/\sigma_x^2}$$
$$+ \mu N\,\frac{\left\{\sigma_\alpha^2/4\mu^2 + \xi_{min}\sigma_\beta^2\left[2 + \frac{\sigma_\beta^2}{\sigma_x^2}\right] + \sigma_x^2\left[\sigma_{\gamma'}^2 + |W_{opt}|^2\sigma_\beta^2 + \sigma_\beta^4\right]\right\}}{1 - D/\sigma_x^2} \quad (43)$$
$$+ |W_{opt}|^2\sigma_\beta^2 + \sigma_{\gamma'}^2$$

avec :

$$D = \mu\left\{(N-1)(\sigma_x^4 + \sigma_\beta^4) + (2N+4)\sigma_x^2\sigma_\beta^2 + \gamma_{4x} + \gamma_{4\beta}\right\} - \sigma_\beta^2 \quad (44)$$

Il est clair, que dans l'équation (43), si l'on raisonne sur un même nombre de bits pour les données et les coefficients la dégradation majeure est apportée par la quantification des coefficients, le terme $(\sigma_\alpha^2/4\mu^2)/(1 - D/\sigma_x^2)$ étant très largement dominant devant tous les autres termes provenant de bruits de quantification. D'autre par ce terme d'erreur est inversement proportionnel au facteur de gain, si bien que dès que l'on pilotera l'algorithme par un facteur de gain plus petit que le facteur de gain nominal la puissance de l'erreur de sortie croîtra. On trouvera des illustrations de ce phénomènè dans [5], [6] et [7].

III-5) CONDITION DE NON BLOCAGE DE L'ALGORITHME

En l'absence de toute quantification, l'équation (36) donnerait :

$$E[M_W(\infty)] \simeq \frac{\mu\,\xi_{min}\,\sigma_x^2}{\sigma_x^2 - \mu(N-1)\sigma_x^4} \quad (45)$$

Une condition de non-blocage de l'algorithme est que la puissance du bruit d'arrondi des composantes du vecteur coefficient soit inférieure à la quantité (45), soit :

$$\sigma_\alpha^2 < \frac{\mu\,\xi_{min}}{1 - \mu(N-1)\sigma_x^2} \quad (46)$$

Connaissant un encadrement de la valeur ξ_{min} on peut estimer la valeur minimale du facteur de gain à ne pas dépasser, soit :

$$\mu > \frac{\sigma_\alpha^2}{\xi_{min} + (N-1)\sigma_x^2\sigma_\alpha^2} \quad (47)$$

Si l'on applique la formule de Sheppard, (47) est parfaitement calculable pour les applications pratiques (Ex : Quantification déterministe $\sigma_\alpha^2 = \frac{q^2}{12}$)

IV CONCLUSION

On a pris en compte l'ensemble des effets de quantification de l'algoritme LMS.

Sous des hypothèses simples, on a donné les équations récurrentes du premier et du second moment de l'écart du vecteur coefficient courant au vecteur de Wiener.

On a exprimé une condition de non-blocage de l'algorithme, valeur minimale du facteur de gain à ne pas dépasser.

Enfin la valeur de l'erreur quadratique a été calculée -(43)-, après simplification de (43) on corrobore les résultats donnés par [5], à l'aide d'hypothèses semblables , mais avec une approche du problème différente.

Bibliogaphie

[1] J.L MOSCHNER, "Adaptive Filter with Clipped Data", Tech. Rpt. 6796-1, Informations Systems Laboratory, Stanford University, Stanford, CA, June 1970.

[2] T. A. C. M. CLAASEN, W. F. G. MECKLENBRÄUKER, "Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters", IEEE Trans. on ASSP, Vol. ASSP-29, June 1981, pp 670-678

[3] D. L. DUTTWEILER, "Adaptive Filter Performance with Non-linearities in the Correlation Multiplier",
IEEE Trans. on ASSP, Vol ASSP-30, N° 4, August 1982.

[4] R. D. GITLIN, S.B. WEINSTEIN, "On the required Tap-Weight Precision for Digitally Implemented Adaptive Filters, The Bell Systems Technical Journal, Vol. 58, pp 301-321, February 1979

[5] C. CARAISCOS and B. LIU, "A Round-Off Error Analysis of the LMS Adaptive Algorithm", IEEE on ASSP, Vol. ASSP-32, N° 1, pp 34-41, February 84.

[6] F. CASTANIE, "Quantification Aléatoire"
Annales des Télécommunications, 41, N°5-6, Mai-Juin 1986.

[7] H. DEDIEU, F. CASTANIE, "Adaptive Filter Improvment Using Randomly Quantized Coefficients", Proc. of ICASSP 85, Tampa, USA.

[8] D. T. SHERWOOD and N. J. BERSHAD, "Quantization Noise Effects in the Complex LMS Algorithm - Linearization with Dither"
Proc. of ICASSP 85, Tampa, USA.

[9] O. MACCHI, "Le Filtrage Adaptatif en Télécommunications", Annales des Télécommunications, 36, N°11-12, Nov.-Décembre 1981.

[10] H. DEDIEU, "Prise en Compte de l'Ensemble des Effets de Quantification dans l'Algorithme LMS", Rappt. Int. GAPSE N°1, 87.