

INFLUENCE DES CARACTERISTIQUES DES REFERENCES-BRUIT SUR LES PERFORMANCES DU SPECTROFILTRE ADAPTATIF

J.-F. GUERRE-CHALEY D. BAUDOIS
C. SERVIERE A. SILVENT (*)

CEPHAG Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble
B.P. 46 38402 SAINT MARTIN D'HERES FRANCE
(*) Attaché aux services techniques des Armées

RÉSUMÉ

Nous montrons que, dans le cas du spectrofiltre adaptatif multivariable, l'utilisation de l'hypothèse d'orthogonalité des références-bruit entre elles diminue considérablement le volume des calculs du traitement. Mais alors le système de soustraction est moins performant vis-à-vis du critère d'optimalité utilisé (erreur quadratique moyenne minimale). Après avoir introduit le problème, nous en donnons la solution générale en présentant les estimateurs correspondants avec facteur d'oubli. Nous voyons ensuite comment utiliser la propriété d'orthogonalité, les résultats théoriques étant validés par des simulations. Nous montrons également la nécessité d'un compromis entre le nombre de références-bruit utilisées et leur qualité caractérisée par leur cohérence avec la voie signal.

ABSTRACT

We show that, in the case of the adaptive multivariate spectrofilter, the utilisation of jointly orthogonal references property involves a great drop off of the computation cost. But the subtractor is then less performant, over against the optimality criteria. In a first part the problem is introduced, then the general solution is given, with the representation of the estimators using a forgetting factor. The utility of the orthogonality property is shown, the theoretical results are validate by simulations. We also show the requisite compromise between the number of references and the information that they bring on noise.

I INTRODUCTION

La soustraction de bruit consiste à identifier, puis à reconstruire, un filtre qui, excité par une référence-bruit, fournit une estimation du bruit perturbateur. Cette estimation retranchée au signal bruité procure une estimation du signal utile [1].

Les canaux de transmission, entre sources de bruit - capteur de signal utile d'une part et, source de bruit - capteur de référence d'autre part, font subir aux signaux des transformations linéaires ou non que l'on peut modéliser par des filtres. Les méthodes classiques consistent à chercher l'estimateur de ces filtres dans la classe des filtres linéaires.

Une technique particulière d'estimation consiste à reconstituer le filtre dans le domaine fréquentiel à l'aide d'un algorithme nommé spectrofiltre [2][3]. Il s'agit d'un algorithme adaptatif non bouclé basé sur la formule de Wiener [4].

L'algorithme du spectrofiltre traite les différents canaux de fréquence indépendamment les uns des autres. Les signaux sont discrets et issus de Transformées de Fourier Rapides (TFR). De façon générale $X(k)$ représente le résultat de la TFR du bloc de rang k du signal à temps discret $x(i)$, pour la fréquence discrète ν (qui n'est pas notée pour simplifier l'écriture).

Le signal observé $Y(k)$ est la somme du signal utile $S(k)$ et du bruit perturbateur $N(k)$:

$$(1) \quad Y(k) = S(k) + N(k)$$

L'utilisateur dispose de n références-bruit scalaires, respectivement $R_1(k), R_2(k) \dots R_n(k)$, indépendantes du signal utile $S(k)$ et corrélées avec $N(k)$. Elles sont notées sous forme vectorielle :

$$(2) \quad \underline{R}(k) \triangleq [R_1(k) R_2(k) \dots R_n(k)]^T$$

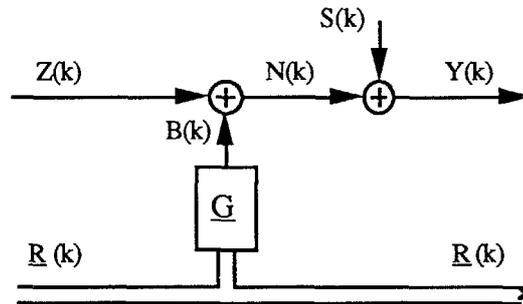


figure 1 : modélisation fréquentielle des différents signaux

Nous appelons $Z(k)$ la composante du bruit $N(k)$ non corrélée avec $\underline{R}(k)$ et $B(k)$ la composante de $N(k)$ totalement corrélée avec $\underline{R}(k)$. Lorsque la dépendance entre la référence-bruit $\underline{R}(k)$ et le bruit $B(k)$ est linéaire, elle peut être modélisée par un filtre \underline{G} qui engendre $B(k)$ à partir de $\underline{R}(k)$ (figure 1).

$$(3) \quad N(k) = Z(k) + B(k) \quad (4) \quad B(k) = \underline{G}^+ \underline{R}(k)$$

On suppose que tous les signaux sont centrés et suivent une loi normale complexe [5][6]. Cette hypothèse intervient pour le calcul de moments d'ordre 4 et trouve sa justification dans les propriétés de la TFR [7] lorsque la longueur des blocs sur lesquels sont réalisées les TFR est suffisamment importante [8].

Les signaux S, \underline{R} et Z sont supposés indépendants, ce qui est vérifié en pratique.

Enfin, on suppose que les signaux, à chaque fréquence ν , issus de blocs temporels de rangs différents sont indépendants entre eux, à savoir :

$$(5) \quad E[\underline{U}(k_1) \underline{V}^+(k_2)] = \underline{\gamma}_{UV} \delta_{k_1-k_2}$$

où $E[X]$ est l'espérance mathématique de X

et où δ_θ représente le symbole de Kronecker (égal à 1 si $\theta=0$, nul sinon), et où $\underline{\gamma}_{UV}$ représente la densité spectrale de



puissance d'interaction du signal \underline{U} avec le signal \underline{V} , à la fréquence ν .

L'hypothèse (5) est admissible lorsque la longueur des blocs temporels sur lesquels sont effectuées les TFR est grande vis-à-vis du support de corrélation des signaux.

A partir des signaux d'entrée $Y(k)$ et $R(k)$ nous cherchons à estimer $B(k)$ en filtrant $R(k)$ à l'aide d'un filtre \hat{G} appartenant à la classe des transformations linéaires (c'est un estimateur de G quand ce dernier est linéaire). Ensuite, l'estimée $\hat{B}(k)$ du bruit $B(k)$ est retranchée au signal bruité $Y(k)$ pour fournir une estimation $\hat{S}(k)$ du signal utile $S(k)$.

Pour chaque canal de fréquence ν nous avons le schéma de soustraction suivant, figure 2 :

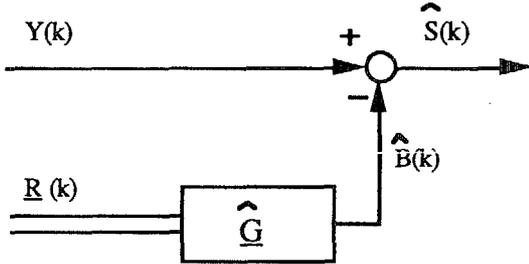


figure 2 : soustraction de bruit dans le domaine fréquentiel

(6) $\hat{B}(k) = \hat{G}^+ R(k)$ (7) $\hat{S}(k) = Y(k) - \hat{B}(k)$

II SOLUTION GENERALE MATRICIELLE

L'expression du gain \hat{G} est obtenue en minimisant l'Erreur Quadratique Moyenne (EQM) en sortie de traitement, c'est-à-dire, à une fréquence donnée ν :

(8) $EQM_{\text{sortie}} = E[|\hat{S}(k) - S(k)|^2]$

On obtient alors classiquement la solution de Wiener, dans le cas où $\det(\underline{\gamma}_R)$ est non nul :

(9) $\hat{G}_{\text{opt}} = (\underline{\gamma}_R)^{-1} \underline{\gamma}_{YR}^+$

où $\underline{\gamma}_R$ représente la matrice autospectrale de la référence-bruit $R(k)$ et $\underline{\gamma}_{YR}$ le vecteur interspectral entre la voie signal bruité $Y(k)$ et la référence-bruit $R(k)$, à la fréquence ν :

(10) $\underline{\gamma}_R \triangleq E[R(k) R^+(k)]$ (11) $\underline{\gamma}_{YR} \triangleq E[Y(k) R^+(k)]$

Les estimateurs sont les suivants :

(12) $\hat{B}(k) = \hat{G}(k-1) R(k)$ (13) $\hat{S}(k) = Y(k) - \hat{B}(k)$

avec :

(14) $\hat{\underline{\gamma}}_R(k) = \alpha \hat{\underline{\gamma}}_R(k-1) + (1 - \alpha) R(k) R^+(k)$

(15) $\hat{\underline{\gamma}}_{YR}(k) = \alpha \hat{\underline{\gamma}}_{YR}(k-1) + (1 - \alpha) Y(k) R^+(k)$

(16) $\hat{G}(k) = (\hat{\underline{\gamma}}_R(k))^{-1} \hat{\underline{\gamma}}_{YR}^+(k)$

Les équations (12) à (16) constituent l'algorithme du spectrofiltre adaptatif avec facteur d'oubli α .

L'erreur d'estimation instantanée peut être écrite :

(17) $\varepsilon(k) = \hat{S}(k) - S(k)$

C'est-à-dire : $\varepsilon(k) = Z(k) - (G - \hat{G}(k-1))^+ R(k)$

La puissance d'erreur d'estimation (ou EQM_{sortie}) est alors :

(18) $P_\varepsilon = E[|\varepsilon(k)|^2]$

Soit : (19) $\delta \hat{G}(k-1) \triangleq \hat{G} - \hat{G}(k-1)$

En utilisant les hypothèses et propriétés énoncées lors de la modélisation du problème il vient :

(20) $P_\varepsilon \cong \gamma_Z + E[\delta \hat{G}^+(k-1) \underline{\gamma}_R \delta \hat{G}(k-1)]$

En supposant que les erreurs d'estimation relatives sont faibles devant 1, il vient [9] :

(21) $P_\varepsilon \cong \gamma_Z + n \frac{1-\alpha}{1+\alpha} (\gamma_Y - \underline{\gamma}_{RY} \underline{\gamma}_R^{-1} \underline{\gamma}_{RY}^+)$

avec (22) $\gamma_Z \triangleq E[|Z(k)|^2]$ et (23) $\gamma_Y \triangleq E[|Y(k)|^2]$

Cette relation est admissible lorsque les erreurs relatives sont faibles, c'est-à-dire lorsque α est proche de 1. Pour α tendant vers 1, ce qui correspond à une fenêtre d'intégration équivalente de longueur tendant vers l'infini, nous obtenons asymptotiquement le terme de puissance d'erreur minimale : γ_Z .

Lorsque les références-bruit sont orthogonales entre elles, a posteriori, l'expression (21) peut être développée et devient :

(24) $P_{\varepsilon g} \cong \gamma_Z + n \frac{1-\alpha}{1+\alpha} \gamma_Y (1 - \sum_{i=1}^n C_{Y,i}^2)$

où $C_{Y,i}$ représente la cohérence entre la référence-bruit scalaire $R_i(k)$ et le signal bruité $Y(k)$.

(25) $C_{Y,i}^2 \triangleq \frac{\gamma_{Yi} \gamma_{Yi}^*}{\gamma_Y \gamma_i}$

(26) $\gamma_i \triangleq E[|R_i(k)|^2]$ (27) $\gamma_{Yi} \triangleq E[Y(k) R_i(k)^*]$

III UTILISATION DE L'HYPOTHESE D'ORTHOGONALITE

Lorsque les références-bruit sont orthogonales entre elles, a priori, il est possible d'utiliser chacune des n références-bruit indépendamment des autres pour un traitement de soustraction de bruit monovariable. Les n soustractions peuvent alors être réalisées en parallèle, figure 3, ou en cascade, figure 4 (pour chaque canal de fréquence).

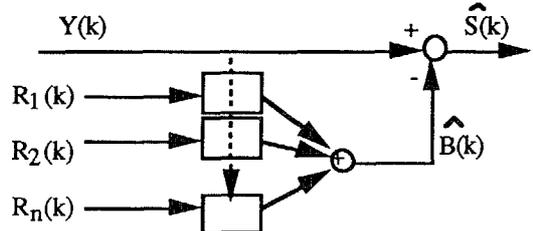


figure 3: structure parallèle

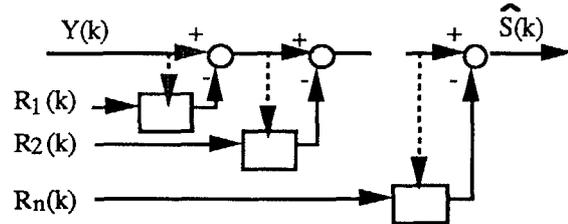


figure 4: structure cascade

Le calcul de P_ε (18) pour les deux structures [9] conduit à :

(28) $P_\varepsilon \cong \gamma_Z + \sum_{i=1}^n \gamma_i E[|G_i(k-1)|^2]$

C'est à dire pour la structure parallèle :

(29) $P_{\varepsilon p} \cong \gamma_Z + \frac{1-\alpha}{1+\alpha} \gamma_Y (n - \sum_{i=1}^n C_{Y,i}^2)$

Pour la structure cascade, un calcul récursif sur le nombre de références des termes $E[|G_i(k-1)|^2]$ permet de déterminer $P_{\varepsilon c}$ [9] :

(30) $P_{\varepsilon c} \cong \gamma_Z + \frac{1-\alpha}{1+\alpha} \gamma_Y (\omega(1) - \sum_{i=1}^n \omega(i) C_{Y,i}^2)$

avec (31) $\omega(i) = \frac{1 - \beta^{n+1-i}}{1 - \beta}$ et (32) $\beta = \frac{2}{1+\alpha}$

Les expressions (29) et (30) montrent que lorsque les n traitements scalaires de soustraction de bruit sont efficaces, $P_{Ec} < P_{Ep}$. La soustraction du bruit liée à une référence bénéficie en effet des apports des soustractions des bruits liés aux références précédentes utilisées.

IV OPTIMALITE DU NOMBRE DE REFERENCES

Les expressions (21), (24), (29) et (30) sont de la forme :

$$(33) P_E = \gamma_Z(n; C_{Y,i}^2) + f(\alpha; n; C_{Y,i}^2)$$

$\gamma_Z(n; C_{Y,i}^2)$ caractérise la puissance d'erreur minimale et tend vers une constante (nulle si les références permettent de caractériser complètement le bruit) lorsque le nombre de références tend vers l'infini.

$f(\alpha; n; C_{Y,i}^2)$ est relatif aux erreurs d'estimation. Il est décroissant selon la longueur de la fenêtre d'intégration, mais croissant selon le nombre de références-bruit utilisées.

Ainsi, il existe un nombre optimal de références-bruit à utiliser (qui rend P_E (33) minimal). Ce nombre est lié aux valeurs de leurs cohérences avec la voie signal. Un compromis doit être réalisé entre le nombre de références (minimal) et leur cohérence généralisée avec la voie signal (maximale).

Afin d'illustrer ce compromis nous présentons une simulation. Le signal utile est perturbé par six bruiteurs indépendants, de puissances décroissantes selon l'indiciage, dont on prend les références à la source. Nous analysons alors le rapport signal à bruit en sortie de traitement (37) en fonction des références prises en compte, avec un traitement matriciel (§II; $\alpha = 0,9$). Les n références utilisées sont représentatives des bruiteurs les plus forts. Ainsi, augmenter n revient à utiliser les références relatives aux bruiteurs plus forts. Nous obtenons alors le graphe figure 5.

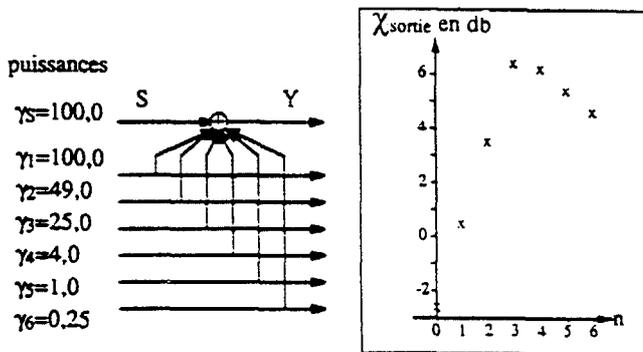


figure 5 : Perturbation et courbe $\chi_{\text{sortie}} = f(n)$

Dans ce cas le nombre optimal de références est trois. Au delà de l'utilisation des trois références les plus cohérentes avec la voie signal, l'emploi d'autres références, même si elles apportent des informations supplémentaires sur le bruit total, dégrade la soustraction. L'utilisation de structures parallèle ou cascade induisent les mêmes conclusions.

V INTERET DE L'UTILISATION DE L'ORTHOGONALITE DES REFERENCES

Le principal intérêt est basé sur la complexité algébrique des structures présentées (figures 2, 3 et 4). La non utilisation de l'hypothèse d'orthogonalité des références-bruit implique un traitement matriciel complexe [10]. Dans le cas de systèmes embarqués, la modification du nombre de références, par exemple, entraîne une modification profonde du système. Par contre, lorsque cette hypothèse est utilisée, les systèmes qui en découlent sont simples à mettre en œuvre car scalaires. Ils sont

également beaucoup plus modulables et le contrôle des singularités (divisions par zéro éventuelles par exemple) est plus aisé à gérer.

Le coût des calculs est également un élément pratique important. Celui-ci est présenté par les tableaux 1 et 2, pour les diverses structures, et comprend le volume des calculs inhérents aux TFR et TFR inverses.

n : nombre de références ; L : ordre des TFR (ordre des filtres)

opération	coût calcul par échantillon de sortie
*	$n (n^2 + 12n + 17)/6 + (n+2) \log_2(L)$
+ -	$n (n^2 + 9n + 8)/6$
/	$n (n + 1)/2$

tableau 1: coût calcul de la solution générale

opération	coût calcul par échantillon de sortie
*	$5n + (n+2) \log_2(L)$
+ -	$3n$
/	n

tableau 2: coût calcul des structures parallèle ou cascade

La solution générale fait intervenir une résolution de système par une technique de Cholesky [10].

Le gain apporté par l'utilisation de l'hypothèse d'orthogonalité des références est immédiat au vu des tableaux présentés. Même si les références sont orthogonalisées par un traitement amont (Gram-Schmidt par exemple), le coût calcul total est nettement inférieur à celui de la solution générale [9].

VI INCONVENIENT DE L'UTILISATION DE L'ORTHOGONALITE DES REFERENCES

Nous allons montrer que les structures parallèle et cascade entraînent une diminution de la qualité de la soustraction, vis-à-vis de la solution générale, en comparant les puissances d'erreur résiduelles correspondantes.

Pour la structure parallèle, il vient :

$$(34) P_{Eg} - P_{Ep} \cong (1-n) \frac{1-\alpha}{1+\alpha} \gamma_Y \sum_{i=1}^n C_{Y,i}^2$$

Ce terme est toujours négatif ou nul. Ainsi le traitement parallèle est systématiquement moins performant que le traitement matriciel, tant que les hypothèses de calcul sont admissibles (erreurs d'estimations relatives faibles devant 1).

La comparaison des expressions (24) et (30) est moins immédiate et nécessite de nouveau une utilisation de l'hypothèse : facteur d'oubli α proche de 1.

Soit $\delta = (1-\alpha)/(1+\alpha)$ ($\delta \ll 1$), alors un développement limité de P_{Ec} conduit à :

$$(35) P_{Eg} - P_{Ec} \cong \delta \gamma_Y \left[-\delta \frac{n(n-1)}{2} - \sum_{i=1}^n C_{Y,i}^2 (i-1) \right]$$

Tout les membres de (35) sont négatifs, à l'exception de celui concernant $C_{Y,1}^2$ qui est positif (développement à un ordre supérieur) mais néanmoins inférieur à $\delta \frac{n(n-1)}{2}$. Ce qui entraîne que, comme (34), la différence (35) est toujours négative (dans la limite des hypothèses). La structure cascade est alors moins performante que la structure matricielle.

Les résultats théoriques sont validés par les simulations suivantes:

La première décrit la variation du rapport signal à bruit en sortie de traitement χ_{sortie} en fonction d'une part du rapport signal à bruit en entrée $\chi_{\text{entrée}}$ et d'autre part du facteur d'oubli α . Le bruiteur Z est de puissance nulle et les rapports signal à bruit sont alors définis par :



$$(36) \chi_{\text{entrée}} = E[|S(k)|^2] / E[|Y(k) - S(k)|^2]$$

$$(37) \chi_{\text{sortie}} = E[|S(k)|^2] / E[|\hat{S}(k) - S(k)|^2]$$

L'un des bruiteurs est deux fois plus puissant que l'autre et les références sont prises à la source de ceux-ci.

Sur la figure 6, sont représentés en trait plein les χ_{sortie} relatifs à la solution générale et, en pointillés ceux relatifs à la structure parallèle, en fonction de $\chi_{\text{entrée}}$ et α .

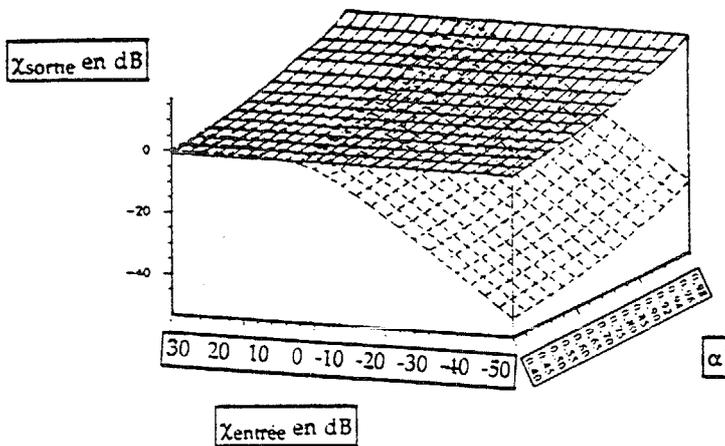


figure 6 : Surface $\chi_{\text{sortie}} = f(\alpha; \chi_{\text{entrée}})$

Les performances de la structure parallèle sont effectivement inférieures à celles de la structure générale. La non correspondance des simulations et de la théorie, pour les faibles valeurs de α ($< 0,5$) et les fortes valeurs de $\chi_{\text{entrée}}$ ($> 20\text{db}$), s'explique par le fait que l'hypothèse consistant à négliger les erreurs relatives d'estimation n'est plus valable.

La seconde simulation présentée met en œuvre 3 références-bruit. La puissance du bruiteur Z est nulle, celle du signal utile est unitaire. Les références sont prises à la source des bruiteurs. La puissance du bruiteur 1 est variable, celle du bruiteur 2 est de 1,0, celle du bruiteur 3 de 0,1. Le rapport signal à bruit en sortie χ_{sortie} (37) est tracé en fonction de la puissance du premier bruiteur γ_1 pour tous les traitements possibles (général, parallèle et les 6 traitements cascades), figure 7. Le facteur d'oubli utilisé est $\alpha = 0,95$.

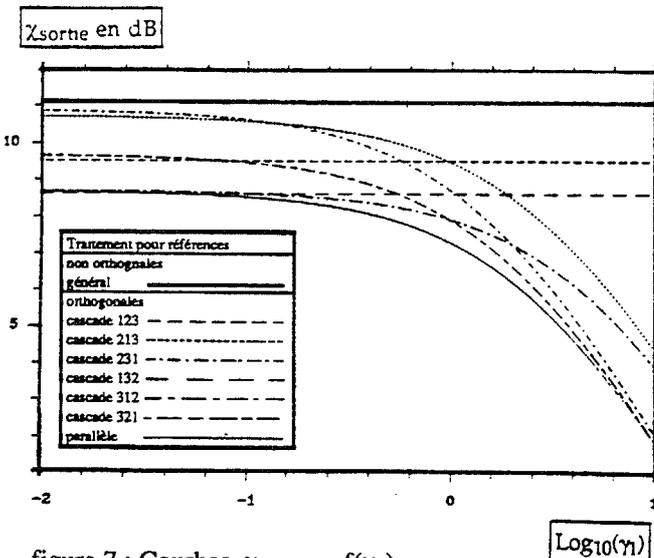


figure 7 : Courbes $\chi_{\text{sortie}} = f(\gamma_1)$

Les simulations confortent la théorie dans la mesure où les traitements utilisant l'hypothèse d'orthogonalité des références sont moins performants que le traitement général.

Il est par ailleurs possible de vérifier les remarques du paragraphe §III concernant la plus grande qualité de soustraction des traitements cascades vis-à-vis du parallèle et, l'ordre d'utilisation des références.

VII CONCLUSION

Les calculs présentés sont relatifs à une technique d'intégration qui met en œuvre un moyennage exponentiel, simple à implanter qui permet en pratique de palier des non stationnarités lentes. Des résultats similaires ont été obtenus théoriquement et en simulation pour d'autres types de moyennages [2] (fenêtre glissante avec ou sans lissage, par exemple).

Après avoir calculé, à convergence des algorithmes, l'EQM_{sortie} relative aux diverses structures envisagées, nous avons montré l'importance d'un compromis à réaliser entre le nombre de références-bruit et la quantité d'information apportée par celles-ci sur le bruit perturbateur, que l'on peut mesurer par leur cohérence généralisée avec la voie signal.

L'utilisation de l'hypothèse d'orthogonalité des références entre elles permet de réduire le volume des calculs.

Lorsque la structure parallèle est employée, les termes diagonaux de (14) et ceux relatifs au vecteur interspectral (15) sont estimés de la même façon que pour la solution matricielle. Les termes non diagonaux de (14), en revanche, ne sont pas estimés mais remplacés par leur valeur statistique exacte, nulle. Cependant, le traitement est moins performant vis-à-vis du critère d'optimalité (8) que lorsque l'hypothèse d'orthogonalité des références n'est pas utilisée. Bien que paradoxal à première vue, ceci s'explique par le fait que les références-bruit, stationnaires, ne sont qu'asymptotiquement orthogonales.

Les performances de la structure cascade ont également été déterminées. L'utilisation en cascade des références-bruit est plus performante que leur utilisation en parallèle (lorsqu'elle est efficace). Néanmoins cet apport ne suffit pas à compenser la perte d'information sur les termes non diagonaux de (14).

VIII REFERENCES BIBLIOGRAPHIQUES

- [1] WIDROW, GLOVER et al, Adaptive noise cancelling : principes and applications *Proc IEEE*, Dec 1975 vol.63 p1692.
- [2] P. COMON J.-L. LACOUME Noise reduction for an estimated Wiener filter using noise references *IEEE trans on info theory* Mars 1986.
- [3] C. SERVIERE D. BAUDOIS A. SILVENT Comparaison de deux méthodes d'estimation du gain complexe d'un filtre optimal au sens de Wiener *Traitement du Signal* vol.6 n°6 1989.
- [4] N. WIENER *Extrapolation interpolation and smoothing of stationary time series with engineering applications* New York Wiley 1949.
- [5] A.M. KSHIRSAGAR *Multivariate Analysis* New-York : Dekker, 1972.
- [6] I.S. REED Moment theorems for complex Gaussian processes *IEEE trans on info theory* Avril 1962 pp194-195.
- [7] D.R. BRILLINGER *Time Serie, Data analysis and Theory* 1981 Holden Day.
- [8] C. SERVIERE *Eléments de comparaison entre différentes méthodes de soustraction de bruit* Thèse INPG Grenoble 1989.
- [9] J-F GUERRE-CHALEY *Etude de structures de soustracteurs de bruit adaptatifs multivariés* Thèse INPG Grenoble 1990.
- [10] GANTMACHER *Théorie des matrices* Dunod 1966.