

# Un nouvel algorithme pour l'estimation adaptative de sous-espaces de type rationnel

I. Fijalkow, Ph. Loubaton  
Telecom Paris, Dept. Signal,  
46 rue Barrault, 75634 Paris Cedex 13, France

Ph. A. Regalia  
Institut National des Telecommunications,  
9 rue Charles Fourier, 91011 Evry Cedex, France

## Résumé

Dans un travail précédent, nous avons proposé une méthode temporelle adaptative d'estimation de sous-espaces de type rationnel. Cependant, la convergence de l'algorithme correspondant n'a pas été démontrée, et il semble que des attracteurs indésirables puissent apparaître. Dans cet article, nous proposons un nouvel algorithme jouissant de propriétés de convergence satisfaisantes. Il est robuste par rapport au modèle rationnel supposé, et a un comportement satisfaisant en présence de faible rapport signal sur bruit.

## Abstract

In a previous work, a time domain algorithm was proposed to adaptively estimate a rational and orthonormal spanning of a rational subspace. However, its convergence was not proved and it seems that asymptotic misbehaviour may occur. A new algorithm is proposed, for which satisfactory convergence properties hold. It provides robustness with regard to the signal model and shows an especially good asymptotical performance for small SNR.

## 1. INTRODUCTION

Dans de nombreuses applications mettant en jeu des traitements multicapteurs, le signal  $q$ -dimensionnel  $y(n)$  dont les composantes représentent les signaux capteurs peut s'écrire sous la forme

$$y(n) = \sum_{k=1}^p [d_k(z)]s_k(n) + w(n) = [D(z)]s(n) + w(n)$$

où les  $(s_k)_{k=1,p}$  sont des signaux sources non corrélés entre eux, et où les  $(d_k(z))_{k=1,p}$  sont des fonctions de transfert  $q \times 1$  que l'on peut supposer rationnelles sans restreindre la généralité du modèle.  $w(n)$  représente un bruit additif, que nous supposons blanc temporellement et spatialement, i.e.  $E[w(n)w^T(m)] = \sigma^2 I_q \delta(n-m)$ . Dans la plupart des cas, le nombre de capteurs  $q$  est supérieur au nombre de sources  $p$ , de sorte que la densité spectrale  $S_s(\omega)$  du signal utile  $y_s(n) = [D(z)]s(n)$  est de rang  $p < q$  pour chaque valeur de  $\omega$ . Dans ce cas, on a coutume d'appeler sous-espace source (noté  $\mathcal{S}(\omega)$ ) l'espace engendré par les colonnes de  $S_s(\omega)$ , et sous-espace bruit ( $\mathcal{B}(\omega)$ ) l'orthogonal de  $\mathcal{S}(\omega)$ . Le problème de l'estimation adaptative de ces sous-espaces pour chaque valeur de  $\omega$  présente un intérêt indéniable dans de nombreuses applications, la plus connue d'entre elles étant la localisation de sources par des méthodes de type MUSIC. Cependant, nous pensons que l'estimation de  $\mathcal{S}(\omega)$  peut être intéressante dans d'autres situations. Ainsi, il peut être utile de générer en temps réel, par filtrage linéaire de  $y$ , un signal  $p$ -dimensionnel  $y'$ , dans lequel la contribution du signal utile soit maximale en un certain sens. Ce problème, qui peut être interprété comme relevant de l'analyse en composantes principales, possède une solution très simple dès lors que  $\mathcal{S}(\omega)$  est connu. En effet, soit  $V(z)$  une fonction de transfert  $p \times q$ , dont les colonnes constituent, pour chaque  $z = e^{j\omega}$ , une base orthonormée de  $\mathcal{S}(\omega)$ ; alors, le signal  $p$ -dimensionnel  $y'(n) = [V(z)]y(n)$ , défini comme la sortie de  $V(z)$  excité par  $y(n)$ , satisfait les conditions précédentes.

Compte tenu des hypothèses que nous avons formulées sur  $w$ ,  $\mathcal{S}(\omega)$  (resp.  $\mathcal{B}(\omega)$ ) coïncide avec l'espace propre associé aux  $p$  plus grandes (resp. aux  $q-p$  plus petites) valeurs propres de la matrice de densité spectrale  $S_y(\omega)$  du signal observé. Dès lors, on pourrait penser à estimer de façon adaptative le sous-espace source (ou le sous-espace bruit) par le

biais de méthodes fréquentielles, consistant à former des estimateurs de  $S_y(\omega)$  à diverses fréquences, puis à y appliquer des algorithmes adaptatifs d'extraction de valeurs propres / vecteurs propres. Cependant, ces méthodes ne sont pas les plus naturelles dans un contexte adaptatif, pour lequel l'utilisation de méthodes fonctionnant dans le domaine temporel est sans doute préférable. Pour cela, nous utilisons le fait que le sous-espace bruit est par définition-même rationnel, en ce sens qu'il possède des bases dont les vecteurs ont toutes leurs composantes rationnelles en  $e^{j\omega}$ . Ainsi, l'estimation adaptative de  $\mathcal{B}(\omega)$  équivaut à l'identification adaptative d'une fonction de transfert matricielle dont les colonnes représentent une base de  $\mathcal{B}(\omega)$  en tout point  $e^{j\omega}$  du cercle unité.

A notre connaissance, l'article [1] est le premier travail dans lequel cette approche est envisagée. Il repose sur le fait que si  $V(z)$  est une fonction de transfert  $q$ -entrées /  $(q-p)$ -sorties, contrainte à vérifier  $V(e^{j\omega})V^T(e^{-j\omega}) = I_{q-p}$ , alors, l'énergie du signal filtré  $[V(z)]y(n)$  est donnée par  $E[\| [V(z)]y(n) \|^2]$  :

$$\text{trace} \left\{ \int_{-\pi}^{\pi} V(e^{j\omega})S_s(\omega)V^T(e^{-j\omega})d\omega \right\} + (q-p)\sigma^2$$

et est minimale si et seulement si (ssi)  $V(e^{j\omega})S_s(e^{j\omega}) = 0$  pour tout  $\omega$ , i.e. ssi les colonnes de  $V^T(e^{-j\omega})$  constituent une base orthonormée de  $\mathcal{B}(\omega)$ . On peut alors identifier une base rationnelle et orthonormée de  $\mathcal{B}(\omega)$  en minimisant de façon adaptative la fonction de coût  $E[\| [V(z)]y(n) \|^2]$  sur l'ensemble des fonctions de transfert, dites sans perte, qui sont rationnelles (de degré suffisant), stables, et vérifient  $V(e^{j\omega})V^T(e^{-j\omega}) = I_{q-p}$  pour tout  $\omega$  (la contrainte de stabilité est obligatoire dans le cadre des problèmes de filtrage adaptatif RII). La difficulté majeure posée par ce problème de filtrage adaptatif provient du fait que le filtre à identifier est contraint d'appartenir à un ensemble dont la structure n'est pas simple. En particulier, la paramétrisation des fonctions de transfert matricielles sans perte (voir [2] par exemple) est d'une complexité rédhibitoire pour la mise en évidence d'un algorithme adaptatif. Une procédure, motivée par l'approche de déflation utilisée dans le cadre de la diagonalisation adaptative des matrices de covariance (voir [3]), a été proposée dans [1]. Elle consiste à identifier ligne par



ligne une solution du problème de minimisation précédent. Pour cela, on adapte un filtre sans perte  $1 \times q$  de fonction de transfert  $u_2(z)$  de telle sorte que  $E \{([u_2(z)]y(n))^2\}$  soit minimal. La paramétrisation de ce type de fonction de transfert (voir figure 1) à valeurs vectorielles est beaucoup tractable que celle des fonctions de transfert matricielles sans perte, et, d'autre part, elle permet de mettre directement en évidence un filtre sans perte  $(q-1) \times q$  de fonction de transfert  $U_1(z)$ , pour lequel  $U(z) = [U_1^T(z), u_2^T(z)]^T$  est sans perte. Ainsi, simultanément à l'adaptation de  $u_2$ , on génère un signal  $(q-1)$ -dimensionnel,  $z_1$ , qui converge vers un processus stationnaire possédant un sous-espace bruit de type rationnel et de dimension  $q-p-1$ . Si l'on extrait un vecteur sans perte  $u_{2,1}(z)$  du sous-espace bruit de  $z_1$ , le vecteur  $u_{2,1}(z)U_1(z)$  correspondra à un second vecteur de  $\mathcal{B}(\omega)$ , orthogonal à  $u_2(z)$  pour chaque  $z = e^{j\omega}$ . En réitérant le processus, on construit une fonction de transfert sans perte  $(q-p) \times q$  engendrant  $\mathcal{B}(\omega)$  pour tout  $\omega$ . Par le biais de cette procédure de déflation, on ramène le problème de l'estimation d'une fonction de transfert matricielle sans perte à celui de l'identification de plusieurs fonctions de transfert vectorielles de tailles décroissantes.

Il apparaît donc que la minimisation adaptative de  $E \{([u_2(z)]y(n))^2\}$  sur l'ensemble des fonctions de transfert sans perte  $1 \times q$  joue un rôle fondamental dans l'approche de [1]. Cependant, la convergence de l'algorithme proposé dans [1] n'a pu être établie, et nous suspectons qu'il possède des attracteurs indésirables. Le but de cet article est de mettre en évidence un nouvel algorithme permettant de minimiser  $E \{([u_2(z)]y(n))^2\}$ , et pour lequel des résultats de convergence plus satisfaisants peuvent être établis. Après avoir présenté notre nouvel algorithme dans la section 2, nous comparons soigneusement ses performances avec celles de la méthode proposée dans [1].

## 2. L'ALGORITHME PROPOSÉ

Pour adapter un filtre sans perte  $u_2(z)$  de degré de Mac-Millan  $M$  fixé, il suffit d'être en mesure d'adapter les  $N = (M+1)q - 1$  angles de rotation de la structure de la figure 1. Pour cela, il convient de rappeler que le système  $z(n) = (z_1^T(n), z_2^T(n))^T = [U(z)]y(n)$  se décrit sous forme d'état par l'intermédiaire de la relation  $(x^T(n+1), z_1^T(n), z_2^T(n))^T = Q(\Theta) (y^T(n), x^T(n))^T$  où  $Q(\Theta)$  est la matrice orthogonale définie par  $Q(\Theta) = Q_1(\theta_1) \cdots Q_N(\theta_N)$ , les  $Q_k(\theta_k)$  représentant les rotations planaires

$$\begin{bmatrix} I_{k-1} & & & \\ & \cos(\theta_k) & & -\sin(\theta_k) \\ & & I_{N-k-1} & \\ & \sin(\theta_k) & & \cos(\theta_k) \end{bmatrix}$$

Pour minimiser, par rapport à  $\Theta$ , la fonction de coût  $E \{([u_2(z)]y(n))^2\}$ , il pourrait être tentant d'utiliser l'algorithme du gradient stochastique. Cependant, ainsi que cela est expliqué dans [1], ceci pose deux problèmes. Tout d'abord, cet algorithme est d'une complexité importante, du fait de la nécessité de calculer récursivement les dérivées de l'état  $x$  par rapport aux paramètres angulaires. Ensuite, la fonction de coût  $E \{([u_2(z)]y(n))^2\}$  possède très certainement des minima locaux parasites vers lesquels l'algorithme du gradient stochastique pourrait converger. Afin d'y remédier, il est proposé dans [1] d'utiliser l'algorithme

$$\Theta(n+1) = \Theta(n) - \mu z_2(n) \begin{pmatrix} x(n+1) \\ z_1(n) \end{pmatrix} \quad (1)$$

où  $(x^T(n+1), z_1^T(n), z_2^T(n))^T = Q(\Theta(n)) (y^T(n), x^T(n))^T$ . La complexité de (1) est très faible. De plus, les performances obtenues sur signaux simulés ont toujours été satisfaisantes. Cependant, la convergence de (1) n'a pu être établie. En particulier, l'ensemble des points stationnaires stables de l'équation différentielle moyenne (ODE) associée à (1) (i.e. les points de convergence possibles de (1)) n'a pu être caractérisé, et il nous semble probable que les points  $\Theta_*$  correspondants aux filtres  $u_2$  du sous-espace bruit ne sont pas les seuls points stationnaires (voir [4]).

Le nouvel algorithme proposé est constitué de deux "étages". Le premier extrait de façon adaptative un filtre RIF  $1 \times q$  de fonction de transfert  $g(z)$  de telle sorte que  $g^T(e^{-j\omega})$  appartienne à  $\mathcal{B}(\omega)$  pour tout  $\omega$ . Simultanément, un algorithme de faible complexité génère un filtre sans perte  $1 \times q$  de fonction de transfert  $u_2(z)$  pour lequel, à la convergence, les vecteurs  $u_2(e^{j\omega})$  et  $g(e^{j\omega})$  sont colinéaires pour tout  $\omega$ . On assure ainsi l'appartenance du filtre limite  $u_2^T(e^{-j\omega})$  à  $\mathcal{B}(\omega)$  pour tout  $\omega$ .

L'extraction d'un filtre RIF  $g(z)$  vérifiant les conditions précédentes est immédiate, et repose sur les propriétés de la matrice de covariance "spatio-temporelle"  $\mathcal{R}_M$  du vecteur aléatoire  $\mathcal{Y}_M(n) = (y^T(n), \dots, y^T(n-M))^T$ . En effet, si  $M$  est suffisamment grand, cette matrice admet  $\sigma^2$ , la variance du bruit, comme plus petite valeur propre, et tout vecteur propre  $g = (g_0, \dots, g_M)$  correspondant est tel que le transconjugé de  $g(e^{j\omega}) = \sum_{k=0}^M g_k e^{-jk\omega}$  appartient à  $\mathcal{B}(\omega)$  pour tout  $\omega$  (voir [5]). L'identification adaptative d'un tel filtre RIF est donc équivalente à l'extraction d'un vecteur propre associé à la plus petite valeur propre de  $\mathcal{R}_M$ . Ceci peut être accompli par des algorithmes bien connus, le plus simple d'entre eux s'écrivant

$$g(n+1) = \frac{g(n) - \mu r(n) \mathcal{Y}_M^T(n)}{\| - \|} \quad (2)$$

où  $g(n)$  est associé au filtre RIF  $g(n, z)$  à l'instant  $n$ , et où  $r(n) = g(n) \mathcal{Y}_M(n) = [g(n, z)]y(n)$ . A ce point de la discussion, il convient de mentionner que l'extraction d'une base de l'espace propre de  $\mathcal{R}_M$  associé à sa plus petite valeur propre  $\sigma^2$  pourrait permettre d'identifier le sous-espace bruit tout entier. Cependant, la dimension du noyau de  $(\mathcal{R}_M - \sigma^2 I)$  est en général beaucoup plus grande que  $q-p$ , puisque l'on peut établir que pour  $M$  assez grand, elle vaut  $(M+1)(q-p) - M'$  où  $M'$  représente le "degré" du sous-espace bruit, i.e. la moitié du degré de Mac-Millan de la matrice, positive et rationnelle, de projection orthogonale sur  $\mathcal{B}(\omega)$ . On serait donc amené à identifier beaucoup plus de vecteurs que nécessaire. D'autre part, les transformées en  $z$  des vecteurs d'une base de  $\ker(\mathcal{R}_M - \sigma^2 I)$  devraient, en théorie, engendrer un espace de dimension  $q-p$ , pour tout  $z = e^{j\omega}$ ; mais tel ne sera pas le cas pour les vecteurs générés par une procédure adaptative, et il serait nécessaire, pour chaque  $\omega$ , d'effectuer une analyse en composantes principales pour en extraire  $q-p$  vecteurs.

Le second étage de la procédure génère une suite de filtres sans perte par le biais de l'algorithme

$$\Theta(n+1) = \Theta(n) + \mu r(n) \begin{pmatrix} x(n+1) \\ z_1(n) \end{pmatrix} \quad (3)$$

où  $\begin{pmatrix} x(n+1) \\ z_1(n) \end{pmatrix} = Q_1(\Theta(n)) \begin{pmatrix} y(n) \\ x(n) \end{pmatrix}$  et où on rappelle que  $r(n)$  coïncide avec  $[g(n, z)]y(n)$ . Afin d'étudier le comportement asymptotique de (3), désignons par  $g_*(z)$  le filtre correspondant à la limite de (2), et par  $r_*(n)$  le signal  $[g_*(z)]y(n)$ . Alors, modulo quelques hypothèses techniques

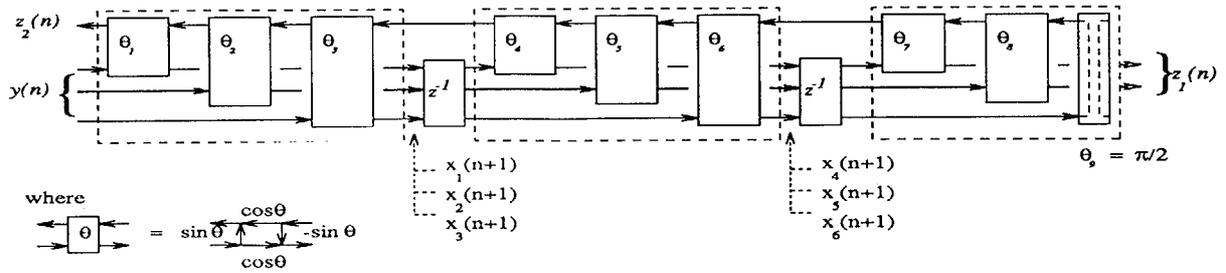


Figure 1: Structure sans perte en cascade, pour  $q = 3$  et  $M = 2$

supplémentaires (voir [6]), le comportement asymptotique de (3) est équivalent à celui de l'algorithme

$$\Theta(n+1) = \Theta(n) + \mu r_*(n) \begin{pmatrix} x(n+1) \\ z_1(n) \end{pmatrix} \quad (4)$$

Pour étudier (4), nous utilisons la méthode de l'équation différentielle moyenne ([7], [6] par exemple). Soit  $h(\Theta)$  la fonction définie par  $h(\Theta) = E \left[ \begin{pmatrix} x(n+1, \Theta) \\ z_1(n, \Theta) \end{pmatrix} r_*(n) \right]$ , où  $x(n+1, \Theta)$  et  $z_1(n, \Theta)$  représentent les signaux générés par la structure de la figure 1 lorsque les  $\theta_k$  sont figés. Alors, il est bien connu que les points de convergence possible de (4) sont les points stationnaires  $\Theta_*$  (i.e.  $h(\Theta_*) = 0$ ) stables (i.e. la matrice des dérivées de  $h$  au point  $\Theta_*$  a ses valeurs propres de partie réelle négative) de l'équation différentielle  $\frac{d\Theta}{dt} = h(\Theta)$ . Le résultat essentiel de cet article est que les seuls points stationnaires de (4) correspondent aux filtres sans pertes  $u_{2,*}$  pour lesquels  $u_{2,*}(e^{j\omega})$  et  $g_*(e^{j\omega})$  engendrent le même espace pour chaque  $\omega$ . Plus précisément, on a le théorème suivant.

**Théorème 1**  $\Theta_*$  est un point stationnaire de (4) ssi il existe un polynôme scalaire  $d(z)$  pour lequel  $g_*(z) = d(z)u_{2,*}(z)$ . Si de plus les diverses composantes de  $g_*(z)$  n'ont aucun zéro commun, (4) admet un unique point stationnaire défini par le fait que  $u_{2,*}(z) = g_*(z)/d_*(z)$ , où  $d_*(z)$  est l'unique polynôme scalaire dépourvu de zéro à l'extérieur du disque unité pour lequel  $g_*(z)g_*^T(z^{-1}) = d_*(z)d_*^T(z^{-1})$

**Démonstration:** Nous ne donnons ici que le principe de la démonstration. Les détails pourront être trouvés dans [4]. Tout repose sur le fait que,  $g_*^T(e^{-j\omega})$  appartenant à  $\mathcal{B}(\omega)$  pour tout  $\omega$ , l'action de  $g_*(z)$  annule le signal utile  $y_s(n)$ , de sorte que  $r_*(n) = [g_*(z)]w(n)$ . Notons  $F_\Theta(z)$  la fonction de transfert pour laquelle  $x^T(n+1, \Theta), z_1^T(n, \Theta)^T = [F_\Theta(z)]y(n)$ ; compte tenu de l'indépendance entre les signaux sources et le bruit,  $h(\Theta) = E \{ [F_\Theta(z)]w(n) r_*(n) \}$ . Puisque le bruit est blanc temporellement et spatialement, les composantes de  $h(\Theta)$  sont égales aux produits scalaires<sup>1</sup> des lignes de  $F_\Theta(z)$  avec  $g_*(z)$ .  $\Theta_*$  est donc point stationnaire de l'ODE associée à (4) ssi  $g_*(z)$  est orthogonal à l'espace engendré par les lignes de  $F_{\Theta_*}(z)$ . On démontre alors en utilisant des arguments assez classiques en théorie des systèmes que ceci implique que  $g_*(z) = d(z)u_{2,*}(z)$  pour un certain polynôme scalaire  $d(z)$ . L'unicité de  $u_{2,*}$ , dans le cas où les composantes de  $g_*(z)$  sont premières entre elles, est immédiate. □

Ce résultat montre que modulo une hypothèse très peu restrictive, (4) possède comme point stationnaire unique le

<sup>1</sup>le produit scalaire considéré ici est  $\langle f(z), g(z) \rangle = \int_{[-\pi, \pi]} f(e^{j\omega})g^T(e^{-j\omega})d\omega$

point  $\Theta_*$  désiré. Bien entendu, ceci ne suffit pas à démontrer la convergence de (4) vers  $\Theta_*$  puisqu'il faudrait pour cela établir que  $\Theta_*$  est un attracteur global pour son ODE, ce qui n'a pu être démontré. Le résultat obtenu est cependant réconfortant. D'autre part, il est intéressant d'interpréter (3) comme un algorithme effectuant implicitement la factorisation causale à phase minimale de la densité spectrale limite des  $g(n, z)g^T(n, z^{-1})$ . En particulier, la charge de calcul correspondante est très faible comparée à celle que l'on obtiendrait en implantant de façon adaptative des algorithmes de factorisation polynomiaux. Enfin, il convient de mentionner que si la variance du bruit est très faible, le signal  $r(n)$  tendra vers 0, et qu'il en résultera des problèmes numériques importants. Pour y remédier, nous proposons de piloter l'algorithme (3) par un bruit temporellement et spatialement blanc  $v$ , de variance  $\sigma_v^2$ . Autrement dit,  $r(n)$  sera remplacé par  $[g(n, z)]v(n)$  tandis que  $x(n+1)$  et  $z_1(n)$  seront générés par le biais de la récurrence  $(x^T(n+1), z_1^T(n))^T = Q_1(\Theta(n)) (v^T(n), x^T(n))^T$ . Il est alors facile de constater que le théorème 1 reste valable et que sa validité ne dépend nullement du caractère rationnel du sous espace bruit comme dans le cas précédent, où le fait que l'action de  $g_*(z)$  annule le signal utile jouait un rôle fondamental. L'utilisation d'un bruit externe permet donc de rendre notre approche robuste vis à vis de l'hypothèse de rationalité. Enfin, le fait de générer  $v$  permet de contrôler la dynamique des signaux intervenant dans l'algorithme. Ainsi, dans ce qui suit, l'algorithme (3) sera supposé piloté par un bruit externe  $v$ . Pour simplifier les notations, nous continuerons cependant à désigner par  $r(n), x(n+1)$  et  $z_1(n)$  les différents signaux qui interviennent.

### 3. COMPARAISON DES ALGORITHMES

Nous allons à présent comparer le nouvel algorithme avec celui proposé dans [1]. Puisque les performances d'un algorithme adaptatif se mesurent par deux facteurs contradictoires, à savoir la variance asymptotique des estimateurs (i.e. les traces des matrices de covariance des erreurs  $\Theta(n) - \Theta_*$ ) et la vitesse de convergence, nous allons comparer la vitesse de convergence de nos deux algorithmes en choisissant les divers paramètres de sorte que leurs variances asymptotiques soient égales. Cela est rendu possible par l'existence dans la littérature de l'approximation stochastique de formules permettant d'évaluer  $E \{ (\Theta(n) - \Theta_*)(\Theta(n) - \Theta_*)^T \}$  lorsque le pas  $\mu$  des algorithmes est choisi suffisamment petit (voir [7], p. 104 par exemple).

Ainsi, dans le cas de l'algorithme de [1],  $T_1 = E \{ \|\Theta(n) - \Theta_*\|^2 \}$  est donné par

$$T_1 = \mu \frac{\sigma^2}{\sigma_s^2} \left( p_s + \frac{\sigma^2}{\sigma_s^2} p_w \right) \sigma_s^2$$

où  $\sigma_s^2$  désigne la puissance du signal utile, et où  $p_s$  et  $p_w$  sont

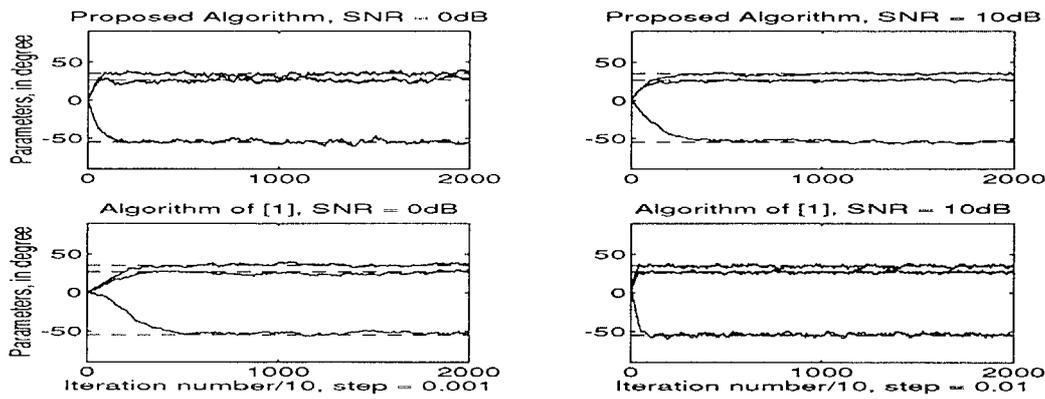


Figure 2: Simulations des Algorithmes pour SNR 0dB (à gauche) and 10dB (à droite)

des termes correspondant au signal utile et au bruit normalisés en puissance. Pour évaluer la variance de l'algorithme (3), il faudrait, en principe, tenir compte du fait que  $r(n)$  est généré à partir d'un polynôme  $g(n, z)$  estimé de façon adaptative par le biais de (2). En particulier, le désajustement de  $g(n, z)$  autour de  $g_*(z)$  contribue à la variance de (3). Cependant, nous avons constaté dans tous les cas étudiés que les performances de (2) étaient bien supérieures à celles de (3), en terme de variance asymptotique comme en terme de vitesse de convergence. Ceci implique que l'on peut réduire le pas de (2), et donc diminuer son influence sur  $E[\|\Theta(n) - \Theta_*\|^2]$ , sans pour autant affecter la vitesse de convergence de (3), du moins dans des gammes de rapport signal à bruit raisonnable. Il est donc légitime de simplifier l'étude des variances asymptotiques de (3) en assimilant son comportement avec celui de (4). On en déduit immédiatement que  $T_2 = E[\|\Theta(n) - \Theta_*\|^2]$  est donné par

$$T_2 = \mu \sigma_v^2 p_v$$

où le facteur  $p_v$ , qui correspond à la valeur de  $T_2$  lorsque  $\sigma_v^2 = 1$ , ne dépend que de  $\Theta_*$  et des statistiques du quatrième ordre du bruit normalisé  $v(n)/\sigma_v$ . Notons qu'il est possible de choisir la loi de probabilité de  $v(n)/\sigma_v$  de telle sorte que le facteur  $p_v$  soit minimal; ainsi, il est établi dans [4] que le minimum de  $p_v$  est atteint lorsque les composantes de  $v(n)/\sigma_v$  suivent une loi de Bernoulli. Ces expressions montrent comment les paramètres  $\sigma_v^2$ ,  $\sigma_s^2$  et  $\sigma^2$  vont influencer sur la vitesse de convergence des algorithmes lorsque l'on impose  $T_1 = T_2$ . En effet, il convient de remarquer que, au moins pour des valeurs suffisamment petites de  $\mu$ , la vitesse de convergence de nos deux algorithmes est plus ou moins liée à la vitesse de convergence des trajectoires  $\Theta(t)$  de leurs ODE vers leurs points d'équilibre, et à la façon dont on associe les valeurs  $\Theta(n)$  aux trajectoires  $\Theta(t)$ . Ainsi que cela est bien connu, les valeurs prises par les  $\Theta(n)$  doivent évoluer dans un tube de confiance contrôlé par le pas  $\mu$ , centré sur des valeurs  $\Theta(n\mu)$  prises par l'une des trajectoires aux instants  $n\mu$ . En examinant la dépendance des ODE de nos deux algorithmes en fonction des paramètres  $\sigma_s^2$  et  $\sigma_v^2$ , on voit immédiatement que le paramètre  $\mu \sigma_s^2$  contrôle la vitesse de l'algorithme de [1], tandis que  $\mu \sigma_v^2$  contrôle celle de (4).

Supposons que le rapport signal sur bruit  $\sigma_s^2/\sigma^2$  soit important. Alors, l'égalité entre  $T_1$  et  $T_2$  implique que le produit  $\mu \sigma_s^2$  va être bien supérieur au produit  $\mu \sigma_v^2$ . Dans ce cas, la vitesse de convergence de l'algorithme de [1] sera supérieure à celle de (4). Par contre, en présence d'un rapport signal sur bruit défavorable, la conclusion sera inversée. Cette discussion est confirmée par les résultats de simulation présentés figure 2, et qui correspondent à un cas simple où

$M = 1$  et  $q = 2$  pour lequel les facteurs  $p_s$ ,  $p_w$  et  $p_v$  peuvent être calculés explicitement. Ainsi, lorsque le rapport signal sur bruit vaut 10dB, l'algorithme proposé dans [1] apparaît bien supérieur à (3), tandis que pour un rapport signal sur bruit de 0dB, l'algorithme proposé jouit d'une vitesse de convergence plus satisfaisante.

#### 4. CONCLUSION

Dans cet article, nous avons présenté un nouvel algorithme permettant d'estimer de façon adaptative des sous-espaces bruit de type rationnel. Son comportement est indépendant du caractère rationnel du sous-espace, de sorte qu'il apparaît comme plus robuste que celui proposé dans [1]. Enfin, ses performances sont meilleures que celles de [1] quand le rapport signal sur bruit est défavorable.

#### References

- [1] Ph.A. Regalia and Ph. Loubaton, *Rational subspace estimation using adaptive lossless filters*, IEEE Trans. on Signal Processing, vol. 40, n°10, pp. 2392–2405, October 1992.
- [2] U.B. Desai, *A state-space approach to orthogonal digital filters*, IEEE Trans. on CAS, vol. 38, n°2, pp. 160–169, February 1991.
- [3] Ph.A. Regalia, *An adaptive unit norm filter with applications to signal analysis and Karhunen-Lohve transformations*, IEEE Trans. on CAS, vol. 37, n°5, pp. 646–649, May 1990.
- [4] I. Fijalkow, Ph. Loubaton and Ph.A. Regalia, *Asymptotic Performances of Rational Subspace Estimation Algorithms based on Cascadable Adaptive Lossless Filters*, en préparation.
- [5] Y. Grenier et Ph. Loubaton, *Localisation de sources large bande par des méthodes temporelles*, In Proc. GRETSI 89, pp. 457–460, 1989.
- [6] H.J. Kushner and D.S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Applied Mathematical Sciences 26, 1978.
- [7] A. Benveniste, M. Metivier and P. Priouret, *Algorithmes Adaptatifs et Approximations Stochastiques*, Masson, 1987.