

Performances asymptotiques de deux algorithmes adaptatifs de structure propre pour l'optimisation de bancs de filtres RIF multicadence sans pertes

Dong-yan Huang⁽¹⁾ et Phillip Regalia⁽²⁾

⁽¹⁾ Université René Descartes-Paris V
UFR de Mathématiques et Informatique
45 rue des Sts Pères, 75270 Paris Cedex 06

⁽²⁾ SIM/INT, 9 rue Charles Fourier, 91011 Evry cedex

RÉSUMÉ

Dans cet article, nous appliquons un résultat de Benveniste pour l'analyse des performances asymptotiques des deux algorithmes proposés dans [3] et réalisons une étude comparative. Nous obtenons des expressions analytiques sur la variance asymptotique des paramètres de bancs de filtres sans pertes et celle de la sortie $(y_2(n, \Theta) - y_2(n, \Theta_*))$. Les simulations montrent que les variances asymptotiques des paramètres $(\Theta(n) - \Theta_*)$ et celle de la sortie $(y_2(n, \Theta) - y_2(n, \Theta_*))$ du deuxième algorithme sont plus petites que celles du premier algorithme. Ceci aide l'utilisateur à choisir un meilleur algorithme.

ABSTRACT

This paper gives a performance analysis of two eigenstructure algorithms proposed in [3] with help of Benveniste's asymptotic analysis results and performs a comparative study. Closed form expression of the parameters of lossless filter banks and the second output $(y_2(n, \Theta) - y_2(n, \Theta_*))$ are given for the case of stochastic inputs and are further analyzed to provide some insights into the behavior of these two algorithms. Simulations show that the asymptotic variance of parameters $(\Theta(n) - \Theta_*)$ and second output $(y_2(n, \Theta) - y_2(n, \Theta_*))$ of the second algorithm are smaller than the first one's, which help the user in choosing a better algorithm.

1 Introduction

La technique adaptative pour l'optimisation de bancs de filtres RIF sans pertes à deux canaux (Fig. 1) a été développée ces dernières années en vue des applications au codage en sous-bandes et à l'analyse en ondelettes [1, 2, 3, 4]. Elle consiste à mettre à jour les angles de rotation $\{\theta_k\}_{k=0}^{M-1}$ pour que la variance de la deuxième sortie $E[y_2^2(n)]$ soit la plus petite possible ou minimisée selon quelque critère prédéfini. Cependant, dans la littérature peu de travaux abordent les problèmes de convergence et de performances, qui peuvent être appréhendées à l'aide de deux facteurs : le comportement transitoire (vitesse de convergence) et l'erreur en régime permanent (finesse) de bancs de filtres RIF sans pertes.

Deux algorithmes de structure propre proposés dans [3] montrent plusieurs avantages par rapport aux algorithmes du type gradient [1, 2], notamment une complexité de calcul réduite et l'absence de minima locaux. Pourtant, la nature non-linéaire et non-gradient de ces deux algorithmes rend l'analyse de convergence et de performances assez difficile. Par conséquent, l'étude de leur convergence est un sujet intéressant mais ouvert. Il faut remarquer qu'il n'y a pas de relation entre la performance asymptotique et la vitesse de convergence. En fait, il n'existe pas de méthode convenable pour étudier la vitesse de convergence. C'est la raison pour laquelle on a obtenu seulement le conditionnement d'entrée

et une condition nécessaire pour laquelle les deux algorithmes convergent [5] en utilisant la méthode d'équation différentielle ordinaire (ODE). En revanche, les travaux de Benveniste *et al.* [6] permettent d'étudier théoriquement les performances asymptotiques de bancs de filtres adaptatifs. Dans cet article, nous allons utiliser un résultat de Benveniste pour étudier et comparer la performance asymptotique des deux algorithmes de structure propre.

Considérons un banc de filtres RIF sans pertes à deux canaux (Fig. 1), on peut écrire la description d'état sous la forme suivante :

$$\begin{bmatrix} \mathbf{x}(n+1) \\ \mathbf{y}(n) \end{bmatrix} = \begin{bmatrix} Q_1(\Theta) \\ Q_2(\Theta) \end{bmatrix} \left\{ \underbrace{\begin{bmatrix} A(\Theta) & B(\Theta) \\ c_1(\Theta) & d_1(\Theta) \\ c_2(\Theta) & d_2(\Theta) \end{bmatrix}}_{Q(\Theta)} \right\} \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{u}(n) \end{bmatrix} \quad (1)$$

où $\mathbf{x}(\cdot) = [x_1(\cdot), \dots, x_M(\cdot)]^T$ est le vecteur d'état, $\mathbf{u}(\cdot) = [u_1(n), u_2(n)]^T$ est le vecteur d'entrée, et $\mathbf{y}(\cdot) = [y_1(\cdot), y_2(\cdot)]^T$, le vecteur de sortie.

A chaque itération n , l'algorithme de structure propre met à jour récursivement $\Theta(\cdot) = [\theta_0, \dots, \theta_{M-1}]^T$, le vecteur des angles de rotation, selon :

$$\Theta(n+1) = \Theta(n) - \mu y_2(n) \Gamma(n) \begin{bmatrix} \mathbf{x}(n+1) \\ y_1(n) \end{bmatrix} \quad (2)$$

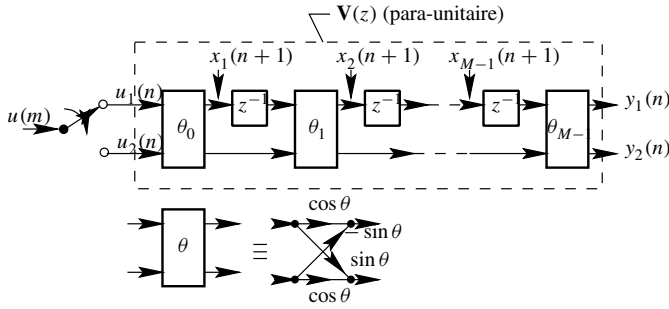


FIG. 1 — Banc de filtres sans pertes à deux canaux

où μ est un gain constant ; $\Gamma(\Theta) = \text{diag}[\gamma_1, \dots, \gamma_M]$ avec $\gamma_M = 1$, $\gamma_k(n) = \gamma_{k+1}(n) \cos \theta_k(n)$, ($1 \leq k \leq M$).

Cependant, l'algorithme de structure propre ne converge pas si les termes d'indice impair de la matrice d'autocorrelation sont négligeables par rapport aux termes d'indice pair. C'est pourquoi on a proposé l'algorithme de modulation [3], qui s'actualise seulement par les termes d'indice impair de la matrice d'autocorrelation :

$$\Theta(n+1) = \Theta(n) - \frac{\mu}{2} \Gamma(n) \times \left\{ y_2(n) \begin{bmatrix} \mathbf{x}(n+1) \\ y_1(n) \end{bmatrix} - \hat{y}_2(n) \begin{bmatrix} \mathbf{x}(n+1) \\ \hat{y}_1(n) \end{bmatrix} \right\} \quad (3)$$

où $\mathbf{x}(n+1)$, $\hat{y}_2(n)$, $\hat{y}_1(n)$ sont obtenus par un banc de filtres (Fig. 1), où le signal d'entrée $u(m)$ est modulé par $(-1)^m$:

$$\mathbf{u}(m) = \begin{bmatrix} u(2n) & -u(2n-1) & u(2(n-1)) & \dots \end{bmatrix}$$

L'analyse des performances asymptotiques de ces deux algorithmes peut se faire à l'aide d'un résultat général de Benveniste *et al* (voir [6, Théorème 2, page 108]). Il concerne l'évaluation de la Matrice de Covariance Asymptotique du Vecteur Ecart (MCAVE) des paramètres $\varepsilon(n) = \Theta(n) - \Theta_*$ et la variance asymptotique de l'écart de la sortie ($y_2(n, \Theta) - y_2(n, \Theta_*)$), où Θ_* est un point stationnaire asymptotique.

Les résultats de simulation montrent que les variances asymptotiques des paramètres ($\Theta(n) - \Theta_*$) et celle de la sortie ($y_2(n, \Theta) - y_2(n, \Theta_*)$) de l'algorithme de modulation sont plus petites que celles de l'algorithme de structure propre, ce qui nous aide à choisir un meilleur algorithme.

L'organisation de cet article est le suivant : on présente brièvement les bases théoriques permettant l'étude de la MCAVE, puis on applique cette base théorique pour analyser la MCAVE des paramètres $\varepsilon(n) = \Theta(n) - \Theta_*$ et la variance asymptotique de l'écart de la sortie ($y_2(n, \Theta) - y_2(n, \Theta_*)$) dans la section 2. Des résultats seront présentés pour valider l'analyse théorique et comparer les performances asymptotiques des deux algorithmes dans la section 3. Finalement, la section 4 donne une conclusion.

2 Analyse des performances

2.1 Rappel théorique sur un résultat d'approximation normale

Afin d'analyser les performances asymptotiques des deux algorithmes, on fait appel à un résultat général de Benveniste *et al.* (voir [6, Théorème 2, page 108]).

Considérons un algorithme sous la forme :

$$\Theta(n+1) = \Theta(n) + \mu H(\Theta, X(n, \Theta)) \quad (4)$$

en régime stationnaire et à gain constant μ . L'étude de convergence de l'algorithme (4) est intimement liée à son ODE associée introduite par Ljung [7] :

$$\frac{d\Theta(t)}{dt} = h(\Theta)$$

où $h(\Theta) = E[H(\Theta, X(n, \Theta))]$, le vecteur champ moyen.

Soit Θ_* un point stationnaire globalement asymptotique stable de l'algorithme (4), on considère l'équation de Lyapunov :

$$RP + PR^T + S = 0 \quad (5)$$

où $R = \frac{dh(\Theta)}{d\Theta}|_{\Theta=\Theta_*}$ est la dérivée du vecteur champ moyen $h(\Theta)$ par rapport à chaque paramètre Θ à la valeur Θ_* et S est définie à partir de la matrice de covariance du vecteur $H(\Theta_*, X(n, \Theta_*))$:

$$S = \sum_{n=-\infty}^{\infty} E_{\Theta_*} [H(\Theta_*, X(n, \Theta_*))(H(\Theta_*, X(0, \Theta_*)))^T]$$

où $X(n, \Theta_*)$ est le processus construit comme $X(n)$ mais avec Θ fixé à la valeur Θ_* . Si toutes les valeurs propres de la dérivée du vecteur champ moyen R ont une partie réelle strictement négative, alors lorsque μ est "suffisamment petit" et n suffisamment grand, la Matrice de Covariance du Vecteur Écart (MCVE) de $(\Theta(n) - \Theta_*)$ converge vers la matrice de covariance P de la loi normale, où P est la solution de l'équation de Lyapunov (5) :

$$\frac{1}{\sqrt{\mu}}(\Theta(n) - \Theta_*) \rightarrow N(0, P)$$

Son utilisation, même quand l'attracteur Θ_* n'est pas global, est motivée par les remarques à propos de l'approximation asymptotique du vecteur écart $\Theta(n) - \Theta_*$ par une loi normale de la matrice de covariance P . (voir Théorème 2, [6, pp. 108]).

Par la suite, on se concentre sur l'analyse des performances des deux algorithmes en se basant sur la méthode de Benveniste *et al.* A cette fin, on fait les hypothèses suivantes :

1. Le signal d'entrée $u(m)$ est stationnaire, gaussien, centré, non-prédictible, m -dépendant et ses moments d'ordre 1, 2 et 4 sont bornés, c.à.d. $\|u(n)\|_p \leq c_p \leq \infty$ pour $1 \leq p \leq 4$;
2. Les algorithmes convergent vers le point stationnaire stable Θ_* de sorte que la variance d'erreur $E[y_2^2(n)]$ soit la plus petite possible ;
3. Aucun des composants de Θ_* ne vaut $\pm\pi/2$. Sinon, le résultat théorique de la MCAVE de $(\Theta(n) - \Theta_*)$ n'est plus valable.

4. $\varepsilon(n)$ et $(y_2(n, \Theta) - y_2(n, \Theta_*))$ sont indépendants.

La condition de m -dépendance est en fait une condition technique. Elle sous-entend que le signal d'entrée satisfait la condition $E\{u(i)u(j)\} = 0$, pour $|i - j| \geq m$ ¹. Cette hypothèse implique que la valeur propre minimale de la matrice d'auto-corrélation du signal d'entrée $R_u = E\{\mathbf{u}(n)\mathbf{u}^T(n)\}$ est strictement positive, $\lambda_{\min}(R_u) > 0$ quelque soit son ordre. La dernière hypothèse n'est pas vraiment restrictive car il est facile de contrôler le cas où $\Theta_* = \pm\pi/2$.

2.2 Performance asymptotique des paramètres $(\Theta(n) - \Theta_*)$ et de la deuxième sortie $(y_2(n, \Theta) - y_2(n, \Theta_*))$

En observant les deux algorithmes (2) et (3), on constate bien que ces deux algorithmes peuvent être résumés par (4). Pour appliquer le résultat de Benveniste *et al* pour l'analyse de leurs performances asymptotiques, il est intéressant d'étudier leur stabilité asymptotique qui est caractérisée par le signe des parties réelles des valeurs propres de la matrice R .

Pour une adaptation lente, la convergence de l'algorithme de structure propre (2) peut être décrite par une solution $\Theta(t)$ de son ODE associée sous la forme

$$\begin{aligned} \frac{d\Theta(t)}{dt} &= -\Gamma(\Theta) E \left\{ \underbrace{\begin{bmatrix} \mathbf{x}(n) \\ y_1(n) \end{bmatrix} y_2(n) |_{\Theta}}_{H(\Theta(n), \mathbf{x}(n))} \right\} \\ &= -\Gamma(\Theta) Q_1(\Theta) K(\Theta) q_2^T(\Theta) \end{aligned} \quad (6)$$

où

$$K(\Theta) \triangleq E \left\{ \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{u}(n) \end{bmatrix} \begin{bmatrix} \mathbf{x}^T(n) & \mathbf{u}(n) \end{bmatrix} |_{\Theta} \right\} \quad (7)$$

Pour un pas d'adaptation μ suffisamment petit, $\Theta(n)$ converge vers un point convergent Θ_* de l'algorithme (2) en probabilité, si et seulement si la même valeur de Θ_* est un point attracteur de l'équation différentielle (6), ce qui donne en particulier, $\frac{d\Theta(t)}{dt}|_{\Theta=\Theta_*} = 0$.

De même, la convergence de l'algorithme de modulation (3) peut être étudiée par son ODE :

$$\begin{aligned} \frac{d\Theta(t)}{dt} &= -\frac{\Gamma(\Theta)}{2} E \left\{ \underbrace{y_2(n) \begin{bmatrix} \mathbf{x}(n) \\ y_1(n) \end{bmatrix} - \hat{y}_2(n) \begin{bmatrix} \mathbf{x}(n) \\ \hat{y}_1(n) \end{bmatrix}}_{H(\Theta(n), \mathbf{x}(n))} \right\} \\ &= -\Gamma(\Theta) \left\{ Q_1(\Theta) \frac{1}{2} \underbrace{[K(\Theta) - \hat{K}(\Theta)] q_2^T(\Theta)}_{K_0(\Theta)} \right\} \\ &= -\Gamma(\Theta) Q_1(\Theta) K_0(\Theta) q_2^T(\Theta) \end{aligned}$$

En remplaçant $K_g(\Theta)$ à $K(\Theta)$ dans (7) ou $K_0(\Theta)$ dans (8), on voit que (6) et (8) peuvent être écrites sous la même forme

$$\frac{d\Theta(t)}{dt} = -\Gamma(\Theta) Q_1(\Theta) K_g(\Theta) q_2^T(\Theta)$$

Un point stationnaire Θ_* des algorithmes, où l'équation ci-dessus est égale à zéro, est attendu, si et seulement si le vecteur de la colonne $K_g(\Theta_*) q_2^T(\Theta_*)$ est orthogonal aux M lignes de $Q_1(\Theta_*)$, c'est-à-dire,

$$K_g(\Theta_*) q_2^T(\Theta_*) = \lambda (K_g(\Theta_*)) q_2^T(\Theta_*) \quad (9)$$

C'est pourquoi on appelle (2) et (3) les algorithmes de structure propre. Au point stationnaire Θ_* , la dérivée du vecteur champ moyen A est résumée par le résultat suivant :

Lemme 2.1 — Soit Θ_* tel que $K_g(\Theta_*) q_2^T(\Theta_*) = \lambda q_2^T(\Theta_*)$, alors

$$\begin{aligned} R &= -\Gamma^2(\Theta_*) Q_1(\Theta_*) \left[(K_g(\Theta_*) - \lambda I) \right] Q_1^T(\Theta_*) \\ &\quad - \Gamma(\Theta_*) Q_1(\Theta_*) \left[\frac{\partial K_g(\Theta_*)}{\partial \theta_0} \dots \frac{\partial K_g(\Theta_*)}{\partial \theta_{M-1}} \right] q_2^T(\Theta_*) \end{aligned}$$

On voit que le signe du premier terme est négatif, cependant, on ne sait pas déduire le signe du deuxième terme à cause de la complexité de calcul. La simulation montre que $\lambda = \lambda_{\min}(K_g(\Theta_*))$ et que le signe des valeurs propres de la matrice R est strictement négatif. Cela assure la stabilité des algorithmes et il en résulte que les matrices R et S des deux algorithmes sont diagonalisables. Dans ce cas, on peut obtenir une expression analytique de la variance asymptotique des paramètres $(\Theta(n) - \Theta_*)$ comme suit :

Propriété 1 — Si la matrice R est diagonalisable avec $Rv_i = \lambda_i v_i$, pour $1 \leq i \leq M$, on a

$$E\|\varepsilon(n)\|^2 \approx \mu \text{Tr}(P) = -\mu \sum_{i=1}^M \frac{v_i^T S v_i}{2\lambda_i}$$

où Tr est la trace de la matrice.

A l'aide du résultat ci-dessus, on peut évaluer la précision de la variance de la deuxième sortie $E[y_2^2(n)]$ autour d'un point stationnaire Θ_* . Le résultat est résumé par la propriété suivante :

Propriété 2 — La variance asymptotique de la deuxième sortie peut s'écrire

$$E[(y_2(n, \Theta) - y_2(n, \Theta_*))^2] \approx \mu \text{Tr}(PH) + o(\sqrt{\mu})$$

où

$$H = \begin{bmatrix} E\left(\frac{\partial y_2}{\partial \theta_0} \frac{\partial y_2}{\partial \theta_0}\right) & E\left(\frac{\partial y_2}{\partial \theta_0} \frac{\partial y_2}{\partial \theta_1}\right) & \dots & E\left(\frac{\partial y_2}{\partial \theta_0} \frac{\partial y_2}{\partial \theta_M}\right) \\ E\left(\frac{\partial y_2}{\partial \theta_1} \frac{\partial y_2}{\partial \theta_0}\right) & E\left(\frac{\partial y_2}{\partial \theta_1} \frac{\partial y_2}{\partial \theta_1}\right) & \dots & E\left(\frac{\partial y_2}{\partial \theta_1} \frac{\partial y_2}{\partial \theta_M}\right) \\ \vdots & \vdots & \ddots & \vdots \\ E\left(\frac{\partial y_2}{\partial \theta_M} \frac{\partial y_2}{\partial \theta_0}\right) & E\left(\frac{\partial y_2}{\partial \theta_M} \frac{\partial y_2}{\partial \theta_1}\right) & \dots & E\left(\frac{\partial y_2}{\partial \theta_M} \frac{\partial y_2}{\partial \theta_M}\right) \end{bmatrix} \Big|_{\Theta=\Theta_*}$$

3 Exemples de simulation

Dans la section précédente, on a obtenu les expressions analytiques à propos des variances asymptotiques des paramètres et de la deuxième sortie des deux algorithmes. Pour comparer les performances asymptotiques de ces deux algorithmes, on recourt à la simulation.

On applique les deux algorithmes pour une séquence générée par une fonction de transfert rationnelle sous la forme :

$$u(n) + au(n-1) = b(n) \quad a = 0.5$$

¹La corrélation à des instants dont l'écart est supérieur à m est nulle.

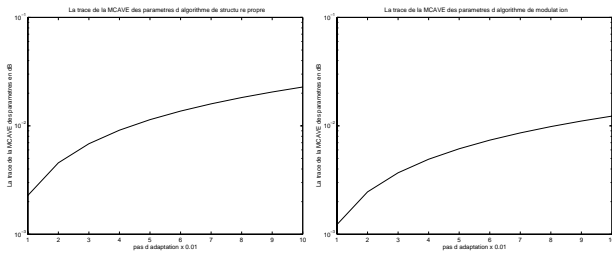


FIG. 2 — La trace de la MCAVE des paramètres $(\Theta(n) - \Theta_*)$ de l'algorithme de structure propre (à gauche) et de l'algorithme de modulation (à droite)

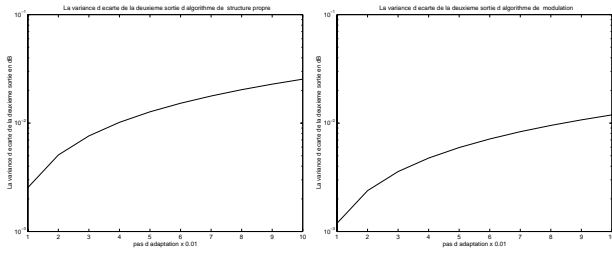


FIG. 3 — La variance asymptotique de l'écart de la sortie $E [y_2(n, \Theta) - y_2(n, \Theta_*)]^2$ de l'algorithme de structure propre (à gauche) et de l'algorithme de modulation (à droite)

où $b(n)$ est un bruit blanc gaussien de variance 1. Pour calculer la matrice S , dans ce cas, on peut supposer que $u(n)$ est m-dépendant avec $m = 180$, c.à.d., $E \{u(i)u(j)\} = 0$, pour $|i - j| \geq 180$.

La trace de MCAVE des paramètres $(\Theta(n) - \Theta_*)$ de l'algorithme de structure propre est égale à 0.0684, et la variance de l'écart de la sortie $E [y_2(n, \Theta) - y_2(n, \Theta_*)]^2$ est 0.0762 pour $\mu = 0.3$. Avec la même valeur μ , on a obtenu la trace de MCAVE des paramètres à $(\Theta(n) - \Theta_*)$ de l'algorithme de modulation qui équivaut 0.0369, et la variance de l'écart de la sortie $E [y_2(n, \Theta) - y_2(n, \Theta_*)]^2$ 0.0357. On présente la trace (en dB) de la MCAVE des paramètres du premier algorithme et du deuxième algorithme sur la figure 2 et la variance asymptotique de l'écart de la sortie $E [y_2(n, \Theta) - y_2(n, \Theta_*)]^2$ de ces deux algorithmes sur la figure 3. La simulation montre que la variance des paramètres $(\Theta(n) - \Theta_*)$ et celle de la sortie $E [y_2(n, \Theta) - y_2(n, \Theta_*)]^2$ du deuxième algorithme sont plus petites que celles du premier algorithme.

4 Conclusion

Dans cet article, on a appliqué un résultat de Benveniste *et al.* pour analyser les performances asymptotiques de deux algorithmes. On a obtenu une expression analytique de la variance des paramètres $(\Theta(n) - \Theta_*)$ et celle de la deuxième sortie $(y_2(n, \Theta) - y_2(n, \Theta_*))$. Les simulations que l'on a effectuées montrent que les valeurs propres de la matrice R des deux algorithmes ont une partie réelle négative, cela assure la stabilité des algorithmes et montre aussi que la variance asymptotique des paramètres $(\Theta(n) - \Theta_*)$ et celle de la sortie

$(y_2(n, \Theta) - y_2(n, \Theta_*))$ du deuxième algorithme sont plus petites que celles du premier algorithme. Ceci aide l'utilisateur à choisir un meilleur algorithme.

Références

- [1] P. Delsarte, B. Macq, and D. T. M. Slock. Signal adapted multiresolution transform for image coding. *IEEE Trans. Information Theory*, 38(2) :897–904, March 1992.
- [2] B. Macq and J. Y. Mertes. Optimization of linear multiresolution transforms for scene adaptive coding. *IEEE Trans. Signal Processing*, 41(12) :3568–3572, Dec. 1993.
- [3] P.A. Regalia and D. Y. Huang. Attainable error bounds in multirate adaptive lossless fir filters. In *Proc. IEEE int. Conf. Acoust., Speech, Signal Processing*, volume V, pages 1460–1463, Detroit, Michigan, May 1995. IEEE.
- [4] D.-Y. Huang and P.A. Regalia. An adaptive projection algorithm for multirate filter bank optimization. *EU-SIPCO*, pp.771-774, Trieste, Italy, Sept. 10-13, 1996.
- [5] D.-Y. Huang, P. A. Regalia and M. Bellanger. Comparison of Two Eigenstructure Algorithms for Lossless Multirate Filter Optimization. In *Proc. IEEE int. Conf. Acoust., Speech, Signal Processing*, Vol. 4, pp. 2437-2441, Munich, Germany, April 1997. IEEE.
- [6] A. Benveniste, M. Metivier, and P. Priouret. *Algorithmes Adaptatifs et Approximation Stochastiques*. Masson, Paris, 1987.
- [7] L. Ljung and T. Söderström. *Theory and practice of recursive identification*. MIT Press, 1983.