

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Nationale Polytechnique



Département d'Electronique

Laboratoire Signal & Communications

Thèse de Doctorat en Electronique

Présentée par :

MADDI Abdelkader

***ALGORITHMES IMPLEMENTABLES D'IDENTIFICATION
PARAMETRIQUE : APPLICATION AU SIGNAL DE PAROLE
BRUITEE***

Devant le jury :

- Président** : M^r F. BOUDJEMA, Professeur à l'ENP.
Examineur : M^r M. BENSEBTI, Professeur à l'USD de Blida.
Examineur : M^r H. SALHI, Maître de conférence à l'USD de Blida.
Examinatrice : M^{me} M. GUERTI, Professeur à l'ENP.
Rapporteur : M^r A. GUESSOUM, Professeur à l'USD de Blida.
M^r D. BERKANI, Professeur à l'ENP.

Février 2008

Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance et mes sincères remerciements à M^r A. GUESSOUM, Professeur à l'Université Saad Dahlab de Blida (USDB) et Directeur du Laboratoire de Traitement de Signal et Imagerie (LATSI), pour son aide déterminante et pour ses conseils précieux en sa qualité de directeur de thèse qui ont largement contribué dans l'avancement de mes travaux de recherche et l'amélioration de cette thèse de Doctorat.

J'adresse également mes remerciements à M^r D. BERKANI, Professeur à l'Ecole Nationale Polytechnique (ENP), pour avoir accepté la direction de cette thèse et d'avoir suivi et contribué avec intérêt à l'évolution de ce travail de recherche. Je lui exprime aussi ma gratitude pour ses précieux conseils.

J'exprime ma gratitude à M^r F. BOUDJEMA, le Président du jury, d'avoir bien voulu présider mon jury.

Je suis très reconnaissant aux membres du jury :

- M^r M. BENSEBTI, Professeur à l'USDB.
- M^r H. SALHI, Maître de conférence à l'USDB.
- M^{me} M. GUERTI, Professeur à l'ENP.

D'avoir accepté d'être membres de la commission d'examen.

Je remercie chacun des membres de l'équipe Parole du Laboratoire LATSI pour leur amitié, leur aide et leur soutien.

Je tiens à exprimer ma gratitude et mon amitié à Mr M. DJENDI, Enseignant Chercheur à l'USDB, pour m'avoir soutenu dans les moments difficiles. Qu'il trouve ici ma reconnaissance et mon respect.

ملخص

إن لخوارزميات المطابقة علاقة جد وطيدة بطرق الإحصاء و طبيعة النماذج الرياضية. لهذا تعتبر المرحلة الأساسية في طريقة المطابقة مبنية على المعرفة الجيدة لكل نظام فيزيائي. و في إطار هذه الأطروحة لنيل شهادة الدكتوراه نقوم بدراسة و انجاز الخوارزميات التالية : LMS RLS و نيوتن لمطابقة الأبعاد القائمة على تبييض الخطأ للحصول على خطأ يؤول إلى الضوضاء البيضاء , علما أن أغلبية الأنظمة الفيزيائية تمثل وفق نموذج ARMA, ARMAX او ARX . وفي الأخير نقترح صيغة متطورة لخوارزمية RELS لمطابقة نموذج ARMAX ثم نقوم بتطبيقها على الإشارة الصوتية التي ستكون ثمرة هذا العمل.

كلمات المفتاح : طريقة المطابقة - نموذج رياضي - خوارزمية R SL - خوارزمية التفاضل - فارق - إشارة صوتية.

Résumé

L'identification de systèmes est fortement liée aux méthodes statistiques et à la construction de modèles décrivant de manière mathématique les processus physiques. La principale étape d'identification se traduit donc par une meilleure connaissance du système.

Dans cette Thèse de Doctorat, nous nous sommes intéressés plus particulièrement, à des méthodes d'identification paramétrique. Les méthodes que nous présentons sont : Gradient Stochastique, Newton Stochastique et Moindres Carrés pour identifier sans biais les paramètres d'un processus physique basées sur le blanchissement de l'erreur de prédiction. Elles consistent à déterminer, d'une façon récursive, un modèle mathématique linéaire, basé sur les données d'observation, pour représenter le système considéré. La plupart des systèmes peuvent être représentés par des modèles ARMAX, ARMA ou ARX.

Enfin, nous proposons la méthode des Moindres Carrés Etendue pour estimer simultanément les paramètres du modèle ARMAX. Une application de cette méthode a été réalisée pour identifier les paramètres du modèle ARX de la parole dans un contexte bruité.

Mots-clés : *Identification, Modèle Mathématique, Moindres Carrés, Gradient Stochastique, Biais, Signal de parole.*

Abstract

The identification of systems is strongly related on the statistical methods and the construction of models describing in a mathematical way the physical processes. The principal stage of identification thus results in a better knowledge of the system.

In this Doctoral Thesis, we present the parametric identification methods, which are the stochastic gradient method, stochastic Newton and recursive least squares method to identify with unbiased estimates parameters of a physical process based on whitening error of prediction. They consist in determining, in a recursive way, a linear mathematical model, based on the data of observation, to represent the system considered. The majority systems can be represented by ARMAX, ARMA or ARX models.

Finally, we propose the Recursive Extended Least Squares method to estimate the parameters of the ARMAX model. An application of this method is realised to identify the parameters of ARX model corresponding to the noisy speech signal.

Keywords : *Identification, Mathematic Model, Least Squares, Stochastic Gradient, Bias, Speech Signal.*

Liste des Abréviations

AR : Autorégressif

ARMA : Autorégressif à Moyenne Ajustée

ARMAX : Autorégressif à Moyenne Ajustée avec entrée auxiliaire

E : Espérance

EM : Expectation Minimisation

EQM : Erreur Quadratique Moyenne

FA : Fonction Aléatoire

FIR : Finite Impulse Response

IIR: Infinite Impulse Response

LMS: Least Mean Squares

LPC : Linear Predictive Coding

MA : Moyenne Ajustée

ML: Maximum Likelihood

MSE: Mean Square Error

NLMS: Normalized Least Mean Squares

SN: Stochastic Newton

RLS: Recursive Least Squares

SBPA : Séquence Binaires Pseudo Aléatoires

SNR : Signal Noise Ration

VAR : Variance

Liste des Figures

Fig.I.1 : Génération d'un signal aléatoire corrélé.....	9
Fig.I.2 : Estimateur biaisé de faible variance.....	13
Fig.I.3 : Estimateur non biaisé de forte variance.....	14
Fig. I.4 : Deux segments de voix, l'un non voisé et l'autre voisé.....	17
Fig. 1.5 : Modèle simplifié de production de la parole.....	19
Fig. 1.6 : Analogie entre le modèle physique et le modèle acoustique.....	19
Fig. II.1 : Schéma d'identification paramétrique.....	27
Fig. III.1 : Représentation graphique de la méthode de Newton.....	34
Fig.III.2 : Structure de réalisation d'un modèle ARMA.....	38
Fig.III.3 : Structure du modèle ARMAX.....	42
Fig. IV.1 : Séquence d'entrée SBPA.....	47
Fig. IV.2 : Bruit additif gaussien et sa fonction d'autocorrelation	47
Fig. IV.3 : Menu simple des algorithmes implémentés sous MATLAB.....	49
Fig. IV.4 : (a) - Evolution des paramètres ($N = 256*4, \mu = 0.0145$), (b) - Autocorrelation de l'erreur de prédiction.....	50
Fig. IV.5 : (a) - Convergence lente $\mu = 0.0097$, (b) – Convergence Rapide $\mu = 0.0193$	51
Fig. IV.6 : (a) - Evolution des paramètres ($N = 256*4, \sigma^2 = 0.01$), (b) - Autocorrelation de l'erreur de prédiction.....	52
Fig. IV.7 : (a) - Convergence lente $\mu = 0.0033$, (b) - Convergence rapide $\mu = 0.0066$	53
Fig. IV.8 : Evolution des paramètres ($N = 256*10, \sigma^2 = 0.25$).....	55
Fig. IV.9 : (a) - Erreur de prédiction ($N = 256*10, \sigma^2 = 0.25$), (b) - Autocorrelation de l'erreur de prédiction.....	55
Fig. IV.10 : Erreur d'identification par l'algorithme SN au cas d'un système à phase minimale.....	56

Fig. IV.11 : Erreur d'identification par l'algorithme SN au cas d'un système à phase non minimale.....	57
Fig. IV.12 : (a) - Evolution des paramètres ($N = 256*4, \sigma^2 = 0.25$), (b) - Autocorrelation de l'erreur de prédiction dans le cas d'un système à phase minimale.....	59
Fig. IV.13 : (a) - Evolution des paramètres ($N = 256*4, \sigma^2 = 0.25$), (b) - Autocorrelation de l'erreur de prédiction dans le cas d'un système à phase non minimale.....	60
Fig. IV.14 : Evolution des paramètres ($N = 256*10, \sigma^2 = 0.25$) dans le cas du modèle ARMAX.....	63
Fig. IV.15 : (a) - Erreur de prédiction ($N = 256*10, \sigma^2 = 0.25$), (b) - Autocorrelation de l'erreur de prédiction. Cas du modèle ARMAX.....	63
Fig. IV.16 : (a), (b) - Erreur d'identification des paramètres par l'algorithme RELS.....	65
Fig.V.1 : Exemples de parole (Signal (x1) : locuteur en haut, locutrice en bas)....	64
Fig.V.2 : Exemples de parole (Signal (x2) : locuteur en haut, locutrice en bas)	67
Fig.V.3 : Exemples de son voisé (droit) et non voisé (gauche).....	68
Fig.V.4 : Tranche voisée de 32 ms du signal (M1) et son spectre de puissance.	69
Fig.V.5 : Tranche non voisée de 32 ms du signal (M1) et son spectre de puissance.	70
Fig.V.6 : Tranche de 32 ms du signal (M1) et son spectre de puissance. Cas : RLS.....	72
Fig.V.7 : Tranche de 32 ms du signal (F1) et son spectre de puissance. Cas : RLS.....	73
Fig.V.8 : Tranche de 32 ms du signal (M2) et son spectre de puissance. Cas : RLS	74
Fig.V.9 : Tranche de 32 ms du signal (F2) et son spectre de puissance. Cas : RLS ...	75
Fig.V.10 : Evolution temporelle du paramètre a_1 et a_2 du signal (M1)	76
Fig.V.11 : Evolution temporelle du paramètre a_4 et a_5 du signal (M1).....	76
Fig. V.12 : Bruit gaussien de moyenne nulle et variance égale à l'unité.....	77
Fig. V.13 : Taux de l'EQM des signaux de parole.....	83

Liste des Tableaux

Tab. IV.1 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.01, \mu = 0.0145$).....	50
Tab. IV.2 : Influence du gain d'adaptation μ , ($\sigma^2 = 0.01, N = 256*4$).....	51
Tab. IV.3 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.01, \mu = 0.0050$).....	52
Tab. IV.4 : Influence du gain d'adaptation μ ($N=256*10, \sigma^2 = 0.01$).....	53
Tab.IV.5 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).....	54
Tab. IV.6 : Influence de la variance σ^2 , ($N = 256*10$).....	54
Tab. IV.7 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)	56
Tab. IV.8 : Influence du variance σ^2 , ($N = 256*10$).....	57
Tab.IV.9 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).....	58
Tab. IV.10 : Influence de la variance σ^2 , ($N=256*10$).....	58
Tab. IV.11 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).....	59
Tab. IV.12 : Influence de la variance σ^2 , ($N=256*4$).....	60
Tab. IV.13 : Biais des estimateurs utilisés, $N = 256*4, \sigma^2 = 0.25$	61
Tab. IV.14 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).....	62
Tab. IV.15 : Influence du variance σ^2 , ($N = 256*10$).....	62
Tab.V.1 : Propriété de l'estimateur RLS. Cas : Tranche de 32 ms du signal (M1)	72
Tab.V.2 : Propriété de l'estimateur RLS. Cas : Tranche de 32 ms du signal (F1)	73
Tab.V.3 : Propriété de l'estimateur RLS. Cas : Tranche de 32 ms du signal (M2)	74
Tab.V.4 : Propriété de l'estimateur RLS. Cas : Tranche de 32 ms du signal (F2)	75
Tab.V.5 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M1), SNR = 40dB.....	78
Tab.V.6 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M1), SNR = 30dB.....	78
Tab.V.7 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M1), SNR = 20dB.....	79

Tab.V.8 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M1), SNR = 10dB.....	79
Tab.V.9 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (F1), SNR = 40dB.....	80
Tab.V.10 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (F1), SNR = 10dB.....	80
Tab.V.11 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M2), SNR = 40dB.....	81
Tab.V.12 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (M2), SNR = 10dB.	81
Tab.V.13 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (F2), SNR = 40dB.	82
Tab.V.14 : Estimation des paramètres du modèle AR par l'algorithme RELS. Cas : Tranche de 32 ms du signal (F2), SNR = 10dB.....	82
Tab.V.15 : Etude comparative des EQM. Cas des signaux faiblement bruités.....	83
Tab.V.16 : Etude comparative des EQM. Cas des signaux fortement bruités.....	83
Tab.V.17 : Modélisation ARMA de la parole par l'estimateur RLS. Cas: Tranche de 32 ms du signal (M1).....	84

Sommaire

Remerciement.....	i
Résumé.....	ii
Liste des Abréviations.....	iii
Liste des Figures.....	iv
Liste des Tableaux.....	vi

INTRODUCTION GENERALE.....	1
----------------------------	---

CHAPITRE I :
NOTIONS DE PROCESSUS STOCHASTIQUES
ET LA PAROLE

I.1. Introduction.....	4
I.2. Notions de processus stochastique.....	4
I.2.1. Densité de probabilité d'une fonction aléatoire en temps continu.....	4
I.2.2. Indépendance.....	5
I.2.3. Moment d'une fonction aléatoire.....	5
I.2.3.1. Moyenne.....	5
I.2.3.2. Variance et covariance.....	5
I.2.3.3. Covariance mutuelle.....	6
I.2.4. Stationnarité et Ergodicité.....	6
I.2.4.1. Stationnarité.....	6
I.2.4.2. Ergodicité.....	7
I.2.5. Densité spectrale de puissance.....	9
I.2.6. Modèles stochastiques.....	9
I.2.7. Processus stochastiques remarquables.....	9
I.2.7.1. Processus Gaussien.....	10
I.2.7.2. Processus Markovien.....	10
I.3. Notions de la théorie d'estimation.....	10
I.3.1. Estimation d'une distribution.....	11
I.3.2. Estimation des paramètre d'une distribution.....	11
I.3.3. Estimation des paramètre d'un modèle.....	11
I.3.4. Estimateurs et estimations.....	12
I.3.5. Propriétés d'un bon estimateur.....	12

I.4. Notions de bruits.....	15
I.4.1. Bruit blanc.....	15
I.4.2. Séquence binaire pseudo aléatoire.....	15
I.4.3. Bruit coloré.....	15
I.5. Notions sur le signal de parole.....	16
I.5.1. Description du signal vocal.....	16
I.5.2. Détection de la fréquence fondamentale de la parole.....	16
I.5.3. Classification des sons.....	16
I.5.4. Modèle de production de la parole.....	18
I.6. Conclusion.....	20

CHAPITRE II :
MODELISATION ET IDENTIFICATION
DES PROCESSUS PHYSIQUES

II.1. Introduction.....	21
II.2. Modélisation des processus physiques.....	21
II.2.1. Différentes approches d'établissement d'un modèle.....	22
II.2.2.1. Modèle de connaissance.....	22
II.2.2.2. Modèle de représentation.....	22
II.2.2. Modélisation autorégressive à moyenne ajustée (ARMA).....	22
II.2.3.1. Modélisation autorégressive (AR).....	23
II.2.3.2. Modélisation à Moyenne Ajustée (MA).....	23
II.2.3. Principales étapes de la modélisation.....	24
II.2.4.1. Système.....	24
II.2.4.2. Modèle.....	24
II.2.4.3. Critère.....	25
II.2.4.4. Optimisation.....	25
II.2.4.5. Incertitude sur les paramètres.....	25
II.2.4.6. Analyse critique des résultats obtenus.....	25
II.2.4. Difficultés de la modélisation.....	25
II.2.4.1. Définition du problème de modélisation.....	25
II.2.4.2. Qualité des données de modélisation.....	25
II.2.4.3. Choix de la technique utilisée.....	26
II.2.4.4. Choix des variables disponibles.....	26
II.2.4.5. Sélection de modèle adéquat.....	26
II.3. Identification des paramètres d'un processus physique.....	27
II.3.1. Identification non paramétrique.....	27
II.3.2. Identification paramétrique.....	27
II.3.3. Etapes d'identification.....	28
II.3.2.1. Acquisition des entrées et sorties.....	28
II.3.2.2. Choix de la complexité du modèle.....	28
II.3.2.3. Estimation des paramètres du modèle.....	29
II.3.2.4. Validation du modèle.....	29
II.3.4. Algorithme d'identification récursif.....	29
II.4. Conclusion.....	29

**CHAPITRE III :
ETUDE ET DEVELOPPEMENT DES ALGORITHMES
D'IDENTIFICATION PARAMETRIQUE**

III.1. Introduction.....	30
III.2. Algorithme du gradient stochastique.....	31
III.2.1. Rappel sur la méthode du gradient.....	31
III.2.2. Identification d'un modèle ARMA.....	31
III.2.3. Les avantages de la méthode du gradient.....	33
III.3. Algorithme de Newton stochastique.....	34
III.3.1. Principe de la méthode.....	34
III.3.2. Equations de mise en œuvre de la méthode de Newton.....	35
III.3.3. Identification du modèle ARMA.....	36
III.4. Algorithme des moindres carrés.....	37
III.4.1. Identification du modèle ARMA.....	37
III.4.2. Solution optimale au sens des moindres carrés.....	38
III.4.3. Algorithme des moindres carrés récursif.....	39
III.4.4. Propriétés de l'estimateur des moindres carrés.....	41
III.5. Algorithme des moindres carrés étendu RELS.....	41
III.5.1. Structure d'un modèle ARMAX.....	42
III.5.2. Détermination des paramètres.....	43
III.5.3. Les équations mise en œuvre de l'algorithme RELS.....	44
III.5.4. Initialisation de l'algorithme RELS.....	44
III.6. Conclusion.....	45

**CHAPITRE IV :
IMPLEMENTATION DES ALGORITHMES
ET RESULTATS DE SIMULATION**

IV.1. Introduction.....	46
IV.2. Etapes de simulation.....	46
IV.2.1. Choix des paramètres du système	46
IV.2.2. Choix des signaux de test.....	46
IV.2.3. Identification des paramètres du modèle.....	48
IV.2.4. Test de validation.....	48
IV.3. Mise en œuvre des algorithmes implémentés.....	49
IV.4. Identification des paramètres par l'algorithme LMS.....	49
IV.4.1. Système à phase minimale.....	49
IV.4.2. Système à phase non minimale.....	52
IV.5. Identification des paramètres par l'algorithme de Newton stochastique.....	54
IV.5.1. Système à phase minimale.....	54
IV.5.2. Système à phase non minimale.....	56
IV.6. Identification des paramètres par l'algorithme RLS.....	58
IV.6.1. Système à phase minimale.....	59
IV.6.2. Système à phase non minimale.....	58

IV.7. Etude comparative des algorithmes d'identification paramétrique.....	61
IV.8. Identification des paramètres d'un modèle ARMAX.....	62
IV.9. Interprétation des résultats obtenus.....	64
IV.10. Conclusion.....	65

CHAPITRE V :
APPLICATION AU SIGNAL DE PAROLE

V.1. Introduction.....	66
V.2. Base de données.....	66
V.3. Exemples de son voisé et non voisé.....	68
V.4. Transformée de Fourier à court terme.....	68
V.5. Modélisation de la parole non bruitée.....	70
V.6. Modélisation de la parole bruitée.....	77
V.7. Etude comparative.....	83
V.8. Modélisation ARMA de la parole non bruitée.....	84
V.9. Conclusion.....	85

CONCLUSION GENERALE ET PERSPECTIVES.....	87
------------------------------------------	----

Références bibliographie.....	89
-------------------------------	----

Biographie

Introduction Générale

1. Problématique

L'intérêt des approches d'identification paramétrique fait l'objet de nombreux développements ces dernières années. Les résultats théoriques établis concernent généralement les performances asymptotiques des méthodes dans le cas de signaux d'entrée connus et suffisamment excitants. Le cas plus réaliste de données faiblement informatives, en nombre fini ou avec des incertitudes non seulement sur les signaux de sortie mais aussi sur les signaux d'entrée a reçu moins d'attention jusqu'à présent ; de ce fait, de nombreux problèmes théoriques et pratiques restent ouverts.

Le sujet proposé est motivé par les problèmes rencontrés lors des récents travaux de recherche effectués sur l'analyse des processus physiques. Dans ce domaine applicatif, la recherche des modèles à temps discrets ayant un sens physique est primordiale. En outre, les récentes évolutions technologiques pour l'acquisition de données permettent aujourd'hui d'envisager le développement de nouvelles techniques de modélisation et d'identification des processus physiques.

Différents travaux concernant cette classe de traitements pour l'identification et/ou la reconstruction de signaux ont été menés. L'algorithme de type LMS permet effectivement un traitement en temps réel mais conduit à une estimation biaisée des paramètres du modèle ARMA. L'algorithme EM est une adaptation d'un algorithme hors ligne, il souffre d'une grande complexité de calcul. Par contre, il fournit une estimation non biaisée des paramètres du modèle ARMA.

D'autres auteurs introduisent une autre catégorie d'algorithmes pour supprimer le bruit [1]. Ils considèrent que chaque segment du signal est stationnaire et modélisé par un système autorégressif d'ordre p . La plupart des modèles temps-fréquences de la parole sont construits sur un modèle linéaire fonctionnel très simple de production de la parole décrit par [Markel & Gray, 1976].

[Makhoul, 1977] a établi les coefficients de prédiction dans le domaine fréquentiel et temporel. Il a traité les cas où le signal est considéré déterministe (voisement) ou non déterministe (bruit, non voisé) en abordant également la question de la stabilité et du gain du système.

[Lim et Oppenheim, 1978 - 1983] cherchent à estimer les paramètres du modèle AR et le signal de parole non bruitée, en utilisant une méthode d'estimation du maximum à posteriori. Une estimation sous optimale est obtenue à l'aide de l'algorithme EM [Dempster et Al., 1977], par un procédé itératif. Les paramètres du modèle AR sont tout d'abord déterminés en supposant le signal non bruité connu. Une estimation du signal bruité est ensuite calculée en utilisant le modèle non bruité estimé et en supposant connue l'estimation de la densité spectrale de puissance du bruit. Cependant, si un vecteur de parole non bruitée ainsi que le modèle AR sont estimés à partir du vecteur correspondant de parole bruitée, le nombre d'inconnues (nombre d'échantillons de parole non bruitée + nombre de coefficients du modèle AR) est supérieur au nombre d'observations bruitées. Il est donc impossible d'obtenir une faible variance à la fois pour l'estimateur des paramètres du modèle AR et pour l'estimateur du signal non bruité. Donc, la méthode de Lim et Oppenheim est limitée au bruit blanc gaussien; [Hansen et Clements, 1985] l'étendent pour des bruits colorés. Une variante de l'approche de Lim et Oppenheim est développée dans [Paliwal et Basu, 1987], où le modèle AR est d'abord estimé à partir du signal bruité, puis utilisé pour estimer le signal non bruité.

[Paliwal et Basu, 1987] supposent que le bruit est blanc, centré, additif et non corrélé avec le signal original. Premièrement, ils estiment les p coefficients du modèle AR par une des méthodes conventionnelles. Ensuite, ils appliquent l'algorithme du filtrage de Kalman en utilisant les p coefficients estimés auparavant pour estimer de nouveau le signal de parole.

[Koo, 1989] et [Baillargeat, 1991] ont considéré une autre approche plus réaliste qui introduit une modélisation autorégressive pour le bruit également. Macaulay [1980] et Yang [1993] ont suggéré d'autres techniques utilisant les notions probabilistes pour estimer le spectre d'énergie.

L'ensemble des algorithmes qu'on a cité jusqu'à présent analyse le signal parvenant d'une seule source. Il existe d'autres catégories d'algorithmes qui utilisent deux sources d'informations ou plus pendant le processus d'analyse.

2. Solutions proposées

L'utilisation des méthodes d'identification traditionnelles sans bruit conduit souvent à des estimations biaisées. C'est pourquoi, le problème d'identification de systèmes dans un contexte dénommé bruité est reconnu pour être beaucoup plus délicat à traiter et reste encore largement ouvert.

Des algorithmes d'identification paramétrique ayant une formulation récursive adaptée aux problèmes d'identification en temps réel et leurs mises en œuvre sur micro-ordinateur, ont été développés. Une première partie de la thèse concerne la mise au point des algorithmes de type Newton stochastique, Gradient stochastique et Moindres carrés ordinaire, dédiés à l'identification des paramètres du modèle ARMA.

Ensuite, nous proposons dans ce travail de thèse de Doctorat un algorithme récursif RELS permettant une estimation non biaisée des paramètres du modèle ARMAX. Cette méthode est fondée sur la famille des moindres carrés récursifs qui généralise l'algorithme RLS au cas des signaux discrets. L'application concernée, est la modélisation d'un signal de parole dans un environnement bruité.

3. Organisation de la thèse

Le premier chapitre de cette thèse est un rappel bref sur des notions et généralités de processus stochastiques et la parole.

Le second chapitre est réservé pour la modélisation et l'identification paramétrique des processus physiques.

Dans le troisième chapitre, nous présentons l'étude et développement des algorithmes d'identification paramétriques tel que l'algorithme de Newton stochastique, Gradient stochastique et Moindres carrés récursifs. Une version étendue RELS est aussi obtenue pour identifier sans biais les paramètres d'un processus physique de type ARMAX.

Le quatrième chapitre est consacré à l'implémentation des algorithmes d'identification et aux résultats obtenus par simulation. Plusieurs tests ont été effectués sur différents processus physiques stables à phase minimale ou non minimale ainsi que leurs interprétations physiques.

Le dernier chapitre, est une application au signal de parole qui sera le fruit de notre travail. Enfin, en conclusion générale de ce travail, nous résumons les méthodes proposées et les résultats obtenus, et présentons des perspectives d'études.

En conclusion, le travail réalisé dans cette thèse a permis d'étudier et implémenter les algorithmes d'identification paramétrique et tester leurs performances. Les perturbations aléatoires qui entachent la sortie mesurée du processus physique peuvent causer un biais sur l'estimation des paramètres. Pour différents types de perturbations, il existe des méthodes appropriées d'identification récursive assurant asymptotiquement une estimation non biaisée des paramètres.

Chapitre I

Notions de Processus Stochastiques Et Parole

I.1. Introduction

Ce chapitre a pour objectif de donner aux lecteurs une autre vision du traitement qui peut être appliqué à des signaux ou des systèmes. Dans les cours du traitement du signal, les signaux et systèmes considérés et modélisés sont supposés connus parfaitement (de manière déterministe). Or, dans de nombreux systèmes réels, les signaux à traiter ne peuvent être modélisés de la sorte. C'est le cas par exemple des signaux de télécommunications qui, par nature, ne sont pas connus de façon certaine, sinon ils ne véhiculeraient aucune information. C'est aussi le cas de nombreux signaux qui proviennent de mesures ou d'observations bruitées, qui font que le signal réellement observé est une forme corrompue du signal voulu. De ce fait, de nombreuses opérations de traitement du signal pensées pour des signaux déterministes doivent être repensées et optimisées en fonction de l'information dont on dispose à propos du signal à traiter [2]. C'est l'objet de ce chapitre.

I.2. Notions de processus stochastique

Un processus stochastique ou processus aléatoire se définit par l'évolution au cours du temps d'un phénomène statistique selon des lois de probabilités. Cette fonction est appelée fonction aléatoire par opposition à une fonction certaine. En d'autres termes, un processus aléatoire est un mode de génération d'une fonction aléatoire. Si la variable indépendante dont dépend la fonction est continue, on parle d'un processus en temps continu. Si cette variable indépendante est l'ensemble des entiers, alors on parle de processus en temps discret. Ce second type de processus n'est donc rien d'autre qu'une séquence de variables aléatoires. Exemples de processus aléatoires: signal de parole, signal radar, bruit, etc. ...

I.2.1. Densité de probabilité d'une fonction aléatoire en temps continu

Soit une fonction aléatoire $X(t)$. En considérant l'instant particulier $t = t_1$, on obtient une VA notée $X(t_1)$. On note la densité de probabilité de cette variable aléatoire par $P_X[x(t_1)]$ pour rappeler la dépendance à l'égard de l'instant considéré. De la même façon on

note la fonction de répartition par $F_X[x(t_1)]$. En considérant des instants t_1, \dots, t_n , on obtient une variable aléatoire à n dimensions. Sa densité de probabilité est notée $P_X[x(t_1), \dots, x(t_n)]$. Cette densité est appelée densité de probabilité du n ème ordre de la fonction aléatoire. A partir d'une densité de probabilité d'un certain ordre, on en déduit celles d'ordres inférieurs par passage aux densités marginales. Les notions de la fonction aléatoire en temps discret sont tout à fait semblables à celles exposées pour des fonctions aléatoires en temps continu. On utilise la notation $P_X[x(n)]$.

I.2.2. Indépendance

On dit d'une FA qu'elle est à valeurs indépendantes si [3]

$$P_X[x(t_1), \dots, x(t_n)] = P_X[x(t_1)] \times \dots \times P_X[x(t_n)] \quad (\text{I.1})$$

A tout instant, le futur est indépendant du présent et du passé puisque

$$P_X[x(t_n) | x(t_1), \dots, x(t_{n-1})] = P_X[x(t_n)] \quad (\text{I.2})$$

I.2.3. Moment d'une fonction aléatoire

La connaissance totale d'une fonction aléatoire requiert que l'on connaisse les densités de probabilités de tous ordres. En pratique on accède très rarement à ce niveau de connaissance de la fonction aléatoire. La plupart des méthodes qui seront vues plus loin requièrent que soient connues les densités de probabilité des deux premiers ordres.

I.2.3.1. Moyenne

On appelle moyenne de la fonction aléatoire en temps continu $X(t)$, la fonction de t qui pour toute valeur de t donne la moyenne de la variable aléatoire obtenue à cet instant, soit

$$m_X(t) = E[X(t)] = \int_{-\infty}^{+\infty} x P_X[x(t)] dx \quad (\text{I.3})$$

I.2.3.2. Variance et covariance

De façon similaire, la variance d'une fonction aléatoire scalaire est une fonction de t définie par [2]:

$$\begin{aligned} \sigma_X^2(t) &= E\{[X(t) - m_X(t)][X(t) - m_X(t)]^*\} \\ &= \int_{-\infty}^{+\infty} [x - m_X(t)][x - m_X(t)]^* P_X[x(t)] dx \end{aligned} \quad (\text{I.4})$$

On appelle fonction de covariance $C_X(t_1, t_2)$ la fonction des 2 variables t_1 et t_2 donnée par :

$$\begin{aligned} C_X(t_1, t_2) &= E\left\{ [X(t_1) - m_X(t_1)][X(t_2) - m_X(t_2)]^* \right\} \\ &= \iint_{x_1, x_2} [x_1 - m_X(t_1)][x_2 - m_X(t_2)]^* P_X[x(t_1), x(t_2)] dx_1 dx_2 \end{aligned} \quad (I.5)$$

On a dès lors $\sigma_X^2 = C_X(t, t)$. Pour une fonction aléatoire réelle, on a $C_X(t_1, t_2) = C_X(t_2, t_1)$.

La fonction aléatoire réelle est non corrélée si $C_X(t_1, t_2) = \sigma_X^2 \delta(t_1 - t_2)$.

I.2.3.3. Covariance mutuelle

Deux fonctions aléatoires $X(t)$ et $Y(t)$ ont la fonction de covariance mutuelle donnée par [3] :

$$\begin{aligned} C_{XY}(t_1, t_2) &= E\left\{ [X(t_1) - m_X(t_1)][Y(t_2) - m_Y(t_2)]^* \right\} \\ &= \iint_{x, y} [x - m_X(t_1)][y - m_Y(t_2)]^* P_{XY}[x(t_1), y(t_2)] dx dy \end{aligned} \quad (I.6)$$

Par ailleurs, on a $C_{YX}(t_1, t_2) = C_{XY}^*(t_2, t_1)$

I.2.4. Stationnarité et Ergodicité

Ces propriétés sont très importantes, car si les fonctions aléatoires que l'on traite n'en bénéficieraient pas, nous serions vraiment limités dans l'interprétation et l'utilisation de celles ci.

I.2.4.1. Stationnarité

Dans le langage courant, la notion de stationnarité est utilisée pour un phénomène dont les caractéristiques essentielles ne se modifient pas au cours du temps : on parle de météo stationnaire, d'état stationnaire d'une maladie, etc. ... Le caractère stationnaire d'une fonction aléatoire est associé à cette acception du mot, mais se définit précisément au moyen de ses propriétés.

I.2.4.1.1. Stationnarité au sens strict

On dit d'une fonction aléatoire qu'elle est stationnaire au sens strict si toutes ses densités de probabilité (de tous ordres) ne dépendent pas de l'origine des temps. Si donc, t étant remplacé par $t + t_0$ avec t_0 quelconque, la densité de probabilité ne change pas. Ceci implique que la densité de probabilité d'ordre 1 ne dépende pas du temps $P_X[x(t)] = P_X(x)$.

I.2.4.1.2. Stationnarité au sens faible

La notion de stationnarité au sens strict est rarement utilisable en pratique car l'on dispose rarement de toutes les densités de probabilité d'une fonction aléatoire. On dit d'une F.A qu'elle est stationnaire au sens faible lorsque la moyenne ne dépend plus du temps, et la fonction de covariance ne dépend plus que de la différence entre les deux arguments, soit [3] :

$$\blacksquare \quad m_X(t) = m_X \quad (\text{I.7})$$

$$\blacksquare \quad C_X(t_1 t_2) = C_X(\tau = t_1 - t_2) \quad (\text{I.8})$$

De ce fait la variance est aussi constante puisque :

$$\sigma_X^2 = C_X(t, t) = C_X(0) = \sigma_X^2 \quad (\text{I.9})$$

I.2.4.2. Ergodicité

L'ergodicité est une propriété importante qui lie les moyennes statistiques et les moyennes temporelles. Tout comme pour la stationnarité, il existe différents types d'ergodisme. Un processus aléatoire est ergodique au sens strict, si tous les moments statistiques sont égaux aux moments temporels. Un processus aléatoire est ergodique au sens large (ou du second ordre), si il y a égalité des moyennes statistiques et temporelles ainsi que des fonctions d'autocorrelation [4]. Dans ce qui suit, on se limitera à étudier l'ergodisme pour des fonctions aléatoires stationnaires.

I.2.4.2.1. Ergodicité relativement à la moyenne

On définit la moyenne temporelle $\eta(t_0)$ sur un intervalle T d'une fonction aléatoire $X(t)$ par :

$$\eta_T(t_0) = \int_{t_0}^{t_0+T} X(t) dt \quad (\text{I.10})$$

Il s'agit clairement d'une V.A dépendant de l'origine choisie t_0 . On peut se demander ce que devient cette valeur pour $T \rightarrow \infty$. On dit d'une F.A qu'elle est ergodique relativement à la moyenne si, lorsque T tend vers l'infini, la moyenne temporelle est indépendante de t_0 et tend vers la moyenne (espérance mathématique, moyenne d'ensemble) :

$$\lim_{T \rightarrow \infty} \eta_T(t_0) = m_X \quad \forall t_0 \quad (\text{I.11})$$

On peut définir divers types de convergence : en probabilité, en moyenne quadratique, etc. La convergence en moyenne quadratique, ce qui signifie :

$$\lim_{T \rightarrow \infty} E \left[\left| \eta_T(t_0) - m_X \right|^2 \right] = 0 \quad (\text{I.12})$$

Pour une FA continue, une condition nécessaire et suffisante d'ergodisme est que :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) C_X(\tau) d\tau = 0 \quad (\text{I.13})$$

Une condition suffisante d'ergodisme relativement à la moyenne est que :

$$\lim_{T \rightarrow \infty} \left(1 - \frac{\tau}{T}\right) C_X(\tau) = 0 \quad (\text{I.14})$$

I.2.4.2.2. Ergodicité relativement à la fonction de corrélation

La fonction de corrélation $\phi_X(\tau, t_0)$ d'une F.A stationnaire $X(t)$ (éventuellement complexe) se définit par [2] :

$$\phi_X(\tau, t_0) = \frac{1}{T} \int_0^{t_0 + T} [X(t + \tau) - \eta_T(t_0 + \tau)] [X(t) - \eta_T(t_0)]^* dt \quad (\text{I.15})$$

On dit de la fonction aléatoire qu'elle est ergodique relativement à la fonction de corrélation lorsque :

- Elle est ergodique relativement à la moyenne
- La limite de la fonction de corrélation pour T tendant vers l'infini existe, est certaine, ne dépend plus de t_0 et est égale à la fonction de covariance, soit :

$$\lim_{T \rightarrow \infty} \phi_X(\tau, t_0) = C_X(\tau) \quad (\text{I.16})$$

I.2.5. Densité spectrale de puissance

La fonction d'autocorrelation est une description dans le domaine temporel d'un processus stochastique stationnaire du second ordre. L'équivalent dans le domaine fréquentiel est la densité spectrale de puissance ou spectre. La densité spectrale de puissance d'un processus continu $X(t)$ stationnaire au sens large est la transformée de Fourier de sa fonction d'autocorrelation. Le spectre et la fonction d'autocorrelation sont reliés comme suit :

$$S_{XX}(\omega) = \int_{-\infty}^{+\infty} \phi_X(\tau) \exp(-j\omega\tau) d\tau \quad (\text{I.17})$$

Si le signal est plus ergodique, la fonction d'autocorrelation peut être obtenue à partir d'une réalisation de la variable aléatoire.

I.2.6. Modèles stochastiques

Un signal est souvent modélisé comme la réponse d'un système à un autre signal de caractéristiques plus simples. L'idée essentielle est qu'une suite temporelle dont les observations sont très corrélées peut être générée à la sortie d'un filtre linéaire qui a pour entrée une suite de nombres statistiquement indépendants.

L'entrée du filtre est souvent supposée être un bruit blanc gaussien de moyenne nulle avec une variance constante [5]. (Fig.I.1)

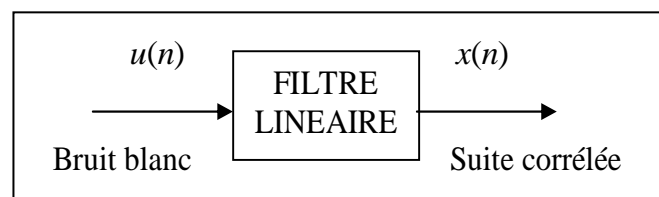


Fig.I.1 : Génération d'un signal aléatoire corrélé.

Le filtre linéaire peut avoir différentes structures qui correspondent à différents modèles pour le signal de sortie. Il existe trois modèles linéaires stochastiques classiques :

- Modèle à moyenne ajustée (MA),
- Modèle autorégressif (AR),
- Modèle autorégressif à moyenne ajustée (ARMA.)

I.2.7. Processus stochastiques remarquables

Parmi les processus stochastiques rencontrés lors de la modélisation des processus physiques sont :

I.2.7.1. Processus Gaussien

Les processus gaussiens sont des processus très importants du fait qu'on les rencontre très souvent en pratique et du fait que de nombreux processus physiques sont approximativement gaussiens [2]. On dit qu'une variable aléatoire réelle X a une distribution normale ou gaussienne si sa loi de probabilité a une densité $P_X(x)$ qui suit une loi normale ou de Gauss:

$$P_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma_X^2}\right] \quad (\text{I.18})$$

Avec m et σ_X représentent la moyenne et l'écart type de la variable aléatoire X respectivement.

Les lois gaussiennes préservent leur caractère gaussien dans toute opération linéaire, convolution, filtrage, dérivation, intégration. Par suite, si l'entrée d'un système linéaire est gaussienne, il en est de même pour la sortie.

I.2.7.2. Processus Markovien

Cette classe de processus stochastiques est importante du fait qu'elle généralise à l'univers stochastique une propriété fondamentale des équations différentielles ordinaires. En effet, la solution $x(t_2) = g(t_2, x(t_1), t_1)$ de l'équation différentielle $\dot{x}(t) = f(x(t))$ est fonction de $x(t_1)$ et ne dépend pas de $x(\tau)$, $\tau < t_1$. Toute l'information sur le passé est concentrée dans le dernier état observé. Un processus de Markov est donc un processus stochastique dont le passé n'a pas d'influence sur le futur si le présent est connu [5].

I.3. Notions de la théorie d'estimation

L'estimation est une des activités principales du statisticien. Les distributions de probabilité ne sont en général connues que par le biais des échantillons qu'elles génèrent, et l'estimation est l'art d'extraire d'un échantillon de l'information utile sur la distribution de probabilité qui l'a engendré. Le terme "Estimation" recouvre plusieurs réalités différentes mais étroitement reliées entre elles, et que nous passons ici brièvement en revue.

I.3.1. Estimation d'une distribution

Le rêve ultime de la statistique est de pouvoir identifier sans ambiguïté la distribution derrière un phénomène aléatoire. Mais cette distribution n'est connue que par le biais d'un échantillon fini et aléatoire, et ne peut donc jamais être identifiée avec certitude. Cependant, il est possible de formuler des conjectures sur la nature de la distribution qui a engendré l'échantillon. Cette question est relativement simple pour les distributions de probabilité discrètes, mais est beaucoup plus difficile pour les distributions de probabilité continues.

Proposer une distribution de probabilité complètement définie comme "distribution candidate pour l'échantillon considéré" est appelé estimation de distribution de probabilité, et plus particulièrement estimation de densité de probabilité dans le cas continu.

I.3.2. Estimation des paramètres d'une distribution

L'estimation des paramètres d'une distribution peut se faire de deux façons :

- On peut renoncer à complètement caractériser la distribution, et se satisfaire de ne caractériser que certains aspects particuliers de cette distribution, comme sa moyenne, sa variance, son mode, ou tout autre quantité définie sur la distribution.
- On peut également faire l'hypothèse que la distribution appartient à une famille de distributions décrite par une expression mathématique contenant un ou plusieurs paramètres numériques. Identifier la distribution revient alors à estimer les valeurs numériques de quelques paramètres.

I.3.3. Estimation des paramètres d'un modèle

Un modèle, qu'il soit prédictif ou descriptif, peut être perçu comme une représentation particulière d'une distribution de probabilité. Un modèle paramétrique contient des paramètres dont les valeurs sont calculées à partir de l'échantillon. Ces paramètres sont donc des variables aléatoires qui ont leurs distributions propres. Identifier ces distributions est une tâche essentielle en modélisation de données, car leur analyse permettra de juger de la fiabilité du modèle : un modèle dont les paramètres ont des distributions larges sont peu fiables.

Nous passons maintenant brièvement en revue certains aspects de l'estimation de paramètres.

I.3.4. Estimateurs et estimations

Soit θ un paramètre de la distribution, dont la vraie valeur (inconnue) est θ_0 . Un estimateur est une fonction des observations de l'échantillon, et dont la valeur sera utilisée en lieu et place de la valeur vraie θ_0 du paramètre θ . La valeur prise par un estimateur sur un échantillon donné est appelée une estimation de θ_0 .

On attend d'une estimation qu'elle ait une valeur proche de la vraie valeur du paramètre. Mais l'échantillon étant aléatoire, l'estimateur $\hat{\theta}$ est une variable aléatoire, dont l'estimation n'est qu'une réalisation. Il ne peut donc jamais être dit avec certitude qu'une estimation est proche de la vraie valeur θ_0 .

Il apparaît ainsi que la théorie de l'estimation ne portera pas sur les estimations, mais sur les estimateurs considérés comme variables aléatoires, c'est à dire, sur leurs distributions, ou sur certains aspects de leurs distributions (essentiellement moyenne et variance).

I.3.5. Propriétés d'un bon estimateur

Rien dans la structure de l'équation définissant une statistique ne permet de dire à priori qu'elle est un estimateur de quoi que ce soit. On peut dire qu'il n'existe pas de définition d'un estimateur. Mais une statistique particulière peut être utilisée pour estimer un paramètre si elle possède certaines bonnes propriétés.

I.3.5.1. Estimateur convergent

Une idée de la statistique est que les grands échantillons donnent une image raisonnablement fidèle de la distribution elle-même. En d'autres termes on espère que la distribution empirique des grands échantillons est proche de la distribution réelle. On est donc en droit d'attendre d'un bon estimateur $\hat{\theta}$ d'un paramètre θ qu'il produise des estimations qui soient de plus en plus proches de la vraie valeur θ_0 quand on considère des échantillons de plus en plus grands.

Si nous tenons pour acquis que la variance de l'estimateur tend vers zéro quand la taille de l'échantillon tend vers l'infini, la notion d'estimateur convergent veut alors simplement dire que la moyenne de la distribution de l'estimateur tend vers θ_0 quand la taille de l'échantillon tend vers l'infini.

I.3.5.2. Biais

La convergence est une propriété asymptotique: définir la convergence demande de considérer des échantillons de taille arbitrairement grande. Dans la réalité, la taille des échantillons est limitée pour des raisons de délais ou de budget. Il est donc naturel de se demander quelle qualité est attendue d'un estimateur limité à des échantillons de taille donnée t . On espère alors certainement que la région centrale de la distribution de cet estimateur soit proche de la valeur vraie θ_0 du paramètre (Fig. I.2). Une façon d'exprimer cette idée est de considérer les estimateurs dont la moyenne de la distribution (l'espérance) soit égale à θ_0 pour toute valeur de t . Un tel estimateur est dit non biaisé, ou sans biais, si cette propriété est vérifiée :

$$E \left[\hat{\theta}(t) \right] = \theta_0 \quad (\text{I.19})$$

Avec E désigne l'espérance mathématique

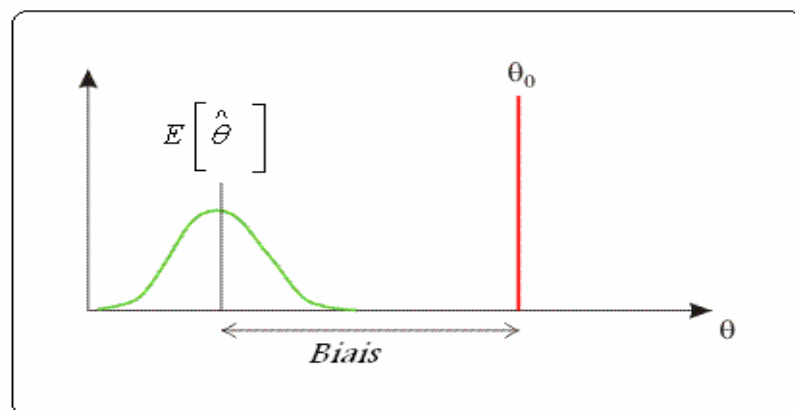


Fig.I.2: Estimateur biaisé de faible variance

I.3.5.3. Efficacité

La définition de l'efficacité relative η de deux estimateurs sans biais $\hat{\theta}_1$ et $\hat{\theta}_2$ d'un même paramètre θ pour une taille d'échantillon donnée n est le rapport de leurs variances :

$$\eta = \frac{\text{Var} \left\{ \hat{\theta}_1(t) \right\}}{\text{Var} \left\{ \hat{\theta}_2(t) \right\}} \quad (\text{I.20})$$

I.3.5.4. Ecart quadratique moyen d'un estimateur

Un bon estimateur est proche de θ_0 en moyenne. Cette proximité est mesurée par la moyenne du carré de l'erreur d'estimation, que l'on appelle: Ecart Quadratique Moyen (EQM) de l'estimateur $\hat{\theta}$:

$$EQM = E\left[(\hat{\theta} - \theta_0)^2\right] \quad (I.21)$$

Nous pouvons facilement montrer que :

$$EQM = Var(\hat{\theta}) + (Biais(\hat{\theta}))^2 \quad (I.22)$$

Cette expression explique pourquoi un estimateur sans biais n'est pas toujours systématiquement préférable à un estimateur biaisé, mais de faible variance. En pratique, cependant, on considère surtout les estimateurs sans biais parce qu'ils sont plus faciles à identifier que les estimateurs d'EQM minimal, et qu'ils ont de "bonnes" propriétés mathématiques.

L'identification d'estimateurs d'erreur quadratique moyenne minimale est cependant malaisée, et la plupart des estimateurs communément utilisés sont simplement des estimateurs sans biais.

Etant donnés deux estimateurs:

- $\hat{\theta}_1$ non biaisé, mais de variance importante
- $\hat{\theta}_2$ biaisé, mais de faible variance.

$\hat{\theta}_2$ peut s'avérer en pratique être un meilleur estimateur que $\hat{\theta}_1$ (Fig. I.3).

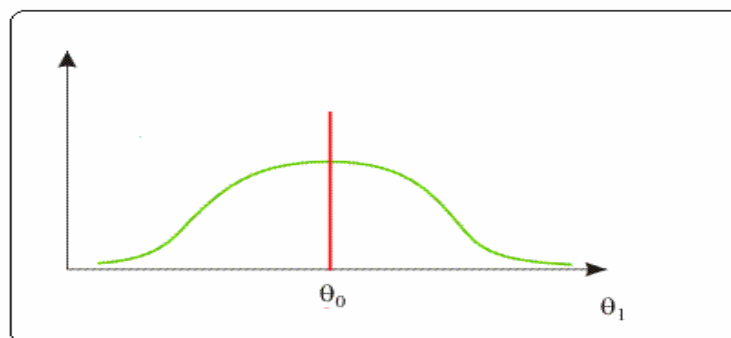


Fig.I.3: Estimateur non biaisé de forte variance

I.4. Notions de bruits

On appelle bruit, tout phénomène perturbateur gênant la perception ou l'interprétation d'un signal par analogie avec les nuisances acoustiques. Le bruit est inhérent à l'environnement naturel et à l'équipement électronique caractérisant aussi le bruit d'origines externe et interne.

I.4.1. Bruit blanc

Un bruit blanc est un processus stochastique utilisé afin de modéliser les bruits intervenant dans toute modélisation de systèmes dynamiques. Une fonction aléatoire faiblement stationnaire $X(t)$ est un bruit blanc si elle est incorrélée. Cela montre qu'un bruit blanc gaussien stationnaire à une densité de puissance identique à toutes les fréquences qui justifie la dénomination de bruit blanc par analogie avec la lumière blanche. Toutefois, si l'on calcule la puissance totale d'un bruit blanc, nous obtenons une valeur infinie qui montre que ce type de processus n'existe pas dans le monde physique [6].

I.4.2. Séquence Binaire Pseudo Aléatoire (SBPA)

Les séquences binaires pseudo aléatoires sont des successions d'impulsions rectangulaires modulées en largeur, qui approchent un bruit blanc discret et donc qui ont un contenu riche en fréquences [2]. Elles s'appellent pseudo aléatoire car elles sont caractérisées par une « longueur de séquence » à l'intérieur de laquelle les variations de la largeur des impulsions varient aléatoirement, mais, sur un grand horizon de temps, elles sont périodiques, la période étant définie par la longueur de la séquence.

En pratique, on utilise le bruit pseudo blanc qui possède une densité spectrale de puissance constante sur une bande fréquentielle.

I.4.3. Bruit coloré

Un bruit coloré est un bruit blanc filtré, ainsi la fonction d'autocorrélation d'un tel bruit ne sera pas une impulsion de Dirac, mais plutôt une courbe étroite [7]. Un bruit coloré de basse fréquence est parfois appelé bruit rose car il ne conserve que les grandes longueurs d'onde, ce qui dans le spectre du visible correspond aux teintes rouges.

I.5. Notions sur le signal de parole

La parole apparaît physiquement comme une variation de la pression de l'air causée et émise par le système articulatoire. La phonétique acoustique étudie ce signal en le transformant dans un premier temps en signal électrique grâce au transducteur approprié : le microphone (lui-même associé à un préamplificateur). De nos jours, le signal électrique résultant est le plus souvent numérisé. Il peut alors être soumis à un ensemble de traitements statistiques qui visent à en mettre en évidence les traits acoustiques : sa fréquence fondamentale, son énergie, et son spectre. Chaque trait acoustique est lui-même intimement lié à une grandeur perceptuelle : pitch, intensité et timbre.

I.5.1. Description du signal vocal

La parole est un signal réel, continu, d'énergie finie, non stationnaire. Sa structure est complexe et variable dans le temps : tantôt périodique (plus exactement pseudopériodique) pour les sons voisés, tantôt aléatoire pour les sons fricatifs, tantôt impulsionnelle dans les phases explosives des sons occlusifs [8]. Cette structure reflète l'organisation temporelle des gestes de production et sur l'onde sonore apparaissent quelques caractéristiques de la source et du conduit tel que la fréquence fondamentale F_0 et la fréquence des formants F_i .

I.5.2. Détection de la fréquence fondamentale

Les sons de la parole dits voisés sont produits avec vibration des cordes vocales ou vibration laryngienne. La mesure de la fréquence fondamentale F_0 par l'analyse de Fourier du signal de parole apparaît comme estimation de la fréquence de vibration laryngée.

Les variations des valeurs de F_0 au cours du temps constituent la courbe mélodique de la phrase. Cette mesure peut se faire à partir du signal de parole dans le domaine temporel, par exemple après filtrage du signal, ou dans le domaine spectral, à partir de la fréquence fondamentale du spectre d'un son voisé [9].

I.5.3. Classification des sons

Les signaux de parole peuvent être classés en deux catégories : signaux voisés caractérisés par des segments quasi-périodiques et d'énergie élevée tels que les voyelles, et signaux non voisés qui présentent généralement des segments de basse énergie tels que les consonnes. Les sons voisés résultent, d'une vibration périodique des cordes vocales. Le larynx est d'abord complètement fermé, ce qui accroît la pression en amont des cordes vocales et force ces dernières à s'ouvrir, ce qui fait tomber la pression en permettant aux cordes vocales

de se refermer. Des impulsions périodiques de pression sont ainsi appliquées au conduit vocal. Le taux auquel les cordes vocales s'ouvrent et se ferment s'appelle la fréquence fondamentale (dénotée par F_0) qui correspond physiquement au pitch perçu, sa valeur change avec la taille du conduit vocal. La fréquence fondamentale peut varier de 80 à 200 Hz pour une voix masculine, de 150 à 450 Hz pour une voix féminine, et de 200 à 600 Hz pour une voix d'enfant.

La Fig.I.4 présente des segments de sons voisés et non voisés avec leurs spectres correspondants.

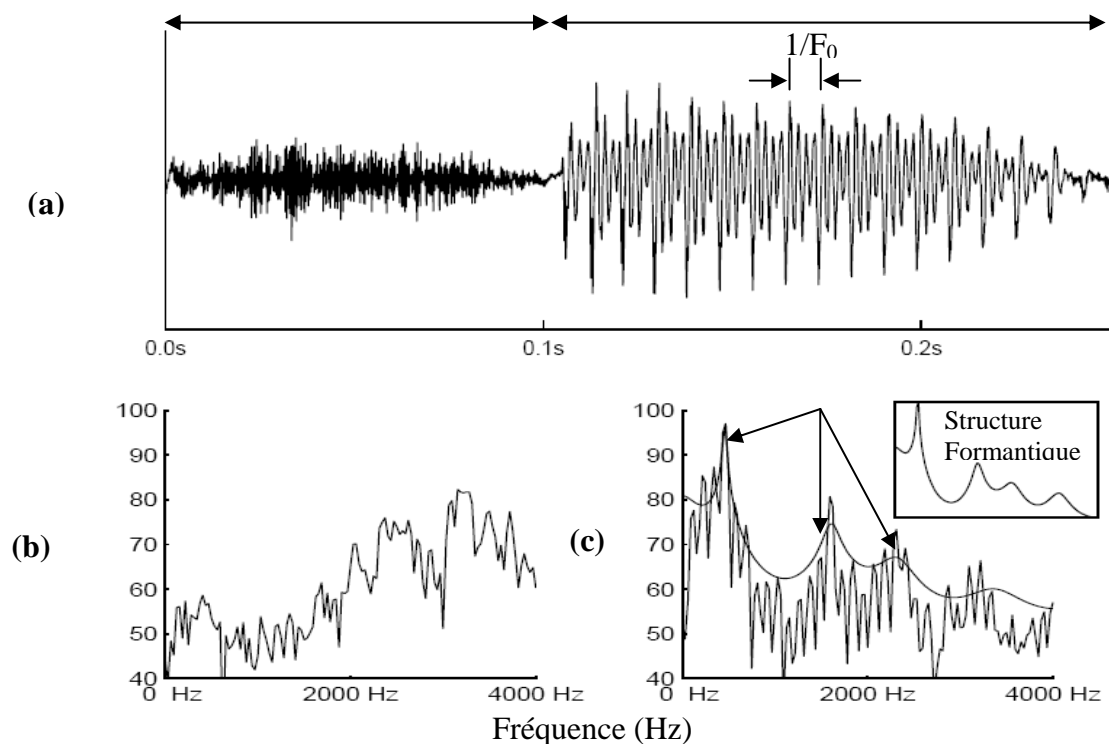


Fig. I.4 : (a) - Deux segments de voix, l'un non voisé et l'autre voisé.
 (b) - Spectre de puissance pour un segment de 32 ms non voisé.
 (c) - Spectre de puissance et sa structure formantique correspondante pour un segment de 32 ms voisé.

Les sons voisés présentent dans le domaine temporel un signal quasi-périodique tandis qu'ils présentent une structure harmonique dans le domaine fréquentiel, l'espacement entre les harmoniques est égal à la fréquence fondamentale F_0 (pitch). L'enveloppe spectrale possède une structure formantique, elle est caractérisée par un nombre de pics (dénotés par F_i), chacun d'eux est appelé Formant, Les trois premiers formants sont essentiels pour caractériser le spectre vocal. Les formants d'ordre supérieur ont une influence plus limitée.

La structure formantique est attribuée au conduit vocal qui agit comme un filtre ayant comme résonances les pôles de la fonction de transfert ou formants, et comme anti-résonances les zéros de la fonction de transfert.

Les sons non voisés sont le résultat du passage du flux d'air par une étroite constriction au niveau du conduit vocal causant des turbulences, c'est-à-dire du bruit. Contrairement aux sons voisés, les sons non voisés ne présentent pas de structure périodique. Ils peuvent être modélisés par un bruit blanc filtré par le conduit vocal.

1.5.4. Modèle de production de la parole

L'analyse de la parole est une étape indispensable à toute application de synthèse, de codage ou de reconnaissance. Elle repose en général sur un modèle. Il existe de nombreux modèles de parole. On distingue les modèles articulatoires, les modèles de production, et les modèles phénoménologiques [10]. Dans le processus de codage, on s'intéresse au modèle de production. On y décrit la parole comme le signal produit par un assemblage de générateurs et de filtres numériques (modèle source filtre). Les paramètres de ces modèles sont ceux des générateurs et filtres qui les constituent. Le modèle Autorégressif (AR) en est l'exemple le plus utilisé.

Fant a proposé en 1960 un modèle de production qui spécifie qu'un signal voisé peut être modélisé par le passage d'un train d'impulsions $u(n)$ à travers un filtre numérique récursif de type tous pôles. On montre que cette modélisation reste valable dans le cas des sons non voisés, à condition que $u(n)$ soit cette fois un bruit blanc. Le modèle final est illustré à la figure I.5. Il est souvent appelé modèle auto régressif (AR), parce qu'il correspond dans le domaine temporel à une régression linéaire de la forme :

$$y(n) = G \cdot u(n) + \sum_{i=1}^p -a_i y(n-i) \quad (\text{I.23})$$

Où $u(n)$, G et p sont respectivement le signal d'excitation, le gain et l'ordre du système.

Chaque échantillon est obtenu en ajoutant un terme d'excitation à une prédiction obtenue par combinaison linéaire des p échantillons précédents. Les coefficients du filtre $\{a_i\}$ sont appelés coefficients de prédiction et le modèle AR est souvent appelé modèle de prédiction linéaire LP (Linear Prediction).

Les paramètres du modèle AR sont : la période du train d'impulsions (sons voisés uniquement), la décision son voisé/non voisé, le gain G et les coefficients du filtre $1/A(z)$, appelé filtre de synthèse.

Le problème de l'estimation d'un modèle AR, souvent appelée analyse LP revient à déterminer les coefficients d'un filtre tous pôles dont on connaît le signal de sortie, mais pas celui de l'entrée. Il est par conséquent nécessaire d'adopter un critère, afin de faire un choix parmi l'ensemble infini de solutions possibles. Le critère généralement utilisé est celui de la minimisation de l'énergie de l'erreur de prédiction.

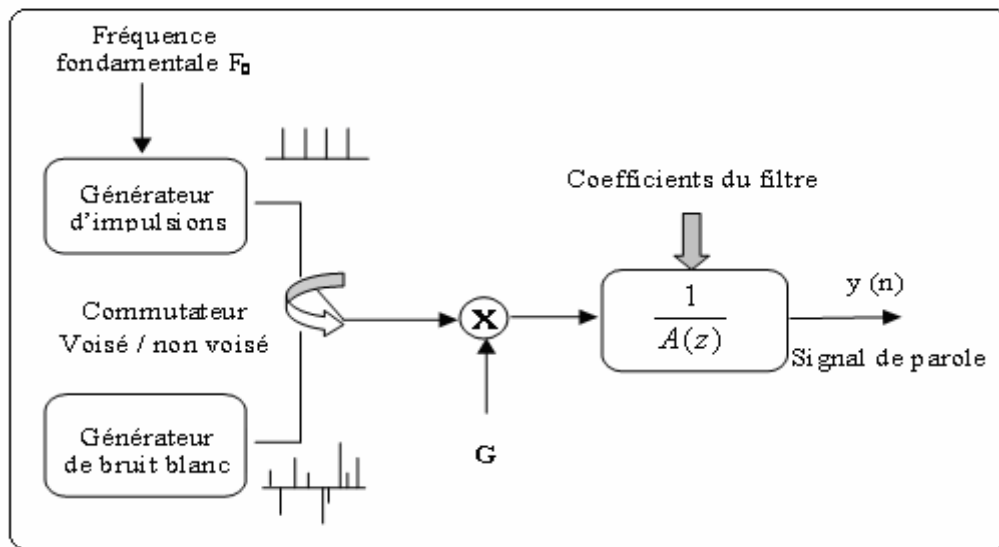


Fig. I.5 : Modèle simplifié de production de la parole.

Par analogie entre le modèle physique et le modèle mathématique, on peut donner les relations d'équivalences illustrées sur la figure I.6.

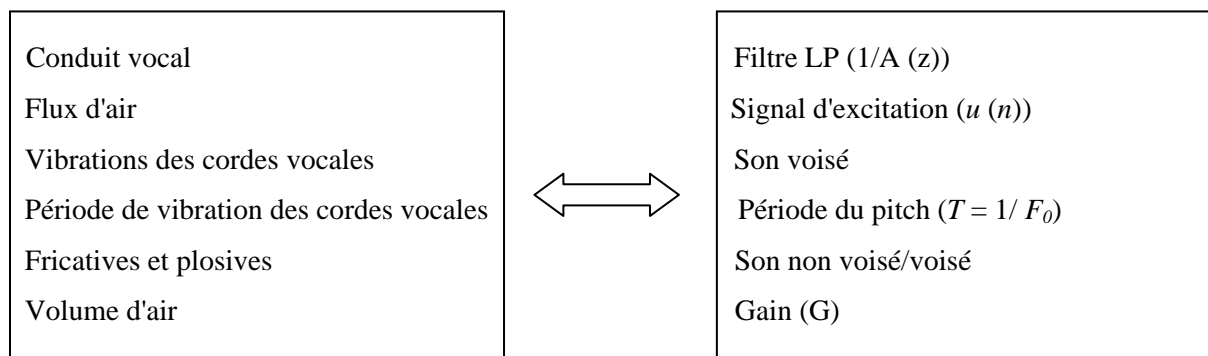


Fig. I.6 : L'analogie entre le modèle physique et le modèle acoustique.

I.6. Conclusion

En conclusion, les fonctions aléatoires jouissent de propriétés remarquables qui en simplifient souvent l'étude et la caractérisation. La propriété la plus usuelle est la stationnarité puisqu'elle consiste à reconnaître une quasi-invariance du comportement statistique de la variable aléatoire quelle que soit la valeur attribuée au paramètre déterministe de l'origine du temps.

Autrement dit, un signal aléatoire est stationnaire si les moments et la densité de probabilité sont indépendants du temps. Une étude plus attentive des fonctions aléatoires montre cependant que la stationnarité n'est pas toujours acquise et qu'il faut parfois distinguer la stationnarité au sens strict de celle qui regarde l'ordre des moments de la fonction. Une autre propriété, étroitement associée à la précédente, s'appelle l'ergodisme. L'ergodisme apporte d'intéressantes simplifications dans la mesure où l'on peut exploiter la convergence existant entre les moments de la variable aléatoire et les valeurs moyennes de la fonction calculée par intégration sur le paramètre déterministe. Cette propriété n'est pas très aisée à justifier sur le plan théorique, aussi faut-il souvent user de considérations intuitives pour la découvrir. Le domaine d'application le plus répandu des fonctions aléatoires est certainement le traitement du signal.

Chapitre **II** **Modélisation et Identification des Processus Physiques**

II.1. Introduction

Ce chapitre est certainement le plus philosophique et le plus stratégique. En effet, la modélisation et l'identification ne se réduisent pas à utiliser un logiciel de traitement de données sur un signal venant d'un processus observé passivement. Représenter le monde réel, bruité, non linéaire, non stationnaire par une équation mathématique forcément limitée est un acte précieux. Il est donc nécessaire de bien saisir les limites de la démarche et la valeur relative de la modélisation. Ajoutons que la grande et bénéfique accessibilité des moyens de calculs modernes augmente le risque pour l'utilisateur non averti, de ne voir que l'aspect algorithmique du problème.

La modélisation consiste à trouver une représentation du comportement d'un système c'est à dire la relation entre l'entrée et la sortie de ce système, et voir les différentes approches pour l'établissement d'un modèle. L'identification consiste aussi à déterminer les caractéristiques dynamiques de ce système, c'est-à-dire, à ajuster les paramètres inconnus du modèle de manière à ce que celui-ci décrive au mieux le fonctionnement du système.

II.2. Modélisation des processus physiques

Le but de la modélisation a toujours été de tenter de plaquer sur une réalité physique mesurable d'un monde extérieur sensible, une représentation rationnelle, en fait logicomathématique. L'objectif de la modélisation n'est donc pas nouveau et s'il doit être rattaché à celui des sciences expérimentales, la modélisation moderne apporte cependant une contribution spécifique. Un système dynamique peut être représenté par un modèle mathématique qui lie les variables mesurées d'entrées et de sorties. Les modèles dynamiques sont devisés en deux types [21], [22] :

- Modèle non paramétrique (exemple : réponse fréquentielle, réponse indicielle).
- Modèle paramétrique (exemple : fonction de transfert, équation différentielle et équation aux différences).

II.2.1. Différentes approches d'établissement d'un modèle

Fondamentalement, au moins deux approches sont possibles pour établir un modèle [23] :

II.2.1.1. Modèle de connaissance

Les modèles de connaissance (ou phénoménologiques) sont familiers à tous ceux qui ont suivi des cours de physique ou de chimie, donc les paramètres de modèle de connaissance ont un sens physique (longueur, résistance électrique,...).

Prenant en compte de manière précise la complexité du problème physique, s'intéressant à des objectifs ambitieux et ne négligeant aucun élément d'information récupérable, cette approche vise donc la compréhension finie de phénomènes, ce qui se traduit par des relations en générale, complexes et coûteuses en temps de calcul. Les modèles associés sont généralement difficiles à simuler et donc rarement utilisables tels quels pour calculer par exemple une commande, ils sont par contre bien adaptés à une simulation détaillée en vue d'une prédiction de comportement à long terme.

II.2.1.2. Modèle de représentation

Ces modèles n'ont aucun pouvoir explicatif de la structure physique de l'objet, leur structure n'est qu'une relation mathématique qui va relier localement les mesures des différentes variables du processus. Ces modèles de représentation sont de type boîte noire, et les paramètres donc n'ont aucun sens physique mais sont suffisants dans les problèmes de traitement de signal.

Les modèles de représentation sont en générale assez simples à simuler et mieux adaptés à la détermination d'une commande, ils sont d'utilisations très fréquentes.

II.2.2. Modélisation Auto-Régressive à Moyenne Ajustée (ARMA)

Un modèle autorégressive à moyenne ajusté d'ordre (n, m), noté ARMA (n, m) est définie par l'équation aux différences suivante [11], [25] :

$$\sum_{i=0}^n a_i y(t-i) = \sum_{i=0}^m b_i u(t-i) \quad (\text{II.1})$$

Avec a_i : représente les paramètres de la partie autorégressive AR. ($a_0 = 1$)

b_i : représente les paramètres de la partie moyenne ajustée MA

n, m: l'ordre du modèle ARMA.

D'où l'expression (II.1) devient :

$$y(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=0}^m b_i u(t-i) \quad (\text{II.2})$$

En prenant la transformée en Z des deux membres de l'équation (II.1), nous aboutissons à la fonction de transfert du modèle ARMA :

$$H(z) = \frac{Y(z)}{U(z)} = \frac{\sum_{i=0}^m b_i z^{-i}}{1 + \sum_{i=1}^n a_i z^{-i}} = \frac{B(z)}{A(z)} \quad (\text{II.3})$$

Où $H(z)$: Fonction de transfert du modèle

$A(z)$: Transformée en Z de la partie AR

$B(z)$: Transformée en Z de la partie MA.

Le modèle ARMA peut être interprété comme un filtre de fonction de transfert $H(z)$, ayant des pôles et des zéros, excité par une entrée $U(z)$ et délivrant à sa sortie un signal $Y(z)$. Les polynômes $A(z)$ et $B(z)$ sont caractérisés par la position de leurs zéros dans le plan complexe.

II.2.2.1. Modélisation Auto-Régressive (AR)

Dans le cas où les b_i sont nuls pour $1 \leq i \leq n$, le modèle dont l'équation (II.2) se réduit à :

$$y(t) = -\sum_{i=1}^n a_i y(t-i) + b_0 u(t) \quad (\text{II.4})$$

Ainsi le polynôme $B(z)$ se réduit à une constante $B(z) = b_0$ et la fonction de transfert $H(z)$ ne contient que des pôles, c'est pour cette raison que ce modèle est aussi appelé modèle tout pôles. La fonction de transfert peut s'écrire sous la forme suivante:

$$H(z) = \frac{b_0}{1 + \sum_{i=1}^n a_i z^{-i}} \quad (\text{II.5})$$

II.2.2.2. Modélisation à Moyenne Ajustée (MA)

Dans le cas où les a_i sont nuls, pour $1 \leq i \leq m$, le modèle dont l'Eq.II.2 se réduit à :

$$y(t) = \sum_{i=0}^m b_i u(t-i) \quad (\text{II.6})$$

Dans le domaine spectral, le modèle sera alors caractérisé par la position de ses zéros dans le plan complexe, ce modèle est appelé tout zéros.

II.2.3. Principales étapes de la modélisation

On distingue six étapes essentielles de la modélisation comme se suit [22] :

II.2.3.1. Système

Un système (où processus) sera pour nous une partie de l'univers qui nous entoure, et que nous avons décidé, plus ou moins arbitrairement d'appréhender comme un tout avec lequel nous interagissons. Nous observons certaines grandeurs caractéristiques du système et le résultat de ces observations forme le vecteur des sorties y et d'autres grandeurs caractéristiques du système forment un vecteur qui peut contenir des grandeurs inaccessibles à la mesure. Nous agissons sur le système par des grandeurs qui peuvent être connues et maîtrisables, ou non maîtrisables et plus ou moins inconnues, ce sont les perturbations ou bruit.

II.2.3.2. Modèle

Le modèle est une règle permettant de calculer, à partir de grandeurs connues ou mesurées, d'autres grandeurs dont nous espérons qu'elles ressembleront aux grandeurs du système qui nous intéresse. Fréquemment, le modèle calcule, à partir de l'entrée u du système une sortie \hat{y} qui ressemble le plus possible à y . Puisque le modèle et le système ont alors la même entrée, on parlera de modèle parallèle. Par contre dans le cas où le modèle calcule à partir de la sortie y un vecteur \hat{u} dont on souhaite qu'il ressemble le plus possible aux entrées u du système, on parlera du modèle série ou inverse.

II.2.3.3. Critère

Pour fixer les idées, nous supposons le modèle de type parallèle, c'est-à-dire, soumis aux mêmes entrées et aux mêmes conditions initiales que le système, on appelle alors erreur de sortie, la différence entre la sortie du système et celle du modèle. L'erreur de sortie peut s'écrire sous la forme:

$$\varepsilon_r = y(t) - \hat{y}(t) \quad (\text{II.7})$$

Le plus souvent, on souhaite que cette erreur de sortie soit aussi proche que possible de zéro, l'échelle de valeurs qui sera utilisée pour effectuer la comparaison prendra la forme d'une fonction scalaire $J(\theta)$ appelée « critère ». Le choix du critère doit traduire le but fixé à la modélisation quelque soit le critère choisi, il convient ensuite de l'optimiser.

II.2.3.4. Optimisation

L'algorithme d'optimisation reçoit toute les informations disponibles et les utilise pour minimiser $J(\theta)$ de façon à calculer $\hat{\theta}$.

II.2.3.5. Incertitude sur les paramètres

La valeur de $\hat{\theta}$ obtenue avec l'algorithme d'optimisation correspond au meilleur modèle possible choisi parmi un ensemble de modèles acceptables qui dépend de l'expérience réalisée pour le recueil des données, donc on pourra faire la planification d'expérience, de façon à essayer de recueillir l'information la plus pertinente possible au sens du critère utilisé.

II.2.3.6. Analyse critique des résultats obtenus

La phase d'analyse critique des résultats obtenus est indispensable, il faut donc soumettre le modèle à un ensemble d'épreuves destinées à l'invalider, si l'on parvient ainsi à détecter des erreurs, alors on a un mauvais choix du modèle.

II.2.4. Difficultés de la modélisation

Les principales difficultés de la modélisation sont :

II.2.4.1. Définition du problème de modélisation

La collecte et la mise en forme des données, la construction, la validation et l'interprétation d'un modèle sont des tâches longues et coûteuses. Le temps passé à définir clairement l'objectif de l'étude, les critères permettant de savoir si celui-ci a ou non été atteint, l'identification de la nature et du volume des données nécessaires à la construction d'un modèle adéquat, est toujours du temps gagné, jamais du temps perdu.

II.2.4.2. Qualité des données de modélisation

Celles-ci ont la fâcheuse habitude d'être rares, chères, mal formatées, incomplètes, non synchronisées, entachées d'erreurs et biaisées. Les données sont pourtant le carburant de la modélisation, et doivent recevoir toute l'attention requise pour atteindre le niveau de qualité requis par l'application, sous peine de rendre vains tous les efforts de modélisation.

II.2.4.3. Choix de la technique utilisée

Le développement considérable de la modélisation met à la disposition du praticien un grand nombre de techniques différentes permettant, théoriquement, d'atteindre un objectif donné. Mais chaque technique a ses avantages et ses inconvénients, et le choix de la technique la plus appropriée est une des conditions essentielles du succès de la modélisation. Malheureusement, en dehors de considérations d'ordre général, seulement une longue pratique permet d'éviter les choix malheureux.

II.2.4.4. Choix des variables disponibles

Pour des raisons fondamentales de statistique (et souvent ignorées des praticiens), il est indispensable de procéder à une sélection rigoureuse des variables parmi les variables disponibles dans le tableau de données. Si trop de variables sont incluses dans le modèle, celui-ci devient exagérément sensible à de petites modifications de l'échantillon, et le modèle obtenu est alors peu crédible (compromis biais - variance). Des objectifs différents, ou même des techniques différentes mais ayant le même objectif, conduiront à des ensembles optimaux de variables différentes.

II.2.4.5. Sélection de modèle adéquat

Tout modèle contient des paramètres (même les modèles dits: "non paramétriques"). Les valeurs de ces paramètres sont :

- Soit calculées par un algorithme d'apprentissage,
- Soit définies arbitrairement.

Dans les deux cas, le choix du nombre de paramètres du modèle est laissé à l'analyste. On se convainc facilement qu'une augmentation du nombre de paramètres du modèle augmente sa souplesse, et donc sa capacité à rendre compte des données d'apprentissage. Mais on découvre également qu'à partir d'un certain point, une augmentation du nombre de paramètres conduit à une dégradation des performances du modèle sur les données nouvelles.

Donc, le modèle une fois construit, il convient d'estimer ses performances réelles sur des données nouvelles (et non pas sur les données qui ont servi à le construire). Ceci se fait en soumettant plusieurs modèles candidats à une série de tests de validation. Le modèle finalement retenu est celui qui aura obtenu les meilleurs résultats lors de ces tests (et ce ne sera vraisemblablement pas celui qui aura obtenu les meilleurs résultats sur les données d'apprentissage).

II.3. Identification des paramètres d'un processus physique

L'identification c'est l'opération de détermination des caractéristiques dynamiques d'un procédé (système). On s'intéresse à l'identification des modèles dynamiques paramétriques échantillonnées qui sont les plus appropriés pour la conception et l'ajustement des systèmes numériques [11], [22].

II.3.1. Identification non paramétrique

Les méthodes non paramétriques se comportent de façon neutre vis-à-vis des données, se refusant à inclure trop d'a priori sur la véritable nature du signal. Elles consistent à déterminer ce qui aurait été fait dans le cas idéal d'un signal déterministe connu et à bâtir des estimateurs point par point de l'autocorrelation et du spectre.

II.3.2. Identification paramétrique

Les méthodes paramétriques consistent à ajuster un modèle aux données observées. Les paramètres des modèles, en nombre faible, caractériseront le signal ; on pourra ainsi injecter des connaissances a priori sur le processus physique qui a engendré le signal.

La procédure standard pour réaliser cet ajustement est l'identification paramétrique dont le schéma est rappelé en Fig.II.1.

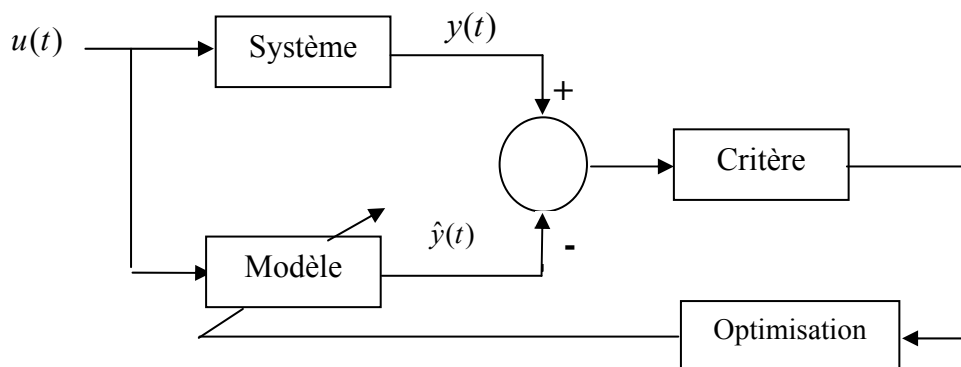


Fig. II.1: Schéma d'identification paramétrique

Cette procédure comporte un très grand nombre de variantes, correspondant au choix à priori, au niveau du modèle, du critère de mesure d'erreur de l'entrée choisie et de l'algorithme d'optimisation. Cette approche, commune à un grand nombre de domaines scientifiques, possède un caractère fondamental marqué et un large champ d'applications.

Les avantages qui peuvent en être attendus sont la souplesse de l'analyse, l'introduction naturelle d'informations à priori, la parcimonie de la représentation et des choix variés d'espèces de représentations paramétriques et offre d'autres possibilités telles que :

- Modélisation des bruits.
- Identification des modèles de perturbation.
- Détection et mesure des fréquences d'oscillations
- Analyse spectrale des signaux.

II.3.3. Etapes d'identification

Pour parvenir à un bon modèle, nous devons généralement suivre les quatre étapes suivantes [22], [25] :

II.3.3.1. Acquisition des entrées et sorties

Il s'agit essentiellement de choisir un signal d'excitation avec une densité spectrale homogène couvrant l'ensemble de la bande passante du procédé à identifier. En pratique, nous utilisons deux catégories de signaux de tests (entrées) :

- Les signaux déterministes tel que l'échelon, la sinusoïde etc., ces signaux sont décrit par une fonction de temps.
- Les signaux aléatoires, complètement décrits par leurs propriétés statistiques, un des signaux les plus utilisés pour l'identification est la séquence binaire pseudo aléatoire.

En pratique l'entrée du système n'est intéressante que si elle est :

- Centrée
- Riche en fréquence
- Déterministe si possible

II.3.3.2. Choix de la complexité du modèle

Le problème typique rencontré dans le cas du modèle paramétrique est le choix de l'ordre des polynômes (numérateurs, dénominateurs) de la fonction de transfert, ce choix de la complexité peut se faire par une procédure essais et erreur. On peut maintenant disposer des algorithmes qui estiment à partir des données, la complexité des modèles.

II.3.3.3. Estimation des paramètres du modèle

Une fois la complexité du modèle fixée, nous pouvons estimer les paramètres du modèle de façon à minimiser un critère de performance. La qualité de cette estimation dépendra de la méthode choisie, et de l'information contenue dans les données d'entrée/sortie.

II.3.3.4. Validation du modèle

Cette étape est certainement la plus importante lors d'une identification, elle consiste à accepter ou rejeter le modèle obtenu. L'ensemble du processus d'élaboration d'un modèle est itératif et le rejet d'un modèle qui ne répond pas à ses objectifs remet en cause l'ensemble des étapes déjà citées.

II.3.4. Algorithme d'identification récursif

Dans les algorithmes récursifs, les paramètres estimés sont optimisés progressivement en utilisant chaque fois une seule paire de données entrée-sortie. Les algorithmes récursifs permettent une identification en temps réel, utilisation d'une taille mémoire réduite et de suivre l'évolution des systèmes lentement variables dans le temps [22].

II.4. Conclusion

Nous avons présenté deux notions très importantes, la modélisation et l'identification. Un modèle de procédé est un ensemble de relations mathématiques permettant de prédire certains aspects de son fonctionnement. L'efficacité des modèles repose sur une analogie, entre le comportement des objets physiques et celui des êtres mathématiques.

La modélisation permet de représenter sous forme synthétique et cohérente un ensemble de connaissances, l'utilisation de matériel informatique impose de travailler avec des valeurs discrètes des différents signaux, nous nous intéressons au modèles qui décrivent l'évolution des procédés au cours du temps.

L'identification consiste à ajuster les paramètres inconnus du modèle de manière à ce que celui-ci décrit au mieux le fonctionnement du procédé.

Les algorithmes d'identification récursifs peuvent être utilisés pour l'estimation des paramètres d'un modèle à partir des données entrée/sortie et présentent plusieurs avantages par rapport aux algorithmes non récursifs.

Chapitre **III** **Etude et Développement des Algorithmes d'Identification**

III.1. Introduction

L'identification de systèmes est fortement liée à la construction de modèles décrivant de manière mathématique les processus physiques. La principale valeur ajoutée apportée par l'étape d'identification se traduit par une meilleure connaissance du système. Ainsi, lorsque l'étape préliminaire relative à l'identification du processus est réalisée, il est possible d'atteindre des objectifs classiques d'automatique comme l'observation de variables non accessibles directement à la mesure ou à la commande de ces processus. Au tout début, l'identification en tant que telle est fortement liée aux méthodes statistiques comme la méthode des Moindres Carrés récursifs ou du Maximum de Vraisemblance, très largement appliquées sur des processus industriels. Ces deux méthodes sont obtenues grâce à l'optimisation d'un critère dépendant des paramètres du système et elles sont très souvent utilisées sur des processus discrétisés. Il est clair que décrire un système en temps discret présente un avantage en termes de calculs informatiques.

Mais, beaucoup de systèmes réels sont modélisés en temps continu. De plus, les paramètres ont, la plupart du temps, une signification physique en temps continu. Les algorithmes peuvent être, soit implémentés tels quels sur des processus réels en considérant la faible fréquence d'échantillonnage de la carte d'acquisition par rapport à la bande passante du système étudié, soit discrétisés si nécessaire par les méthodes habituelles.

Dans le cadre de cette thèse, nous nous sommes intéressés plus particulièrement, à l'identification paramétrique par les algorithmes suivants:

- Algorithme du gradient stochastique (Least Mean Squares (LMS))
- Algorithme de Newton stochastique (Stochastic Newton (SN))
- Algorithme des moindres carrés récursif (Recursive Least Squares (RLS))
- Algorithme des moindres carrés étendu (Recursive Extended Least Squares (RELS))

III.2. Algorithme du gradient stochastique

L'algorithme LMS introduit par Widrow en 1970 est une approximation de l'algorithme du gradient déterministe. L'algorithme LMS est certainement l'algorithme adaptatif le plus populaire qui existe en raison de sa simplicité. Il pilote les paramètres du modèle ajustable de prédiction à partir des informations recueillies sur le système à chaque pas d'échantillonnage [11], [41]. L'objectif de l'algorithme du gradient stochastique est de minimiser un critère quadratique en terme d'erreur de prédiction.

III.2.1. Rappel sur la méthode du gradient

Soit une fonctionnel quadratique $J(x)$. Et on veut trouver l'optimal de $J(x)$ c'est-à-dire qu'on cherche \hat{x} de telle façon que $J(x)$ est minimal. La méthode du gradient est donnée par :

$$\hat{x}(t) = \hat{x}(t-1) + \mu \left[-\frac{\partial J(x)}{\partial(x)} \right]_{x=\hat{x}(t-1)} \quad (\text{III.1})$$

Avec μ représente le pas d'adaptation.

III.2.2. Identification du modèle ARMA

La modélisation auto-régressive à moyenne ajustée d'ordre (n, m) noté ARMA (n, m) est définie par l'équation aux différences suivante [51], [21] :

$$\sum_{i=0}^n a_i y(t-i) = \sum_{i=0}^m b_i u(t-i) + e(t) \quad (\text{III.2})$$

Avec $a_0 = 1$, $b_0 = 1$

$u(t)$: Entrée du système

$y(t)$: Sortie du système

$e(t)$: Bruit blanc.

En développant l'équation (III.2), on obtient :

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) + b_1 u(t-1) + \dots + b_m u(t-m) + e(t) \quad (\text{III.3})$$

Comme on peut représenter cette dernière équation sous une forme matricielle:

$$y(t) = \theta^T \varphi(t) + e(t) \quad (\text{III.4})$$

Avec T : Transposé

Posons pour la commodité des notations [11] :

$$\theta^T = [a_1 \ a_2 \ \dots \ a_n, b_1 \ b_2 \ \dots \ b_m]$$

$$\varphi^T(t) = [-y(t-1) \ \dots \ -y(t-n), u(t-1) \ \dots \ u(t-m)]$$

Avec

θ^T : Vecteur des paramètres à identifier

φ^T : Vecteur des données.

L'identification du vecteur paramètre θ^T au sens de l'erreur quadratique moyenne est :

$$J(\theta) = \frac{1}{2} E[\varepsilon^2(t)] \quad (\text{III.5})$$

Avec $E[.]$ désigne l'espérance mathématique.

Or l'erreur de prédiction est définie comme se suit :

$$\varepsilon(t) = [y(t) - \hat{\theta}^T \varphi(t)] \quad (\text{III.6})$$

La solution optimale aux sens de l'erreur quadratique moyenne est donnée par:

$$\left[\frac{\partial J(\theta)}{\partial \theta} \right]^T = -E\{\varphi(t)[y(t) - \theta^T \varphi(t)]\}_{\theta=\hat{\theta}(t-1)} \quad (\text{III.7})$$

D'où l'algorithme du gradient s'écrira :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mu \left[-\frac{\partial J(\theta)}{\partial \theta} \right]^T \Big|_{\theta=\hat{\theta}(t-1)} \quad (\text{III.8})$$

Avec μ représente le pas d'adaptation.

On remplaçant l'équation (III.7) dans l'équation (III.8), nous obtenons :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mu E\{\varphi(t)[y(t) - \hat{\theta}^T(t-1)\varphi(t)]\} \quad (\text{III.9})$$

Cet algorithme est porte le nom : *Algorithme du gradient déterministe*, en général les stochastiques de $((\varphi(t), y(t)))$ sont inconnus. Cependant, on remplace la valeur moyenne par sa valeur instantanée, et dans ce cas, l'algorithme est appelé par l'algorithme du gradient stochastique, et on écrit :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mu \varphi(t)[y(t) - \hat{\theta}^T(t-1)\varphi(t)] \quad (\text{III.10})$$

Pour que l'algorithme converge [1], [16], il faut que :

$$0 < \mu \leq \frac{2}{\lambda_{\max}} \quad \text{Avec } \lambda_{\max} : \text{ la plus grande valeur propre de la matrice } E\{\varphi(t)\varphi^T(t)\}.$$

III.2.3. Les avantages de la méthode du gradient

La méthode du gradient est simple et facile à mettre en œuvre. Elle permet de diminuer rapidement la valeur du critère lorsque le point initial est situé loin du point recherché. Cependant, l'utilisation d'informations du premier ordre uniquement rend la convergence très lente au voisinage de ce point.

L'algorithme du gradient présente plusieurs avantages tel que:

- L'algorithme LMS est très simple
- Les performances du LMS dépendent de trois facteurs :
 - le pas d'adaptation μ ,
 - les valeurs propres λ_i de la matrice $E\{\varphi(t)\varphi^T(t)\}$
 - la longueur du filtre L .
- Avec un pas d'adaptation petite, le LMS converge lentement mais l'erreur quadratique moyenne EQM excédentaire est petite
- Avec un pas d'adaptation grande, le LMS converge rapidement mais l'erreur quadratique moyenne excédentaire est grande
- Le temps de convergence de l'algorithme LMS dépend du conditionnement de la matrice $E\{\varphi(t)\varphi^T(t)\}$

III.3. Algorithme de Newton stochastique

L'algorithme de Newton stochastique est désigné pour l'identification des paramètres du processus physique. Il permet de minimiser l'erreur d'une manière plus efficace. Cet algorithme est utilisé dans plusieurs applications du traitement de signal.

III.3.1. Principe de la méthode

La méthode de Newton est parmi les plus utilisées pour la résolution des équations. Considérons une fonction $f(x)$, et on cherche x tel que $f(x) = 0$.

En commençant par une valeur initiale x_0 qui est assez proche de x en extrapolant la tangente en x_0 jusqu'à son intersection avec l'axe OX , on continue jusqu'à obtenir une valeur de x tel que $f(x) = 0$.

La figure suivante nous donne une description graphique de la méthode de Newton.

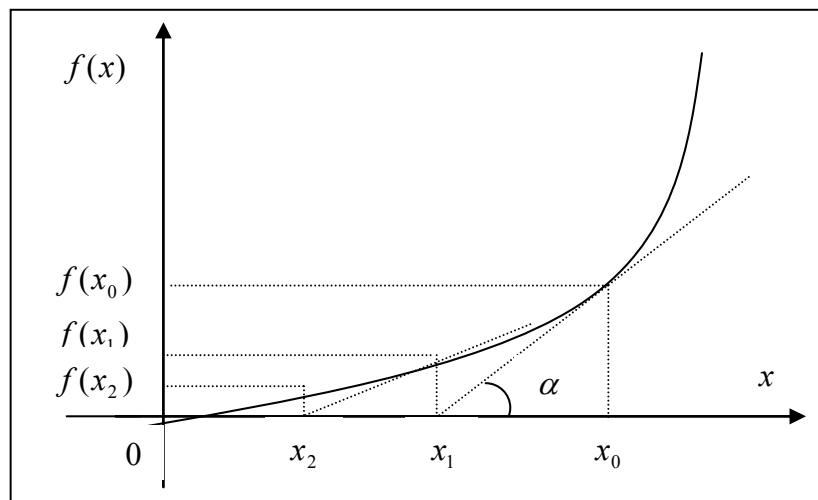


Fig. III.1 : Représentation graphique de la méthode de Newton

D'après la figure précédente, on écrit:

$$\operatorname{tg}(\alpha) = f'(x_0) = \frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{f(x_0)}{x_0 - x_1} \quad (\text{III.11})$$

A partir de cette dernière équation, on peut écrire les termes suivants:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}, \quad x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}, \quad x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}, \quad \dots \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

D'où la formule générale s'écrira:

$$x(t) = x(t-1) - \frac{f[x(t-1)]}{f'[x(t-1)]} \quad (\text{III.12})$$

III.2.2. Equations de mise en œuvre de la méthode de Newton

Soit $J(x)$ une fonction quadratique, et on veut la minimiser, en utilisant la méthode de Newton alors on écrit :

$$\text{Min } J(x) \Rightarrow \frac{\partial J(x)}{\partial x} \Big|_{x=\hat{x}} = 0 \quad (\text{III.13})$$

On pose :

$$f(x) = \frac{\partial J(x)}{\partial x} = J'(x) \quad (\text{III.14})$$

$$f'(x) = \frac{\partial^2 J(x)}{\partial x^2} = J''(x) \quad (\text{III.15})$$

En remplaçant dans l'équation III.12 on a :

$$x(t) = x(t-1) - \frac{J'[x(t-1)]}{J''[x(t-1)]} \quad (\text{III.16})$$

Cette dernière équation porte le nom : *Algorithme de Newton*, cas scalaire.

Comme il peut être représenté dans le cas vectoriel par l'équation (III.17):

$$\hat{x}(t) = \hat{x}(t-1) - [J''(\hat{x}(t-1))]^{-1} [J'(\hat{x}(t-1))]^T \quad (\text{III.17})$$

Où J^T représente un vecteur ligne.

III.3.3. Identification du modèle ARMA

Rappelant les équations (III.4, 5, 6), la minimisation du critère $J(\theta)$ consiste de trouver l'optimum, c'est-à-dire de trouver le point où la dérivée s'annule, d'où on écrit :

$$\left[\frac{\partial J(\theta)}{\partial \theta} \right]^T = -E\{\varphi(t)[y(t) - \theta^T \varphi(t)]\} \quad (\text{III.18})$$

$$\left[\frac{\partial^2 J(\theta)}{\partial \theta^2} \right] = E[\varphi(t)\varphi^T(t)] = R \quad (\text{III.19})$$

L'utilisation de la procédure d'approximation stochastique, permet d'écrire l'équation (III.17) sous la forme:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \gamma_1(t)R^{-1}(t)\varphi(t)(y(t) - \hat{\theta}^T(t-1)\varphi(t)) \quad (\text{III.20})$$

Avec $\gamma_1(t)$ gain scalaire, $\gamma_1(t) \rightarrow 0$ lorsque $t \rightarrow \infty$

Si R est inconnue, alors R peut être estimée par :

$$E(\varphi(t)\varphi^T(t)) = R \Rightarrow E(\varphi(t)\varphi^T(t) - R) = 0 \quad (\text{III.21})$$

En utilisant la procédure d'approximation. D'où :

$$R(t) = R(t-1) + \gamma_2(t)(\varphi(t)\varphi^T(t) - R(t-1)) \quad (\text{III.22})$$

Avec $\gamma_2(t)$ représente un gain scalaire

De l'équation (III.20) et (III.22), on peut tirer l'algorithme de Newton stochastique

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \gamma_1(t)R^{-1}(t)\varphi(t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \quad (\text{III.23})$$

$$R(t) = R(t-1) + \gamma_2(t)(\varphi(t)\varphi^T(t) - R(t-1)) \quad (\text{III.24})$$

Remarques

- Si $R(t) = I$ (matrice d'identité) alors on obtient l'algorithme du gradient stochastique
- Si $\gamma_1(t) = \gamma_2(t) = \frac{1}{t}$ alors on obtient l'algorithme RLS.

III.4. Algorithme des moindres carrés

La méthode des moindres carrés a été introduite par Karl Gauss en 1809. Elle a été à la base de toutes les méthodes d'identification et d'estimation des paramètres, cette méthode est basée sur la minimisation d'une fonction quadratique J définie comme [42] :

$$J_N(\theta) = \frac{1}{N} \sum_{t=1}^N [\varepsilon(t)]^2 \quad (\text{III.25})$$

Où

$\varepsilon(t)$: représente l'erreur de prédiction.

N : nombre d'échantillons

III.4.1. Identification du modèle ARMA

La modélisation auto-régressive à moyenne ajustée d'ordre (n, m) noté ARMA (n, m) est définie par l'équation aux différences. En utilisant la transformée en Z , l'équation (III.2) peut s'écrire :

$$\sum_{i=0}^n a_i Y(z) z^{-i} = \sum_{i=0}^m b_i U(z) z^{-i} + E(z) \quad (\text{III.26})$$

On pose:

$$A(z) = \sum_{i=0}^n a_i z^{-i}$$

$$B(z) = \sum_{i=0}^m b_i z^{-i}$$

Donc, l'équation (III.26) s'écrira sous cette forme :

$$A(z) Y(z) = B(z) U(z) + E(z) \quad (\text{III.27})$$

Ce qui donne comme erreur de prédiction :

$$E(z) = A(z) Y(z) - B(z) U(z) \quad (\text{III.28})$$

A partir de l'équation (III.28), on peut tirer le schéma suivant:

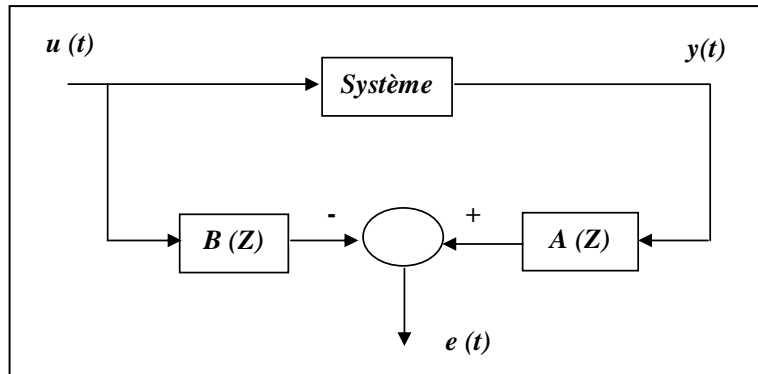


Fig.III.2 : Structure de réalisation d'un modèle ARMA

III.4.2. Solution optimale au sens des moindres carrés

On peut donc, penser de manière intuitive que si l'on augmente le nombre d'observations, le problème se ramènera à la résolution d'un système d'équations linéaires. Dans le cas d'un système de type ARMA, nous effectuons N mesures d'observations, nous pouvons écrire [42], [54] d'après l'équation (III.3), une forme matricielle :

$$y(t) = \theta^T \varphi(t) + e(t) \quad (\text{III.29})$$

Avec

$$\theta^T = [a_1 \quad a_2 \quad \dots \quad a_n, \quad b_1 \quad b_2 \quad \dots \quad b_m]$$

$$\varphi^T(t) = [-y(t-1) \quad \dots \quad -y(t-n), \quad u(t-1) \quad \dots \quad u(t-m) \quad]$$

On définit l'erreur de prédiction comme étant la différence entre la sortie du système et la sortie du modèle :

$$\varepsilon(t) = y(t) - \hat{y}(t) \quad (\text{III.30})$$

Sachant que :

$$\hat{y}(t) = \varphi^T(t) \hat{\theta}(t-1) \quad (\text{III.31})$$

Avec

$\hat{\theta}(t-1)$: Vecteur des paramètres estimés.

La méthode des moindres carrés est basée sur la détermination des meilleurs paramètres, c'est à dire ceux qui minimiseront un certain critère d'optimalité [43]. Il représente la somme des carrés des erreurs de prédictions, et qui est mentionné par l'Eq.III.25.

La minimisation du critère $J_N(\theta)$ consiste à trouver un optimum, c'est à dire de calculer sa dérivée :

$$\left[\frac{\partial J_N(\theta)}{\partial \theta} \right]_{\theta=\hat{\theta}(N)} = 0 \quad (\text{III.32})$$

$$\left[\frac{\partial J_N(\theta)}{\partial \theta} \right]_{\theta=\hat{\theta}(N)} = -\frac{2}{N} \left\{ \sum_{t=1}^N \varphi(t) [y(t) - \theta^T \varphi(t)] \right\}_{\theta=\hat{\theta}(N)} \quad (\text{III.33})$$

A partir de ces deux dernières équations (Eq.III.32) et (Eq.III.33), on déduit la solution optimale au sens des moindres carrés de la forme suivante :

$$\hat{\theta}(N) = \left[\sum_{t=1}^N \varphi(t) \varphi^T(t) \right]^{-1} \cdot \sum_{t=1}^N \varphi(t) y(t) \quad (\text{III.34})$$

Nous constatons que la matrice $(\varphi(t) \varphi^T(t))$ est grande, si le nombre d'échantillons N est important, d'où le calcul de son inverse n'est pas conseillé, pour cela on utilise l'estimation récursive des moindres carrés .

III.4.3. Algorithme des moindres carrés récursif (RLS)

Pour la mise en œuvre de l'algorithme récursif, on pose :

$$R(t) = \sum_{k=1}^t \varphi(k) \varphi^T(k) = R(t-1) + \varphi(t) \varphi^T(t) \quad (\text{III.35})$$

D'après les équations (III.34) et (III.35), on a :

$$\hat{\theta}(t) = R^{-1}(t) \sum_{k=1}^t \varphi(k) y(k) \quad (\text{III.36})$$

$$\hat{\theta}(t) = R^{-1}(t) \left[\sum_{k=1}^{t-1} \varphi(k) y(k) + \varphi(t) y(t) \right] \quad (\text{III.37})$$

$$\hat{\theta}(t) = R^{-1}(t) \left[R(t-1) \hat{\theta}(t-1) + \varphi(t) y(t) \right] \quad (\text{III.38})$$

$$\hat{\theta}(t) = R^{-1}(t) \left[R(t) \hat{\theta}(t-1) - \varphi(t) \varphi^T(t) \hat{\theta}(t-1) + \varphi(t) y(t) \right] \quad (\text{III.39})$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + R^{-1}(t) \varphi(t) \left[y(t) - \hat{\theta}^T(t-1) \varphi(t) \right] \quad (\text{III.40})$$

D'après cette dernière équation (III.40), on remarque que la solution des moindres carrés récursive contient le terme $R^{-1}(t)$ qui nécessite une inversion matricielle à chaque instant t .

Rappelons, le lemme d'inversion matricielle qui se présente sous la forme [11] :

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1}DA^{-1} \quad (\text{III.41})$$

Nous posons:

$$A = R(t-1), B = \varphi(t), C = 1, D = \varphi^T(t)$$

Or

$$R^{-1}(t) = [R(t-1) + \varphi(t) \varphi^T(t)]^{-1} \quad (\text{III.42})$$

En utilisant le lemme d'inversion matricielle [1], l'équation (III.42) peut se réécrire :

$$R^{-1}(t) = R^{-1}(t-1) - \frac{R^{-1}(t-1) \varphi(t) \varphi^T(t) R^{-1}(t-1)}{1 + \varphi^T(t) R^{-1}(t-1) \varphi(t)} \quad (\text{III.43})$$

L'introduction de la matrice du gain d'adaptation $P(t) = R^{-1}(t)$, permet la mise en œuvre de l'algorithme des moindres carrés récursif RLS de la forme suivante :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t) \varphi(t) \left(y(t) - \hat{\theta}^T(t-1) \varphi(t) \right) \quad (\text{III.44})$$

$$P(t) = P(t-1) - \frac{P(t-1) \varphi(t) \varphi^T(t) P(t-1)}{1 + \varphi^T(t) P(t-1) \varphi(t)} \quad (\text{III.45})$$

III.4.4. Propriétés de l'estimateur des moindres carrés

L'estimateur $\hat{\theta}$ est non biaisé si cette condition est vérifiée :

$$E \left[\hat{\theta}(t) \right] = \theta_0 \quad (\text{III.46})$$

En remplaçant l'équation (III.29) dans (III.34), on obtient :

$$\hat{\theta}(N) = \left[\sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \sum_{k=1}^N \varphi(k) [\varphi^T(k) \theta_0 + e(k)] \quad (\text{III.47})$$

$$\hat{\theta}(N) = \theta_0 + \left(\left[\sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \cdot \sum_{k=1}^N \varphi(k) e(k) \right) \quad (\text{III.48})$$

Donc, pour que la méthode des moindres carrés fournisse une estimation non biaisée c'est-à-dire l'équation (III.46) est vérifiée, il faut que:

- $e(k)$ et $\varphi(k)$ sont statistiquement indépendants (non corrélés),
- $e(k)$ est une séquence aléatoire centrée (moyenne nulle).

III.5. Algorithme des moindres carrés étendu (RELS)

Cette méthode a été développée pour pouvoir identifier sans biais des modèles « système + perturbation » de type ARMAX dont la structure suivante :

$$A(z) y(t) = B(z) u(t) + C(z) e(t) \quad (\text{III.49})$$

Rappelons que l'algorithme RLS suppose que le bruit additif en sortie est un bruit blanc par contre dans le cas du RELS considère que le bruit ajouté en sortie est un bruit coloré. Ce bruit peut être considéré comme étant le résultat du filtrage d'un bruit blanc.

L'idée est d'identifier simultanément le modèle du procédé et le modèle de la perturbation, pour pouvoir obtenir une erreur de prédiction asymptotiquement "blanche" [44].

III.5.1. Structure d'un modèle ARMAX

Le modèle ARMAX ayant la structure suivante :

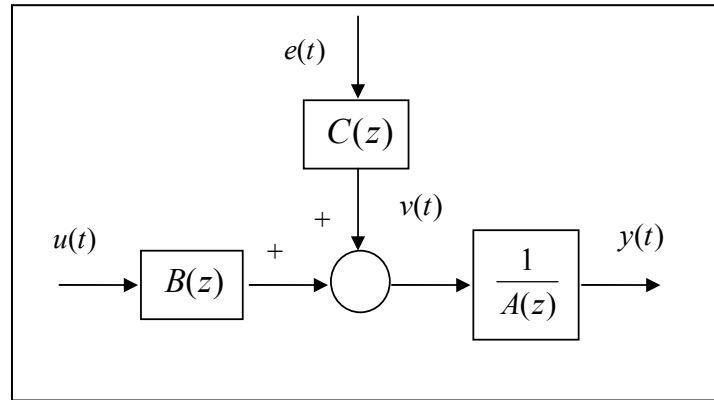


Fig.III.3 : Structure du modèle ARMAX

Reprenons le modèle ARMAX décrit par l'équation suivante [7] :

$$y(t) = \sum_{i=0}^m b_i u(t-i) - \sum_{i=1}^n a_i y(t-i) + v(t) \quad (\text{III.50})$$

Avec

n, m : Ordre du modèle

$u(t), y(t)$: Entrée, Sortie du système respectivement

$v(t)$: Bruit coloré.

La perturbation $v(t)$ est considérée comme le résultat d'un bruit blanc filtré, et on écrit:

$$v(t) = e(t) + c_1 e(t-1) + c_2 e(t-2) + \dots + c_p e(t-p) \quad (\text{III.51})$$

En utilisant la transformée en Z, l'Eq.III.50 peut s'écrire sous cette forme:

$$\sum_{i=0}^n a_i z^{-i} Y(z) = \sum_{i=0}^m b_i z^{-i} U(z) + \sum_{i=0}^p c_i z^{-i} E(z) \quad (\text{III.52})$$

Nous posons:

$$A(z) = \sum_{i=0}^n a_i z^{-i}, \quad B(z) = \sum_{i=0}^m b_i z^{-i}, \quad C(z) = \sum_{i=0}^p c_i z^{-i}$$

Ce qui permet de réécrire l'équation (III.52) sous la forme:

$$Y(z) = \frac{B(z)}{A(z)}U(z) + \frac{C(z)}{A(z)}E(z) \quad (\text{III.53})$$

Or

$$H_1(z) = \frac{B(z)}{A(z)} \quad ; \quad H_2(z) = \frac{C(z)}{A(z)}$$

Avec

$H_1(z)$: Fonction de transfert du système.

$H_2(z)$: Fonction de transfert du filtre.

III.5.2. Détermination des paramètres du modèle ARMAX

On peut donc, penser de manière intuitive que si l'on augmente le nombre d'observations, le problème se ramènera à la résolution d'un système d'équations linéaires. Dans le cas d'un système bruité, nous effectuons N mesures d'observations, nous pouvons écrire d'après l'équation III.50, une forme matricielle :

$$y(t) = \theta^T \varphi(t) + e(t) \quad (\text{III.54})$$

Posons pour la commodité des notations [5] :

$$\theta^T = [a_1 \ a_2 \ \dots \ a_n, \ b_1 \ b_2 \ \dots \ b_m, \ c_1 \ c_2 \ \dots \ c_p]$$

$$\varphi^T(t) = [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n), \ u(t-1)u(t-2) \ \dots \ u(t-m), \ e(t-1)e(t-2) \ \dots \ e(t-p)]$$

On définit l'erreur de prédiction a priori comme étant la différence entre la sortie du système et la sortie du modèle :

$$\varepsilon(t) = y(t) - \hat{y}(t) \quad (\text{III.55})$$

Sachant que :

$$\hat{y}(t) = \hat{\theta}^T(t-1)\varphi(t) \quad (\text{III.56})$$

Avec $\hat{\theta}(t-1)$ représente les paramètres estimés.

III.5.3. Les équations de la mise en œuvre de l'algorithme RELS

Le problème de la méthode des moindres carrés ordinaire c'est que $e(t)$ n'est pas mesurable, donc il faut l'estimer. D'après les équations (III.50), et (III.51), on écrit :

$$e(t) = y(t) + a_1 y(t-1) + \dots + a_n y(t-n) - b_1 u(t-1) - \dots - b_m u(t-m) - c_1 e(t-1) \dots - c_p e(t-p) \quad (\text{III.57})$$

On appelle $\varepsilon(t)$ l'estimateur de $e(t)$, d'où l'Eq.III.57 s'écrira sous la forme :

$$\varepsilon(t) = y(t) + \hat{a}_1 y(t-1) + \dots + \hat{a}_n y(t-n) - \hat{b}_1 u(t-1) - \dots - \hat{b}_m u(t-m) - \hat{c}_1 \varepsilon(t-1) \dots - \hat{c}_p \varepsilon(t-p) \quad (\text{III.58})$$

La prédiction ajustable a priori dans le cas des moindres carrés étendus s'obtient de l'équation (III.56), en remplaçant les paramètres connus par les paramètres estimés :

$$\hat{\theta}^T(t) = [\hat{a}_1(t) \hat{a}_2(t) \dots \hat{a}_n(t), \hat{b}_1(t) \hat{b}_2(t) \dots \hat{b}_m(t), \hat{c}_1(t) \hat{c}_2(t) \dots \hat{c}_p(t)]$$

$$\varphi^T(t) = [-y(t-1) -y(t-2) \dots -y(t-n), u(t-1) u(t-2) \dots u(t-m), \varepsilon(t-1) \varepsilon(t-2) \dots \varepsilon(t-p)]$$

En résumé les équations permettant la mise en œuvre de l'algorithme RELS seront :

$$\varepsilon(t) = y(t) - \hat{\theta}^T(t-1) \varphi(t) \quad (\text{III.59})$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t) \varphi(t) \varepsilon(t) \quad (\text{III.60})$$

$$P(t) = P(t-1) - \frac{P(t-1) \varphi(t) \varphi^T(t) P(t-1)}{1 + \varphi^T(t) P(t-1) \varphi(t)} \quad (\text{III.61})$$

III.5.4. Initialisation de l'algorithme RELS

L'initialisation concernant la matrice $P(t)$ et le vecteur des paramètres $\hat{\theta}(t)$ se fait souvent comme suit :

$$\hat{\theta}(0) = \begin{cases} \theta_p & \text{à priori} \\ 0 & \text{si non} \end{cases}$$

$P(0) = C.I$ tel que C : Constante et I : matrice identité.

III.6. Conclusion

Nous avons présenté dans ce chapitre les méthodes récursives, d'identification de modèles paramétriques des signaux et systèmes. L'accent a été mis sur la présentation de la structure des différentes méthodes, leur raison d'être, les conditions de convergence dans un environnement stochastique et sur leur domaine d'application.

La structure des algorithmes d'adaptation paramétrique commune à toutes les méthodes d'identification a été présentée en détail indiquant l'influence du choix des différents paramètres sur la précision de l'estimation et la capacité de suivi de variation des modèles. Les algorithmes d'identification paramétrique développés sont :

- Algorithme du Gradient stochastique
- Algorithme de Newton stochastique
- Algorithme des Moindres carré récursif
- Algorithme des Moindres carrés étendu.

Dans le chapitre qui suit, nous aborderons l'implémentation de ces algorithmes sur MATLAB ainsi que les résultats obtenus par simulation pour tester les performances de chaque algorithme.

Chapitre **IV**

Implémentation des Algorithmes et Résultats de Simulation

IV.1. Introduction

Dans ce chapitre, nous allons présenter les résultats obtenus par l'application des algorithmes d'identification paramétrique pour identifier les paramètres d'un processus physique tel que l'algorithme du gradient, Newton stochastique et moindres carrés récursif d'une part et d'autre part une étude comparative entre ces algorithmes pour envisager les avantages et les inconvénients de ces algorithmes.

IV.2. Etapes de simulation

Pour bien identifier les paramètres inconnus d'un processus physique, il est meilleur de suivre les étapes suivantes :

IV.2.1. Choix des paramètres du système

Les paramètres du système que nous allons utiliser tout le long de notre travail sont choisis de telle façon que notre système est stable, que ce soit le système à phase minimale ou non minimale. A titre d'exemple de simulation, on prend :

1^{er} Cas: Système à phase minimale:

$$H_1(z) = z^{-1} \frac{1 + 0.5z^{-1}}{1 + 0.3z^{-1} + 0.8z^{-2}} \quad (\text{IV.1})$$

2^{ème} Cas: Système à phase non minimale :

$$H_2(z) = z^{-1} \frac{1 + 2z^{-1}}{1 + 0.3z^{-1} + 0.8z^{-2}} \quad (\text{IV.2})$$

IV.2.2. Choix des signaux de test

L'objectif de notre travail de simulation est d'identifier les paramètres inconnus d'un processus physique. Pour cela, on utilise des signaux de test qui sont très riches en fréquences de la Fig.IV.1 et Fig.IV.2.

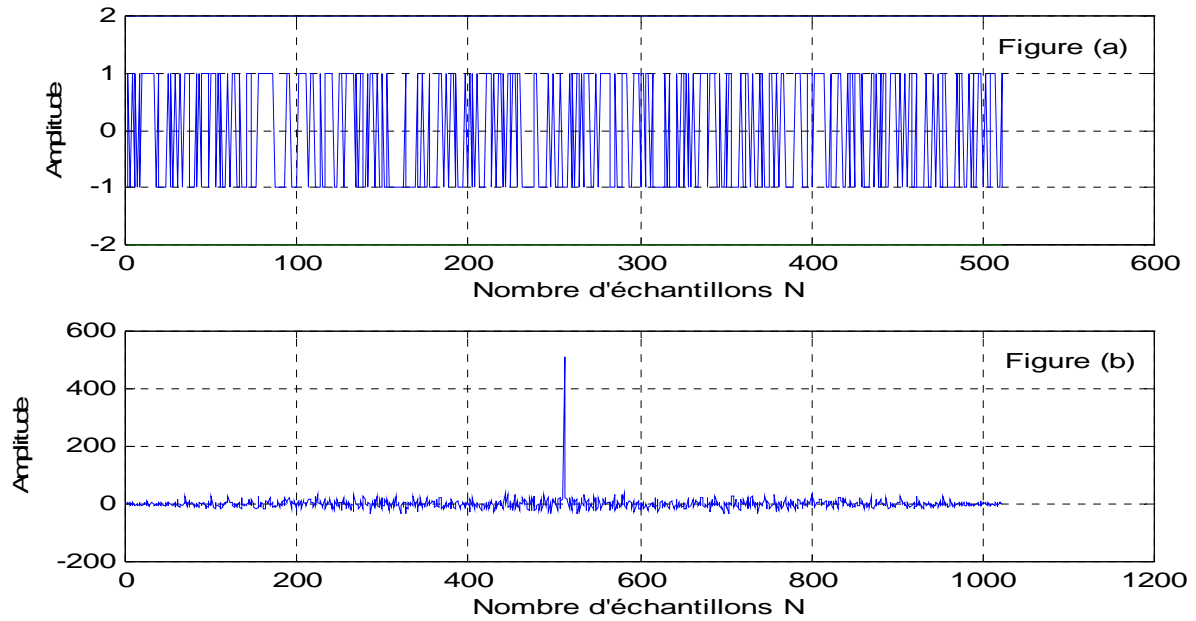


Fig. IV.1: Séquence d'entrée SBPA et sa fonction d'autocorrelation
 $(m_u = 0, \sigma_u^2 = 1)$

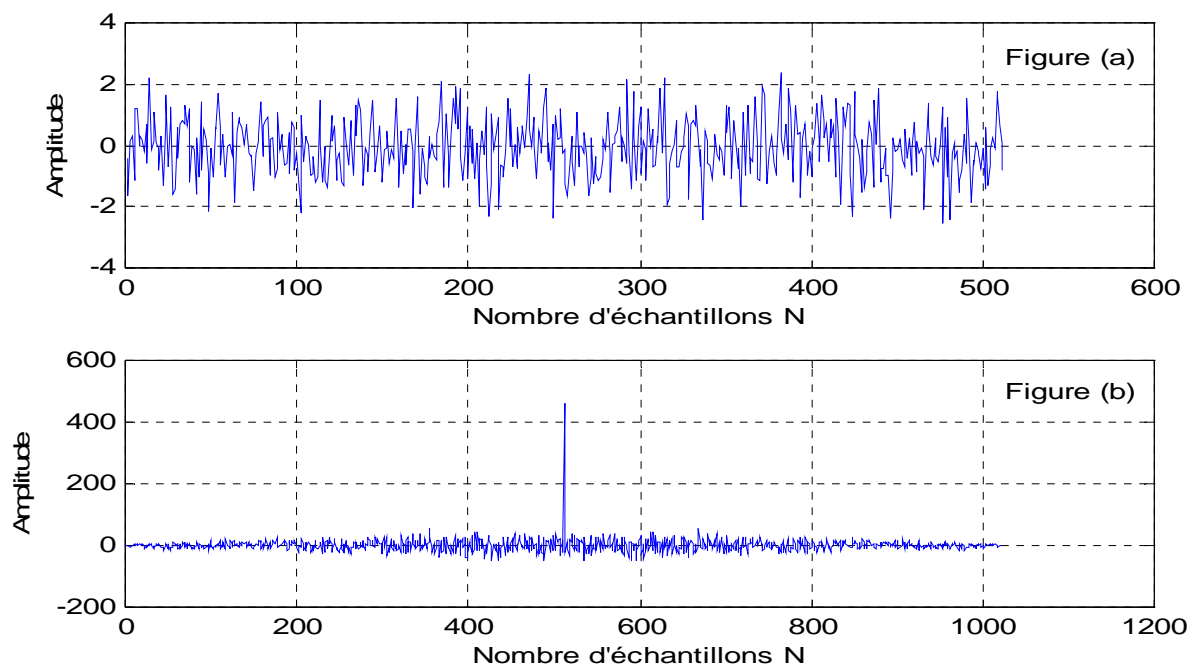


Fig. IV.2 : Bruit additif gaussien et sa fonction d'autocorrelation
 $m_v = 0, \sigma_v^2 = 1$

La figure en haut, nous montre une Séquence Binaire Pseudo Aléatoire (SBPA) et sa fonction d'autocorrelation qui rapproche à un bruit blanc. Par contre, la figure en bas représente un bruit additif Gaussien de moyenne nulle et de variance égale à l'unité. Ce dernier est généré par la fonction RANDN du logiciel MATLAB.

IV.2.3. Identification des paramètres du modèle

Les algorithmes d'identification paramétrique étudiés sont appliqués pour identifier les paramètres d'un système stable, pour cela, plusieurs tests de simulations ont été effectués afin d'établir les performances de chaque algorithme.

IV.2.4. Test de validation

La validation du modèle identifié a une grande importance car elle permet de juger son aptitude à représenter un certain type de signal. La validation porte essentiellement sur deux aspects:

- La variance de l'erreur de prédiction résiduelle
- Les propriétés statistiques des erreurs de prédiction résiduelle.

En effet si la méthode d'estimation utilisée est appropriée pour un type de signal, les propriétés statistiques de l'erreur de prédiction résiduelle sont celles qui correspondent à la convergence de l'algorithme. En concret, on teste:

- Soit la blancheur des erreurs résiduelles, pour les méthodes d'estimation basées sur le blanchissement asymptotique de l'erreur de prédiction, par calcul des autocorrélations.
- Soit la décorrélation entre l'erreur de prédiction résiduelle et le signal prédit par le modèle, pour les méthodes basées sur la décorrélation asymptotique du vecteur des observations et de l'erreur de prédiction (variable instrumentale, erreur de sortie), par calcul des intercorrélations.

IV.3. Mise en œuvre des algorithmes implémentés

Un simple menu de la figure IV.3 a fait l'objet d'une implémentation des algorithmes d'identification paramétrique sur MATLAB. Il permet de traiter des fichiers préalablement enregistrés sous forme des fonctions, et d'exécuter la fonction correspondante à l'algorithme d'identification (LMS, SN, RLS et RELS). Enfin, un ensemble de représentations graphiques permet de visualiser l'évolution temporelle des paramètres estimés du modèle, l'erreur d'identification et la fonction d'autocorrelation de l'erreur de prédiction, ce qui en fait un véritable outil de recherche utilisé lors de la simulation.

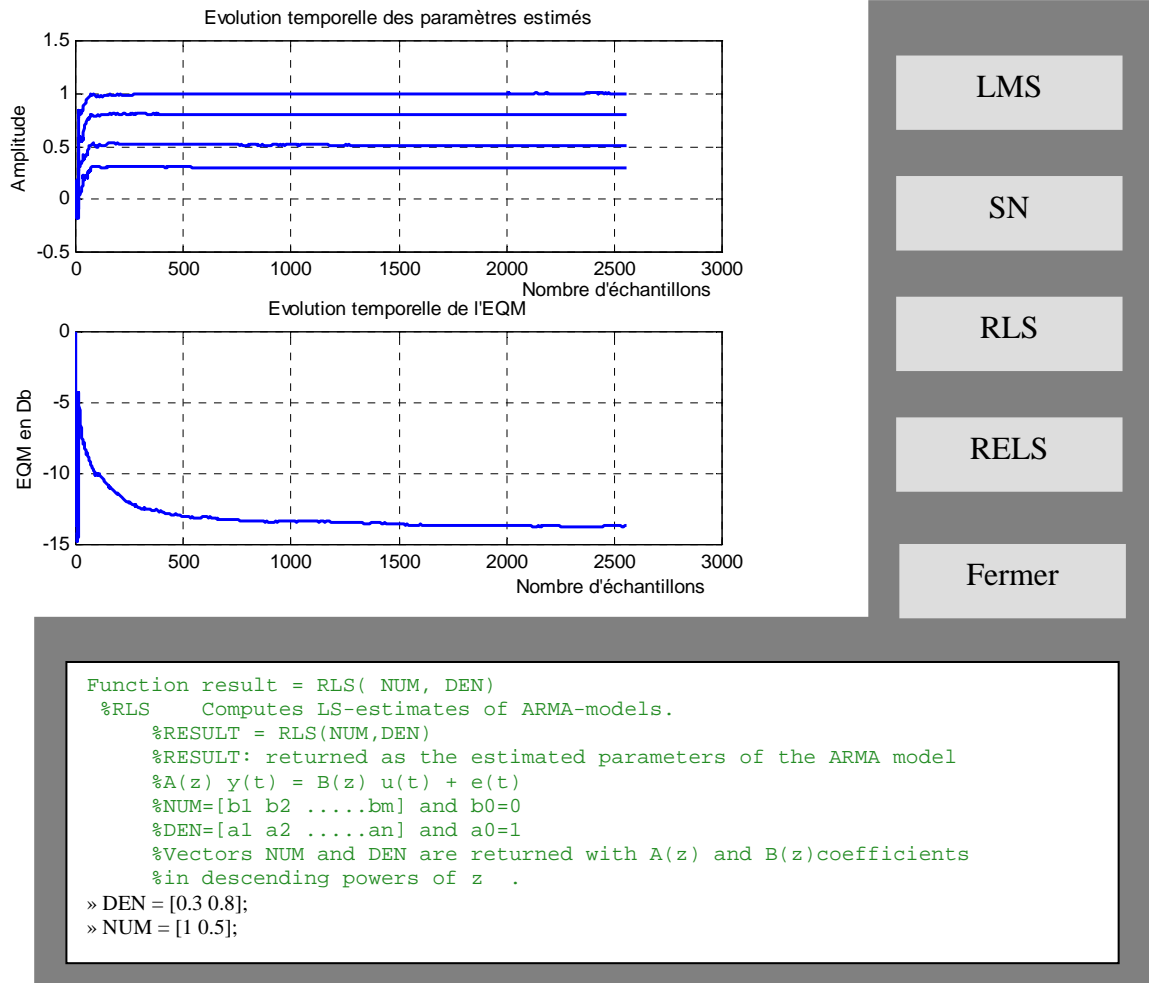


Fig. IV.3: Menu simple des algorithmes implémentés sous MATLAB.

IV.4. Identification des paramètres par l'algorithme LMS

Les résultats de simulation obtenus par l'algorithme LMS sont illustrés sur les tableaux et les figures qui suivent. En prenant deux exemples différents, le premier exemple correspond à un système à phase minimale et le second correspond à système à phase non minimale.

IV.4.1. Système à phase minimale

Considérons un système stable d'ordre 2 dont les zéros de la fonction de transfert sont à l'intérieur du cercle unité. Equation IV.1, s'écrira donc sous cette forme :

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) + b_1 u(t-1) + b_2 u(t-2) + v(t) \quad (\text{IV.3})$$

Avec $a_1 = 0.3$, $a_2 = 0.8$, $b_1 = 1$, $b_2 = 0.5$

IV.4.1.1. Influence du nombre d'échantillons

Pour tester l'influence du nombre d'échantillons N sur le biais d'estimation des paramètres du système, on dresse le tableau ci-dessous suivi par l'évolution des paramètres du système ainsi que l'auto corrélation de l'erreur de prédiction (Figure IV.4).

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
256*2	0.2217	0.7263	0.9539	0.4130	0.2629
256*4	0.3115	0.8056	0.9979	0.4925	0.0585
256*10	0.3039	0.7959	0.9997	0.5008	0.0251
256*20	0.3011	0.8211	0.9919	0.4990	0.0122
256*50	0.2856	0.8121	1.0156	0.4802	0.0060

Tab. IV.1 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.01$, $\mu = 0.0145$)

Les résultats obtenus montrent que lorsque le nombre d'échantillons N est grand alors les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est asymptotiquement blanche.

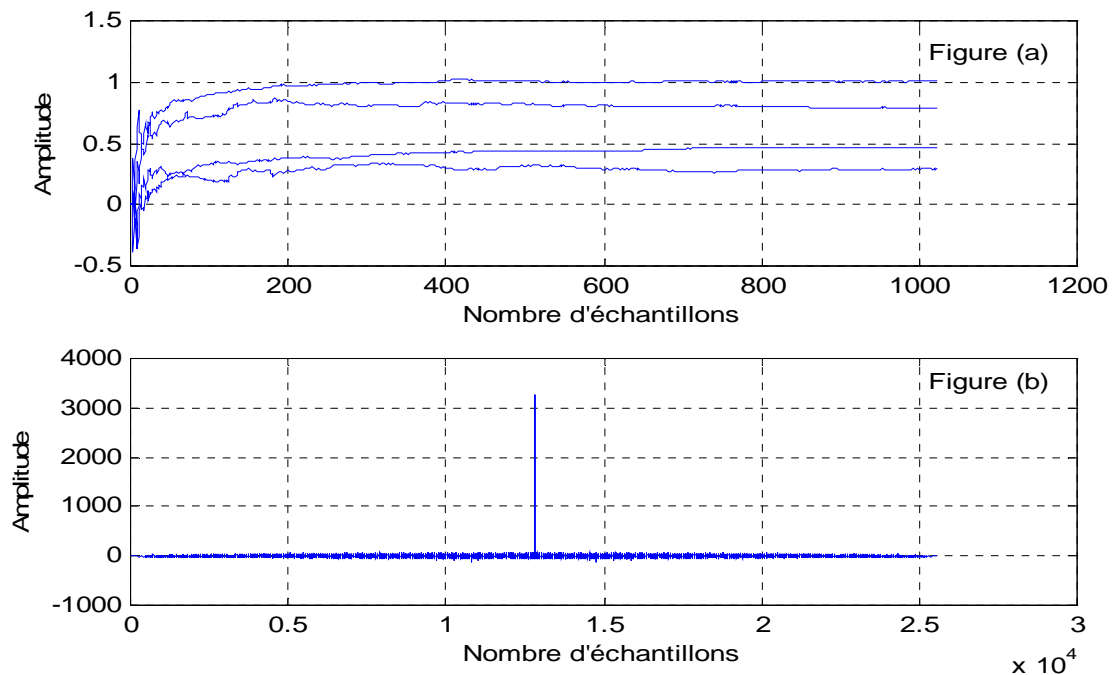


Fig. IV.4 : (a) - Evolution des paramètres ($N = 256*4$, $\mu = 0.0145$)
(b) - Autocorrelation de l'erreur de prédiction

IV.4.1.2. Influence du gain d'adaptation

Pour différentes valeurs du gain d'adaptation μ , les résultats obtenus par simulation sont illustrés dans le tableau IV.2, suivi par l'évolution temporelle des paramètres du système aux cas de la convergence lente et rapide (Fig. IV.5).

Facteur μ	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
0.0049	0.2774	0.7873	0.9877	0.4889	0.1468
0.0098	0.2949	0.8021	1.0005	0.5035	0.0782
0.0144	0.3017	0.7968	0.9877	0.4924	0.0467
0.0194	0.2987	0.7933	0.9793	0.5169	0.0429

Tab. IV.2 : Influence du gain d'adaptation μ , ($\sigma^2 = 0.01$, $N = 256*4$)

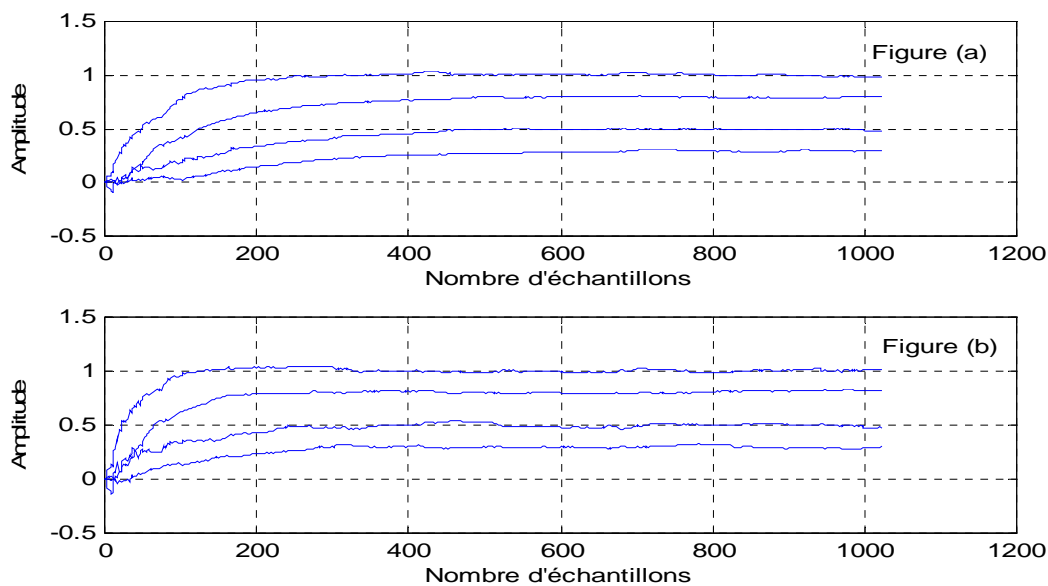


Fig. IV.5 : (a) – Convergence lente $\mu = 0.0097$
(b) – Convergence rapide $\mu = 0.0193$

La figure ci-dessus montre un exemple de convergences (lente et rapide) pour deux valeurs différentes du pas d'adaptation de l'algorithme LMS. Si μ augmente alors le temps de convergence des paramètres estimés vers les vraies valeurs diminue. Pour $\mu = 0.0097$, le temps de convergence correspond à 400 échantillons, par contre si $\mu = 0.0193$ alors la convergence est assurée après 200 échantillons.

IV.4.2. Système à phase non minimale

Considérons un système stable d'ordre 2 dont les zéros de la fonction de transfert sont situés à l'extérieur du cercle unité (Eq. IV.2), avec $a_1 = 0.3$, $a_2 = 0.8$, $b_1 = 1$, $b_2 = 2$.

IV.4.2.1. Influence du nombre d'échantillons

Pour différentes valeurs de N , on obtient les résultats du tableau IV.3, suivi par l'évolution des paramètres et l'autocorrelation de l'erreur de prédiction (Figure IV.6).

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
256*2	0.2878	0.7899	0.9297	1.8581	0.2830
256*4	0.2957	0.8061	0.9835	2.0068	0.1602
256*10	0.2992	0.8024	0.9932	1.9925	0.0616
256*20	0.2966	0.7998	0.9896	1.9994	0.0338
256*50	0.3025	0.7931	0.9823	2.0100	0.0135

Tab. IV.3 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.01$, $\mu = 0.0050$)

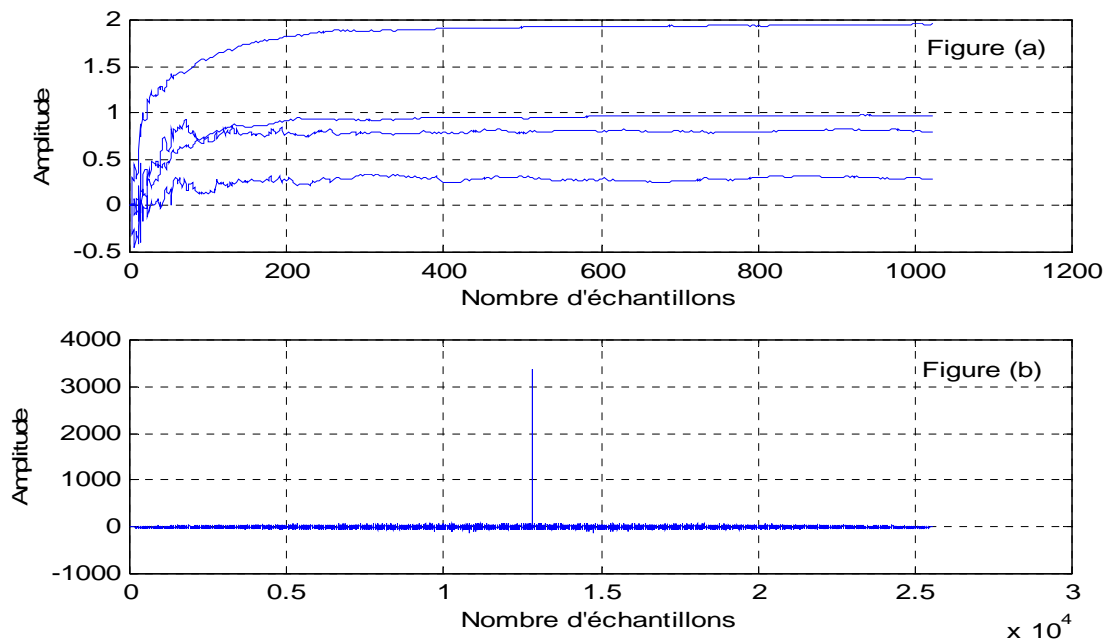


Fig. IV.6 : (a) - Evolution des paramètres ($N = 256*4$, $\sigma^2 = 0.01$).
(b) - Autocorrelation de l'erreur de prédiction.

Nous remarquons que les paramètres estimés convergent vers les paramètres réels du système et la position des zéros du système dans la carte des pôles et des zéros n'influe pas sur la convergence de l'algorithme. L'erreur de prédiction est considérée asymptotiquement blanche.

IV.4.2.2. Influence du gain d'adaptation

Pour différentes valeurs de μ , on obtient les résultats du tableau IV.4, suivi par l'évolution des paramètres du système en cas de la convergence lente et rapide (Figure IV.7).

Facteur μ	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
0.0033	0.3016	0.8013	1.0129	1.9921	0.1047
0.0049	0.2995	0.8039	0.9976	2.0019	0.0611
0.0066	0.2988	0.7960	1.0160	1.9866	0.0426
0.0098	0.3137	0.8114	0.9922	1.9861	0.0225

Tab. IV.4 : Influence du gain d'adaptation μ ($N=256*10, \sigma^2 = 0.01$)

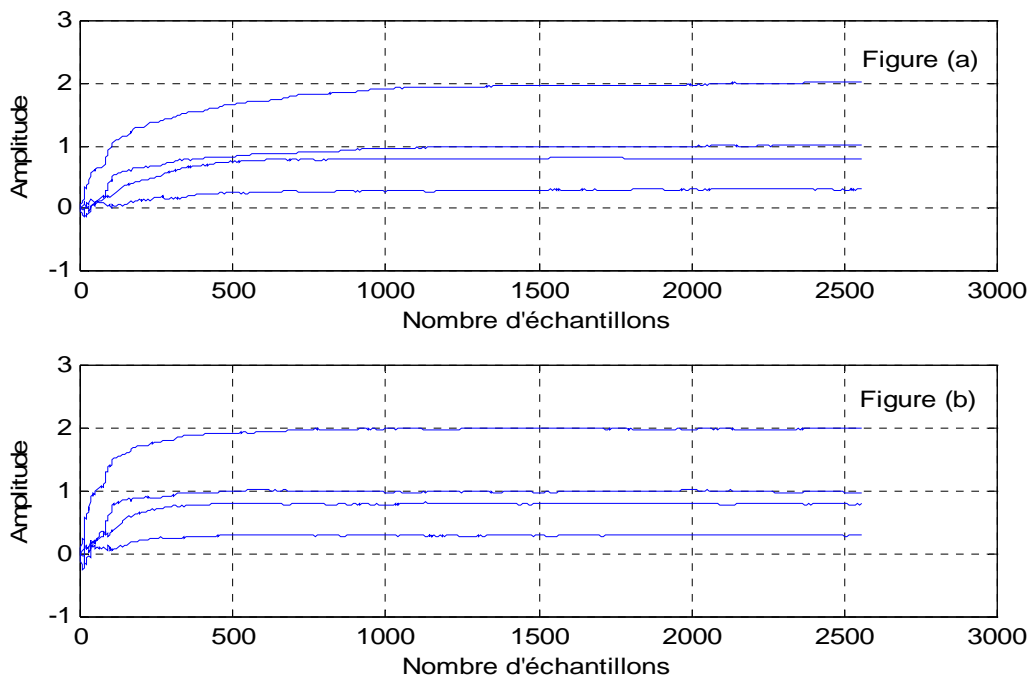


Fig. IV.7 : (a) – Convergence lente $\mu = 0.0033$
(b) – Convergence rapide $\mu = 0.0066$

D'après les résultats obtenus, nous remarquons que l'algorithme LMS permet une convergence suffisamment rapide pour un bon choix de μ et la présence des zéros en dehors du cercle unité n'influe pas sur le temps de convergence. Cependant, l'estimateur LMS est considéré asymptotiquement non biaisé.

IV.5. Identification des paramètres par l'algorithme de Newton stochastique

Les résultats de simulation par l'algorithme de Newton stochastique sont donnés par les tableaux et les figures qui suivent.

IV.5.1. Système à phase minimale

Considérons le même système physique décrit par l'équation IV.1 dont les zéros de la fonction de transfert sont à l'intérieur du cercle unité.

L'influence du nombre d'échantillons N et la variance σ^2 du bruit additif sur le biais de l'estimateur sont données dans les tableaux IV.5 et IV.6. L'évolution des paramètres en fonction du temps et l'erreur de prédiction sont illustrés sur les figures IV.8 et IV.9. On note aussi que l'erreur d'identification converge vers zéro telle qu'elle est montrée sur la figure IV.10.

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
256*2	0.2706	0.7880	1.0317	0.4587	0.0394
256*4	0.3069	0.7884	0.9981	0.5067	0.0210
256*10	0.3067	0.7970	0.9910	0.5007	0.0101
256*20	0.2991	0.8011	0.9881	0.5085	0.0063
256*50	0.2974	0.8014	0.9972	0.4958	0.0059

Tab.IV.5 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)

variance σ^2	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
1.00	0.2905	0.8016	1.0343	0.5018	0.0195
0.25	0.2987	0.8057	1.0078	0.5142	0.0100
0.04	0.2992	0.8013	0.9958	0.4956	0.0096
0.01	0.3014	0.8034	1.0020	0.4995	0.0073

Tab. IV.6 : Influence de la variance σ^2 , ($N = 256*10$)

Les tableaux ci-dessus montrent lorsque le nombre d'échantillons augmente ou la variance du bruit additif diminue alors le biais de l'estimateur diminue, et les paramètres estimés sont plus proches aux valeurs réelles.

Par exemple pour un choix de $N = 256 \cdot 10$, $\sigma^2 = 0.25$ un nombre de 500 échantillons est suffisamment large pour assurer la convergence aux valeurs réelles avec une erreur d'identification de l'ordre de un millième comme nous montrent les figures ci-dessous.

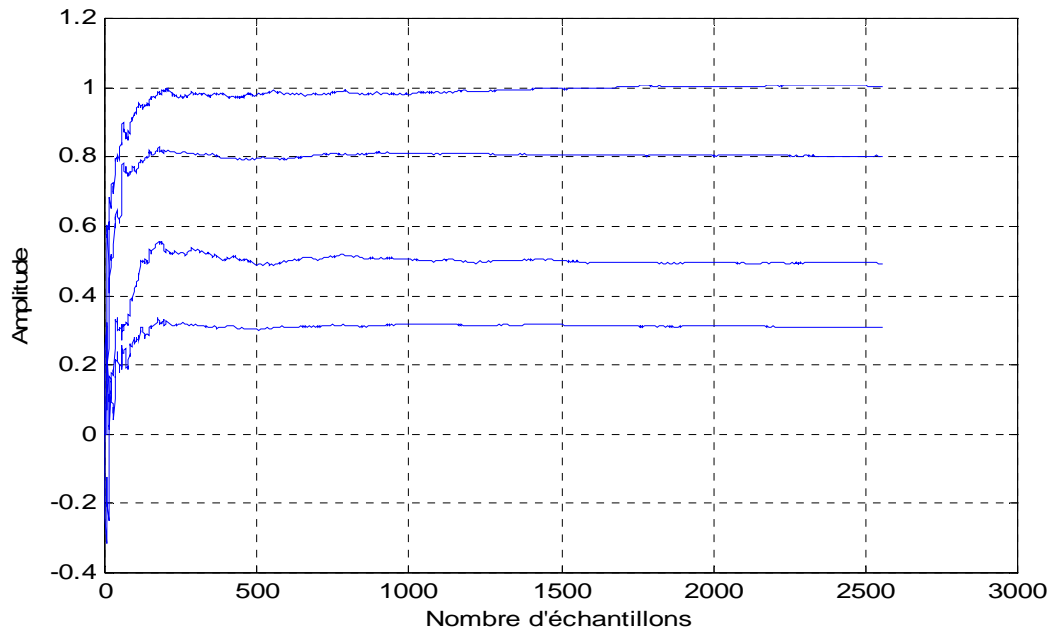


Fig. IV.8 : Evolution des paramètres ($N = 256 \cdot 10$, $\sigma^2 = 0.25$)

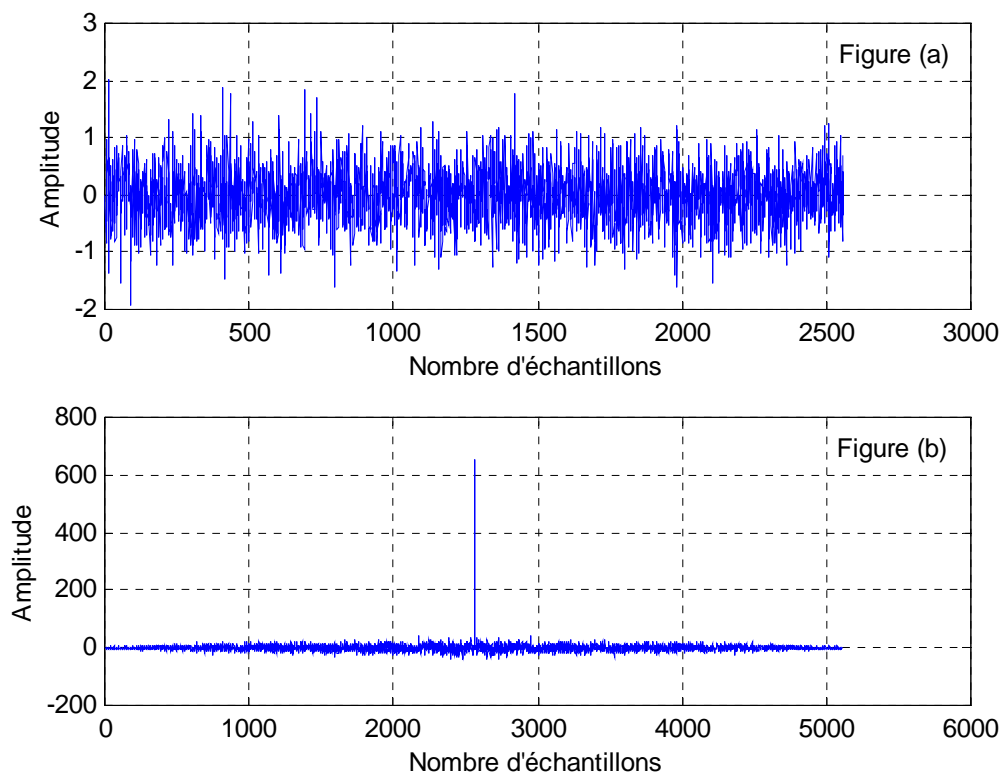


Fig. IV.9 : (a) - Erreur de prédiction ($N = 256 \cdot 10$, $\sigma^2 = 0.25$)
(b) - Autocorrelation de l'erreur de prédiction.

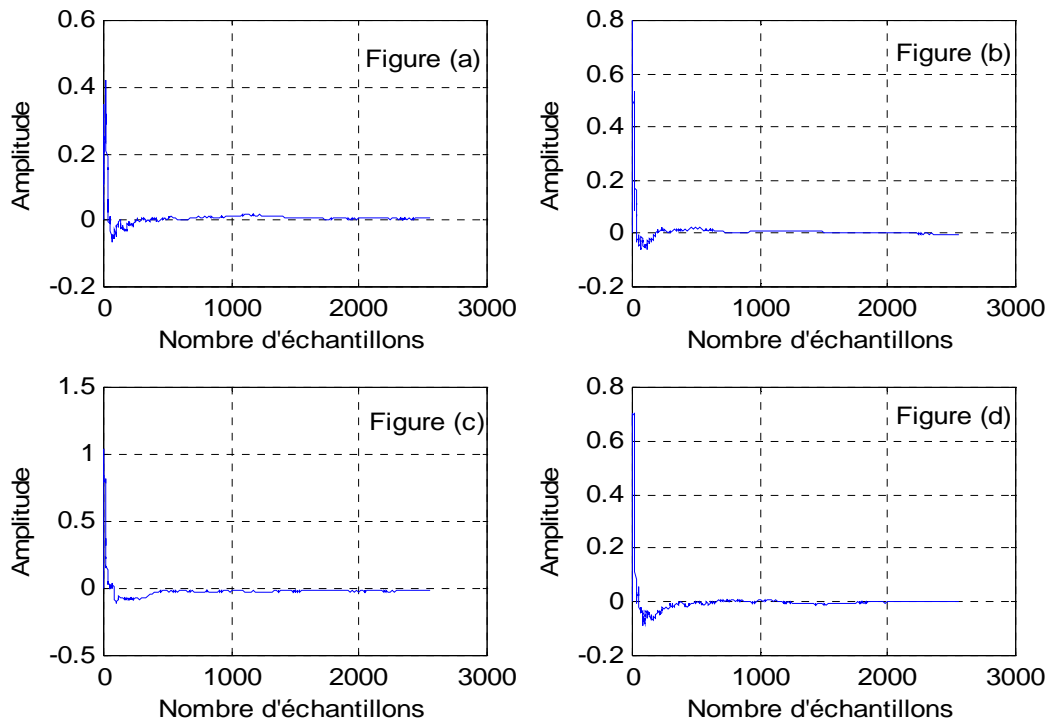


Fig. IV.10 : (a) - Erreur d'identification du paramètre a_1
 (b) - Erreur d'identification du paramètre a_2
 (c) - Erreur d'identification du paramètre b_1
 (d) - Erreur d'identification du paramètre b_2

IV.5.2. Système à phase non minimale

Considérons le même système décrit par l'équation IV.2 dont les zéros de la fonction de transfert sont à l'extérieur du cercle unité (phase non minimale). Les paramètres estimés par la méthode de Newton stochastique sont donnés par les tableaux IV.7, IV.8. L'erreur d'identification de ces paramètres est illustrée sur la figure IV.11.

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
256*2	0.2821	0.7967	1.0041	1.9707	0.0474
256*4	0.2924	0.7953	0.9974	2.0027	0.0300
256*10	0.3004	0.7994	0.9988	2.0070	0.0103
256*20	0.2989	0.8003	0.9881	2.0078	0.0103
256*50	0.3002	0.8004	1.0002	2.0090	0.0032

Tab. IV.7 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)

Variance σ^2	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
1.00	0.3004	0.8028	1.0349	2.0104	0.0163
0.25	0.2987	0.7994	0.9970	1.9983	0.0115
0.04	0.2986	0.7971	0.9970	2.0026	0.0102
0.01	0.2995	0.8007	1.0001	1.9984	0.0100

Tab. IV.8 : Influence du variance σ^2 , ($N = 256*10$)

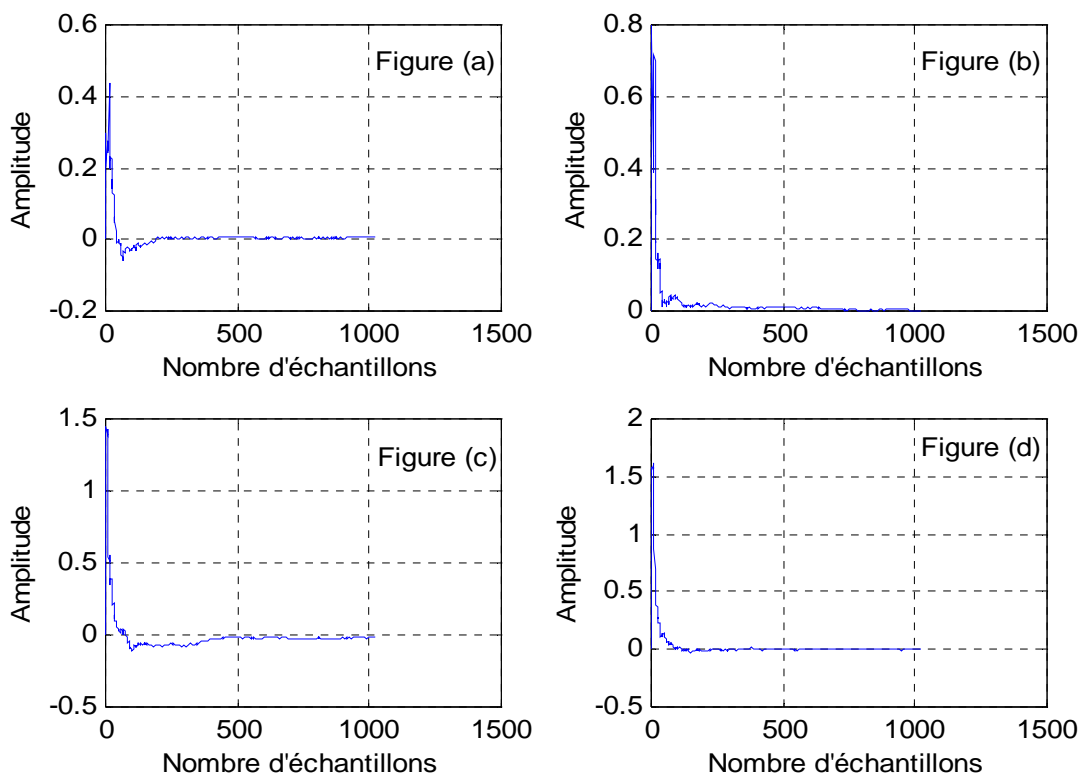


Fig. IV.11 : (a) - Erreur d'identification du paramètre a_1
 (b) - Erreur d'identification du paramètre a_2
 (c) - Erreur d'identification du paramètre b_1
 (d) - Erreur d'identification du paramètre b_2

Sur les tableaux IV.7 et IV.8 on voit que le fait de prendre le nombre d'échantillons $N = 256*10$ et la variance du bruit additif varie dans l'intervalle $[0.01 \quad 1]$ entraîne un biais de l'estimateur de l'ordre de 0.01.

IV.6. Identification des paramètres par l'algorithme RLS

Les résultats de simulation par l'algorithme des moindres carrés récursifs sont illustrés sur les tableaux et les figures qui suivent.

IV.6.1. Système à phase minimale

Considérons le même système décrit par l'équation IV.1 dont les zéros de la fonction de transfert sont à l'intérieur du cercle unité (phase minimale).

Le tableau IV.9 montre les résultats de simulation obtenus lorsque nous faisons varier le nombre d'échantillons N . Par contre le tableau IV.10 est réservé pour les résultats obtenus en fonction de la variance σ^2 avec N constant.

La figure IV.12 montre l'évolution temporelle des paramètres estimés suivi par la fonction d'autocorrelation de l'erreur de prédiction.

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
2*256	0.2901	0.8072	0.9745	0.4519	0.0302
4*256	0.2956	0.8002	0.9909	0.4802	0.0241
10*256	0.2960	0.8002	0.9885	0.4992	0.0097
20*256	0.2953	0.8029	1.0134	0.4852	0.0050
50*256	0.2984	0.7997	0.9971	0.4968	0.0023

Tab.IV.9 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)

variance σ^2	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
1.00	0.3024	0.7986	1.0069	0.4966	0.0107
0.25	0.3001	0.7930	0.9942	0.4982	0.0094
0.04	0.2972	0.8015	0.9979	0.4947	0.0065
0.01	0.3017	0.8001	0.9982	0.5016	0.0047

Tab. IV.10 : Influence de la variance σ^2 , ($N=256*10$)

Nous remarquons que si le nombre d'échantillons augmente ou la variance du bruit additif diminue alors le biais de l'estimateur RLS diminue et le temps de convergence sera minimal de l'ordre de 200 échantillons comparativement aux algorithmes précédents.

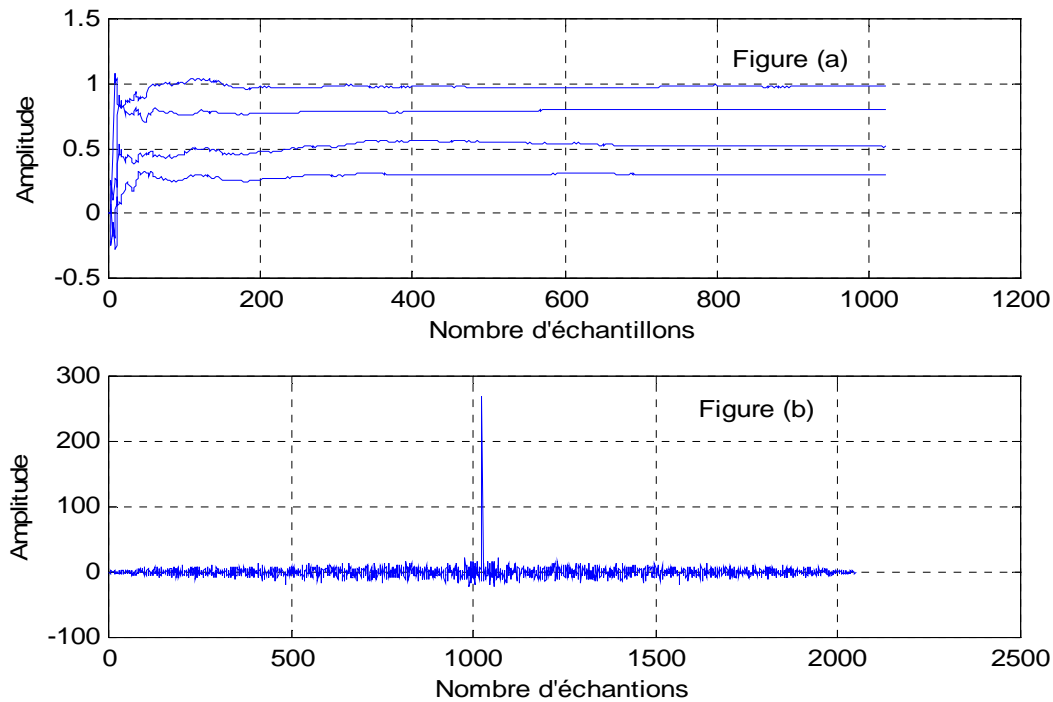


Fig. IV.12 : (a) - Evolution des paramètres ($N = 256 \cdot 4$, $\sigma^2 = 0.25$).
 (b) - Autocorrelation de l'erreur de prédiction.

IV.6.2. Système à phase non minimale

Considérons le même système décrit par l'équation IV.2 dont les zéros de la fonction de transfert sont à l'extérieur du cercle unité (phase non minimale). Les résultats obtenus sont donnés par les tableaux IV.11, IV.12 et suivis par l'évolution des paramètres et la fonction d'autocorrelation de l'erreur de prédiction (Figure IV.13).

Nombre d'échantillons	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
2*256	0.2957	0.8017	0.9740	1.9551	0.0466
4*256	0.2994	0.8002	0.9909	1.9824	0.0263
10*256	0.2979	0.8017	1.0078	1.9832	0.0110
20*256	0.3007	0.8018	1.0040	1.9872	0.0076
50*256	0.3011	0.8002	0.9988	1.9975	0.0033

Tab. IV.11 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)

Variance σ^2	Paramètres estimés				Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	
1.00	0.3005	0.8028	1.0353	2.0113	0.0171
0.25	0.3008	0.8011	0.9911	1.9919	0.0110
0.04	0.3008	0.8013	1.0026	1.9998	0.0078
0.01	0.3000	0.8009	0.9975	1.9992	0.0076

Tab. IV.12 : Influence de la variance σ^2 , ($N=256*4$)

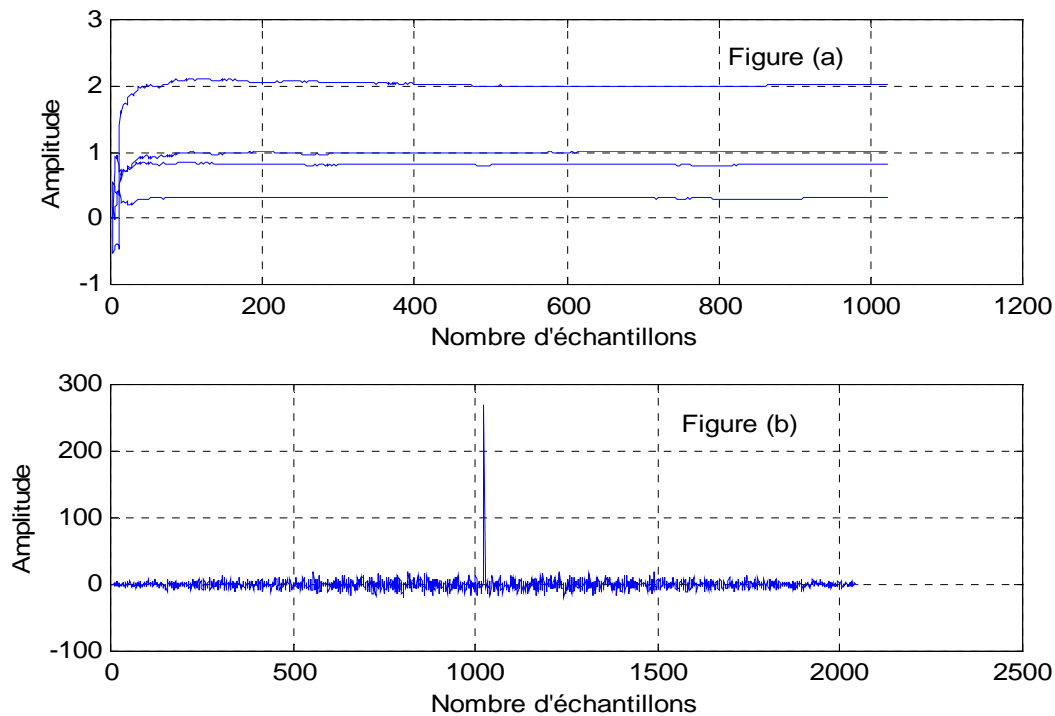


Fig. IV.13 : (a) - Evolution des paramètres ($N = 256*4$, $\sigma^2 = 0.25$).
(b) - Autocorrelation de l'erreur de prédiction.

Nous venons de montrer que, par simulation, l'algorithme RLS est plus performant que le LMS. En revanche, la comparaison théorique des deux algorithmes pour des évolutions permanentes déterministes ou aléatoires ne donne pas toujours la bonne poursuite au RLS. La réponse dépend des paramètres qui interviennent dans les résidus d'erreurs de sortie, à savoir le nombre d'échantillons N , la variance et l'ordre du système. Il peut donc exister des cas où l'algorithme LMS poursuit mieux que le RLS.

IV.7. Etude comparative des algorithmes d'identification paramétrique

La méthodologie de comparaison que nous venons de donner peut être utilisée dans des situations non stationnaires variées. Dans la théorie, il est possible de bien classer ces situations. Par contre, dans la pratique, c'est beaucoup plus difficile, lorsqu'on observe par exemple, la variabilité du signal de parole dans la même phrase prononcée par le même locuteur. C'est pourquoi, les conclusions des auteurs pourront parfois apparaître comme mitigées. Pourtant, le passage par l'outil théorique est une première étape incontournable pour pouvoir se représenter et comprendre un peu ce qui se passe dans une situation pratique.

Rappelons l'estimation des paramètres du modèle dans le cas de la méthode des moindres carrés récursifs, c'est-à-dire la mise à jour du vecteur $\hat{\theta}$ dépend de façon critique du signal résiduel ε , qui est en fait l'erreur de prédiction. Ce mécanisme conduit à une bonne précision de l'estimation car l'erreur de prédiction, reflète la différence entre le modèle et le système, est en principe toujours minimisée. Elle peut même attendre une valeur nulle s'il n'y a pas de bruit et si le modèle est parfaitement correct.

Dans ce paragraphe, nous présentons un tableau explicatif et comparatif des biais de chaque estimateur appliqué au système physique à phase minimale.

Biais	Algorithmes utilisés		
	Moindre carrés	Newton stochastique	Gradient stochastique
$E\left[\hat{\theta}-\theta_0\right]$	0.0018	0.0102	0.0177
	0.0048	0.0008	0.0176
	0.0054	-0.0031	0.0187
	-0.0034	0.0056	0.0397

Tab. IV.13 : Biais des estimateurs utilisés, $N = 256*4$, $\sigma^2 = 0.25$

De toute évidence, nous pouvons prévoir que cette imprécision de l'estimation se manifesterait plus clairement dans le cas où la structure du spectre est relativement complexe comme, par exemple, le modèle ARMAX.

IV.8. Identification des paramètres d'un modèle ARMAX

Considérons un système physique stable de type ARMAX décrit par l'équation suivante :

$$y(t) = z^{-d} \frac{B(z^{-1})}{A(z^{-1})} u(t) + \frac{C(z^{-1})}{A(z^{-1})} e(t) \quad (\text{IV.4})$$

Avec $A(z^{-1}) = a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots$

$B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots$

$C(z^{-1}) = c_0 + c_1 z^{-1} + c_2 z^{-2} + \dots$

IV.8.1. Résultats de simulation

A titre d'exemple de simulation nous prenons le cas d'un système physique stable à phase minimale dont :

$a_0 = 1, a_1 = 0.3, a_2 = 0.8$

$b_0 = 0, b_1 = 1, b_2 = 0.5$

$c_0 = 1, c_1 = 0.1, c_2 = 0.4$

L'implémentation de l'algorithme RELS sur PC, en utilisant le langage de programmation MATLAB, nous a donné les résultats suivants :

Nombre d'échantillons	Paramètres estimés						Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	\hat{c}_1	\hat{c}_2	
2*256	0.3241	0.7868	0.9730	0.5152	0.1006	0.3081	0.0581
4*256	0.2981	0.7983	1.0174	0.5192	0.0381	0.3700	0.0276
10*256	0.2969	0.7984	0.9993	0.4760	0.1077	0.3931	0.0230
20*256	0.2971	0.8024	0.9934	0.4978	0.1045	0.3943	0.0124
50*256	0.2993	0.7986	1.0111	0.4966	0.1089	0.3881	0.0088

Tab. IV.14 : Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$)

Variance σ^2	Paramètres estimés						Moyenne du Biais
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2	\hat{c}_1	\hat{c}_2	
1.00	0.2972	0.7843	1.0219	0.5065	0.0863	0.3429	0.0303
0.25	0.3042	0.8047	0.9867	0.5053	0.1181	0.3627	0.0236
0.04	0.2781	0.7939	0.9763	0.4623	0.0617	0.3806	0.0298
0.01	0.2990	0.8009	1.0013	0.4997	0.0837	0.3213	0.0355

Tab. IV.15 : Influence de la variance σ^2 , ($N = 256*10$)

Les paramètres estimés obtenus par l'algorithme RELS appliqué au système à phase minimale sont plus proches aux valeurs réelles avec un biais d'estimation de l'ordre de 2 % pour un nombre d'échantillons $N = 256 \cdot 10$ et la variance du bruit additif $\sigma^2 = 0.25$. On note aussi que, le biais de l'estimateur est devenu de plus en plus important si les zéros du modèle stochastique du bruit sont situés en dehors du cercle unité.

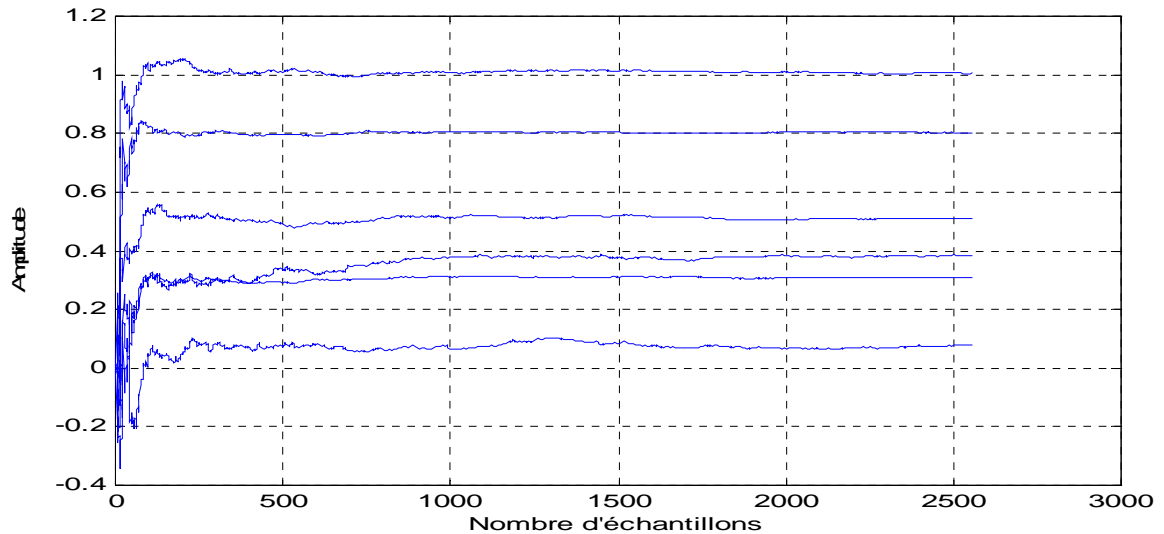


Fig. IV.14 : Evolution des paramètres ($N = 256 \cdot 10$, $\sigma^2 = 0.25$)

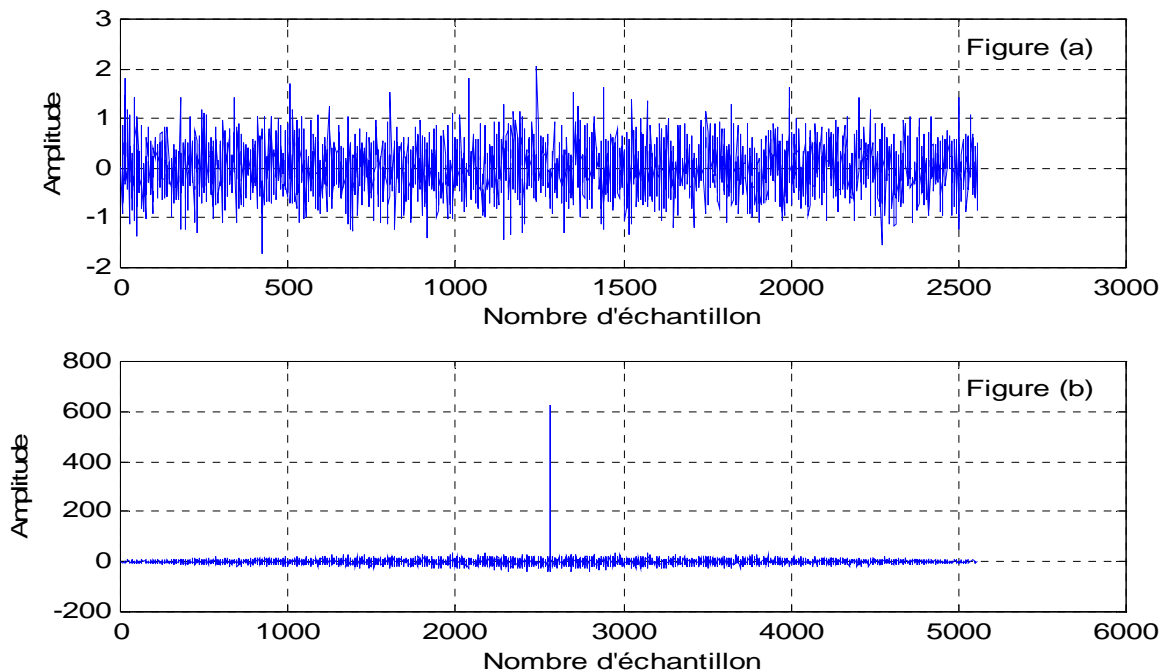


Fig. IV.15 : (a) - Erreur de prédiction ($N = 256 \cdot 10$, $\sigma^2 = 0.25$)

(b) - Autocorrelation de l'erreur de prédiction.

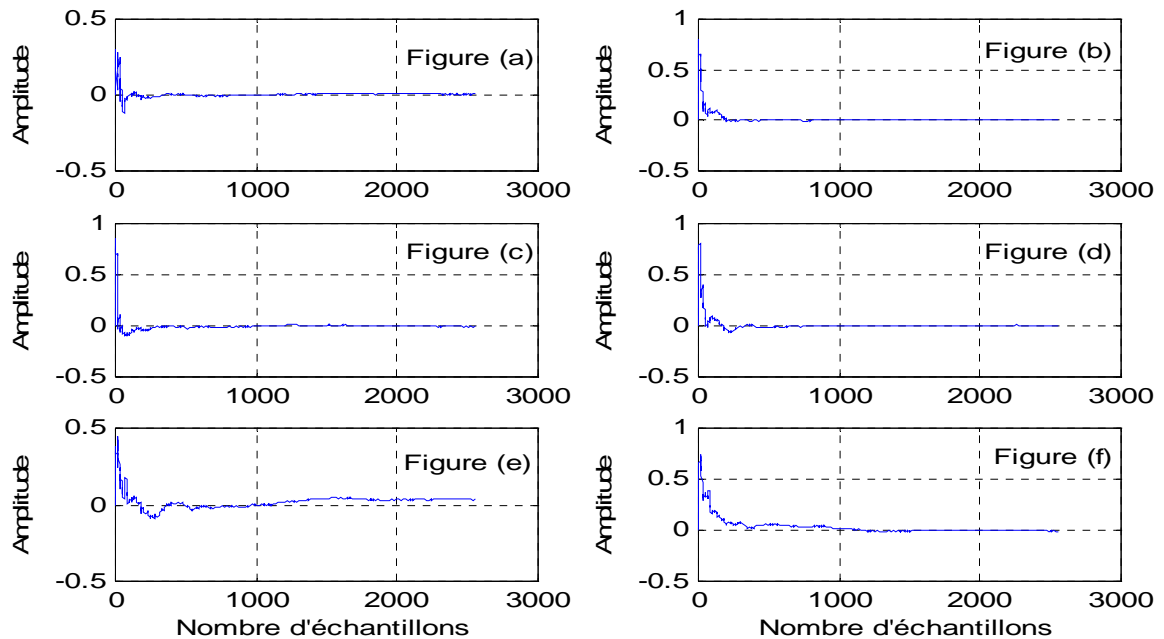


Fig. IV.16 : Erreur d'identification des paramètres a_i , b_i et c_i

IV.9. Interprétation des résultats obtenus

On remarque, d'après les résultats obtenus dont seule une partie été montrée que lorsque le nombre d'échantillons N augmente ou la variance σ^2 du bruit additif $v(t)$ diminue alors l'écart type de l'erreur de prédiction diminue, et les paramètres estimés sont plus proches aux valeurs réelles. Et la position des zéros du système n'influe pas sur l'erreur d'estimation des paramètres, par contre la position des pôles doit être située à l'intérieur du cercle unité pour assurer la stabilité du système, sinon l'algorithme diverge.

L'évolution temporelle des paramètres estimés par l'algorithme LMS montre l'effet du facteur d'adaptation μ sur la convergence. Une valeur plus grande de μ mène à une convergence plus rapide. Cependant, quand la valeur de μ est trop haute elle mène à l'instabilité de l'algorithme et mène à un résultat incorrect.

Nous venons de conclure que la vitesse de convergence est une propriété importante concernant l'estimation des paramètres du modèle par rapport à l'erreur d'identification. C'est pourquoi avant de conclure à la supériorité des algorithmes RLS il faut préciser quel critère de performance est utilisé. Il est clair que pour les problèmes transitoires (acquisition), c'est bien la vitesse de convergence, mais pour les problèmes permanents (poursuite).

IV.10. Conclusion

Dans ce chapitre nous avons implémenté sous MATLAB les algorithmes d'identification paramétrique basés sur le blanchissement de l'erreur de prédiction, qui sont : l'algorithme du gradient stochastique, Newton stochastique et l'algorithme des moindres carrés récursif.

Comme il est bien connu, il y a essentiellement deux familles d'algorithmes d'adaptation, les gradients, encore dénommés (LMS) et les moindres carrés récursifs (RLS). Les premiers sont appréciés pour leurs simplicités et performances. Puisque ce sont des algorithmes itératifs ils peuvent être appliqués à un système très variant dans le temps. Ils ont une exécution stable et robuste contre différents états de signal. Cependant ils peuvent ne pas avoir une vitesse vraiment rapide de convergence.

Les seconds sont connus comme ayant une vitesse de convergence supérieure. Pour cette dernière raison beaucoup d'auteurs ont pensé que les RLS sont capables d'une meilleure poursuite dans un contexte non stationnaire.

Chapitre

V

Application au Signal de Parole

V.1. Introduction

Le modèle AR (autorégressif) a prouvé son efficacité dans l'analyse spectrale du signal de parole. La raison en est que la parole est, dans la plupart des cas, produite par l'excitation des résonances acoustiques du conduit vocal, et que celles-ci correspondent bien à une fonction de transfert tout pôle. Mais la production de la parole est en réalité bien plus complexe. La source d'excitation glottale et le couplage source/conduit peuvent de toute évidence modifier la fonction de transfert du conduit vocal. Cette dernière peut elle-même présenter une structure complexe, notamment une ou plusieurs paires pôle zéro, dans le cas où plusieurs cavités sont couplées acoustiquement [61, 62]. C'est pourquoi on peut observer dans le spectre du signal de parole des creux marqués, des variations brusques ou des parties plates, qui sont difficiles à modéliser par un simple filtre tout pôle d'ordre limité [25].

Afin d'améliorer la modélisation du signal de parole, on a été amené à appliquer des algorithmes des moindres carrés récursifs RLS et RELS pour estimer les paramètres d'un modèle AR qui ont été développés [42, 44]. Ils permettent avantageusement de suivre les variations des paramètres.

Rappelons que ces algorithmes d'identification sont normalement conçus pour estimer les paramètres d'un système inconnu, ou modèle, soumis à un signal d'excitation connu. Or pour la parole, le signal d'entrée ne nous est pas accessible.

Nous voulons montrer dans ce travail de thèse, que lorsqu'on ne connaît pas le signal d'excitation, la précision de l'estimation des paramètres est très limitée. Il est clair que l'ajustement des paramètres du modèle dépend essentiellement du signal résiduel. Dans le cas où le signal d'entrée est inconnu, celui-ci se retrouve dans le signal résiduel et participe donc à l'erreur de prédiction. Cette présence du signal d'entrée inconnu dans le résidu perturbe le mécanisme d'ajustement, et diminue fortement la précision de l'estimation.

V.2. Base de données

Les expériences ont été effectuées sur la base de données TIMIT. Les deux locuteurs (Homme et Femme) prononcent chacun deux phrases phonétiquement équilibrées.

Les signaux de parole sont d'abord filtrés passe-bas à une fréquence 3.2 KHz, puis échantillonnés à 8 KHz qui sont représentés sur les figures V.1, 2

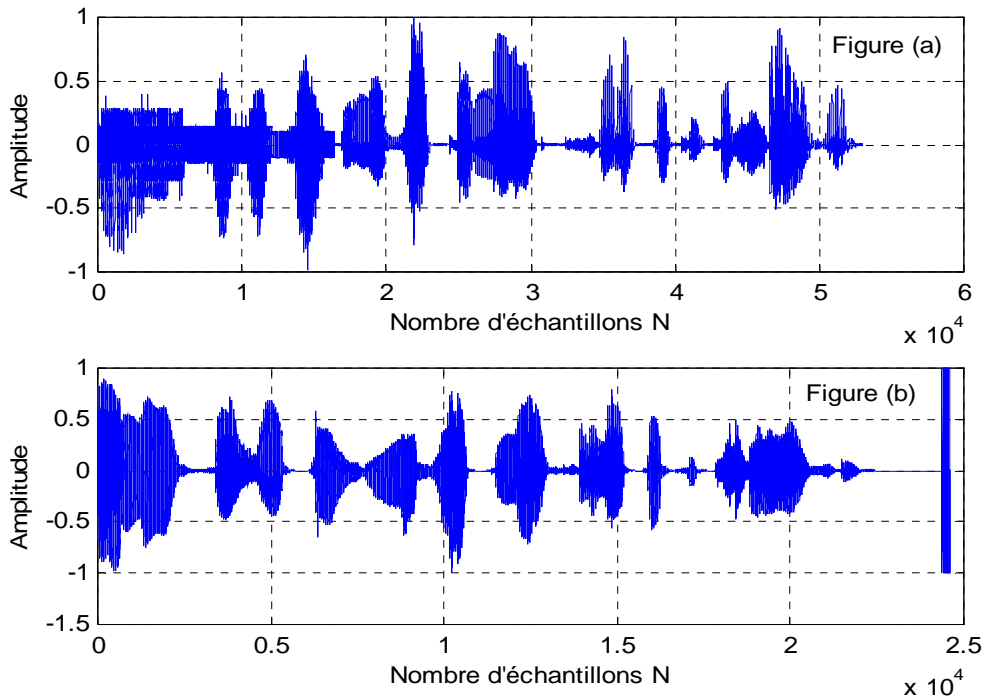


Fig.V.1 : Exemples de parole (Signal (x1): locuteur en haut, locutrice en bas) prononce la phrase: « *Un loup s'est jeté immédiatement sur la petite chèvre* ». Avec $x \in \{M, F\}$, or M et F désigne un locuteur masculin et féminin respectivement.

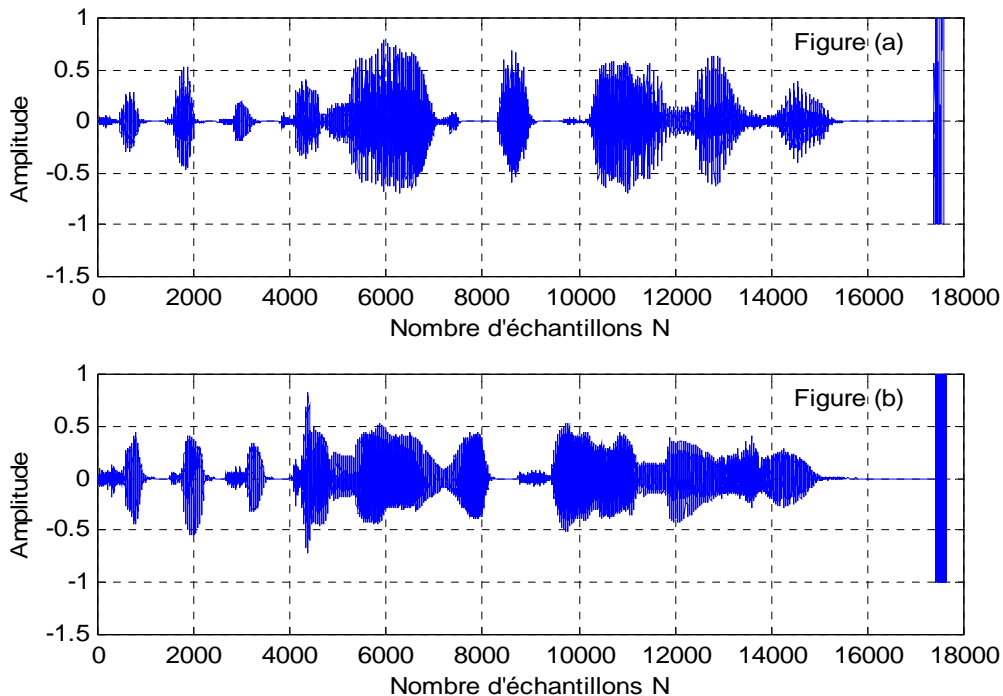


Fig.V.2 : Exemples de parole (Signal (x2): locuteur en haut, locutrice en bas) prononce la phrase: « *C'est petit canard apprend à nager* ». Avec $x \in \{M, F\}$, or M et F désigne un locuteur masculin et féminin respectivement.

La figure V.1 et V.2 représentent l'évolution temporelle, ou audiogramme, des signaux vocaux. On y constate une alternance de zones assez périodiques et de zones bruitées, appelées zones voisées et non voisées.

V.3. Exemples de son voisé et non voisé

La figure V.3 donne une représentation plus fine de tranches de signaux voisés et non voisés. L'évolution temporelle ne fournit cependant pas directement les traits acoustiques du signal. Il est nécessaire, pour les obtenir, de mener à bien un ensemble de transformations.

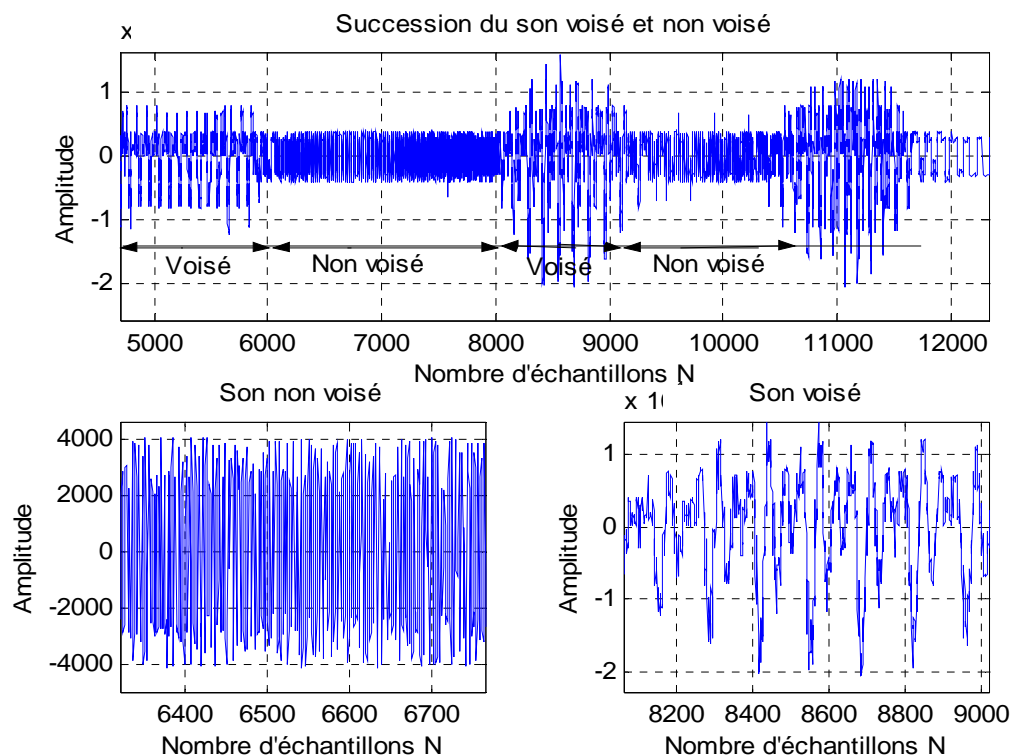


Fig.V.3 : Exemples de son voisé (droit) et non voisé (gauche).

V.4. Transformée de Fourier à court terme

La transformée de Fourier à court terme est obtenue en extrayant de l'audiogramme une 30aine de ms de signal vocal échantillonné à 16 KHz, en pondérant ces échantillons par une fenêtre de pondération et en effectuant une transformée de Fourier discrète sur une tranche voisée de la Fig.V.4. Les parties voisées du signal apparaissent sous la forme de successions de pics spectraux marqués, dont les fréquences centrales sont multiples de la fréquence fondamentale.

Par contre, le spectre d'un signal non voisé (Fig.V.5) ne présente aucune structure particulière.

La forme générale de ces spectres, appelée enveloppe spectrale, présente elle-même des pics et des creux qui correspondent aux résonances et aux anti-résonances du conduit vocal et sont appelés formants et anti-formants. L'évolution temporelle de leur fréquence centrale et de leur largeur de bande détermine le timbre du son.

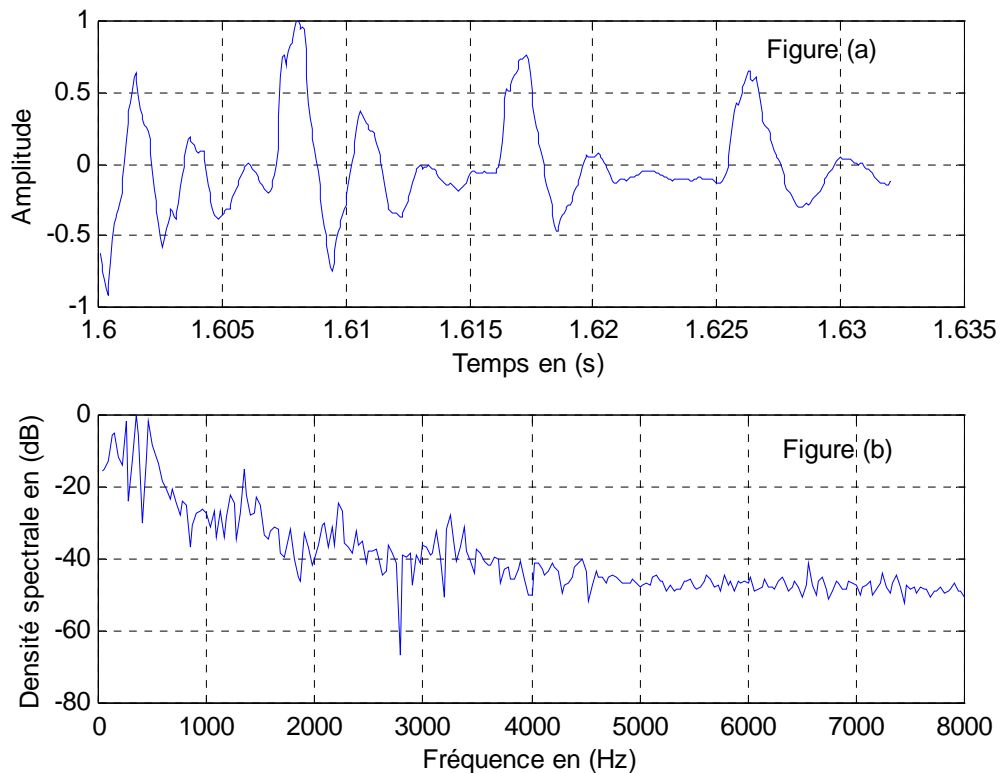


Fig.V.4 : (a) - Tranche voisée de 32 ms, $F_e = 16 \text{ KHz}$
(b) - Spectre du signal.

La Fig.V.4 montre que l'enveloppe spectrale des sons voisés est de type passe bas, avec environ un formant par kHz de bande passante, et dont seuls les trois ou quatre premiers contribuent de façon importante au timbre.

Par contre, les sons non voisés de la Fig.V.5 présentent souvent une accentuation vers les hautes fréquences.

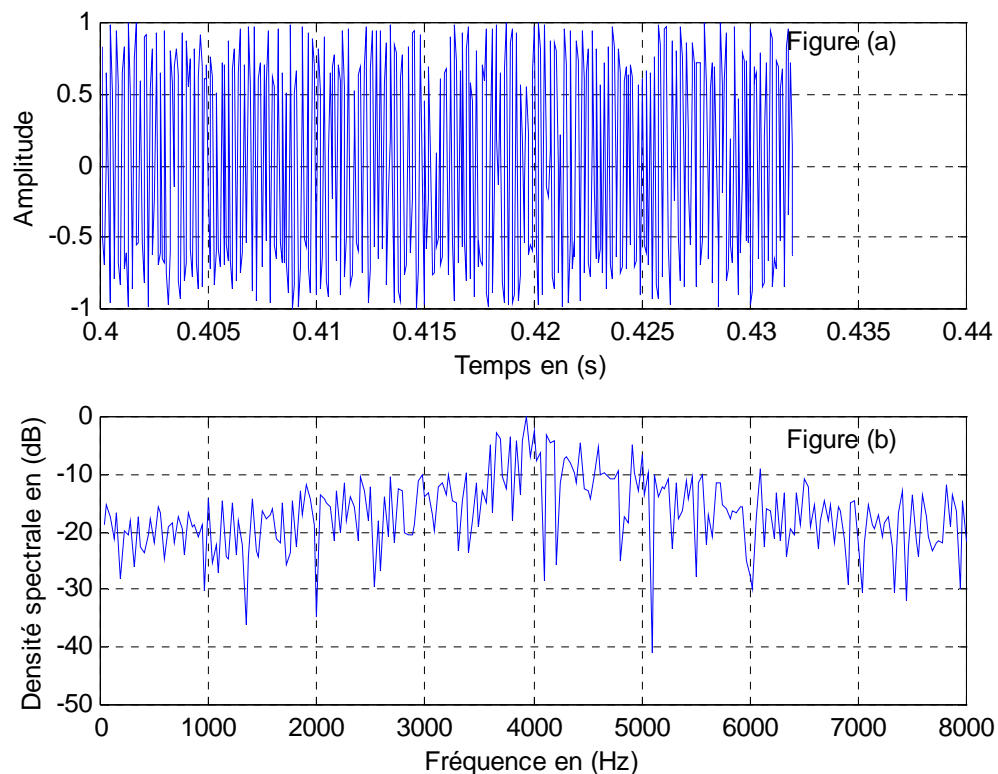


Fig.V.5 : (a) - Tranche non voisée de 32 ms, $F_e = 16$ KHz
(b) - Spectre du signal.

V.5. Modélisation de la parole non bruitée

Les procédés de traitement numérique du signal appliqués aux signaux de parole se basent souvent sur une modélisation paramétrique du signal de parole. Ce modèle comporte généralement quatre paramètres [63], [66] :

- Une information de « voisement », qui décrit le caractère plus ou moins périodique (sons voisés) ou aléatoire (sons non voisés) du signal.
- La fréquence fondamentale ou « PITCH », pour les sons voisés.
- L'évolution temporelle de « l'énergie » du signal de parole.
- Son enveloppe spectrale, ou « spectre », qui est généralement obtenue par une modélisation autorégressive (filtre de prédiction linéaire ou LPC) ou par une analyse de Fourier à court terme synchrone avec le pitch.

Ces quatre paramètres sont estimés périodiquement sur le signal de parole (de une à plusieurs fois par trame selon le paramètre, pour une durée trame typiquement comprise entre 10 et 30 ms).

Le problème de l'estimation d'un modèle AR correspondant au signal de parole dans un milieu non bruité revient à déterminer les coefficients d'un filtre tout pôle [67] de $H(z)$ d'ordre p , et on écrit:

$$H(z) = \frac{Y(z)}{U(z)} = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}}$$

Avec $Y(z)$, $H(z)$ et $U(z)$ les transformées en z du signal, du filtre du conduit vocal et de l'excitation.

Considérations pratiques

Pour mener à une bonne simulation et analyse, il faut bien choisir :

- La fréquence d'échantillonnage
- La méthode d'analyse
- L'ordre p du modèle AR
- Le nombre d'échantillons N par tranche

Le choix de la fréquence d'échantillonnage est fonction de l'application visée et de la qualité du signal à analyser. L'ordre d'analyse conditionne le nombre de formants que l'analyse est capable de prendre en compte. On estime en général que la parole présente un formant par kHz de bande passante, ce qui correspond à une paire de pôles pour $A(z)$. Si on y ajoute une paire de pôles pour la modélisation de l'excitation glottique, on obtient les valeurs classiques de $p = 10, 12,$ et 18 pour $F_c = 8, 10$ et 16 kHz respectivement. Elles trouvent d'ailleurs une justification expérimentale dans le fait que l'énergie de l'erreur de prédiction diminue rapidement lorsqu'on augmente p à partir de 1, pour tendre vers une asymptote autour de ces valeurs : il devient inutile d'encore augmenter l'ordre, puisqu'on ne prédit rien de plus. La durée des tranches d'analyse et leur décalage sont souvent fixés à 30 et 10 ms respectivement. Ces valeurs ont été choisies empiriquement; elles sont liées au caractère quasi-stationnaire du signal de parole.

Résultats de simulation

Pour bien analyser les performances de l'estimateur RLS, nous avons effectué des simulations sur des tranches de 32 ms pour les différents signaux de parole (M1, F1, M2, F2), les résultats obtenus sont illustrés dans les tableaux: Tab. V.1, 2, 3, 4 respectivement. L'ordre du modèle est choisi égale à 10.

La Fig.V.6 donne la représentation temporelle du signal M1, ainsi que son spectre de puissance. L'analyse est menée sur des tranches de 32 ms (256 échantillons), et les résultats de simulation obtenus par l'algorithme RLS sont illustrés dans le Tab.V.1.

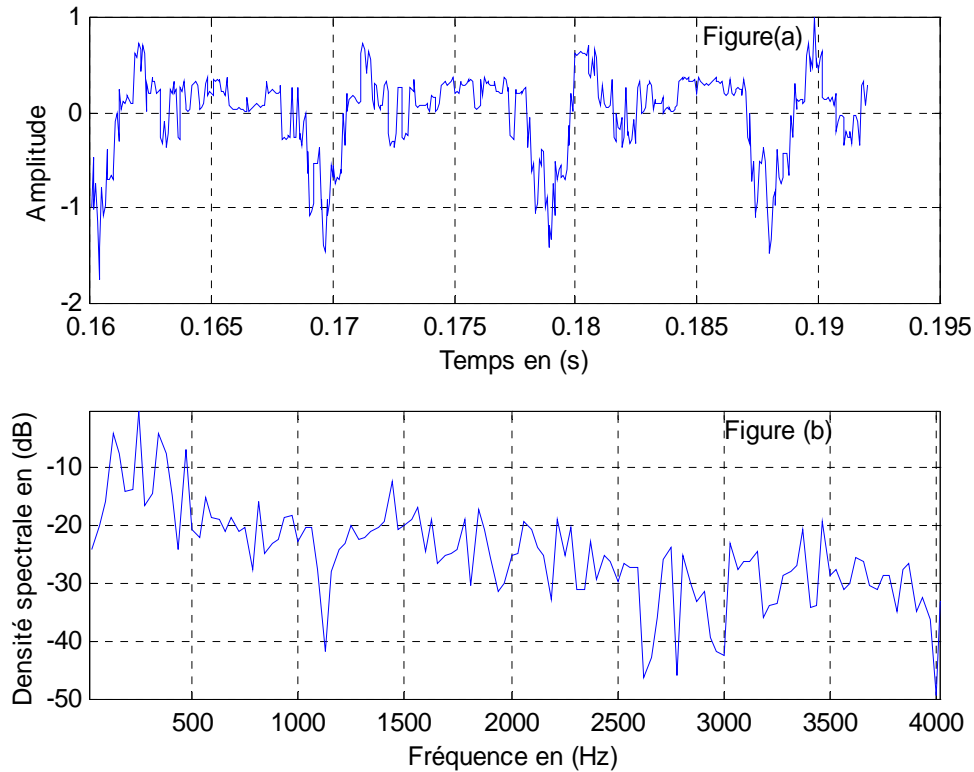


Fig.V.6 : (a) - Tranche de 32 ms du signal (M1)
(b) - Spectre du signal

Modèle AR (10)	Signal de Parole (M1)				
	MATLAB	RLS	BIAIS	VAR	EQM
a_0	1.0000	1.0000	0.0000	0.0000	0.0000
a_1	-0.6832	-0.6817	-0.0473	0.0108	0.0130
a_2	-0.2362	-0.2544	-0.0012	0.0097	0.0097
a_3	0.0552	0.0499	0.1054	0.0097	0.0208
a_4	0.0084	0.0122	-0.0297	0.0011	0.0020
a_5	0.0070	0.0065	-0.0241	0.0085	0.0090
a_6	-0.0024	-0.0094	0.0021	0.0048	0.0047
a_7	-0.0341	0.0077	-0.0366	0.0322	0.0335
a_8	-0.0457	-0.0751	0.0097	0.0023	0.0024
a_9	-0.1520	-0.1510	-0.0363	0.0109	0.0122
a_{10}	0.1922	0.1985	0.0328	0.0069	0.0079

Tab.V.1 : Propriétés de l'estimateur RLS. Cas: Tranche de 32 ms du signal (M1)

La modélisation par l'algorithme RLS d'un modèle AR d'ordre 10, réalisée sur le signal de parole F1 échantillonné à 8 KHz, a permis de calculer les 10 premiers coefficients du filtre tout pôles donnés par le tableau Tab.V.2.

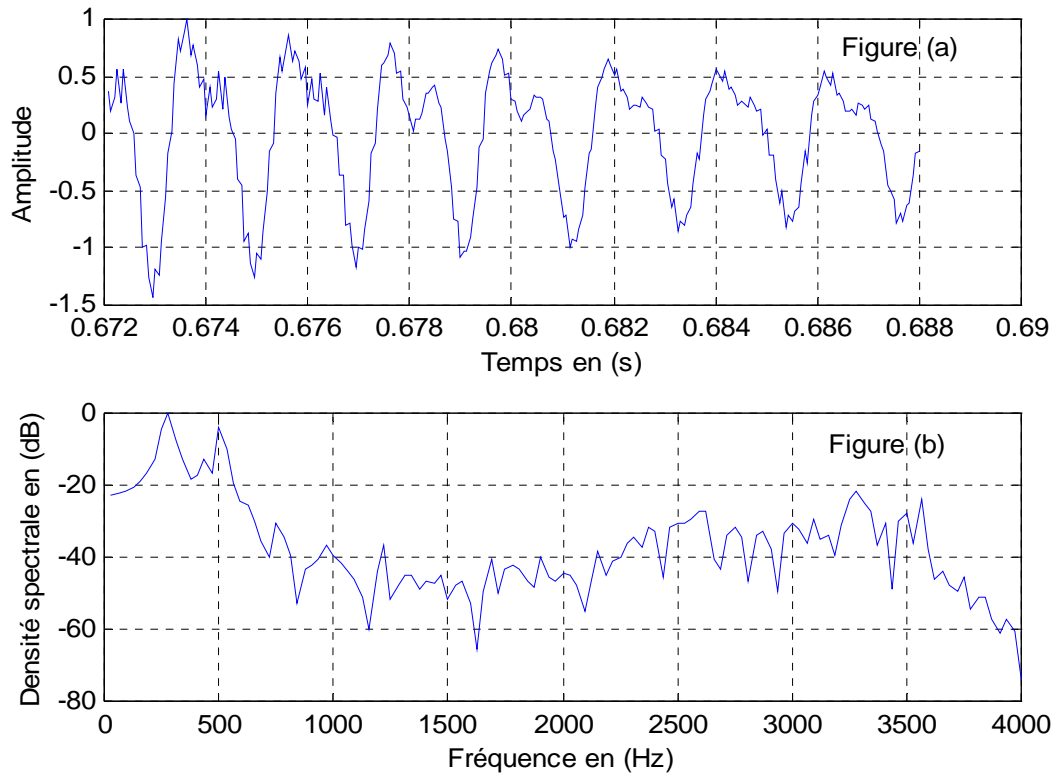


Fig.V.7 : (a) - Tranche de 32 ms du signal (F1)
(b) - Spectre du signal

Modèle AR (10)	Signal de Parole (F1)				
	MATLAB	RLS	BIAIS	VAR	EQM
a_0	1.0000	1.0000	0.0000	0.0000	0.0000
a_1	-0.5763	-0.5889	0.0035	0.0175	0.0174
a_2	-0.9770	-0.9407	-0.1514	0.0447	0.0674
a_3	-0.7179	-0.7039	-0.0374	0.0268	0.0280
a_4	1.1541	1.1353	0.0597	0.0827	0.0860
a_5	0.9373	0.8614	0.2713	0.0768	0.1501
a_6	0.2681	0.2898	-0.0567	0.0231	0.0262
a_7	-1.0967	-1.0403	-0.1709	0.1030	0.1318
a_8	-0.2428	-0.2193	-0.1158	0.0209	0.0342
a_9	0.0398	-0.0097	0.1011	0.0275	0.0376
a_{10}	0.3306	0.3376	0.0321	0.0127	0.0137

Tab.V.2 : Propriétés de l'estimateur RLS. Cas : Tranche de 32 ms du signal (F1)

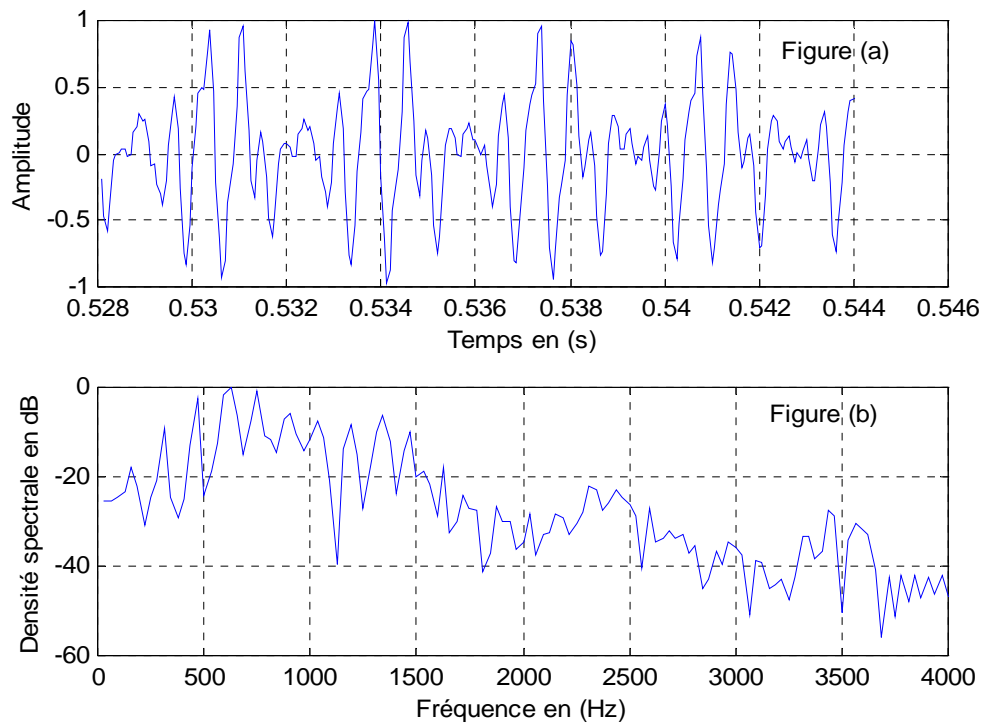


Fig.V.8 : (a) - Tranche de 32 ms du signal (M2)
(b) - Spectre du signal

Modèle AR (10)	Signal de Parole (M2)				
	MATLAB	RLS	BIAIS	VAR	EQM
a_0	1.0000	1.0000	0.0000	0.0000	0.0000
a_1	-1.7946	-1.7865	-0.0713	0.1636	0.1680
a_2	1.5466	1.5342	0.0752	0.2265	0.2312
a_3	-0.5707	-0.5615	-0.0540	0.1479	0.1502
a_4	0.2610	0.2223	0.1714	0.1018	0.1308
a_5	-0.0883	-0.0286	-0.1366	0.1275	0.1457
a_6	0.0597	0.0064	-0.0939	0.6631	0.6692
a_7	-0.2615	-0.2245	0.3022	1.8183	1.9023
a_8	1.0209	0.9750	-0.1620	1.7793	1.7983
a_9	-1.0749	-1.0326	-0.0103	0.9689	0.9650
a_{10}	0.5023	0.4802	0.0508	0.5484	0.5488

Tab.V.3 : Propriétés de l'estimateur RLS. Cas: Tranche de 32 ms du signal (M2)

Le Tab.V.3 montre bien les propriétés de l'estimateur RLS de point vue Biais, Variance et EQM. Les résultats de simulation obtenus par l'algorithme RLS sont plus proches aux paramètres du signal de parole (M2) obtenus par la fonction du MATLAB.

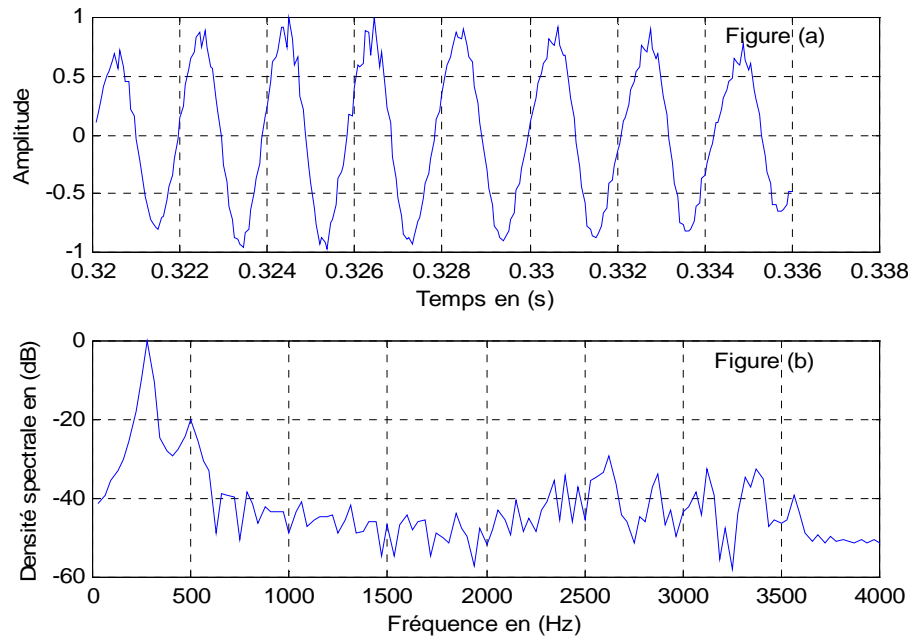


Fig.V.9 : (a) - Tranche de 32 ms du signal (F2)
(b) - Spectre du signal

Modèle AR (10)	Signal de Parole (F2)				
	MATLAB	RLS	BIAIS	VAR	EQM
a_0	1.0000	1.0000	0.0000	0.0000	0.0000
a_1	-0.6391	-0.6875	0.0854	0.0831	0.0901
a_2	-0.4405	-0.4241	-0.0668	0.0250	0.0293
a_3	-1.1401	-1.0705	-0.3397	0.3425	0.4565
a_4	0.9734	0.9889	0.1263	0.1641	0.1794
a_5	0.5053	0.4415	0.2709	0.0462	0.1194
a_6	0.7659	0.7070	0.2326	0.1594	0.2128
a_7	-0.9283	-0.8996	-0.2823	0.1538	0.2329
a_8	-0.0470	0.0195	-0.2326	0.1625	0.2160
a_9	-0.3334	-0.3326	0.0987	0.8721	0.8783
a_{10}	0.4232	0.3859	0.0730	0.4746	0.4780

Tab.V.4 : Propriétés de l'estimateur RLS. Cas: Tranche de 32 ms du signal (F2)

L'estimateur RLS appliqué au signal de parole démontre une très grande variabilité. Il est en effet peu probable de mesurer deux signaux de parole totalement semblables même si des phrases identiques sont prononcées par le même locuteur. Cette variabilité rend le processus de traitement automatique de parole excessivement complexe. Pour se faire, une analyse à court terme est réalisée. Les tableaux précédents (Tab.V.1, 2, 3, 4) expliquent pourquoi un estimateur sans biais n'est pas toujours systématiquement préférable à un

estimateur biaisé, mais de faible variance. Les figures suivantes: Fig.V.10 et Fig.V.11 sont une démonstration du résultat relatif à l'EQM de l'estimateur RLS des paramètres a_1 , a_2 , a_4 et a_5 obtenus a partir du tableau Tab.V.1.

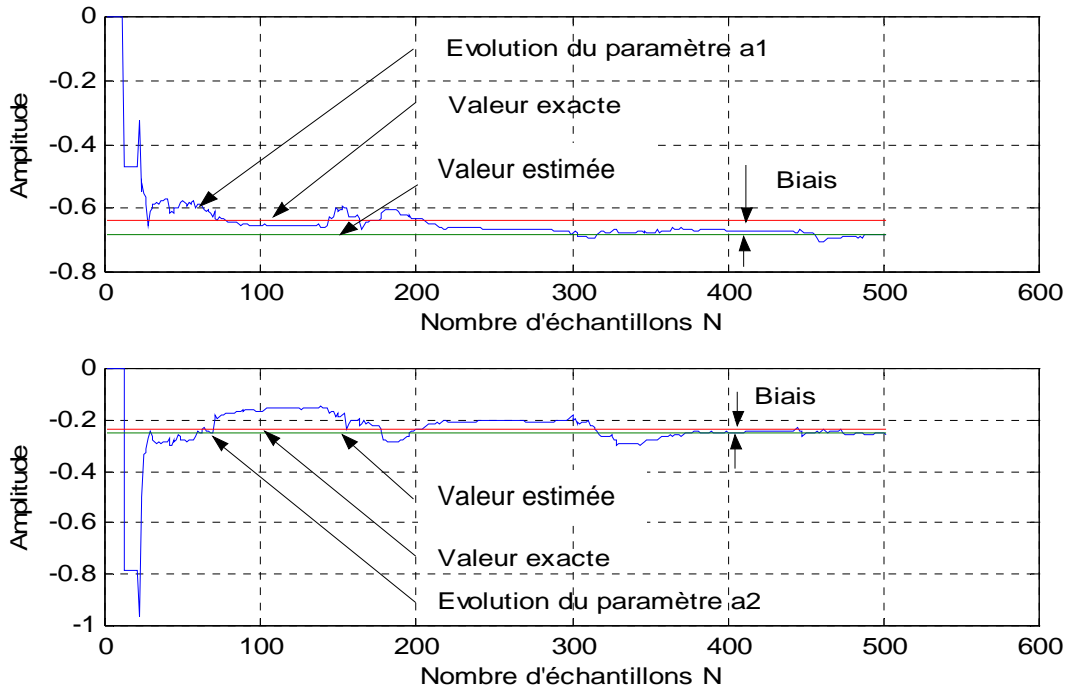


Fig.V.10 : Evolution temporelle du paramètre a_1 et a_2 du signal (M1).

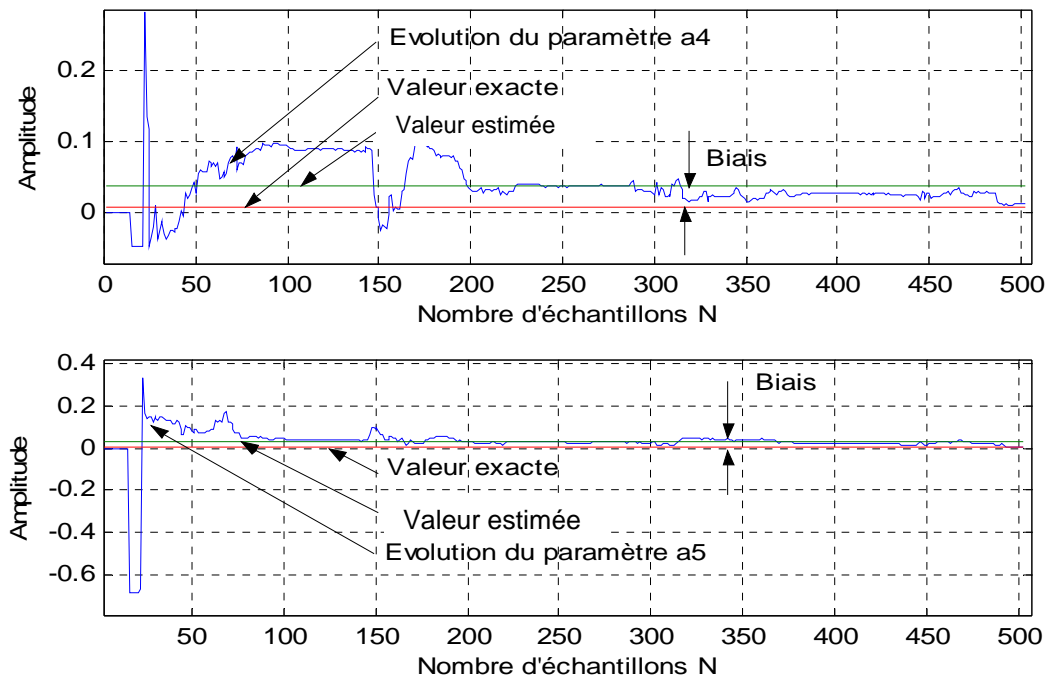


Fig.V.11 : Evolution temporelle du paramètre a_4 et a_5 du signal (M1).

V.6. Modélisation de la parole bruitée

La parole bruitée fait l'objet de plusieurs travaux en télécommunications. Des méthodes pour réduire le bruit et pour annuler l'écho ont été proposées dans plusieurs références. En général, tous les algorithmes appliqués sont basés sur des filtres à coefficients adaptatifs qui fonctionnent assez bien.

Une grande partie des méthodes existantes aborde le problème de la réduction de bruit sous son aspect fréquentiel afin de profiter des caractéristiques de ce domaine d'étude. Parallèlement aux nombreuses variantes de la soustraction spectrale, les méthodes s'inspirant d'un filtrage optimal se sont développées depuis les travaux réalisés par Wiener. Celles-ci prennent en compte souvent la statistique du signal de parole. Ces méthodes appliquent des modifications sur la transformée de Fourier court terme du signal bruité, en utilisant les informations *a priori* disponibles sur le bruit.

Nous proposons donc d'étudier, la méthode des moindres carrés étendus RELS qui permet de mieux estimer les paramètres du modèle autorégressif correspondant au signal de parole dans un environnement bruité. Le bruit additif Gaussien $v(t)$ utilisé le long de ce travail est illustré sur la Fig.V.12.

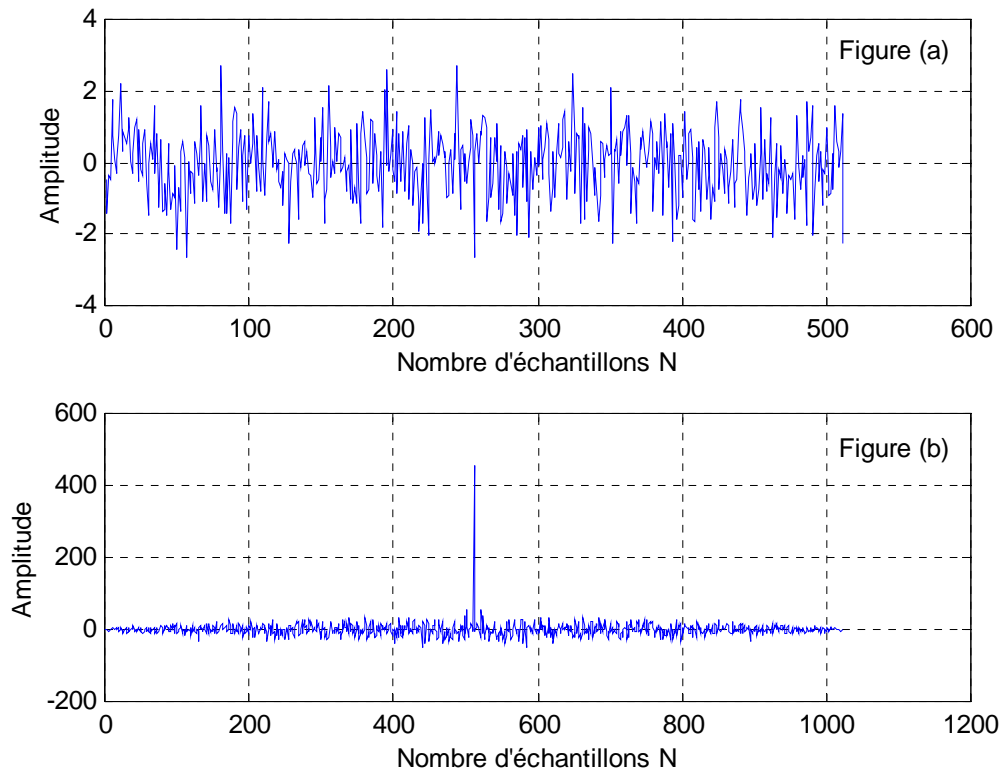


Fig.V.12 : Bruit Gaussien de moyenne nulle et variance égale à l'unité

Les tableaux IV.5, 6, 7,...,14 présentent les propriétés de l'estimateur RELS (Biais, Variance et Ecart Quadratique Moyenne) sur 4 phrases du corpus (2 hommes et 2 femmes) pour des valeurs du SNR allant de 40 dB à 10 dB.

Modèle AR (10)	Signal de Parole (M1) faiblement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.6979	-0.0350	0.0171	0.0183
a ₂	-0.2525	0.0119	0.0469	0.0470
a ₃	0.0594	0.0947	0.0093	0.0182
a ₄	0.0182	-0.0421	0.0068	0.0086
a ₅	0.0029	-0.0257	0.0157	0.0164
a ₆	-0.0089	-0.0060	0.0110	0.0110
a ₇	0.0116	-0.0453	0.0367	0.0387
a ₈	-0.0800	0.0102	0.0308	0.0308
a ₉	-0.1572	-0.0222	0.0238	0.0243
a ₁₀	0.2006	0.0568	0.1653	0.1682

Tab.V.5 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M1), SNR = 40 dB.

Modèle AR (10)	Signal de Parole (M1) moyennement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.6898	-0.0443	0.0168	0.0187
a ₂	-0.2506	0.0148	0.0462	0.0463
a ₃	0.0526	0.0933	0.0084	0.0171
a ₄	0.0061	-0.0253	0.0061	0.0068
a ₅	0.0136	-0.0349	0.0156	0.0168
a ₆	-0.0076	-0.0041	0.0113	0.0113
a ₇	0.0102	-0.0485	0.0367	0.0389
a ₈	-0.0790	0.0020	0.0325	0.0325
a ₉	-0.1540	-0.0140	0.0218	0.0219
a ₁₀	0.1955	0.0565	0.1666	0.1694

Tab.V.6 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M1), SNR = 30 dB.

Pour un $SNR \geq 30$ dB, l'estimateur RELS donne des résultats meilleurs, et l'EQM varie légèrement dans un contexte bruité.

Modèle AR (10)	Signal de Parole (M1) moyennement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.6642	-0.0782	0.0155	0.0216
a ₂	-0.2666	0.0294	0.0380	0.0388
a ₃	0.0302	0.1365	0.0118	0.0304
a ₄	0.0078	-0.0408	0.0044	0.0061
a ₅	0.0485	-0.0468	0.0128	0.0149
a ₆	-0.0133	-0.0210	0.0084	0.0088
a ₇	-0.0281	0.0010	0.0328	0.0327
a ₈	-0.0381	-0.0422	0.0235	0.0253
a ₉	-0.1769	-0.0021	0.0132	0.0132
a ₁₀	0.1991	0.0533	0.0950	0.0977

Tab.V.7 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M1), SNR = 20 dB.

Modèle AR (10)	Signal de Parole (M1) Fortement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.5621	-0.1747	0.0279	0.0583
a ₂	-0.2085	-0.0369	0.0553	0.0565
a ₃	-0.0460	0.1359	0.0083	0.0268
a ₄	-0.0332	0.0631	0.0046	0.0086
a ₅	-0.0417	0.0967	0.0119	0.0212
a ₆	0.0897	-0.1462	0.0106	0.0320
a ₇	0.0325	-0.0343	0.0103	0.0114
a ₈	-0.1096	0.0288	0.0494	0.0502
a ₉	-0.1035	-0.1197	0.0252	0.0395
a ₁₀	0.1101	0.1361	0.1322	0.1504

Tab.V.8 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M1), SNR = 10 dB.

Le Tab.V.8 démontre la dégradation totale et rapide des performances de l'estimateur RELS dans un contexte fortement bruité.

Modèle AR (10)	Signal de Parole (F1) faiblement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.5851	0.0319	0.2441	0.2441
a ₂	-0.9681	-0.1240	0.1963	0.2109
a ₃	-0.6938	-0.0586	0.0449	0.0481
a ₄	1.1475	0.0278	0.2162	0.2161
a ₅	0.8823	0.2196	0.4953	0.5415
a ₆	0.2692	-0.0252	0.0226	0.0231
a ₇	-1.0468	-0.1476	0.1968	0.2178
a ₈	-0.2263	-0.0629	0.5993	0.6009
a ₉	0.0010	0.0863	0.0380	0.0453
a ₁₀	0.3369	0.0604	0.1193	0.1225

Tab.V.9 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (F1), SNR = 40 dB.

Modèle AR (10)	Signal de Parole (F1) fortement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.5306	-0.1366	0.0890	0.1073
a ₂	-0.5511	-0.3568	0.0660	0.1930
a ₃	-0.1012	-0.5067	0.2853	0.5409
a ₄	0.1171	0.9392	0.1840	1.0654
a ₅	0.1212	0.8506	0.1377	0.8608
a ₆	0.2073	-0.1138	0.2872	0.2990
a ₇	-0.0371	-1.1441	0.2696	1.5775
a ₈	0.0729	-0.1625	0.5810	0.6051
a ₉	0.0102	0.1643	0.8536	0.8771
a ₁₀	-0.0933	0.2588	2.9638	3.0187

Tab.V.10 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (F1), SNR = 10 dB.

Le Tab.V.10 présente certains des paramètres du signal de parole (F1) qui sont masqués par le bruit et dans un tel cas les paramètres du signal de parole non bruitée ne peuvent être estimés à partir du signal bruité et sont donc considérés comme manquants ou incertains.

La même chose se présente aux signaux de parole (M2 et F2). Les tableaux V.12 et V.14 démontrent la dégradation totale et rapide des performances de l'estimateur RELS dans un contexte fortement bruité.

Modèle AR (10)	Signal de Parole (M2) faiblement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-1.7938	-0.0459	0.1782	0.1795
a ₂	1.5763	-0.0475	0.3340	0.3349
a ₃	-0.6255	0.1508	0.5667	0.5871
a ₄	0.2778	-0.0462	1.2845	1.2815
a ₅	-0.0559	-0.0764	0.7204	0.7233
a ₆	0.0154	0.0024	0.3787	0.3771
a ₇	-0.2147	0.0626	0.1631	0.1663
a ₈	0.9577	0.1014	0.9855	0.9917
a ₉	-1.0191	-0.1916	1.2683	1.2999
a ₁₀	0.4791	0.1237	0.3817	0.3954

Tab.V.11 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M2), SNR = 40 dB.

Modèle AR (10)	Signal de Parole (M2) fortement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.8883	-0.8309	0.2423	0.9317
a ₂	0.2015	1.2480	0.5754	2.1305
a ₃	0.1941	-0.7843	0.5256	1.1386
a ₄	0.1260	0.3893	1.0773	1.2244
a ₅	-0.0094	-0.2377	0.6948	0.7485
a ₆	0.0554	-0.0067	0.8087	0.8055
a ₇	-0.1342	0.0533	1.0971	1.0955
a ₈	0.3532	0.4774	6.0642	6.2675
a ₉	-0.0402	-0.8013	9.6257	10.2287
a ₁₀	-0.0547	0.4101	4.6694	4.8187

Tab.V.12 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (M2), SNR = 10 dB.

Modèle AR (10)	Signal de Parole (F2) faiblement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0.0000	0.0000	0.0000
a ₁	-0.6878	0.0469	0.0405	0.0425
a ₂	-0.4519	0.0050	0.0262	0.0261
a ₃	-1.0277	-0.3590	0.2361	0.3640
a ₄	0.9730	0.1248	0.1151	0.1303
a ₅	0.4679	0.1654	0.0420	0.0692
a ₆	0.6575	0.3396	0.2715	0.3857
a ₇	-0.8814	-0.2813	0.2827	0.3607
a ₈	0.0126	-0.2144	0.1673	0.2126
a ₉	-0.2997	0.0333	0.6212	0.6198
a ₁₀	0.3624	0.1217	0.1622	0.1764

Tab.V.13 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (F2), SNR = 40 dB.

Modèle AR (10)	Signal de Parole (F2) fortement bruité			
	RELS	BIAIS	VAR	EQM
a ₀	1.0000	0	0	0
a ₁	-0.3089	-0.4452	0.0716	0.2696
a ₂	-0.3884	0.0621	0.2451	0.2480
a ₃	-0.1194	-0.6899	0.8593	1.3318
a ₄	-0.1524	1.0674	0.3259	1.4640
a ₅	-0.0871	0.3612	1.9091	2.0318
a ₆	0.0214	0.6341	0.8296	1.2283
a ₇	0.1514	-1.0830	0.5344	1.7050
a ₈	-0.0009	-0.0006	2.4739	2.4639
a ₉	0.1157	-0.5769	1.1280	1.4562
a ₁₀	0.1705	0.4819	2.2276	2.4508

Tab.V.14 : Estimation des paramètres du modèle AR par l'algorithme RELS.
Cas: Tranche de 32 ms du signal (F2), SNR = 10 dB.

L'estimateur RELS n'utilise aucune information a priori sur le bruit de mesure, et nous avons montré que si le signal de parole est fortement bruité, l'estimation des paramètres est biaisée. La qualité de l'estimateur se dégrade rapidement lorsque le SNR diminue, et limite de ce fait la qualité des traitements en présence d'un niveau de bruit élevé.

V.7. Etude comparative

Les tableaux ci-dessous se résument l'EQM de l'estimateur RELS appliqué aux signaux de parole (M1, F1, M2 et F2) pour des SNR = 40 dB puis 10 dB.

Modèle AR (10)	Estimateur RELS SNR = 40 dB			
	EQM (M1)	EQM (F1)	EQM (M2)	EQM (F2)
a_0	0.0000	0.0000	0.0000	0.0000
a_1	0.1088	0.4062	0.1381	0.0686
a_2	0.2794	0.3510	0.2576	0.0421
a_3	0.1082	0.0800	0.4517	0.5873
a_4	0.0511	0.3596	0.9858	0.2102
a_5	0.0975	0.9011	0.5564	0.1116
a_6	0.0654	0.0384	0.2901	0.6223
a_7	0.2301	0.3625	0.1279	0.5820
a_8	0.1831	1.0000	0.7629	0.3430
a_9	0.1445	0.0754	1.0000	1.0000
a_{10}	1.0000	0.2039	0.3042	0.2846

Tab.V.15 : EQM des signaux faiblement bruités.

Modèle AR (10)	Estimateur RELS SNR = 10 dB			
	EQM (M1)	EQM (F1)	EQM (M2)	EQM (F2)
a_0	0.0000	0.0000	0.0000	0.0000
a_1	0.3876	0.0355	0.0911	0.1094
a_2	0.3757	0.0639	0.2083	0.1007
a_3	0.1782	0.1792	0.1113	0.5405
a_4	0.0572	0.3529	0.1197	0.5942
a_5	0.1410	0.2852	0.0732	0.8246
a_6	0.2128	0.0990	0.0787	0.4985
a_7	0.0758	0.5226	0.1071	0.6920
a_8	0.3338	0.2005	0.6127	1.0000
a_9	0.2626	0.2906	1.0000	0.5910
a_{10}	1.0000	1.0000	0.4711	0.9947

Tab.V.16 : EQM des signaux fortement bruités.

Une étude comparative dans la représentation des taux des EQM sur la Fig.IV.13 du signal bruité, montre que certains des paramètres du signal de parole sont masqués par le bruit et les paramètres du signal de parole non bruitée ne peuvent être estimés à partir du signal bruité (SNR d'ordre 10 dB). Cependant, le débruitage par filtre de Wiener adaptatif donne généralement des meilleurs résultats [28], [29].

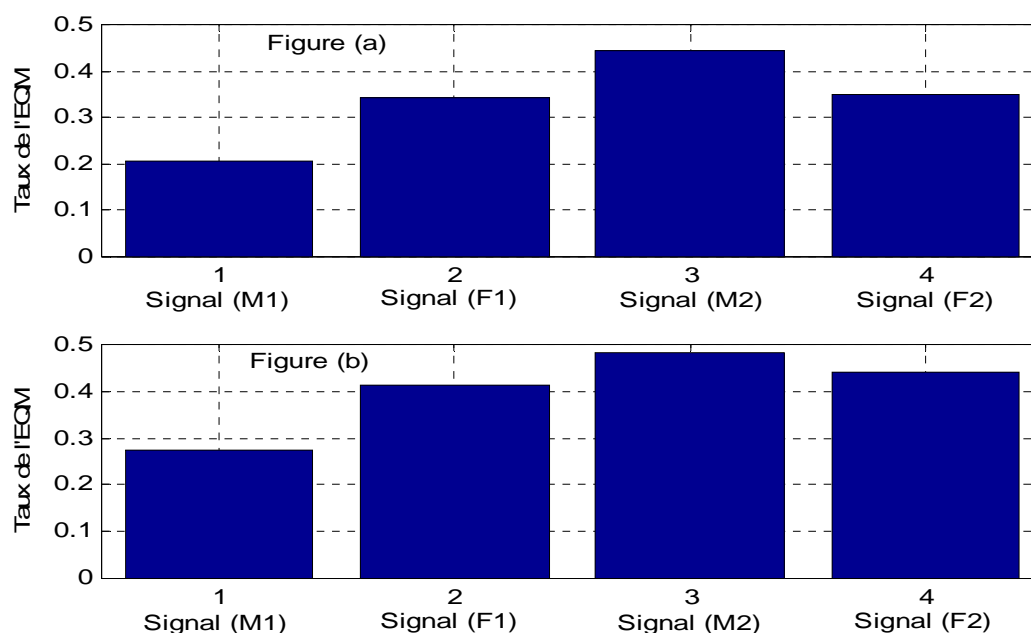


Fig.V.13 : (a) Taux des EQM pour des signaux faiblement bruités
(b) Taux des EQM pour des signaux fortement bruités.

V.8. Modélisation ARMA de la parole non bruitée

Dans l'algorithme RLS, la mise à jour pas à pas des paramètres du modèle permet de suivre les paramètres réels du système. Le modèle ARMA possède des zéros et des pôles, ces derniers ne sont pas contraints de se placer à l'intérieur du cercle unité. Nous pouvons espérer que le résidu sera plus proche d'un train d'impulsions et que la méthode décrite précédemment pourra être appliquée.

Un premier test a été réalisé sur un signal de parole (M1) incorporant l'influence de la source d'excitation et du couplage source/conduit. Nous avons utilisé une forme d'excitation typique de la source glottale.

Pour simuler le couplage, nous avons fait varier les bandes passantes des résonances proportionnellement à l'ouverture de la glotte [63]. En ce qui concerne le choix de l'ordre du modèle, il est évident que ce dernier doit être supérieur à l'ordre de la partie du système lié au conduit vocal, car le modèle doit pouvoir représenter également la source glottale. Dans notre travail, l'ordre de la partie AR a été fixé à 10, et l'ordre de la partie MA à 6.

Nous présentons les résultats obtenus par l'algorithme RLS sur le tableau ci-dessous.

Modèle ARMA (10,6)		Signal de Parole (M1)		
		Fonction MATLAB	Algorithme RLS	
			Partie AR	Partie MA
a ₀	b ₀	1.0000	1.0000	0.0000
a ₁	b ₁	-0.6832	-0.6841	-9.3885
a ₂	b ₂	-0.2362	-0.2402	-151.4724
a ₃	b ₃	0.0552	0.0558	3.8186
a ₄	b ₄	0.0084	0.0115	32.1925
a ₅	b ₅	0.0070	0.0138	26.5402
a ₆	b ₆	-0.0024	-0.0109	135.7180
a ₇	b ₇	-0.0341	-0.0397	-----
a ₈	b ₈	-0.0457	-0.0434	-----
a ₉	b ₉	-0.1520	-0.1473	-----
a ₁₀	b ₁₀	0.1922	0.1906	-----

Tab.V.17: Modélisation ARMA de la parole par l'estimateur RLS.
Cas: Tranche de 32 ms du signal (M1)

Dans le Tab.V.17, on tient compte du couplage avec la cavité nasale responsable d'une autre classe de sons appelée: sons nasals. Le couplage entre l'excitation glottale et le conduit vocal rend difficile la mesure précise du fondamental.

Notons aussi que le modèle ARMA est souvent retenu pour modéliser la parole. Plusieurs méthodes d'estimations des modèles ARMA sont décrites dans les références [16] ou [67] : méthode de corrélation (algorithmes de Levinson ou de Schur), de covariance (algorithme de Cholesky), de Burg.

V.9. Conclusion

L'identification des paramètres d'un signal de parole non bruitée par l'algorithme RLS donne des résultats meilleurs mais ces performances se dégradent rapidement en présence du bruit. De manière générale, Pour un SNR inférieur à 20 dB, l'ensemble des signaux de parole échantillonnés à 8 KHz apporte une dégradation totale par rapport aux signaux de parole échantillonnés à 16 KHz. De plus, on observe une variation brusque du biais de l'estimateur RELS.

L'application au signal de parole démontre une très grande variabilité. Pour se faire, une analyse à court terme est réalisée. Le signal de parole est observé au travers d'une fenêtre d'analyse d'une durée d'environ de 30 millisecondes. Pour chaque position de la fenêtre d'analyse, une estimation des paramètres, c'est-à-dire les p coefficients LP, du signal de parole observé sont calculés.

En résumé, l'analyse de la parole est une étape indispensable à toute application de synthèse, de codage, ou de reconnaissance. Elle repose en général sur un modèle. Celui-ci possède un ensemble de paramètres numériques, dont les plages de variation définissent l'ensemble des signaux couverts par le modèle. Pour un signal et un modèle donné, l'analyse consiste en l'estimation des paramètres du modèle dans le but de lui faire correspondre le signal analysé.

Les erreurs de modélisation peuvent être causées par le modèle lui-même, puisqu'il restreint implicitement l'ensemble des signaux couverts, de sorte qu'un signal qui n'en fait pas partie ne pourra jamais être modélisé correctement (quelque soit la valeur choisie pour les paramètres). On parlera alors d'erreur intrinsèque. D'autre part, un signal couvert par le modèle ne sera effectivement bien modélisé que dans la mesure où l'algorithme d'analyse est effectivement capable de trouver les paramètres qui annulent l'erreur de modélisation. Si ce n'est pas le cas, on parlera d'erreur extrinsèque, ou biais d'analyse.

Les modèles de production ne cherchent à reproduire que le schéma de principe du mécanisme phonatoire, par le biais de son équivalent électrique. On y décrit la parole comme le signal produit par un assemblage de générateurs et de filtres numériques. Les paramètres de

ces modèles sont ceux des générateurs et filtres qui les constituent. Le modèle Auto - Régressif (AR), que nous avons étudié et simulé plus spécialement dans ce chapitre en raison de l'utilisation répandue en traitement de la parole, en est l'exemple le plus simple.

Un autre modèle connu sous autorégressif à moyenne ajustée ARMA est parfois utilisé. Il se distingue par la présence à la fois des pôles et des zéros. Ce dernier est plus délicat à estimer qu'un modèle AR. Cela amène parfois à préférer, pour une qualité donnée de la modélisation, un modèle AR avec un ordre un peu surestimé. Mais la principale limitation réside dans l'hypothèse de stationnarité du signal acoustique qui est faite [66]. Il existe un compromis entre la longueur de la fenêtre d'analyse et la durée pendant laquelle l'hypothèse de stationnarité est raisonnable. Ce compromis est réalisable pendant les zones stables (voyelles), mais il n'est pas satisfaisant durant les phases transitoires et non justifié sur les plosives.

Conclusion Générale et Perspectives

Dans ce travail de thèse, un panorama des méthodes récursives, d'identification de modèles paramétriques des signaux et systèmes, a été utilisé. L'accent a été mis sur la présentation de la structure des différentes méthodes, leur raison d'être, les conditions de convergence dans un environnement stochastique et sur leur domaine d'application. La structure des algorithmes d'adaptation paramétrique commune à toutes les méthodes d'identification a été présentée en détail indiquant l'influence du choix des différents paramètres sur la précision de l'estimation et la capacité de suivi de variation des modèles. L'aspect validation du modèle identifié qui est spécifique aux méthodes paramétriques de modélisation du signal a été décrit. L'interprétation des critères de validation des modèles paramétriques a été donnée, permettant ainsi d'établir un lien entre les méthodes d'identification paramétriques.

Les algorithmes d'identification paramétrique développés sont :

- Gradient stochastique (LMS)
- Newton stochastique (NS)
- Moindres carrés récursifs (RLS)
- Moindres carrés étendus (RELS)

L'implémentation des algorithmes LMS, NS et RLS sur PC, en utilisant le langage de programmation MATLAB, les résultats obtenus nous a montré que lorsque le nombre d'échantillons N augmente ou la variance du bruit additif diminue alors le biais de l'estimateur diminue, et les paramètres estimés sont plus proches à des valeurs réelles. Et la position des zéros du système n'influe pas sur l'erreur d'estimation des paramètres, par contre la position des pôles doit être située à l'intérieur du cercle unité pour assurer la stabilité du système, sinon l'algorithme diverge.

La méthode des moindres carrés étendus (RELS) nous a permis d'identifier sans biais des modèles (système + perturbation). L'idée est d'identifier simultanément le modèle du système et le modèle de la perturbation pour obtenir une erreur de prédiction asymptotiquement blanche, en présence d'une perturbation aléatoire correspondant au modèle

ARMAX, l'erreur de prédiction $\varepsilon(t)$ tend asymptotiquement vers un bruit blanc, ce qui garantit une estimation non biaisée des paramètres du système.

Notre contribution a porté sur l'utilisation de la méthode des moindres carrés étendue (RELS) au signal de parole en présence du bruit coloré.

Les résultats obtenus par l'algorithme RLS sont meilleurs, mais ses performances se dégradent rapidement en présence du bruit. Tandis que dans le contexte de la parole bruitée l'application de l'algorithme RELS donne de bons résultats jusqu'à un SNR de 20 dB, au dessous cette limite, les résultats sont moins satisfaisants. Plus précisément, la difficulté se manifeste au niveau des signaux échantillonnés à la fréquence de 8 KHz. La raison de ce résultat, est due au nombre d'échantillons N qui est insuffisant à la convergence de l'algorithme et l'énergie du signal se trouve complètement noyée ou écrasée dans le bruit, ce qui a rendu l'estimateur RELS biaisé dans ce cas. Dans la représentation temps fréquence du signal bruité, certains des paramètres du modèle de la parole sont masqués par le bruit, et les paramètres du signal non bruité ne peuvent être estimés à partir du signal bruité et sont donc considérés comme manquants ou incertains.

Notons aussi que le modèle ARMA est souvent retenu pour modéliser la parole, il est plus délicat à estimer qu'un modèle AR. Pour remédier à ce problème, une approche originale connue par le *cepstre* a été proposée par Noll (1964) pour extraire la fréquence fondamentale. Oppeheim (1968) l'a exploitée pour les systèmes de codage et décodage. Finalement, on trouve les systèmes d'analyse et de la reconnaissance automatique qui ont bénéficié largement de cette technique.

En résumé, la méthode des moindres carrés, n'utilise aucune information a priori sur le bruit de mesure, et nous avons montré que si le bruit n'est pas à valeur moyenne nulle, l'estimation des paramètres est biaisée. D'une manière générale, nous n'avons dans aucune de ces méthodes utilisées, la loi de répartition statistique du bruit et nous ne l'avons même pas supposée connue. La qualité de l'estimation des pôles se dégrade rapidement lorsque le SNR diminue, et limite de ce fait la qualité des traitements en présence d'un niveau de bruit élevé.

Parmi les perspectives de ce travail, nous souhaitons améliorer la qualité de l'estimateur par l'application de la méthode de vraisemblance au signal de parole bruité ainsi que leur implémentation sur la carte DSP spécialisée. Un second objectif porte sur l'extension des méthodes précédentes au cas des modèles non linéaires à temps continu dans un contexte bruité. Finalement, Nous espérons que notre travail va enrichir la recherche du domaine de processus stochastiques et traitement de la parole.

Références Bibliographiques

- [1] O. Siohan, « Reconnaissance automatique de la parole continue en environnement bruité : Application à des modèles stochastiques de trajectoires », Thesis of Doctorate, university Henri Poincaré, Nancy I, September 1995.
- [2] B. Picinbono, « Signaux aléatoires et bases du traitement statistique du signal », Tome 3 Dunod, Paris 1993.
- [3] T. Chonavel, « Statistical signal processing, Modelling and estimation », Springer, January 2002
- [4] F. Mix, « Random signal processing », prentice hall, 1995.
- [5] D. Arzelier, « Introduction à l'étude des processus stochastiques et au filtrage optimal », version 2, INSA, 1998.
- [6] D. Declercq, A.Quinqui, « Signaux et systemes en questions », Hermes, Paris 1996.
- [7] J. Bernussou, « Processus stochastique – filtrage », Cours ENSICA, Toulouse, 1994.
- [8] J.P. Tubach, « La parole et son traitement automatique », Masson, 1989.
- [9] P. Renevey, « Speech recognition in noisy conditions using missing feature approach », Thesis of doctorate, Federal Polytechnic School of Lausanne, 2000
- [10] T. Dutoit, « Introduction au Traitement Automatique de la Parole », Première édition, Faculté Polytechnique de Mons, Année 2000.
- [11] M. Kunt, M. Bellanger, « Techniques modernes de traitement numérique des signaux », collection électricité, traitement de l'information, Volume1, Année 1991.
- [12] M.D.Srinath, P.k. Rajasekaran,R.Viswanathan, « Introduction to statistical signal processing with applications », Prentice hall editor 1996.
- [13] J.G. Wade, « Codage et traitement du signal », Masson, 1991.
- [14] M. Bellanger, « Traitement numérique du signal, Paris, Collection technique et scientifique des télécommunications », MASSON, 1990.
- [15] M. Bellanger, « Analyse des Signaux et Filtrage Numérique Adaptatif », Masson, 1989.
- [16] R. Boite et M. Kunt, « Traitement de la parole », Lausanne, Presses Polytechniques Romandes, 1987.

- [17] J. D. Markel and A.H. Gray, « Linear Prediction of Speech », Springer-Verlag, 1976
- [18] A.H. Gray and J.D. Markel, « Distance Measures for Speech Processing », IEEE Trans. on ASSP, Vol. 24, No. 5, Oct. 1976.
- [19] J. Makhoul, « Stable and lattice methods for linear prediction », IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 25, pp. 423–428, October 1977.
- [20] A. Buzo, R.M. Gray, A.H. Gray, Jr., and J.D. Markel, « Optimal Quantization of Coefficient Vectors in LPC Speech », 1978 Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, HI, Dec. 1978.
- [21] R.Ben Abdennour, P.Borne, « Identification et commande numérique des procédés industriels », April 2001.
- [22] D. Landau, « Identification et commande des systèmes, Hermes, Paris, 1998.
- [23] Jacques Richalet, « Pratique de l'identification », Hermes, Paris 1998.
- [24] A. Benveniste, M. Metivier, P.Priouret, « Algorithmes adaptatifs et approximations stochastiques, théorie et application », Masson, Paris 1987.
- [25] A. Borrie, « Stochastic System for Engineering, Modelling, Estimation and Control », Prentice hall, 1992
- [26] J. Jean pierre, « Identification et modélisation des processus de génération des crues », Thèse pour l'obtention du grade de docteur en sciences techniques, N° 101, Ecole polytechnique fédérale de Lausanne, 1992.
- [27] Landau, « Commande des systèmes, méthodes d'identification », Année 2002.
- [28] H. Unbehauen, G.P. Rao, « A review of identification in continuous time systems », Annual reviews in Control, pp: 145-171, 1998.
- [29] L. Ljung, « System Identification: Theory for the User », Second edition, Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [30] R. Johansson, « System Modeling and Identification », information and system sciences series, Prentice Hall, 1993.
- [31] E. Walter, L. Pronzato, « Identification de modèles paramétriques à partir de données expérimentales », Masson, 1994.
- [32] G. Rao and N. Sinha, « Identification of Continuous Time Systems, Methodology and Computer Implementation », chapter Continuous-time models and approaches, pages: 1-15, Kluwer Academic Publishers, Edition, 1991.
- [33] P.C. Young and H. Garnier, « Identification and Estimation of Continuous Time Data Based Mechanistic Models for Environmental Systems, Environmental Modeling and Software », pages: 1055–1072, August 2006.

- [34] P.C. Young et H. Garnier, « Identification and estimation of continuous time rainfall-flow models », 14th IFAC Symposium on System Identification (SYSID'2006), pages 1276–1281, Newcastle, Australia, March 2006.
- [35] S. Thil, M. Gilson, et H. Garnier, « Méthodes de compensation de biais pour l'identification de modèles erreurs en les variables », Journées doctorales du GdR MACS, Lyon, France, septembre 2005.
- [36] R. Lozano et D. Taoutaou, Identification et commande adaptative, Hermes, 2001.
- [37] B.R.J. Haverkamp, C.T. Chou, M. Verhaegen, and R. Johansson, « Identification of Continuous Time MIMO State Space Models from Sampled Data in the Presence of Process and Measurement Noise », 35th Conference on Decision and Control, Kobe - Japan, December 1996.
- [38] P. Vanden Hof, « Closed Loop Issues in System Identification », Annual Reviews in Control, 22:173–186, 1998.
- [39] M. Gilson and P. Van den Hof, « Instrumental Variable Methods for Closed Loop System Identification », Automatica, 41(2): 241–249, 2005.
- [40] M. Gilson, H. Garnier, P.C. Young and P. Van den Hof, « Identification of Continuous Time Models from Sampled Data », Springer-Verlag, edition, 2007.
- [41] B. Widrow et al. « Adaptive Noise canceling: Principles and Applications », Proc. IEEE, 63, N° 12, December 1975.
- [42] A. Maddi, A. Guessoum, D. Berkani, « Application the recursive extended least squares method for modelling a speech signal », Second International Symposium on Communications, Control and Signal Processing, ISCCSP'06, Marrakech, Morocco, March 13-15, 2006.
- [43] A. Nongnuch and A. Sorarerks, « An algorithm for vector quantization using density estimation », Proceedings of the 5th WSEAS Int. Conf. on Signal, Speech and Image Processing, Corfu, Greece, August 17-19, 2005.
- [44] A. Maddi, A. Guessoum, D. Berkani, O. Belkina, « Etude de la méthode des moindres carrés récursive et application au signal de parole », Sousse, Tunisie, SETIT, 17-21 Mars 2005.
- [45] M. Djendi, « Amélioration du comportement des algorithmes des moindres carrés transversaux rapides. Application a l'identification de réponse impulsionnelle acoustique longues », thèse de doctorat, école polytechnique d'El Harrach, Alger, Septembre 2006.
- [46] S. Mirsaidi, G. Fleury and J. Oksman, « LMS like AR modelling in the case of missing observations », IEEE Transactions on Signal Processing, vol. 45, n°6, 1997.
- [47] X. Qui, C. Hannsen, « A Modified Filtered-x LMS Algorithm for Active Noise Control of Periodic Noise With On-line Cancellation Path Modeling », Journal of Low Frequency Noise, Vibration and Active Control 19(1), pages 25-46, 2000.

- [48] A. Benveniste, « Design of Adaptive Algorithms for the Tracking of Time- Varying Systems », *Journal of Adaptive Control and Signal Processing*, vol. 1, pp. 3 - 29, 1987.
- [49] A. Benallal, A. Gilloire, « A New Method to stabilize Fast RLS Algorithms based on a first order Model of Propagation of Numerical errors », *Proc. ICASSP, New York*, pp. 1373-1376, 1988.
- [50] A. Gilloire, T. Petillon, « A comparison of NLMS and Fast RLS algorithm for the identification of time - varying systems with noisy outputs, application to acoustic echo cancellation », *EUSIPCO, Barcelona, Spain*, September 1990.
- [51] A. Maddi, A. Guessoum, D. Berkani, « A Comparative Study of Speech Modeling Methods », *11th WSEAS International Conference on SYSTEMS, Agios Nikolaos, Crete Island, Greece*, pages: 104-109, July 23-25 2007.
- [52] A. Maddi, A. Guessoum, D. Berkani, « Using Stochastic Gradient Method for Modeling a Speech Signal », *SIGMAP 2007 the International Conference on Signal Processing and Multimedia Applications, Barcelona, Spain*, 28 - 31 July, 2007.
- [53] S. Cruces, A. Cichocki, « Globally Convergent Newton Algorithms for Blind Decorrelation », *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan*, April 2003.
- [54] S.C. Douglas, « Numerically-robust RLS algorithms using Least Squares pre-whitening », in *Proceedings of ICASSP, Istanbul, Turkey*, pp. 412-415, 2000.
- [55] S. Karaboyas and N. Kalouptsidis, « Efficient adaptive algorithms for ARX identification », *IEEE Trans. Signal Processing*, vol. 39, no. 3, pp. 571-582, Mar. 1991.
- [56] G. V. Moustakides and S. Theodoridis, « Fast Newton transversal filters, a new class of adaptive estimation algorithms », *IEEE Trans. Signal Processing*, vol. 39, no 10, pp. 2184-2193, Oct. 1991.
- [57] S. Theodoridis. G. V. Moustakides, and K. Berberidis, « A class of fast adaptive algorithms for multichannel system identification and signal processing », *CTI Tech. Rep. 930936*, Sept. 1993.
- [58] K.P. Wong, D. Feng, et W.C. Siu, « Generalized linear least squares algorithm for non uniformly sampled biomedical system identification with possible repeated eigenvalues », *Computer methods and programs in biomedicine*, 57 :167–177, 1998.
- [59] H. Garnier, M. Gilson, and W.X. Zheng, « A bias-eliminated least squares method for continuous-time model identification of closed-loop systems », *International Journal of Control*, 73(1): 38–48, 2000.
- [60] W.X. Zheng, « A bias correction method for identification of linear dynamic errors in variables models », *IEEE Transactions on Automatic Control*, 47(7):1142–1147, July 2002.
- [61] G. Feng, « Modélisation acoustique et traitement du signal de parole », *Thèse de Docteur, INP, Grenoble*, 1986.

- [62] H. Morikawa, « Adaptive analysis of speech based on a pole zero representation », IEEE Trans. On ASSP, ASSP-30, n° 1, pp. 77-87,1982.
- [63] G. FENG, « Modélisation acoustique et traitement du signal de parole. Le cas des voyelles nasals », thèse de Docteur, INP, Grenoble, Juillet 1984.
- [64] M. Gales, « Model based techniques for noise robust speech recognition », Ph. D. Thesis, University of Cambridge, September 1995.
- [65] L. Benaroya, « Séparation de plusieurs sources sonores avec un seul microphone », Ph. D. Thèse, Université de Rennes 1, 2003.
- [66] G. Gabnoulet, D. Jovet, « Reconnaissance de la parole et modélisation statistique : expérience du CNET », Traitement du signal, Vol 7, n° 4, pp. 267-274, 1990.
- [67] Calliope, « La parole et son traitement automatique. Collection Technique et Scientifique des Télécommunications », Masson 1989.
- [68] A. Maddi, A. Guessoum, D. Berkani, « Noisy Speech Modelling Using Recursive Extended Least Squares Method », WSEAS TRANSACTIONS on SIGNAL PROCESSING Issue 9, Volume 2, ISSN 1790-5022, pages: 1268-1274, September 2006.
- [69] A. Maddi, A. Guessoum, D. Berkani, « Using Recursive Least Squares Estimator for Modelling a Speech Signal », The 6th WSEAS International Conference on SIGNAL, SPEECH AND IMAGE PROCESSING SSIP '06 Lisbon, Portugal, September 22-24, 2006.
- [70] WAKITA, « Linear predictive coding and its application to speech processing », Bonas, France, June 1981
- [71] R.Sanchis, P. Albertos, « Recursive identification under scarce measurements convergence analysis », Automatica, 38, pp.: 535-544, 2002.
- [72] A. Isaksson, « A recursive EM algorithm subject to missing data », in SYSID, vol.2, pp. 679-684, Copenhagen, Denmark, 1994.
- [73] L. Plévert, « L'électronique sait aujourd'hui se connecter au corps humain ». Électronique International Hebdo, N° 294, Février 1998, pages 20-21
- [74] P. Gournay, G. Quagliari, A. Goye, G. Guilmin, F. Chartier, « Procédé de correction auditive utilisant un modèle paramétrique du signal de parole », Dix-septième colloque GRETSI, Vannes, 13-17 septembre 1999.
- [75] J. Makhoul, « Linear Prediction: A Tutorial Review ». Proc. of the IEEE 1975, Vol. 63, No. 4, pp. 561-580.
- [76] R. J. Macaulay, M.L Malpass, « Speech enhancement using decision noise suppression filter », IEEE trans. on ASSP, Vol. 28, No 2, pp. 137-145,1980.
- [77] E. Koo, J. D. Gibson et S. D. Gray, « Filtering of colored noise for speech enhancement and coding », IEEE ICASSP 1989, pp. 349-352.

- [78] C. Baillargeat, « Contribution à l'amélioration des performances d'un radiotéléphone mains-libres à commandes vocales ». Thèse de l'université Paris VI. Année 1991.
- [79] J. Yang, « Frequency domain noise suppression approaches in mobile telephone systems », IEEE ICASSP'93, pp. 363-366, 1993.
- [80] A. Maddi, A. Guessoum, D. Berkani, « A Comparative Study of Speech Modeling Methods », The 11th WSEAS International Conference on SYSTEMS, Agios Nikolaos, Crete Island, Greece, pages: 104-109, July 23-25 2007.
- [81] A. Maddi, A. Guessoum, D. Berkani, « Using Stochastic Gradient Method for Modeling a Speech Signal », SIGMAP 2007 the International Conference on Signal Processing and Multimedia Applications, Barcelona, Spain; 28-31 July, 2007.
- [82] H. Garnier, M. Gilson and W.X. Zheng, « A bias eliminated least-squares method for continuous-time model identification of closed-loop systems », International Journal of Control, pp: 38-48, 2000.
- [83] H. Ezzaidi, « Détection de la double parole dans le contexte de la radio-téléphone main libre en véhicule », Mémoire pour la maîtrise de l'ingénierie, Université de Québec à Chicoutimi, Décembre 97.
- [84] M. Firouzmand & L. Girin, « Perceptually weighted long-term modeling of sinusoidal speech amplitude trajectories », Proc. IEEE Int. Conf. on Acoustics, Speech & Signal Proc. (ICASSP 2005), Philadelphia, USA, 2005.
- [85] R. Boite, H. Boulard, T. Dutoit, J. Hancq, H. Leich, « Traitement de la parole », Presses Polytechniques et Universitaires Romandes, 2000.
- [86] J. Nolzco, S. Young, « Continuous speech recognition in noise using spectral subtraction and HMM adaptation », The IEEE International Conference on Acoustic, Speech and Signal Processing, 1994
- [87] J. S. Garofolo et al. « DARPA TIMIT Acoustic phonetic Continuous Speech Database », NIST, Gaithersburg, October 1988
- [88] M. Kuo, R. Morgan, « Active Noise Control Systems, Algorithms and DSP Implementations », Wiley Inter science, 1996
- [89] L. Buniet, « Traitement automatique de la parole en milieu bruité : étude de modèles connexionnistes statiques et dynamiques », Thèse présentée pour l'obtention du Doctorat de l'Université Henri Poincaré - Nancy 1 spécialité informatique, février 1997.
- [90] C. Alessandro et C. Demars, « Représentations temps - fréquence du signal de parole, Traitement du signal », Vol 9, n° 2, pp. 153-172, 1992.
- [91] M. Zhuy, F. Peyrin, R. Goutte, « Transformation de pseudo Wigner discrète en temps et en fréquence, Annales des Télécommunications », Vol 46, n° 5-6, 1991, pp. 301-309.

- [92] A. Waibel , T. Hanazawa , G. Hinton , « Phoneme recognition using Time-Delay Neural Networks », ATR Research Report, TR-I-0006, Octobre 1987.
- [93] J.F. Roux, « Transformée de Fourier discrète récursive avec fenêtre exponentielle dissymétrique », Traitement du signal, Vol 10, n° 2, pp. 185-192,1993.
- [94] G. Pernnou, « La reconnaissance vocale », Le courrier du CNRS, n°80, pp. 95-96, 1994.
- [95] J.C. Junqua, « Utilisation d'un modèle d'audition et de connaissances phonétiques en reconnaissance automatique de la parole », Traitement du signal, Vol 7, n° 4, 1990, pp. 275-284.
- [96] V. Barreaud, « Reconnaissance automatique de la parole continue : compensation des bruits par transformation de la parole », Thèse de doctorat de l'université Henri Poincaré, Novembre 2004.
- [97] B. Doval, C.Alessandro, and B. Diard, « Spectral methods for voice source parameters estimation », Euro-speech, pp: 533 – 536, 1997.
- [98] E. Bilange, « Dialogue personne - machine, modélisation et réalisation informatique », Paris, Hermes, 1992.
- [99] S. Lauglin, B. Mulgrew, C. Cowan, « Performance Comparison of LS and LMS as HF Channel Estimators », Proc. ICASSP, pp. 49, 1987.
- [100] B. Doval, « Estimation de la fréquence fondamentale des signaux sonores », Thèse de PHD, Université de Paris VI, 1994.
- [101] D. C. Farden, « Tracking Properties of Adaptive Signal Processing Algorithms », IEEE Transactions ASSP, Vol. 29, pp. 439 - 446, 1981
- [102] O. Macchi, M. Bellanger, « Le Filtrage Adaptatif Transverse », Annales des Télécommunications, vol. 5, pp. 115-132, 1988.
- [103] A. B. Kain and J. P. H. van Santen, « A speech model of acoustic inventories based on asynchronous interpolation », Euro-speech, pages 329–332, 2003.
- [104] A. Benallal, « Etude des algorithmes des moindres carrés transversaux rapides et application à l'identification des réponses impulsionnelles acoustiques », Thèse de Doctorat, Université de Rennes-1, Janvier 1989.
- [105] Y. Engel, S. Mannor, R. Meir, « The Kernel Recursive Least Squares Algorithm », IEEE Trans. Signal Proc.2004, 52 (8), 2275-2285.
- [106] Ch. 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, No 6, 1989.
- [107] L. Ljung, « System identification toolbox for use with MATLAB », User's guide, version 6.0, the math works, Inc., November 2003.

- [108] C.T. Chou, M. Verhaegen, et R. Johansson, « Continuous-time identification of SISO systems using Laguerre functions », *IEEE Transactions on Signal Processing*, 47(2):349–362, 1999.
- [109] H. Garnier, M. Gilson, T. Bastogne, et H. Zbali, « CONTSID : un outil logiciel pour l'identification de modèles paramétriques à temps continu à partir de données expérimentales », *Journées Identification et Modélisation Expérimentale (JIME'2006)*, Poitiers (France), Novembre 2006
- [110] M. Gilson, H. Garnier and P. Van den Hof, « Instrumental Variable Methods for Continuous Time Model Identification in Closed Loop », *American Control Conference*, Boston, USA, 2004.
- [111] S.V. Huffel and P. Lemmerling, « Total Least Squares and Errors Invariables Modeling », *Kluwer Academic Publishers*, edition, 2002.
- [112] K. Paliwal, A. Basu, « A Speech enhancement method based on kalman filtering », *IEEE ICASSP'87*, pp. 177-180, 1987.

Nom: Maddi
Prénom : Abdelkader
Adresse : Département d'Electronique, USD de Blida
Fonction : Enseignant/chercheur à l'USDB
Email : a_maddi@hotmail.com



Curriculum vitae

FORMATIONS (ETUDES UNIVERSITAIRES)

- Mars 1997 diplôme de magister en électronique option contrôle. Université de Blida.
- Octobre 1990 diplôme d'ingénieur d'état en électronique option contrôle. Université de Blida.
- Juin 1985 Baccalauréat série Mathématiques. Lycée mixte de Meftah.

EXPERIENCE PROFESSIONNELLE

- 1992-1998 : Assistant au niveau de département d'Electronique à l'USDB.
- 1999-2001 : Maître assistant à l'USDB.
- 2002-2008 : Chargé de cours à l'USDB.

COMMUNICATIONS ET PUBLICATIONS

- A. Maddi, « Modélisation et contrôle du vol latéral d'un avion », First International Conference on Electrical Engineering ICEE'2000, University of Boumerdes, November 04-06, 2000, Boumerdes, Algeria.
- A. Maddi, A. Guessoum, D. Berkani, O. Belkina, « Etude de la méthode des moindres carrée récursive et application au signal de parole », SETIT 2005, 17-21, 2005, Sousse, Tunisie.
- A. Maddi, A. Guessoum, D. Berkani, O. Belkina, « Etude de la méthode des moindres carrée étendue et application au signal de parole », SETIT 2005, 17-21, Mars 2005 Sousse, Tunisie.
- A. Maddi, A. Guessoum, D. Berkani, « Application the recursive extended least squares method for modeling a speech signal », Second International Symposium on Communications, Control and Signal Processing, ISCCSP'06, March 13-15, 2006 Marrakech, Morocco.
- A. Maddi, A. Guessoum, D. Berkani, « Using Recursive Least Squares Estimator for Modeling a Speech Signal », The 6th WSEAS International Conference on SIGNAL, SPEECH AND IMAGE PROCESSING SSIP '06 Lisbon, Portugal, September 22-24 2006, pages:182-186.

- A. Maddi, A. Guessoum, D. Berkani, « Noisy Speech Modeling Using Recursive Extended Least Squares Method », WSEAS TRANSACTIONS ON SIGNAL PROCESSING Issue 9, Volume 2, September 2006 ISSN 1790-5022, pages: 1268-1274.
- A. Maddi, A. Guessoum, D. Berkani, « A Comparative Study of Speech Modelling Methods », 11th WSEAS International Conference on SYSTEMS. Agios Nikolaos, Crete Island, Greece, July 23-25 2007, pages: 104-109.
- A. Maddi, A. Guessoum, D. Berkani, « Using Stochastic Gradient Method for Modeling a Speech Signal », SIGMAP 2007 the International Conference on Signal Processing and Multimedia Applications, Barcelona, Spain; 28 - 31 July, 2007.

ENCADREMENT DES PROJETS DE FIN D'ETUDE

- «Blanchissement de l'erreur de prédiction par la méthode de vraisemblance. Application au signal de parole», PFE d'ingénieur d'état, Département d'électronique, USDB, Octobre 2007.
- «Identification paramétrique par les algorithmes de type gradient et Newton stochastique et application au signal de parole», PFE d'ingénieur d'état, Département d'électronique, USDB, Juillet 2006.
- « Etude et réalisation d'une serrure à carte à puce basée d'un microcontrôleur PIC16F876 », PFE de DEUA, Département d'électronique, USDB, Juin 2005.
- « Identification paramétrique par les méthodes basées sur le blanchissement de l'erreur de prédiction et application au signal de parole ». PFE d'ingénieur d'état, Département d'électronique, USDB, Octobre 2004.
- « Etude et réalisation d'un système de protection contre les coupures du courant de secteur », PFE de DEUA, Département d'électronique, USDB, Octobre 2004.
- « Etude et implémentation des algorithmes d'identification paramétriques et application au signal de parole ». PFE d'ingénieur d'état, Département d'électronique, USDB, Octobre 2002.
- « Etude et réalisation d'un simulateur de tension pour mesurer la capacité », PFE de DEUA, Département d'Electronique, USDB, Octobre 2001.

CONNAISSANCES EN INFORMATIQUE

- Maîtrise de la programmation en Langage MATLAB
- Maîtrise de la programmation en Langage SIMNON
- Maîtrise de la programmation en Langage PASCAL

ACTUELLEMENT (2007/2008)

- Chargé de cours du module « Mesures Electriques » au niveau de département d'Electronique, USDB.
- Chargé de cours du module « Théories des probabilités et statistiques » au niveau de département d'électronique, USDB.
- Chargé de cours du module « Identification des Processus » au niveau de département d'électronique, USDB.
- Chargé de cours du module « Systèmes Multivariables » au niveau de département d'électronique, USDB.
- Membre chargé de recherche dans l'équipe de recherche ayant le code J0200420060010, « Conception de programmes et systèmes adaptatifs ».

DOMAINE D'INTERET

- Traitement du signal
- Identification et modélisation des systèmes
- Contrôle adaptatif