



Université ZIANE ACHOUR de Djelfa
Faculté des Sciences et de la Technologie
Département des Sciences et des Technologiques

Mémoire de Magister en Electronique

Option : Signaux et Systèmes

Présenté par :

LATRECHE Boubakeur

Ingénieur en Electronique

Thème :

*Identification de modèles ARMA utilisant
les observations quantifiées à la sortie.*

Devant le Jury :

M BOUKEZZI Laarbi	MCA à UZA Djelfa	Président
M MITICHE Lahcène	MCA à UZA Djelfa	Rapporteur
M ^{me} ADAMOU-MITICHE A.B.H.	MCA à UZA Djelfa	Co-rapporteur
M AHFIR Maamar	MCA à UAT Laghouat	Examineur
M GUEDDIM Ahmed	MCA à UZA Djelfa	Examineur

2013-2014

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Remerciements

*Je tiens à exprimer mon vif remerciement à mon promoteur **Dr. Lahcène MITICHE** à l'université Ziane-Achour de Djelfa pour m'avoir encadré durant cette mémoire et me conseiller tout au long de mon travail.*

*J'adresse également mes remerciements à **M^{me} ABH. ADAMOÛ MITICHE**, Maître de conférences à l'université Ziane-Achour de Djelfa, pour son encouragement et son soutien.*

*J'exprime ma gratitude à **Mr L. BOÛKEZI**, Maître de conférence à l'université Ziane-Achour de Djelfa, le Président du jury, d'avoir bien voulu présider mon jury.*

Je remercie chaleureusement les membres du jury :

- *Mr M. AHFIR*
- *Mr A. GUEDDIM*

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

*Je remercie également **Mr S. SLAMI** Docteur à l'université Ziane-Achour de Djelfa, **Mr A. MADDI** Maître de conférence à l'USDB, pour m'avoir soutenu dans les moments difficiles. Qu'ils trouvent ici ma reconnaissance et mon respect.*

Enfin, je remercie tout à tous ceux qui nous ont soutenu ou qui, d'une manière ou d'une autre, ont contribué à l'élaboration de ce travail.

Dédicaces

A ma très chère mère.

A mon très cher père.

A mes frères et mes sœurs.

A toute ma famille

A tous mes amis.

...

...

...

...

A tous ceux qui m'aiment.

Boubakeur.

ملخص

إنّ لخوارزميات الإحصاء علاقة وطيدة جدا بطرق الإحصاء وطبيعة النماذج الرياضية. لهذا تعتبر المرحلة الأساسية في طريقة المطابقة مبنية على المعرفة الجيدة لكل نظام فيزيائي. نتطرق في عملنا هذا إلى خوارزمية المطابقة للنماذج من نوع ARMA التي تكون بمخارج مكمّمة، سنستعمل هنا خوارزمية التمكين الأعظم. سنقوم كذلك بدراسة والخوارزميات التالية: RLS، LMS ونيوتن لمطابقة الأبعاد القائمة على تبييض الخطأ للحصول على الضوضاء البيضاء، علما أنّ أغلبية الأنظمة الفيزيائية تمثّل وفق النماذج: ARMA، ARMAX و ARX.

كلمات مفتاح: المطابقة، نموذج ARMA، مخارج مكمّمة، التمكين الأعظم.

Abstract

The identification of systems is strongly related at 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.

We present here identification method for ARMA model with quantized output, using Maximum Likelihood estimation to estimate the parameters of the ARMA model.

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.

We present the Maximum Likelihood Algorithm, ML composed of two methods, and Expectation Maximization algorithm and Quasi-Newton gradient search algorithm, finally, examples for examine the performance of ML algorithm.

Keywords: Identification, ARMA, Estimation, Quantification, Maximum Likelihood.

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.

Nous présentons dans ce travail une méthode d'identification de modèle ARMA dans le cas où la sortie est quantifiée, nous allons utiliser la méthode de maximum de vraisemblance pour estimer les paramètres de ce modèle (ARMA).

Nous présentons aussi, 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.

Ces méthodes 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.

Nous présentons la méthode maximum de vraisemblance, elle se décompose en deux méthodes essentielles, EM et Quasi-Newton. A la fin, on présent plusieurs exemples de simulation pour juger les performances de ces algorithmes.

Mots clés : Identification, Modèle ARMA, Estimation, Quantification, Maximum de vraisemblance.

Liste des Abréviations

AR: Autorégressif.

ARMA : Autorégressif à Moyenne Ajustée.

BFGS : Broyden-Fletcher-Goldfarb-Shanno.

CAN : Convertisseur Analogique Numérique.

E: Espérance Mathématique.

EM : Expectation Maximisation.

EQM : Ecart Quadratique Moyen.

FA : Fonction Aléatoire.

GS: Gradient Stochastique.

IID : Indépendantes et Identiquement Distribuées.

MA : Moyenne Ajustée.

MCR: Moindre carrés réursive.

MVR: Maximum de Vraisemblance.

NLMS: Gradient Stochastique Normalisée.

RIF : Réponse Impulsionnelle Finie.

RII: Réponse Impulsionnelle Infinie.

SBPA: Séquence Binaires Pseudo Aléatoires.

SN: Newton Stochastique.

Liste des Figures

Fig. I.1: Génération d'un signal aléatoire corrélé.....	09
Fig. I.2: Estimateur biaisé de faible variance.....	14
Fig. I.3: Estimateur non biaisé de forte variance.....	15
Fig. II.1: Schéma d'identification paramétrique.....	26
Fig. III.1: Représentation graphique de la méthode de Newton.....	34
Fig.III.2: Structure de réalisation d'un modèle ARMA.....	37
Fig.IV.1: Séquence d'entrée SBPA	47
Fig.IV.2: Bruit additif gaussien	48
Fig.IV.3: Lieu des pôles (à gauche) et lieu des zéros (à droite).....	49
Fig.IV.4: Lieu des pôles (à gauche) et lieu des zéros (à droite).....	49
Fig.IV.5: Spectres d'amplitude du modèle original et du modèle estimé.....	52
Fig.IV.6: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	52
Fig.IV.7: - Fonction d'autocorrélation d'erreur de prédiction.....	53
Fig.IV.8: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c),- Erreur d'identification du paramètre b_2 (figure d).....	53
Fig.IV.9: - Lieu des pôles du modèle estimé (à gauche) ,- Lieu des zéros du modèle estimé (à droite)	54
Fig.IV.10: Spectres d'amplitude du modèle original et du modèle estimé.....	55
Fig.IV.11: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	56
Fig.IV.12: - Fonction d'autocorrélation d'erreur de prédiction.....	56

Fig.IV.13: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c),- Erreur d'identification du paramètre b_2 (figure d).....	57
Fig.IV.14: - Lieu des pôles du modèle estimé (à gauche) , - Lieu des zéros du modèle estimé (à droite)	57
Fig.IV.15: Spectres d'amplitude du modèle original et du modèle estimé.....	59
Fig.IV.16: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	60
Fig.IV.17: - Fonction d'autocorrélation d'erreur de prédiction.....	60
Fig.IV.18: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c),- Erreur d'identification du paramètre b_2 (figure d).....	61
Fig.IV.19: - Lieu des pôles du modèle estimé (à gauche) , - Lieu des zéros du modèle estimé (à droite)	61
Fig.IV.20: Spectres d'amplitude du modèle original et du modèle estimé.....	63
Fig.IV.21: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	63
Fig.IV.22: - Fonction d'autocorrélation d'erreur de prédiction.....	64
Fig.IV.23: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c), - Erreur d'identification du paramètre b_2 (figure d).....	64
Fig.IV.24: - Lieu des pôles du modèle estimé (à gauche) , - Lieu des zéros du modèle estimé (à droite)	65
Fig.IV.25: Spectres d'amplitude du signal d'entrée et celui du signal quantifié (B=1 bit).....	66
Fig.IV.26: Spectres d'amplitude du signal d'entrée et celui du signal quantifié (B=2 bit).....	67
Fig.IV.27 : Spectres d'amplitude du signal d'entrée et celui du signal quantifié (B=3 bit).....	67
Fig.IV.28: Spectres d'amplitude du modèle original et du modèle estimé.....	69
Fig.IV.29: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	69

Fig.IV.30: - Fonction d'autocorrélation d'erreur de prédiction.....	70
Fig.IV.31: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c),- Erreur d'identification du paramètre b_2 (figure d).....	70
Fig.IV.32: - Lieu des pôles du modèle estimé (à gauche) , - Lieu des zéros du modèle estimé (à droite)	71
Fig.IV.33: Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).....	72
Fig.IV.34: - Fonction d'autocorrélation d'erreur de prédiction.....	73
Fig.IV.35: - Erreur d'identification du paramètre a_1 (figure a),- Erreur d'identification du paramètre a_2 (figure b),- Erreur d'identification du paramètre b_1 (figure c),- Erreur d'identification du paramètre b_2 (figure d).....	73

Liste des Tableaux

Tab. IV.1: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25)$	51
Tab. IV.2: Influence de la variance $\sigma^2, (N = 1024)$	51
Tab. IV.3: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25)$	55
Tab. IV.4: Influence de la variance $\sigma^2, (N = 1024)$	55
Tab. IV.5: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25)$	58
Tab. IV.6: Influence de la variance $\sigma^2, (N = 1024)$	59
Tab. IV.7: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25)$	62
Tab. IV.8: Influence de la variance $\sigma^2, (N = 1024)$	62
Tab. IV.9: Biais et EQM (RLS et ML) $\sigma^2 = 0.25$, exemple à phase minimale.....	65
Tab. IV.10: Biais et EQM (RLS et ML) $\sigma^2 = 0.25$, exemple à phase non minimale.....	66
Tab. IV.11: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25, b = 2)$	68
Tab. IV.12: Influence de la variance $\sigma^2, (N = 1024, b = 2)$	68
Tab. IV.13: Influence du nombre d'échantillons $N, (\sigma^2 = 0.25, b = 2)$	71
Tab. IV.14: Influence de la variance $\sigma^2, (N = 1024, b = 2)$	72

Table des matières

Remerciements.....	i
Dédicaces.....	ii
Résumé.....	iii
Liste des Abréviations.....	iv
Liste des Figures.....	v
Liste des Tableaux.....	viii
INTRODUCTION GENERALE.....	01
CHAPITRE I : NOTIONS SUR LES PROCESSUS STOCHASTIQUES ET L'ESTIMATION	
I.1 Introduction.....	04
I.2 Notions sur les processus stochastique.....	04
I.2.1 Densité de probabilité d'une fonction aléatoire en temps continu.....	04
I.2.2 Indépendance.....	05
I.2.3 Moment d'une fonction aléatoire.....	05
I.2.3.1 Moyenne.....	05
I.2.3.2 Variance et covariance.....	05
I.2.3.3 Covariance mutuelle.....	06
I.2.4 Stationnarité et Ergodicité.....	06
I.2.4.1 Stationnarité.....	06
I.2.4.1.1 Stationnarité au sens strict.....	07
I.2.4.1.2 Stationnarité au sens faible.....	07
I.2.4.2 Ergodicité.....	07
I.2.4.2.1 Ergodicité relativement à la moyenne.....	07
I.2.4.2.2 Ergodicité relativement à la fonction de corrélation.....	08
I.2.5 Densité spectrale de puissance.....	09
I.2.6 Modèles stochastiques.....	09
I.2.7 Processus stochastiques remarquables.....	10
I.2.7.1 Processus Gaussien.....	10
I.2.7.2 Processus Markovien.....	10
I.3 Notions de la théorie d'estimation.....	11
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.....	12
I.3.4 Estimateurs et estimations.....	12
I.3.5 Propriétés d'un bon estimateur.....	13
I.3.5.1 Estimateur convergent.....	13
I.3.5.2 Biais.....	13
I.3.5.3 Efficacité.....	14
I.3.5.4 Ecart quadratique moyen d'un estimateur.....	14
I.4 Notions de bruits.....	15
I.4.1 Bruit blanc.....	15
I.4.2 Séquence binaire pseudo aléatoire.....	16
I.4.3 Bruit coloré.....	16
I.5 Conclusion.....	16

CHAPITRE II : MODELISATION ET IDENTIFICATION DES PROCESSUS PHYSIQUES

II.1 Introduction.....	19
II.2 Modélisation des processus physiques.....	19
II.2.1 Différentes approches d'établissement d'un modèle.....	19
II.2.1.1 Modèle de connaissance.....	20
II.2.1.2 Modèle de représentation.....	20
II.2.2 Modélisation autorégressive à moyenne ajustée (ARMA).....	20
II.2.2.1 Modélisation autorégressive (AR).....	21
II.2.2.2 Modélisation à Moyenne Ajustée (MA).....	21
II.2.3 Principales étapes de la modélisation.....	22
II.2.3.1 Système.....	22
II.2.3.2 Modèle.....	22
II.2.3.3 Critère.....	22
II.2.3.4 Optimisation.....	23
II.2.3.5 Incertitude sur les paramètres.....	23
II.2.3.6 Analyse critique des résultats obtenus.....	23
II.2.4 Difficultés de la modélisation.....	23
II.2.4.1 Définition du problème de modélisation.....	23

II.2.4.2	Qualité des données de modélisation.....	24
II.2.4.3	Choix de la technique utilisée.....	24
II.2.4.4	Choix des variables disponibles.....	24
II.2.4.5	Sélection de modèle adéquat.....	24
II.3	Identification des paramètres d'un processus physique.....	25
II.3.1	Identification non paramétrique.....	25
II.3.2	Identification paramétrique.....	25
II.3.3	Etapes d'identification.....	26
II.3.3.1	Acquisition des entrées et sorties	26
II.3.3.2	Choix de la complexité du modèle.....	27
II.3.3.3	Estimation des paramètres du modèle.....	27
II.3.3.4	Validation du modèle.....	27
II.3.4	Algorithme d'identification récursif.....	27
II.4	Conclusion.....	28

CHAPITRE III : ETUDE DES ALGORITHMES D'IDENTIFICATION

III.1	Introduction.....	30
III.2	Algorithme du gradient stochastique.....	30
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.....	33
III.3.1	Principe de la méthode.....	33
III.3.2	Equations de mise en œuvre de la méthode de Newton.....	34
III.3.3	Identification du modèle ARMA.....	35
III.4	Algorithme des moindres carrés.....	36
III.4.1	Identification du modèle ARMA.....	36
III.4.2	Solution optimale au sens des moindres carrés.....	37
III.4.3	Algorithme des moindres carrés récursif.....	38
III.4.4	Propriétés de l'estimateur des moindres carrés.....	39
III.5	Algorithme de Maximum de vraisemblance (MVR).....	39
III.5.1	Identification du modèle ARMA.....	39
III.5.2	Maximum de vraisemblance.....	40

III.5.2.1 La méthode d'Expectation Maximisation(EM) récursive.....	41
III.5.2.2 La méthode de Quasi-Newton récursive.....	43
III.5.3 Sommaire de l'algorithme d'Identification.....	44
III.6 Conclusion.....	45

CHAPITRE IV : SIMULATIONS ET INTERPRETATION

IV.1 Introduction.....	47
IV.2 Etapes de simulation.....	47
IV.2.1 Choix des signaux de test.....	47
IV.2.2 Choix des paramètres du système.....	48
IV.2.3 Identification des paramètres du modèle.....	50
IV.3 Implémentation des algorithmes.....	50
IV.4 Identification des paramètres par l'algorithme MCR.....	50
IV.4.1 Système à phase minimale.....	50
IV.4.1.1 Influence du nombre d'échantillons et de la variance.....	51
IV.4.2 Système à phase non minimale.....	54
IV.4.2.1 Influence du nombre d'échantillons et de la variance.....	54
IV.5 Identification par Maximum de Vraisemblance (MVR).....	58
IV.5.1 Système à phase minimale.....	58
IV.5.1.1 Influence du nombre d'échantillons et de la variance.....	58
IV.5.2 Système à phase non minimale.....	62
IV.5.2.1 Influence du nombre d'échantillons et de la variance.....	62
IV.6 Etude comparative des algorithmes d'identification MVR et MCR.....	65
IV.7 Identification par Maximum de Vraisemblance à sortie quantifiée.....	66
IV.7.1 Système à phase minimale.....	67
IV.7.1.1 Influence du nombre d'échantillons et de la variance.....	67
IV.7.2 Système à phase non minimale.....	71
IV.7.2.1 Influence du nombre d'échantillons et de la variance.....	71
IV.8 Conclusion.....	74
CONCLUSION GENERALE	76
Références bibliographique.....	79

INTRODUCTION GENERALE

Introduction générale

Depuis les années 60, le filtrage adaptatif a suscité un développement sans précédent. Ce développement du filtrage adaptatif est né de l'essor du traitement numérique, de la croissance soutenue de la puissance des processeurs de traitement qui permettent la mise en œuvre en temps réel, d'algorithmes de plus en plus complexes et qui vont à des cadences de plus en plus élevées [1].

Suite à cet essor du filtrage adaptatif, de nombreuses applications telles que : l'identification des systèmes [2,3], l'égalisation de canaux, la modélisation des systèmes [4], la résolution de problème inverse [5] ont vu le jour. Des applications mentionnées plus haut, certaines comme l'identification des systèmes, sont un problème important dans des domaines aussi variés que le traitement numérique du signal [6], le contrôle de procédé [7], ou encore les communications. L'identification de système est réalisée par l'ajustement de paramètres pour un modèle donné jusqu'à ce que ses sorties, pour une entrée particulière, coïncident autant que possible avec la sortie mesurée du système identifié pour la même entrée.

L'identification est aujourd'hui une composante fondamentale des sciences de l'ingénieur située au traitement du signal numérique, cette discipline scientifique est liée au développement des moyens et des techniques de télécommunications [8], traitement de l'information [9], le codage et le décodage (compression) [10], la conversion analogique/numérique (CAN). Le développement continu de ces techniques à aider les chercheurs à mieux comprendre l'identification si les sorties (observations) et/ou les entrées sont quantifiées (**quantized identification**) [3].

L'identification des systèmes linéaires est effectuée depuis des décennies à l'aide des algorithmes du gradient stochastique tels que : le LMS (Least Mean Square), sa méthode normalisée NMLS (Normalized LMS), Newton stochastique (SN). L'identification s'est également beaucoup effectuée à l'aide de méthode de type récursif comme le moindres carrés récursif 'MCR' et la méthode de maximum de vraisemblance (MVR).

Notre travail consiste à l'identification d'un modèle ARMA (Autorégressive à Moyenne Ajustée) [11] à sortie quantifiée. A titre d'exemple, soit un signal d'entrée $u(t)$ est modélisé par un signal déterministe ou un processus aléatoire qui excite le modèle ARMA, la sortie est bruitée par un bruit qui est une séquence indépendantes et identiquement distribuées

(IID) de distribution normale $N(0, \sigma^2)$. La sortie bruyante $y(t)$ est quantifiée avec un quantificateur scalaire de Lloyd-Max Q_t de K niveau variant dans le temps t

Notre objectif est d'estimer les paramètres du modèle ARMA. Nous le faisons en utilisant le critère du maximum de vraisemblance (Maximum Likelihood Estimation).

Le mémoire est ordonné comme suit :

Le premier chapitre de ce mémoire est un bref rappel sur des notions et généralités de processus stochastiques, des notions sur la théorie d'estimation et quelque notion sur les bruits.

Le second chapitre est réservé à la modélisation des processus physiques de modèles ARMA et ses principale étapes (système, modèle, critère, optimisation...), 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 (SN), Gradient stochastique (LMS) et Moindres carrés récursifs (RLS), et par la suite, la méthode de maximum de vraisemblance, qui est l'axe de notre travail

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

Chapitre I :
NOTIONS SUR LES
PROCESSUS STOCHASTIQUE
ET L'ESTIMATION

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 [12]. C'est l'objet de ce chapitre.

I.2 Notions sur les processus stochastiques

Un processus stochastique ou processus aléatoire se définit par l'évolution au cours du temps d'un phénomène statistique selon des lois de probabilités [13]. 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 on t_1, \dots, t_n , on obtient une variable

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHASTIQUE ET L'ESTIMATION

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 nième 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 Fonction Aléatoire (FA) qu'elle est à valeurs indépendantes si [14]

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

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

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

I.2.3 Moments 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 (t), la fonction de t qui pour toute valeur de t donne la moyenne de la variable aléatoire obtenue à cet instant, soit

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

I.2.3.2 Variance et covariance

De façon similaire, la variance d'une fonction aléatoire scalaire est une fonction de t définie par [12]:

$$\sigma_X^2(t) = E\{[X(t) - m_X(t)][X(t) - m_X(t)]^*\}$$

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

$$\sigma_X^2(t) = \int_{-\infty}^{+\infty} [x - m_X(t)][x - m_X(t)]^* P_X[x(t)] dx \quad (I.4)$$

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

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

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 [14]

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

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

I.2.4 Stationnarité et Ergodicité

Ces propriétés sont très importantes, car si les fonctions aléatoires que l'on traite n'en bénéficiaient 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 [14] :

- $m_X(t) = m_X$ (I.7)

- $C_{XY}(t_1, t_2) = C_{XY}(\tau = t_1 - t_2)$ (I.8)

De ce fait la variance est aussi constante puisque :

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

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'auto corrélation[15]. 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 $X(t)$

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

par :

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

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) :

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

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

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

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

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

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 [12] :

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

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 :

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

I.2.5 Densité spectrale de puissance

La fonction d'auto corrélation 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'auto-corrélation. Le spectre et la fonction d'autocorrélation sont reliés comme suit :

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

Si le signal est plus ergodique, la fonction d'autocorrélation peut être obtenu 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 [16]. (Fig. I.1)



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

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

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 [12]. 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 :

$$P_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp \left[-\frac{(x-m)^2}{2\sigma_X^2} \right] \quad (I.19)$$

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 $x(t) = f(x(t))$ 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

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

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 [16].

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 [17].

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 :

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

Avec E désigne l'espérance mathématique

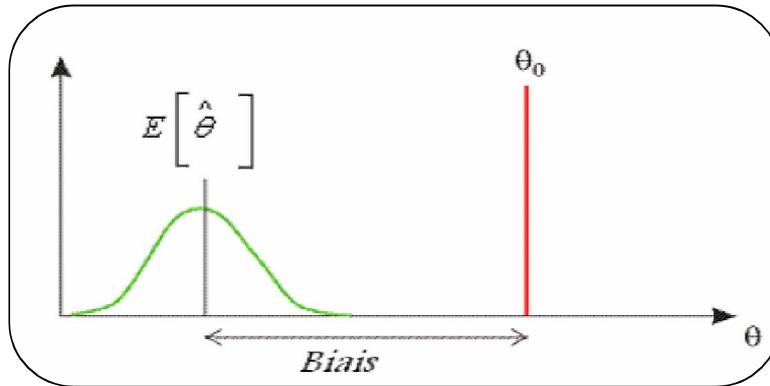


Fig. I.2: Estimateur biaisé de faible variance

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 :

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

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 $\hat{\theta}$:

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

Nous pouvons facilement montrer que :

$$EQM = \text{Var}(\hat{\theta}) - (\text{biais}(\hat{\theta}))^2 \quad (\text{I.22})$$

Cette expression explique pourquoi un estimateur sans biais n'est pas toujours systématiquement préférable à un estimateur biaisé, mais de faible variance. En pratique,

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

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).

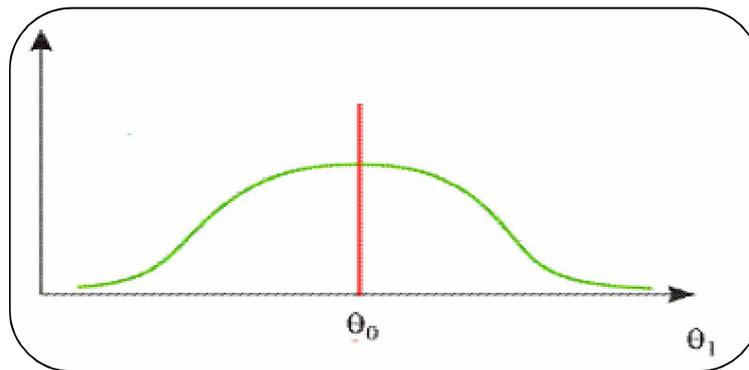


Fig. I.3: Estimateur non biaisé de forte variance

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

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

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 [18].

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 [12]. 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 [19]. 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 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

CHAPITRE I NOTIONS SUR LES PROCESSUS STOCHATIQUE ET L'ESTIMATION

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 :
MODELISATION 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 [20,21] :

- 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 [22] :

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 Autorégressive 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 [23,24] :

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

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 :

Chapitre II MODELISATION ET IDENTIFICATION DES PROCESSUS PHYSIQUES

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

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 Autorégressive (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'équation (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 [20] :

II.2.3.1 Système

Un système (ou 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 [25]. 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 » [26]. Le choix du critère doit traduire le but fixé à la modélisation quel que soit le critère choisi, il convient ensuite de l'optimiser.

II.2.3.4 Optimisation

L'algorithme d'optimisation [27] reçoit toutes 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 $\hat{\theta}$ de 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

Chapitre II MODELISATION ET IDENTIFICATION DES PROCESSUS PHYSIQUES

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 [21,23].

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'à 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'autocorrélation 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 à 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.

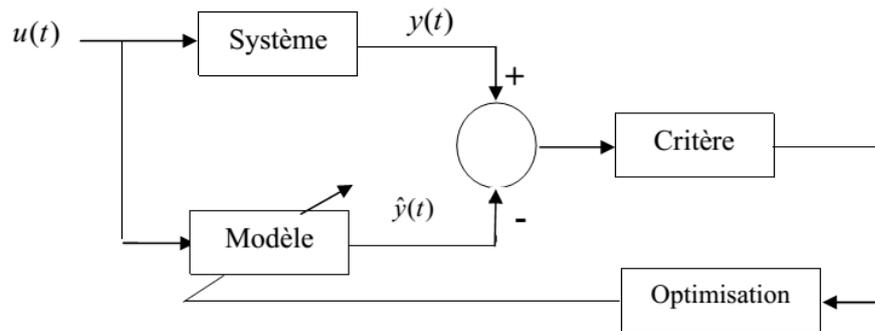


Fig. II.1: Schéma d'identification paramétrique

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 [21,24] :

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 [21].

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

Chapitre II MODELISATION ET IDENTIFICATION DES PROCESSUS PHYSIQUES

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 :

*ETUDES 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 ce chapitre, nous nous sommes intéressés plus particulièrement, à l'identification paramétrique par les algorithmes suivants :

- Algorithme du Gradient Stochastique (GS)
- Algorithme de Newton Stochastique (NS)
- Algorithme des Moindres Carrés Récursif (MCR)
- Algorithme de Maximum de Vraisemblance (MVR)

III.2 Algorithme du gradient stochastique

L'algorithme GS 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 [23], [28]. L'objectif

de l'algorithme du gradient stochastique est de minimiser un critère quadratique en termes 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 :

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

Avec μ représente le pas d'adaptation.

III.2.2 Identification du modèle ARMA

La modélisation autorégressive à moyenne ajustée d'ordre (n, m) noté ARMA (n, m) est définie par l'équation aux différences suivante [20], [29] :

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

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 :

$$y(t) = -a_1 y(t - 1) - a_2 y(t - 2) - \dots - a_n y(t - n) + b_0 u(t) + b_1 u(t - 1) + \dots + b_m u(t - m) + e(t) \tag{III.3}$$

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

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

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 :

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

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, il faut que :

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

III.2.3 Les avantages de la méthode du gradient

La méthode du gradient est simple et facile à mettre en œuvre [30]. 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 tels que [31,32] :

- L'algorithme GS est très simple
- Les performances du GS 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 GS converge lentement mais l'erreur quadratique moyenne excédentaire est petite
- Avec un pas d'adaptation grande, le GS converge rapidement mais l'erreur quadratique moyenne excédentaire est grande
- Le temps de convergence de l'algorithme GS 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 [33]. 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 [33]. 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.

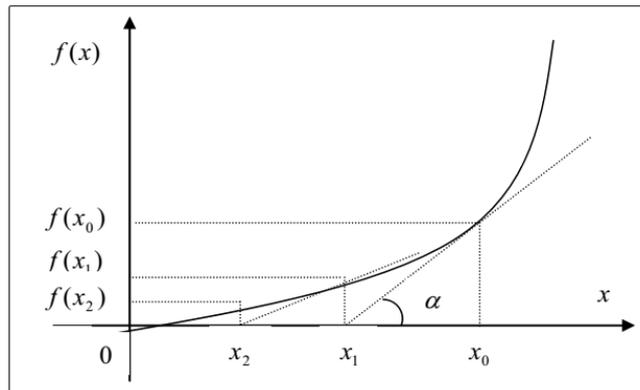


Fig. III.1 : Représentation graphique de la méthode de Newton

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'écrit :

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

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 :

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

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

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

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

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 **MCR**.

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 [34] :

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

Où

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

N : Nombre d'échantillons

III.4.1 Identification du modèle ARMA

La modélisation autorégressive 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 :

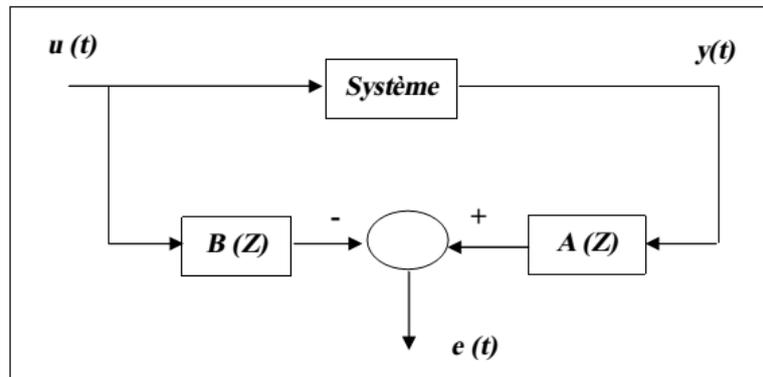


Fig.III.2: Structure de réalisation d'un modèle ARMA

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 [35], [36] d'après l'équation (III.3), une forme matricielle :

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

Avec

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

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

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

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

Sachant que :

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

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 [37]. 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 :

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

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

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 :

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

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 (MCR)

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 :

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

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

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

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

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

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 [23] :

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

Nous posons :

$$A = R(t-1),$$

$$B = \varphi(t),$$

$$C = 1,$$

$$D = \varphi^T(t).$$

Or

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

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 MCR de la forme suivante :

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

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

III.4.4 Propriétés de l'estimateur des moindres carrés

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

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

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

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

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

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 de Maximum de vraisemblance (MVR)

III.5.1 Identification du modèle ARMA

Considérons le modèle ARMA suivante :

$$\begin{aligned} x(t) &= \frac{B(t)}{A(t)} u(t); \\ y(t) &= x(t) + w(t); \\ Z(t) &= Q_t[y(t)]. \end{aligned} \quad (\text{III.49})$$

avec :

$u(t)$ est le signal d'entrée modélisé par un signal déterministe ou un processus aléatoire $A(z)$ et la partie autorégressive (AR) :

$$A(z) = \sum_{i=1}^n a_i z^{-i} \text{ avec } (a_0 = 1) ;$$

$B(z)$ et la partie moyenne ajusté (MA):

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

$w(t)$ est une séquence indépendantes et identiquement distribuées (IID) de distribution normale $N(0, \sigma^2)$. La sortie bruyante $y(t)$ est quantifiée avec un quantificateur scalaire Q_t de K niveau variant dans le temps t

$$Q_t: R \rightarrow \{v_{t,1}, \dots, v_{t,k}\}, t \in Z$$

Défini par des pas de quantification :

$$[b_{t,k-1}, b_{t,k}] = Q^{-1}[v_{t,k}] ; k = 1, \dots, K.$$

Avec $b_{t,0} = -\infty$ et $b_{t,k} = +\infty$, pour $t \in Z$.

Notre objectif est d'estimer θ_* , avec la connaissance de U_N et Z_N . Nous le faisons en utilisant le critère du maximum de vraisemblance (MVR).

III.5.2 Maximum de vraisemblance

En utilisant le critère MVR [40] [41], la valeur estimé $\hat{\theta}_N$ à temps N est obtenue par

$$\hat{\theta}_N \in \arg \max_{\theta} p_{\theta} (Z_N | U_N) = \arg \max_{\theta} l(\theta | U_N, Z_N), \quad (\text{III.50})$$

avec p_{θ} désigne la distribution de probabilité donnée, et

$$l(\theta | U_N, Z_N) = \log p_{\theta} (Z_N | U_N) \quad (\text{III.51})$$

est la fonction de log-likelihood de θ qui nous donne U_N et Z_N .

Pour résoudre (III.50), on utilise un algorithme récursif (on-line algorithme) basée sur la méthode Expectation Maximisation (EM). L'avantage de cette méthode est que l'on peut dériver un algorithme qui ne nécessite pas une initialisation. Cependant, la méthode EM est connu d'essuyer d'un faible taux de convergence [42]. A cette, l'estimation obtenue après quelques itérations EM peut être utilisé pour initialiser une méthode de quasi-Newton pour recherche de gradient. Ainsi, on utilise un algorithme récursif basé sur la méthode quasi-Newton. Le but de l'utilisation de la méthode EM est de fournir une bonne estimation initiale pour la méthode quasi-Newton.

III.5.2.1 La méthode d'Expectation Maximisation (EM) récursive

Avant de dériver la méthode requise, nous donnons une brève explication intuitive de la méthode EM. Pour une présentation détaillée, voir [41]. Un inconvénient de (III.50) est que $\log p_\theta(Z_N)$ est difficile à maximiser par rapport à θ en raison de l'induction d'un quantificateur non linéaire.

Supposons que la sortie $Y_N = \{y(t) : t = 1, \dots, N\}$ avant le quantificateur était connue. Notez que ce serait beaucoup plus facile de maximiser $\log p_\theta(Y_N)$. tell que Y_N est indisponible, nous remplaçons $\log p_\theta(Z_N)$ par la moyenne de $\log p_\theta(Z_N, Y_N) = \log p_\theta(Z_N|Y_N) + \log p_\theta(Y_N)$ sur toutes les valeurs possibles de Y_N . Pour ce faire, nous utilisons la moyenne conditionnelle de distribution $p_{\hat{\theta}}(Y_N|Z_N)$ de Y_N compte tenu des observations Z_N , et des estimations précédente $\hat{\theta}$ de θ . Cela conduit à la méthode EM, ce qui résout le problème de MVR (III.50) en utilisant la procédure itérative suivante :

$$\hat{\theta}_N^{(i)} \in \arg \max_{\theta} Q_N(\theta, \theta_N^{(i-1)}), \quad (III.52)$$

$$Q_N(\theta, \hat{\theta}) = \int \log p_\theta(Z_N, Y_N) p_{\hat{\theta}}(Y_N|Z_N) dY_N. \quad (III.53)$$

Les itérations (III.52) et (III.53) permettent de calculer $\hat{\theta}_N$, pour un nombre fixe de N . Pour obtenir un algorithme récursif, nous pouvons calculer une itération pour chaque nouvel échantillon disponible. Ce faisant, nous obtenons l'algorithme itératif suivant :

$$\hat{\theta}_N \in \arg \max_{\theta} Q_N(\theta, \hat{\theta}_{N-1}). \quad (III.54)$$

Lemme 1

la fonction $Q_N(\dots)$ dans (III.53) est donnée par :

$$Q_N(\theta, \hat{\theta}) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^N (\bar{y}(t, \hat{\theta}) - x(t, \theta))^2 - \frac{1}{2\sigma^2} \sum_{t=1}^N (\overline{y^2}(t, \hat{\theta}) - \bar{y}^2(t, \hat{\theta})),$$

Où

$$\overline{y^2}(t, \theta) = \mathcal{E}_\theta\{y^2(t)|z(t)\}$$

Et

$$\bar{y}(t, \theta) = \mathcal{E}_\theta\{y(t)|z(t)\}. \quad (III.55)$$

Note :

Si a et b sont des variables aléatoires, et $f(\cdot)$ est une fonction, expectation conditionnel $\mathcal{E}_\theta\{f(a)|b\}$ de $f(a)$ qui nous donne b est définie par :

$$\mathcal{E}_\theta\{f(a)|b\} = \int f(a)p_\theta\{a|b\} da.$$

Remarque : Notez que (III.55) peut être calculé en :

$$\begin{aligned} \bar{y}(t, \theta) &= \int_{Q_t^{-1}[z(t)]} y(t)p_\theta(y(t))dy(t) \\ &= \frac{1}{2} \left[\operatorname{erf} \left(\frac{b(t)-x(t,\theta)}{\sqrt{2\sigma^2}} \right) - \operatorname{erf} \left(\frac{a(t)-x(t,\theta)}{\sqrt{2\sigma^2}} \right) \right], \end{aligned}$$

Où

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \text{ désigne la fonction d'erreur}$$

Et $[a(t), b(t)] = Q_t^{-1}[z(t)]$ désigne la pas de quantification correspondant à (t) .

En combinant (III.53) avec le lemme 1, on obtient que :

$$\hat{\theta}_N \in \operatorname{arg} \min_{\theta} \sum_{t=1}^N (\bar{y}(t, \hat{\theta}_{N-1}) - x(t, \theta))^2. \quad (\text{III.56})$$

$$\hat{\theta}_N \in \operatorname{arg} \min_{\theta} \sum_{t=1}^N \left(A(z, \theta) \frac{\bar{y}(t, \hat{\theta}_{N-1})}{A(z, \hat{\theta}_{N-1})} - B(z, \theta) \frac{u(t)}{A(z, \hat{\theta}_{N-1})} \right)^2.$$

Après, on calcule $\hat{\theta}_N$ comme suivant :

$$\hat{\theta}_N = \left(\sum_{t=1}^N \tilde{\phi}(t, \hat{\theta}_{N-1}) \tilde{\phi}^T(t, \hat{\theta}_{N-1}) \right)^\dagger \times \left(\sum_{t=1}^N \tilde{\phi}(t, \hat{\theta}_{N-1}) \tilde{y}(t, \hat{\theta}_{N-1}) \right), \quad (\text{III.57})$$

Où l'exposant de \dagger désigne le pseudo-inverse Moore-Penrose, et

$$\tilde{y}(t, \theta) = \frac{\bar{y}(t, \theta)}{A(z, \theta)}, \quad (\text{III.58})$$

$$\tilde{\phi}(t, \theta) = \frac{1}{A(z, \theta)} [u(t), \dots, u(t-n), -\bar{y}(t-1, \theta), \dots, -\bar{y}(t-m, \theta)]^T. \quad (\text{III.59})$$

En suivant les étapes décrites dans [43, Sec. 11], (III.8) peut être écrite sous une forme récursive.

$$\hat{\theta}_N = \hat{\theta}_{N-1} + L_N (\tilde{y}(N, \hat{\theta}_{N-1}) - \tilde{\phi}^T(N, \hat{\theta}_{N-1}) \hat{\theta}_{N-1}), \quad (\text{III.60})$$

Avec

$$L_N = \frac{P_{N-1} \tilde{\phi}(N, \hat{\theta}_{N-1})}{1 + \tilde{\phi}^T(N, \hat{\theta}_{N-1}) P_{N-1} \tilde{\phi}(N, \hat{\theta}_{N-1})} \quad (\text{III.61})$$

$$P_N = P_{N-1} - \frac{P_{N-1} \tilde{\phi}(N, \hat{\theta}_{N-1}) \tilde{\phi}^T(N, \hat{\theta}_{N-1}) P_{N-1}}{1 + \tilde{\phi}^T(N, \hat{\theta}_{N-1}) P_{N-1} \tilde{\phi}(N, \hat{\theta}_{N-1})}. \quad (\text{III.62})$$

III.5.2.2 La méthode de Quasi-Newton récursive

Pour une nombre fixe de N , le problème MVR (III.52) peut être résolu en utilisant un algorithme de recherche de gradient. A cet effet, nous utilisons quasi-Newton méthode. Comme ci-dessus, pour obtenir un algorithme récursif, nous calculons une itération quasi-Newton pour chaque nouvel échantillon disponible. Cela donne les itérations suivantes :

$$\hat{\theta}_{N+1} = \hat{\theta}_N - \mu_N T_N g_N, \quad (\text{III.63})$$

où le scalaire μ_N désigne le pas de taille à l'itération N , la matrice T_N dénote une approximation de l'inverse de la Hessien de $l(\theta|Z_N)$ à $\hat{\theta}_N$, et le vecteur g_N désigne le gradient de $l(\theta|Z_N)$ à $\hat{\theta}_N$, c'est à dire ,

$$g_N = \left. \frac{\partial}{\partial \theta} l(\theta|Z_N) \right|_{\hat{\theta}_N}. \quad (\text{III.64})$$

Nous avons choisi de calculer T_N en utilisant la formule de **Broyden-Fletcher-Goldfarb-Shanno (BFGS)** [27], qui est initialisé par $T_{N_0} = I$ (N_0 désigne le premier échantillon après les itérations EM passer à itérations quasi-Newton) et se déroule comme suit :

$$T_{N+1} = T_N + \left(1 + \frac{q_N^T T_N q_N}{s_N^T q_N} \right) \frac{s_N s_N^T}{s_N^T q_N} - \frac{s_N q_N^T T_N + T_N q_N s_N^T}{s_N^T q_N}, \quad (\text{III.65})$$

Avec

$$s_N = \theta_{N+1} - \theta_N,$$

Et

$$q_N = g_{N+1} - g_N.$$

De plus, le paramètre ; μ_N est obtenue à partir d'un algorithme de recherche récursive. Nous utilisons l'algorithme de retour arrière est décrit dans [43], qui est formée par les sous- itérations des itérations principales (III.63). Soit $\alpha = 0.01$ et $\beta = 0,5$. Au sous- itération, le pas μ_N est

mis à jour en utilisant $\mu_{N,i} = \beta \mu_{N,i-1}$, à partir de la valeur initiale de $\mu_{N,i} = 1$. Les sous-itérations sont arrêtées lorsque

$$l(\hat{\theta}_N - \mu_{N,i} T_N g_N | Z_N) < l(\hat{\theta}_N | Z_N) + \alpha \mu_{N,i} g_N^T s_N. \quad (\text{III.66})$$

Pour implémenter les itérations (III.63), nous devons fournir des expressions pour $l(\theta | Z_N)$ et son gradient ;

$$l(\theta | Z_N) = \sum_{t=1}^N \log p_{\theta}(z(t)). \quad (\text{III.67})$$

Lemme 2

Le gradient de la fonction de log-Vraisemblance $l(\theta | Z_N)$ est donnée par :

$$\frac{\partial}{\partial \theta} l(\theta | Z_N) = \frac{1}{\sigma^2} \sum_{t=1}^N (\bar{y}(t, \theta) - x(t, \theta)) \dot{x}(t, \theta),$$

Où

$$\dot{x}(t, \theta) \triangleq \frac{\partial}{\partial \theta} x(t, \theta) = \phi(z, \theta) u(t), \quad (\text{III.68})$$

$$\phi(z, \theta) = \left[\frac{\Omega_n^T(z)}{A(z, \theta)}, \frac{z^{-1} B(z, \theta) \Omega_{m-1}^T(z)}{A^2(z, \theta)} \right]^T, \quad (\text{III.69})$$

et $\Omega_n(z) = [1, z^{-1}, \dots, z^{-n}]^T$.

III.5.3 Sommaire de l'algorithme d'Identification

L'algorithme d'identification décrit ci-dessus peut être résumé dans ce paragraphe. Choisissez une θ_0 estimation initiale de sorte que $(z, \hat{\theta}_0) = A(z, \hat{\theta}_0) = 1$, et choisir un nombre N_0 pour que l'algorithme d'estimation récursive passe de la méthode fondée sur EM à la méthode fondée sur quasi-Newton. Puis :

Si $N < N_0$:

1. Utilisation de (III.55) pour calculer $\bar{y}(N, \hat{\theta}_{N-1})$;
2. Utilisation de (III.58) et (III.59) pour calculer $\tilde{y}(N, \hat{\theta}_{N-1})$ et $\tilde{\phi}(N, \hat{\theta}_{N-1})$;
3. Utilisation de (III.60), (III.61) et (III.62) pour calculer $\hat{\theta}_N$.

Sinon, si $\geq N_0$, définissons la valeur initiale des itérations (III.65) comme $T_{N_0} = I$. Puis :

1. Usage (III.64) et le lemme 2 pour calculer le de gradient g_N ;
2. Usage (III.66) et (III.67) pour calculer μ_N ;
3. Usage (III.63) calculer $\hat{\theta}_N$;
4. Usage (III.65) calculer T_N .

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.

Dans le chapitre suivant, nous examinons deux méthode parmi ces méthode MVR et MCR, par différentes exemples de simulation, avec de modèle ARMA stable à phase minimale et à phase minimale,..., pour tester la performance de chaque algorithme.

Chapitre IV :
SIMULATIONS ET
INTERPRETATION

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 moindres carrés récursif (MCR) et la méthode de maximum de vraisemblance (MVR) d'une part et d'autre part une étude comparative entre ces algorithmes pour envisager les avantages et les inconvénients de ces algorithmes.

Les simulations sont faites sous le logiciel **MATLAB** (ver 7.10.0).

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 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 comme :

➤ **Séquence Binaire Pseudo Aléatoire (SBPA) :**

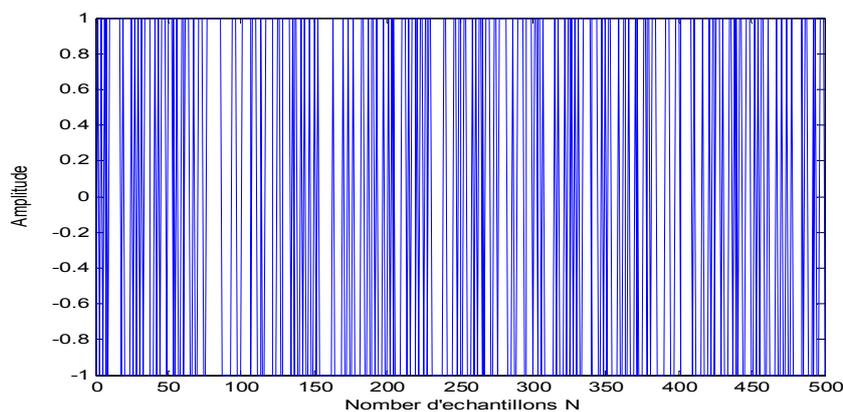


Fig.IV.1 Séquence d'entrée SBPA.

Cette figure la, précédent, nous montre une Séquence Binaire Pseudo Aléatoire, qui est une comme un signal d'entrée tris riche en fréquence.

➤ **Bruit additif Gaussien :**

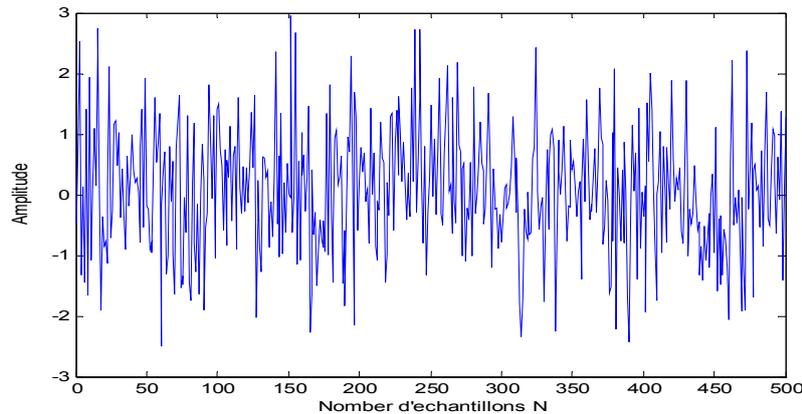


Fig.IV.2 Bruit additif gaussien .

Cette figure représente un bruit additif Gaussien. Ce dernier est généré par la fonction '**randn**' du logiciel Matlab.

IV.2.2 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 soit stable, que ce soit le système à phase minimale ou non minimale. A titre d'exemple de simulation, on prend :

➤ **système à phase minimale :**

$$H_1(z) = \frac{1+0.62 z^{-1}}{1+0.73z^{-1}-0.24 z^{-2}} \quad (\text{IV.1})$$

$$B_i = [1 \ 0.62], \ b_0 = 0 ; \ A_i = [0.73 \ -0.24], \ a_1 = 1.$$

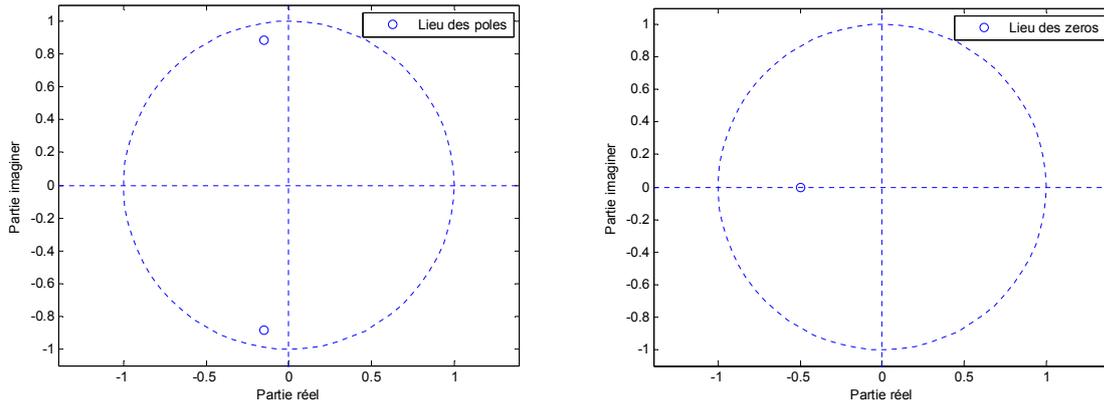


Fig.IV.3 Lieu des pôles (à gauche) et lieu des zéros (à droite).

Cette figure montre que notre système est stable et à phase minimale (les pôles et les zéros sont toujours à l'intérieur du cercle unité).

➤ système à phase non minimale :

$$H_2(z) = \frac{1+1.2z^{-1}}{1+0.89z^{-1}+0.45z^{-2}} \tag{IV.2}$$

$B_i = [1 \ 1.2]$, $b_0 = 0$; $A_i = [0.89 \ 0.45]$, $a_1 = 1$.

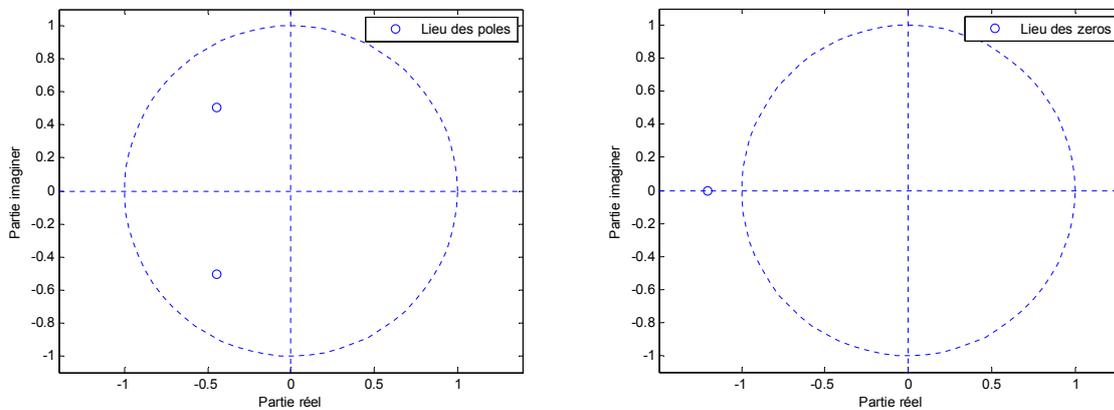


Fig.IV.4 Lieu des pôles (à gauche) et lieu des zéros (à droite)

Cette figure montre que système est stable et à phase non minimale (les pôles sont toujours à l'intérieur du cercle unité mais les pôles non).

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.3 Implémentation des algorithmes

Dans cette partie nous implémentons les algorithmes d'identification paramétrique (moindre carrée récurrente 'MCR' et Maximum vraisemblance 'MVR') sur MATLAB [44,45], on a choisi deux modèles de tests à phase minimale et à phase non minimale comme on a dit, aussi sans observations quantifiées et avec observation quantifiées de quantificateur scalaire un bit, deux bits....

On a comparé la performance entre les deux méthodes (MCR et MVR) pour les deux systèmes sans observations quantifiées.

Enfin, un ensemble de représentations graphiques permet de visualiser l'amplitude des deux systèmes, l'original et l'estimé ainsi que l'évolution temporelle des paramètres estimés du modèle, l'erreur d'identification et la fonction d'autocorrélation de l'erreur de prédiction aussi teste de stabilité et de phase.

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

Les résultats de simulation obtenus par l'algorithme MCR 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 :

$$y(t) = -a_1y(t-1) - a_2y(t-2) + b_1u(t-1) + b_2u(t-2) + e(t) \quad (\text{IV.3})$$

Avec :

$$a_1 = 0.73, a_2 = -0.24, b_1 = 1 \text{ et } b_2 = 0.62.$$

IV.4.1.1. Influence du nombre d'échantillons et de la variance

L'influence du nombre d'échantillons N et la variance σ^2 du bruit additif sur le biais et l'Ecart Quadratique Moyenne (EQM) d'estimateur (MCR), sont données dans les tableaux IV.1 et IV.2, suivi par les spectres d'amplitude de système initial et le système estimé presque identiques (Fig.IV.5), puis l'évolution des paramètres du système en fonction du temps (Fig.IV.6), ainsi que l'autocorrélation de l'erreur de prédiction est considérée asymptotiquement blanche (Fig.IV.7). On note que l'erreur d'identification converge vers zéro telle qu'elle est montrée sur la figure (Fig.IV.8). En fin un test de stabilité et de phase (Fig.IV.9).

Tab. IV.1 Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).

Nombre d'échantillons	Paramètres estimés (MCR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.7059	-0.2691	0.9557	0.5933	0.1069	0.0342
512	0.7171	-0.2478	0.9726	0.6391	0.0591	0.0172
1024	0.7203	-0.2522	1.0055	0.6122	0.0354	0.0087
2048	0.7293	-0.2390	1.0018	0.6243	0.0195	0.0044
4096	0.7269	-0.2430	1.0016	0.6138	0.0121	0.0022

Tab. IV.2 Influence de la variance σ^2 , ($N = 1024$).

Variance σ^2	Paramètres estimés (MCR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.6984	-0.2797	1.0239	0.5983	0.0655	0.0156
0.25	0.7203	-0.2522	1.0055	0.6122	0.0354	0.0087
0.04	0.7254	-0.2420	1.0001	0.6152	0.0216	0.0061
0.01	0.7271	-0.2415	0.9993	0.6166	0.0186	0.0059

Les résultats obtenus montrent que lorsque le nombre d'échantillons N est augmenté (ou la variance du bruit additif diminue) ; alors le biais de l'estimateur diminue (même que pour l'EQM), les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est tend vers un bruit blanc.

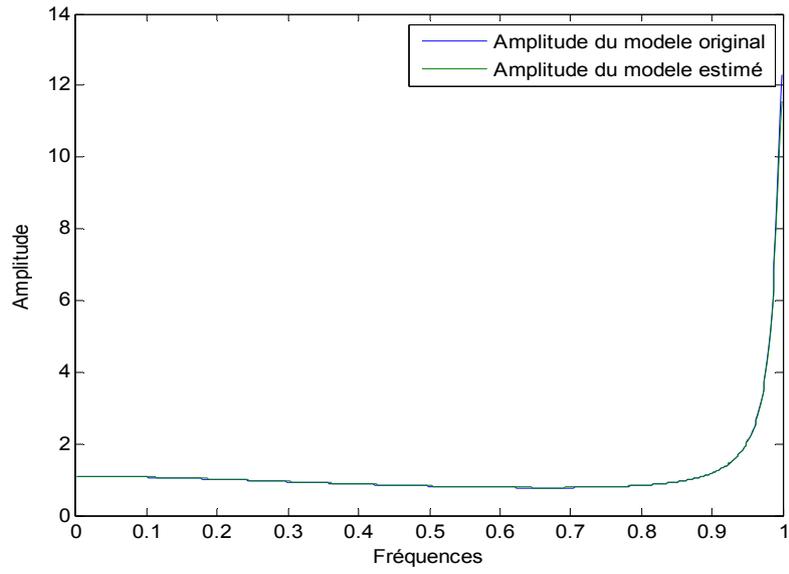


Fig.IV.5 Spectres d'amplitude du modèle original et du modèle estimé.

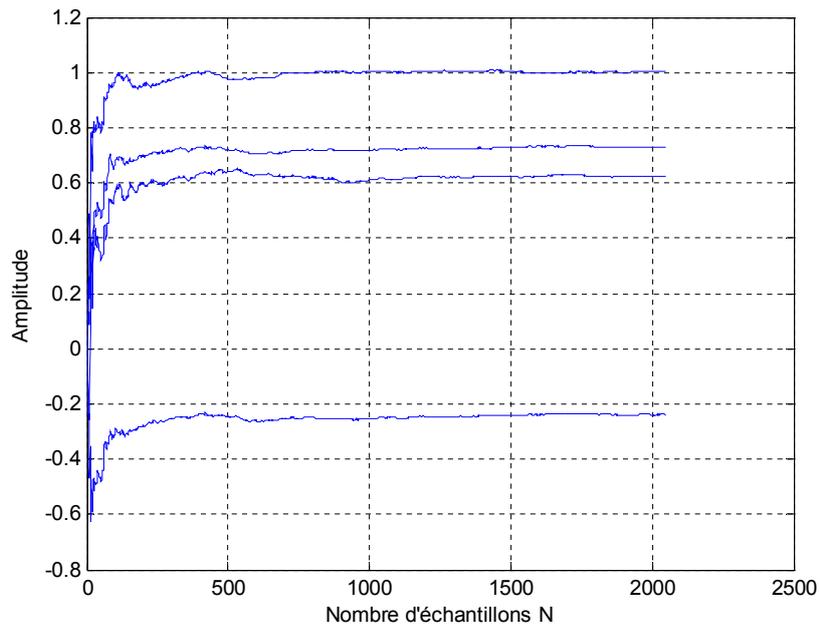


Fig.IV.6 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

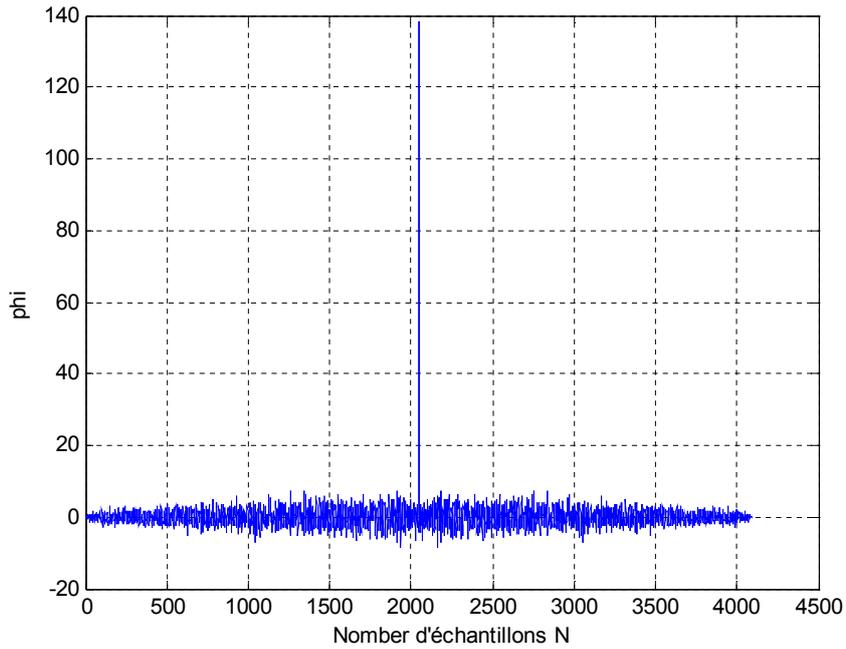
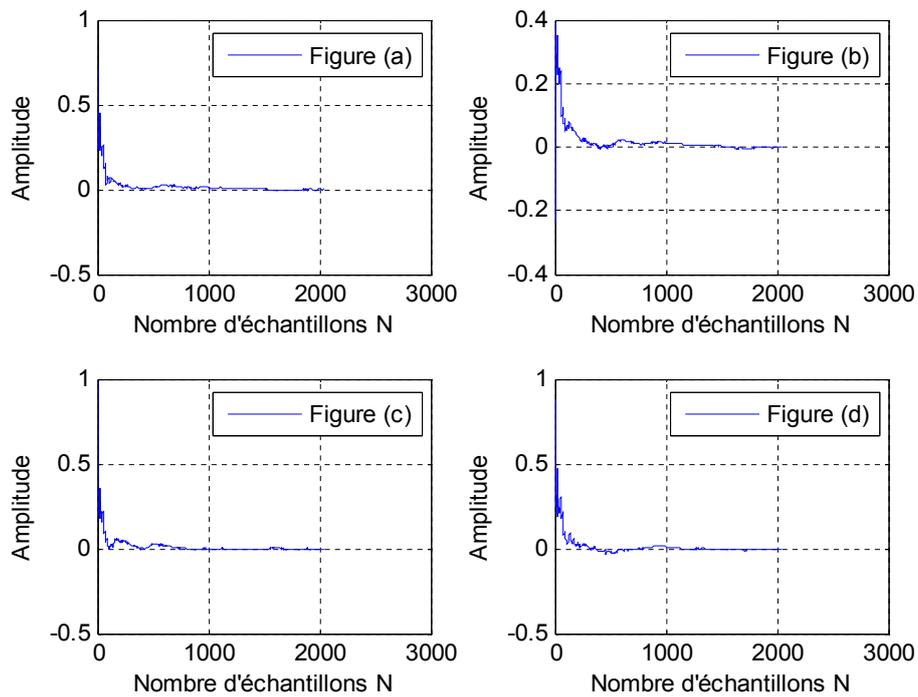
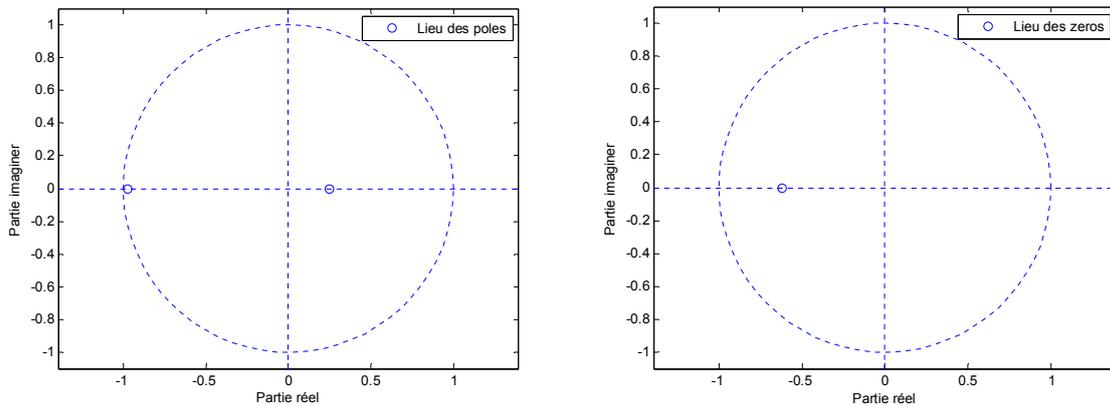


Fig.IV.7 – Fonction d'autocorrélation d'erreur de prédiction.



**Fig.IV.8 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).**



**Fig.IV.9 - Lieu des pôles du modèle estimé (à gauche),
Lieu des zéros du modèle estimé (à droite).**

IV.4.2. Système à phase non minimale

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

$$y(t) = -a_1y(t-1) - a_2y(t-2) + b_1u(t-1) + b_2u(t-1) + e(t) \quad (\text{IV.4})$$

Avec :

$$a_1 = 0.89, a_2 = 0.45, b_1 = 1 \text{ et } b_2 = 1.2.$$

IV.4.2.1. Influence du nombre d'échantillons et de la variance

Le tableau IV.3 montre les résultats de simulation obtenus lorsque nous faisons varier le nombre d'échantillons N . Par contre le tableau IV.4 est réservé pour les résultats obtenus en fonction de la variance σ^2 avec N constant. Nous testons son influence sur le biais et l'Ecart Quadratique Moyenne (**EQM**) d'estimateur (MCR), des représentations graphique montre L'évolution des paramètres du système ainsi que l'autocorrélation de l'erreur de prédiction sont illustré sur les figure (Fig.IV.11) et (Fig.IV.12), aussi que pour les spectres d'amplitude de deux modèles (Fig.IV.10), l'erreur d'identification de chaque paramètre (Fig.IV.13), en fin lieu des pôles et zéros (Fig.IV.14).

Tab. IV.3 Influence du nombre d'échantillons N , ($\sigma^2 = 0.25$).

Nombre d'échantillons	Paramètres estimés (MCR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.2528	0.4099	0.9553	1.1586	0.1121	0.0508
512	0.8687	0.4222	0.9723	1.2090	0.0688	0.0258
1024	0.8818	0.4315	1.0053	1.1931	0.0415	0.0130
2048	0.8929	0.4435	1.0016	1.2078	0.0233	0.0065
4096	0.8851	0.4465	1.0016	1.1918	0.0143	0.0033

Tab. IV.4 Influence de la variance σ^2 , ($N = 1024$).

Variance σ^2	Paramètres estimés (MCR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.8492	0.4036	1.0234	1.1683	0.0850	0.0206
0.25	0.8818	0.4315	1.0053	1.1931	0.0415	0.0130
0.04	0.8885	0.4468	1.0001	1.1977	0.0224	0.0110
0.01	0.8879	0.4487	0.9993	1.1968	0.0210	0.0109

Les tableaux ci-dessus montrent que soit le système à phase minimale où non lorsque le nombre d'échantillons N est grand (la variance démunie) alors les paramètres estimés sont plus proches aux valeurs réelles, et le biais (aussi l'EQM) toujours diminue

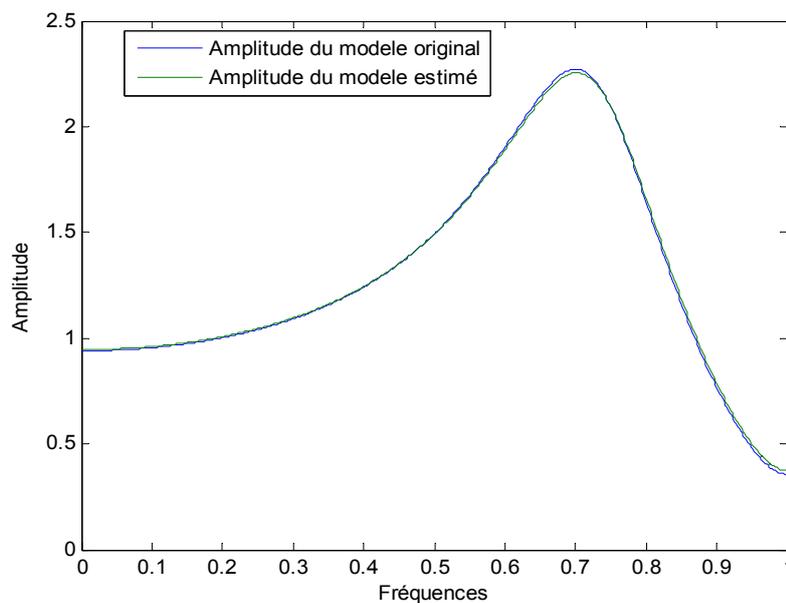


Fig.IV.10 Spectres d'amplitude du modèle original et du modèle estimé.

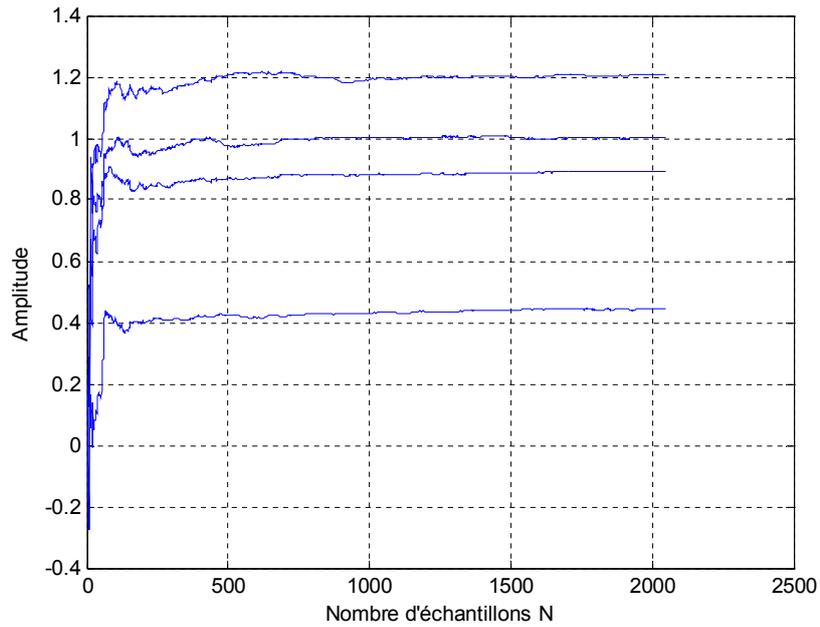


Fig.IV.11 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

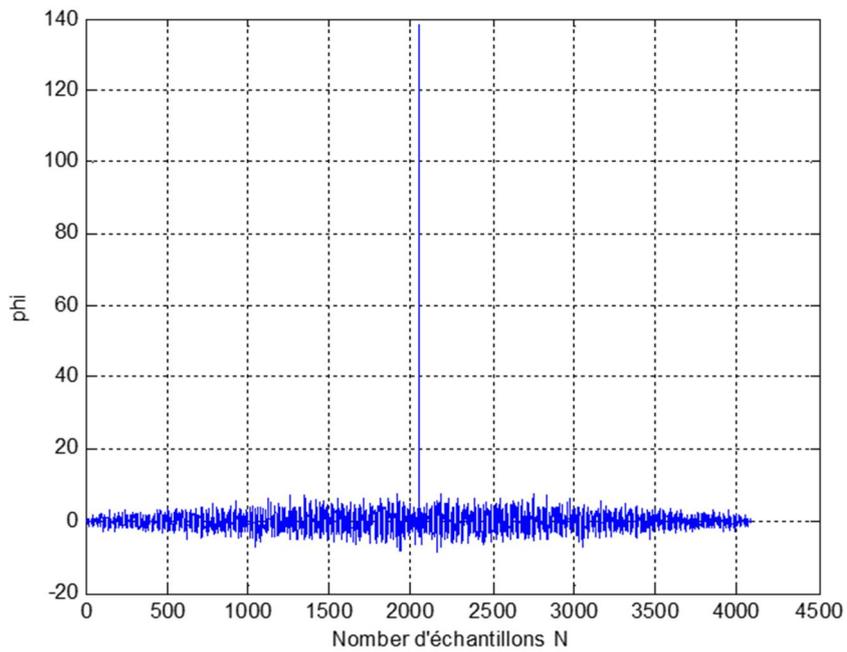
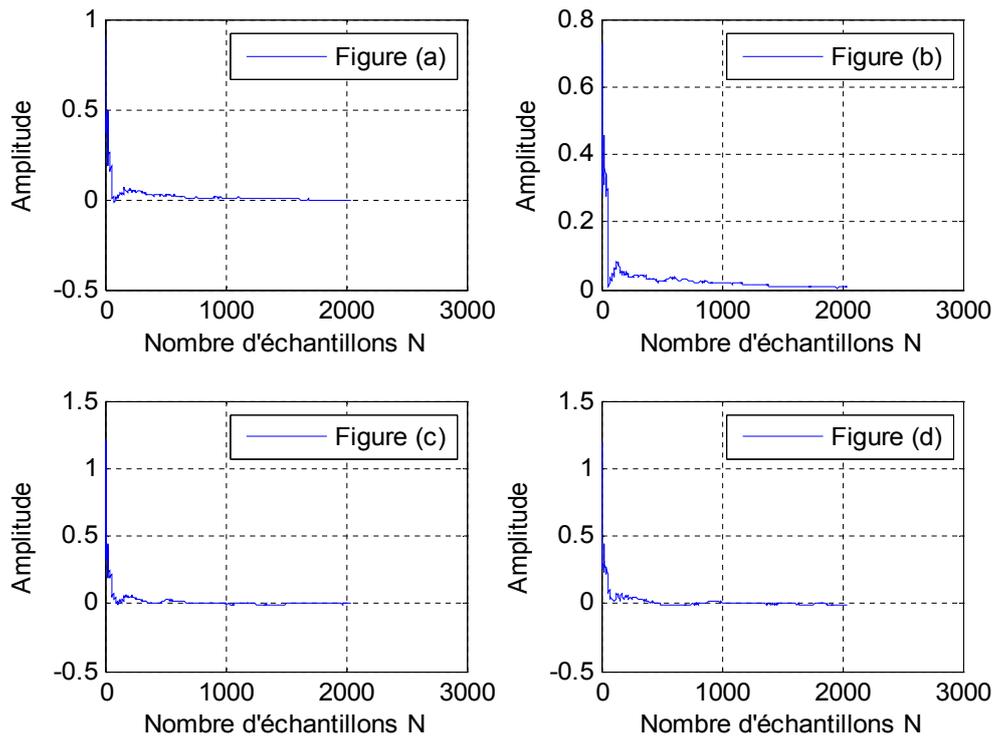
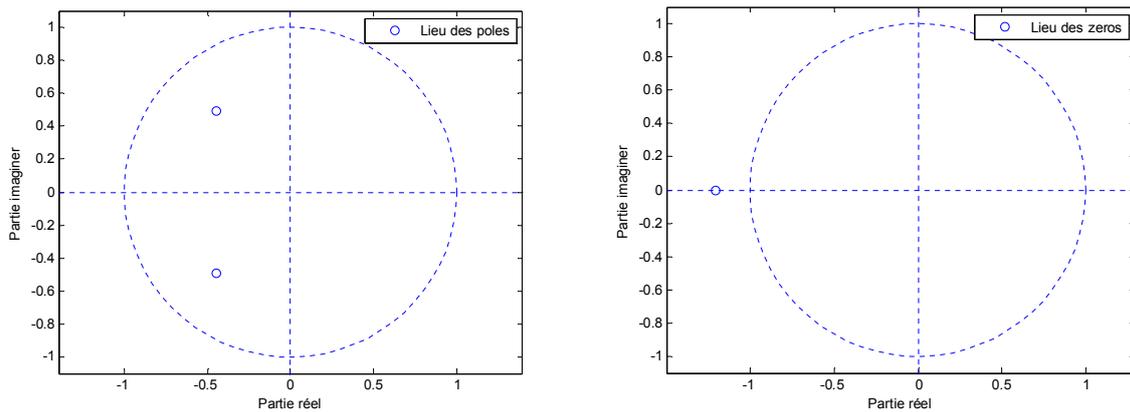


Fig.IV.12 – Fonction d'autocorrélation d'erreur de prédiction.



**Fig.IV.13 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).**



**Fig.IV.14 - Lieu des pôles du modèle estimé (à gauche),
 Lieu des zéros du modèle estimé (à droite).**

IV.5 Identification par Maximum de Vraisemblance (MVR)

Dans la partie précédente (VI.4) nous allons utiliser l'estimateur MCR sur deux exemple, maintenant, on utilise l'estimateur de maximum de vraisemblance avec les même exemples précédents (IV.2.2), Les résultats de simulation sont illustrés sur les tableaux et les figures qui suivent.

IV.5.1 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).

IV.5.1.1 Influence du nombre d'échantillons et de la variance

L'influence du nombre d'échantillons N et la variance σ^2 du bruit additif sur le biais et l'Ecart Quadratique Moyenne (EQM) d'estimateur (MVR), sont données dans les tableaux IV.5 et IV.6, suivi par les spectres d'amplitude de système initial et le système estimé presque identiques (Fig.IV.15), puis l'évolution des paramètres du système en fonction du temps (Fig.IV.16), ainsi que l'autocorrélation de l'erreur de prédiction est considérée asymptotiquement blanche (Fig.IV.17). On note que l'erreur d'identification converge vers zéro telle qu'elle est montrée sur la figure (Fig.IV.18). En fin un test de stabilité et de phase (Fig.IV.19).

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

Nombre d'échantillons	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.6988	-0.2556	0.9745	0.5879	0.0849	0.0278
512	0.7104	-0.2450	0.9854	0.6157	0.0476	0.0140
1024	0.7238	-0.2449	1.0023	0.6143	0.0276	0.0070
2048	0.7272	-0.2393	1.0007	0.6195	0.0149	0.0035
4096	0.7284	-0.2410	1.0007	0.6167	0.0089	0.0018

Tab. IV.6 Influence de la variance σ^2 , ($N = 1024$).

Variance σ^2	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.7094	-0.2667	1.0116	0.6039	0.0501	0.0117
0.25	0.7238	-0.2449	1.0023	0.6143	0.0276	0.0070
0.04	0.7265	-0.2416	0.9995	0.6161	0.0196	0.0060
0.01	0.7274	-0.2414	0.9992	0.6169	0.0181	0.0056

Nous remarquons que si le nombre d'échantillons N est augmenté -où la variance diminue- ; alors le biais de l'estimateur (MVR) toujours diminue -même que pour l'EQM-, les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est tend vers un bruit blanc.

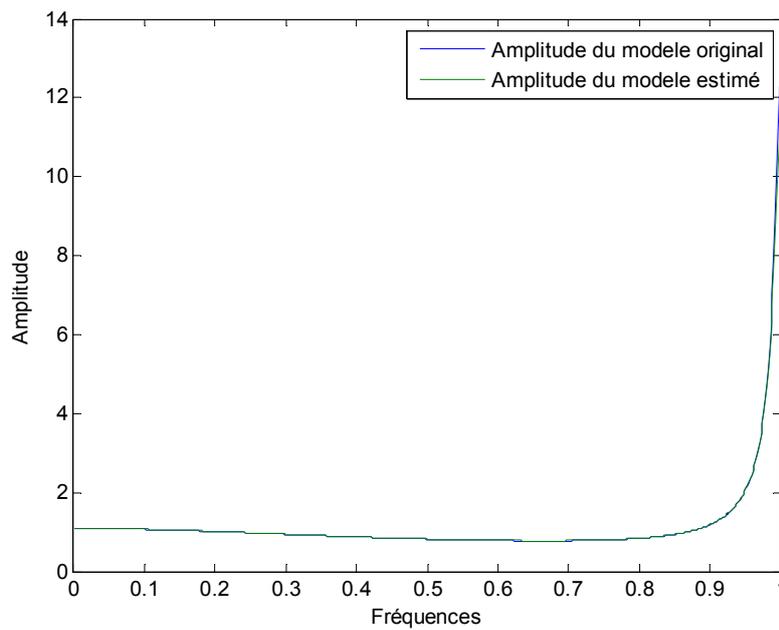


Fig.IV.15 Spectres d'amplitude du modèle original et du modèle estimé.

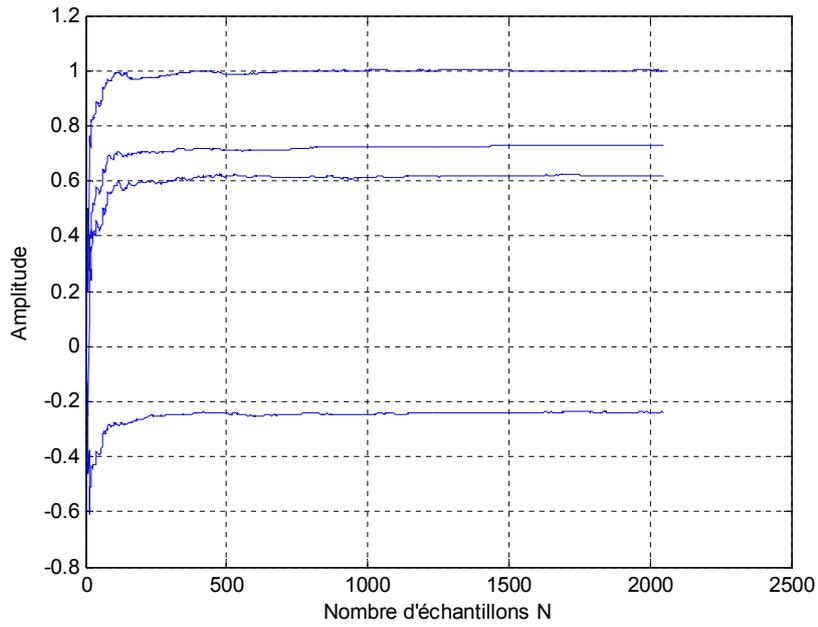


Fig.IV.16 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

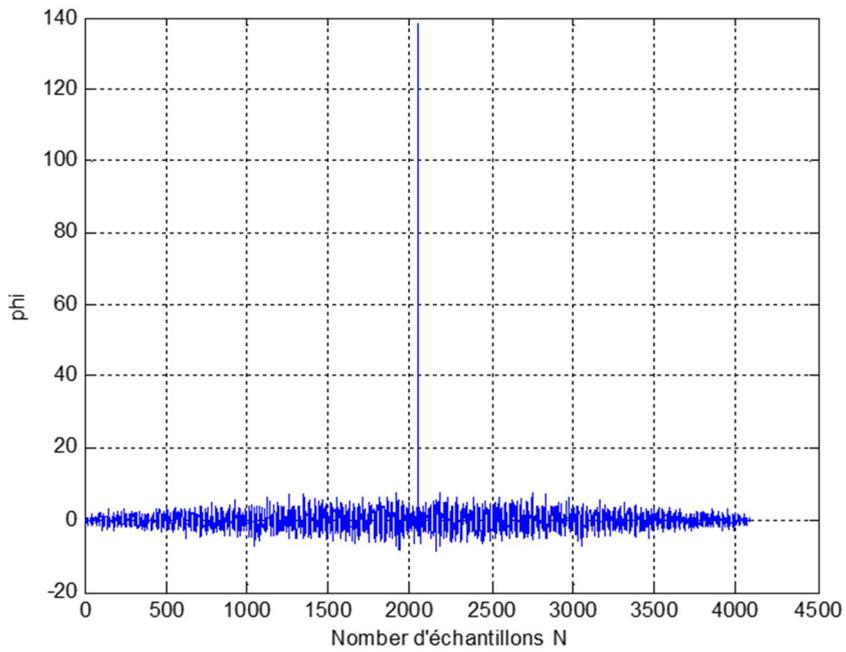
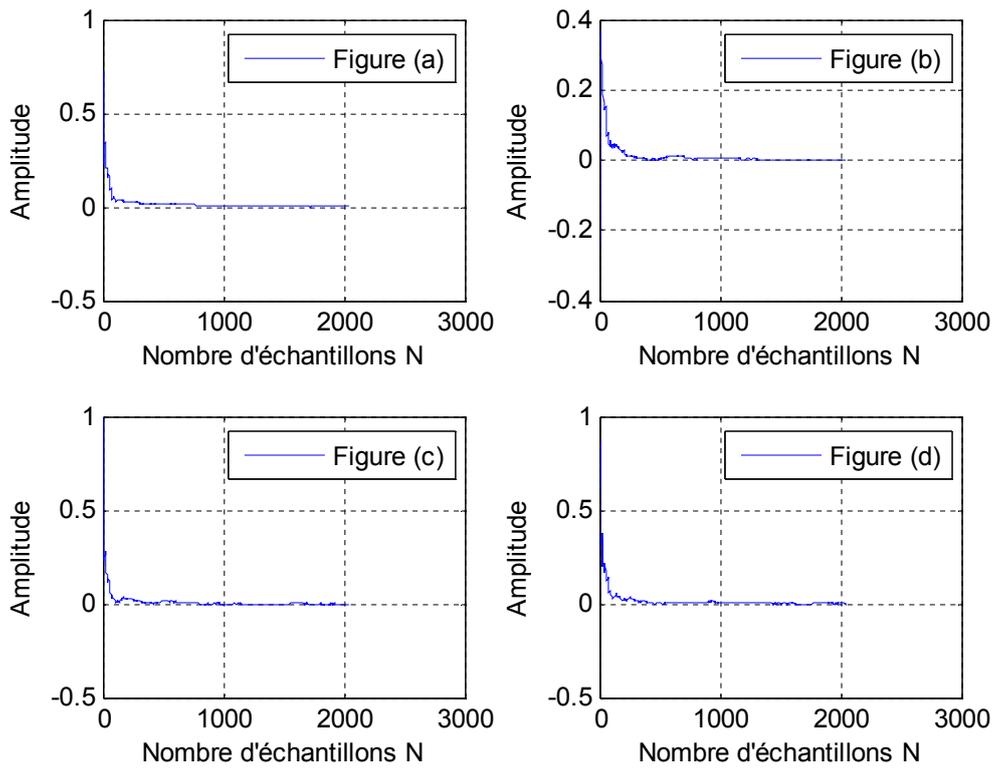
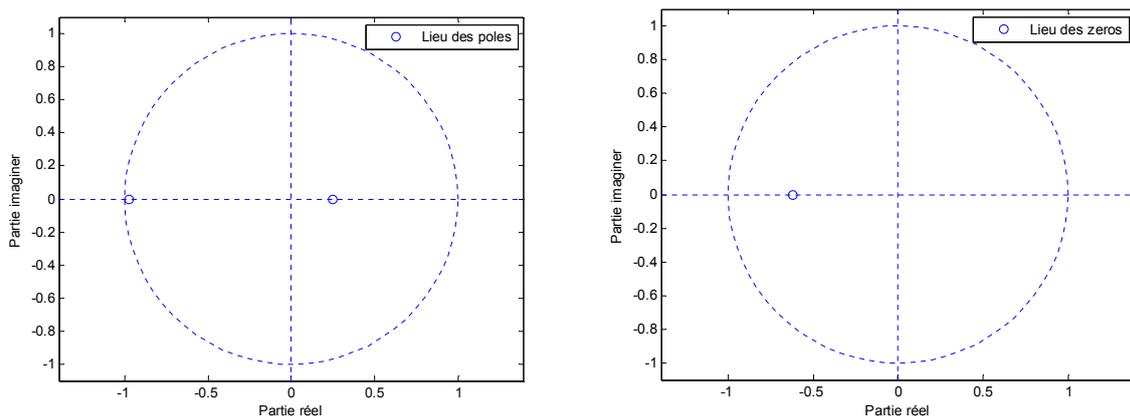


Fig.IV.17 – Fonction d'autocorrélation d'erreur de prédiction.



**Fig.IV.18 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).**



**Fig.IV.19 - Lieu des pôles du modèle estimé (à gauche),
 Lieu des zéros du modèle estimé (à droite).**

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).

IV.5.2.1 Influence du nombre d'échantillons et de la variance

Pour tester l'influence du nombre d'échantillons N sur le biais et l'Ecart Quadratique Moyenne (EQM) d'estimation des paramètres du système, on dresse le tableau ci-dessous suivi par l'évolution des paramètres du système ainsi que l'autocorrélation de l'erreur de prédiction sont illustré sur les figure (Fig.IV.21) et (Fig.IV.22) , aussi que pour les spectre d'amplitude de deux modèles (Fig.IV.20), l'erreur d'identification de chaque paramètre (Fig.IV.23),en fin lieu des pôles et zéros (Fig.IV.24).

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

Nombre d'échantillons	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.8733	0.4297	0.9764	1.1785	0.0883	0.0457
512	0.8787	0.4355	0.9853	1.2025	0.0509	0.0230
1024	0.8879	0.4407	1.0022	1.1980	0.0289	0.0115
2048	0.8929	0.4473	1.0006	1.2050	0.0158	0.0058
4096	0.8870	0.4488	1.0007	1.1952	0.0093	0.0029

Tab. IV.8 Influence de la variance σ^2 , ($N = 1024$).

Variance σ^2	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.8657	0.4163	1.0114	1.1796	0.0638	0.0170
0.25	0.8879	0.4407	1.0022	1.1980	0.0289	0.0115
0.04	0.8881	0.4481	0.9995	1.1972	0.0214	0.0110
0.01	0.8877	0.4490	0.9992	1.1966	0.0208	0.0109

Les résultats obtenus montrent que lorsque le nombre d'échantillons N est augmenté ; alors les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est asymptotiquement tend vers un bruit blanc.

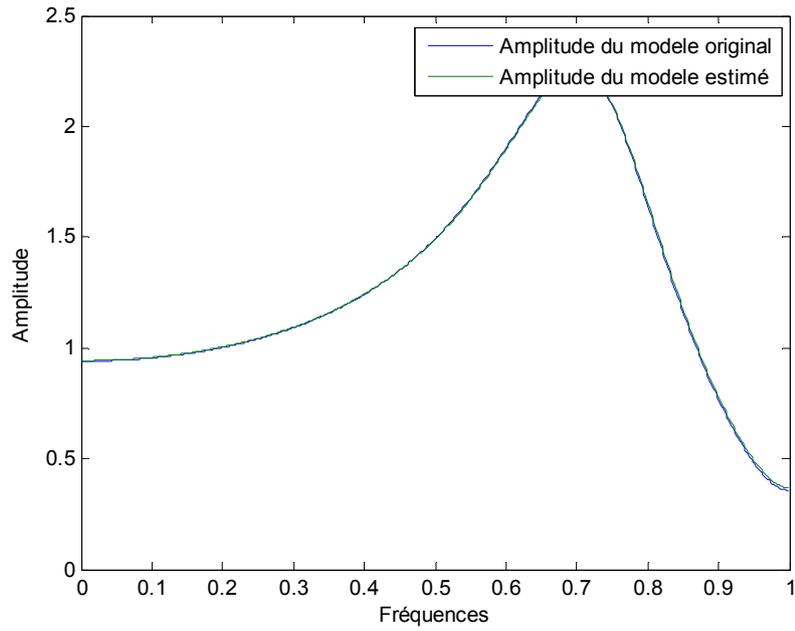


Fig.IV.20 Spectres d'amplitude du modèle original et du modèle estimé.

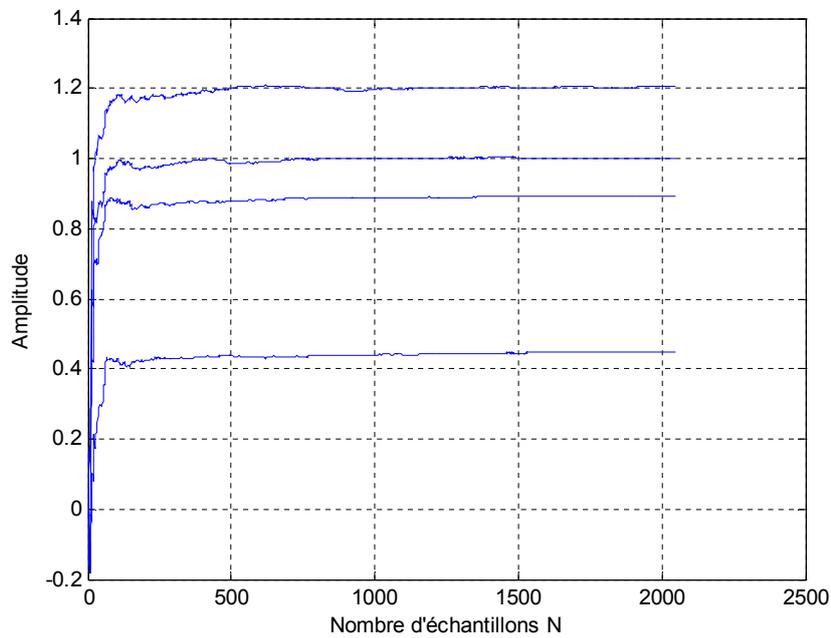


Fig.IV.21 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

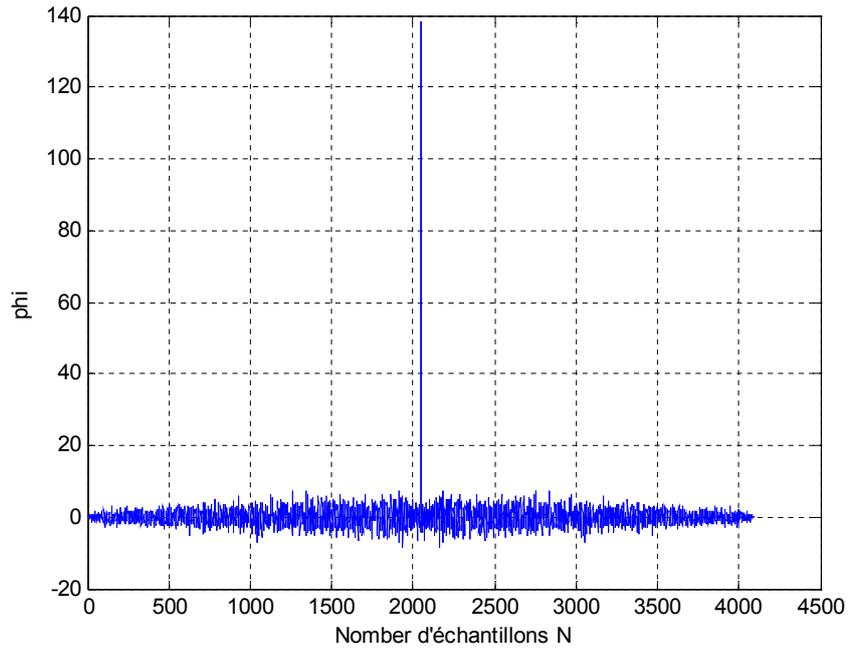


Fig.IV.22 – Fonction d'autocorrélation d'erreur de prédiction.

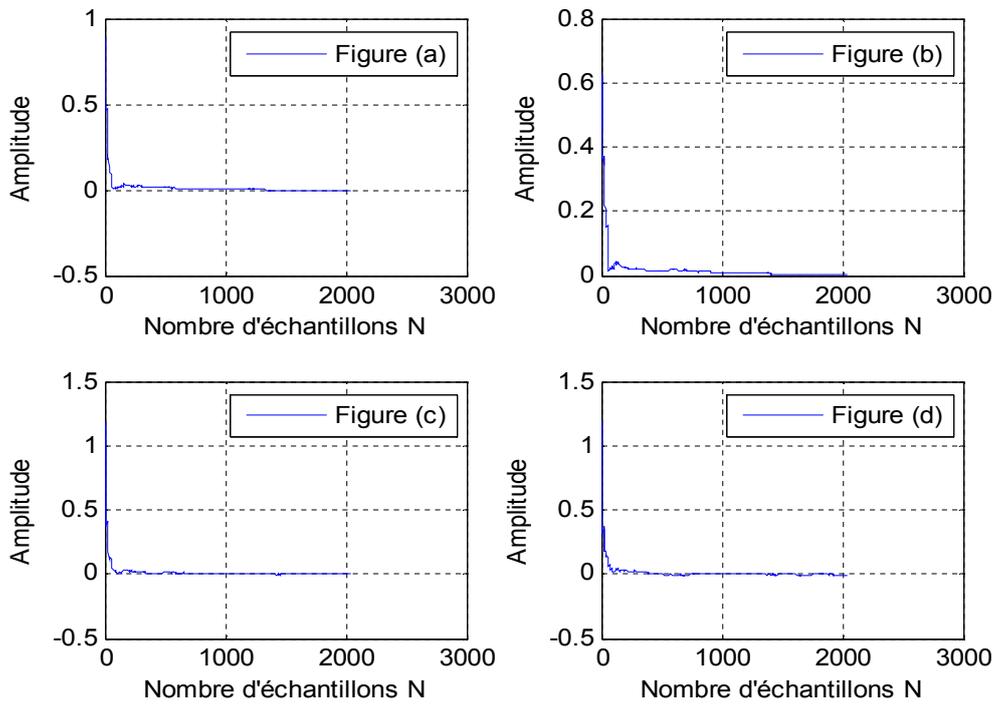
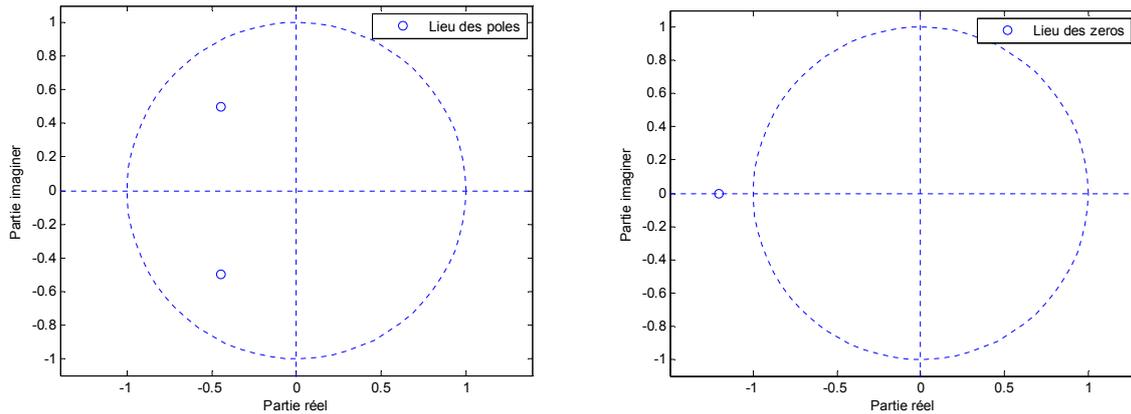


Fig.IV.23 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).



**Fig.IV.24 - Lieu des pôles du modèle estimé (à gauche),
Lieu des zéros du modèle estimé (à droite).**

IV.6 Etude comparative des algorithmes d'identification MCR et MVR

A partir des résultats obtenus dans les exemples de simulation précédents, nous remarquons un bon résultat d'estimation soit le système à phase minimale ou à phase non minimale, le système initial et le système estimé sont de même spectre d'amplitude ayant aussi la même stabilité, l'erreur d'identification des paramètres estimés est presque nulle et l'erreur de prédiction asymptotiquement blanche.

Nous remarquons que l'algorithme MVR est meilleur que MCR, car le biais et l'écart quadratique moyen de l'estimateur MVR sont inférieurs à ceux de l'estimateur MCR (Tab. IV.9 et IV.10).

Tab. IV.9 Biais et EQM (MCR et MVR) $\sigma^2 = 0.25$, exemple à phase minimale.

Nombre d'échantillons	Moyenne du Biais		Moyenne de l'EQM	
	MCR	MVR	MCR	MVR
256	0.1069	0.0849	0.0342	0.0278
512	0.0591	0.0476	0.0172	0.0140
1024	0.0354	0.0276	0.0087	0.0070
2048	0.0195	0.0149	0.0044	0.0035
4096	0.0121	0.0089	0.0022	0.0018

Tab. IV.10 Biais et EQM (MCR et MVR) $\sigma^2 = 0.25$, exemple à phase non minimale.

Nombre d'échantillons	Moyenne du Biais		Moyenne de l'EQM	
	MCR	MVR	MCR	MVR
256	0.1121	0.0883	0.0508	0.0457
512	0.0688	0.0509	0.0258	0.0230
1024	0.0415	0.0289	0.0130	0.0115
2048	0.0233	0.0158	0.0065	0.0058
4096	0.0143	0.0093	0.0033	0.0029

IV.7 Identification par Maximum de Vraisemblance à sortie quantifiée

Dans cette partie, on utilise les mêmes exemples précédents mais cette fois avec des sorties quantifiées avec un quantificateur scalaire (III.5.1) de Lloyd-Max. Les figures ci-dessous représentent le spectre d'amplitude du signal d'entrée au quantificateur et celui du signal quantifié pour un nombre de bits différents.

Soit le système de quantification suivant :

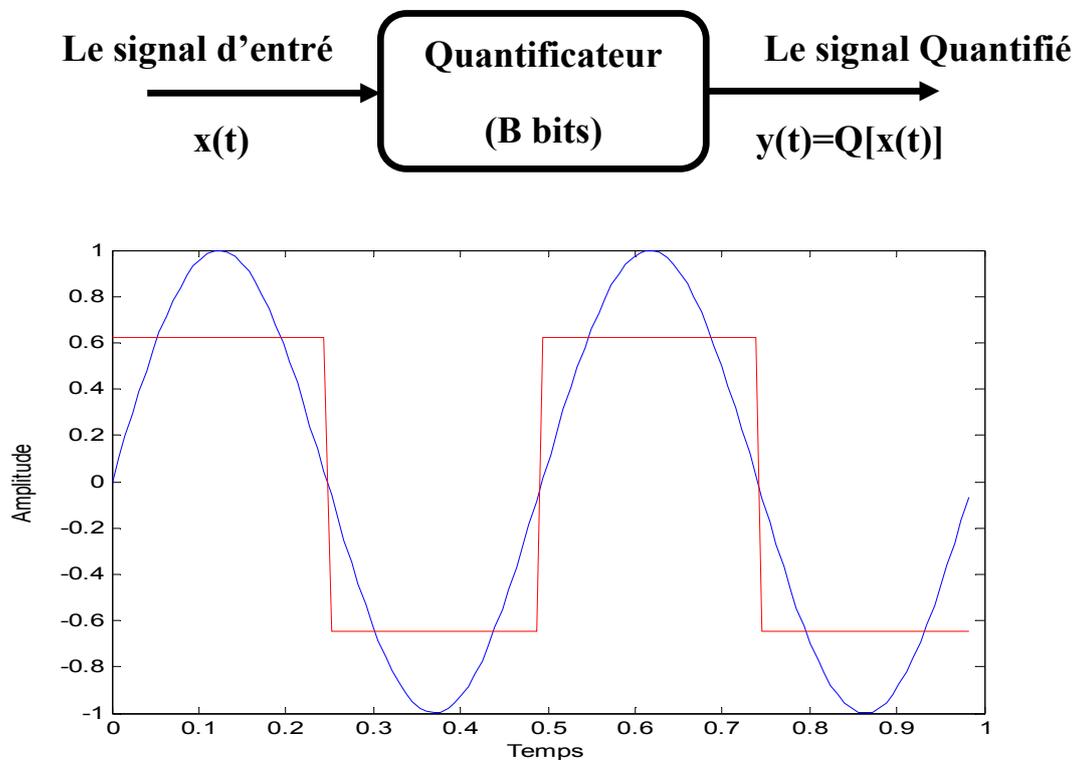


Fig.IV.25 Spectres d'amplitude du signal d'entrée et celui du signal quantifié (B=1 bit).

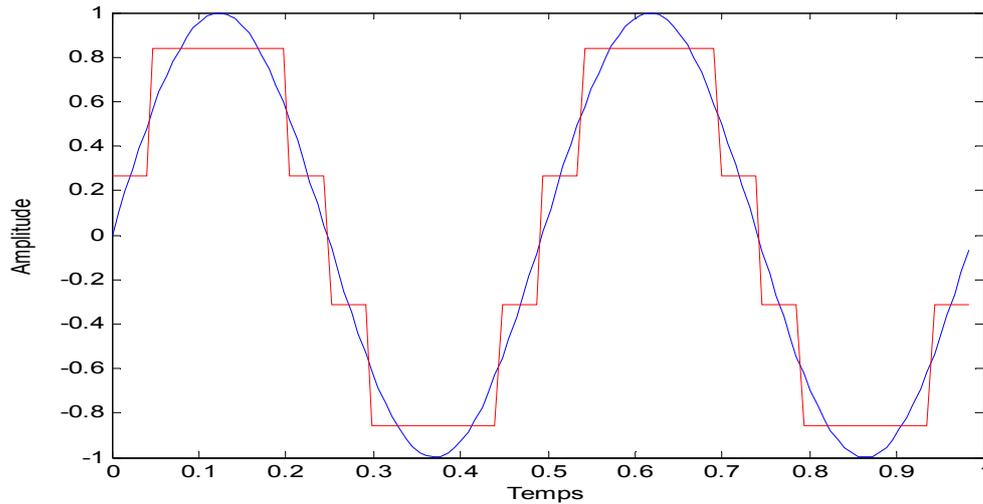


Fig.IV.26 Spectres d'amplitude du signal d'entré et celui du signal quantifié (B=2 bits).

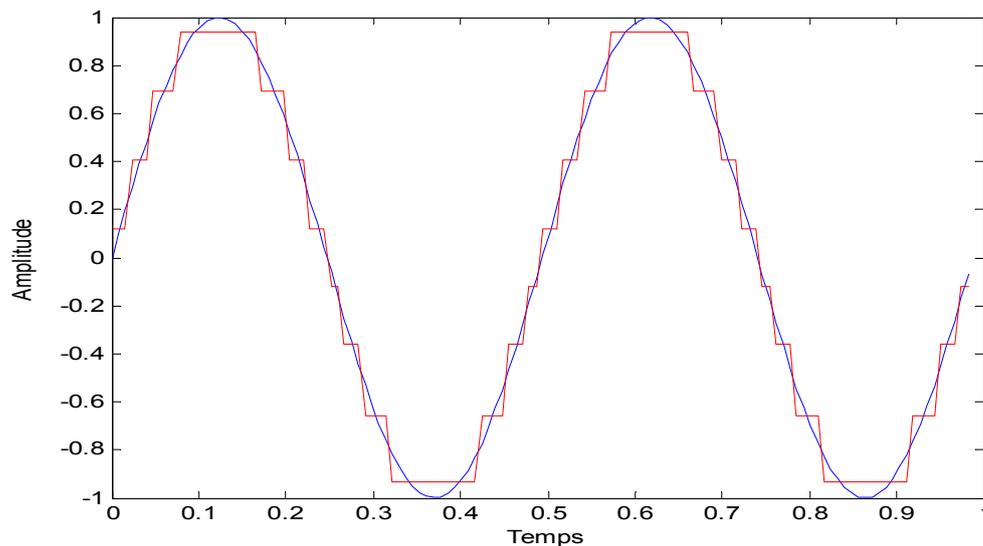


Fig.IV.27 Spectres d'amplitude du signal d'entré et celui du signal quantifié (B=3 bits).

IV.7.1 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), la sortie de ce système est quantifiée a travers un quantificateur (III.5.1) de Lloyd-max (2 bits).

IV.7.1.1 Influence du nombre d'échantillons et de la variance

L'influence du nombre d'échantillons N et la variance σ^2 du bruit additif sur le biais et l'Ecart Quadratique Moyenne (EQM) d'estimateur (MVR), sont données dans les tableaux

IV.11 et IV.12, suivi par les spectres d'amplitude de système initial et le système estimé presque identiques (Fig.IV.28), puis l'évolution des paramètres du système en fonction du temps (Fig.IV.89), ainsi que l'autocorrélation de l'erreur de prédiction est considérée asymptotiquement blanche (Fig.IV.30). On note que l'erreur d'identification converge vers zéro telle qu'elle est montrée sur la figure (Fig.IV.31). En fin un test de stabilité et de phase (Fig.IV.32).

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

Nombre d'échantillons	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.6962	-0.2552	0.9741	0.5859	0.0851	0.0279
512	0.7095	-0.2448	0.9853	0.6149	0.0485	0.0143
1024	0.7236	-0.2449	1.0023	0.6142	0.0282	0.0072
2048	0.7271	-0.2393	1.0007	0.6194	0.0151	0.0036
4096	0.7283	-0.2410	1.0007	0.6167	0.0090	0.0018

Tab. IV.12 Influence de la variance σ^2 , ($N = 1024, b = 2$).

Variance σ^2	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.7094	-0.2666	1.0117	0.6039	0.0500	0.0123
0.25	0.7236	-0.2449	1.0023	0.6142	0.0282	0.0072
0.04	0.7264	-0.2416	0.9995	0.6160	0.0196	0.0060
0.01	0.7274	-0.2414	0.9991	0.6169	0.0180	0.0059

Nous remarquons que : si le nombre d'échantillons N est augmenté, alors le biais de l'estimateur (MVR) toujours diminue, les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est tend vers un bruit blanc.

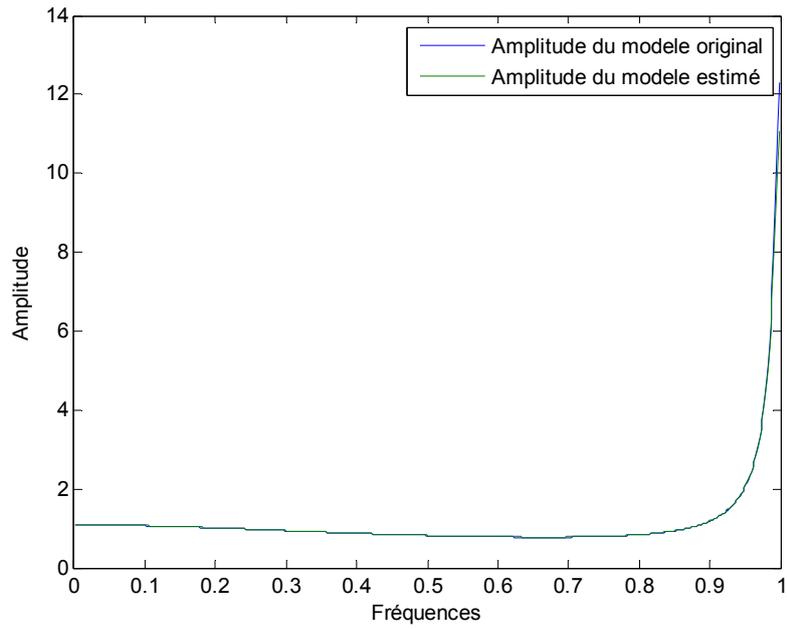


Fig.IV.28 Spectres d'amplitude du modèle original et du modèle estimé.

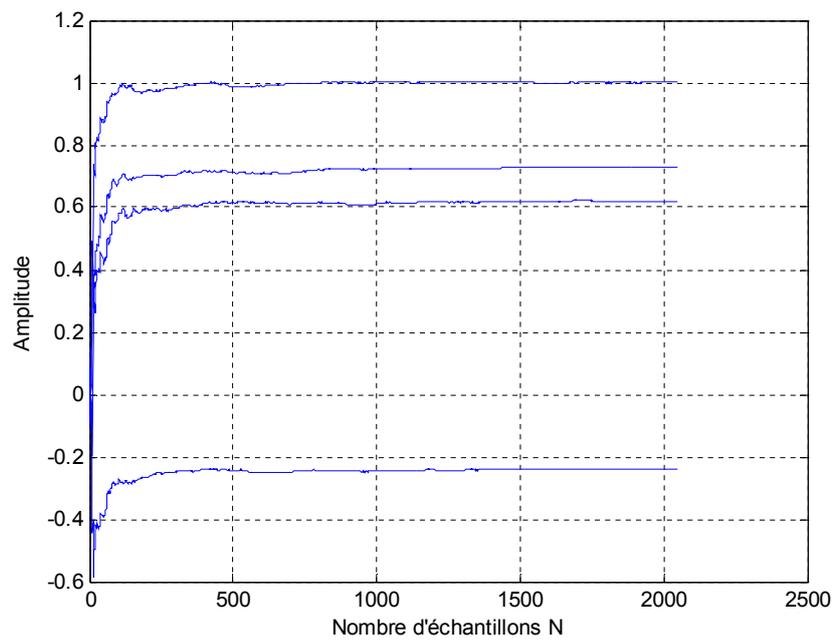


Fig.IV.29 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

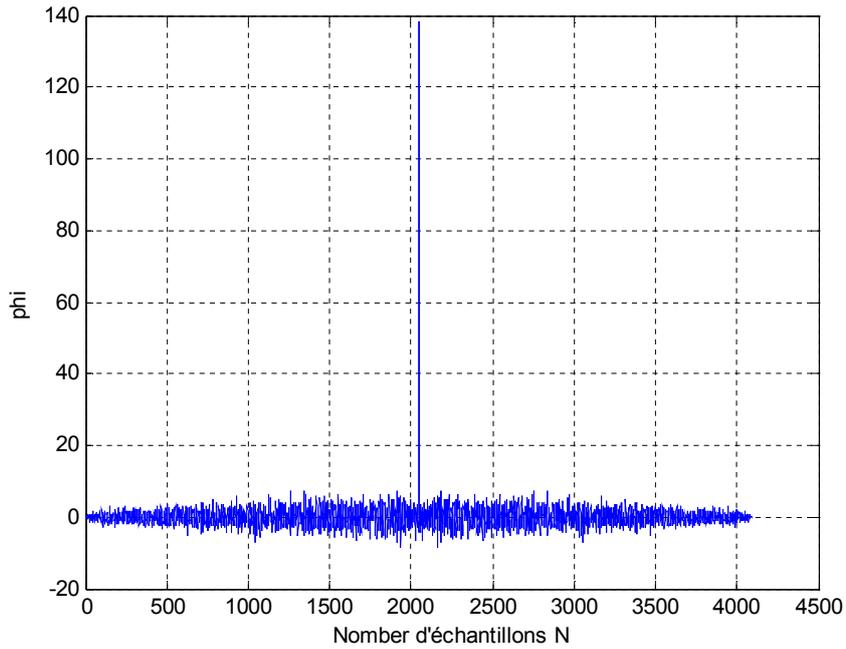


Fig.IV.30 – Fonction d'autocorrélation d'erreur de prédiction.

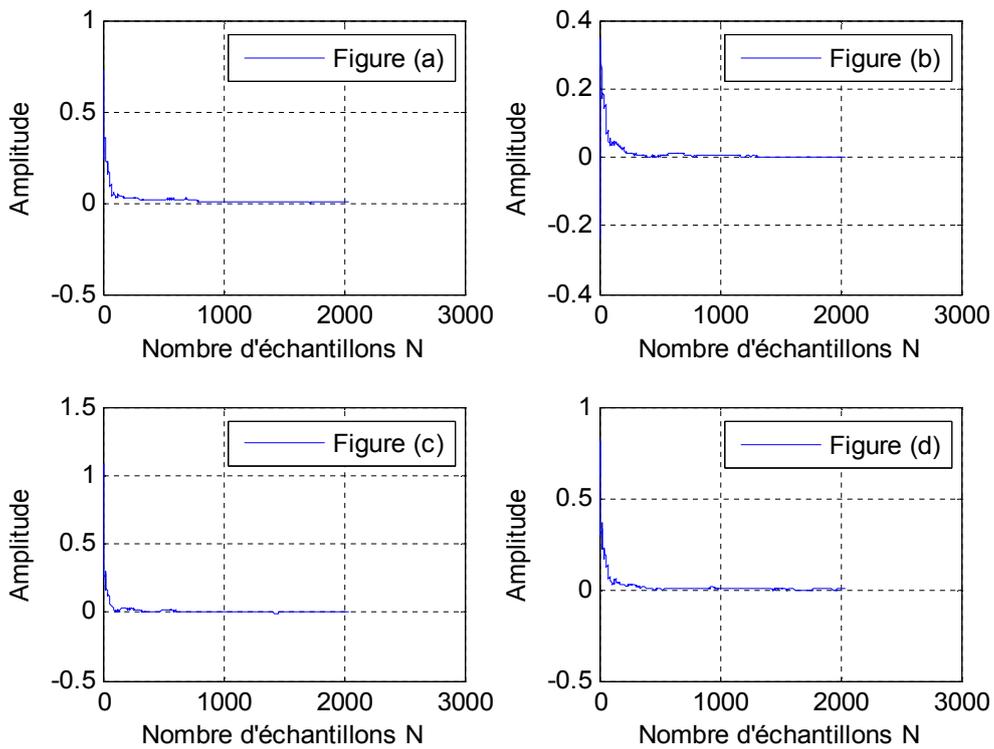
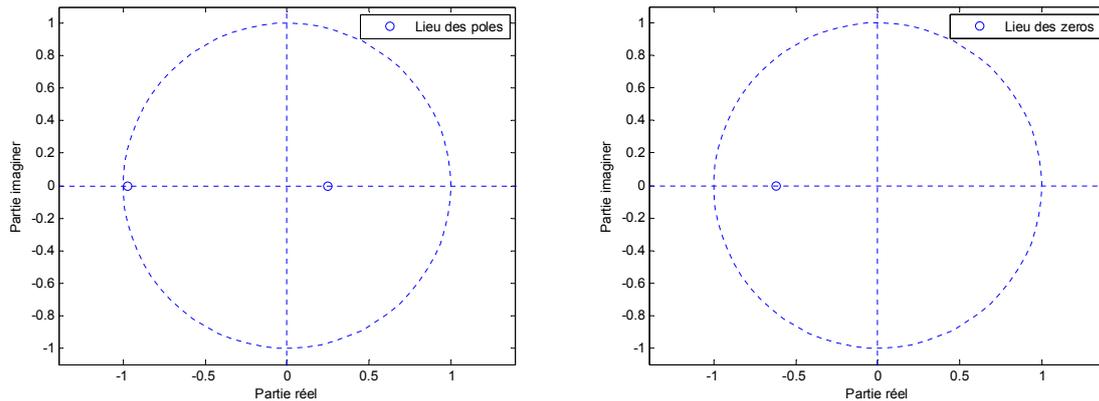


Fig.IV.31 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).



**Fig.IV.32 - Lieu des pôles du modèle estimé (à gauche),
Lieu des zéros du modèle estimé (à droite).**

IV.7.2 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), la sortie de ce système est quantifiée à travers un quantificateur (III.5.1) de Lloyd-max (2 bits).

IV.7.2.1 Influence du nombre d'échantillons et de la variance

Pour tester l'influence du nombre d'échantillons N sur le biais et l'Ecart Quadratique Moyenne (EQM) d'estimation des paramètres du système, on dresse le tableau ci-dessous suivi par l'évolution des paramètres du système ainsi que l'autocorrélation de l'erreur de prédiction sont illustré sur les figure (Fig.IV.33) et (Fig.IV.34), l'erreur d'identification de chaque paramètre (Fig.IV.35).

Tab. IV.13 Influence du nombre d'échantillons $N, (\sigma^2 = 0.25, b = 2)$.

Nombre d'échantillons	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
256	0.8729	0.4297	0.9765	1.1780	0.0898	0.0467
512	0.8785	0.4354	0.9854	1.2023	0.0519	0.0235
1024	0.8878	0.4407	1.0023	1.1979	0.0293	0.0118
2048	0.8928	0.4473	1.0006	1.2050	0.0160	0.0059
4096	0.8870	0.4488	1.0007	1.1952	0.0094	0.0030

Tab. IV.14 Influence de la variance σ^2 , ($N = 1024, b = 2$).

Variance σ^2	Paramètres estimés (MVR)				Moyenne du Biais	Moyenne de l'EQM
	\hat{a}_1	\hat{a}_2	\hat{b}_1	\hat{b}_2		
1	0.8658	0.4167	1.0114	1.1798	0.0582	0.0138
0.25	0.8878	0.4407	1.0023	1.1979	0.0293	0.0118
0.04	0.8881	0.4481	0.9995	1.1971	0.0217	0.0111
0.01	0.8874	0.4490	0.9992	1.1966	0.0210	0.0110

Les résultats obtenus montrent que lorsque le nombre d'échantillons N est augmenté alors les paramètres estimés sont plus proches aux valeurs réelles et l'erreur de prédiction est et tend vers un bruit blanc blanche.

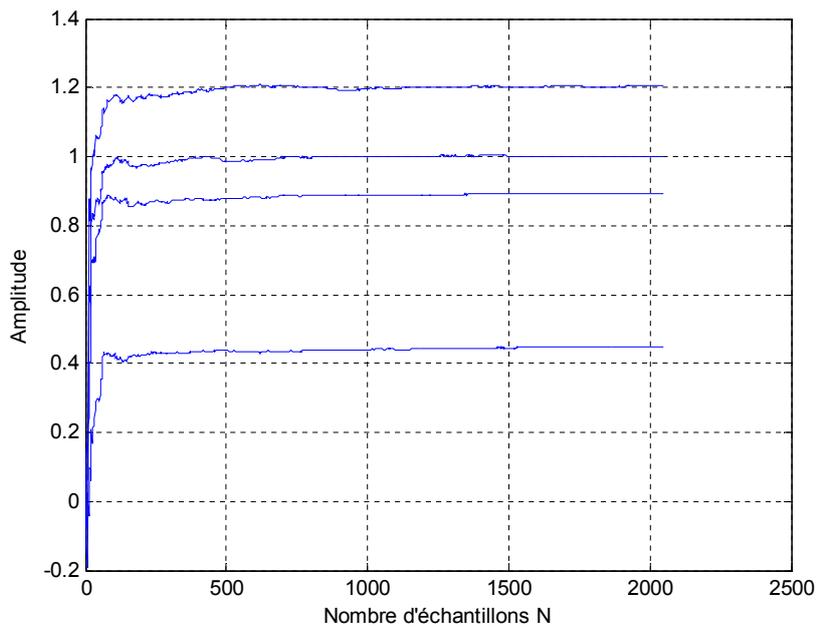


Fig.IV.33 Evolution des paramètres ($N = 2048, \sigma^2 = 0.25$).

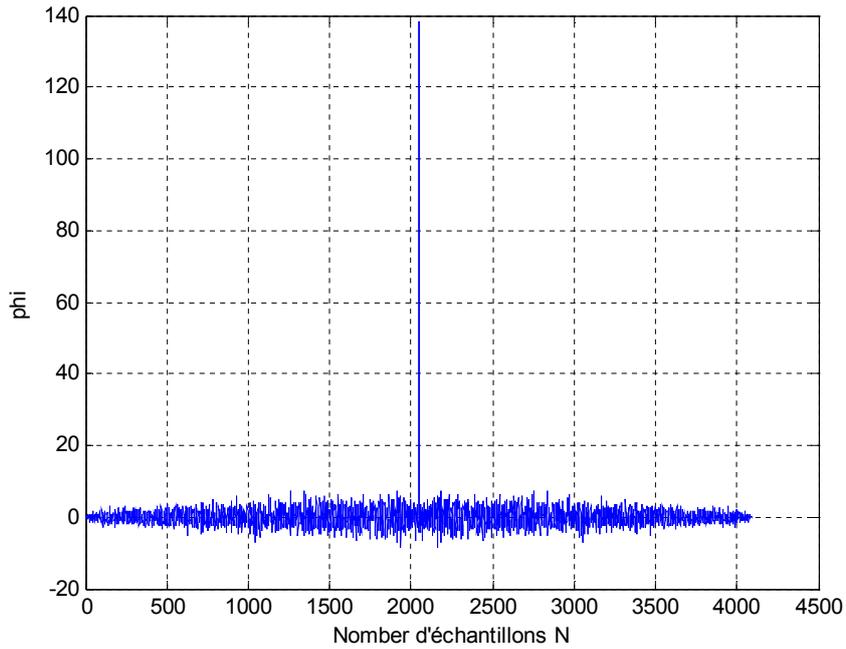


Fig.IV.34 – Fonction d'autocorrélation d'erreur de prédiction.

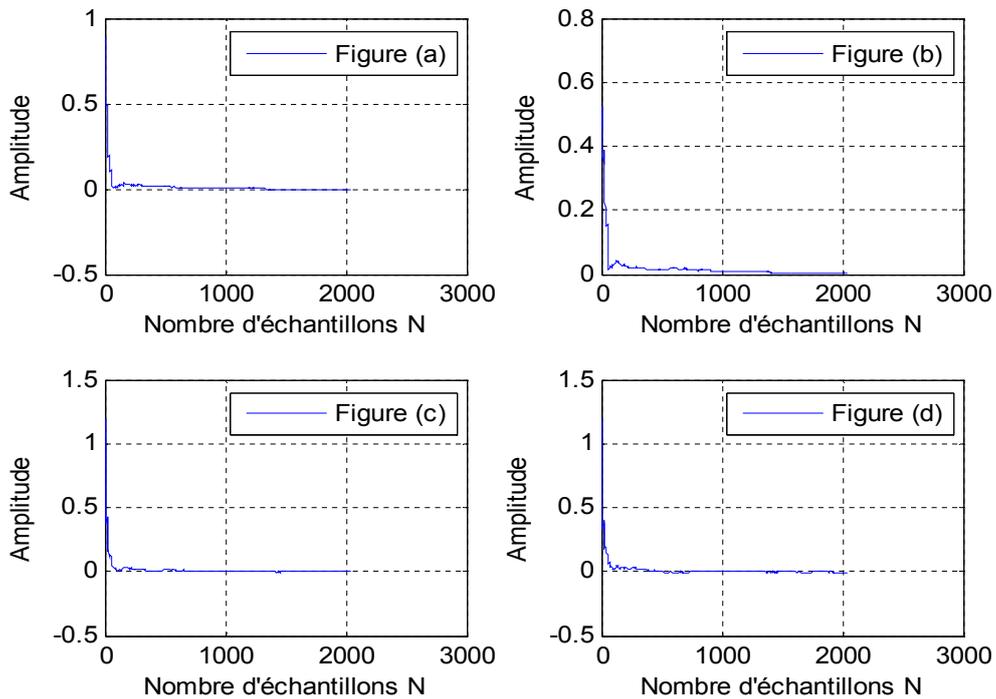


Fig.IV.35 - Erreur d'identification du paramètre a_1 (figure a),
 Erreur d'identification du paramètre a_2 (figure b),
 Erreur d'identification du paramètre b_1 (figure c),
 Erreur d'identification du paramètre b_2 (figure d).

IV.8 Conclusion

Dans ce chapitre nous avons présenté les résultats de l'implémentation sous MATLAB les algorithmes d'identification paramétrique basés sur le blanchissement de l'erreur de prédiction, qui sont : l'algorithme du moindre carré récursif et l'algorithme de maximum de vraisemblance (MVR). Ce sont des algorithmes récursifs qui sont connus comme ayant une vitesse de convergence supérieure.

D'après les résultats de simulation, on tire les conclusions suivantes :

- La méthode MVR est plus performante que la méthode MCR, car son biais (même que l'EQM) est toujours inférieur que le biais du MCR (voir Tab. IV.10).
- En cas d'estimation un modèle à sortie quantifiée, les paramètres estimés par MVR sont plus proches aux valeurs réelles. Avec une erreur de prédiction tend vers un bruit blanc.
- Les modèles estimés sont toujours stables (voir les figures : IV.19, IV.24, IV.32).

CONCLUSION GENERALE

Conclusion générale

Dans le présent travail, nous nous sommes fixés comme objectif l'identification de modèles ARMA à sortie quantifiée ; pour cela, il y a divers algorithmes d'identification paramétrique a été étudié et expliqué, Comme ;

- Gradient stochastique (GS)
- Newton stochastique (NS)
- Moindres carrés récurrents (MCR)
- Maximum de vraisemblance (MVR)

La structure des algorithmes d'identification 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

Nous implémentons deux algorithmes 'MCR' et 'MVR' sur PC (pour éviter une mémoire volumineux), nous choisissons l'algorithme MCR car son aptitude d'estime les paramètres et son temps de convergence rapide à titre de comparaison avec MVR, en utilisant le langage de programmation MATLAB (ver 7.10.0).

Les résultats obtenus nous ont montré que lorsque le nombre d'échantillons N augmente ou la variance σ^2 du bruit additif diminue alors le biais de l'estimateur diminue même que l'écart quadratique moyenne, 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 de Maximum de vraisemblance nous a permis d'identifier des modèles de sortie quantifiée (observations quantifiées).

La solution directe de l'algorithme MVR est connue par sa difficulté, l'idée est de décomposition sur deux étapes :

- Etape 1 : La méthode EM (Expectation Maximisation).
- Etape 2 : La méthode de Quasi-Newton récurrente.

Conclusion générale

La première étape est l'algorithme EM ; son avantage est qu'elle ne demande pas d'initialisation, mais est un algorithme avec convergence lente, le deuxième algorithme résoudre ce problème, parce que sa convergence est rapide.

Enfin, et d'après les résultats de simulations, on dira que l'algorithme MVR est performant que les autres algorithmes étudiés dans ce mémoire, il a l'aptitude d'identifier différents systèmes, à phase minimale, à phase non minimale, et les systèmes à sortie quantifiées .On note que sa convergence est rapide, et l'erreur de prédiction est asymptotiquement blanche.

Bibliographie

Bibliographie

- [1] F. Michaut, M. Bellanger, " Filtrage adaptatif ", théorie et algorithmes, Paris : Lavoisier, 2005.
- [2] L. Ljung, « System Identification: Theory for the User », Second edition, Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [3] L. Y. Wang, G. G. Yin, J. F. Zhang, Y. Zhao, " System Identification with Quantized Observations " Birkhauser, Boston, 2010.
- [4] F. van der Heijden, R.P.W. Duin, D. de Ridder, " Parameter Estimation and State Estimation , An Engineering Approach using MATLAB " Wiley , 2004.
- [5] Y-P .Sheng, L-S. Wang, "The Least Squares Solutions of Inverse Problem for Symmetric Orthogonal Anti-symmetric Matrices", IEEE Conference on Machine learning and Cybernetics, 2006, pp. 631-634.
- [6] T. Soderstrom and P. Stoica, "On some system identification techniques for adaptive filtering," IEEE Trans. Circuits Syst, vol. 35, pp. 457-461,1988.
- [7] D. R. Coughanowr, " Process systems analysis and control", McGraw-Hill, Inc. New York, 1991.
- [8] J. Frang, H. Li, " Distributed adaptive quantization for wireless sensor networks : From delta modulation to maximum likelihood", IEEE Transaction on Signal Processing, 56(10) : 5246-5257, 2008.
- [9] H. C. Papadopoulos, G .W. Wornell, A. V. Oppenheim,"Sequential signal encoding from noisy measurements using quantizers with dynamic bias control", IEEE Transactions on Information Theory, 47(3) : 978-1002 , 2002.
- [10] J. Xiao, A. Ribeiro, Z. Luo, G. Giannakis, " Distributed compression-estimation using wireless sensor networks ", IEEE Signal Processing Magazine, 23(4) : 27-41, 2006.
- [11] B. Choi, " ARMA Model Identification ", Springer, New York,1992.
- [12] B. Picinbono, " Signaux aléatoires et bases du traitement statistique du signal ", Tome3, Dunod, Pris 1993.

- [13] I. G. Ivanov, " Stochastic Modeling and Control ", Croatia, 2012.
- [14] T. Chonavel, " Statistical signal processing, Modelling and estimation ", Springer, January 2002.
- [15] F. Mix, " Random signal processing ", prentice hall, 1995.
- [16] D. Arzelier, " Introduction à l'étude des processus stochastiques et au filtrage optimal" , version 2, INSA, 1998.
- [17] B. Frank, " Item response theory: parameter estimation techniques ", Assessment Systems Corporation, 1992.
- [18] D. Declercq, A. Quinqu, " Signaux et systemes en questions ", Hermes, Paris 1996.
- [19] J. Bernussou, " Processus stochastique– filtrage ", Cours ENSICA, Toulouse, 1994.
- [20] R. Ben Abdennour, P .Borne, " Identification et commande numérique des procédés industriels ", April 2001.
- [21] D. Landau, " Identification et commande des systèmes" , Hermes, Paris, 1998.
- [22] Jacques Richalet, " Pratique de l'identification ", Hermes, Paris 1998.
- [23] M. Kunt, M. Bellanger, "Techniques modernes de traitement numérique des signaux", collection électricité, traitement de l'information, Volume1, Année 1991.
- [24] A. Borrie, " Stochastic System for Engineering, Modelling, Estimation and Control ", Prentice hall, 1992.
- [25] J. J. Distefano, A. R. Stubberud, I. J. Williams, " Feedback and Control Systems " , second edition, McGraw-Hill, Inc. New York, 1990.
- [26] K. Kitz, " Evolutionary Algorithms ", Rijeka, Croitia, Mars 2011.
- [27] R. Fletcher, " Partial Methods of optimization ", Jhon Wiley et Sons, Inc, 1987.
- [28] B. Widrow et al. " Adaptive Noise canceling: Principles and Applications ", Proc. IEEE, 63, N° 12, December 1975.
- [29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] A. Maddi, A. Guessoum, D. Berkani, O. Belkina, " Etude de la méthode des moindres carrée récursive et application au signal de parole ", Sousse, Tunisie, SETIT, 17-21 Mars 2005.

[35] S.C. Douglas, " Numerically-robust RLS algorithms using Least Squares pre-whitening", in Proceedings of ICASSP, Istanbul, Turkey, pp. 412-415, 2000.

[36] 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.

[37] A. Nongnuch, 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.

[38] M. Djendi, " Amélioration du comportement des algorithmes des moindres carrés transversaux rapides. Application à l'identification de réponse impulsionnelle acoustique longues ", thèse de doctorat, école polytechnique d'El Harrach, Alger, Septembre 2006.

[39] P.P.B. Eggermont, V. N. L a Riccia, " Maximum Penalized Likelihood Estimation ",

Springer, New York, 2009.

[40] J. Schoukens, R. Pintelon, Y. Rolain, "Mastering System Identification in 100 Exercises", IEEE Press et Wiley, Canada, 2012.

[41] G. J. Mclachlan, T. Krishnan, "The EM Algorithm and Extensions", Wiley-Interscience, 2^e édition, 2008.

[42] D. Cappé, E. Moulines, T. Rydén, "Inference in Hidden Markov Models", Springer, 2005.

[43] S. Boyd, L. Vandenberghe, "convex optimisation", Cambridge Univ, 2004.

[44] S. T. Karris, "Signals and Systems with MATLAB Computing and Simulink Modeling", Orchard, USA, 3^e édition, 2007.

[45] S. T. Karris, "Introduction to Simulink with Engineering Applications", Orchard, USA, 2006.

