



**HAL**  
open science

# Statistique asymptotique de certaines séries chronologiques à mémoire

Marius Soltane

► **To cite this version:**

Marius Soltane. Statistique asymptotique de certaines séries chronologiques à mémoire. Statistiques [math.ST]. Le Mans Université, 2020. Français. NNT : 2020LEMA1027 . tel-03092320

**HAL Id: tel-03092320**

**<https://theses.hal.science/tel-03092320v1>**

Submitted on 2 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Thèse de doctorat de

LE MANS UNIVERSITÉ  
Comue Université Bretagne Loire

Ecole Doctorale n° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : Mathématiques

Par  
**Marius SOLTANE**

« **Statistique asymptotique de certaines séries chronologiques  
à mémoire** »

Thèse présentée et soutenue à Le Mans, le 10 Novembre 2020  
Unité de recherche : Laboratoire Manceau de Mathématiques  
Thèse N° : 2020LEMA2017

## Rapporteurs avant soutenance :

Alexander AUE, Professeur, Université de Davis, États-Unis d'Amérique  
Mathieu ROSENBAUM, Professeur, École Polytechnique, France

## Composition du Jury :

Président : Anne PHILIPPE, Professeur, Université de Nantes, France

Examineurs : Marina KLEPTSZYNA, Professeur, Le Mans Université, France  
Céline LACAUX, Professeur, Université d'Avignon, France  
Agnès LAGNOUX, Maître de conférences, Université Toulouse Jean Jaurès, France

Dir. de thèse : Alexandre BROUSTE, Professeur, Le Mans Université, France

Co-dir. de thèse : Frédéric PROÏA, Maître de conférences, Université d'Angers, France

# Remerciements

Je tiens en premier lieu à remercier mes directeurs de thèse Alexandre Brouste et Frédéric Proïa pour leurs conseils avisés tout au long de thèse me permettant d'aboutir aux travaux présentés dans ce manuscrit.

Je souhaiterais également remercier Alexander Aue et Mathieu Rosenbaum d'avoir accepté de rapporter ce manuscrit. Je tiens également à remercier Céline Lacaux, Agnès Lagnoux, Marina Kleptsyna et Anne Philippe pour avoir accepté d'examiner cette thèse.

J'adresse également mes remerciements aux membres du département de mathématiques et du laboratoire de mathématiques. L'ambiance chaleureuse qui règne au sein des locaux est très agréable et propice au travail. Je remercie Chunhao Cai, Eirini Votsi et Longmin Wang pour quelques collaborations fructueuses.

Je tiens pour finir à remercier mes proches ainsi que mes amis pour le soutien apporté durant cette thèse.

# Table des matières

<b>Introduction</b>	<b>6</b>
<b>Introduction</b>	<b>8</b>
<b>1 Sur les séries chronologiques</b>	<b>10</b>
1.1 Outils fondamentaux	10
1.1.1 Le processus stochastique	10
1.1.2 La stationnarité	11
1.1.3 L'autocorrélation	12
1.1.4 La densité spectrale	13
1.1.5 Ratio de vraisemblance localement asymptotiquement normal (LAN)	13
1.2 Le processus ARMA stationnaire	14
1.2.1 Présentation du modèle	15
1.2.2 Causalité, inversibilité et stationnarité	16
1.2.3 Autocorrélations	17
1.2.4 Densité spectrale	19
1.3 Inférence pour les processus autorégressifs	19
1.3.1 L'estimateur des moindres carrés	20
1.3.2 Consistance forte	20
1.3.3 Normalité asymptotique	21
1.4 Détection d'une rupture pour les processus autorégressifs	22
1.4.1 Présentation du test	22
1.4.2 Comportement asymptotique du test	23
1.4.3 Illustrations numériques	24
1.5 Processus RCAR(1)	25
1.5.1 Présentation du modèle	25
1.5.2 Consistance forte	26
1.5.3 Normalité asymptotique	27
<b>2 Sur les processus autorégressifs à bruits Gaussiens stationnaires</b>	<b>28</b>
2.1 Historique et motivation	28
2.2 Le modèle et les propriétés connues	29
2.2.1 Présentation du modèle ainsi que de la transformation	29
2.2.2 Construction de l'estimateur du maximum de vraisemblance	31
2.3 Résultats principaux	32

2.3.1	Propriétés asymptotiques presque sûres du MLE . . . . .	32
2.3.2	Propriété LAN et applications . . . . .	33
2.4	Simulations . . . . .	35
2.4.1	Simulation pour un processus AR(1) avec $\theta_0 = 0, 6$ . . . . .	36
2.4.2	Simulation pour un processus AR(2) avec $\theta_0 = (0, 3; 0, 5)^*$ . . . . .	36
2.4.3	Simulation pour un processus AR(3) avec $\theta_0 = (0, 3; 0, 6; -0, 4)^*$ . . . . .	37
2.5	Conclusion et Perspectives . . . . .	37
2.6	Résultats Auxiliaires . . . . .	38
2.6.1	Preuve de la Proposition 2.11 . . . . .	42
2.7	Preuve des Résultats principaux . . . . .	44
2.7.1	Preuve du Théorème 2.1 . . . . .	44
2.7.2	Preuve du Théorème 2.2 . . . . .	45
2.7.3	Preuve de la Proposition 2.3 . . . . .	48
2.7.4	Preuve du Théorème 2.5 . . . . .	49
2.7.5	Preuve de la Proposition 2.6 . . . . .	49
2.7.6	Preuve du Théorème 2.10 . . . . .	50
<b>3</b>	<b>Détection d'une rupture au sein d'un processus autorégressifs à bruits gaussiens stationnaires</b> . . . . .	<b>51</b>
3.1	Historique et motivation . . . . .	51
3.2	Notations et rappels . . . . .	52
3.3	Résultat principal . . . . .	54
3.4	Simulations . . . . .	55
3.5	Conclusion et Perspectives . . . . .	58
3.6	Résultats Auxiliaires . . . . .	58
3.6.1	Lemmes techniques pour établir la Proposition 3.2 . . . . .	60
3.6.2	Lemmes techniques pour établir la Proposition 3.4 . . . . .	63
3.6.3	Preuve de la Proposition 3.3 . . . . .	67
3.6.4	Preuve de la Proposition 3.4 . . . . .	68
3.7	Preuve du résultat principal . . . . .	68
<b>4</b>	<b>Un modèle chronologique à coefficients aléatoires corrélés</b> . . . . .	<b>70</b>
4.1	Historique et motivation . . . . .	70
4.2	Stationnarité et fonction d'autocorrélation . . . . .	73
4.3	Comportement asymptotique de la moyenne empirique et de l'estimateur usuel dans ce modèle . . . . .	74
4.4	Un test de détection de mémoire dans les coefficients . . . . .	78
4.5	Une généralisation vectorielle . . . . .	79
4.6	Conclusion et Perspectives . . . . .	82
4.7	Simulations . . . . .	82
4.8	Preuve des résultats principaux . . . . .	84
4.8.1	Au sujet des Remarques des Sections 4.2 et 4.3 . . . . .	85
4.8.2	Preuve du Théorème 4.1 . . . . .	86
4.8.3	Preuve du Théorème 4.2 . . . . .	86
4.8.4	Preuve du Théorème 4.3 . . . . .	89

4.8.5	Preuve du Théorème 4.4	91
4.8.6	Preuve du Théorème 4.5	98
4.8.7	Preuve du Théorème 4.7	99
4.8.8	Preuve de la Proposition 4.8	102
4.8.9	Preuve du Théorème 4.10	103
4.8.10	Preuve du Théorème 4.11	105
4.9	Décompositions Martingales intervenant à la Section 4.5	105
4.10	Intégrabilité du modèle introduit à la Section 4.5	107
4.10.1	Moments d'ordre deux	107
4.10.2	Moments d'ordre quatre	108
4.11	Notations additionnelles	109
<b>5</b>	<b>Procédure One-Step pour un bruit Gaussien fractionnaire</b>	<b>111</b>
5.1	Historique et motivation	111
5.2	Bruit gaussien fractionnaire dans le schéma haute-fréquence	113
5.2.1	Rappels sur le fGn	113
5.2.2	Propriétés asymptotiques du MLE	114
5.3	Efficacité asymptotique de l'estimateur one-step MLE	115
5.3.1	Estimateur initial via la méthode des variations quadratiques généralisées	116
5.3.2	Propriétés asymptotiques de l'estimateur one-step MLE	117
5.3.3	Simulations	117
5.4	Conclusion et perspectives	121
5.5	Preuve des résultats principaux	121
5.5.1	Preuve du Théorème 5.1	121
5.5.2	Preuve du Théorème 5.2	123

# Introduction

We generally designate by time series, the modeling of a series of random events, generally on a time scale. Many real data flows admit such a representation, we can for example think of economic and industrial indicators, the evolution of a climatic phenomenon . . .

The main results that contributed to the formalization of the time series come from innovative work of Box and Jenkins [17, 16]. The ARMA process (Autoregressive Moving Average), on which we will have the opportunity to return in detail in the first part of this thesis, relies exclusively on past observations from a time series. The main feature of these models is the strong dependence between two consecutive observations. Thus, the analysis of linear time series experienced a great development at the end of the last century. More recently, Nicholls and Quinn [78, 79, 80] have introduced and then formulated the first rigorous results on the RCAR model (Random Coefficient Autoregressive) which is an extension of the autoregressive model. The flexibility of this model is based on the fact that the coefficients of the drift are random.

Random shocks in the autoregressive process are generally assimilated to a strong white noise in the most common models or to a weak white noise in the more general models. These usual working hypotheses guarantee both the consistency of the least squares estimator of the parameters and its asymptotic normality. Likewise, the random coefficients in the RCAR model are assumed to be independent and identically distributed to guarantee the consistency and asymptotic normality of the least squares estimator.

When these hypotheses are called into question, it is shown in [46, 74] that the least squares estimator is no longer consistent in the ARMA model and therefore no longer estimates the desired parameter. For RCAR, again, when the random coefficients have memory, the least square estimator leads to erroneous parameter estimates. We propose in the following consistent and asymptotically normal estimators of parameters in the case of autoregressive processes with dependence in the coefficients.

The case of the autoregressive model disturbed by a stationary Gaussian process is studied in the second part of the thesis. In line of [21], we study in [91] the almost-sure asymptotic properties of the maximum likelihood estimator in the case where the covariance structure is known. The maximum likelihood estimator has the virtue of being consistent and asymptotically efficient in regular models. With very general conditions [96] the existence, the consistency as well as the existence of an asymptotic distribution for this estimator are gua-

ranteed.

We also propose for this model, in the third part, a procedure to detect a change of regime of the autoregressive parameter. In the classical case a procedure was proposed in [35] which is based on the use of a strong invariance principle combined with the Darling-Erdős theorem. We will see that we cannot apply the exact same method but we will nevertheless obtain a procedure to detect a break as described in [22].

The example of the RCAR model with random coefficients presenting a short memory is studied in the fourth part. We propose in [84] a method to correct the bias problem to estimate the autoregressive parameter in the scalar case. A generalization of this model in the vector case is also studied.

Finally, in a last part, we study the asymptotic properties of a joint variance estimator and of the Hurst exponent of a fractional Gaussian noise observed at high frequency. The efficiency of the maximum likelihood estimator has been proven in this singular statistical experiment [23]. However, the maximum likelihood estimator does not have an explicit form. Numerical optimization is very costly from a computational point of view and can, in certain cases, prove to be unstable. The asymptotic properties of the estimator based on the Le Cam procedure described in [68] and allow, starting from an initial estimator, to construct a new estimator having the same asymptotic variance as the maximum likelihood.

# Introduction

Nous désignons généralement par série chronologique, ou série temporelle, la modélisation d'une suite d'évènements aléatoires, généralement sur une échelle temporelle. De nombreux flux de données réelles admettent une telle représentation, nous pouvons par exemple penser aux indicateurs économiques et industriels, à l'évolution d'un phénomène climatique ...

Les principaux résultats ayant contribué à la formalisation des séries chronologiques reviennent aux travaux fondateurs de Box et Jenkins [17, 16]. Le processus ARMA (Autoregressive Moving Average), sur lequel nous aurons l'occasion de revenir en détail dans la première partie de cette thèse, s'appuie exclusivement sur les observations passées d'une série chronologique. La caractéristique principale de ces modèles est la dépendance forte liant deux observations consécutives. Ainsi, l'analyse des séries chronologiques linéaires a connu un vaste essor à la fin du siècle dernier. Plus récemment, Nicholls et Quinn [78, 79, 80] ont introduit puis formulé les premiers résultats rigoureux sur le modèle RCAR (Random Coefficient Autoregressive) qui est une extension du modèle autorégressif. La souplesse de ce modèle repose sur le fait que les coefficients de la dérive sont aléatoires.

Les chocs aléatoires dans le processus autorégressif sont généralement assimilés à un bruit blanc fort dans les modèles les plus courants ou à un bruit blanc faible dans les modèles plus généraux. Ces hypothèses usuelles de travail garantissent à la fois la consistance de l'estimateur des moindres carrés des paramètres et la normalité asymptotique de celui-ci. De même, les coefficients aléatoires dans le modèle RCAR sont supposés indépendants et identiquement distribués pour garantir la consistance et la normalité asymptotique de l'estimateur des moindres carrés.

Lorsque ces hypothèses sont remises en cause, il est démontré [46, 74] que l'estimateur des moindres carrés n'est plus consistant dans le modèle ARMA et n'estime donc plus le paramètre voulu. Pour le RCAR, là encore, lorsque les coefficients aléatoires présentent de la mémoire, l'estimateur des moindres carrés conduit à des estimations erronées des paramètres. Nous proposons dans la suite des estimateurs consistants et asymptotiquement normaux des paramètres dans le cas de processus autorégressifs présentant de la dépendance dans les aléas.

Le cas du modèle autorégressif perturbé par un processus gaussien stationnaire est étudié dans la deuxième partie de la thèse. Dans la continuité de [21], nous étudions dans [91] les propriétés asymptotiques presque-sûres de l'estimateur du maximum de vraisemblance dans le cas où la structure de covariance est connue. L'estimateur de maximum de vraisemblance

possède la vertu d'être consistant et asymptotiquement efficace dans les modèles réguliers. Des conditions très génériques [96] garantissent l'existence, la consistance ainsi que l'existence d'une distribution asymptotique pour cet estimateur.

Nous proposons également pour ce modèle, dans la troisième partie, une procédure afin de détecter un changement de régime du paramètre autorégressif. Dans le cas classique une procédure fût proposée dans [35] qui repose sur l'utilisation d'un principe d'invariance fort conjuguée au théorème de Darling-Erdős. Nous verrons que nous ne pouvons pas appliquer exactement la même méthode mais nous obtiendrons néanmoins une procédure pour détecter une rupture comme cela est décrit dans [22].

L'exemple du modèle RCAR avec des coefficients aléatoires présentant une mémoire courte est étudié dans la quatrième partie. Nous proposons dans [84] une méthode afin de corriger le problème de biais pour estimer le paramètre autorégressif dans le cas scalaire. Une généralisation de ce modèle dans le cas vectoriel est également étudiée.

Enfin, dans une dernière partie, nous étudions les propriétés asymptotique d'un estimateur conjoint de la variance et de l'exposant de Hurst d'un bruit gaussien fractionnaire observé à haute fréquence. L'efficacité de l'estimateur du maximum de vraisemblance a été prouvée dans cette expérience statistique singulière [23]. L'estimateur du maximum de vraisemblance a cependant le défaut de ne pas posséder de forme explicite. L'optimisation numérique est très couteuse sur le plan computationnel et peut, dans certains cas, s'avérer être instable. Les propriétés asymptotiques de l'estimateur basé sur la procédure de Le Cam décrites dans [68] et permettent, à partir d'un estimateur initial, de construire un nouvel estimateur ayant la même variance asymptotique que le maximum de vraisemblance.

# Chapitre 1

## Sur les séries chronologiques

Ce chapitre est une présentation des principales définitions et propriétés des séries chronologiques linéaires, comme la notion de stationnarité qui est l'un des socles de l'étude. Nous abordons ensuite la modélisation ARMA dont le processus autorégressif est un cas particulier. Nous rappelons ses propriétés usuelles ainsi que celles de ses estimateurs des moindres carrés puis nous rappelons également une procédure statistique permettant de détecter un changement de régime au sein du série chronologique générée par un processus autorégressif. Nous terminerons ce chapitre par un historique sur un modèle chronologique généralisant un processus autorégressif.

### 1.1 Outils fondamentaux

Nous rappelons, en guise de section introductive, certains concepts fondamentaux des séries chronologiques.

#### 1.1.1 Le processus stochastique

Considérons un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , un ensemble d'indices  $T$  et un espace métrique  $S$  muni de la tribu borélienne  $\mathcal{B}(S)$ .

**Définition 1.1.** *Un processus stochastique est une suite de variable aléatoires  $(X_t)$  définies sur  $(\Omega, \mathcal{F}, \mathbb{P})$ , indexées par  $T$  et à valeurs dans  $S$ . Pour toute réalisation  $\omega \in \Omega$ , la famille  $(X_t(\omega))$  est une trajectoire du processus.*

Commençons cette section par restreindre l'espace  $T$  tel que décrit dans la Définition 1.1 sur lequel est indexé le processus stochastique. Nous nous intéressons ici aux processus définis sur  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ , indexés par  $T = \mathbb{Z}$  et à valeurs dans  $S = \mathbb{R}$ .

C'est à la trajectoire d'un processus observé sur un sous-ensemble de  $T$  que l'on associera la notion de série chronologique. Cependant, pour ne pas alourdir les notations, nous ne distinguerons généralement pas le processus  $(X_t)$  d'une de ses réalisations  $(X_t(\omega))$ , lorsqu'il n'y aura aucune ambiguïté vis à vis de la quantité que l'on manipule. Nous travaillerons par la suite sur des processus discrets indexés par  $\mathbb{N}$  ou  $\mathbb{Z}$  et à valeurs réelles. Nous ferons

d'ailleurs référence à cette sous-classe de manière récurrente à travers le terme générique de processus. L'ensemble  $T$  sera en conséquence l'espace des temps,  $\mathbb{N}$  ou bien  $\mathbb{Z}$  et ce tout au long de nos travaux.

### 1.1.2 La stationnarité

Intuitivement, la stationnarité reflète la capacité d'un processus à ne pas dépendre (au sens des lois jointes finies dimensionnelles) de l'indice temporel. Ce dernier est alors entièrement décrit par sa loi stationnaire qui, par construction, ne dépend pas du temps. Une telle propriété présente un intérêt pratique, mais possède également un impact théorique considérable puisqu'elle figure comme une hypothèse formulée au sein de nombreux travaux. Nous distinguons traditionnellement deux notions de stationnarité, la stationnarité au sens stricte et celle au sens faible. Pour les formaliser, considérons un processus  $(X_t)$  défini sur  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ , l'ensemble des variables aléatoires de carré intégrable.

**Définition 1.2.** *Le processus  $(X_t)$  est dit strictement stationnaire si pour tout  $n \in \mathbb{N}^*$  et tout décalage temporel  $k \in \mathbb{Z}$ , la loi du vecteur  $(X_1, \dots, X_{1+n})$  est la même que celle du vecteur  $(X_{1+h}, \dots, X_{1+n+h})$ .*

La stationnarité stricte est une hypothèse assez forte, nécessairement difficile à vérifier en pratique surtout lorsque le processus n'est pas gaussien. Nous introduisons donc une notion de stationnarité moins contraignante.

**Définition 1.3.** *On dit que le processus  $(X_t)$  est faiblement stationnaire si, pour tout décalage temporel  $k \in \mathbb{Z}$ ,  $\mathbb{E}(X_t)$  et  $\mathbb{V}ar(X_t)$  sont constants, et  $\mathbb{C}ov(X_{t+k}, X_t)$  ne dépend que de  $k$ .*

Nous dirons également que le processus est stationnaire au second ordre, en lien avec la stabilisation de sa variance. Nous ferons par ailleurs implicitement référence à cette notion de stationnarité dans la suite de nos rappels. Notons que la stationnarité stricte conjuguée à l'existence du moment d'ordre 2 implique bien entendu la stationnarité faible. L'exemple le plus trivial de processus stationnaire est le bruit blanc.

**Définition 1.4.** *Un processus  $(\varepsilon_t)$  est qualifié de bruit blanc si, pour tous  $t_1, t_2 \in T$ , on a simultanément  $\mathbb{E}(\varepsilon_t) = 0$ ,  $\mathbb{V}ar(\varepsilon_t) = \sigma_2 < +\infty$  et  $\mathbb{C}ov(\varepsilon_{t_1}, \varepsilon_{t_2}) = 0$  dès que  $t_1 \neq t_2$ .*

Si  $(\varepsilon_t)$  est générée par des variables aléatoires indépendantes et identiquement distribuées, on parlera alors de bruit blanc fort. Il est très courant de supposer que le processus des chocs aléatoires perturbant un modèle statistique forme un bruit blanc. Un test de blancheur résiduelle peut ainsi être considéré comme un indicateur de la qualité de la modélisation, dans la mesure où l'absence de mémoire résiduelle implique que toute l'information statistique est contenue dans notre modélisation. La perturbation aléatoire inobservable aura donc un caractère purement erratique. Par ailleurs, il est facile d'établir que toute combinaison linéaire finie de variables aléatoires issues d'un bruit blanc est, au sens de la Définition 1.3, stationnaire. Cela nous permet d'arriver à la notion de moyenne mobile.

**Définition 1.5.** Un processus  $(X_t)$  est qualifié de moyenne mobile d'ordre  $q \in \mathbb{N}^*$  si, pour tout  $t \in T$ ,

$$X_t = \varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t-q}$$

où  $a \in \mathbb{R}^q$  est un paramètre réel, et  $(\varepsilon_t)$  forme un bruit blanc.

Les valeurs initiales du processus sont traditionnellement choisies dans  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ , lorsque  $T$  est fermé à gauche. Nous utiliserons par la suite l'abréviation  $MA(q)$ , provenant de l'anglais *moving average*, afin de faire référence au processus moyenne mobile d'ordre  $q$ . Il serait judicieux pour être plus rigoureux d'évoquer ici la "stationnarité asymptotique". Les paramètres de dispersion (espérance et variance) du processus ne stabilisent qu'une fois l'influence des variables aléatoires initiales estompée, sauf si ces dernières sont générées selon la loi stationnaire (inconnue dans la majeure partie des cas). Par analogie avec la Définition 1.3, cela se matérialise par le fait que  $\mathbb{E}(X_t)$ ,  $\text{Var}(X_t)$  et  $\text{Cov}(X_{t+k}, X_t)$  convergent vers des limites finies lorsque  $t$  tend vers l'infini.

### 1.1.3 L'autocorrélation

L'un des principaux intérêts de la modélisation chronologique d'une expérience aléatoire est sa structure de corrélation temporelle, en d'autres termes sa structure d'autocorrélation. Il existe plusieurs outils permettant de caractériser l'autocorrélation d'une série chronologique. Ils seront pour nous extrêmement intéressants puisque nous verrons dans la section suivante qu'ils possèdent des propriétés remarquables dans le cadre du processus ARMA stationnaire. Revenons pour le moment au cadre plus général du processus  $(X_t)$  stationnaire, indexé par  $T$  et défini sur  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ . Nous commençons par formaliser la notion de fonction d'autocorrélation qui quantifie l'influence d'un décalage temporel entre deux observations du processus via un calcul standard de corrélation.

**Définition 1.6.** On appelle fonction d'autocorrélation associée au processus stationnaire  $(X_t)$ , la fonction  $\rho$  définie, pour tout décalage temporel  $k \in \mathbb{Z}$ , par

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

où  $\gamma$  est la fonction d'autocovariance donnée par  $\gamma(k) = \text{Cov}(X_k, X_0)$ .

Elle est donc à valeurs dans  $[-1, +1]$ , par construction. Il est alors possible de lui donner une interprétation pour  $k \in \mathbb{Z}$ , elle forme alors une fonction paire. Une représentation graphique, de l'évolution de  $\rho$  avec  $k$  est un autocorrélogramme.

Nous voulons désormais quantifier l'influence exacte d'une observation passée sur la valeur courante du processus en omettant de l'étude toutes les observations intermédiaires. Cette mesure de corrélation est beaucoup moins intuitive donc plus délicate à interpréter, mais nous allons malgré tout la formaliser. Nous bâtissons dans un premier temps la suite  $(\phi_{k,j})$  où l'indice  $k \in \mathbb{N}^*$  joue le rôle de décalage temporel et l'indice  $j$  varie entre 1 et  $k$ . On fixe  $\phi_{1,1} = \rho(1)$  et, pour tout  $k \geq 2$ , on pose

$$\phi_{k,k} = \left(1 - \sum_{i=1}^{k-1} \phi_{k-1,i} \rho(i)\right)^{-1} \left(\rho(k) - \sum_{i=1}^{k-1} \phi_{k-1,i} \rho(i-k)\right)$$

où, lorsque  $i < k$ ,  $\phi_{k,i} = \phi_{k-1,i} - \phi_{k,k}\phi_{k-1,h-j}$ . La valeur de  $\phi_{k,k}$  coïncide avec celle du coefficient de corrélation linéaire entre les résidus issus de la régression linéaire de  $X_k$  et de  $X_0$  sur les observations intermédiaires  $X_1, \dots, X_{k-1}$  respectivement. Formellement, pour tout  $k \in \mathbb{N}^*$ ,

$$\phi_{k,k} = \text{Corr} \left( Y_k - P_{ev\{Y_1, \dots, Y_{k-1}\}}(Y_k), Y_0 - P_{ev\{Y_1, \dots, Y_{k-1}\}}(Y_0) \right) \quad (1.1)$$

où  $P_{ev\{Y_1, \dots, Y_{k-1}\}}$  désigne le projecteur orthogonal de toute variable aléatoire de  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  sur le sous espace vectoriel fermé engendré par  $\{Y_1, \dots, Y_{k-1}\}$ . Cette construction récursive est également appelée algorithme de Durbin-Levinson sur lequel nous reviendrons au cours des premiers chapitres.

**Définition 1.7.** *On appelle fonction d'autocorrélation partielle associée au processus stationnaire  $(X_t)$ , la fonction  $\beta$  définie par  $\beta_0 = 1$  et, pour tout décalage temporel  $k \in \mathbb{N}^*$ , par*

$$\beta_k = \phi_{k,k}.$$

*Cette fonction est également à valeurs dans  $[-1; 1]$ . Lorsque l'on représente graphiquement l'évolution de  $\beta$  avec  $k$ , on construit un autocorrélogramme partiel.*

### 1.1.4 La densité spectrale

Nous introduisons dans cette section la notion de densité spectrale pour des processus stationnaires, intimement liée avec leurs propriétés de corrélation. De fait les autocovariances d'un processus stationnaire  $(X_t)$  coïncident avec les coefficients de Fourier d'une mesure positive, dénommée mesure spectrale, que l'on notera  $\mu_X$  et dont on peut supposer qu'elle est absolument continue par rapport à la mesure de Lebesgue sur le tore  $\mathbb{T} = [-\pi; +\pi]$ . Par souci de simplification, nous considérons ici les processus dont la fonction d'autocovariance est de carrée sommable, car ce sont précisément ceux qui nous occuperont durant la majeure partie de ce travail. Cela implique par ailleurs que  $\gamma(k)$  tend vers 0 lorsque  $k$  tend vers l'infini.

**Définition 1.8.** *Soit un processus  $(X_t)$  stationnaire de fonction d'autocovariance  $\gamma$ . Alors, pour tout  $\lambda \in \mathbb{T}$ , on appelle densité spectrale du processus, la fonction*

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \gamma(k) \exp(-\lambda ik).$$

*De manière réciproque par transformation de Fourier inverse, on a également, pour tout  $k \in \mathbb{Z}$ ,*

$$\gamma(k) = \int_{\mathbb{T}} f(\lambda) \exp(\lambda ik) d\lambda.$$

### 1.1.5 Ratio de vraisemblance localement asymptotiquement normal (LAN)

La propriété LAN (localement asymptotiquement normal) est une propriété très importante en statistique et dépasse largement le cadre des séries chronologiques. Elle fût introduite dans les années 60 par le mathématicien français L.Le Cam et en voici une définition formelle

**Définition 1.9.** Une famille de mesures de probabilités  $\mathbb{P}_\theta^{(n)}$  paramétrées par  $\theta$  est dite LAN en  $\theta_0$  où  $\theta \in M \subset \mathbb{R}^d$  lorsque les conditions suivantes portant sur le ratio de vraisemblance sont toutes vérifiées :

$$L_n(u) = \frac{d\mathbb{P}_{\theta_0 + \phi_n(\theta_0)u}^{(n)}}{d\mathbb{P}_{\theta_0}^{(n)}},$$

$$L_n(u) = \exp \left( \langle u, Z_n(\theta_0) \rangle - \frac{1}{2} \langle u, J(\theta_0)u \rangle + R_n(\theta_0, u) \right),$$

où

$$Z_n(\theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, J(\theta_0)), \quad (1.2)$$

et

$$R_n(\theta_0, u) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} 0 \quad (1.3)$$

sous la mesure  $\mathbb{P}_{\theta_0}^{(n)}$ .

La suite  $(\phi_n(\theta_0))$  vérifie

$$\phi_n(\theta_0) \xrightarrow[n \rightarrow \infty]{} 0.$$

Dans cette définition  $u \in K \subset \mathbb{R}^d$ ,  $\phi_n(\theta_0)$  sont des matrices régulières, et  $J(\theta_0)$  est également une matrice régulière qui sont toutes de taille  $d \times d$ .

Bien que cette propriété semble dénuée d'intérêt au premier abord la vérification de cette dernière entraîne bon nombre de propriétés au sein de l'expérience statistique étudiée. Ces dernières seront présentées mais surtout illustrées au cours des prochains chapitres dans divers contextes. La vérification de la propriété LAN entraîne une notion d'efficacité asymptotique dans l'expérience statistique considérée ce qui, lors de la partie inférence, rend intéressant le recherche d'estimateurs efficaces. Elle permet également de donner une vitesse maximale aux estimateurs qui vaut  $(\phi_n(\theta_0))^{-1}$ .

Nous souhaitons à présent introduire le processus ARMA et, par corollaire, le processus autorégressif. Nous allons réaliser un rappel des résultats classiques existants, en faisant souvent référence aux travaux de Browkwell et Davis dans les années 90. Cela nous permettra d'aborder les chapitres suivants avec un cadre travail clair et bien défini en disposant des outils nécessaires

## 1.2 Le processus ARMA stationnaire

Introduite par Box et Jenkins, ainsi que mentionné dans la section précédente, la modélisation ARMA est issue de l'abréviation anglaise *autoregressive moving average*. L'hypothèse de base de ces travaux affirme que la valeur générique du processus peut être entièrement décrite par une combinaison linéaire de ses  $p$  dernières valeurs, aléatoirement perturbée par une moyenne mobile d'ordre  $q$ , comme décrite dans la Définition 1.5. Avant d'aller plus loin,

pour ne pas alourdir les notations, on a souvent recours à trois opérateurs chronologiques : le retard noté  $L$  comme *lag*, la différenciation notée  $\Delta$ , et la sommation notée  $\nabla$ .

**Définition 1.10.** On définit l'opérateur « retard », sur l'ensemble des processus  $(X_t)$  indexés par  $\mathbb{Z}$ , comme associant à toute valeur générique du processus sa valeur précédente,

$$L(X_t) = X_{t-1}.$$

**Définition 1.11.** Nous définissons également les opérateurs « différenciation » et « sommation », sur la classe des processus  $(X_t)$  indexés par  $\mathbb{Z}$ , comme associant à toute valeur générique du processus les combinaisons

$$\Delta(X_t) = (1 - L)(X_t) = X_t - X_{t-1} \text{ et } \nabla(X_t) = (1 + L)(X_t) = X_t + X_{t-1}$$

où 1 désigne l'opérateur identité. Nous généralisons facilement ces opérateurs à tout décalage temporel  $k \in \mathbb{N}$  par l'intermédiaire de la convention  $L^k(X_t) = (L \circ \dots \circ L)(X_t) = X_{t-k}$ . Ainsi,  $\Delta^k(X_t) = (1 - L)^k(X_t)$ ,  $\Delta_k(X_t) = (1 - L^k)(X_t) = X_t - X_{t-k}$ ,  $\nabla^k(X_t) = (1 + L)^k(X_t)$  et  $\nabla_k(X_t) = (1 + L^k)(X_t) = X_t + X_{t-k}$ .

### 1.2.1 Présentation du modèle

Soit  $(X_t)$  un processus de  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  stationnaire indexé par  $\mathbb{Z}$ , et deux paramètres  $(p, q) \in \mathbb{N}$ .

**Définition 1.12.** Le processus  $(X_t)$  admet une représentation ARMA( $p, q$ ) si, pour tout  $t \in \mathbb{Z}$ , il vérifie la relation récursive

$$X_t - \sum_{i=1}^p \theta_i X_{t-i} = \mu + \sum_{i=1}^q \phi_i \varepsilon_{t-i} + \varepsilon_t$$

où  $(\varepsilon_t)$  est un bruit blanc de variance  $\sigma_2 > 0$ ,  $\mu \in \mathbb{R}$ ,  $\theta \in \mathbb{R}^p$  et  $\phi \in \mathbb{R}^q$ . Lorsque  $q = 0$ , le modèle admet alors la représentation AR( $p$ ) donnée, pour tout  $t \in \mathbb{Z}$ , par la relation récursive

$$X_t = \mu + \sum_{i=1}^p \theta_i X_{t-i} + \varepsilon_t$$

et l'on dit que  $(X_t)$  est un processus autorégressif d'ordre  $p$ . Lorsque  $p = 0$ , le modèle admet la représentation MA( $q$ ) donnée, pour tout  $t \in \mathbb{Z}$ , par

$$X_t = \mu + \sum_{i=1}^q \phi_i \varepsilon_{t-i} + \varepsilon_t$$

et l'on dit que  $(X_t)$  est une moyenne mobile d'ordre  $q$ .

Nous pouvons en outre, sans aucune perte de généralité, considérer que le processus  $(X_t)$  est centré et donc que  $\mu = 0$ . En effet, si l'on note  $m$  l'espérance de la loi stationnaire, on a

$$m = \frac{\mu}{1 - \sum_{i=1}^p \theta_i}$$

et nous constatons alors immédiatement que le changement de variables  $Y_t = X_t - m$  nous ramène par simple translation à la modélisation  $ARMA(p, q)$  centrée, engendrant  $(Y_t)$ . Notons également que nous avons défini la moyenne mobile dans la section précédente, à travers la Définition 1.5, dans un cadre plus restrictif. De manière compacte, à l'aide de l'opérateur retard, cette modélisation s'écrit encore, pour tout  $t \in \mathbb{Z}$ ,

$$\mathcal{A}(L)(X_t) = \mathcal{B}(L)(\varepsilon_t) \quad (1.4)$$

où l'on définit, pour tout  $z \in \mathbb{C}$ , les polynômes  $\mathcal{A}$  et  $\mathcal{B}$  par

$$\mathcal{A}(z) = 1 - \theta_1 z - \dots - \theta_p z^p \text{ et } \mathcal{B}(z) = 1 + \phi_1 z + \dots + \phi_q z^q.$$

Dans le cas du modèle  $AR(p)$ , on a  $\mathcal{B}(z) = 1$  alors que symétriquement, dans le cas du modèle  $MA(q)$ , on a  $\mathcal{A}(z) = 1$ . La modélisation est qualifiée de « minimale » si  $\theta_p \neq 0$ ,  $\phi_q \neq 0$  et si  $\mathcal{A}$  et  $\mathcal{B}$  n'ont pas de racine commune. Sans cela, nous pouvons toujours extraire une formulation  $ARMA(p', q')$  équivalente avec  $p' \leq p$  et  $q' \leq q$  engendrant  $(X_t)$ .

### 1.2.2 Causalité, inversibilité et stationnarité

Notons  $(\lambda_i)$  l'ensemble des racines complexes du polynôme  $\mathcal{A}$ . Si  $|\lambda_i| > 1$  pour tout  $1 \leq i \leq p$ , la modélisation est alors qualifiée de « causale » et il est bien connu que son polynôme autorégressif est alors inversible. En vertu de la Définition 3.1.3 de [20], nous savons par ailleurs que la suite des coefficients du polynôme inverse est absolument sommable. Nous avons alors les relations

$$\mathcal{A}^{-1}(z) = \sum_{k=0}^{\infty} a_k z^k \text{ et } \mathcal{A}^{-1}(z) \mathcal{A}(z) = \sum_{i=0}^{\infty} a_i z^i \prod_{j=1}^p \left(1 - \frac{z}{\lambda_j}\right),$$

ce qui nous permet d'identifier la suite  $(a_k)$ , avec en particulier  $a_0 = 1$ . Il en résulte qu'un processus  $ARMA(p, q)$  causal possède une représentation  $MA(\infty)$  définie, pour tout  $t \in \mathbb{Z}$ , par

$$X_t = \mathcal{A}(L)^{-1} \mathcal{B}(L)(\varepsilon_t). \quad (1.5)$$

La causalité du polynôme autorégressif constitue donc une condition suffisante conférant la stationnarité du processus engendré par la modélisation ARMA. Nous avons également un résultat d'unicité de la solution (1.5) dont le lecteur trouvera une démonstration détaillée à la suite du Théorème 3.1.1 de [20] et de la Remarque 2 associée.

**Proposition 1.13.** *Si  $\mathcal{A}(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| \leq 1$ , alors la représentation ARMA minimale  $\mathcal{A}(L)(Y_t) = \mathcal{B}(L)(\varepsilon_t)$  admet une unique solution stationnaire, causale donnée par (1.5). Les coefficients  $(\psi_k)$  du polynôme  $\mathcal{A}^{-1}(L) \mathcal{B}(L)$  vérifient de plus*

$$\sum_{k=0}^{\infty} |\psi_k| < \infty.$$

Il est possible d'établir via un raisonnement analogue que la causalité du polynôme  $\mathcal{B}$  que la modélisation  $ARMA(p, q)$  admet une représentation  $AR(\infty)$  définie, pour tout  $t \in \mathbb{Z}$ , par

$$\mathcal{B}(L)^{-1} \mathcal{A}(L)(X_t) = \varepsilon_t. \quad (1.6)$$

Notre modélisation est alors qualifiée "inversible". Cette dernière représentation présente un fort intérêt pratique, dans la mesure où, comme cela sera mentionnée, l'estimation des paramètres est beaucoup plus abordable dans un modèle purement autorégressif. Nous prenons dès à présent l'habitude de désigner via le terme générique "stable" un processus solution d'une modélisation ARMA dont le polynôme autorégressif est causal. Par opposition, nous le qualifierons « instable » lorsqu'il est issu d'une modélisation dont le polynôme autorégressif possède au moins une racine sur le cercle unité, et toutes les autres à l'extérieur. Lorsque l'une au moins des racines est située à l'intérieur du cercle unité, nous parlerons alors de processus « explosif », les deux derniers cas étant exclus de nos travaux.

### 1.2.3 Autocorrélations

Nous souhaitons maintenant obtenir les fonctions d'autocorrélation simple et partielle, telles que décrites au sein des Définitions 1.6 et 1.7, respectivement, dans le cadre d'un processus stationnaire  $(X_t)$  générée par une modélisation  $ARMA(p, q)$  causale. Nous avons donc

$$X_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad (1.7)$$

où les coefficients  $(\psi_k)$  sont identifiables par le biais de la relation polynomiale sur  $\mathbb{C}$ ,

$$\mathcal{A}(L)^{-1} \mathcal{B}(L)(z) = \sum_{k=0}^{\infty} \psi_k z^k.$$

A l'exception du cas trivial  $\psi_0 = 1$ , nous chercherons pas à expliciter les coefficients, et ce dans la mesure où leur complexité analytique s'accroît extrêmement rapidement à mesure que  $p$  et  $q$  grandissent. La présentation (1.7) est alors fort agréable dans le calcul de la fonction d'autocovariance  $\gamma$  du processus. Comme  $(\varepsilon_t)$  est un bruit blanc de variance  $\sigma_2$ , un calcul usuel de covariance entraîne bon nombre de simplifications. Nous obtenons alors que, pour tout décalage temporel  $k \in \mathbb{Z}$ ,

$$\gamma(k) = \sigma_2 \sum_{i=0}^{\infty} \psi_i \psi_{i+|k|} \text{ et } \rho(k) = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+|k|}}{\sum_{i=0}^{\infty} \psi_i^2} \quad (1.8)$$

Dans le cadre des processus  $MA(q)$  et  $AR(p)$ , cette écriture se simplifie encore plus. Intéressons nous dans un premier temps à la modélisation  $MA(q)$  pour laquelle  $\mathcal{A}(z) = 1$ , quel que soit  $z \in \mathbb{C}$ . Nous pouvons donc identifier  $\psi_0 = 1$ ,  $\psi_k = \phi_k$  pour tout  $1 \leq k \leq q$  et  $\psi_k = 0$  sinon. Il vient que l'autocovariance et l'autocorrélation de la moyenne mobile d'ordre  $q$  sont données, pour  $|k| \leq q$ , par

$$\gamma(k) = \sigma_2 \sum_{i=0}^{q-|k|} \phi_i \phi_{i+|k|} \text{ et } \rho(k) = \frac{\sum_{i=0}^{q-|k|} \phi_i \phi_{i+|k|}}{\sum_{i=0}^{q-|k|} \phi_i^2} \quad (1.9)$$

avec comme convention  $\phi_0 = 1$ . De plus, dès que  $|k| > q$ , alors  $\gamma(k) = \rho(k) = 0$ . Cette propriété tout à fait remarquable des processus  $MA(q)$  stationnaires s'énonce de manière exhaustive comme suit.

**Proposition 1.14.** *Le processus stationnaire centré  $(X_t)$  est engendré par une modélisation minimale  $MA(q)$  si et seulement si  $\rho(q) \neq 0$  et  $\rho(k) = 0$  pour tout  $|k| > q$ .*

Cette proposition comporte un énoncé plus fort que ceux énoncés ci-dessus lors de nos rappels en formulant l'aspect réciproque. C'est à dire un processus stationnaire centré dont la fonction d'autocorrélation est nulle à partir d'un certain rang s'identifie nécessairement à une moyenne mobile. On se référera à la Proposition 3.2.1 de [20] pour une preuve détaillée de ce résultat. Le processus  $AR(p)$ , correspond au cas  $\mathcal{B}(z) = 1$  pour tout  $z \in \mathbb{C}$  pour lequel une écriture similaire à (1.8) peut être obtenue après identification des coefficients, non dépourvue d'une certaine complexité analytique puisqu'il s'agit toujours d'inverser le polynôme  $\mathcal{A}$ . Nous allons rappeler une procédure beaucoup plus astucieuse reposant sur l'exploitation de la relation linéaire entre les valeurs de  $\gamma(k)$  dans le cadre purement autorégressif. Nous avons d'une part

$$\begin{aligned}\gamma(0) &= \text{Cov}(\theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} + \varepsilon_t, X_t) \\ &= \theta_1 \gamma(1) + \theta_2 \gamma(2) + \cdots + \theta_p \gamma(p) + \sigma_2\end{aligned}$$

grâce au caractère stationnaire de  $(X_t)$  et de la parité de  $\gamma$ . D'autre part, pour tout  $1 \leq k \leq p$ , nous pouvons évaluer  $\gamma(k)$  en fonction de  $\gamma(0), \dots, \gamma(p)$  de façon analogue. À titre informatif, cela nous conduit à, pour  $k = 1$ ,

$$\begin{aligned}\gamma(1) &= \text{Cov}(\theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} + \varepsilon_t, X_{t-1}) \\ &= \theta_1 \gamma(0) + \theta_2 \gamma(1) + \cdots + \theta_p \gamma(p-1)\end{aligned}$$

On construit alors un système de  $p+1$  équations linéaires à  $p+1$  inconnues se formulant

$$\begin{pmatrix} 1 & -\theta_1 & -\theta_2 & \cdots & -\theta_{p-1} & -\theta_p \\ -\theta_1 & 1 - \theta_2 & -\theta_3 & \cdots & -\theta_p & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -\theta_p & -\theta_{p-1} & -\theta_{p-2} & \cdots & -\theta_1 & 1 \end{pmatrix} \begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \vdots \\ \gamma(p) \end{pmatrix} = \begin{pmatrix} \sigma_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (1.10)$$

Cette formulation matricielle est mieux bine connue sous l'appellation « équations de Yule-Walker » que l'on trouvera dans la Section 8.1 de [20], bien que sous une formulation duale. Nous précisions que ce système n'est pas singulier et qu'il admet une solution unique pour  $\gamma(0), \dots, \gamma(p)$  dès lors que la condition de stabilité  $\|\theta\|_1 < 1$  vérifiée, cette dernière entraînant de plus la causalité de la modélisation et donc la stationnarité du processus. Le calcul de  $\gamma(p+1)$  est des valeurs successives s'obtient alors

$$\begin{aligned}\gamma(p+1) &= \text{Cov}(\theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} + \varepsilon_t, X_{t-p-1}) \\ &= \theta_1 \gamma(p) + \theta_2 \gamma(p-1) + \cdots + \theta_p \gamma(1),\end{aligned}$$

et ainsi de suite. La fonction d'autocorrélation est entièrement déterminée en renormalisant  $\gamma(k)$  par  $\gamma(0)$  pour tout  $k \in \mathbb{N}$ . L'espérance nulle et la fonction  $\gamma$  permettent de caractériser les moments d'ordre 2 de la loi stationnaire du processus. L'algorithme de Durbin-Levinson, évoqué dans la Section I.1.3, détaille la construction de manière itérative, à partir des valeurs

de la fonction d'autocorrélation  $\rho$ , de celles de la fonction d'autocorrélation partielle  $\beta$  qui ne possède malheureusement pas d'écriture explicite dans le cas général, à l'exception du cadre autorégressif pur. Nous avons néanmoins la Proposition suivante, par analogie avec la Proposition 1.14, qui découle directement de la Définition (1.1) de  $\beta_k$  dans le cadre d'un processus dont les coefficients de la représentation  $AR(\infty)$  finissent par s'annuler.

**Proposition 1.15.** *Le processus stationnaire centré  $(X_t)$  est engendré par une modélisation minimale  $AR(p)$  si et seulement si  $\beta_p \neq 0$  et  $\beta_k = 0$  pour tout  $h > p$ .*

Nous avons là encore un énoncé réciproque, formulé par exemple au sein du Théorème 3 de [85]. Les Propositions 1.14 et 1.15, à propos desquelles on pourra trouver davantage de détails dans le Chapitre 3 de [20], sont couramment utilisées en pratique afin d'évaluer l'ordre d'une moyenne mobile ou d'une autorégression, par l'étude empirique des valeurs empiriques de  $\rho$  et de  $\beta$ . Sans trop nous attarder sur ces thématiques, nous tenons à rappeler que de nombreux résultats existent en marge des fonctions d'autocorrélation (voir le Chapitre 7 de [20]) : estimateurs consistants, normalité asymptotique, intervalles de confiance, etc. Nous savons de plus que les conditions  $\|\theta\|_1 < 1$  et  $\|\phi\|_1 < 1$  sont suffisantes pour assurer la causalité des polynômes  $\mathcal{A}$  et  $\mathcal{B}$ , et qu'en conséquence,  $\beta_k$  et  $\rho(k)$  tendent à vitesse géométrique vers 0 avec des motifs éventuellement sinusoïdaux, phénomène parfaitement décrit par les Chapitres 3 et 6 de [16]. Nous connaissons exactement leur comportement pour des petites valeurs de  $p$  et de  $q$ . Cela constitue donc un outil graphique supplémentaire pour étudier la stationnarité du processus considéré et les ordres  $p$  et  $q$  d'une éventuelle modélisation ARMA.

### 1.2.4 Densité spectrale

Un processus généré par une modélisation  $ARMA(p, q)$ , non nécessairement causale ni inversible, possède une densité spectrale telle que nous l'avons présentée dans la Section 1.1.4.

**Proposition 1.16.** *Soit le processus  $(X_t)$  stationnaire engendré par la modélisation  $ARMA(p, q)$  minimale  $\mathcal{A}(L)(X_t) = \mathcal{B}(L)(\varepsilon_t)$ , où  $(\varepsilon_t)$  est un bruit blanc de variance  $\sigma_2 > 0$ . Alors, pour  $\lambda \in \mathbb{T} = [-\pi, \pi]$  sa densité spectrale est donnée par*

$$f_X(\lambda) = \frac{\sigma_2}{2\pi} \frac{|\mathcal{B}(\exp(-i\lambda))|^2}{|\mathcal{A}(\exp(-i\lambda))|^2}.$$

Le caractère stationnaire de  $(X_t)$  est directement relié au fait que le polynôme  $\mathcal{A}$  ne s'annule pas sur le cercle unité, garantissant ainsi l'existence de  $f_X$  sur tout le tore  $\mathbb{T}$ .

## 1.3 Inférence pour les processus autorégressifs

Nous allons maintenant nous intéresser à l'aspect estimation des paramètres inconnus associés aux processus  $ARMA(p, q)$ . Nous précisons qu'il est bien connu que l'estimateur du maximum de vraisemblance (sous l'approximation gaussienne) de  $\theta$  et de  $\rho$  issu d'une modélisation ARMA causale et inversible est fortement consistant, comme l'atteste par exemple le Théorème 10.8.1 de [20]. Il satisfait de plus une propriété de normalité asymptotique (voir par

exemple le Théorème 10.8.2) dont la matrice de covariance peut être explicitée facilement, intimement liée à l'information de Fisher dans le cas gaussien. La stratégie classique des moindres carrés n'est pas appropriée à l'estimation du paramètre  $\phi$  car ce dernier gouverne l'évolution d'un processus aléatoire partiellement observable. Elle est en revanche adaptée dans une modélisation autorégressive pure pour laquelle  $\mathcal{B}(z) = 1$ , et c'est ce que nous allons désormais présenter.

### 1.3.1 L'estimateur des moindres carrés

Soit  $(X_t)$  un processus de  $\mathbb{L}^2(\Omega, \mathcal{T}, \mathbb{P})$  tel que, pour tout  $t \in \mathbb{Z}$ , on a

$$X_t = \sum_{i=1}^p \theta_i X_{t-i} + \varepsilon_t \quad (1.11)$$

où  $(\varepsilon_t)$  est un bruit blanc de variance  $\sigma_2$ . Nous supposons que  $\mathcal{A}(z) = 1 - \theta_1 z - \dots - \theta_p z^p$  est un polynôme causal pour  $z \in \mathbb{C}$ . En conséquence,  $(X_t)$  est un processus autorégressif stationnaire d'ordre  $p$ , et l'on observe une trajectoire de ce processus sur l'intervalle de temps discret  $\{0, \dots, n\}$ . Cette dernière est donc décrite par la filtration  $\mathcal{F}(X_0, \dots, X_{p-1}, \varepsilon_0, \dots, \varepsilon_n)$ , pour  $n \geq 0$ . On s'intéresse à la minimisation de la fonctionnelle

$$\Delta_n(\theta) = \sum_{i=0}^n \varepsilon_i^2 = \sum_{i=p}^n \left( X_i - \sum_{j=1}^p \theta_j X_{i-j} \right)^2$$

Dans ce cadre, l'estimateur des moindres carrés de  $\theta$  est donné, pour tout  $n \geq p$ , par

$$\hat{\theta}_n = \left( \sum_{i=p}^n \Phi_{i-1}^p \Phi_{i-1}^{p'} \right)^{-1} \sum_{i=p}^n X_i \Phi_{i-1}^p \quad (1.12)$$

en notant  $\Phi_i^p = (X_i, X_{i-1}, \dots, X_{i-p+1})'$ , pour  $p-1 \leq i \leq n$ . Le caractère inversible de la matrice intervenant dans le calcul de  $\hat{\theta}_n$  peut éventuellement être garanti par l'ajout de la matrice identité d'ordre  $p$ . L'estimateur des moindres carrés possède un comportement analogue à l'estimateur de Yule-Walker obtenu par le biais des estimateurs consistants usuels des autocovariances  $\gamma$ , présentés dans la Section 7.2 de [20], puis par la résolution du système (1.10). Plus précisément, ils coïncident à l'exception des termes de bord qui sont asymptotiquement négligeables dans le cas stable. Ils partagent donc dans ce cas la même normalité asymptotique. Nous tenons à préciser que, dans le cas gaussien, l'expression de cette estimateur coïncide à celui du maximum de vraisemblance qui sera largement étudié au cours des prochains chapitres.

### 1.3.2 Consistance forte

Le résultat de consistance forte que nous allons présenter est établi dans le Théorème 1 de [67], puis dans le Corollaire 1 de [20] dans un contexte plus général.

**Proposition 1.17.** *Soit le processus  $(X_t)$  stationnaire associé au modèle autorégressif (1.11) minimal causal, tel que  $\mathbb{E}(\varepsilon_0) = \sigma_2 < \infty$ . Alors, on a la convergence presque sûre*

$$\lim_{n \rightarrow \infty} \widehat{\theta}_n = \theta \text{ p.s.}$$

Nous supposons ici que  $(\varepsilon_t)$  est un bruit blanc, mais le résultat ci-dessus reste également vrai lorsque  $(\varepsilon_t)$  est une différence de martingale admettant un moment conditionnel d'ordre strictement supérieur à 2 fini. La Proposition précédente tombe malheureusement en défaut dès lors qu'une mémoire est présente dans le bruit. L'estimateur décrit dans 1.12 n'estime plus la quantité voulue et il est alors intéressant de construire de nouveaux estimateurs consistants envers le paramètre autorégressif. Nous considérons à cet égard l'estimateur du maximum de vraisemblance dans un processus autorégressif généré par des bruits gaussiens stationnaires. Intuitivement, nous pouvons pressentir ce problème via le calcul suivant

$$\widehat{\theta}_n = \theta + \left( \sum_{i=p}^n \Phi_{i-1}^p \Phi_{i-1}^{p'} \right)^{-1} \sum_{i=p}^n \Phi_{i-1}^p \varepsilon_i$$

D'où, sous une condition de stationnarité, le terme  $\frac{\sum_{i=p}^n \Phi_{i-1}^p \varepsilon_i}{n}$  ne convergera pas vers 0 lorsque la suite  $(\varepsilon_i)$  ne sera plus un bruit blanc et l'estimateur des moindres carrés sera donc non consistant envers  $\theta$ . Nous montrons aisément que la variance empirique est un estimateur fortement consistant de  $\sigma_2$ . Formellement, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=p}^n \left( X_i - \widehat{\theta}_n \Phi_{i-1}^{p'} \right)^2 = \sigma_2 \text{ p.s.} \quad (1.13)$$

### 1.3.3 Normalité asymptotique

Comme évoqué dans le paragraphe précédent, les estimateurs des moindres carrés et de Yule-Walker partagent la même normalité asymptotique dans le cas stable. Le résultat suivant ainsi qu'une démonstration détaillée sont présentés dans la Section 8.10 de [20].

**Proposition 1.18.** *Soit le processus  $(X_t)$  stationnaire associé au modèle autorégressif (1.11) minimal causal, tel que  $\mathbb{E}(\varepsilon_0) = \sigma_2 < \infty$ . Alors, on a la normalité asymptotique*

$$\sqrt{n} \left( \widehat{\theta}_n - \theta \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( 0, \sigma_2 \Gamma_p^{-1} \right)$$

où  $\Gamma_p$  est la matrice de covariance d'ordre  $p$  du processus stationnaire.

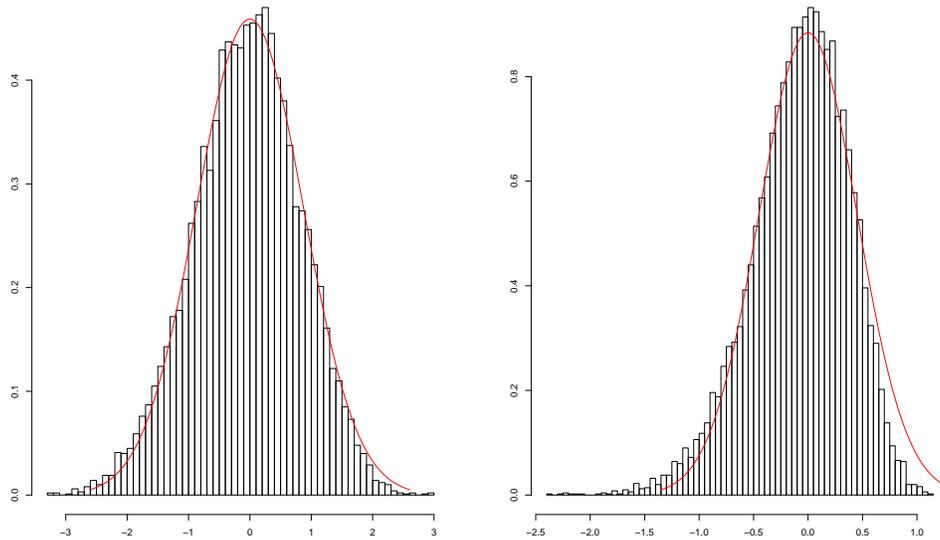


FIGURE 1.1 – Exemple de normalité asymptotique pour un processus  $AR(1)$ . Nous effectuons 10000 simulations Monte-Carlo pour des échantillons de taille 500. Le paramètre autorégressif vaut 0,5 à gauche et 0,9 à droite.

## 1.4 Détection d'une rupture pour les processus autorégressifs

Nous allons à présent présenter le problème de la détection d'une rupture dans  $AR(p)$ . Nous présentons ici le test issu de l'article de Davis [35] basé sur le ratio de vraisemblance. Considérons dans cette section le processus  $(X_t)$  satisfaisant pour tout  $t \in \mathbb{Z}$  la relation récursive

$$X_t = \mu + \sum_{i=1}^p \theta_i X_{t-i} + \varepsilon_t$$

où  $(\varepsilon_t)$  est un bruit blanc fort de variance  $\sigma_2 < \infty$ .

### 1.4.1 Présentation du test

Nous supposons uniquement dans cette sous-section que pour tout  $t \in \mathbb{Z}$ ,  $\varepsilon_t$  est distribuée selon une loi normale d'espérance 0 et de variance  $\sigma_2 = 1$ . Supposons que nous observons une trajectoire sur un intervalle temps discret  $\{0, 1, \dots, n\}$ , nous souhaitons construire un test pour discriminer les hypothèses

$$\mathcal{H}_0 := \forall t \in \{p, \dots, n\}, X_t = \mu + \sum_{i=1}^p \theta_i X_{t-i} + \varepsilon_t$$

contre l'alternative suivante :

$$\mathcal{H}_1 := \exists k \in \{p, \dots, n-1\} \text{ tel que } \begin{cases} \forall t \in \{p, \dots, k\} X_t = \mu + \sum_{i=1}^p \theta_i X_{t-i} + \varepsilon_t \\ \forall t \in \{k+1, \dots, n\} X_t = \tilde{\mu} + \sum_{i=1}^p \tilde{\theta}_i X_{t-i} + \varepsilon_t \end{cases}$$

où  $(\mu, \theta_1, \dots, \theta)^* \neq (\tilde{\mu}, \tilde{\theta}_1, \dots, \tilde{\theta}_p)^*$ . Il est aisé de montrer que sur l'échantillon  $X_{(m)}^{(n)} = (X_m, \dots, X_n)$ , la fonction de log-vraisemblance est donnée par

$$\mathcal{L}(\mu, \theta, X_{(m)}^{(n)}) = -\frac{1}{2} \sum_{t=m}^n \left( X_t - \mu - \theta \phi_{t-1}^{p'} \right)^2 - \frac{n-m+1}{2} \log(2\pi). \quad (1.14)$$

Notons  $(\hat{\mu}_{m,n}, \hat{\theta}_{m,n})$  l'estimateur du maximum de vraisemblance du couple  $(\mu, \theta)$  qui, au passage, coïncide avec l'estimateur des moindres carrés présenté dans la section précédente lorsque  $\mu = 0$ .

Posons

$$\begin{aligned} R(m, n) &= -2\mathcal{L}(\hat{\mu}_{m,n}, \hat{\theta}_{m,n}, X_{(m)}^{(n)}) \\ &= \sum_{i=m}^n X_i^2 - \sum_{i=m}^n X_i \Phi_{i-1}^{p'} \left( \sum_{i=m}^n \Phi_{i-1}^p \Phi_{i-1}^{p'} \right)^{-1} \sum_{i=m}^n X_i \Phi_{i-1}^p \\ &\quad + \frac{(n-m+1)}{2} \log(2\pi) \end{aligned}$$

puis considérons la variable aléatoire

$$\Lambda_n(k) = R(p, k) + R(k+1, n) - R(p, n) \quad (1.15)$$

où  $p \leq k \leq n-1$ .

## 1.4.2 Comportement asymptotique du test

Nous nous affranchissons à présent de l'hypothèse de normalité du vis à vis du processus de nuisance effectuée au sein de la sous-section précédente. Nous souhaitons maintenant donner le comportement asymptotique de la variable aléatoire  $\max_{p \leq k < n} \Lambda_n(k)$  convenablement normalisée. Nous introduisons à cet égard les suites numériques suivantes

$$b_n(p) = \frac{\left( 2 \log \log(n) + \frac{p}{2} \log \log \log(n) - \log \left( \Gamma \left( \frac{p}{2} \right) \right) \right)^2}{2 \log \log(n)},$$

$$a_n(p) = \sqrt{\frac{b_n(p)}{2 \log \log(n)}}$$

où  $\Gamma$  désigne la fonction Gamma et  $p$  l'ordre du processus autorégressif. Nous avons alors le résultat suivant qui donne la convergence en loi de la variable aléatoire  $\max_{p \leq k < n} \Lambda_n(k)$  vers une loi de Gumbel.

**Théorème 1.19.** *Supposons que sous  $\mathcal{H}_0$  le processus  $(X_t)$  est stationnaire associé au modèle autorégressif tel que  $\mathbb{E}(\varepsilon_0^{4+\delta}) < \infty$  où  $\delta > 0$ . Supposons en outre que la suite  $(\varepsilon_t)$  est un bruit blanc fort dont la loi parente est continue, alors sous  $\mathcal{H}_0$  :*

$$\mathbb{P} \left( \sigma_2^{-1} \frac{\max_{p \leq k < n} \Lambda_n(k) - b_n(p+1)}{a_n(p+1)} \leq x \right) \xrightarrow[n \rightarrow \infty]{} \exp \left( -2 \exp \left( -\frac{x}{2} \right) \right)$$

Une preuve détaillée de ce résultat figure dans [35] et repose sur l'utilisation d'un principe d'invariance fort conjuguée à l'application d'un corollaire du Théorème de Darling-Erdős. Nous verrons au cours du Chapitre 3 qu'un résultat similaire peut être obtenu lorsque le processus autorégressif est généré par un bruit gaussien stationnaire. Il n'est malheureusement pas possible d'appliquer la même méthode que dans [35] à cause de l'absence de stationnarité que nous évoquerons au cours du prochain Chapitre.

### 1.4.3 Illustrations numériques

Nous illustrons au sein de ce paragraphe les propriétés numériques du test présenté dans la Théorème ci-dessus. La fonction de répartition empirique de la statistique de test est en noir et celle de la loi limite est en rouge. Nous effectuons à chaque fois 10000 simulations pour des trajectoires de taille  $n = 500$ .

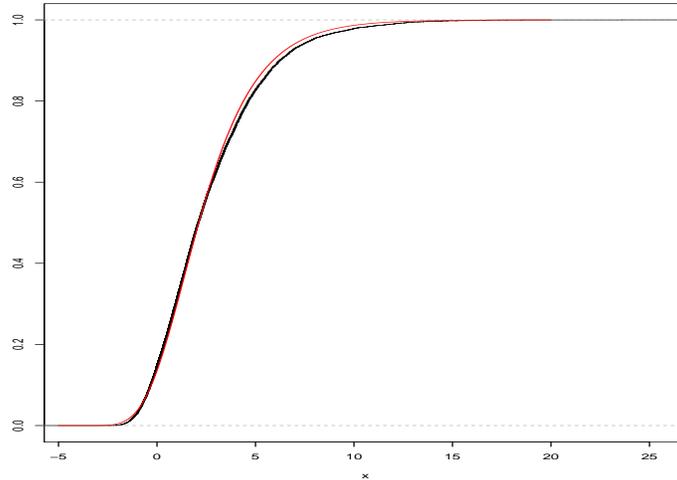


FIGURE 1.2 – Fonction de répartition empirique (en noire) pour 10000 échantillons de taille  $n = 500$  issue d'un processus  $AR(1)$  centré dirigé par un bruit gaussien, centré et réduit. Le paramètre autorégressif est fixé à  $\theta = 0,5$ .

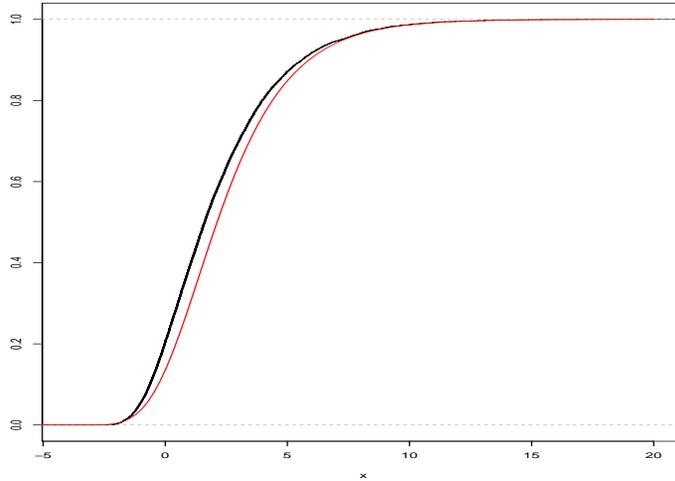


FIGURE 1.3 – Fonction de répartition empirique (en noire) pour 10000 échantillons de taille  $n = 500$  issue d'un processus  $AR(1)$  centré dirigé par un bruit gaussien, centré et réduit. Le paramètre autorégressif est fixé à  $\theta = 0,9$ .

Nous observons un petit décalage entre la fonction de répartition empirique et théorique au sein de la Figure 1.3. Cela pourrait être induit par une convergence assez lente de la fonction de répartition de notre statistique vers celle de sa loi limite. La vitesse de convergence du maximum convenablement normalisé d'une suite *i.i.d.* de  $n$  variables gaussiennes est bien connu et est de l'ordre de  $\frac{1}{\log(n)}$ . Ce dernier résultat est démontré en détails dans [50].

## 1.5 Processus RCAR(1)

### 1.5.1 Présentation du modèle

**Définition 1.20.** *Le processus  $(X_t)$  admet une représentation RCAR(1) si, pour tout  $t \in \mathbb{Z}$ , il est donné par la relation récursive*

$$X_t = (\theta + \eta_t) X_{t-1} + \varepsilon_t$$

où  $(\varepsilon_t)$  est un bruit blanc fort de variance  $\sigma_2 > 0$ , et  $(\eta_t)$  est un bruit blanc fort de variance  $\tau_2 > 0$ , indépendant du processus  $(\varepsilon_t)$ .

Le modèle  $RCAR(1)$  constitue donc une extension du modèle chronologique  $AR(1)$  dans lequel est ajouté un terme d'aléa dans le coefficient. De tels processus furent introduits et étudiés dans [78], [79] et [80]. Nous donnons à présent un premier résultat concernant la stationnarité au premier ainsi qu'au second ordre d'un tel processus.

**Proposition 1.21.** *Supposons que le processus  $(X_t)$  admet une représentation RCAR(1) et que de plus,  $\theta^2 + \tau_2 < 1$  et  $\sigma_2 < +\infty$ . Alors, le processus  $(X_t)$  est strictement stationnaire, causal et également stationnaire au second ordre.*

Une preuve de ce résultat figure dans [79]. Il est également possible de montrer que les processus  $RCAR(1)$  et  $AR(1)$  partagent la même fonction d'autocorrélation lorsqu'ils sont tous deux stationnaires au second ordre. Récemment Aue, Horvath et Steinebach ont donné dans [5] une condition nécessaire et suffisante pour assurer l'existence d'une solution causale, strictement stationnaire et ergodique. Voici un énoncé précis du résultat en question.

**Théorème 1.22.** *Soit  $(X_t)$  un processus vérifiant pour tout  $t \in \mathbb{Z}$  la relation récursive*

$$X_t = (\theta + \eta_t) X_{t-1} + \varepsilon_t$$

où  $(\varepsilon_t)$  est une suite i.i.d et  $(\eta_t)$  est également une suite i.i.d indépendante du processus  $(\varepsilon_t)$ . Supposons que

$$\mathbb{E}(\log^+(|\varepsilon_0|)) < \infty \text{ et } \mathbb{E}(\log(|\theta + \eta_0|)) < \infty.$$

Alors :

- Si  $-\infty \leq (\log^+(|\theta + \eta_0|)) < 0$ , alors le processus  $(X_t)$  admet une solution causale, strictement stationnaire et ergodique.
- Réciproquement, si  $\mathbb{P}(\varepsilon_0 = 0) < 1$  et  $\mathbb{P}(|X_t| < \infty) = 1$ , alors  $-\infty \leq (\log^+(|\theta + \eta_0|)) < 0$ .

## 1.5.2 Consistance forte

Nous reprenons l'estimateur des moindres carrés tel que décrit dans (1.12). Nous donnons un premier résultat garantissant la consistance forte de ce dernier, montrant ainsi que cet estimateur possède la même limite dans ce modèle. Nous rappelons au passage que d'autres procédures d'estimation basée sur la méthode de quasi-vraisemblance furent étudiées dans [5],[11] et [4].

**Théorème 1.23.** *Soit  $(X_t)$  un processus  $RCAR(1)$  vérifiant les conditions  $\theta^2 + \tau_2 < 1$  et  $\sigma_2 < +\infty$ . Supposons que nous observons une trajectoire discrète  $\{0, 1, \dots, n\}$  où  $X_0$  est distribuée selon la loi stationnaire du processus. Alors, l'estimateur des moindres carrés est fortement consistant, i.e.*

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta.$$

Une preuve détaillée de ce résultat figure dans [79]. Le théorème précédent tombe également en défaut lorsqu'une mémoire est présente dans les coefficients ou bien dans le processus de nuisance. Nous allons étudier ce phénomène au cours du Chapitre 4 en considérant un processus autorégressif où les coefficients sont aléatoires et présentent une courte mémoire. Nous verrons alors que l'estimateur des moindres carrés n'estime plus le paramètre autorégressif et nous proposerons une méthode afin de corriger l'estimation.

Nous pouvons également construire un estimateur du couple  $(\tau_2, \sigma_2)$  en remarquant que

$$\text{Var}(X_t - \theta X_{t-1} | X_{t-1}) = \tau_2 X_{t-1}^2 + \sigma_2.$$

Il reste alors à construire le processus résiduel comme nous l'avons fait dans l'estimateur (1.13) puis régresser linéairement ce dernier par rapport à  $(X_{t-1}^2)$  où  $t \in \{1, \dots, n\}$ .

### 1.5.3 Normalité asymptotique

Ce paragraphe donne la normalité asymptotique de notre estimateur à vitesse usuelle à savoir  $\sqrt{n}$ . Contrairement à un processus autorégressif stable, la normalité asymptotique de l'estimateur des moindres carrés dans un modèle  $RCAR(1)$  stationnaire requiert l'existence des moments d'ordre 4 du processus  $(X_t)$ .

**Lemme 1.24.** *Considérons un processus  $(X_t)$  satisfaisant les conditions de la Proposition 1.21. Supposons que d'une part  $\mathbb{E}(\eta_0^4) = \tau_4 < \infty, \mathbb{E}(\varepsilon_0^4) = \sigma_4 < \infty, \mathbb{E}(\eta_0^3) = 0, \mathbb{E}(\varepsilon_0^3) = 0$ , d'autre part  $\theta^4 + 6\theta^2\tau_2 + \tau_4 < 1$ . Alors  $(X_t)$  possède des moments d'ordre 4.*

**Théorème 1.25.** *Soit le processus  $(X_t)$  satisfaisant les conditions du Théorème 1.23. Supposons en outre que les conditions du Lemme 1.24 sont vérifiées. Alors, nous avons la normalité asymptotique suivante*

$$\sqrt{n} \left( \hat{\theta}_n - \theta \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( 0, \frac{\mathbb{E}(X_0^4) \tau_2 + \mathbb{E}(X_0^2) \sigma_2}{\mathbb{E}(X_0^2)^2} \right).$$

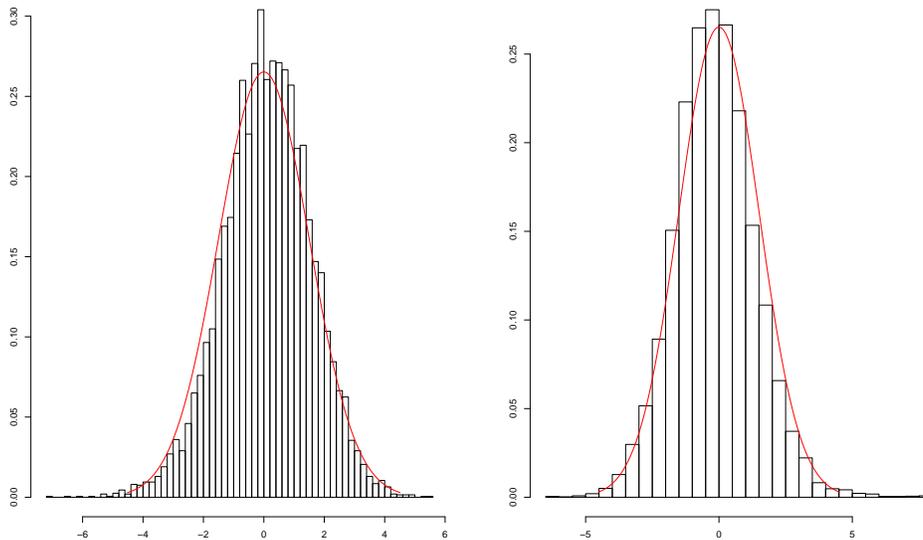


FIGURE 1.4 – Exemple de normalité asymptotique pour un processus  $RCAR(1)$ . Nous effectuons 10000 simulations Monte-Carlo pour des échantillons de taille 500. Le paramètre autorégressif vaut  $-0,3$  à gauche et  $0,3$  à droite.

# Chapitre 2

## Sur les processus autorégressifs à bruits Gaussiens stationnaires

Les résultats que nous allons présenter dans ce chapitre sont résumés dans [91] sous un point de vue très similaire.

### 2.1 Historique et motivation

Nous avons rappelé au sein de la section 1.3 la procédure d'estimation dans un processus autorégressif classique via la méthode des moindres carrés et, en particulier, la consistance forte de l'estimateur des moindres carrés envers le paramètre autorégressif. Outre l'hypothèse de stationnarité formulée au sein de la Proposition 1.18, la conclusion de cette dernière repose également sur le caractère *i.i.d.* de la suite  $(\varepsilon_t)$  intervenant dans la relation récursive (1.11). Il est bien connu que lorsque les perturbations générant le processus  $AR(p)$  sont corrélées, l'estimation fournie par l'estimateur des moindres carrés est alors non consistante pour le paramètre autorégressif. Ce phénomène fut tout d'abord illustré par Mallinvaud [74] puis par Griliches [46]. Ce phénomène fut récemment, de nouveau mis en exergue par Bercu et Proïa dans [9] puis par Proïa dans [82]. Lorsque la structure des perturbations aléatoires est connue, il est alors envisageable de proposer une correction de l'estimation afin d'estimer le paramètre autorégressif comme cela fut proposé dans [9].

Récemment une approche fut développée dans [21] pour estimer le paramètre autorégressif dans une série de données générées par un processus autorégressif dirigé par des bruits Gaussiens stationnaires. Elle repose sur une transformation linéaire des données afin d'obtenir une nouvelle série de données générées par un processus autorégressif où les bruits sont Gaussiens et indépendants. Il est alors possible de construire la fonction de vraisemblance afin d'en extraire l'estimateur par maximum de vraisemblance du paramètre autorégressif. Les propriétés asymptotiques de ce dernier furent établies dans [21] où, il est démontré, sous des hypothèses que nous expliciterons au cours de ce chapitre que cet estimateur est faiblement consistant pour le paramètre autorégressif et qu'il est asymptotiquement normal à vitesse  $\sqrt{n}$ .

Nous nous concentrons au cours de ce chapitre sur deux problématiques distinctes. La première consiste en l'obtention de propriétés asymptotiques presque sûres de l'estimateur du maximum de vraisemblance. Nous montrons que sous une hypothèse supplémentaire, l'estimateur du maximum de vraisemblance est fortement consistant envers le paramètre autorégressif. Nous établissons de plus une loi forte quadratique ainsi qu'une loi du logarithme itérée pour l'estimateur du maximum de vraisemblance nous permettant ainsi d'obtenir une estimation presque sûre sur de la vitesse de convergence de notre estimateur.

Le seconde problématique consiste à obtenir la propriété LAN au sens de la Définition 1.9 afin de construire un test optimal, en un sens que nous préciserons ultérieurement, nous permettant de tester la significativité du paramètre autorégressif. Nous verrons par la suite que, en dépit du caractère Gaussien du bruit du processus post-filtrage, nous ne pouvons appliquer directement les résultats de [31] afin d'établir la propriété LAN. Nous présentons dans ce chapitre l'un des grands intérêts de la propriété LAN, à savoir la possibilité de caractériser un test optimal parmi une classe de tests au sein d'une expérience paramétrique. La question d'optimalité des test fût étudiée dans un premier temps par Wald [100] dans des expériences statistiques générées par des variables aléatoires indépendantes et identiquement distribuées. Ce ne fût véritablement que dans les années 90 que ces problématiques furent traitées dans des expériences statistiques plus générales, toutes assujetties à la propriété LAN, par Hall et Matthiasson [51] puis par Choi, Hall et Schick [28].

## 2.2 Le modèle et les propriétés connues

### 2.2.1 Présentation du modèle ainsi que de la transformation

Nous considérons dans ce chapitre un processus  $(X_n)$  indexé sur  $\mathbb{N}$  satisfaisant la relation récurrente

$$X_n = \sum_{i=1}^p \theta_i X_{n-i} + \xi_n \quad (2.1)$$

et ce, pour tout  $n \in \mathbb{N}$ . Le processus de nuisance  $(\xi_n)$  est un processus Gaussien stationnaire, centré et réduit. Nous supposons de plus que les conditions initiales vérifient  $X_{-p} = \dots = X_{-1} = 0$ . Nous pouvons via la relation (2.1) reformuler notre processus sous forme vectorielle. Posons à cet égard

$$A_0 = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \dots & \theta_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix} \text{ et } Y_n = (X_n, X_{n-1}, \dots, X_{n-p+1})^*.$$

Nous avons donc pour tout  $n$ ,

$$Y_n = A_0 Y_{n-1} + b \xi_n, \quad (2.2)$$

où  $b = (1, 0_{1 \times (p-1)})^*$ .

Nous considérons tout au long de notre étude les hypothèses suivantes :

- $(H_1)$  (condition de stabilité)  $\rho(A_0) < 1$ , où  $\rho$  désigne le rayon spectral de  $A_0$ . L'espace paramétrique est donc  $\Theta = \{\theta \in \mathbb{R}^p | \rho(A_0) < 1\}$ .
- $(H_2)$  La fonction de covariance  $r$  du processus de nuisance vérifie  $r(n) = O(n^{-\alpha})$  lorsque  $n \rightarrow \infty$  pour  $\alpha > 0$  avec  $r(0) = 1$ .
- $(H_3)$  Soit  $(\beta_n)$  la fonction d'autocorrélation partielle (PACF) de  $(\xi_n)$ , nous supposons alors que  $\beta_n^2 = O(n^{-\gamma})$  pour  $\gamma > 1$ .

La dernière hypothèse est légèrement plus forte que  $(\beta_n^2) \in \ell^1(\mathbb{N})$ , qui est vérifiée, mais elle est requise au cours d'un lemme technique. Soit  $f_\xi$  la densité spectrale du processus  $(\xi_n)$ . Nous précisons que  $(H_2)$  peut être reformulée (comme dans [21]) sous une condition formulée sur la densité spectrale. C'est à dire, la densité spectrale  $f_\xi$  appartient à la classe de Szegö, donc  $\log f_\xi$  est intégrable,

$$\int_{-\pi}^{\pi} |\log f_\xi(\lambda)| d\lambda < \infty. \quad (2.3)$$

Intéressons nous à présent à la transformation linéaire de notre processus afin d'obtenir, comme nous le verrons un peu plus tard, un processus de Markov dirigé par des bruits gaussiens indépendants. Posons  $\sigma_1 \varepsilon_1 = \xi_1$ , et pour tout  $n \geq 2$ ,

$$\sigma_n \varepsilon_n = \xi_n - \mathbb{E}(\xi_n | \xi_1, \dots, \xi_{n-1}), \quad (2.4)$$

où  $(\varepsilon_n)$  est une suite *i.i.d.* telle que  $\varepsilon_n \sim \mathcal{N}(0, 1)$ . En vertu du Théorème de la corrélation normale (Théorème 13.1 dans [73]), nous avons la décomposition suivante

$$\sigma_n \varepsilon_n = \sum_{i=1}^n k(n, i) \xi_i, \quad (2.5)$$

où  $(k(n, i))_{\{1 \leq i \leq n, n \in \mathbb{N}^*\}}$  est une suite triangulaire déterministe et  $(\sigma_n^2)$  correspond à la variance des innovations.

Posons

$$\beta_{n-1} = -k(n, 1). \quad (2.6)$$

L'algorithme de Durbin-Levinson tel que décrit dans [40], conduit aux relations itératives suivantes, nous permettant ainsi de réaliser le calcul des coefficients intervenants dans (2.5) :

$$\sigma_n^2 = \prod_{i=1}^{n-1} (1 - \beta_i^2), \quad n \geq 2, \quad \sigma_1 = 1, \quad (2.7)$$

$$\sum_{i=1}^n k(n, i) r(i) = \beta_n \sigma_n^2, \quad k(n, n) = 1, \quad (2.8)$$

$$k(n+1, n+1-i) = k(n, n-i) - \beta_n k(n, i), \quad 1 \leq i \leq n-1. \quad (2.9)$$

Considérons à présent

$$Z_n = \sum_{i=1}^n k(n, i) Y_i, \quad (2.10)$$

ainsi que

$$\zeta_n = \left( \begin{array}{c} Z_n \\ \sum_{k=1}^{n-1} \beta_k Z_k \end{array} \right). \quad (2.11)$$

Il est démontré dans [24] que  $(\zeta_n)$  est un processus de Markov  $2p$ -dimensionnel. Plus précisément, pour tout  $n \in \mathbb{N}^*$ ,

$$\zeta_n = \tilde{A}_{n-1} \zeta_{n-1} + \ell \sigma_n \varepsilon_n, \quad (2.12)$$

où

$$\tilde{A}_n = \left( \begin{array}{cc} A_0 & \beta_n A_0 \\ \beta_n Id_p & Id_p \end{array} \right), \quad \ell = (1, 0_{1 \times (2p-1)})^* \text{ and } \zeta_0 = 0_{2p \times 1}.$$

Le problème initial concernant l'estimation du paramètre  $\theta$  est ainsi remplacé par l'estimation du paramètre  $\theta$  à partir des observations  $\zeta = (\zeta_n, n \geq 1)$ . Nous allons dans le paragraphe suivant rappeler comment nous obtenons l'estimateur du maximum de vraisemblance du paramètre précité ainsi que des propriétés obtenues dans [21], que nous utiliserons par la suite dans nos démonstrations.

## 2.2.2 Construction de l'estimateur du maximum de vraisemblance

La relation (2.12) nous permet donc de construire la fonction de log-vraisemblance qui est donnée par

$$\log \mathcal{L}(\theta, X^{(n)}) = -\frac{1}{2} \sum_{i=1}^n \left( \frac{\ell^*(\zeta_i - \tilde{A}_{i-1} \zeta_{i-1})}{\sigma_i} \right)^2 - \frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^n \sigma_i^2 \quad (2.13)$$

où  $\theta = (\theta_1, \dots, \theta_p)$  et  $X^{(n)} = (X_0, X_1, \dots, X_n)$ .

En utilisant la fonction de log-vraisemblance telle que décrite dans (2.13), il est alors possible d'obtenir une expression analytique explicite de l'estimateur du maximum de vraisemblance de  $\theta$ , donnée par

$$\hat{\theta}_n = \left( \sum_{i=1}^n \frac{a_{i-1}^* \zeta_{i-1} \zeta_{i-1}^* a_{i-1}}{\sigma_i^2} \right)^{-1} \left( \sum_{i=1}^n \frac{a_{i-1}^* \zeta_{i-1} \ell^* \zeta_i}{\sigma_i^2} \right), \quad (2.14)$$

où  $a_n = (Id_p, \beta_n Id_p)^*$ .

Nous définissons la matrice  $I(\theta)$  comme l'unique solution de l'équation de Lyapunov donnée par

$$I(\theta) = A_0^* I(\theta) A_0 + b b^*, \quad (2.15)$$

et nous rappelons les résultats suivants qui sont démontrés dans le Théorème 1 de [21].

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}_\theta^{(n)}} \theta, \quad (2.16)$$

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, I(\theta)^{-1}) \quad (2.17)$$

sous  $\mathbb{P}_\theta^{(n)}$ .

## 2.3 Résultats principaux

### 2.3.1 Propriétés asymptotiques presque sûres du MLE

Dans ce paragraphe, la consistance forte du MLE envers le paramètre autorégressif ainsi qu'une loi forte quadratique sont établies. Nous obtenons également une loi du logarithme itéré (LLI) pour cet estimateur nous permettant d'obtenir une estimation presque sûre sur sa vitesse de convergence. L'ensemble des résultats présentés dans ce paragraphe sont vérifiés sous  $(H_1)$ ,  $(H_2)$  et  $(H_3)$ .

**Théorème 2.1.** *Le MLE est fortement consistant, i.e.*

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta. \quad (2.18)$$

**Preuve.** Voir Section 2.7.1. □

**Théorème 2.2.** *Nous avons la loi forte quadratique suivante pour le MLE :*

$$\frac{1}{\log n} \sum_{k=1}^n (\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^* \xrightarrow[n \rightarrow \infty]{p.s.} I(\theta)^{-1}. \quad (2.19)$$

*La limite est exactement la matrice de covariance asymptotique dans la relation (2.17), où  $I(\theta)$  est définie dans (2.15).*

**Preuve.** Voir Section 2.7.2. □

Pour achever ce paragraphe, nous formulons la LLI du MLE, ce qui nous permet d'en déduire une estimation de la vitesse de convergence du MLE.

**Proposition 2.3.** *Pour tout  $v \in \mathbb{R}^p$ , nous avons :*

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{\frac{1}{2}} v^*(\widehat{\theta}_n - \theta) &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{\frac{1}{2}} v^*(\widehat{\theta}_n - \theta) \\ &= (v^* I(\theta)^{-1} v)^{\frac{1}{2}} \text{ p.s.} \end{aligned}$$

*En conséquence,*

$$\|\widehat{\theta}_n - \theta\|^2 = O\left(\frac{\log \log n}{n}\right) \text{ p.s.} \quad (2.20)$$

**Preuve.** Voir Section 2.7.3. □

### 2.3.2 Propriété LAN et applications

Comme évoqué dans le paragraphe 1.1.5, la propriété LAN est une notion importante via laquelle il est possible de définir une notion d'efficacité asymptotique pour les estimateurs. Avant d'énoncer les résultats obtenus dans la thématique précitée, nous rappelons les propriétés vérifiées par une expérience statistique assujettie à la propriété LAN. La propriété LAN pour des processus Gaussiens stationnaires fut obtenue dans [31] sous des conditions sur la densité spectrale. Ici, nous l'obtenons via un calcul direct sur le ratio de vraisemblance afin d'obtenir la propriété LAN.

**Théorème 2.4.** *Nous supposons que la famille de mesures  $\{\mathbb{P}_\theta^{(n)}, \theta \in M \subset \mathbb{R}^d\}$  est LAN en  $\theta_0$ . Alors pour tout  $\delta > 0$ ,*

$$\liminf_{n \rightarrow \infty} \sup_{\|\phi_n(\theta_0)^{-1}(\hat{\theta}_n - \theta_0)\| \leq \delta} \mathbb{E}_{\theta_0}^{(n)} \left( f \left( \phi_n(\theta_0)^{-1}(\hat{\theta}_n - \theta_0) \right) \right) \geq \int_{\mathbb{R}^d} f \left( J(\theta_0)^{-\frac{1}{2}} x \right) \Phi_d(x) dx, \quad (2.21)$$

ce pour tout estimateur  $\hat{\theta}_n$  et pour toute fonction coût  $f$  telle que  $f$  est continue, symétrique, quasi-convexe et  $f(z) \exp(-\frac{\|z\|^2}{2}) \rightarrow 0$  quand  $\|z\| \rightarrow \infty$ . Dans cette définition,  $\Phi_d$  est la densité d'un vecteur Gaussien  $d$ -dimensionnel standard.

Une preuve détaillée du résultat est formulée dans le Théorème 12.1, Chapitre 2 de [60]. Nous pouvons à présent énoncer la propriété LAN pour le modèle en question.

**Théorème 2.5.** *Sous les notations de la Définition 1.9, si  $\phi_n(\theta_0) = \frac{1}{\sqrt{n}} Id_p$ , alors*

$$\log \left( \frac{\mathcal{L}(\theta_0 + \phi_n(\theta_0)u, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) = \left\langle u, \frac{M_n}{\sqrt{n}} \right\rangle - \frac{1}{2} \langle u, I(\theta_0)u \rangle + R_n(\theta_0, u), \quad (2.22)$$

où  $(\frac{M_n}{\sqrt{n}})$  vérifie la condition (1.2) sous  $\mathbb{P}_{\theta_0}^{(n)}$  avec  $J(\theta_0) = I(\theta_0)$ , et  $R_n(\theta_0, u)$  vérifie la condition (1.3) sous  $\mathbb{P}_{\theta_0}^{(n)}$ . Dans ce Théorème,  $u \in B(0; R)$  pour tout  $R > 0$ .

**Preuve.** Voir Section 2.7.4. □

Nous sommes à présent en mesure de formuler un résultat concernant l'efficacité asymptotique du MLE.

**Proposition 2.6.** *Sous  $(H_1)$  et  $(H_2)$ , le MLE est asymptotiquement efficace. Plus précisément, la borne inférieure donnée par le Théorème 2.4 est atteinte.*

**Preuve.** Voir Section 2.7.5. □

Nous souhaitons désormais construire un test optimal pour tester la significativité du paramètre autorégressif. Nous devons dans un premier temps conférer un sens précis au mot "optimal" dans la cadre d'un test. A cet égard, nous rappelons des notions ainsi que des résultats formulés dans [28]. Nous supposons dans tout ce qui suit que la famille de mesures  $\mathbb{P}_\theta^{(n)}$  est LAN en  $\theta_0$ , et nous souhaitons donc, comme mentionné précédemment construire une procédure optimale afin de tester  $\theta = \theta_0$  contre  $\theta \neq \theta_0$ .

**Définition 2.7.** Un test  $\phi_n^1$  est dit AUMP( $\alpha$ ) (asymptotiquement uniformément le plus puissant de niveau  $\alpha$ ) lorsque,

$$\limsup_{n \rightarrow \infty} \mathbb{E}_{\theta_0}^{(n)}(\phi_n^1) \leq \alpha, \quad (2.23)$$

et pour tout autre test  $\phi_n^2$  de niveau asymptotique  $\alpha$ ,

$$\limsup_{n \rightarrow \infty} \mathbb{E}_{\theta_0 + \phi_n(\theta_0)u}^{(n)}(\phi_n^2) \leq \liminf_{n \rightarrow \infty} \mathbb{E}_{\theta_0 + \phi_n(\theta_0)u}^{(n)}(\phi_n^1). \quad (2.24)$$

**Remarque 2.3.1** Nous énonçons un lemme formulé dans [28] afin de formaliser la prochaine Définition. Nous le formulons dans le cadre de notre étude (sans les paramètres de nuisance), car dans notre cas nous les calculons (via l'algorithme de Durbin-Levinson).

**Lemme 2.8.** Sous les notations de la Définition 1.9, pour tout test  $\phi_n^1$  et toute sous suite  $n'$ , nous pouvons extraire une sous suite de cette dernière  $n''$ , et un test  $\phi$  de  $\mathbb{R}^p$  dans  $[0; 1]$  tels que pour tout  $u \in K$ ,

$$\lim_{n'' \rightarrow \infty} \mathbb{E}_{\theta_0 + \phi_{n''}(\theta_0)u}^{(n'')}(\phi_{n''}^1) = \int_{\mathbb{R}^d} \phi(x) \Phi_d(x - J(\theta_0)u) dx, \quad (2.25)$$

où  $\Phi_d$  est définie dans le Théorème 2.4.

Nous restreignons à présent la classe des test optimaux à ceux satisfaisant un principe d'invariance par rotation formulé rigoureusement dans la définition suivante.

**Définition 2.9.** Un test  $\phi_n^1$  est AUMPI( $\alpha$ ) lorsque la condition de la Définition 2.7 est satisfaite et, de plus, pour toute sous suite  $n'$ , le test correspondant  $\phi$  (obtenu via le Lemme 2.8) vérifie  $\phi(Ru) = \phi(u)$  pour toute rotation  $R$  de  $\mathbb{R}^d$  dans  $\mathbb{R}^d$ .

Pour terminer ce paragraphe, nous formulons une procédure AUMPI pour tester la significativité du paramètre autorégressif.

**Théorème 2.10.** Le test

$$\tilde{\phi}_n = \mathbb{1} \left\{ 2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) \geq C_{1-\alpha} \right\} \quad (2.26)$$

pour tester  $\theta = \theta_0$  contre  $\theta \neq \theta_0$ , est AUMPI( $\alpha$ ), où  $C_{1-\alpha}$  est la quantile d'ordre  $1 - \alpha$  d'une variable aléatoire distribuée selon un  $\chi_p^2$ .

**Preuve.** Voir Section 2.7.6 □

**Remarque 2.3.2** Nous supposons dans ce chapitre que la covariance du bruit doit être connue afin de procéder au filtrage. Ce modèle peut par exemple être appliqué en mathématiques financières où est étudié un modèle AR(1) dirigé par un fGn d'exposant de Hurst connu (voir [45]).

**Remarque 2.3.3** Dans le cas où la variance  $\rho(0) = \sigma^2$  est inconnue, nous tenons à signaler que l'estimateur  $\hat{\theta}_n$  ne dépend pas de  $\sigma$  dans le filtrage. En conséquence l'estimateur  $\hat{\sigma}_n^2 = \arg \max_{\sigma^2 > 0} L(N, \hat{\theta}_n)$  peut être utilisé afin d'estimer la variance après le filtrage comme dans un processus autorégressif standard.

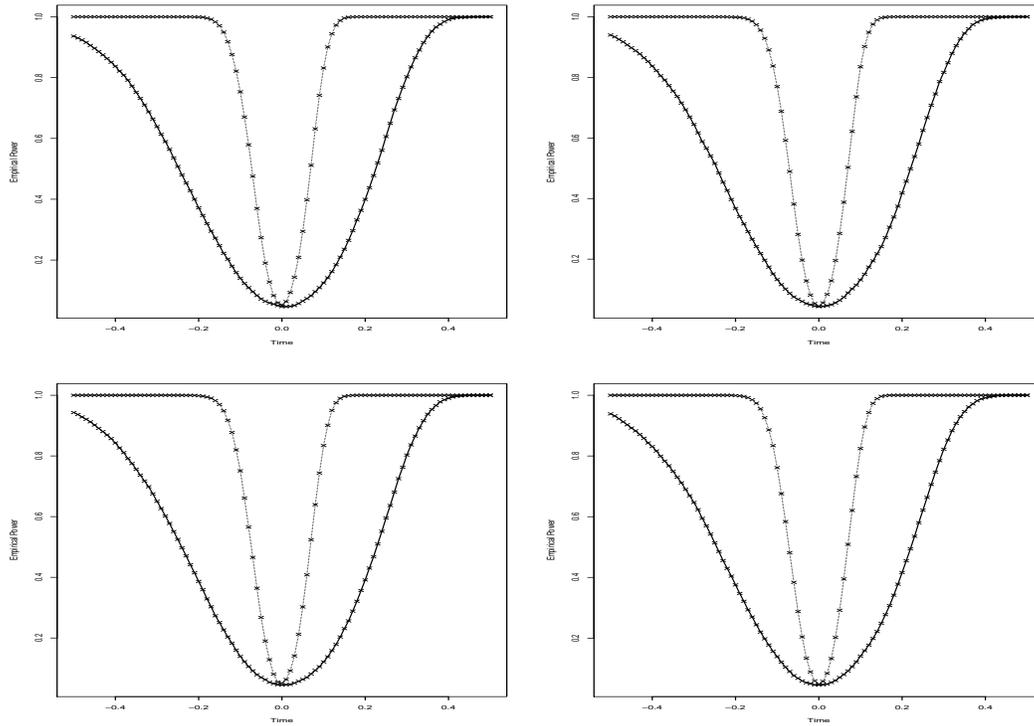
## 2.4 Simulations

Dans cette section, nous illustrons les propriétés asymptotiques du test  $\tilde{\phi}_n$  formulées au sein du Théorème 2.10 en évaluant sa puissance empirique dans différents cas. Nous notons  $\theta_0$  le vrai paramètre utilisé pour effectuer les simulations, puis nous choisissons 100 alternatives appartenant à  $\prod_{i=1}^p [\theta_0^{(i)} - 0,5; \theta_0^{(i)} + 0,5]$  où  $p$  est l'ordre d'autorégression et  $\theta_0^{(i)}$  est la  $i$ -ième composante de  $\theta_0$ . La composante de chaque alternative varie selon un pas régulier de 0.01 et nous effectuons 10000 simulations dans chacun des cas. L'axe des abscisses qui est noté "Time" correspond à la variation de chaque alternative (en norme 1) par rapport au bon paramètre  $\theta_0$  et l'axe "Empirical Power" représente naturellement la puissance empirique de la valeur correspondante. Nous réalisons les simulations pour des échantillons de taille  $n = 50$  (ligne pleine) ainsi que pour des échantillons de taille  $n = 500$  (ligne en pointillée). Nous réalisons les simulations pour divers processus de nuisance en fixant à chaque fois le risque à  $\alpha = 0,05$ .

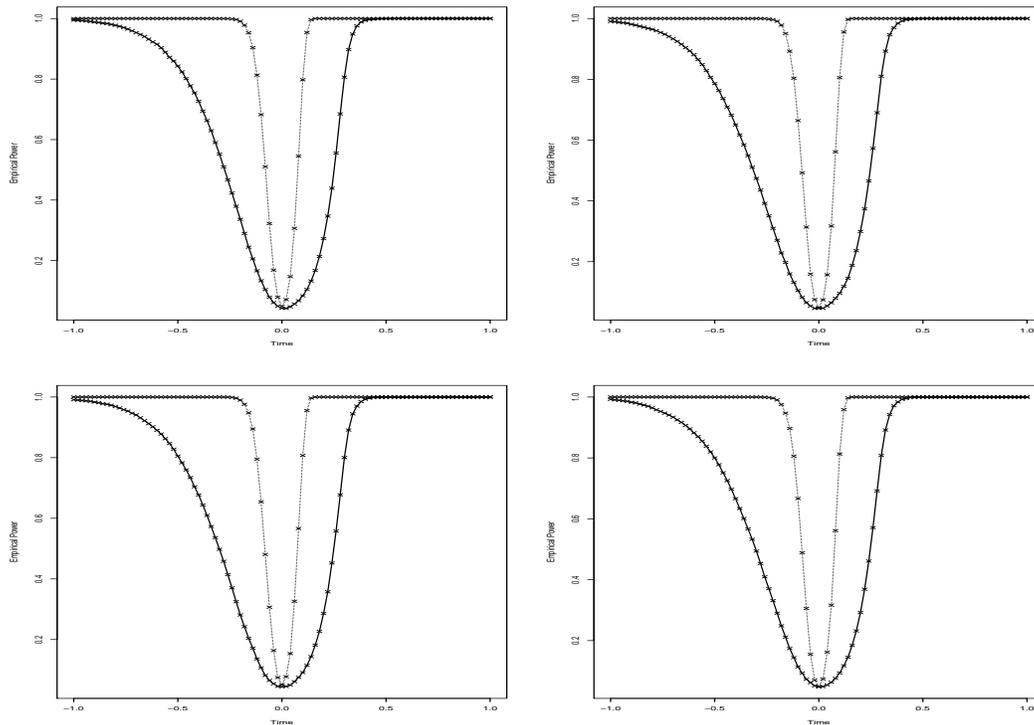
Nous utilisons les processus de nuisance suivant :

- (en haut à gauche) Le processus de nuisance est une moyenne mobile d'ordre 1, MA(1). Concrètement  $\xi_n = \varepsilon_n + a\varepsilon_{n-1}$  où  $(\varepsilon_n)$  est une suite *i.i.d.* distribuée suivant une loi normale values de variance  $d = 0,8$  et  $a = -0,5$ .
- (en haut à droite) Le processus de nuisance est AR(1) gaussien d'où  $\xi_n = \rho\xi_{n-1} + \varepsilon_n$  où la suite  $(\varepsilon_n)$  est encore une fois *i.i.d.* de loi parente Normale de variance  $d = 1 - \rho^2$  avec  $\rho = 0,4$ .
- (en bas à gauche) Le processus de nuisance est un ARMA(1,1) gaussien. Nous avons donc  $\xi_n = \rho\xi_{n-1} + z_n$  où  $(z_n)$  est un processus MA(1) gaussien dont les paramètres sont  $a = -0,5$  et  $d = \frac{1}{1 + \frac{(a+\rho)^2}{1-\rho^2}}$ . Ici  $\rho = 0,6$ .
- (en bas à droite) Le processus de nuisance est un bruit gaussien Fractionnaire. Le processus  $(\xi_n)$  est donc stationnaire et le paramètre de Hurst est fixé à  $H = 0,7$ .

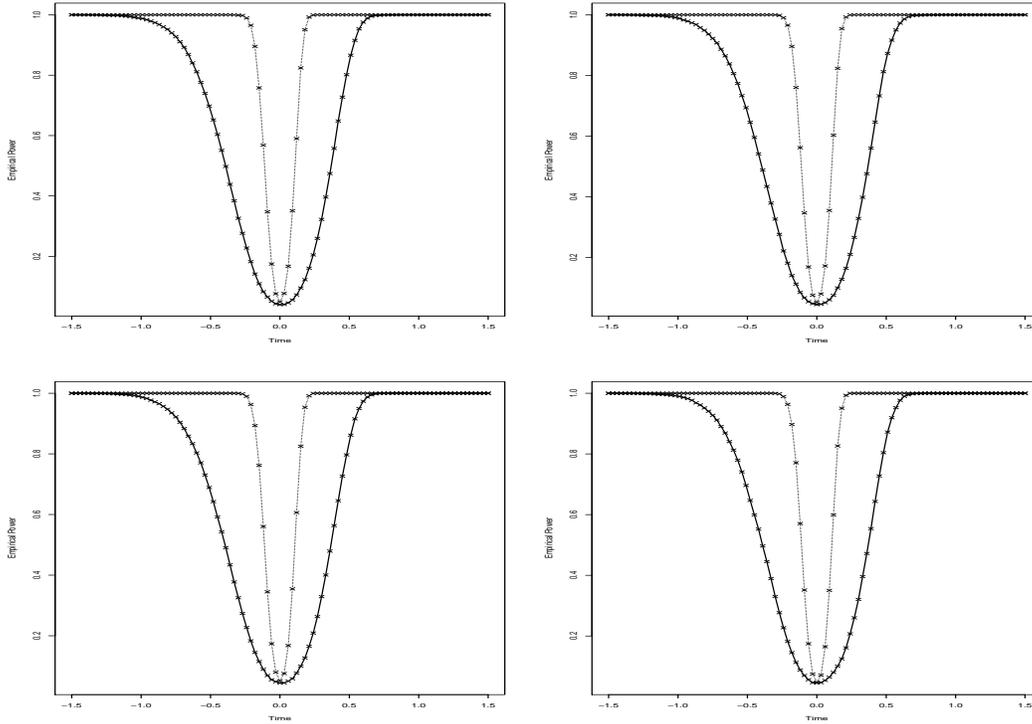
### 2.4.1 Simulation pour un processus AR(1) avec $\theta_0 = 0,6$



### 2.4.2 Simulation pour un processus AR(2) avec $\theta_0 = (0, 3; 0, 5)^*$



### 2.4.3 Simulation pour un processus AR(3) avec $\theta_0 = (0, 3; 0, 6; -0, 4)^*$



Les simulations sont assez chrono-phage et nécessitent l'utilisation d'un serveur de calculs . Nous observons que la puissance empirique ne dépend pas du processus  $(\xi_n)$  ce qui est du au filtrage permettant de blanchir les perturbations aléatoires. Nous observons également que la puissance empirique s'approche assez rapidement de 1 à mesure que l'alternative s'éloigne de  $\theta_0$ .

## 2.5 Conclusion et Perspectives

Nous venons de montrer, sous une hypothèse supplémentaire portant sur la PACF, que l'estimateur du maximum de vraisemblance est fortement consistant envers le paramètre autorégressif. Nous avons également obtenu une estimation sur la vitesse de convergence de cet estimateur. Dans un autre registre, nous avons obtenu la propriété LAN pour cette expérience statistique qui nous a permis de construire une procédure optimale afin de tester la significativité du paramètre autorégressif. Il reste cependant un problème en suspens, nous devons connaître la fonction de covariance afin de procéder au filtrage ce qui, en pratique, n'est pas toujours le cas. Dans le cadre du bruit gaussien fractionnaire nous pensons avoir construit, même si la formalisation reste à venir, un estimateur de l'exposant de Hurst en utilisant la densité spectrale conjuguée à méthode présentée dans [86]. Il serait alors intéressant d'étudier l'estimateur du paramètre autorégressif en injectant l'estimation de  $H$  dans le filtrage.

## 2.6 Résultats Auxiliaires

Cette Section est dévolue aux nombreux résultats techniques que nous invoquerons ensuite dans la Section suivante afin de démontrer les résultats présentés dans 2.3. Pour commencer nous rappelons des résultats techniques établis dans [21]. L'erreur statistique admet la décomposition suivante

$$\widehat{\theta}_n - \theta = (\langle M \rangle_n)^{-1} M_n, \quad (2.27)$$

où

$$\langle M \rangle_n = \sum_{i=1}^n \frac{a_{i-1}^* \zeta_{i-1} \zeta_{i-1}^* a_{i-1}}{\sigma_i^2} \text{ et } M_n = \sum_{i=1}^n \frac{a_{i-1}^* \zeta_{i-1} \varepsilon_i}{\sigma_i}.$$

Posons  $\mathcal{F}_n = \mathcal{F}(X_0, X_1, \dots, X_n)$  qui est la  $\sigma$ -algèbre engendrée par les valeurs du processus jusqu'à l'instant  $n$ . Il vient que  $(M_n)$  est une  $\mathcal{F}_n$ -martingale, et que  $(\langle M \rangle_n)$  est sa variation quadratique prévisible.

Nous avons en outre,

$$\sum_{i=1}^{\infty} \beta_i^2 < \infty, \quad (2.28)$$

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}_\theta^{(n)}} I(\theta), \quad (2.29)$$

où pour rappel  $I(\theta)$  est définie dans (2.15),

$$\sigma_n^2 \xrightarrow[n \rightarrow \infty]{} \sigma_\infty^2 > 0, \quad (2.30)$$

$$\frac{M_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, I(\theta)) \quad (2.31)$$

sous  $\mathbb{P}_\theta^{(n)}$ .

**Remarque 2.6.1** *Le processus  $(\zeta_n)$  obtenu dans (2.11) est un processus de Markov, malheureusement inhomogène, ce qui obère la stationnarité de  $(\zeta_n)$  et donc l'utilisation des arguments classiques liés à l'ergodicité dans un processus  $AR(p)$  stable. La première étape dans nos preuves consiste à montrer que les  $p$  premières composantes de  $(\zeta_n)$  ont le même comportement asymptotique que un processus  $AR(p)$  stable, au sens de la convergence presque sûre.*

**Proposition 2.11.** *Notons pour la suite  $(\zeta_n^{(1)})$  les  $p$  premières composantes de  $(\zeta_n)$  tel que décrit dans (2.12), et  $(\zeta_n^{(2)})$  les  $p$  dernières composantes. Considérons*

$$\gamma_n = A_0 \gamma_{n-1} + \ell^1 \sigma_\infty \varepsilon_n, \quad \forall n \geq 1,$$

où la condition initiale  $\gamma_0$  est distribuée selon la loi strictement stationnaire et ergodique associée au processus autorégressif classique, et  $\ell^1 = (1, 0 \dots, 0)^*$  est un vecteur de taille  $p$ . Alors,

$$\|\zeta_n^{(1)} - \gamma_n\| \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (2.32)$$

**Preuve.** Voir Section 2.6.1. □

**Remarque 2.6.2** Comme  $\rho(A_0) < 1$ , le processus  $(\gamma_n)$  admet une unique représentation possédant les propriétés suivantes : causalité, stationnarité stricte et ergodicité (Des détails à ce sujet sont donnés dans [20]).

Ces propriétés seront utilisées par la suite pour obtenir les convergences presque sûre via des arguments d'ergodicité. Comme  $(\gamma_n)$  est un processus Gaussien ergodique,  $\mathbb{E}(\|\gamma_n\|^p) = K_p < \infty$  pour tout  $p$ .

**Lemme 2.12.** Considérons un vecteur aléatoire  $G_n \in \mathbb{R}^d$  tel que pour tout  $n \geq 1$ ,

$$G_n \sim \mathcal{N}(0, A_n),$$

où la matrice de covariance vérifie  $\|A_n\| = O(n^{-\delta})$  pour un  $\delta > 0$ . Alors,

$$\|G_n\| \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (2.33)$$

**Preuve.** La preuve suit le même cheminement que celle du Lemme A.1 de [31]. Pour tout  $\varepsilon > 0$ , nous avons

$$\begin{aligned} \mathbb{P}(\|G_n\| > \varepsilon) &= \mathbb{P}(\|G_n\|^2 > \varepsilon^2) \\ &= \mathbb{P}(\langle A_n \mu_n, \mu_n \rangle > \varepsilon^2) \leq \mathbb{P}\left(\|\mu_n\|^2 > \frac{\varepsilon^2}{\|A_n\|}\right), \end{aligned}$$

où  $\|\mu_n\|^2 \sim \chi^2(d)$ , ce qui, en conséquence, implique

$$\mathbb{P}(\|\mu_n\|^2 > \varepsilon^2 \|A_n\|^{-1}) = \int_{\varepsilon^2 \|A_n\|^{-1}}^{\infty} c(d) x^{\frac{d}{2}-1} \exp\left(-\frac{x}{2}\right) dx \leq \int_{\varepsilon^2 \|A_n\|^{-1}}^{\infty} c_\beta(d) x^{-\beta} dx$$

pour tout  $\beta > 1$  et  $n$  assez grand, où  $c(d)$  and  $c_\beta(d)$  sont deux constantes positives indépendantes de  $x$  et  $n$ . En utilisant l'hypothèse portant sur  $\|A_n\|$ , nous obtenons

$$\int_{\varepsilon^2 \|A_n\|^{-1}}^{\infty} c_\beta(d) x^{-\beta} dx = \frac{c_\beta(d) \varepsilon^{2(1-\beta)} \|A_n\|^{\beta-1}}{\beta-1} = O(n^{\delta(1-\beta)})$$

résultant de  $\beta > 1$ . Choisissons  $\beta$  de sorte que  $\delta(1-\beta) < -1$ , i.e.  $\beta > \frac{1+\delta}{\delta}$ . Il ne reste plus qu'à appliquer le Lemme de Borel-Centelli afin d'obtenir le résultat voulu. □

Le Lemme suivant nous donne un contrôle de la norme de la matrice de transition  $\tilde{A}_n$ .

**Lemme 2.13.** Soit  $\tilde{A}_n$  la matrice de transition définie dans (2.12), c'est à dire

$$\forall n \geq 1, \tilde{A}_n = \begin{pmatrix} A_0 & \beta_n A_0 \\ \beta_n Id_p & Id_p \end{pmatrix} \text{ avec } \tilde{A}_0 = \begin{pmatrix} A_0 & A_0 \\ Id_p & Id_p \end{pmatrix}.$$

Alors,

$$\sup_{n \in \mathbb{N}} \left\| \prod_{i=1}^n \tilde{A}_{n-i} \right\| < \infty. \quad (2.34)$$

**Preuve.** Soit  $T_0 = Id_{2p}$ , et pour tout  $n \geq 1$ ,

$$T_n = \prod_{i=1}^n \tilde{A}_{n-i}.$$

En travaillant block par block, nous obtenons aisément via la relation récursive  $T_{n+1} = \tilde{A}_n T_n$  que

$$\begin{cases} T_{n+1}^{(11)} = A_0 T_n^{(11)} + \beta_n A_0 T_n^{(21)} \\ T_{n+1}^{(12)} = A_0 T_n^{(12)} + \beta_n A_0 T_n^{(22)} \\ T_{n+1}^{(21)} = \beta_n T_n^{(11)} + T_n^{(21)} \\ T_{n+1}^{(22)} = \beta_n T_n^{(12)} + T_n^{(22)} \end{cases}$$

où nous utilisons la notation

$$T_n = \begin{pmatrix} T_n^{(11)} & T_n^{(12)} \\ T_n^{(21)} & T_n^{(22)} \end{pmatrix}.$$

Puisque  $\rho(A_0) < 1$ , il existe une norme matricielle  $\|\cdot\|_* = \sup(|\cdot u|_*; u \in \mathbb{C}^p, |u|_* = 1)$  satisfaisant  $\|A_0\|_* < 1$  (voir Proposition 2.3.15 de [39]). Il vient que

$$\begin{cases} \|T_{n+1}^{(11)}\|_* \leq a_0 \|T_n^{(11)}\|_* + |\beta_n| a_0 \|T_n^{(21)}\|_* \\ \|T_{n+1}^{(21)}\|_* \leq |\beta_n| \|T_n^{(11)}\|_* + \|T_n^{(21)}\|_* \end{cases}$$

où, pour une meilleure lisibilité, nous avons posé  $a_0 = \|A_0\|_*$ . Il existe donc une constante  $C > 0$  telle que

$$\|T_{n+1}^{(11)}\|_* + \|T_{n+1}^{(21)}\|_* \leq C \prod_{i=1}^n \|H_i\|, \quad (2.35)$$

où

$$\forall i \geq 1, H_i = \begin{pmatrix} a_0 & |\beta_i| a_0 \\ |\beta_i| & 1 \end{pmatrix}.$$

Pour déterminer  $\|H_i\|$ , nous allons déterminer les valeurs propres de  $H_i^*H_i$ . Soit

$$P_i(\lambda) = \lambda^2 - \lambda(1 + a_0^2)(1 + \beta_i^2) + a_0^2(1 - \beta_i^2)^2$$

le polynôme caractéristique de  $H_i^*H_i$  défini pour tout  $\lambda \in \mathbb{C}$ . Alors, un calcul nous indique que le discriminant de l'équation  $P_i(\lambda) = 0$  est

$$\Delta_i = (1 - a_0^2)^2 + (1 + 6a_0^2 + a_0^4)2\beta_i^2 + (1 - a_0^2)^2\beta_i^4 > 0, \quad (2.36)$$

conduisant ainsi aux valeurs propres

$$\lambda_{k,i} = \frac{(1 + a_0^2)(1 + \beta_i^2) \pm \sqrt{\Delta_i}}{2}, \quad k = 1, 2.$$

Puisque  $\beta_n \rightarrow 0$ , un développement de Taylor de  $\sqrt{\Delta_i}$  s'écrit

$$\lambda_{k,i} = \frac{(1 + a_0^2)(1 + \beta_i^2)}{2} \pm \frac{1 - a_0^2}{2} [1 + K_i\beta_i^2 + o(\beta_i^2)]$$

pour un terme aisément identifiable  $|K_i| = O(1)$  lorsque  $i \rightarrow \infty$ . En conséquence,

$$\lambda_{1,i} = a_0^2 + O(\beta_i^2) \text{ et } \lambda_{2,i} = 1 + O(\beta_i^2),$$

impliquant  $\|H_i\|^2 = 1 + O(\beta_i^2)$ . Cette estimation conjuguée à (2.35) et à (2.28) implique  $\|T_{n+1}^{(11)}\|_* + \|T_{n+1}^{(21)}\|_* < \infty$ . Le raisonnement symétrique étant valide pour  $\|T_{n+1}^{(12)}\|_* + \|T_{n+1}^{(22)}\|_*$ , le résultat attendu est donc prouvé.  $\square$

Le Lemme suivant traite la convergence presque-sûre de  $(\beta_n\zeta_n)$ .

**Lemme 2.14.** *Considérons le processus  $(\zeta_n)$  défini dans (2.12). Alors,*

$$\|\beta_n\zeta_n\| \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (2.37)$$

**Preuve.** Nous déduisons de (2.12) que

$$\zeta_n = \sum_{k=1}^{n-1} \left( \prod_{i=1}^k \tilde{A}_{n-i} \right) \ell \sigma_{n-k} \varepsilon_{n-k} + \ell \sigma_n \varepsilon_n.$$

Alors,

$$\begin{aligned} \mathbb{E}(\beta_n^2 \zeta_n \zeta_n^*) &= \beta_n^2 \sum_{k=1}^n P_{k,n} \ell \sigma_{n-k}^2 \mathbb{E}(\varepsilon_{n-k}^2) \ell^* P_{k,n}^* + \beta_n^2 \ell \ell^* \sigma_n^2 \mathbb{E}(\varepsilon_n^2) \\ &= \beta_n^2 \left( \sum_{k=1}^n P_{k,n} \ell \sigma_{n-k}^2 \ell^* P_{k,n}^* + \ell \sigma_n^2 \ell^* \right), \end{aligned}$$

où pour tout  $n - 1 \geq k \geq 1$ ,  $P_{k,n} = \tilde{A}_{n-1} \dots \tilde{A}_{n-k}$ . En vertu du Lemme 2.13,

$$\|\mathbb{E}(\beta_n^2 \zeta_n \zeta_n^*)\| \leq \beta_n^2 \left( \sum_{k=1}^n \|P_{k,n}\|^2 \sigma_{n-k}^2 + \sigma_n^2 \right) \leq Kn\beta_n^2$$

pour une constante  $K > 0$  et  $n$  assez grand. Alors, l'hypothèse  $(H_3)$  nous conduit à

$$\|\mathbb{E}(\beta_n^2 \zeta_n \zeta_n^*)\| \leq \frac{K}{n^{\gamma-1}}$$

et le Lemme 2.12 achève la démonstration.  $\square$

Nous pouvons désormais utiliser les Lemmes précédents afin de démontrer la Proposition 2.11.

### 2.6.1 Preuve de la Proposition 2.11

**Preuve.** Un calcul direct montre que

$$\zeta_n^{(1)} - \gamma_n = A_0(\zeta_{n-1}^{(1)} - \gamma_{n-1}) + \beta_{n-1}A_0\zeta_{n-1}^{(2)} + \ell^1(\sigma_n - \sigma_\infty)\varepsilon_n.$$

En vertu du Lemme 2.14,  $\|\beta_{n-1}A_0\zeta_{n-1}^{(2)}\| \xrightarrow[n \rightarrow \infty]{p.s.} 0$ . En utilisant (2.30), (2.7) ainsi que  $(H_3)$ , nous avons

$\sigma_n - \sigma_\infty = O(n^{1-\gamma})$ . La normalité de la suite  $(\varepsilon_n)$  conjuguée au Lemme 2.12 implique que  $\|\ell^1(\sigma_n - \sigma_\infty)\varepsilon_n\| \xrightarrow[n \rightarrow \infty]{p.s.} 0$ . D'où, en utilisant la même norme que celle invoquée dans le Lemme 2.13, pour tout  $\eta > 0$ , il existe une variable aléatoire  $n_0$  telle que pour tout  $n \geq n_0$ ,

$$\|\zeta_n^{(1)} - \gamma_n\|_* \leq \|A_0\|_* \|\zeta_{n-1}^{(1)} - \gamma_{n-1}\|_* + \eta \quad p.s.$$

Comme  $\|A_0\|_* < 1$ , nous obtenons  $\|\zeta_n^{(1)} - \gamma_n\| \xrightarrow[n \rightarrow \infty]{p.s.} 0$ .  $\square$

**Lemme 2.15.** Sous  $\mathbb{P}_{\theta_0}^{(n)}$ , nous avons

$$2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_p^2. \quad (2.38)$$

**Preuve.** En utilisant la décomposition obtenue dans la preuve du Théorème 2.5, nous via (2.27) que sous  $\mathbb{P}_{\theta_0}^{(n)}$

$$\begin{aligned} 2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) &= 2 \langle \sqrt{n}(\hat{\theta}_n - \theta_0), \frac{M_n}{\sqrt{n}} \rangle - \langle \sqrt{n}(\hat{\theta}_n - \theta_0), I(\theta_0)\sqrt{n}(\hat{\theta}_n - \theta_0) \rangle \\ &\quad - \langle \sqrt{n}(\hat{\theta}_n - \theta_0), \left( \frac{\langle M \rangle_n}{n} - I(\theta_0) \right) \sqrt{n}(\hat{\theta}_n - \theta_0) \rangle \\ &= \langle \langle M \rangle_n^{-\frac{1}{2}} M_n, \langle M \rangle_n^{-\frac{1}{2}} M_n \rangle. \end{aligned}$$

D'où, (2.38) découle immédiatement de (2.29) et (2.31).  $\square$

Le Lemme suivant peut être vu comme un Lemme de Toeplitz pour les matrices, (voir Théorème 1.1 in [72] pour une référence).

**Lemme 2.16.** *Soient  $(B_{k,n})$  et  $(A_n)$  deux suites de matrices carrées telles que*

- $\|A_n - A\| \xrightarrow{n \rightarrow \infty} 0$ ,
- $\|\sum_{k=1}^n B_{k,n} - B\| \xrightarrow{n \rightarrow \infty} 0$ ,
- $\sum_{k=1}^n \|B_{k,n}\|$  est bornée,
- pour tout  $n_0 > 0$ ,  $\sum_{k=1}^{n_0} \|B_{k,n}\| \xrightarrow{n \rightarrow \infty} 0$ .

Alors,

$$\left\| \sum_{k=1}^n A_k B_{k,n} - AB \right\| \xrightarrow{n \rightarrow \infty} 0. \quad (2.39)$$

**Preuve.** Nous avons

$$\sum_{k=1}^n A_k B_{k,n} - AB = \sum_{k=1}^n A_k B_{k,n} - \sum_{k=1}^n AB_{k,n} + \sum_{k=1}^n AB_{k,n} - AB.$$

Pour tout  $\varepsilon > 0$ , il existe  $n_0 > 0$  tel que pour tout  $n \geq n_0$ ,  $\|A_n - A\| < \varepsilon$ . D'où,

$$\begin{aligned} \left\| \sum_{k=1}^n A_k B_{k,n} - AB \right\| &\leq \left\| \sum_{k=1}^n (A_k - A) B_{k,n} \right\| + \left\| \sum_{k=1}^n A (B_{k,n} - B) \right\| \\ &\leq \sum_{k=1}^{n_0-1} \|A_k - A\| \|B_{k,n}\| + \varepsilon \sum_{k=n_0}^n \|B_{k,n}\| + \|A\| \left\| \sum_{k=1}^n B_{k,n} - B \right\|. \end{aligned}$$

Les hypothèses du Lemme permettent donc de montrer que le terme de droite de la dernière égalité peut être rendu arbitrairement petit dès lors que  $n$  tend vers l'infini.  $\square$

**Remarque 2.6.3** *Le Lemme précédent est encore valide si les hypothèses sont vérifiées au sens de la convergence presque-sûre.*

**Remarque 2.6.4** *Soit deux suites  $(u_n)$  et  $(v_n) \in \mathbb{R}^d$ , nous avons que  $u_n v_n^* = \bar{u}_n \bar{v}_n^*$ , où*

$$\bar{u}_n = \begin{pmatrix} u_{11} & 0 & \dots & \dots & 0 \\ u_{21} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ u_{(d-1)1} & 0 & \dots & \dots & 0 \\ u_{d1} & 0 & \dots & \dots & 0 \end{pmatrix}.$$

Plus précisément,  $\bar{u}_n$  est une matrice de taille  $d \times d$  avec  $u_n$  comme première colonne et des 0 ailleurs. Alors  $\|u_n v_n^*\| \leq \|\bar{u}_n\| \|\bar{v}_n^*\|$ , et, comme nous travaillons en dimension finie,  $\|\bar{u}_n - \bar{u}\| \xrightarrow{n \rightarrow \infty} 0$  si et seulement si chaque composante de  $\bar{u}_n - \bar{u}$  converge vers 0.

**Lemme 2.17.**  $\Theta$  est un sous ensemble ouvert de  $\mathbb{R}^p$ .

**Preuve.** Soit

$$P(z) = 1 - \theta_1 z - \dots - \theta_p z^p$$

le polynôme autorégressif défini sur  $\mathbb{C}$ .  $(H_1)$  est totalement équivalente à la condition suivante (voir [20]) :

$$\text{si } |z| \leq 1 \text{ alors } |P(z)| > 0. \quad (2.40)$$

Soit

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p) \in \mathbb{R}^p \text{ et } Q_\alpha(z) = 1 - (\theta_1 + \alpha_1)z - \dots - (\theta_p + \alpha_p)z^p.$$

Alors pour tout  $|z| \leq 1$ ,

$$\begin{aligned} |Q_\alpha(z)| &= |P(z) - \alpha_1 z - \dots - \alpha_p z^p| \\ &\geq |P(z)| - |\alpha_1 z + \dots + \alpha_p z^p| \\ &\geq |P(z)| - |\alpha_1| - \dots - |\alpha_p|. \end{aligned}$$

La condition (2.40) assure que

$$\inf_{z \in D(0;1)} |P(z)| = \delta_{min} > 0.$$

Il suffit alors de choisir  $|\alpha_1| + \dots + |\alpha_p| < \delta_{min}$  pour obtenir le résultat voulu.  $\square$

**Remarque 2.6.5** *Le Lemme précédent assure que lorsque  $\theta \in \Theta$ , alors pour tout  $u \in \mathbb{R}^p$  et  $n$  assez grand,  $\theta + \frac{u}{\sqrt{n}} \in \Theta$ . Sous les notations de la preuve précédente avec  $\alpha \in \mathbb{R}^p$  tel que  $\|\alpha\| < \delta_{min}$ , nous avons pour toute rotation  $R$  de  $\mathbb{R}^p$  dans  $\mathbb{R}^p$  et  $n$  assez grand, que  $\theta + \frac{Ru}{\sqrt{n}} \in \Theta$ .*

## 2.7 Preuve des Résultats principaux

### 2.7.1 Preuve du Théorème 2.1

**Preuve.** Il est montré dans [21] que

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}_{\theta_0}^{(n)}} I(\theta).$$

Nous souhaitons à présent améliorer ce résultat en obtenant la convergence *p.s.* vers  $I(\theta)$ .

Nous avons

$$\begin{aligned} \frac{\langle M \rangle_n}{n} &= \frac{1}{n} \sum_{i=1}^n \frac{(\zeta_{i-1}^{(1)} + \beta_{i-1} \zeta_{i-1}^{(2)} - \gamma_{i-1} + \gamma_{i-1})(\zeta_{i-1}^{(1)} + \beta_{i-1} \zeta_{i-1}^{(2)} - \gamma_{i-1} + \gamma_{i-1})^*}{\sigma_i^2} \\ &= S_n + r_n, \end{aligned}$$

où

$$S_n = \frac{1}{n} \sum_{k=1}^n \frac{\gamma_{i-1} \gamma_{i-1}^*}{\sigma_i^2}.$$

Le terme de reste  $r_n$  tend *p.s.* vers 0 grâce au Théorème de Cesàro, l'ergodicité de  $(\gamma_n)$ , le Lemme 2.14, la Proposition 2.11, le Lemme 2.16 et la Remarque 2.6.4. Une application directe du Théorème ergodique conjuguée au Lemme 2.16 nous conduit à

$$S_n \xrightarrow[n \rightarrow \infty]{p.s.} \frac{1}{\sigma_\infty^2} \mathbb{E}(\gamma_0 \gamma_0^*),$$

et finalement (2.29),

$$S_n \xrightarrow[n \rightarrow \infty]{p.s.} I(\theta).$$

D'où,

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} I(\theta). \quad (2.41)$$

En utilisant le fait que  $\varepsilon_n$  est indépendante de  $\gamma_{n-1}$  et des arguments similaires à cette preuve, nous avons

$$\frac{M_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} 0, \quad (2.42)$$

qui donne la consistance forte du MLE.

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta. \quad \square$$

Posons  $\bar{\mathcal{F}}_n$  le  $\sigma$ -algèbre  $\bar{\mathcal{F}}_n = \bar{\mathcal{F}}_n(X_0, \dots, X_n, \gamma_0, \dots, \gamma_n)$ , où  $(\gamma_n)$  est le processus défini au sein de la Proposition 2.11.

**Remarque 2.7.1**  $(M_n)$  est également une  $\bar{\mathcal{F}}_n$ -martingale, et son introduction est nécessaire pour la preuve suivante.

## 2.7.2 Preuve du Théorème 2.2

**Preuve.** Afin de prouver le Théorème 2.2, nous allons utiliser la loi forte quadratique pour les martingales (voir Théorème 2.1 dans [27]). Soit  $V_n = \sqrt{n} Id_p$  une suite à croissance régulière au sens de Chaabane et Maouia. Nous devons étudier le comportement asymptotique de la variation totale de la martingale, i.e.

$$\frac{[M]_n}{n} = \frac{1}{n} \sum_{i=1}^n \left( \frac{(\zeta_{i-1}^{(1)} + \beta_{i-1} \zeta_{i-1}^{(2)} - \gamma_{i-1} + \gamma_{i-1}) (\zeta_{i-1}^{(1)} + \beta_{i-1} \zeta_{i-1}^{(2)} - \gamma_{i-1} + \gamma_{i-1})^*}{\sigma_i^2} \right) \varepsilon_i^2. \quad (2.43)$$

En utilisant des arguments similaires à la preuve précédente, nous avons

$$\frac{[M]_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E} \left( \frac{\gamma_1 \gamma_1^*}{\sigma_\infty^2} \right). \quad (2.44)$$

Nous devons ensuite montrer que la condition de Lindeberg , c'est à dire pour tout  $\varepsilon > 0$ ,

$$L_n = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( \|\Delta M_i\|^2 \mathbb{1}_{\{\|\Delta M_i\| \geq \varepsilon \sqrt{n}\}} \middle| \bar{\mathcal{F}}_{i-1} \right) \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (2.45)$$

Soit  $M > 0$  et

$$L_{n,M} = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( \|\Delta M_i\|^2 \mathbb{1}_{\{\|\Delta M_i\| \geq M\}} \middle| \bar{\mathcal{F}}_{i-1} \right).$$

En utilisant (2.27) et la définition de  $a_n$ , nous avons  $\Delta M_1 = M_1$ , et pour tout  $n \geq 2$ ,

$$\Delta M_n = \frac{(\zeta_{n-1}^{(1)} + \beta_{n-1} \zeta_{n-1}^{(2)}) \varepsilon_n}{\sigma_n}.$$

Il vient que

$$\begin{aligned} \|\Delta M_n\|^2 &= \left\| \frac{(\zeta_{n-1}^{(1)} - \gamma_{n-1} + \beta_{n-1} \zeta_{n-1}^{(2)} + \gamma_{n-1}) \varepsilon_n}{\sigma_n} \right\|^2 \\ &\leq \frac{2 \left\| \zeta_{n-1}^{(1)} - \gamma_{n-1} + \beta_{n-1} \zeta_{n-1}^{(2)} \right\|^2 \varepsilon_n^2}{\sigma_n^2} + \frac{2 \|\gamma_{n-1}\|^2 \varepsilon_n^2}{\sigma_n^2}. \end{aligned}$$

Alors,

$$L_{n,M} \leq 2R_{1,n} + 2R_{2,n} + 2R_{3,n},$$

où

$$R_{1,n} = \frac{1}{n} \sum_{i=1}^n \frac{\left\| \zeta_{i-1}^{(1)} - \gamma_{i-1} + \beta_{i-1} \zeta_{i-1}^{(2)} \right\|^2}{\sigma_i^2},$$

$$R_{2,n} = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( \frac{\|\gamma_{i-1}\|^2}{\sigma_i^2} \varepsilon_i^2 \mathbb{1}_{\left\{ \frac{\|\gamma_{i-1}\| \varepsilon_i}{\sigma_i} \geq \frac{M}{2} \right\}} \middle| \bar{\mathcal{F}}_{i-1} \right),$$

et

$$R_{3,n} = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( \frac{\|\gamma_{i-1}\|^2}{\sigma_i^2} \varepsilon_i^2 \mathbb{1}_{\left\{ \frac{\left\| \zeta_{i-1}^{(1)} - \gamma_{i-1} + \beta_{i-1} \zeta_{i-1}^{(2)} \right\| \varepsilon_i}{\sigma_i} \geq \frac{M}{2} \right\}} \middle| \bar{\mathcal{F}}_{i-1} \right).$$

Le même raisonnement que précédemment montre que  $R_{1,n}$  tend vers 0 *p.s.* Intéressons nous aux termes  $R_{2,n}$  et  $R_{3,n}$ . Premièrement, nous savons via (2.30) que pour  $0 < \tilde{m} < \sigma_\infty$ , il existe  $n_0$  tel que pour tout  $n \geq n_0$ ,  $|\sigma_n| \geq \tilde{m}$ . D'où,

$$R_{2,n} \leq R_{2,n}^{(0)} + \frac{1}{n} \sum_{i=n_0}^n \mathbb{E} \left( \frac{\|\gamma_{i-1}\|^2}{\tilde{m}^2} \varepsilon_i^2 \mathbb{1}_{\left\{ \frac{\|\gamma_{i-1}\| |\varepsilon_i|}{\tilde{m}} \geq \frac{M}{2} \right\}} \middle| \bar{\mathcal{F}}_{i-1} \right),$$

d'où  $R_{2,n}^{(0)} \xrightarrow[n \rightarrow \infty]{p.s.} 0$ . Le processus  $(\gamma_{n-1} \varepsilon_n)$  est ergodique car  $(\varepsilon_n)$  est *i.i.d.* il existe  $\phi$  indépendant de  $n$  telle que  $\gamma_{n-1} = \phi(\varepsilon_{n-1}, \varepsilon_{n-2}, \dots)$  (voir Théorème 5.3.8 dans [94]). En conséquence,

$$\lim_{n \rightarrow \infty} R_{2,n} \leq \mathbb{E} \left( \frac{\|\gamma_0\|^2}{\tilde{m}^2} \varepsilon_1^2 \mathbb{1}_{\left\{ \frac{\|\gamma_0\| |\varepsilon_1|}{\tilde{m}} \geq \frac{M}{2} \right\}} \right),$$

et lorsque  $M \rightarrow \infty$ ,  $R_{2,n} \xrightarrow[n, M \rightarrow \infty]{p.s.} 0$ . Nous avons en vertu de l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} R_{3,n}^2 &\leq \frac{1}{n^2} \sum_{i=1}^n \mathbb{E} \left( \frac{\|\gamma_{i-1}\|^4}{\sigma_i^4} \varepsilon_i^4 \middle| \bar{\mathcal{F}}_{i-1} \right) \sum_{i=1}^n \mathbb{E} \left( \mathbb{1}_{\left\{ \frac{\|\zeta_{i-1}^{(1)} - \gamma_{i-1} + \beta_{i-1} \zeta_{i-1}^{(2)}\| |\varepsilon_i|}{|\sigma_i|} \geq \frac{M}{2} \right\}} \middle| \bar{\mathcal{F}}_{i-1} \right) \\ &\leq \frac{3}{n^2} \sum_{i=1}^n \frac{\|\gamma_{i-1}\|^4}{\sigma_i^4} \sum_{i=1}^n \left[ \mathbb{1}_{\left\{ \frac{\|\zeta_{i-1}^{(1)} - \gamma_{i-1} + \beta_{i-1} \zeta_{i-1}^{(2)}\|}{|\sigma_i|} \geq \frac{\sqrt{M}}{\sqrt{2}} \right\}} + \mathbb{P} \left( |\varepsilon_1| \geq \frac{\sqrt{M}}{\sqrt{2}} \right) \right]. \end{aligned}$$

Puisque  $\|\zeta_{i-1}^{(1)} - \gamma_{i-1} + \beta_{i-1} \zeta_{i-1}^{(2)}\| \xrightarrow[n \rightarrow \infty]{p.s.} 0$  via la Proposition 2.11 et le Lemme 2.14, et comme  $(\varepsilon_n)$  est Gaussienne, une nouvelle application du Théorème ergodique nous conduit à  $R_{3,n} \xrightarrow[n \rightarrow \infty]{p.s.} 0$ , dès lors que  $M \rightarrow \infty$ . D'où,

$$L_{n,M} \xrightarrow[n, M \rightarrow \infty]{p.s.} 0 \text{ et donc } L_n \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

En vertu de la loi forte quadratique pour les martingales,

$$\frac{1}{p \log n} \sum_{k=1}^n \left( 1 - \frac{k^p}{(k+1)^p} \right) \frac{M_k M_k^*}{k} \xrightarrow[n \rightarrow \infty]{p.s.} I(\theta). \quad (2.46)$$

Soit  $\lambda^{(k)} = \rho(M_k M_k^*) = \|M_k M_k^*\| \leq \text{Tr}(M_k M_k^*)$ , où la relation précédente est satisfaite du fait que  $M_k M_k^*$  symétrique positive.

Soit  $v_i = (0, \dots, 0, 1, 0, \dots, 0)^*$  un vecteur de taille  $p$  avec 1 sur la  $i$ -ième composante et 0 ailleurs. Alors,

$$\frac{1}{p \log n} \sum_{k=1}^n \left( 1 - \frac{k^p}{(k+1)^p} \right) \frac{\lambda^{(k)}}{k} \leq \frac{1}{p \log n} \sum_{i=1}^p \sum_{k=1}^n \left( 1 - \frac{k^p}{(k+1)^p} \right) \frac{\langle v_i, M_k M_k^* v_i \rangle}{k} \text{ p.s.} \quad (2.47)$$

Puisque le terme de droite de (2.47) est borné, les conditions du Lemme 2.16 sont vérifiées, et nous avons donc

$$\frac{1}{\log n} \sum_{k=1}^n (\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^* = \frac{1}{\log n} \sum_{k=1}^n \langle M \rangle_k^{-1} M_k M_k^* \langle M \rangle_k^{-1}. \quad (2.48)$$

Comme  $\left(1 - \frac{k^p}{(k+1)^p}\right) \sim \frac{p}{k}$ , en utilisant (2.46), (2.48), (2.41) conjuguées au Lemme 2.16, nous obtenons

$$\frac{1}{\log n} \sum_{k=1}^n (\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^* \xrightarrow[n \rightarrow \infty]{p.s.} I(\theta)^{-1}. \quad (2.49) \quad \square$$

### 2.7.3 Preuve de la Proposition 2.3

**Preuve.** Pour obtenir une loi du logarithme itéré, nous allons appliquer le Lemme C.2 de [7]. Nous avons déjà montré que (2.41); en conséquence, il reste à montrer que

$$\sum_{n=1}^{\infty} \left( \frac{\|a_n^* \zeta_n\|}{|\sigma_{n+1}| \sqrt{n}} \right)^\beta < \infty \text{ p.s.}$$

pour  $\beta > 2$ . En vertu de la Proposition 2.11 et du Lemme 2.14, il existe une suite  $(\tau_n)$  telle que

$$\|a_n^* \zeta_n\| = \|\gamma_n + \tau_n\| \text{ et } \tau_n \rightarrow 0 \text{ p.s.}$$

Par conséquent, pour  $n$  assez grand,

$$\left( \frac{\|a_n^* \zeta_n\|}{|\sigma_{n+1}| \sqrt{n}} \right)^\beta \leq \frac{\|\tau_n + \gamma_n\|^\beta}{\widetilde{m}^{\frac{\beta}{2}} n^{\frac{\beta}{2}}} \leq \frac{(1 + \|\gamma_n\|)^\beta}{\widetilde{m}^{\frac{\beta}{2}} n^{\frac{\beta}{2}}} \text{ p.s.}$$

où  $\widetilde{m}$  est la même constante que celle de la preuve précédente. Comme  $(\gamma_n)$  est un processus Gaussien ergodique,  $\|\gamma_n\| = O(n^\alpha)$  p.s. pour tout  $0 < \alpha < \frac{1}{2}$ . La dernière inégalité implique que pour  $n$  assez grand,

$$\left( \frac{\|a_n^* \zeta_n\|}{|\sigma_{n+1}| \sqrt{n}} \right)^\beta \leq \frac{C}{\widetilde{m}^{\frac{\beta}{2}} n^{\beta \frac{1-2\alpha}{2}}} \text{ p.s.}$$

pour une constante  $C > 0$ . Le résultat voulu est obtenu dès lors que  $\beta > \frac{2}{1-2\alpha} > 2$ . D'où,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{\frac{1}{2}} v^* \langle M \rangle_n^{-1} M_n &= - \liminf_{n \rightarrow \infty} \left( \frac{n}{2 \log \log n} \right)^{\frac{1}{2}} v^* \langle M \rangle_n^{-1} M_n \\ &= (v^* I(\theta)^{-1} v)^{\frac{1}{2}} \text{ p.s.} \end{aligned}$$

Soit  $\widehat{\theta}_n^i$  la  $i$ -ième composante de  $\widehat{\theta}_n$ , et posons  $v_i = (0, \dots, 0, 1, 0, \dots, 0)^*$  un vecteur de taille  $p$  avec 1 à la  $i$ -ième composante et 0 ailleurs. La dernière inégalité implique immédiatement (2.20).  $\square$

## 2.7.4 Preuve du Théorème 2.5

**Preuve.** Soit

$$\phi_n(\theta_0) = \frac{Id_p}{\sqrt{n}}, u = (u_1, u_2, \dots, u_p)^*,$$

$$U_n = \begin{pmatrix} u_1 & u_2 & \dots & u_p & \beta_n u_1 & \beta_n u_2 & \dots & \beta_n u_p \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \text{ and } \ell^* U_n = u^* a_n^*.$$

Alors,

$$\begin{aligned} \log \left( \frac{\mathcal{L}(\theta_0 + \phi_n(\theta_0)u, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) &= -\frac{1}{2} \sum_{k=1}^n \frac{\ell^*(2\zeta_k - 2\tilde{A}_{k-1}\zeta_{k-1} - \frac{U_{k-1}}{\sqrt{n}}\zeta_{k-1})\ell^*(-\frac{U_{k-1}}{\sqrt{n}}\zeta_{k-1})}{\sigma_k^2} \\ &= -\frac{1}{2} \sum_{k=1}^n \frac{\ell^*(2\ell\sigma_k\varepsilon_k - \frac{U_{k-1}}{\sqrt{n}}\zeta_{k-1})\ell^*(-\frac{U_{k-1}}{\sqrt{n}}\zeta_{k-1})}{\sigma_k^2} \\ &= \sum_{k=1}^n \left( \frac{u^* a_{k-1}^* \zeta_{k-1} \varepsilon_k}{\sqrt{n}\sigma_k} - \frac{1}{2} \frac{u^* a_{k-1}^* \zeta_{k-1} \zeta_{k-1}^* a_{k-1} u}{n\sigma_k^2} \right) \\ &= \left\langle u, \frac{M_n}{\sqrt{n}} \right\rangle - \frac{1}{2} \left\langle u, \frac{\langle M \rangle_n}{n} u \right\rangle \\ &= \left\langle u, \frac{M_n}{\sqrt{n}} \right\rangle - \frac{1}{2} \left\langle u, I(\theta_0)u \right\rangle - \frac{1}{2} \left\langle u, \left( \frac{\langle M \rangle_n}{n} - I(\theta_0) \right) u \right\rangle. \end{aligned}$$

Posons  $R_n(\theta_0, n) = \frac{1}{2} \left\langle u, (I(\theta_n) - \frac{\langle M \rangle_n}{n})u \right\rangle$ . Alors  $|R_n(\theta_0, n)| \leq \frac{1}{2} \|I(\theta_0) - \frac{\langle M \rangle_n}{n}\| R^2$  et la propriété LAN en  $\theta_0$  découle immédiatement de (2.31) et (2.29).  $\square$

**Remarque 2.7.2** Dans [60], la propriété LAN est définie pour des expériences statistiques admettant la décomposition (1.9) avec  $J(\theta_0) = Id_d$ . Cette définition n'est pas restrictive car (1.9) peut être reformulée comme dans la Définition 2.1 du Chapitre 2 dans [60] en posant  $v = J(\theta_0)^{-\frac{1}{2}}u$  où bien, de manière équivalente, en considérant une nouvelle suite de matrices  $\tilde{\phi}_n(\theta_0) = \phi_n(\theta_0)J(\theta_0)^{-\frac{1}{2}}$ .

## 2.7.5 Preuve de la Proposition 2.6

**Preuve.** La Proposition découle immédiatement de (2.17). Plus précisément, la condition suivante formulée au sein du Lemme 13.1, Chapitre 2 de [60], afin d'obtenir l'efficacité asymptotique est vérifiée

$$I(\theta_0)^{\frac{1}{2}} \sqrt{n}(\hat{\theta}_n - \theta_0) - I(\theta_0)^{-\frac{1}{2}} \frac{M_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathbb{P}_{\theta_0}^{(n)}} 0 \quad (2.50)$$

$\square$

**Définition 2.18.** *Un test  $\phi_n^1$  est asymptotiquement équivalent à  $\tilde{\phi}_n^1$  si*

$$\phi_n^1 - \tilde{\phi}_n^1 \xrightarrow[n \rightarrow \infty]{\mathbb{P}_{\theta_0}^{(n)}} 0. \quad (2.51)$$

Nous formulons un Théorème caractérisant les tests AUMPI( $\alpha$ ) démontré dans Théorème 3 de [28].

**Théorème 2.19.** *Sous les notations de la Définition 1.9, soit  $C_{1-\alpha}$  le quantile d'ordre  $1 - \alpha$  d'un  $\chi_d^2$ . Alors, le test*

$$\phi_n = \mathbb{1}_{\{(J(\theta_0)^{-1}Z_n(\theta_0), Z_n(\theta_0)) \geq C_{1-\alpha}\}} \quad (2.52)$$

*pour tester  $\theta = \theta_0$  contre  $\theta \neq \theta_0$  est AUMPI( $\alpha$ ), et donc tout autre test asymptotiquement équivalent possédera cette propriété.*

## 2.7.6 Preuve du Théorème 2.10

**Preuve.** Pour démontrer ce Théorème, nous devons montrer que le test  $\tilde{\phi}_n$  est asymptotiquement équivalent au test formulé dans le Théorème 2.19 sous  $\mathbb{P}_{\theta_0}^{(n)}$  (l'hypothèse nulle). Nous avons

$$|\tilde{\phi}_n - \phi_n| = \tilde{\phi}_n + \phi_n - 2\tilde{\phi}_n\phi_n. \quad (2.53)$$

Posons  $\delta_n = \langle I(\theta_0)^{-1} \frac{M_n}{\sqrt{n}}, \frac{M_n}{\sqrt{n}} \rangle - 2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right)$ , nous avons pour tout  $\varepsilon > 0$ ,

$$|\tilde{\phi}_n - \phi_n| \leq \tilde{\phi}_n + \phi_n - 2\mathbb{1}_{\left\{2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) \geq C_{1-\alpha} + \varepsilon\right\}} \mathbb{1}_{\{\delta_n \geq -\varepsilon\}}. \quad (2.54)$$

Il est aisé de voir que  $\delta_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} 0$  sous  $\mathbb{P}_{\theta_0}^{(n)}$ . Le Théorème de Slutsky ainsi que le Lemme 2.15 et le continuous mapping Théorème (voir Théorème 2.1 dans [15]) impliquent que

$$2\mathbb{1}_{\left\{2 \log \left( \frac{\mathcal{L}(\hat{\theta}_n, X^{(n)})}{\mathcal{L}(\theta_0, X^{(n)})} \right) \geq C_{1-\alpha} + \varepsilon\right\}} \mathbb{1}_{\{\delta_n \geq -\varepsilon\}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} 2\mathbb{1}_{\{Z \geq C_{1-\alpha} + \varepsilon\}} \quad (2.55)$$

sous  $\mathbb{P}_{\theta_0}^{(n)}$ , où  $Z \sim \chi_p^2$ . Puisque le terme de droite de (2.54) est uniformément borné par rapport à  $n$ , nous avons

$$\limsup_{n \rightarrow \infty} \mathbb{E}_{\theta_0}^{(n)}(|\tilde{\phi}_n - \phi_n|) \leq 2\mathbb{P}(Z \geq C_{1-\alpha}) - 2\mathbb{P}(Z \geq C_{1-\alpha} + \varepsilon), \quad (2.56)$$

et les inégalités précédentes nous conduisent à

$$\mathbb{E}_{\theta_0}^{(n)}(|\tilde{\phi}_n - \phi_n|) \xrightarrow[n \rightarrow \infty]{} 0. \quad (2.57)$$

La dernière condition nous assure donc que les tests  $\tilde{\phi}_n$  et  $\phi_n$  sont asymptotiquement équivalents.  $\square$

# Chapitre 3

## Détection d'une rupture au sein d'un processus autorégressifs à bruits gaussiens stationnaires

### 3.1 Historique et motivation

Les résultats présentés dans ce chapitre sont résumés dans [22] et sont le fruit d'un travail mené avec mon directeur de thèse et d'une collaboration avec Chunhao Cai et Longmin Wang que je tiens à remercier.

Nous nous intéressons dans ce chapitre à la détection d'une rupture du paramètre autorégressif au sein d'une série chronologique générée par le modèle présenté dans le chapitre précédent. Nous rappelons que ce problème fût considéré dans un processus  $AR(p)$  classique dans [35] où il suppose que le processus de nuisance est un bruit blanc fort vérifiant une condition d'intégrabilité à un certain ordre. Le cas d'une rupture au sein d'un processus  $AR(1)$  fût particulièrement étudié dans les contributions suivantes [30, 36, 37].

Dans divers champs d'applications des statistiques, les scientifiques ont observé un phénomène de longue mémoire où la corrélation entre les observations décroît plus lentement que lorsque les données sont indépendantes où générées par un processus  $AR(p)$  classique. Ce phénomène apparaît dans des données chronologiques de nature macro-économique mais également en assurance (voir [2] pour une référence complète). Plus particulièrement, le processus  $AR(1)$  généré par des bruits dépendants (bruit Gaussien fractionnaire) est apparu récemment dans une large littérature au sein des mathématiques financières (voir par exemple [45]).

Le problème de la détection d'une rupture au sein de données générées par un processus  $AR(1)$  dirigé par un bruit gaussien stationnaire est ainsi considéré dans ce chapitre. Soit  $(X_1, X_2, \dots, X_N)$   $N$  observations successives vérifiant les équations récursives suivantes

$$X_n = \theta_1 X_{n-1} + \xi_n, \quad 1 \leq n \leq \tau, \quad X_0 = 0, \quad (3.1)$$

$$= \theta_2 X_{n-1} + \xi_n, \quad \tau + 1 \leq n \leq N, \quad (3.2)$$

où  $1 \leq \tau \leq N$  est la position inconnue de la rupture,  $(\theta_1, \theta_2) \in \Theta^2 \subset (-1, 1)^2$  sont les paramètres autogrèssifs inconnus et le processus de nuisance  $\xi = (\xi_n)_{n \geq 1}$  est gaussien stationnaire, centré, régulier, *i.e.*

$$\int_{-\pi}^{\pi} |\ln f_{\xi}(\lambda)| d\lambda < \infty, \quad (3.3)$$

où comme dans le chapitre précédent  $f_{\xi}(\lambda)$  est la densité spectrale de  $\xi$ . Nous supposons que la fonction de covariance  $c = (c(m, n), m, n \geq 1)$  définie par

$$\mathbb{E}(\xi_m \xi_n) = c(m, n) = \rho(|n - m|), \quad \rho(0) = 1, \quad (3.4)$$

est définie positive.

L'objectif de ce chapitre est d'étudier le comportement asymptotique du test par ratio de vraisemblance sous l'hypothèse nulle  $H_0 : \tau = N$  (aucune rupture ne s'est produite) contre  $H_1 : \tau < N$  (une rupture a eu lieu).

Dans le cadre d'un processus autorégressif classique, le comportement asymptotique du test est obtenu via l'application d'un principe d'invariance fort pour les processus vérifiant une condition de mélange (voir [65]) conjugué à l'application du Théorème de Darling-Erdős étendu par Horvath dans [56]. Nous ne pouvons dans notre cas utiliser la même méthode dans la mesure où le processus obtenu après le filtrage n'est pas stationnaire. La technique de preuve repose sur l'utilisation d'un principe d'invariance pour les processus ne vérifiant pas nécessairement une condition de mélange (voir [41]), l'obtention d'estimations presque sûre sur les convergences établies dans le chapitre précédent.

Nous rappelons brièvement avant d'énoncer les résultats obtenus la transformation linéaire effectuée afin d'obtenir un modèle Markovien. Le résultat principal concernant la convergence en loi du test par ratio de vraisemblance vers une loi de Gumbel est énoncé dans la Section 3.3. Des simulations sont effectuées au sein de la Section 3.4. Les preuves sont présentées dans la Section 3.7.

## 3.2 Notations et rappels

Nous rappelons que nous pouvons associer au processus  $\xi = (\xi_n)_{n \geq 1}$  une suite d'innovations  $(\sigma_n \varepsilon_n)_{n \geq 1}$  définie par les relations :

$$\sigma_1 \varepsilon_1 = \xi_1, \quad \sigma_n \varepsilon_n = \xi_n - \mathbb{E}(\xi_n | \xi_1, \xi_2, \dots, \xi_{n-1}), \quad n \geq 2,$$

où  $(\varepsilon_n, n \geq 1)$  est une suite *i.i.d.* distribuée selon une loi normale standard. Nous avons donc via le Théorème de la corrélation normale [73, Théorème 13.1] qu'il existe un noyau déterministe  $k = (k(n, m), n \geq 1, m \leq n)$  tel que  $k(n, n) = 1$  et

$$\sigma_n \varepsilon_n = \sum_{m=1}^n k(n, m) \xi_m. \quad (3.5)$$

Nous rappelons que pour tout  $n \geq 1$ , nous notons  $\beta_{n-1}$  le coefficient d'autocorrélation partielle

$$\beta_{n-1} = -k(n, 1). \quad (3.6)$$

Les relations suivantes entre la fonction de covariance  $\rho(\cdot)$  définie dans (3.4), les coefficients d'autocorrélation partielle  $(\beta_n)_{n \geq 1}$  et la variance des innovations  $(\sigma_n^2)_{n \geq 1}$  (voir l'algorithme de Durbin-Levinson dans [40]) sont valides :

$$\sigma_n^2 = \prod_{m=1}^{n-1} (1 - \beta_m^2), \quad n \geq 2, \quad \sigma_1 = 1, \quad (3.7)$$

$$\sum_{m=1}^n k(n, m) \rho(m) = \beta_n \sigma_n^2, \quad (3.8)$$

$$k(n+1, n+1-m) = k(n, n-m) - \beta_n k(n, m). \quad (3.9)$$

Puisque la covariance  $c(\cdot, \cdot)$  du processus de nuisance est définie positive, il existe également un noyau déterministe  $K = (K(n, m), n \geq 1, m \leq n)$  tel que

$$\xi_n = \sum_{m=1}^n K(n, m) \sigma_m \varepsilon_m. \quad (3.10)$$

La relation entre les différents noyaux  $k$  et  $K$  est donnée dans [21]. Introduisons à présent le processus  $Z = (Z_n)_{n \geq 1}$  défini comme dans le chapitre précédent i.e.

$$Z_n = \sum_{m=1}^n k(n, m) X_m, \quad n \geq 1, \quad (3.11)$$

d'où  $X_n = \sum_{m=1}^n K(n, m) Z_m$ , et la filtration naturelle induite par  $Z$  et celle induite par  $X = (X_n)_{n \geq 1}$  coïncident. Nous rappelons qu'il est montré dans [24] que  $Z$  correspond à la première composante d'un processus de Markov bidimensionnel  $\zeta = (\zeta_n, n \geq 1)$  défini par

$$\zeta_n = \begin{pmatrix} Z_n \\ \sum_{r=1}^{n-1} \beta_r Z_r \end{pmatrix}$$

satisfaisant les équations récursives suivantes :

$$\zeta_n = A_{n-1}^{\theta_1} \zeta_{n-1} + \sigma_n b \varepsilon_n, \quad 0 \leq n \leq \tau, \quad \zeta_0 = \mathbf{0}_{2 \times 1}, \quad (3.12)$$

$$= A_{n-1}^{\theta_2} \zeta_{n-1} + \sigma_n b \varepsilon_n, \quad \tau + 1 \leq n \leq N, \quad (3.13)$$

où

$$A_n^\theta = \begin{pmatrix} \vartheta & \vartheta \beta_n \\ \beta_n & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.14)$$

et  $(\varepsilon_n, n \geq 1)$  est une suite *i.i.d.* distribuée selon une loi normale standard. La méthode pour détecter une rupture à partir des observations  $(\zeta_1, \zeta_2, \dots, \zeta_N)$  est présentée dans la Section suivante.

### 3.3 Résultat principal

Dans cette section, nous construisons le test par ratio de vraisemblance afin de déterminer si une rupture du paramètre autorégressif a eu lieu puis nous donnons son comportement asymptotique sous  $H_0$ . Nous rappelons que ce test sert formellement à discriminer les hypothèses suivantes  $H_0 : \tau = N$  (aucune rupture) contre  $H_1 : \tau = k$  (une rupture a eu lieu à l'instant  $k$ ). Posons

$$L(N, \vartheta_1) = \left( \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma_n^2}} \right) \exp \left( -\frac{1}{2} \sum_{n=1}^N \left( \frac{b^*(\zeta_n - A_{n-1}^{\vartheta_1} \zeta_{n-1})}{\sigma_n} \right)^2 \right)$$

et

$$L(k, \theta_1, \theta_2) = \left( \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma_n^2}} \right) \exp \left( -\frac{1}{2} \sum_{n=1}^k \left( \frac{b^*(\zeta_n - A_{n-1}^{\theta_1} \zeta_{n-1})}{\sigma_n} \right)^2 - \frac{1}{2} \sum_{n=k+1}^N \left( \frac{b^*(\zeta_n - A_{n-1}^{\theta_2} \zeta_{n-1})}{\sigma_n} \right)^2 \right).$$

Nous rappelons que la notation  $*$  désigne la transposition. Le ratio de vraisemblance renormalisé à l'instant  $k$  est alors donné par

$$\Lambda_N(k) = \min_{\theta_1 \in \Theta} (-2 \ln L(N, \theta_1)) - \min_{(\theta_1, \theta_2) \in \Theta^2} (-2 \ln L(k, \theta_1, \theta_2)). \quad (3.15)$$

Nous déduisons des calculs précédents une statistique de test définie comme suit

$$\Lambda_N = \max_{1 < k \leq N} \Lambda_N(k). \quad (3.16)$$

Le Théorème suivant explicite le comportement asymptotique de notre statistique de test  $\Lambda_N$  sous l'hypothèse nulle.

**Théorème 3.1.** *Supposons que  $\beta_n = O(n^{-\rho})$  avec  $\rho > \frac{1}{2}$ . Alors, sous  $H_0$ , quand  $N \rightarrow \infty$ ,*

$$\mathbb{P} \left( \frac{\Lambda_N - \beta(N)}{\alpha(N)} \leq x \right) \longrightarrow \exp(-2e^{-x/2}), \quad (3.17)$$

où  $\beta(N) = (2 \ln \ln N + \frac{1}{2} \ln \ln \ln N - \ln \Gamma(\frac{1}{2}))^2 / (2 \ln \ln N)$  and  $\alpha(N) = \sqrt{\beta(N) / (2 \ln \ln N)}$  sont les constantes de normalisation et  $\Gamma(\cdot)$  est la fonction Gamma standard.

**Remarque 3.3.1** *Nous rappelons que la condition sur la densité spectrale (3.3) implique que*

$$\sum_{n \geq 1} \beta_n^2 < \infty. \quad (3.18)$$

*Nous supposons de plus  $\beta_n = O(n^{-\rho})$  où  $\rho > 1/2$  qui est vérifiée théoriquement pour les processus Gaussiens ARMA stables. Pour le bruit Gaussien fractionnaire (voir équation (5.1) pour une forme explicite de la fonction de covariance) la forme explicite de la fonction d'autocorrélation partielle n'est pas connue et nous avons vérifiée numériquement cette hypothèse.*

**Remarque 3.3.2** Dans le cadre d'un modèle  $AR(p)$  standard, l'estimateur du maximum de vraisemblance et celui des moindres carrés sont tout deux consistants envers le paramètre autorégressif. Pour un processus Gaussien longue mémoire (un bruit Gaussien fractionnaire par exemple avec  $H > 1/2$ ), le LSE n'est plus consistant envers le paramètre autorégressif et ne peut en conséquence être utilisé pour l'estimer.

**Remarque 3.3.3** La preuve du Théorème 3.1 peut être étendue dans le cas  $AR(p)$ . Les estimations fournies dans les Lemmes 3.7, 3.8 et 3.9 devront être formulées de manière vectorielle.

## 3.4 Simulations

Dans cette section, nous présentons des simulations Monte-Carlo afin d'observer le comportement du ratio de vraisemblance sous l'hypothèse nulle pour des échantillons de différentes taille. Nous réalisons ces simulations pour les processus de nuisance suivant :

- processus  $AR(1)$  Gaussien

$$\xi_{n+1} = \sqrt{1 - \alpha^2} \epsilon_{n+1} + \alpha \xi_n, \quad n \geq 1,$$

où  $(\epsilon_n, n \geq 1)$  est bien entendu une suite *i.i.d.* distribuée selon une loi normale standard. Le fonction de covariance est donc donnée par  $\rho(|n - m|) = \alpha^{|n-m|}$ . De plus,  $K(n, m) = \alpha^{n-m}$  et  $k(n, m) = \delta_{n-m} + \alpha \delta_{n-m-1}$ . Le processus de nuisance en question est stationnaire dès lors que  $|\alpha| < 1$ . La condition  $\beta_n = -k(n, 1) = O(n^{-\rho})$ ,  $\rho > 1/2$ , est dans cas de loin satisfaite.

- Un bruit Gaussien fractionnaire (fGn) où la fonction de covariance est donnée par

$$\rho(|n - m|) = \frac{1}{2} (|m - n + 1|^{2H} - 2|n - m|^{2H} + |m - n - 1|^{2H}) \quad (3.19)$$

pour un exposant de Hurst connu  $H \in (0, 1)$ . Le processus de nuisance est stationnaire pour tout  $H \in (0, 1)$ . Le cas où  $H = 0.5$  correspond bien entendu au cas *i.i.d.* Comme cela est décrit dans la Remarque 3.3.1, la suite  $\beta_n$  n'a pas de forme explicite mais la vitesse de décroissance est vérifiée numériquement.

Dans la Figure 3.1, la fonction de répartition empirique de la statistique  $(\Lambda_N - \beta(N))/\alpha(N)$  est représentée graphiquement et nous superposons également la fonction de répartition de la loi limite, telle que décrite au sein du Théorème 3.1. Nous effectuons ces simulations pour deux différentes structures de corrélations ( $AR(1)$  et fGn). Ces figures mettent en exergue la convergence en loi lente de notre statistique vers la distribution limite (ligne gris uni) à mesure que la taille l'échantillon  $N$  augmente. Ces figures peuvent être comparées à celles du cas classique figurant dans [35, Figure 4 page 295] qui correspondent aux cas  $\alpha = 0$  and  $H = 0.5$ .

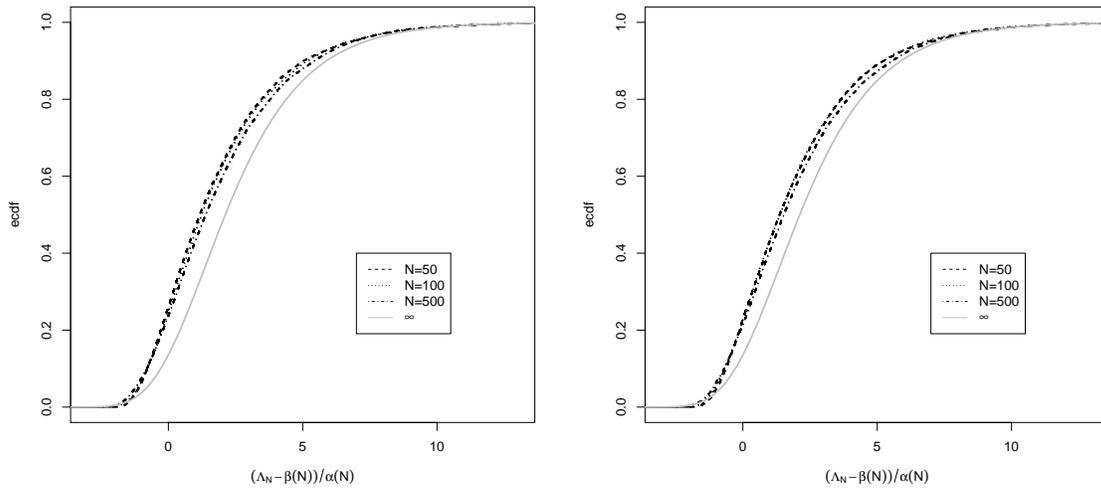


FIGURE 3.1 – Fonction de répartition empirique de notre statistique de test pour un bruit AR(1) gaussien  $\alpha = 0.5$  (à gauche) et un fGn où  $H = 0.7$  (à droite). Ici, le nombre de simulations Monte-Carlo est de  $M = 10000$  et  $\theta_1 = \theta_2 = 0.5$ .

Dans la Figure 3.2, le fonction de répartition empirique du test  $(\Lambda_N - \beta(N))/\alpha(N)$  est représentée graphiquement puis comparée à celle de la loi limite donnée dans le Théorème 3.1 pour le fGn comme processus de nuisance avec trois paramètres autorégressifs différents  $\theta_1 \in \{0, 0.5, 0.9\}$  et trois paramètres de Hurst différents  $H \in \{0.3, 0.5, 0.7\}$  pour un échantillon de taille  $N = 100$ . Comme dans le cas classique, plus  $\theta$  se rapproche de zero et plus la convergence de la statistique vers la loi limite est lente. De plus, dans notre cas, nous constatons que la paramètre de Hurst qui gouverne la mémoire de notre processus de nuisance n’affecte pas le comportement de notre statistique et ce grâce au filtrage.

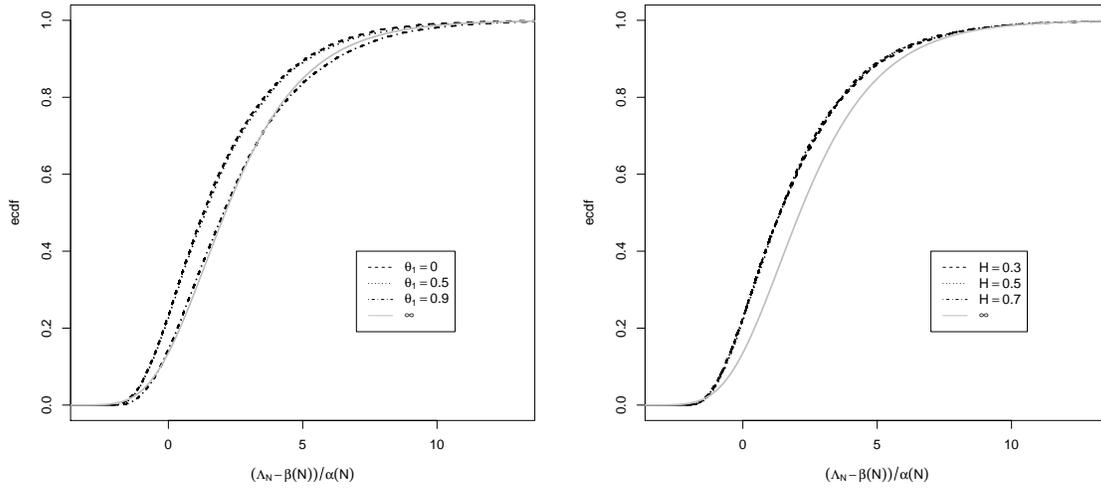


FIGURE 3.2 – Fonction de répartition empirique de notre statistique pour le fGn avec différents paramètres autorégressifs  $\theta_1 \in \{0, 0.5, 0.9\}$  pour  $H = 0.7$  (à gauche) et différent exposant de Hurst  $H \in \{0.3, 0.5, 0.7\}$  pour  $\theta_1 = 0.5$  (à droite). Ici  $N = 100$  et le nombre de simulations Monte Carlo est de  $M = 10000$ .

Dans la Table 3.1, la puissance empirique du test est calculée via la moyenne empirique de nombre de rejets lors des simulations Monte-Carlo. L'erreur de première espèce est fixée à  $\alpha = 0.05$  et la borne supérieur de notre intervalle de confiance pour  $(\Lambda_N - \beta(N))/\alpha(N)$  est fixée à  $x = \log 4 - 2 \log \log \frac{1}{1-\alpha}$  et ce grâce au Théorème 3.1. Nous effectuons  $M = 10000$  simulations Monte-Carlo avec le fGn comme processus de nuisance (3.19) où  $H = 0.7$ . Nous rappelons que grâce au filtrage les performances du test ne seront pas impactées par  $H$ . Comme cela est illustré dans la Table 3.1, la puissance empirique du test augmente à mesure que  $N$  augmente, la différence entre  $|\theta_2 - \theta_1|$  augmente.

$\theta_2 - \theta_1$	0.1	0.3	-0.3	-0.1	0.1	0.3	0.4	-0.3	-0.1
$1/4 \cdot N$	0.04	0.07	0.08	0.04	0.04	0.08	0.13	0.16	0.07
$1/2 \cdot N$	0.04	0.10	0.11	0.05	0.04	0.12	0.22	0.23	0.08
$3/4 \cdot N$	0.03	0.08	0.09	0.04	0.04	0.12	0.24	0.20	0.08
$\theta_2 - \theta_1$	0.1	0.3	-0.3	-0.1	0.1	0.3	0.4	-0.3	-0.1
$1/4 \cdot N$	0.04	0.12	0.16	0.05	0.04	0.14	0.28	0.34	0.10
$1/2 \cdot N$	0.05	0.18	0.20	0.06	0.05	0.24	0.52	0.41	0.11
$3/4 \cdot N$	0.04	0.14	0.14	0.05	0.05	0.22	0.48	0.29	0.10
$\theta_2 - \theta_1$	0.1	0.3	-0.3	-0.1	0.1	0.3	0.4	-0.3	-0.1
$1/4 \cdot N$	0.09	0.63	0.67	0.10	0.10	0.82	0.99	0.96	0.31
$1/2 \cdot N$	0.10	0.80	0.81	0.14	0.13	0.95	1.00	0.99	0.37
$3/4 \cdot N$	0.09	0.66	0.65	0.10	0.10	0.88	1.00	0.94	0.26

TABLE 3.1 – Puissance empirique du test par rapport à la localisation de la rupture (premier quart, moitié ou troisième quart) et la différence (en valeur absolue) entre les paramètres autorégressifs avant et après rupture  $\theta_2 - \theta_1$ . La taille de l'échantillon est fixée à  $N = 50$  (en haut),  $N = 100$  (au milieu) et  $N = 500$  (en bas). Le paramètre autorégressif est fixé  $\theta_1 = 0.2$  (à gauche),  $\theta_1 = 0.5$  (au milieu) et  $\theta_1 = 0.9$  (à droite).

### 3.5 Conclusion et Perspectives

Il est démontré dans ce chapitre que, sous les hypothèses formulées au sein du Chapitre 1, que la statistique de test obtenue par ratio de vraisemblance et convenablement renormalisée se comporte asymptotiquement comme celle présentée dans [35] sous  $H_0$ . Il serait alors intéressant dans un premier temps de généraliser les résultats dans le cas d'un modèle  $AR(p)$ . Il serait également intéressant dans un second temps de construire une procédure permettant de détecter une rupture lorsque la fonction de covariance du bruit n'est pas connue.

Dans tout ce qui suit, nous formulons les preuves des Proposition 5.1 et Proposition 5.2. Les Lemmes techniques permettant d'aboutir à ces preuves sont donnés dans les Résultats Auxiliaires.

### 3.6 Résultats Auxiliaires

Les symboles  $o(\cdot)$  et  $O(\cdot)$  vis à vis des variables aléatoires sont utilisés de manière répétitives dans un sens très similaire à ceux utilisés dans la cas déterministe : lorsque  $n \rightarrow +\infty$ , et pour une suite positive  $(v_n)$ ,  $X_n = o(v_n)$  *p.s.* signifie que  $X_n/v_n$  converge presque sûrement vers 0 tandis que  $X_n = O(v_n)$  *p.s.* signifie, au sens de la définition de [39], que pour presque tout  $\omega$ ,  $X_n(\omega) = O(v_n)$ , i.e.  $|X_n(\omega)| \leq C(\omega) v_n$  pour une constante finie  $C(\omega) \geq 0$  et  $n \geq N(\omega)$ .

Comme mentionné dans la Remarque 2.6.1, le processus  $\zeta = (\zeta_n)_{n \geq 1}$  n'est pas stationnaire et ne vérifie pas en conséquence la condition de mélange qui est le point clé dans [35]. Pour démontrer le Théorème 3.1, nous devons ainsi dans un premier temps établir les trois Propositions suivantes :

**Proposition 3.2.** Soit  $W(t)$  un mouvement Brownien vérifiant  $\mathbf{E}(W^2(t)) = \mathcal{I}(\vartheta)t$ . Supposons que la fonction d'autocorrélation partielle vérifie  $\beta_n = O(n^{-\alpha})$ , alors

$$\sum_{k \leq t} F_k - W(t) \ll t^{1/2-\kappa}, \quad t \rightarrow \infty, \quad p.s. \quad (3.20)$$

pour une constante  $\kappa > 0$ .

**Proposition 3.3.** Considérons le processus  $Q_N(t) = \Lambda_N([Nt])$ , où  $\Lambda_N(k)$  est défini dans (3.47), nous avons pour tout  $0 < t_1 < t_2 < 1$ ,

$$Q_N(t) \xrightarrow{\mathcal{L}} \frac{W(t)^2}{t} + \frac{(W(1) - W(t))^2}{1-t} - W(1)^2 = \frac{(W(t) - tW(1))^2}{t(1-t)}, \quad (3.21)$$

où  $W(t)$  est le mouvement Brownien standard.

**Proposition 3.4.** Pour tout  $\delta > 0$ , nous avons les convergences suivantes :

$$\limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left[ \max_{1 < k \leq \varepsilon N} \Lambda_N(k) - \max_{1 < k \leq \varepsilon N} \frac{M_k^2}{\langle M \rangle_k} \right] / \alpha(N) > \delta \right] \rightarrow 0, \quad \varepsilon \rightarrow 0, \quad (3.22)$$

$$\limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left[ \max_{(1-\varepsilon)N < k \leq N} \Lambda_N(k) - \max_{(1-\varepsilon)N < k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} \right] / \alpha(N) > \delta \right] \rightarrow 0, \quad \varepsilon \rightarrow 0, \quad (3.23)$$

et pour tout  $\varepsilon > 0$

$$\mathbb{P} \left[ \left( \max_{1 < k \leq \varepsilon N} \frac{M_k^2}{\langle M \rangle_k} - \beta(N) \right) / \alpha(N) \leq x \right] \rightarrow \exp(-e^{-x/2}), \quad N \rightarrow \infty, \quad (3.24)$$

$$\mathbb{P} \left[ \left( \max_{(1-\varepsilon)N < k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} - \beta(N) \right) / \alpha(N) \leq x \right] \rightarrow \exp(-e^{-x/2}), \quad N \rightarrow \infty, \quad (3.25)$$

où  $\beta(N) = (2 \ln \ln N + \frac{1}{2} \ln \ln \ln N - \ln \Gamma(\frac{1}{2}))^2 / (2 \ln \ln N)$  et  $\alpha(N) = \sqrt{\beta(N) / (2 \ln \ln N)}$ .

Soit  $\gamma = (\gamma_n)_{n \geq 1}$  un processus autorégressif standard d'ordre 1 :

$$\gamma_n = \theta \gamma_{n-1} + \varepsilon_n \quad (3.26)$$

où  $\vartheta$  et le bruit  $\varepsilon_n$  sont les mêmes que dans (3.46). Posons maintenant

$$F_k = \frac{a_{k-1}^* \zeta_{k-1}}{\sigma_k} \varepsilon_k$$

et  $M_n(m)$  donné par

$$M_n(m) = \sum_{k=m+1}^{m+n} F_k, \quad M_n = M_n(0). \quad (3.27)$$

### 3.6.1 Lemmes techniques pour établir la Proposition 3.2

Comme précédemment, nous supposons que  $\beta_n = O(n^{-\alpha})$  avec  $\alpha > 1/2$ . A partir de maintenant, nous supposons que  $1/2 < \alpha < 1$ . Cela n'est pas restrictif car si  $\alpha \geq 1$  nous aurons également  $\beta_n = O(n^{-\lambda_1})$  pour  $1/2 < \lambda_1 < 1$ . Les estimation suivantes seront donc valide en remplaçant  $\alpha$  par  $\lambda_1$ .

Dans notre modèle, le processus obtenu après la filtrage ne vérifie pas de condition de mélange afin d'appliquer le principe d'invariance fort de [65]. Nous utilisons en conséquence la martingale  $(M_k, k \geq 1)$  définie dans (3.27) afin d'appliquer un principe d'invariance fort décrit dans [41]. Nous noterons par la suite  $(\langle M \rangle_k, k \geq 1)$  la variation quadratique prévisible de  $(M_k, k \geq 1)$ .

Posons

$$\Phi_n = (\Phi_n^{(ij)})_{i,j=1,2} = \mathbb{E}(\zeta_n \zeta_n^*).$$

Alors

$$\Phi_n = A_{n-1}^\theta \Phi_{n-1} (A_{n-1}^\theta)^* + \begin{pmatrix} \sigma_n^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad \Phi_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.28)$$

Puisque  $\Phi_n^{(12)} = \Phi_n^{(21)}$ , nous avons

$$\begin{aligned} \Phi_n^{(11)} &= \theta^2 \Phi_{n-1}^{(11)} + 2\theta^2 \beta_{n-1} \Phi_{n-1}^{(12)} + \theta^2 \beta_{n-1}^2 \Phi_{n-1}^{(22)} + \sigma_n^2, \\ \Phi_n^{(12)} &= \theta(1 + \beta_{n-1}^2) \Phi_{n-1}^{(12)} + \theta \beta_{n-1} \Phi_{n-1}^{(11)} + \theta \beta_{n-1} \Phi_{n-1}^{(22)}, \\ \Phi_n^{(22)} &= \Phi_{n-1}^{(22)} + \beta_{n-1}^2 \Phi_{n-1}^{(11)} + 2\beta_{n-1} \Phi_{n-1}^{(12)}. \end{aligned} \quad (3.29)$$

**Lemme 3.5.** *Sous  $H_0$ , nous avons*

$$\Phi_n^{(11)} - \sigma_n^2 \mathbb{E}(\gamma_n^2) = O(n^{1-3\alpha}), \quad \Phi_n^{(12)} = O(n^{1-2\alpha}), \quad \Phi_n^{(22)} = O(n^{2-3\alpha}).$$

**Preuve.** Soit

$$\Psi_n = \begin{pmatrix} \sigma_n^2 \mathbb{E}(\gamma_n^2) & 0 \\ 0 & 0 \end{pmatrix}$$

nous remarquons que

$$\mathbb{E}(\gamma_n^2) = \theta^2 \mathbb{E}(\gamma_{n-1}^2) + 1. \quad (3.30)$$

Alors

$$\Phi_n - \Psi_n = A_{n-1}^\theta (\Phi_{n-1} - \Psi_{n-1}) (A_{n-1}^\theta)^* + V_n$$

avec

$$V_n = \begin{pmatrix} \vartheta^2 \beta_{n-1}^2 \sigma_{n-1}^2 \mathbb{E}(\gamma_{n-1}^2) & \vartheta \beta_{n-1} \sigma_{n-1}^2 \mathbb{E}(\gamma_{n-1}^2) \\ \vartheta \beta_{n-1} \sigma_{n-1}^2 \mathbb{E}(\gamma_{n-1}^2) & \beta_{n-1}^2 \sigma_{n-1}^2 \mathbb{E}(\gamma_{n-1}^2) \end{pmatrix}.$$

Il vient donc que

$$\Phi_n - \Psi_n = \sum_{k=1}^{n-1} A_{n-1}^\theta \cdots A_{n-k}^\theta V_{n-k} (A_{n-k}^\theta)^* \cdots (A_{n-1}^\theta)^* + V_n.$$

Il est montré dans [21], qu'il existe une constante positive  $c_1$  telle que  $\sup_{n \geq 1} \left\| \prod_{j=1}^n A_{n-j}^\theta \right\| < c_1$ . Ici  $\|\cdot\|$  est une norme d'opérateur. Nous avons donc que  $\|V_k\| \leq c_2 \beta_{k-1}$ . D'où,  $\|\Phi_n - \Psi_n\| \leq c_3 \sum_{k=1}^n \beta_k = O(n^{1-\alpha})$ .

En vertu de (3.29), nous avons que

$$\Phi_n^{(12)} = \theta(1 + \beta_n^2) \Phi_{n-1}^{(12)} + O(n^{1-2\alpha}),$$

et donc  $\Phi_n^{(12)} = O(n^{1-2\alpha})$ . De manière analogue,  $\Phi_n^{(22)} = \Phi_{n-1}^{(22)} + O(n^{1-3\alpha})$  et  $\Phi_n^{(22)} = O(n^{2-3\alpha})$ . Nous avons donc via (3.29) et (3.30) que

$$\Phi_n^{(11)} - \sigma_n^2 E(\gamma_n^2) = \theta^2 \left( \Phi_{n-1}^{(11)} - \sigma_{n-1}^2 \mathbb{E}(\gamma_{n-1}^2) \right) + O(n^{1-3\alpha}),$$

ce qui achève la preuve du Lemme. □

**Lemme 3.6.** *Nous avons*

$$\frac{1}{n} \mathbb{E} [M_n(m)^2] - I(\theta) = O(n^{1-3\alpha}). \quad (3.31)$$

**Preuve.** Nous remarquons que

$$\mathbb{E} [M_n(m)^2] = \sum_{k=m+1}^{m+n} \frac{\mathbb{E} [(a_{k-1}^* \zeta_{k-1})^2]}{\sigma_k^2} = \sum_{k=m+1}^{m+n} \frac{a_{k-1}^* \Phi_{k-1} a_{k-1}}{\sigma_k^2}.$$

En vertu du Lemme 3.5,  $a_{k-1}^* \Phi_{k-1} a_{k-1} - \sigma_{k-1}^2 \mathbb{E}(\gamma_{k-1}^2) = O(k^{1-3\alpha})$ . D'où,

$$\frac{1}{n} \sum_{k=m+1}^{m+n} \frac{a_{k-1}^* \Phi_{k-1} a_{k-1} - \sigma_{k-1}^2 \mathbb{E}(\gamma_{k-1}^2)}{\sigma_k^2} \leq \frac{c_1}{n} \sum_{k=m+1}^{m+n} k^{1-3\alpha} = O(n^{1-3\alpha}).$$

Le Lemme est démontré du fait que

$$\sum_{k=m+1}^{m+n} \frac{\sigma_{k-1}^2 \mathbb{E}(\gamma_{k-1}^2)}{\sigma_k^2} - I(\theta) = O(1). \quad \square$$

**Lemme 3.7.** *Il existe une constante  $K > 0$  telle que  $\mathbb{E}(|F_n|^4) \leq K$  pour tout  $n$ .*

**Preuve.** En utilisant le Lemme 3.6,  $\mathbb{E} [(a_k^* \zeta_k)^2] = \sigma_k^2 \mathbb{E}(\gamma_k^2) + O(k^{1-3\alpha})$  est bornée. Comme le suite  $(\varepsilon_n)_{n \geq 1}$  est *i.i.d.* distribuée selon une loi normale standard,

$$\mathbb{E} [(a_n^* \zeta_n)^4] = 3 \left\{ \mathbb{E} [(a_n^* \zeta_n)^2] \right\}^2$$

est bornée. D'où,

$$6 \sum_{k=1}^{n-1} [a_k^* A_{k-1}^\theta \cdots A_k^\theta b]^4 \sigma_k^4 + 2 \sum_{1 \leq k < l \leq n-1} [a_k^* A_{k-1}^\theta \cdots A_k^\theta b]^2 [a_l^* A_{l-1}^\theta \cdots A_l^\theta b]^2 \sigma_k^2 \sigma_l^2$$

est également bornée. □

**Lemme 3.8.** *Nous avons que*

$$\|a_k^* A_{k-1}^\vartheta \cdots A_m^\vartheta\| = O(k^{-\delta}) \quad (3.32)$$

*uniformément en  $m$  où  $\delta > 0$ .*

**Preuve.** Soit

$$T_{k,m} = \prod_{i=1}^{k-m} A_{k-i}^\vartheta.$$

En travaillant block par block, nous obtenons aisément via la relation récursive  $T_{k+1,m} = A_k^\vartheta T_{k,m}$  que

$$\begin{cases} T_{k+1,m}^{(11)} = \vartheta T_{k,m}^{(11)} + \beta_k \vartheta T_{k,m}^{(21)} \\ T_{k+1,m}^{(12)} = \vartheta T_{k,m}^{(12)} + \beta_k \vartheta T_{k,m}^{(22)} \\ T_{k+1,m}^{(21)} = \beta_n T_{k,m}^{(11)} + T_{k,m}^{(21)} \\ T_{k+1,m}^{(22)} = \beta_n T_{k,m}^{(12)} + T_{k,m}^{(22)} \end{cases}$$

où nous utilisons la notation

$$T_{k,m} = \begin{pmatrix} T_{k,m}^{(11)} & T_{k,m}^{(12)} \\ T_{k,m}^{(21)} & T_{k,m}^{(22)} \end{pmatrix}.$$

Grâce à [21], il existe une constante positive  $K$  telle que  $\sup_{n \geq 1} \left\| \prod_{j=1}^n A_{n-j}^\vartheta \right\| < K$ . D'où,

$$\begin{cases} |T_{k+1,m}^{(11)}| \leq |\vartheta T_{k,m}^{(11)}| + |\beta_k \vartheta T_{k,m}^{(21)}| \\ |T_{k+1,m}^{(12)}| \leq |\vartheta T_{k,m}^{(12)}| + |\beta_k \vartheta T_{k,m}^{(22)}| \\ |T_{k+1,m}^{(21)}| \leq |\beta_k| K + |T_{k,m}^{(21)}| \\ |T_{k+1,m}^{(22)}| \leq |\beta_k| K + |T_{k,m}^{(22)}| \end{cases}$$

et donc,

$$\begin{cases} T_{k+1,m}^{(11)} = O(k^{-\alpha}) \\ T_{k+1,m}^{(12)} = O(k^{-\alpha}) \\ T_{k+1,m}^{(21)} = O(k^{1-\alpha}) \\ T_{k+1,m}^{(22)} = O(k^{1-\alpha}) \end{cases}$$

Puisque  $a_k^* T_{k,m} = \left( T_{k,m}^{(11)} + \beta_k T_{k,m}^{(21)}, T_{k,m}^{(12)} + \beta_k T_{k,m}^{(22)} \right)$  le Lemme est démontré.  $\square$

**Lemme 3.9.** *Nous avons*

$$\mathbb{E} \left\{ \left| \mathbb{E} [M_n(m)^2 | \mathcal{F}_m] - \mathbb{E} [M_n(m)^2] \right| \right\} = O(n^{1-2\delta}) \quad (3.33)$$

*uniformément en  $m$  où  $\mathcal{F}_n$  est définie par  $\mathcal{F}_n = \sigma(F_1, \dots, F_n)$ .*

**Preuve.** Les incréments de notre martingale étant décorrélés, nous avons que

$$\begin{aligned}\mathbb{E} [M_n(m)^2 | \mathcal{F}_m] &= \mathbb{E} [(M_{m+n} - M_m)^2 | \mathcal{F}_m] = \sum_{k=m+1}^{m+n} \mathbb{E} [F_k^2 | \mathcal{F}_m] \\ &= \sum_{k=m+1}^{m+n} \sigma_k^{-2} a_{k-1}^* \mathbb{E} [\zeta_{k-1} \zeta_{k-1}^* | \mathcal{F}_m] a_{k-1}.\end{aligned}$$

Posons  $\Theta_k = \mathbb{E} [\zeta_k \zeta_k^* | \mathcal{F}_m]$  pour  $m+1 \leq k \leq m+n$ . Alors

$$\Theta_k = A_{k-1}^\theta \Theta_{k-1} (A_{k-1}^\theta)^* + \begin{pmatrix} \sigma_k^2 & 0 \\ 0 & 0 \end{pmatrix},$$

et

$$\Theta_k - \mathbb{E}(\Theta_k) = A_{k-1}^\theta \cdots A_m^\theta (\Theta_m - \mathbb{E}(\Theta_m)) (A_m^\theta)^* \cdots (A_{k-1}^\theta)^*. \quad (3.34)$$

La suite  $\mathbb{E} [\|\Theta_m - \mathbb{E}(\Theta_m)\|]$  est bornée. En vertu du Lemme 3.8,

$$\mathbb{E} \left\{ \left| \mathbb{E} [M_n(m)^2 | \mathcal{F}_m] - \mathbb{E} [M_n(m)^2] \right| \right\} \leq c_1 \sum_{k=m+1}^{m+n} \|a_k^* A_{k-1}^\theta \cdots A_m^\theta\|^2 \leq c_2 n^{1-2\delta}.$$

Ce qui achève la preuve. □

En utilisant les Lemmes 3.6, 3.7, 3.9 et la Remarque 3.6.1, la Proposition 3.2 s'obtient en appliquant le Théorème [41, Théorème 1]

**Remarque 3.6.1** *Comme le processus obtenu après le filtrage est gaussien, nous avons que au sein du nouvel espace de probabilité (celui donné par le Lemme 3.2) que l'équation (3.46) est vérifiée. En conséquence, le Lemme 3.11 sera vérifié dans ce nouvel espace et*

$$\frac{M_k}{\langle M \rangle_k^{1/2}} - \frac{W(k)}{k^{1/2}} = O(k^{-\kappa}) \text{ p.s.}$$

où  $W(k)$  est le mouvement Brownien standard.

### 3.6.2 Lemmes techniques pour établir la Proposition 3.4

**Lemme 3.10.** *Considérons un vecteur aléatoire  $B_n \in \mathbb{R}^d$  tel que pour tout  $n \geq 1$ ,*

$$B_n \sim \mathcal{N}(0, \Sigma_n),$$

où la matrice de covariance vérifie  $\|\Sigma_n\| = O(1)$ . Alors,

$$\left\| \frac{B_n}{\ln n} \right\| = o(1) \text{ p.s.} \quad (3.35)$$

**Preuve.** Pour tout  $\varepsilon > 0$ , nous avons

$$\begin{aligned} \mathbf{P}\left(\left\|\frac{B_n}{\ln n}\right\| > \varepsilon\right) &= \mathbf{P}\left(\|B_n\|^2 > \ln n^2 \varepsilon^2\right) \\ &= \mathbf{P}\left(\langle \Sigma_n \mu_n, \mu_n \rangle > \ln n^2 \varepsilon^2\right) \leq \mathbf{P}\left(\|\mu_n\|^2 > \frac{\ln n^2 \varepsilon^2}{\|\Sigma_n\|}\right), \end{aligned}$$

où  $\|\mu_n\|^2 \sim \chi^2(d)$ , ce qui, implique

$$\begin{aligned} \mathbf{P}(\|\mu_n\|^2 > \ln n^2 \varepsilon^2 \|\Sigma_n\|^{-1}) &= \int_{\ln n^2 \varepsilon^2 \|\Sigma_n\|^{-1}}^{\infty} c_1(d) x^{\frac{d}{2}-1} \exp\left(-\frac{x}{2}\right) dx \\ &\leq \int_{\ln n^2 \varepsilon^2 \|\Sigma_n\|^{-1}}^{\infty} c_2(d) \exp\left(-\frac{x}{4}\right) dx \end{aligned}$$

pour  $n$  assez grand, où  $c_1(d)$  et  $c_2(d)$  sont deux constantes positives indépendantes de  $x$  et  $n$ . Grâce à l'hypothèse  $\|\Sigma_n\|$  et l'inégalité  $8 \ln n \leq \ln n^2 \varepsilon^2 \|\Sigma_n\|^{-1}$  pour  $n$  assez grand, nous obtenons

$$\int_{\ln n^2 \|\Sigma_n\|^{-1} \varepsilon^2}^{\infty} c_2(d) \exp\left(-\frac{x}{4}\right) dx \leq 4c_2(d)n^{-2} = O(n^{-2}).$$

Il ne reste plus qu'à appliquer le Lemme de Borel-Cantelli pour achever la démonstration.  $\square$

**Lemme 3.11.** *Sous  $H_0$ , nous avons*

$$\zeta_n^{(1)} - \sigma_n \gamma_n = O(n^{1-2\alpha+\lambda}) \text{ p.s.}, \quad \zeta_n^{(2)} = O(n^{1-\alpha+\lambda}) \text{ p.s.} \quad (3.36)$$

pour tout  $\lambda > 0$ .

**Preuve.** Soit,

$$T_n = \begin{pmatrix} \vartheta & 0 \\ 0 & 1 \end{pmatrix} T_{n-1} + b\sigma_n \varepsilon_n \quad (3.37)$$

où  $T_0$  est un vecteur Gaussien dégénéré de seconde composante nulle et notons

$$\Delta_n = \zeta_n - T_n. \quad (3.38)$$

Nous obtenons

$$\Delta_n = A_{n-1}^{\vartheta} \Delta_{n-1} - \begin{pmatrix} 0 \\ \beta_{n-1} T_{n-1}^{(1)} \end{pmatrix} \quad (3.39)$$

Nous rappelons de nouveau que dans [21], il existe une constante positive  $K$  telle que  $\sup_{n \geq 1} \left\| \prod_{j=1}^n A_{n-j}^{\vartheta} \right\| < K$ . Par (3.39)

$$\zeta_n^{(2)} = \Delta_n^{(2)}. \quad (3.40)$$

Le Lemme 3.10, nous permet d'obtenir l'estimation suivante  $T_n = O(\ln n)$  p.s. Grâce à (3.40), il existe  $C := C(\omega) > 0$  telle que  $|\zeta_n^{(2)}| \leq KC \ln n \sum_{k=1}^{n-1} |\beta_k|$ . Cette inégalité conjuguée à l'estimation  $\beta_n = O(n^{-\alpha})$  implique

$$|\zeta_n^{(2)}| = O(n^{1-\alpha+\lambda}) \text{ p.s.} \quad (3.41)$$

pour tout  $\lambda > 0$ .

Il reste à obtenir une estimation sur  $\zeta_n^{(1)}$ . Puisque

$$\sigma_n \gamma_n = \vartheta \sqrt{(1 - \beta_{n-1}^2) \sigma_{n-1} \gamma_{n-1}} + \sigma_n \varepsilon_n$$

et

$$\zeta_n^{(1)} = \vartheta \zeta_{n-1}^{(1)} + \vartheta \beta_{n-1} \zeta_{n-1}^{(2)} + \sigma_n \varepsilon_n,$$

nous avons via un développement de Taylor que

$$\zeta_n^{(1)} - \sigma_n \gamma_n = \vartheta \left( \zeta_{n-1}^{(1)} - \sigma_{n-1} \gamma_{n-1} \right) + \vartheta O(\beta_{n-1}^2) \sigma_{n-1} \gamma_{n-1} + \vartheta \beta_{n-1} \zeta_{n-1}^{(2)}. \quad (3.42)$$

Grâce à (3.41), nous obtenons que  $\vartheta \beta_{n-1}^2 \sigma_{n-1} \gamma_{n-1} + \vartheta \beta_{n-1} \zeta_{n-1}^{(2)} = O(n^{1-2\alpha+\lambda})$  *p.s.* pour tout  $\lambda > 0$ . Par conséquent il existe  $C := C(\omega) > 0$  telle que

$$|\zeta_n^{(1)} - \sigma_n \gamma_n| \leq C \sum_{k=1}^{n-1} \vartheta^{n-k} k^{1-2\alpha+\lambda} = O(n^{1-2\alpha+\lambda}) \text{ p.s.}, \quad (3.43)$$

ce qui achève la preuve du Lemme. □

Nous rappelons que la martingale s'écrit  $M_k = M_k(0) = \sum_{n=2}^k \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \varepsilon_n$  et que sa variation quadratique prévisible s'écrit  $\langle M \rangle_k = \sum_{n=2}^k \left( \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \right)^2$ , nous avons les Lemmes suivants qui sont les ingrédients clés des preuves des Propositions.

**Lemme 3.12.** *Considérons  $U_k = \sum_{n=1}^{k-1} Z_n$  où  $Z_n, n = 1, 2, \dots, k-1$  est une suite i.i.d distribuée selon une loi normale centrée de variance  $I(\theta) = \frac{1}{1-\theta^2}$ , nous avons*

$$\frac{M_k^2}{I(\theta)} - \frac{U_k^2}{I(\theta)} = O(k^{1-\lambda^4}) \text{ p.s.}$$

as  $k \rightarrow \infty$  *p.s.* pour  $\lambda^4 > 0$ .

**Preuve.** La différence vaut

$$\frac{M_k^2}{I(\theta)} - \frac{U_k^2}{I(\theta)} = (M_k + U_k)(M_k - U_k) / I(\vartheta)$$

Grâce à l'équation (3.20) nous avons  $M_k - U_k = O(k^{1/2-\kappa})$  lorsque  $k \rightarrow \infty$ , nous obtenons d'un autre côté grâce à la loi du logarithme itéré que,  $U_k / I(\theta) = O((k \ln \ln k)^{1/2})$  *p.s.* ce qui achève la preuve. □

**Lemme 3.13.** *Nous avons*

$$\frac{M_k^2}{\langle M \rangle_k} - \frac{U_k^2}{kI(\theta)} \rightarrow 0, \quad k \rightarrow \infty, \text{ p.s.}$$

**Preuve.** En utilisant le Lemme 3.12, il suffit de montrer que

$$\frac{M_k^2}{\langle M \rangle_k} - \frac{M_k^2}{kI(\vartheta)} = \frac{1}{I(\vartheta)} \left( \frac{M_k}{k^{1/2}} \right)^2 \left( I(\vartheta) - \frac{\langle M \rangle_k}{k} \right) \frac{k}{\langle M \rangle_k} \rightarrow 0, \quad k \rightarrow \infty, \quad p.s.$$

En vertu de la loi du logarithme itéré (voir [93]) nous avons que  $\frac{1}{k} \sum_{n=2}^k \gamma_{n-1}^2 - I(\vartheta) = O((\ln \ln k/k)^{1/2})$  *p.s.* et grâce au Lemme 3.10,  $\gamma_n = O(\ln n)$  *p.s.*

Nous avons donc via le Lemme 3.11 et la Remarque 3.6.1 que

$$\begin{aligned} \left| \frac{1}{k} \sum_{n=2}^k \frac{(a_n^* \zeta_n - \sigma_n \gamma_n + \sigma_n \gamma_n)^2}{\sigma_{n+1}^2} - I(\theta) \right| &= \left| \frac{1}{k} \sum_{n=2}^k \frac{(a_n^* \zeta_n - \sigma_n \gamma_n)^2}{\sigma_{n+1}^2} - \frac{2}{k} \sum_{n=2}^k \frac{(a_n^* \zeta_n - \sigma_n \gamma_n) \sigma_n \gamma_n}{\sigma_{n+1}^2} \right. \\ &\quad \left. + \frac{1}{k} \sum_{n=2}^k \frac{\sigma_n^2}{\sigma_{n+1}^2} \gamma_n^2 - I(\theta) \right| \\ &= O(k^{-\delta_2}) \quad p.s. \end{aligned}$$

D'où,

$$\frac{\langle M \rangle_k}{k} - I(\theta) = O(k^{-\delta_2}) \quad p.s. \quad (3.44)$$

pour  $\delta_2 > 0$  et donc la convergence presque sûre vers 0 as  $k \rightarrow \infty$ , est vérifiée.  $\square$

**Lemme 3.14.** *Pour tout  $\epsilon' > 0$  lorsque  $n \rightarrow \infty$ ,*

$$\mathbb{P} \left[ \max_{1 < k \leq N} \frac{M_k^2}{\langle M \rangle_k} = \max_{1 < k \leq \epsilon' N} \frac{M_k^2}{\langle M \rangle_k} \right] \rightarrow 1.$$

**Preuve.** Grâce à la preuve de la Proposition 3.3,

$$\max_{\epsilon' N \leq k \leq N} \frac{M_k^2}{\langle M \rangle_k} \xrightarrow{\mathcal{L}} \max_{t \in [\epsilon', 1]} \frac{W^2(t)}{t}.$$

Les estimations  $\max_{\epsilon' N \leq k \leq N} \frac{M_k^2}{\langle M \rangle_k} = O_{\mathbb{P}}(1)$  et  $\max_{1 < k \leq \epsilon' N} \frac{M_k^2}{\langle M \rangle_k} \xrightarrow{\mathbb{P}} \infty$  achèvent la preuve.  $\square$

**Lemme 3.15.** *Pour tout  $0 < \epsilon < 1/2$  les variables aléatoires  $\max_{1 \leq k \leq \epsilon N} \frac{M_k^2}{\langle M \rangle_k}$  et  $\max_{(1-\epsilon)N \leq k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k}$  sont asymptotiquement indépendantes.*

**Preuve.** De par le Lemme 3.2 et la remarque 3.6.1 nous avons que :

$$\max_{1 \leq k \leq \epsilon N} \frac{M_k^2}{\langle M \rangle_k} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \max_{0 \leq t \leq \epsilon} \frac{W(t)^2}{t}$$

et

$$\max_{(1-\epsilon)N \leq k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \max_{1-\epsilon \leq t \leq 1} \frac{(W(1) - W(t))^2}{1-t}$$

Le Lemme découle du fait que variables aléatoires  $W(t)$  and  $W(1) - W(t)$  sont indépendantes.  $\square$

### 3.6.3 Preuve de la Proposition 3.3

Premièrement le ratio de vraisemblance gaussien tel que décrit dans (3.15) peut se reformuler

$$\begin{aligned}\Lambda_N(k) &= \min_{\vartheta \in \Theta} \sum_{n=1}^N \left( \frac{b^*(\zeta_n - A_{n-1}^\vartheta \zeta_{n-1})}{\sigma_n} \right)^2 - \min_{\theta_1 \in \Theta} \sum_{n=1}^k \left( \frac{b^*(\zeta_n - A_{n-1}^{\theta_1} \zeta_{n-1})}{\sigma_n} \right)^2 \\ &\quad - \min_{\theta_2 \in \Theta} \sum_{n=k+1}^N \left( \frac{b^*(\zeta_n - A_{n-1}^{\theta_2} \zeta_{n-1})}{\sigma_n} \right)^2. \\ &= Q_1 - Q_2 - Q_3\end{aligned}$$

où

$$\begin{aligned}Q_1 &= \sum_{n=1}^N \left( \frac{b^* \zeta_n}{\sigma_n} \right)^2 - \frac{\left( \sum_{n=1}^N \frac{b^* \zeta_n a_{n-1}^* \zeta_{n-1}}{\sigma_n^2} \right)^2}{\sum_{n=1}^N \left( \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \right)^2} \\ Q_2 &= \sum_{n=1}^k \left( \frac{b^* \zeta_n}{\sigma_n} \right)^2 - \frac{\left( \sum_{n=1}^k \frac{b^* \zeta_n a_{n-1}^* \zeta_{n-1}}{\sigma_n^2} \right)^2}{\sum_{n=1}^k \left( \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \right)^2} \\ Q_3 &= \sum_{n=k+1}^N \left( \frac{b^* \zeta_n}{\sigma_n} \right)^2 - \frac{\left( \sum_{n=k+1}^N \frac{b^* \zeta_n a_{n-1}^* \zeta_{n-1}}{\sigma_n^2} \right)^2}{\sum_{n=k+1}^N \left( \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \right)^2}.\end{aligned}\tag{3.45}$$

avec  $a_n = \begin{pmatrix} 1 \\ \beta_n \end{pmatrix}$  et  $b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  comme dans (3.14). Alors, nous obtenons via la relation récursive

$$\zeta_n = A_{n-1}^\theta \zeta_{n-1} + b \sigma_n \varepsilon_n,\tag{3.46}$$

où  $A_n^\theta$  et  $b$  sont définis dans (3.14), que

$$\frac{b^* \zeta_n}{\sigma_n} = \frac{\theta a_{n-1}^* \zeta_{n-1}}{\sigma_n} + \varepsilon_n, \quad n \geq 1$$

et  $\zeta_0 = \mathbf{0}_{2 \times 2}$  mais, également,

$$Q_1 = \sum_{n=1}^N \varepsilon_n^2 - \frac{\left( \sum_{n=2}^N \frac{a_{n-1}^* \zeta_{n-1} \varepsilon_n}{\sigma_n} \right)^2}{\sum_{n=2}^N \left( \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \right)^2}.$$

Les calculs précédents montrent donc, que, sous  $H_0$

$$\Lambda_N(k) = \frac{\left(\sum_{n=2}^k \frac{a_{n-1}^* \zeta_{n-1} \varepsilon_n}{\sigma_n}\right)^2}{\sum_{n=2}^k \left(\frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n}\right)^2} + \frac{\left(\sum_{n=k+1}^N \frac{a_{n-1}^* \zeta_{n-1} \varepsilon_n}{\sigma_n}\right)^2}{\sum_{n=k+1}^N \left(\frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n}\right)^2} - \frac{\left(\sum_{n=2}^N \frac{a_{n-1}^* \zeta_{n-1} \varepsilon_n}{\sigma_n}\right)^2}{\sum_{n=2}^N \left(\frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n}\right)^2}. \quad (3.47)$$

La Proposition 3.3 découle de la limite en distribution de (3.47) obtenue via la Proposition 3.2.

### 3.6.4 Preuve de la Proposition 3.4

Puisque la suite  $U_k/\sqrt{I(\theta)}$  du Lemme 3.12 vérifie les hypothèses du Corollaire A.2 dans [35], il vient que

$$\mathbb{P}\left[\left(\max_{1 < k \leq N} \frac{U_k^2}{kI(\theta)} - \beta(N)\right) / \alpha(N) \leq x\right] \rightarrow \exp(-e^{-x/2}), \quad N \rightarrow \infty. \quad (3.48)$$

Nous pouvons à présent, achever la preuve qui s'articule de la même façon à celle de [35] qui constituent les point clés pour établir la Proposition 3.4.

Nous via la Proposition 3.3 que

$$\begin{aligned} \left| \frac{\max_{1 < k \leq \epsilon' N} \Lambda_N(k) - \max_{1 < k \leq \epsilon' N} \frac{M_k^2}{\langle M \rangle_k}}{\alpha(N)} \right| &\leq \max_{1 < k \leq \epsilon' N} \left| \frac{\Lambda_N(k) - \frac{M_k^2}{\langle M \rangle_k}}{\alpha(N)} \right| \\ &\leq \max_{1 < k \leq \epsilon' N} \left| \frac{\frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} - \frac{M_N^2}{\langle M \rangle_N}}{\alpha(N)} \right| \\ &\xrightarrow{\mathcal{L}} \max_{0 < t \leq \epsilon'} \left| \frac{(W(1) - W(t))^2}{1 - t} - W^2(1) \right|, \quad N \rightarrow \infty \\ &\rightarrow 0, \quad \epsilon' \rightarrow 0, \quad p.s. \end{aligned}$$

car  $\alpha(N) \rightarrow 1$ . Cela prouve (3.22) et (3.23) s'obtient exactement de la même manière. Nous avons également via le Lemme 3.13 que

$$\left| \max_{M \leq k \leq N} \frac{M_k^2}{\langle M \rangle_k} - \max_{M \leq k \leq N} \frac{U_k^2}{kI(\theta)} \right| \leq \sup_{k \geq M} \left| \frac{M_k^2}{\langle M \rangle_k} - \frac{U_k^2}{kI(\theta)} \right| \rightarrow 0 \quad (3.49)$$

quand  $M \rightarrow \infty$ , *p.s.* Le Lemme 3.14, conjuguées aux relations (3.48) et (3.49) implique (3.24). Les mêmes arguments nous conduisent à la relation (3.25).

## 3.7 Preuve du résultat principal

Pour tout  $0 < \varepsilon < 1/2$ , nous divisons le maximum décrit par  $\Lambda_N$  en trois termes sur les ensembles  $(1 < k \leq \varepsilon N)$ ,  $(\varepsilon N < k \leq (1 - \varepsilon)N)$  et  $((1 - \varepsilon)N < k \leq N)$  respectivement. En conséquence l'ensemble suivant peut être présenté de la manière suivante

$$\left\{ \frac{\Lambda_N - \beta(N)}{\alpha(N)} \leq x \right\} = L_1 \cap L_2 \cap L_3$$

où

$$L_1 = \left\{ \left( \max_{1 < k \leq \varepsilon N} \Lambda_N(k) - \max_{1 < k \leq \varepsilon N} \frac{M_k^2}{\langle M \rangle_k} \right) / \alpha(N) + \left( \max_{1 < k \leq \varepsilon N} \frac{M_k^2}{\langle M \rangle_k} - \beta(N) \right) / \alpha(N) \leq x \right\},$$

$$L_2 = \left\{ \left( \max_{\varepsilon N < k \leq (1-\varepsilon)N} \Lambda_N(k) - \beta(N) \right) / \alpha(N) \leq x \right\},$$

et

$$L_3 = \left\{ \left( \max_{(1-\varepsilon)N < k \leq N} \Lambda_N(k) - \max_{(1-\varepsilon)N < k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} \right) / \alpha(N) + \left( \max_{(1-\varepsilon)N < k \leq N} \frac{(M_N - M_k)^2}{\langle M \rangle_N - \langle M \rangle_k} - \beta(N) \right) / \alpha(N) \leq x \right\}.$$

En utilisant la Proposition 3.3 et la Remarque 3.6.1 de la sous Section 3.6.1, nous obtenons

$$\max_{\varepsilon N \leq k \leq (1-\varepsilon)N} \Lambda_N(k) - \beta(N) \xrightarrow[N \rightarrow \infty]{\mathbf{P}} -\infty.$$

Nous déduisons de la mutuelle indépendance asymptotique (voir Lemme 3.15) et de la Proposition 3.4 que

$$\limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left( \max_{1 \leq k \leq N} \Lambda_N(k) - \beta(N) \right) / \alpha(N) \leq x \right] = A_{1,x} \times A_{2,x}$$

où

$$A_{1,x} = \lim_{\varepsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left( \max_{1 < k \leq \varepsilon N} \Lambda_N(k) - \beta(N) \right) / \alpha(N) \leq x \right]$$

et

$$A_{2,x} = \lim_{\varepsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left( \max_{(1-\varepsilon)N < k \leq N} \Lambda_N(k) - \beta(N) \right) / \alpha(N) \leq x \right].$$

Nous souhaitons à présent identifier les termes limites  $A_{1,x}$  et  $A_{2,x}$ . Les équations (3.22) et (3.24) nous permettent d'appliquer le Théorème 4.2 de [15] afin d'obtenir

$$\limsup_{N \rightarrow \infty} \mathbb{P} \left[ \left( \max_{1 \leq k \leq \varepsilon N} \Lambda_N(k) - \beta(N) \right) / \alpha(N) \leq x \right] = \exp \left( -e^{-\frac{x}{2}} \right).$$

Puisque la valeur limite ne dépend pas de  $\varepsilon$  nous obtenons  $A_{1,x} = \exp \left( -e^{-\frac{x}{2}} \right)$ . Nous obtenons également la valeur de  $A_{2,x}$  via un même raisonnement, ce qui démontre le Théorème 3.1.

# Chapitre 4

## Un modèle chronologique à coefficients aléatoires corrélés

### 4.1 Historique et motivation

Les résultats présentés dans ce chapitre sont résumés dans [84] et sont le fruit d'un travail mené avec mon co-directeur de thèse que je tiens à remercier.

En économétrie, les séries chronologiques non linéaires sont très utilisées. Nous nous intéressons à une généralisation du processus autorégressif d'ordre 1 où les coefficients sont aléatoires. Nous rappelons que le processus autorégressif d'ordre 1 RCAR(1) est défini pour tout  $t \in \mathbb{Z}$  par la relation récursive

$$X_t = (\theta + \eta_t)X_{t-1} + \epsilon_t$$

où  $(\epsilon_t)$  et  $(\eta_t)$  sont des bruits blancs décorrélés. De par les travaux fondateurs de Anděl [1] et Nicholls-Quinn [79], diverses conditions suffisantes portant sur les moments de  $(\epsilon_t)$  et  $(\eta_t)$  furent exhibées afin d'assurer la stationnarité de ce modèle. Nous rappelons qu'il est démontré que ce modèle est stationnaire au second ordre dès lors que  $\theta^2 + \tau_2 < 1$  où  $\tau_2$  désigne la variance de  $(\eta_t)$ . Très récemment, Aue *et al.* [5] ont donné des conditions nécessaires et suffisantes afin d'assurer l'existence d'une solution strictement stationnaire dans un modèle RCAR(1), en s'inspirant d'un travail plus général de Brandt [18], quelques conditions techniques s'inspirent donc de ce travail. Cependant la souplesse d'un modèle RCAR est induite par l'absence de mémoire entre deux valeurs consécutives des coefficients aléatoires. Dans le cadre d'une modélisation chronologique, cela nous semble peu pertinent. Notre objectif principal est de montrer que la violation de l'hypothèse d'indépendance entre les coefficients, qui est vraisemblable dans une expérience aléatoire, induit un biais dans l'ensemble de la procédure d'estimation, et donc une mauvaise interprétation des indicateurs statistiques. Nous suggérons pour cette raison d'étudier dans ce chapitre un modèle RCAR(1) où les coefficients possèdent une mémoire finie, qui se matérialise sous la forme d'une moyenne mobile d'ordre 1, pour laquelle nous estimerons sa moyenne comme si il s'agissait d'un modèle RCAR(1) classique. Posons à cet égard, pour tout  $t \in \mathbb{Z}$ , un processus autorégressif d'ordre 1 s'écrivant

$$X_t = \theta_t X_{t-1} + \epsilon_t \tag{4.1}$$

où  $\theta_t$  est un coefficient aléatoire engendré par une moyenne mobile

$$\theta_t = \theta + \alpha \eta_{t-1} + \eta_t. \quad (4.2)$$

Ce choix comme structure de dépendance entre les coefficients est motivé par la Proposition 3.2.1 de [20] qui stipule que tout processus stationnaire ayant une mémoire finie peut se présenter comme une moyenne mobile. En d'autres termes, il existe un bruit blanc tel que le coefficient aléatoire admet la décomposition donnée ci-dessus, ce qui justifie (4.2). Nous pouvons trouver l'étude d'un modèle similaire dans Koubkovà [64] où dans un contexte plus général dans Brandt [18], mais nous verrons que les objectifs de ce chapitre sont différents. Tandis que ces travaux s'intéressent aux propriétés de la solution stationnaire, une grande partie de chapitre s'attache à l'inférence statistique. Les hypothèses que nous retiendront seront présentées à la fin de cette introduction, la Section 4.2 est dévolue à la présentation de conditions permettant de garantir l'existence, l'unicité, d'une solution stationnaire de notre modèle. Ces conditions préliminaires nous permettront ensuite de nous intéresser à la fonction d'autocorrélation. Dans la Section 4.3, la moyenne empirique et les estimateurs usuels de  $\theta$  et  $\sigma_2$  sont étudiés, où  $\sigma_2$  désigne la variance de  $(\epsilon_t)$ .

En particulier, nous obtenons des convergences presque sûres, des normalités asymptotiques et des estimations sur les vitesses de convergence. Nous avons besoin pour cela d'une condition assurant l'existence des moments d'ordre quatre que nous expliciterons. Le corollaire surprenant de ces calculs est que l'estimation renvoyée par le LSE est non consistante envers  $\theta$  dès lors que  $\alpha \neq 0$ , tandis qu'il est bien connu que ce même estimateur préserve la consistance envers  $\theta$  dans un modèle RCAR(1). Cela nous conduit donc dans la Section 4.4 à construire une estimation consistante et d'obtenir sa normalité asymptotique, pour en extraire une procédure statistique afin de détecter une corrélation dans les coefficients. Nous démontrons l'ensemble de ces résultats dans la Section 4.8. L'estimation des paramètres d'un processus RCAR fût intensivement étudiée dans le cas stationnaire, par exemple Nicholls-Quinn [78] et plus tard [90], utilisent la méthode des moindres carrés ou le maximum de quasi-vraisemblance. Le point clé des ces travaux est que les estimateurs considérés sont consistants envers  $\theta$  tandis qu'il apparaît dans nos résultats que l'introduction de la corrélation dans les coefficients obère cette consistance. D'une manière générale, notre objectif est de montrer l'impact de l'introduction de cette mémoire dans la procédure d'estimation d'un modèle ordinaire, afin d'ouvrir des perspectives pour des structures de dépendances plus complexes. Tout au long de ce chapitre, nous rappellerons des résultats connus concernant le processus *RCAR* au premier ordre qui coïncident à nos résultats dès lors que  $\alpha = 0$ . Le lecteur pourra trouver une large littérature dans Nicholls-Quinn [80] et nous mentionnons également les travaux [87], [62], [58], [59], [11] au sujet de l'inférence au sein d'un RCAR, mais également le travail de Aue et Horváth [4] au sujet d'une procédure d'estimation unifiée dans le cas stationnaire comme non stationnaire.

Pour tout  $a > 0$ , nous notons les moments

$$\sigma_a = \mathbb{E}(\epsilon^a), \quad \tau_a = \mathbb{E}(\eta_0^a).$$

Nous considérons afin de simplifier les calculs la famille de vecteurs donnée par

$$U_0 = \begin{pmatrix} 1 \\ 0 \\ \tau_2 \end{pmatrix}, U_1 = \begin{pmatrix} 0 \\ \tau_2 \\ 0 \end{pmatrix}, U_2 = \begin{pmatrix} \tau_2 \\ 0 \\ \tau_4 \end{pmatrix}. \quad (4.3)$$

Nous ferons régulièrement également appel à la matrice suivante de taille  $3 \times 3$  caractérisant les moments d'ordre 2 basée sur la concaténation de  $\{U_0, U_1, U_2\}$  de la manière suivante

$$M = \begin{pmatrix} \theta^2 + \tau_2 & 2\alpha\theta & \alpha^2 \\ 2\theta\tau_2 & 2\alpha\tau_2 & 0 \\ \theta^2\tau_2 + \tau_4 & 2\alpha\theta\tau_2 & \alpha^2\tau_2 \end{pmatrix} \text{ avec } \begin{cases} M_1 = \theta^2 U_0 + 2\theta U_1 + U_2 \\ M_2 = 2\alpha(\theta U_0 + U_1) \\ M_3 = \alpha^2 U_0. \end{cases} \quad (4.4)$$

Nous décrirons de manière analogue les moments d'ordre quatre du processus via les vecteurs  $\{V_0, \dots, V_4\}$  où

$$V_0 = \begin{pmatrix} 1 \\ 0 \\ \tau_2 \\ 0 \\ \tau_4 \end{pmatrix}, V_1 = \begin{pmatrix} 0 \\ \tau_2 \\ 0 \\ \tau_4 \\ 0 \end{pmatrix}, V_2 = \begin{pmatrix} \tau_2 \\ 0 \\ \tau_4 \\ 0 \\ \tau_6 \end{pmatrix}, V_3 = \begin{pmatrix} 0 \\ \tau_4 \\ 0 \\ \tau_6 \\ 0 \end{pmatrix}, V_4 = \begin{pmatrix} \tau_4 \\ 0 \\ \tau_6 \\ 0 \\ \tau_8 \end{pmatrix}. \quad (4.5)$$

Nous utilisons ces derniers afin de bâtir la matrice  $H$  de taille  $5 \times 5$  dont les colonnes sont définies par

$$\begin{cases} H_1 = \theta^4 V_0 + 4\theta^3 V_1 + 6\theta^2 V_2 + 4\theta V_3 + V_4 \\ H_2 = 4\alpha(\theta^3 V_0 + 3\theta^2 V_1 + 3\theta V_2 + V_3) \\ H_3 = 6\alpha^2(\theta^2 V_0 + 2\theta V_1 + V_2) \\ H_4 = 4\alpha^3(\theta V_0 + V_1) \\ H_5 = \alpha^4 V_0. \end{cases} \quad (4.6)$$

Elle s'écrit explicitement

$$H = \begin{pmatrix} \theta^4 + 6\theta^2\tau_2 + \tau_4 & 4\alpha(\theta^3 + 3\theta\tau_2) & 6\alpha^2(\theta^2 + \tau_2) & 4\alpha^3\theta & \alpha^4 \\ 4\theta^3\tau_2 + 4\theta\tau_4 & 4\alpha(3\theta^2\tau_2 + \tau_4) & 12\alpha^2\theta\tau_2 & 4\alpha^3\tau_2 & 0 \\ \theta^4\tau_2 + 6\theta^2\tau_4 + \tau_6 & 4\alpha(\theta^3\tau_2 + 3\theta\tau_4) & 6\alpha^2(\theta^2\tau_2 + \tau_4) & 4\alpha^3\theta\tau_2 & \alpha^4\tau_2 \\ 4\theta^3\tau_4 + 4\theta\tau_6 & 4\alpha(3\theta^2\tau_4 + \tau_6) & 12\alpha^2\theta\tau_4 & 4\alpha^3\tau_4 & 0 \\ \theta^4\tau_4 + 6\theta^2\tau_6 + \tau_8 & 4\alpha(\theta^3\tau_4 + 3\theta\tau_6) & 6\alpha^2(\theta^2\tau_4 + \tau_6) & 4\alpha^3\theta\tau_4 & \alpha^4\tau_4 \end{pmatrix}.$$

Diverses hypothèses concernant les paramètres sont requises (pas toujours simultanément) tout au long du chapitre, intimement liées aux perturbations dirigeant notre processus.

- (H<sub>1</sub>) Les processus  $(\epsilon_t)$  et  $(\eta_t)$  sont des bruits blancs forts mutuellement indépendants tels que  $\mathbb{E}(\log^+ |\epsilon|) < \infty$  et  $\mathbb{E}(\log |\theta + \alpha\eta_0 + \eta_1|) < 0$ .
- (H<sub>2</sub>)  $\sigma_{2k+1} = \tau_{2k+1} = 0$  pour tout  $k \in \mathbb{N}$  dès lors que ces moments existent.
- (H<sub>3</sub>)  $\sigma_2 > 0$ ,  $\tau_2 > 0$ ,  $\sigma_2 < \infty$ ,  $\tau_4 < \infty$  et  $\rho(M) < 1$ .
- (H<sub>4</sub>)  $\sigma_4 < \infty$ ,  $\tau_8 < \infty$  and  $\rho(H) < 1$ .
- (H<sub>5</sub>) Il existe des applications continues  $g$  et  $h$  telles que  $\sigma_4 = g(\sigma_2)$  et  $\tau_4 = h(\tau_2)$ .

**Remarque 4.1.1** ( $H_2$ ) peut bien évidemment être remplacée par la condition suivante moins contraignante  $\sigma_1 = \tau_1 = 0$ . Le fait de considérer que chaque moment d'ordre impair de  $(\epsilon_t)$  et  $(\eta_t)$  est nul relève simplement d'une simplification des calculs, qui comme nous le verrons sont assez fastidieux. Une étude plus générale (et possible) pourrait inclure les contributions de  $\sigma_3, \tau_3, \tau_5$  et  $\tau_7$  dans l'ensemble des calculs.

**Remarque 4.1.2** ( $H_5$ ) est vérifiée dans le cas d'une loi normale centrée où  $g(t) = h(t) = 3t^2$ . Elle est vérifiée dans la plupart des distributions perturbant un modèle de régression (loi uniforme centrée, de Student, de Laplace, etc.). Il s'agit d'une hypothèse très forte employée seulement à la fin de notre étude pour bâtir une procédure statistique.

Nous donnons des informations complémentaires au sujet des remarques apparaissant dans les Sections 4.2 et 4.3 dès le début de la Section 4.8.

## 4.2 Stationnarité et fonction d'autocorrélation

Il est aisé d'établir que la suite des coefficients aléatoires  $(\theta_t)$  donnée par (4.2) est strictement stationnaire et ergodique de moyenne  $\theta$  et de fonction d'autocovariance donnée par

$$\gamma_\theta(0) = \tau_2(1 + \alpha^2), \quad \gamma_\theta(1) = \alpha \tau_2, \quad \gamma_\theta(h) = 0 \quad (|h| > 1).$$

Toute solution de (4.1) vérifie une relation de récurrence, et le premier résultat est donc de formuler une condition assurant l'existence d'une solution causale, strictement stationnaire et ergodique.

**Théorème 4.1.** *Supposons que ( $H_1$ ) est vérifiée. Alors presque sûrement, pour tout  $t \in \mathbb{Z}$ ,*

$$X_t = \epsilon_t + \sum_{k=1}^{\infty} \epsilon_{t-k} \prod_{\ell=0}^{k-1} (\theta + \alpha \eta_{t-\ell-1} + \eta_{t-\ell}). \quad (4.7)$$

*En conséquence,  $(X_t)$  est strictement stationnaire et ergodique.*

**Preuve.** Voir Section 4.8.2. □

Par extension, le même type de conclusion peut être obtenu pour un processus de la forme  $(\epsilon_t^a \eta_t^b X_t^c)$  où  $a, b, c \geq 0$ , en formulant des conditions adéquates sur les moments. Cela a pour Corollaire, d'assurer l'existence de  $\mathbb{E}(\epsilon_t^a \eta_t^b X_t^c)$  afin d'identifier le comportement asymptotique (lorsque  $n \rightarrow \infty$ ) d'une moyenne empirique de la forme

$$\frac{1}{n} \sum_{t=1}^n \epsilon_t^a \eta_t^b X_t^c.$$

En vertu de la représentation causale issue du Théorème 4.7, le processus  $(X_t)$  est adapté à la filtration suivante définie par

$$\mathcal{F}_t = \sigma((\epsilon_s, \eta_s), s \leq t). \quad (4.8)$$

Nous nous intéressons à présent aux propriétés de second ordre du processus, sous des hypothèses additionnelles. Nous déduisons de cela sa fonction d'autocorrélation en utilisant les notations précédentes et en posant

$$N = \begin{pmatrix} \theta & \alpha & 0 \\ \tau_2 & 0 & 0 \\ \theta \tau_2 & \alpha \tau_2 & 0 \end{pmatrix} \text{ avec } \begin{cases} N_1 = \theta U_0 + U_1 \\ N_2 = \alpha U_0 \\ N_3 = 0, \end{cases} \quad (4.9)$$

et nous déduisons de ce résultat l'existence mais aussi l'unicité d'une solution stationnaire au second-ordre.

**Théorème 4.2.** *Supposons que les hypothèses  $(H_1)$ – $(H_3)$  sont vérifiées. Alors,  $(X_t)$  est strictement et faiblement stationnaire de moyenne nulle et de fonction d'autocovariance donnée par*

$$\gamma_X(h) = \sigma_2 [N^{|h|} (I_3 - M)^{-1} U_0]_1 \quad (4.10)$$

pour  $h \in \mathbb{Z}$ . Sa fonction d'autocorrélation est définie par

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}. \quad (4.11)$$

De plus, il s'agit de l'unique solution causale, ergodique, strictement et faiblement stationnaire.

**Preuve.** Voir Section 4.8.3. □

**Remarque 4.2.1** *Supposons que le processus est stationnaire avec des moments d'ordre deux vérifiant  $2\alpha\tau_2 = 1$ . Alors, (4.10) conduit à  $\gamma_X(0) = 0$ , ce qui implique que  $(X_t)$  est déterministe. Ce cas pathologique est naturellement exclu de notre étude, de même que  $\sigma_2 = 0$  conduisant à une conclusion analogue.*

**Remarque 4.2.2** *Pour  $\alpha = 0$ , le spectre de  $M$  est  $\{\theta^2 + \tau_2, 0, 0\}$ . D'où, l'hypothèse  $\rho(M) < 1$  est réduite à  $\theta^2 + \tau_2 < 1$ , qui est une condition suffisante très connue assurant la stationnarité d'un processus RCAR(1).*

## 4.3 Comportement asymptotique de la moyenne empirique et de l'estimateur usuel dans ce modèle

Supposons que nous disposons d'une série chronologique  $(X_t)$  générée par (4.1)–(4.2) et observable sur l'intervalle  $t \in \{0, \dots, n\}$ , où  $n \geq 1$ . Nous supposons de plus que  $X_0$  est distribuée selon la loi strictement stationnaire et ergodique du processus.

**Remarque 4.3.1** *L'hypothèse selon laquelle  $X_0$  est distribuée selon la loi strictement stationnaire et ergodique du processus est une hypothèse permettant de simplifier les calculs et d'invoquer directement le Théorème ergodique. Pour être précis, supposons que la série  $(Y_t)$  est générée via la même relation récursive avec une condition initiale  $Y_0$ . Alors pour tout  $t \geq 1$ ,*

$$X_t - Y_t = (X_0 - Y_0) \prod_{\ell=1}^t (\theta + \alpha \eta_{\ell-1} + \eta_\ell).$$

Posons  $\kappa = \mathbb{E}(\ln |\theta + \alpha\eta_0 + \eta_1|) < 0$ , il est montré pour  $t$  assez grand (voir Section 4.8.1 pour plus de détails) que, presque sûrement,

$$|X_t - Y_t| \leq |X_0 - Y_0| \exp\left(\frac{\kappa t}{2}\right).$$

D'où  $Y_0$  peut être une variable aléatoire quelconque vérifiant  $|X_0 - Y_0| < \infty$  p.s. et s'intégrant au moins au même ordre que  $X_0$ .

Notons la moyenne empirique de notre échantillon

$$\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t. \quad (4.12)$$

Alors, nous avons le résultat suivant, où la variance asymptotique  $\kappa^2$  est explicitée dans (4.64).

**Théorème 4.3.** *Supposons que  $(H_1)$ – $(H_2)$  sont vérifiées. Alors lorsque  $n$  tend vers l'infini, nous avons la convergence presque sûre*

$$\bar{X}_n \rightarrow 0. \quad (4.13)$$

*Supposons de plus que  $(H_3)$  est vérifiée, alors, nous avons la normalité asymptotique*

$$\sqrt{n} \bar{X}_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \kappa^2). \quad (4.14)$$

**Preuve.** Voir Section 4.8.4. □

**Remarque 4.3.2** *Pour  $\alpha = 0$ , nos calculs donnent*

$$\kappa_0^2 = \frac{\sigma_2(1 - \theta^2)}{(1 - \theta)^2(1 - \theta^2 - \tau_2)}. \quad (4.15)$$

*Si de plus  $\tau_2 = 0$ , nous obtenons*

$$\kappa_{00}^2 = \frac{\sigma_2}{(1 - \theta)^2} \quad (4.16)$$

*qui peut être directement déduit du Théorème 7.1.2 de [20].*

Considérons à présent l'estimateur suivant donnée par

$$\hat{\theta}_n = \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2}. \quad (4.17)$$

Il est important de garder à l'esprit que  $\hat{\theta}_n$  n'est pas l'estimateur des moindres carrés de  $\theta$  dès lors que  $\alpha \neq 0$ . Le choix de cet estimateur résulte de notre objectif : montrer qu'une estimation de  $\theta$  via le LSE dans un modèle RCAR(1) conduit à des conclusions erronées (du fait de la présence de mémoire dans les coefficients). Nous verrons au sein de cette Section

que cet estimateur n'est pas consistant pour  $\theta$  dès lors que  $\alpha \neq 0$ , en explicitant sa valeur limite. Nous démontrons également qu'il est asymptotiquement normal. Posons

$$\theta^* = \frac{\theta}{1 - 2\alpha\tau_2} \quad (4.18)$$

et rappelons que  $2\alpha\tau_2 \neq 1$ . La variance asymptotique  $\omega^2$  intervenant dans le théorème central limite sera construite de manière détaillée dans la Section 4.8.5 et formulée de manière concise dans (4.85).

**Théorème 4.4.** *Supposons que les hypothèses  $(H_1)$ – $(H_3)$  sont vérifiées. Alors lorsque  $n$  tend vers l'infini, nous avons la convergence presque sûre*

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta^*. \quad (4.19)$$

De plus, si  $(H_4)$  est vérifiée, nous avons la normalité asymptotique

$$\sqrt{n} \left( \widehat{\theta}_n - \theta^* \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} (0; \omega^2). \quad (4.20)$$

**Preuve.** Voir Section 4.8.5. □

**Remarque 4.3.3** *Lorsque  $\alpha = 0$ ,  $\theta^* = \theta$  et, il est bien connu que, cet estimateur est consistant envers  $\theta$ . De plus, les coefficients de la matrice  $K$  définis dans (4.111) prennent une forme simplifiée où la majeure partie des coefficients sont nuls à l'exception de  $K_{11} = \sigma_2$  et  $K_{22} = \tau_2$ . De façon analogue, seulement la première colonne de  $M$  et celle de  $H$  sont non nulles. D'où, en posant  $\lambda_0 = \mathbb{E}(X_t^2) = \gamma_X(0)$  et  $\delta_0 = \mathbb{E}(X_t^4)$  comme dans la preuve associée, la variance asymptotique s'écrit*

$$\omega_0^2 = \frac{\sigma_2}{\lambda_0} + \frac{\tau_2 \delta_0}{\lambda_0^2}.$$

*Ce résultat est énoncé dans le Théorème 4.1 de [78], dans le cas d'un processus RCAR(1) classique mais en formulant une hypothèse plus agréable (ils supposent seulement que  $\mathbb{E}(X_t^4) < \infty$  tandis que nous obtenons cette condition à partir des perturbations). Elle s'écrit alors explicitement*

$$\omega_0^2 = \frac{(1 - \theta^2 - \tau_2) (\tau_2 \sigma_4 (\theta^2 + \tau_2 - 1) + \sigma_2^2 (\theta^4 + \tau_4 - 6\tau_2^2 - 1))}{\sigma_2^2 (\theta^4 + \tau_4 + 6\theta^2 \tau_2 - 1)}. \quad (4.21)$$

*Si nous supposons en plus que  $\tau_2 = \tau_4 = 0$ , nous obtenons que*

$$\omega_{00}^2 = 1 - \theta^2 \quad (4.22)$$

*qui est un résultat décrit dans la Proposition 8.10.1 de [20], par exemple.*

**Remarque 4.3.4** *Pour  $\alpha = 0$ , l'ensemble des valeurs propres de  $H$  est  $\{\theta^4 + 6\theta^2\tau_2 + \tau_4, 0, 0, 0, 0\}$ . D'où, l'hypothèse  $\rho(H) < 1$  se résume à la contrainte  $\theta^4 + 6\theta^2\tau_2 + \tau_4 < 1$ , qui peut être assimilée à une condition assurant l'existence du moment d'ordre quatre dans un processus RCAR(1) classique.*

**Théorème 4.5.** *Supposons que les hypothèses  $(H_1)$ – $(H_4)$  sont vérifiées. Alors lorsque  $n$  tend vers l'infini, nous avons les estimation suivantes*

$$\frac{1}{\log n} \sum_{t=1}^n (\hat{\theta}_t - \theta^*)^2 \xrightarrow[n \rightarrow \infty]{p.s.} \omega^2 \quad (4.23)$$

et

$$\limsup_{n \rightarrow +\infty} \frac{n}{2 \log \log n} (\hat{\theta}_n - \theta^*)^2 = \omega^2 \text{ p.s.} \quad (4.24)$$

**Preuve.** Voir Section 4.8.6. □

**Remarque 4.3.5** *Le Théorème précédent nous conduit à la même estimation sur la vitesse de convergence de notre estimateur que dans un modèle stable,*

$$(\hat{\theta}_n - \theta^*)^2 = O\left(\frac{\log \log n}{n}\right) \text{ p.s.} \quad (4.25)$$

**Remarque 4.3.6** *Même si cela est dénué d'intérêt statistique, la même estimation peut être obtenue pour  $\bar{X}_n$ .*

Construisons le processus résiduel, défini pour tout  $1 \leq t \leq n$ , par

$$\hat{\epsilon}_t = X_t - \hat{\theta}_n X_{t-1}. \quad (4.26)$$

L'estimateur usuel de  $\sigma_2$  est défini par

$$\hat{\sigma}_{2,n} = \frac{1}{n} \sum_{t=1}^n \hat{\epsilon}_t^2. \quad (4.27)$$

Posons

$$\sigma_2^* = (1 - (\theta^*)^2) \gamma_X(0). \quad (4.28)$$

**Théorème 4.6.** *Supposons que  $(H_1)$ – $(H_3)$  sont vérifiées. Alors lorsque  $n$  tend vers l'infini, nous avons la convergence presque sûre*

$$\hat{\sigma}_{2,n} \xrightarrow[n \rightarrow \infty]{p.s.} \sigma_2^*. \quad (4.29)$$

**Preuve.** Le développement de  $\hat{\sigma}_{2,n}$  dans (4.27) nous conduit via un argument d'ergodicité à

$$\hat{\sigma}_{2,n} \xrightarrow[n \rightarrow \infty]{p.s.} (1 + (\theta^*)^2) \gamma_X(0) - 2\theta^* \gamma_X(1).$$

La définition de  $\hat{\theta}_n$  dans (4.17) implique pour la même raison  $\gamma_X(1) = \theta^* \gamma_X(0)$ , qui donne  $\sigma_2^*$ . □

**Remarque 4.3.7** *Pour  $\alpha = 0$ , (4.28) devient*

$$\sigma_{2,0}^* = \frac{\sigma_2 (1 - \theta^2)}{1 - \theta^2 - \tau_2}. \quad (4.30)$$

*Dans leur travail, Nicholls et Quinn [78] ont pris en considération le fait que l'estimateur usuel  $\sigma_2$  n'est pas consistant, pour cette raison ils ont suggéré de nouveaux estimateurs que nous présenterons dans la Section suivante. Si à présent  $\tau_2 = 0$ , nous retrouvons la consistance de cet estimateur envers  $\sigma_2$ .*

## 4.4 Un test de détection de mémoire dans les coefficients

Nous allons à présent utiliser une méthode type Yule-Walker envers la fonction d'auto-corrélation à l'ordre deux. En utilisant les notations du Théorème 4.2 et en posant  $\gamma = \alpha \tau_2$ , nous avons

$$\begin{cases} (1 - 2\rho_X^2(1))\theta &= (1 - 2\rho_X(2))\rho_X(1) \\ (1 - 2\rho_X^2(1))\gamma &= \rho_X(2) - \rho_X^2(1). \end{cases}$$

Par ergodicité, une estimation consistante de  $\theta^* = \rho_X(1)$  et  $\vartheta^* = \rho_X(2)$  est obtenue *via*

$$\hat{\theta}_n = \frac{\sum_{t=1}^n X_{t-1}X_t}{\sum_{t=1}^n X_{t-1}^2} \text{ et } \hat{\vartheta}_n = \frac{\sum_{t=2}^n X_{t-2}X_t}{\sum_{t=2}^n X_{t-2}^2} \quad (4.31)$$

respectivement. Définissons une fonction de  $[-1; 1] \setminus \{\pm \frac{1}{\sqrt{2}}\} \times [-1; 1]$  dans  $\mathbb{R}^2$  par

$$f : (x, y) \mapsto \left( \frac{(1 - 2y)x}{1 - 2x^2}, \frac{y - x^2}{1 - 2x^2} \right) \quad (4.32)$$

et le nouveau couple d'estimateurs

$$(\tilde{\theta}_n, \tilde{\gamma}_n) = f(\hat{\theta}_n, \hat{\vartheta}_n). \quad (4.33)$$

Pour être cohérent vis à vis de (4.32), nous supposons dans la suite que  $\sqrt{2}\theta \neq \pm(1 - 2\alpha\tau_2)$ . Nous supposons également que  $\psi_0^0 \neq 0$ , où  $\psi_0^0$  sera décrit ultérieurement. Ces expressions étant extrêmement lourdes, nous ne donnons pas d'expression analytique dans les hypothèses, nous considérons plus simplement  $\Theta^* = \{\sqrt{2}\theta = \pm(1 - 2\alpha\tau_2)\} \cup \{\psi_0^0 = 0\}$  l'ensemble contentant les cas pathologiques et choisissons nos paramètres sur le complémentaire de  $\Theta^*$  pour achever notre étude. Il vient aisément que  $\tilde{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta$  et  $\tilde{\gamma}_n \xrightarrow[n \rightarrow \infty]{p.s.} \gamma$ . Dans le Théorème suivant, nous établissons la normalité asymptotique jointe des ces estimateurs, qui sera utile pour construire un test. Nous notons  $\nabla f$  la matrice Jacobienne de  $f$ .

**Théorème 4.7.** *Supposons que  $(H_1)$ – $(H_4)$  sont vérifiées. Alors lorsque  $n$  tend vers l'infini, nous avons la normalité asymptotique*

$$\sqrt{n} \begin{pmatrix} \tilde{\theta}_n - \theta \\ \tilde{\gamma}_n - \gamma \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Psi) \quad (4.34)$$

où  $\Sigma$  est la matrice de covariance décrite dans (4.102) et

$$\Psi = \nabla^T f(\theta^*, \vartheta^*) \Sigma \nabla f(\theta^*, \vartheta^*). \quad (4.35)$$

**Preuve.** Voir Section 4.8.7. □

Supposons que les coefficients sont aléatoires (i.e.,  $\tau_2 > 0$ ), d'où  $\gamma = 0 \Leftrightarrow \alpha = 0$ . Notre dernier objectif est de construire une procédure statistique pour discriminer les hypothèses

$$\mathcal{H}_0 : “\alpha = 0” \text{ vs } \mathcal{H}_1 : “\alpha \neq 0”. \quad (4.36)$$

Comme cela est mentionné dans la Remarque 4.8.1, en dépit de sa structure complexe,  $\Psi$  dépend uniquement des paramètres. Soit  $\psi = \psi(\theta, \alpha, \{\tau_k\}_{2,4,6,8}, \{\sigma_\ell\}_{2,4})$  le coefficient en bas

à droite de  $\Psi$ , et  $\psi^0 = \psi(\theta, 0, \{\tau_k\}_{2,4,6,8}, \{\sigma_\ell\}_{2,4})$ . Les calculs sont plus agréables sous  $\mathcal{H}_0$  et donnent  $\theta^* = \theta$ ,  $\vartheta^* = \theta^2$  ainsi que

$$\psi^0 = \frac{\psi_0^0}{(1 - 2\theta^2)^2 \sigma_2^2 (\theta^4 + 6\theta^2 \tau_2 + \tau_4 - 1)} \quad (4.37)$$

où le numérateur est donné par

$$\begin{aligned} \psi_0^0 = & (\tau_2 + \theta^2 - 1) [\sigma_4 \tau_2 ((6\theta^2 - 1) \tau_2^2 + (8\theta^4 - 9\theta^2 + 1) \tau_2 \\ & + 2\theta^2 (\theta^2 - 1)^2) + \sigma_2^2 \tau_2 (-36 \tau_2^2 \theta^2 + 6 \tau_2^2 - 12 \tau_2 \theta^4 \\ & + 12 \tau_2 \theta^2 - 6 \theta^6 + 17 \theta^4 + 6 \tau_4 \theta^2 - 12 \theta^2 - \tau_4 + 1) \\ & + \sigma_2^2 (\theta^6 - \theta^4 + \theta^2 \tau_4 - \theta^2 - \tau_4 + 1)] \end{aligned}$$

et est supposé non nul (en excluant  $\Theta^*$ ). Comme Corollaire,  $\psi^0$  dépend de manière continue des paramètres (voir Remarque 4.3.4). Supposons de plus que  $(H_5)$  est vérifiée, d'où  $\psi^0 = \psi^0(\theta, \tau_2, \sigma_2)$ , et considérons

$$\widehat{\psi}_n^0 = \psi^0(\bar{\theta}_n, \bar{\tau}_{2,n}, \bar{\sigma}_{2,n})$$

où  $\bar{\theta}_n$  est soit  $\widehat{\theta}_n$  soit  $\widetilde{\theta}_n$ , et  $(\bar{\tau}_{2,n}, \bar{\sigma}_{2,n})$  est le couple d'estimateurs suggéré par [78] dans les formules (3.6) et (3.7) respectivement, également donné dans [62]. Ils sont définis par

$$\bar{\tau}_{2,n} = \frac{\sum_{t=1}^n (Z_t - \bar{Z}_n) \widehat{\epsilon}_t^2}{\sum_{t=1}^n (Z_t - \bar{Z}_n)^2} \text{ et } \bar{\sigma}_{2,n} = \widehat{\sigma}_{2,n} - \bar{Z}_n \bar{\tau}_{2,n} \quad (4.38)$$

où  $(\widehat{\epsilon}_t)$  est le processus résiduel tel que décrit dans (4.26),  $\widehat{\sigma}_{2,n}$  est donné dans (4.27) et pour  $t \in \{1, \dots, n\}$ ,  $Z_t = X_t^2$ . Le Théorème 4.2 de [78] garantit la consistance de ces derniers dès lors que notre processus RCAR(1) possède un moment d'ordre quatre. Notre étude garantit en plus la consistance de  $\bar{\theta}_n$  sous  $\mathcal{H}_0$ . Nous déduisons du Lemme de Slutsky que

$$\widehat{\psi}_n^0 \xrightarrow[n \rightarrow \infty]{p.s.} \psi^0 > 0 \text{ et } \frac{n (\widetilde{\gamma}_n)^2}{\widehat{\psi}_n^0} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_1^2 \quad (4.39)$$

si  $\mathcal{H}_0$  est valide, où  $\chi_1^2$  est une variable aléatoire distribuée selon la loi du  $\chi^2$  à un degré de liberté, tandis que sous  $\mathcal{H}_1$  le test statistique diverge (presque sûrement). L'introduction de  $(H_5)$  nous permet de considérer

$$\bar{\sigma}_{4,n} = g(\bar{\sigma}_{2,n}) \text{ et } \bar{\tau}_{4,n} = h(\bar{\tau}_{2,n})$$

comme estimateurs consistants des moments précités. La comparaison de notre test statistique aux quantiles d'un  $\chi_1^2$  constitue donc une procédure standard afin de détecter de la mémoire dans les coefficients de notre processus autorégressif.

## 4.5 Une généralisation vectorielle

Les résultats présentés dans cette section figurent également dans un travail [83] mené avec Frédéric Proïa que je tiens à remercier.

L'objectif de cette courte section est de présenter un modèle autorégressif à l'ordre  $p$  dont les coefficients sont aléatoires et corrélés généralisant ainsi 4.1. Considérons alors un processus  $(Y_t)$  satisfaisant la relation récursive

$$Y_t = \sum_{k=1}^p (\theta_k + \alpha_k \eta_{k,t-1} + \eta_{k,t}) Y_{t-k} + \epsilon_t \quad (4.40)$$

où  $(\epsilon_t)$  et  $(\eta_{k,t})$  sont des bruits blancs forts mutuellement indépendants avec des moments d'ordre impair nulle. Il est plus agréable de réécrire ce modèle sous forme vectorielle

$$\Phi_t = (C_\theta + N_{t-1} D_\alpha + N_t) + E_t \quad (4.41)$$

où  $\Phi_t^T = (Y_t, Y_{t-1}, \dots, Y_{t-p+1})$ ,  $E_t^T = (\epsilon_t, 0, \dots, 0)$ ,  $N_t$  est une matrice carrée de taille  $p$  avec  $\eta_t = (\eta_{1,t}, \dots, \eta_{p,t})$  comme première ligne et 0 ailleurs,  $D_\alpha = \text{diag}(\alpha_1, \dots, \alpha_p)$  et  $C_\theta$  est la matrice compagnon, c'est à dire

$$C_\theta = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \theta_p \\ & I_{p-1} & & 0 \end{pmatrix}.$$

Nous considérons deux espaces paramétriques suivants que nous utiliserons comme condition dans nos résultats. Ces derniers portent sur une restriction des paramètres  $\theta, \alpha, \sigma_q$  et  $\eta_{k,q}$ . Posons

$$\Theta_2 = \{ \{ \theta, \alpha, \sigma_q, \eta_{k,q} \} \mid \mathbb{E}(\eta_{k,t}^a Y_t^2) < +\infty \text{ pour tout } a \in \{0, 1, 2\} \}$$

et

$$\Theta_4 = \{ \{ \theta, \alpha, \sigma_q, \eta_{k,q} \} \mid \mathbb{E}(\eta_{k,t}^a Y_t^4) < +\infty \text{ pour tout } a \in \{0, 1, 2, 3, 4\} \}.$$

Nous formulons à présent une première condition sur la distribution des perturbations qui est

$$\mathbb{E}(\log \|C_\theta + N_0 D_\alpha + N_1\|) < 0 \text{ et } \mathbb{E}(\log^+ |\varepsilon_0|) < \infty. \quad (4.42)$$

Considérons la filtration

$$\mathcal{F}_t^{(1)} = \sigma((\varepsilon_s, \eta_{1,s}, \dots, \eta_{p,s}), s \leq t). \quad (4.43)$$

Nous avons un premier résultat au sujet de l'existence d'une solution strictement stationnaire et ergodique

**Proposition 4.8.** *Supposons que la condition (4.42) est vérifiée, alors pour tout  $t \in \mathbb{Z}$ ,*

$$\Phi_t = E_t + \sum_{k=1}^{\infty} E_{t-k} \prod_{\ell=0}^{k-1} (C_\theta + N_{t-\ell-1} D_\alpha + N_{t-\ell}) \text{ p.s.} \quad (4.44)$$

*En conséquence,  $(Y_t)$  est strictement stationnaire, ergodique et  $\mathcal{F}_t^{(1)}$ -mesurable.*

**Preuve.** Voir Section 4.8.8. □

Supposons à présent que nous disposons d'une trajectoire observable  $(Y_{-p+1}, \dots, Y_n)$  où  $\Phi_0$  suit la loi strictement stationnaire et ergodique du processus. L'estimateur des moindres carrés dans la cas vectoriel s'écrit donc sous nos notations

$$\widehat{\theta}_n = S_n^{-1} \sum_{t=1}^n \Phi_{t-1} Y_t \text{ où } S_n = \sum_{t=1}^n \Phi_t \Phi_t^T. \quad (4.45)$$

Posons à présent  $\bar{\ell}_i = \mathbb{E}(Y_i Y_0)$  puis convenons des notations suivantes

$$\bar{\Lambda}_0 = \begin{pmatrix} \bar{\ell}_0 & \bar{\ell}_1 & \dots & \bar{\ell}_{p-1} \\ \bar{\ell}_1 & \bar{\ell}_0 & \dots & \bar{\ell}_{p-2} \\ \dots & \dots & \dots & \dots \\ \bar{\ell}_{p-1} & \bar{\ell}_{p-2} & \dots & \bar{\ell}_0 \end{pmatrix} \text{ et } \bar{L}_1 = \begin{pmatrix} \bar{\ell}_1 \\ \bar{\ell}_2 \\ \dots \\ \bar{\ell}_p \end{pmatrix} \quad (4.46)$$

**Théorème 4.9.** *Supposons que la condition (4.42) est vérifiée et que nous restreignons nos paramètres à l'espace paramétrique  $\Theta = \Theta_2 \setminus \Theta^*$ . Alors, nous avons la convergence presque sûre*

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta^* = \bar{\Lambda}_0^{-1} \bar{L}_1.$$

**Théorème 4.10.** *Supposons que la condition (4.42) est vérifiée et que nous restreignons nos paramètres à l'espace paramétrique  $\Theta = \Theta_4 \setminus \Theta^*$ . Alors, nous avons la normalité asymptotique*

$$\sqrt{n} \left( \widehat{\theta}_n - \theta^* \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( 0; \bar{\Lambda}_0^{-1} \bar{L} \bar{\Lambda}_0^{-1} \right)$$

où  $\bar{L}$  est une matrice symétrique que nous expliciterons au cours de la preuve.

**Preuve.** Voir Section 4.8.9. □

**Théorème 4.11.** *Supposons que la condition (4.42) est vérifiée et que nous restreignons nos paramètres à l'espace paramétrique  $\Theta = \Theta_4 \setminus \Theta^*$ . Alors, nous avons l'estimation suivante sur la vitesse de convergence*

$$\limsup_{n \rightarrow \infty} \frac{n}{2 \log \log n} \|\widehat{\theta}_n - \theta^*\|^2 < \infty \text{ p.s.}$$

où  $\bar{L}$  est une matrice symétrique que nous expliciterons au cours de la preuve.

**Preuve.** Voir Section 4.8.10. □

Ce dernier théorème nous conduit à l'estimation usuelle sur la vitesse de convergence de l'estimateur

$$\|\widehat{\theta}_n - \theta^*\| = O \left( \sqrt{\frac{\log \log n}{n}} \right) \text{ p.s.}$$

## 4.6 Conclusion et Perspectives

Pour conclure nous venons de montrer que, dans un modèle assez simple, l'introduction d'une corrélation dans les coefficients impacte grandement la procédure d'inférence. Il est, dans le cadre d'une modélisation chronologique, pertinent de prendre en compte une structure d'autocorrélation dans les coefficients qui incite donc à étudier dans ce cadre le comportement des indicateurs usuels et à corriger éventuellement les estimations. Il serait intéressant dans un premier temps d'obtenir des résultats similaires au cas scalaire dans le cas vectoriel qui comporte des lacunes. Il serait intéressant dans un second temps d'étendre cette étude à un processus où la mémoire dans les coefficients serait plus longue et/ou l'ordre d'autorégression serait plus élevé. Nous pourrions également étudier le comportement de ces estimateurs lorsque le modèle est instable voire explosif.

## 4.7 Simulations

Nous commençons par illustrer numériquement le comportement de l'estimateur des moindres carrés dans différents cas. Nous simulons  $M = 10000$  échantillons d'erreurs statistiques  $\sqrt{b}(\hat{\theta}_n - \theta)$  où  $N = 500$  afin de mettre exergue le décalage entre l'estimation obtenue grâce à l'estimateur usuel et le bon paramètre  $\theta$  lors de la présence de mémoire dans les coefficients.

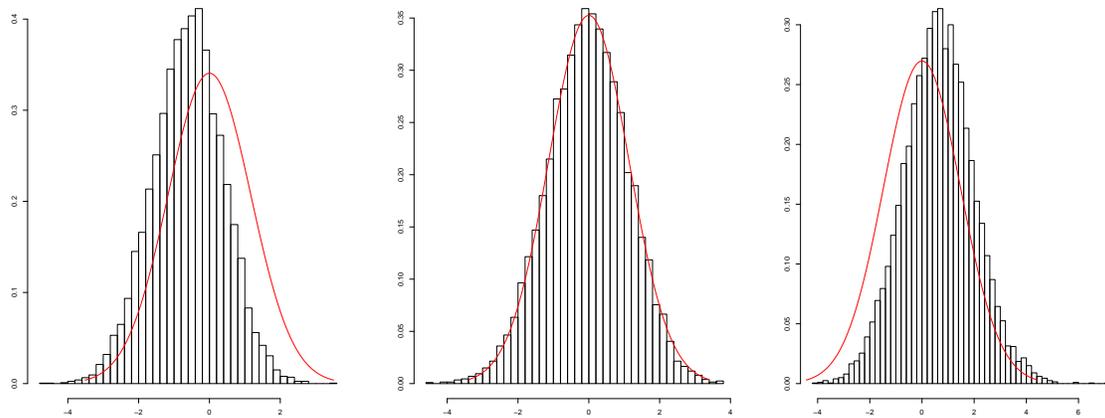


FIGURE 4.1 – Comportement de l'erreur statistique normalisée entre l'estimateur des moindres carrés et  $\theta$  pour différentes valeurs de  $\alpha$ .  $\alpha = -0,3$  à gauche,  $0$  au centre et  $0,3$  à droite.

Nous illustrons ensuite la normalité asymptotique des estimateurs de  $\theta$  et  $\alpha\tau_2$  obtenus via l'approche Yule-Walker développée à la section 4.4 en simulant 10000 échantillons d'erreurs statistique de taille 500. Dans ces simulations,  $\theta = 0.3$ ,  $\alpha = 0.2$ ,  $\tau_2 = 0.25$  et  $\sigma_2 = 1$ .

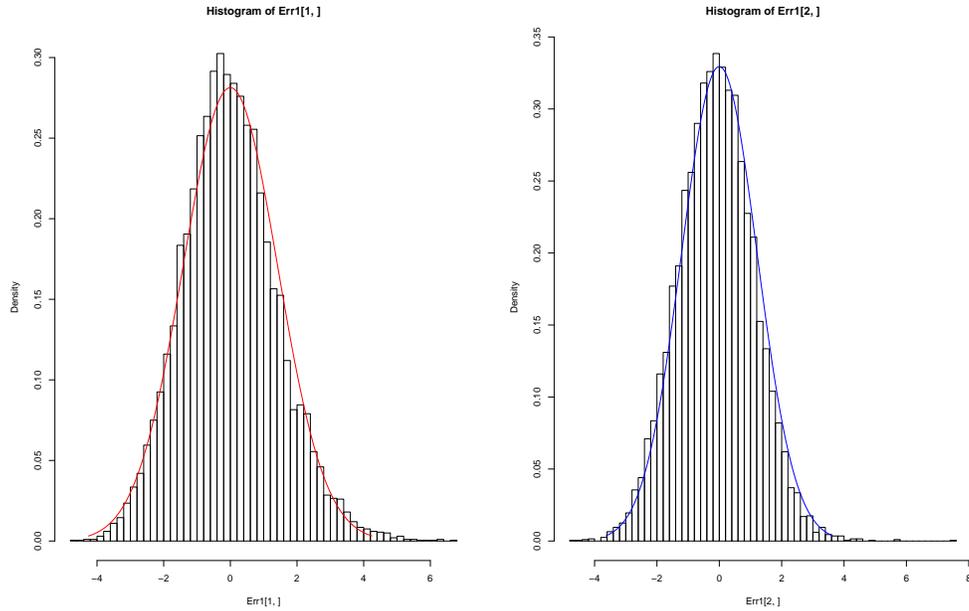


FIGURE 4.2 – Comportement de l’erreur statistique  $\sqrt{n}(\tilde{\theta} - \theta)$  ( $\sqrt{n}(\tilde{\gamma}_n - \gamma)$ ) à gauche (à droite).

Pour terminer les simulations, nous illustrons la puissance empirique du test permettant de détecter la présence d’une mémoire (i.e.  $\alpha \neq 0$ ) dans les coefficients. Nous fixons  $\theta = 0.4$ ,  $\tau_2 = 0.36$ ,  $\sigma_2 = 1$  puis nous simulons 10000 trajectoires de taille 500 en faisant varier le paramètre  $\alpha$  dans l’intervalle  $[-0.5; 0.5]$  avec un pas de 0.01. Nous fixons le risque de première espèce à 0.05 puis nous calculons la fréquence de rejet pour chaque valeur de  $\alpha$ .

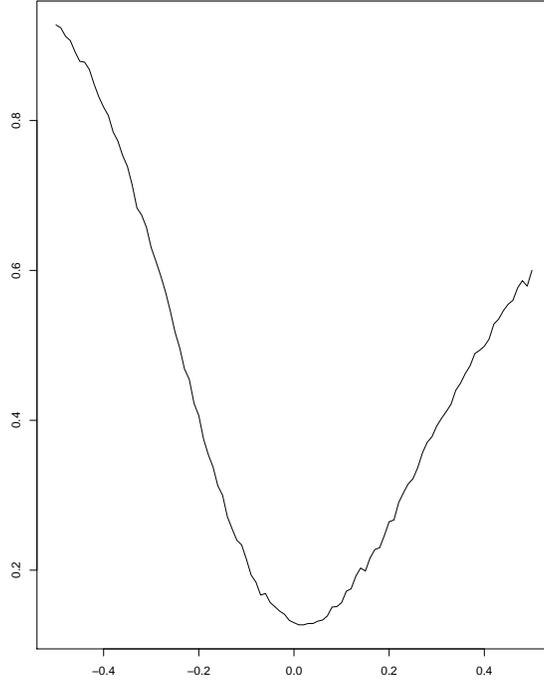


FIGURE 4.3 – Puissance empirique du test de mémoire dans les coefficients pour 10000 trajectoires de taille 500.

## 4.8 Preuve des résultats principaux

Cette section est dévolue aux démonstrations de l'ensemble des résultats présentés dans ce chapitre. Les outils fondamentaux liés à l'ergodicité sont donnés dans le Théorème 3.5.8 de [94] ou bien dans le Théorème 1.3.3 de [97]. Nous aurons à faire de manière répétitive à des espérances de la forme  $\mathbb{E}(\eta_t^a(\theta + \eta_t)^b)$  for  $a, b \in \{0, \dots, 4\}$ , nous donnons donc pour plus de clarté la valeur de ces dernières sous  $(H_2)$  au sein de la Table 4.1 ci-contre. Pour alléger les

$a \setminus b$	0	1	2	3	4
0	1	$\theta$	$\theta^2 + \tau_2$	$\theta^3 + 3\theta\tau_2$	$\theta^4 + 6\theta^2\tau_2 + \tau_4$
1	0	$\tau_2$	$2\theta\tau_2$	$3\theta^2\tau_2 + \tau_4$	$4\theta^3\tau_2 + 4\theta\tau_4$
2	$\tau_2$	$\theta\tau_2$	$\theta^2\tau_2 + \tau_4$	$\theta^3\tau_2 + 3\theta\tau_4$	$\theta^4\tau_2 + 6\theta^2\tau_4 + \tau_6$
3	0	$\tau_4$	$2\theta\tau_4$	$3\theta^2\tau_4 + \tau_6$	$4\theta^3\tau_4 + 4\theta\tau_6$
4	$\tau_4$	$\theta\tau_4$	$\theta^2\tau_4 + \tau_6$	$\theta^3\tau_4 + 3\theta\tau_6$	$\theta^4\tau_4 + 6\theta^2\tau_6 + \tau_8$

TABLE 4.1 –  $\mathbb{E}(\eta_t^a(\theta + \eta_t)^b)$  for  $a, b \in \{0, \dots, 4\}$ .

preuves, nous formulons les nombreuses constantes que nous utiliserons par la suite au sein

d'une appendice . Nous commençons par donner quelques détails aux sujets des Remarques 4.2 et 4.3.

### 4.8.1 Au sujet des Remarques des Sections 4.2 et 4.3

#### Remarque 4.2.1

Le calcul de  $\gamma_X(0)$  à partir de (4.10) conduit à

$$\gamma_X(0) = \frac{\sigma_2 (2\alpha\tau_2 - 1)}{d(\theta, \alpha, \tau_2, \tau_4)}$$

pour un dénominateur vérifiant  $d(\theta, \alpha, \tau_2, \tau_4) = 2\theta^2$  lorsque  $2\alpha\tau_2 = 1$ . Il en découle que même si cette égalité est vérifiée sous réserve que le processus soit faiblement stationnaire, il sera alors déterministe.

#### Remarque 4.3.1

Il s'agit de montrer que la différence entre le processus démarrant à  $X_0$  étant distribuée sous la loi strictement stationnaire et ergodique et le processus démarrant à  $Y_0$  quelconque est (p.s.) négligeable sous des conditions faibles portant sur  $Y_0$ . En s'inspirant du Lemme 1 de [5] et en utilisant le Théorème ergodique, nous obtenons que pour  $t$  assez grand, presque sûrement

$$\frac{1}{t} \sum_{\ell=1}^t \ln |\theta + \alpha \eta_{\ell-1} + \eta_\ell| \leq \frac{\kappa}{2} < 0.$$

D'où le produit  $\prod_{\ell=1}^t |\theta + \alpha \eta_{\ell-1} + \eta_\ell|$  décroît à vitesse géométrique par rapport à  $t$  sous  $(H_1)$  et nous avons l'estimation suivante  $|X_t - Y_t| \leq |X_0 - Y_0| \exp\left(\frac{\kappa t}{2}\right)$  nous permettant de retenir des conditions faibles sur  $Y_0$  pour finalement en déduire  $X_t - Y_t = o(1)$  p.s.

#### Remarque 4.3.2

Dans le cas particulier ou  $\alpha = \tau_2 = 0$  (ce qui est, ni plus ni moins qu'un processus AR(1)), Le Théorème 7.1.2 de [20] stipule que  $\sqrt{n} \bar{X}_n$  est asymptotiquement normal de moyenne 0 et de variance

$$\sum_{h \in \mathbb{Z}} \gamma_X(h) = \sigma_2 \left( \sum_{k=0}^{+\infty} \theta^k \right)^2 = \frac{\sigma_2}{(1-\theta)^2}.$$

D'où,  $\kappa_{00}^2$  obtenu est cohérent de ce point de vue.

#### Remarque 4.3.3

Comme dans la Remarque précédente, la Proposition 8.10.1 de [20] stipule que, lorsque  $\alpha = \tau_2 = 0$ , le LSE de  $\theta$  est asymptotiquement normal à vitesse  $\sqrt{n}$ , moyenne nulle et variance donnée par  $1 - \theta^2$ , qui correspond à  $\omega_{00}^2$ . Maintenant si  $\tau_2 > 0$ , le Théorème 4.1 de [78], et plus particulièrement la formule (4.1), exprime la variance asymptotique comme fonction de  $\mathbb{E}(X_t^2)$  et  $\mathbb{E}(X_t^4)$  comme mentionné dans la Remarque 4.3.3. Notre étude permet d'explicitier  $\omega_0^2$  comme fonction des paramètres en injectant  $\alpha = 0$  dans  $\lambda_0$  et  $\delta_0$  qui sont calculés dans (4.56) et (4.76), respectivement.

### 4.8.2 Preuve du Théorème 4.1

L'existence d'une solution strictement stationnaire, causale, qui coïncide avec  $(X_t)$  p.s. sous  $(H_1)$  résulte du Corollaire du Théorème 1 de [18]. De plus,  $(\theta_t)$  est un processus MA(1) strictement stationnaire et ergodique indépendant de  $(\epsilon_t)$ , lui même strictement stationnaire et ergodique. Nous donnons maintenant plus de détails. Premièrement, l'hypothèse  $(H_1)$  nous permet d'employer une preuve similaire à [5] où le Théorème ergodique remplace la loi forte de grands nombres invoquée pour obtenir la formule (6), et de justifier que (4.7) est la limite d'une série convergente (avec probabilité 1). Alors pour tout  $t \in \mathbb{Z}$ ,

$$\begin{aligned} \theta_t X_{t-1} &= (\theta + \alpha \eta_{t-1} + \eta_t) \left[ \epsilon_{t-1} + \sum_{k=1}^{\infty} \epsilon_{t-k-1} \prod_{\ell=0}^{k-1} (\theta + \alpha \eta_{t-\ell-2} + \eta_{t-\ell-1}) \right] \\ &= \sum_{k=1}^{\infty} \epsilon_{t-k} \prod_{\ell=0}^{k-1} (\theta + \alpha \eta_{t-\ell-1} + \eta_{t-\ell}) = X_t - \epsilon_t \end{aligned}$$

ce qui implique que (4.7) est une solution de l'équation récursive. Finalement, la stationnarité stricte ainsi que l'ergodicité de  $(X_t)$  peuvent être obtenues en suivant le même raisonnement que [78]. La représentation causale (4.7) montre qu'il existe  $\phi$  indépendant de  $t$  tel que pour tout  $t \in \mathbb{Z}$ ,

$$X_t = \phi((\epsilon_t, \eta_t), (\epsilon_{t-1}, \eta_{t-1}), \dots).$$

L'ensemble  $((\epsilon_t, \eta_t), (\epsilon_{t-1}, \eta_{t-1}), \dots)$  est constitué de variables aléatoires i.i.d.,  $(X_t)$  est donc strictement stationnaire. L'ergodicité découle du Théorème 1.3.3 de [97].

### 4.8.3 Preuve du Théorème 4.2

L'ergodicité et la stationnarité stricte proviennent du Théorème 4.1. Nous considérons la représentation causale (4.7). Premièrement, puisque  $(\epsilon_t)$  et  $(\eta_t)$  sont des bruits blancs indépendants, pour tout  $t \in \mathbb{Z}$ ,

$$\mathbb{E}(X_t) = 0. \quad (4.47)$$

Pour expliciter la fonction d'autocovariance de  $(X_t)$ , nous devons dans un premier temps établir un Lemme technique traitant des propriétés du second ordre du processus. Pour tout  $k, h \in \mathbb{N}^*$ , considérons la suite

$$\begin{aligned} u_{0,h}^{(a)} &= \mathbb{E}(\eta_h^a \theta_h \dots \theta_1), \\ u_{k,0}^{(a)} &= \mathbb{E}(\eta_k^a \theta_k^2 \dots \theta_1^2), \\ u_{k,h}^{(a)} &= \mathbb{E}(\eta_{k+h}^a \theta_{k+h} \dots \theta_{k+1} \theta_k^2 \dots \theta_1^2), \end{aligned}$$

où  $a \in \{0, 1, 2\}$ , et construisons

$$U_{k,h} = \begin{pmatrix} u_{k,h}^{(0)} \\ u_{k,h}^{(1)} \\ u_{k,h}^{(2)} \end{pmatrix}. \quad (4.48)$$

Dans la suite,  $M$ ,  $N$  et  $U_0$  correspondent à (4.4), (4.9) et (4.3), respectivement.

**Lemme 4.12.** *Supposons que  $(H_1)$ – $(H_3)$  sont vérifiées. Alors, pour tout  $h, k \in \mathbb{N}$ ,*

$$U_{k,h} = N^h M^k U_0 \quad (4.49)$$

avec la convention  $U_{0,0} = U_0$ .

**Preuve.** Dans tout le preuve,  $(\mathcal{F}_t)$  est la filtration définie dans (4.8) et la Table 4.1 est utilisée afin d'explicitier les coefficients apparaissant dans les calculs. Les coefficients  $\theta_{k+h-1}$ ,  $\theta_{k+h-2}$ ,  $\dots$  sont  $\mathcal{F}_{k+h-1}$ -mesurables. D'où pour  $h \geq 1$ ,

$$\begin{aligned} u_{k,h}^{(0)} &= \mathbb{E}(\theta_{k+h-1} \dots \theta_{k+1} \theta_k^2 \dots \theta_1^2 \mathbb{E}(\theta_{k+h} | \mathcal{F}_{k+h-1})) \\ &= \theta u_{k,h-1}^{(0)} + \alpha u_{k,h-1}^{(1)}, \\ u_{k,h}^{(1)} &= \mathbb{E}(\theta_{k+h-1} \dots \theta_{k+1} \theta_k^2 \dots \theta_1^2 \mathbb{E}(\eta_{k+h} \theta_{k+h} | \mathcal{F}_{k+h-1})) \\ &= \tau_2 u_{k,h-1}^{(0)}, \\ u_{k,h}^{(2)} &= \mathbb{E}(\theta_{k+h-1} \dots \theta_{k+1} \theta_k^2 \dots \theta_1^2 \mathbb{E}(\eta_{k+h}^2 \theta_{k+h} | \mathcal{F}_{k+h-1})) \\ &= \theta \tau_2 u_{k,h-1}^{(0)} + \alpha \tau_2 u_{k,h-1}^{(1)}. \end{aligned}$$

Les dernières égalités se reformulent sous forme matricielle de manière compacte comme suit  $U_{k,h} = N U_{k,h-1}$ . Il vient que, pour  $h \in \mathbb{N}$ ,

$$U_{k,h} = N^h U_{k,0}. \quad (4.50)$$

La dernière étape consiste à calculer  $U_{k,0}$ , et nous utiliserons le même cheminement. Pour  $k \geq 1$ ,

$$\begin{aligned} u_{k,0}^{(0)} &= \mathbb{E}(\theta_{k-1}^2 \dots \theta_1^2 \mathbb{E}(\theta_k^2 | \mathcal{F}_{k-1})) \\ &= (\theta^2 + \tau_2) u_{k-1,0}^{(0)} + 2\alpha \theta u_{k-1,0}^{(1)} + \alpha^2 u_{k-1,0}^{(2)}, \\ u_{k,0}^{(1)} &= \mathbb{E}(\theta_{k-1}^2 \dots \theta_1^2 \mathbb{E}(\eta_k \theta_k^2 | \mathcal{F}_{k-1})) \\ &= 2\theta \tau_2 u_{k-1,0}^{(0)} + 2\alpha \tau_2 u_{k-1,0}^{(1)}, \\ u_{k,0}^{(2)} &= \mathbb{E}(\theta_{k-1}^2 \dots \theta_1^2 \mathbb{E}(\eta_k^2 \theta_k^2 | \mathcal{F}_{k-1})) \\ &= (\theta^2 \tau_2 + \tau_4) u_{k-1,0}^{(0)} + 2\alpha \theta \tau_2 u_{k-1,0}^{(1)} + \alpha^2 \tau_2 u_{k-1,0}^{(2)}. \end{aligned}$$

D'où, (4.50) se reformule

$$U_{k,h} = N^h M^{k-1} U_{1,0}$$

où le vecteur initial  $U_{1,0}$  est donné par

$$\begin{aligned} u_{1,0}^{(0)} &= \mathbb{E}(\theta_1^2) = (\theta^2 + \tau_2) + \alpha^2 \tau_2, \\ u_{1,0}^{(1)} &= \mathbb{E}(\eta_1 \theta_1^2) = 2\theta \tau_2, \\ u_{1,0}^{(2)} &= \mathbb{E}(\eta_1^2 \theta_1^2) = (\theta^2 \tau_2 + \tau_4) + \alpha^2 \tau_2^2. \end{aligned}$$

Nous déduisons des calculs précédents que, pour tout  $k \in \mathbb{N}^*$  et  $h \in \mathbb{N}$ ,

$$U_{k,h} = N^h M^k U_0.$$

Pour  $k = 0$ , un calcul similaire partant de la valeur initiale  $u_{0,h}^{(a)}$  pour  $a \in \{0, 1, 2\}$  conduit à  $U_{0,h} = N^h U_0$ , ce qui implique que (4.49) est valide pour tout  $k, h \in \mathbb{N}$ .  $\square$

**Corollaire 4.13.** *Supposons que  $(H_1)$ – $(H_3)$  sont vérifiées. Alors, les propriétés du second ordre impliquant  $(X_t)$  sont telles que, pour tout  $a \in \{0, 1, 2\}$ ,*

$$\mathbb{E}(\eta_t^a X_t^2) < \infty.$$

**Preuve.** Pour tout  $t \in \mathbb{Z}$  et  $k \geq 1$ , posons

$$\Lambda_t = \begin{pmatrix} 1 \\ \eta_t \\ \eta_t^2 \end{pmatrix} \quad P_{t,k} = \prod_{i=0}^{k-1} \theta_{t-i} \quad (4.51) \quad \square$$

avec  $P_{t,0} = 1$ . Puisque  $(\epsilon_t)$  et  $(\eta_t)$  sont des bruits blancs, en utilisant la représentation causale (4.7) et en posant  $h = 0$ ,

$$\mathbb{E}(\Lambda_t X_t^2) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \mathbb{E}(\Lambda_t P_{t,k} P_{t,\ell} \epsilon_{t-k} \epsilon_{t-\ell}) = \sigma_2 \sum_{k=0}^{\infty} M^k U_0 = \sigma_2 (I_3 - M)^{-1} U_0$$

comme conséquence de la stationnarité stricte de  $(\theta_t)$ . Nous rappelons que, sous  $(H_3)$ , il est démontré (voir par exemple [55]) que  $I_3 - M$  est inversible et

$$\sum_{k=0}^{\infty} M^k = (I_3 - M)^{-1}.$$

Revenons à la preuve du Théorème 4.2. En vertu du Lemme 4.12 et du Corollaire 4.13, nous sommes en position d'évaluer la fonction d'autocovariance de  $(X_t)$ . Pour  $h \in \mathbb{N}$ ,

$$\mathbb{C}ov(X_t, X_{t-h}) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \mathbb{E}(P_{t,k} P_{t-h,\ell} \epsilon_{t-k} \epsilon_{t-h-\ell}).$$

Nous obtenons ainsi

$$\gamma_X(h) = \sigma_2 \sum_{k=0}^{\infty} \mathbb{E}(P_{t,k+h} P_{t-h,k}) = \sigma_2 \left( \mathbb{E}(P_{t,h}) + \sum_{k=1}^{\infty} u_{k,h}^{(0)} \right) = \sigma_2 \left[ \sum_{k=0}^{\infty} U_{k,h} \right]_1.$$

De part le Lemme 4.12,

$$\gamma_X(h) = \sigma_2 [N^h (I_3 - M)^{-1} U_0]_1.$$

Il ne reste plus qu'à conclure en remarquant que  $\gamma_X$  ne dépend pas de  $t$ . Pour tout  $t \in \mathbb{Z}$  et  $h \in \mathbb{N}$ ,  $\gamma_X(h) = \mathbb{C}ov(X_{t-h}, X_t) = \mathbb{C}ov(X_t, X_{t+h})$ , qui montre que le même raisonnement reste valide pour  $h \in \mathbb{Z}$ , en remplaçant  $h$  par  $|h|$ . Supposons à présent que nous disposons d'une autre solution causale, strictement stationnaire, ergodique  $(W_t)$  possédant un moment d'ordre deux. Il existe alors  $\varphi$  indépendant de  $t$  tel que pour tout  $t \in \mathbb{Z}$ ,

$$X_t - W_t = \varphi((\epsilon_t, \eta_t), (\epsilon_{t-1}, \eta_{t-1}), \dots)$$

et nécessairement,  $(X_t - W_t)$  est un également un processus strictement stationnaire possédant un moment d'ordre deux. Soit  $e^{(a)} = \mathbb{E}(\eta_t^a (X_t - W_t)^2)$ , pour  $a \in \{0, 1, 2\}$ . Via les mêmes calculs et en utilisant la stationnarité faible de  $(X_t - W_t)$ , il vient que

$$\begin{pmatrix} e^{(0)} \\ e^{(1)} \\ e^{(2)} \end{pmatrix} = M \begin{pmatrix} e^{(0)} \\ e^{(1)} \\ e^{(2)} \end{pmatrix}$$

impliquant, si  $(e^{(0)} e^{(1)} e^{(2)}) \neq 0$ , que 1 est une valeur propre de  $M$ . Cela contredit  $\rho(M) < 1$  faisant partie de  $(H_3)$ . D'où,  $\mathbb{E}((X_t - W_t)^2)$  vaut zero et  $X_t = W_t$  p.s.

#### 4.8.4 Preuve du Théorème 4.3

La converge vers zero est une application directe du Théorème ergodique, puisque nous avons montré dans (4.47) que  $\mathbb{E}(X_t) = 0$ . Dans cette preuve, uniquement  $(H_1)$  et  $(H_2)$  sont nécessaires. Convenons de la notation suivante,

$$\begin{aligned}\bar{M}_n^{(1)} &= \sum_{t=1}^n X_{t-1} \left( (1 + \alpha \theta) \eta_t + \alpha (\eta_t^2 - \tau_2) \right), \\ \bar{M}_n^{(2)} &= \alpha^2 \sum_{t=1}^n \eta_{t-1} X_{t-1} \eta_t, \\ \bar{M}_n^{(3)} &= \sum_{t=1}^n (1 + \alpha \eta_t) \epsilon_t.\end{aligned}$$

Considérons la filtration  $(\mathcal{F}_n^*)$  engendrée par  $\mathcal{F}_0^* = \sigma(X_0, \eta_0)$  et, pour  $n \geq 1$ , par

$$\mathcal{F}_n^* = \sigma(X_0, \eta_0, (\epsilon_1, \eta_1), \dots, (\epsilon_n, \eta_n)) \quad (4.52)$$

et posons

$$\bar{M}_n = \begin{pmatrix} \bar{M}_n^{(1)} \\ \bar{M}_n^{(2)} \\ \bar{M}_n^{(3)} \end{pmatrix}. \quad (4.53)$$

Sous nos hypothèses,  $\bar{M}_n$  est une  $(\mathcal{F}_n^*)$ -martingale localement de carré intégrable. Nous allons utiliser le Théorème central limite pour les martingales décrit dans le Corollaire 2.1.10 de [39]. Dans un premier temps, nous devons étudier le comportement asymptotique de la variation quadratique prévisible de  $\bar{M}_n$ . Pour tout  $n \geq 1$ , posons

$$\langle \bar{M} \rangle_n = \sum_{t=1}^n \mathbb{E} \left( (\Delta \bar{M}_t) (\Delta \bar{M}_t)^T \mid \mathcal{F}_{t-1}^* \right), \quad (4.54)$$

avec  $\Delta \bar{M}_1 = \bar{M}_1$ . Nous introduisons quelques notations afin de simplifier les calculs. Les moments d'ordre deux du processus sont notés

$$\mathbb{E} \left( \Lambda_t X_t^2 \right) = \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \Lambda \quad (4.55)$$

où  $\Lambda_t$  est donnée (4.51), avec  $\lambda_0 = \gamma_X(0)$ . En utilisant la stationnarité stricte afin d'établir, via le Corollaire 4.13 sous l'hypothèse additionnelle  $(H_3)$ , que

$$\Lambda = \sigma_2 (I_3 - M)^{-1} U_0 \quad (4.56)$$

et nous obtenons donc immédiatement via un argument d'ergodicité

$$\frac{1}{n} \sum_{t=1}^n \Lambda_t X_t^2 \xrightarrow[n \rightarrow \infty]{p.s.} \Lambda. \quad (4.57)$$

Nous allons maintenant étudier le comportement asymptotique de  $\langle \bar{M} \rangle_n / n$ . Premièrement, sous les hypothèses formulées

$$\langle \bar{M}^{(1)}, \bar{M}^{(3)} \rangle_n = \langle \bar{M}^{(2)}, \bar{M}^{(3)} \rangle_n = 0.$$

Les autres calculs étant similaires nous détaillons seulement le premier terme du crochet,

$$\begin{aligned} \langle \bar{M}^{(1)} \rangle_n &= \sum_{t=1}^n X_{t-1}^2 \mathbb{E} \left( \left( (1 + \alpha \theta) \eta_t + \alpha (\eta_t^2 - \tau_2) \right)^2 \right) \\ &= \left( (1 + \alpha \theta)^2 \tau_2 + \alpha^2 (\tau_4 - \tau_2^2) \right) \sum_{t=1}^n X_{t-1}^2. \end{aligned}$$

Nous obtenons via  $\bar{K}$  donnée dans (4.109) que

$$\langle \bar{M} \rangle_n = \bar{K} \circ \sum_{t=1}^n \begin{pmatrix} X_t^2 & \eta_t X_t^2 & 0 \\ \eta_t X_t^2 & \eta_t^2 X_t^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \bar{R}_n \quad (4.58)$$

où le produit de Hadamard  $\circ$  est utilisé afin d'alléger la formule, et où le terme de reste  $\bar{R}_n$  est constitué de termes isolés tels que, via (4.57),

$$\frac{\bar{R}_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (4.59)$$

Nous obtenons ce résultat en calculant  $\langle \bar{M}^{(i)}, \bar{M}^{(j)} \rangle_n$  pour  $i, j \in \{1, 2, 3\}$  comme nous l'avons fait pour quelques termes, puis en re-normalisant chaque somme, sans les termes isolés qui figurent dans le terme de reste. Par exemple,

$$\sum_{t=1}^n X_{t-1}^2 = \sum_{t=1}^n X_t^2 + (X_0^2 - X_n^2).$$

Une application directe du Théorème ergodique donne une estimation en  $O(n)$  de la somme (et en conséquence en  $o(n)$  pour le terme de reste), ce qui nous permet via (4.57), (4.58) et (4.59), d'identifier la valeur limite

$$\frac{\langle \bar{M} \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \bar{K} \circ \bar{\Gamma} \quad (4.60)$$

où  $\bar{\Gamma}$  est donnée par

$$\bar{\Gamma} = \begin{pmatrix} \lambda_0 & \lambda_1 & 0 \\ \lambda_1 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.61)$$

Dans un second temps, nous devons prouver la condition de Lindeberg, c'est à dire pour tout  $\epsilon > 0$ ,

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E} (\|\Delta \bar{M}_t\|^2 \mathbb{1}_{\{\|\Delta \bar{M}_t\| \geq \epsilon \sqrt{n}\}} | \mathcal{F}_{t-1}^*) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \quad (4.62)$$

lorsque  $n$  tend vers l'infini. Comme les incréments  $(\Delta \bar{M}_t)$  sont strictement stationnaire et ergodique il vient que, sous l'hypothèse sur  $X_0$ , pour tout  $M > 0$ ,

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E} (\|\Delta \bar{M}_t\|^2 \mathbb{1}_{\{\|\Delta \bar{M}_t\| \geq M\}} | \mathcal{F}_{t-1}^*) \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E} (\|\Delta \bar{M}_1\|^2 \mathbb{1}_{\{\|\Delta \bar{M}_1\| \geq M\}}).$$

Le Corollaire 4.13 implique que  $\mathbb{E} (\|\Delta \bar{M}_1\|^2) < \infty$  et le terme de droite peut être rendu aussi petit que l'on souhaite, ce qui vérifie la condition de Lindeberg. Nous déduisons de (4.60) et (4.62) que

$$\frac{\bar{M}_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} (0, \bar{K} \circ \bar{\Gamma}) \quad (4.63)$$

qui n'est rien d'autre qu'une application directe du Théorème central limite décrit précédemment. On peut remarquer que le raisonnement ci-dessus est en fait une extension vectorielle du résultat principal de [14], relatif au Théorème central limite pour des martingales à incrément strictement stationnaire et ergodique. Nous obtenons finalement via une étape calculatoire que

$$\sqrt{n} \bar{X}_n = \frac{\Omega_3^T \bar{M}_n + \bar{r}_n}{(1 - \theta - \alpha \tau_2) \sqrt{n}}$$

où  $\Omega_3^T = (1 \ 1 \ 1)$  et  $\bar{r}_n = o(\sqrt{n})$  ap.s. via (4.57). Il ne reste plus qu'à appliquer le Lemme de Slutsky afin de conclure que

$$\sqrt{n} \bar{X}_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} (0, \kappa^2)$$

avec

$$\kappa^2 = \frac{\Omega_3^T (\bar{K} \circ \bar{\Gamma}) \Omega_3}{(1 - \theta - \alpha \tau_2)^2} \quad (4.64)$$

construite via les notations précédentes.

#### 4.8.5 Preuve du Théorème 4.4

La convergence presque sûre est liée à l'ergodicité du processus. Le Théorème 4.2 conjuguée au Théorème ergodique nous conduit directement à

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \frac{\gamma_X(1)}{\gamma_X(0)} = \frac{[N (I_3 - M)^{-1} U_0]_1}{[(I_3 - M)^{-1} U_0]_1}$$

lorsque  $n$  tend vers l'infini, mais nous aimerions identifier la valeur limite. En utilisant les expressions (4.1)–(4.2), nous obtenons le développement

$$\sum_{t=1}^n X_{t-1} X_t = \theta \sum_{t=1}^n X_{t-1}^2 + \alpha \sum_{t=1}^n \eta_{t-1} X_{t-1}^2 + \sum_{t=1}^n X_{t-1}^2 \eta_t + \sum_{t=1}^n X_{t-1} \epsilon_t. \quad (4.65)$$

Nous déduisons du Corollaire 4.13 que, pour tout  $t \in \mathbb{Z}$ ,

$$\begin{aligned}\mathbb{E}(\eta_t X_t^2) &= \mathbb{E}(\theta_t^2 X_{t-1}^2 \eta_t) + \mathbb{E}(\epsilon_t^2 \eta_t) + 2\mathbb{E}(\theta_t X_{t-1} \epsilon_t \eta_t) \\ &= 2\alpha\tau_2 \mathbb{E}(\eta_{t-1} X_{t-1}^2) + 2\theta\tau_2 \mathbb{E}(X_{t-1}^2).\end{aligned}$$

D'où, par ergodicité,

$$\frac{1}{n} \sum_{t=1}^n \eta_{t-1} X_{t-1}^2 \xrightarrow[n \rightarrow \infty]{p.s.} \frac{2\theta\tau_2 \gamma_X(0)}{1 - 2\alpha\tau_2}. \quad (4.66)$$

De manière analogue,  $\mathbb{E}(X_{t-1}^2 \eta_t) = \mathbb{E}(X_{t-1} \epsilon_t) = 0$  et le Théorème ergodique implique,

$$\frac{1}{n} \sum_{t=1}^n X_{t-1}^2 \xrightarrow[n \rightarrow \infty]{p.s.} \gamma_X(0), \quad \frac{1}{n} \sum_{t=1}^n X_{t-1}^2 \eta_t \xrightarrow[n \rightarrow \infty]{p.s.} 0, \quad \frac{1}{n} \sum_{t=1}^n X_{t-1} \epsilon_t \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (4.67)$$

L'expression de  $\widehat{\theta}_n$  décrite dans (4.17) combinée à la décomposition (4.65) ainsi qu'à (4.66) et (4.67) implique

$$\widehat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta + \frac{2\alpha\theta\tau_2}{1 - 2\alpha\tau_2} = \frac{\theta}{1 - 2\alpha\tau_2}.$$

Intéressons nous maintenant à la normalité asymptotique de cet estimateur. Premièrement, nous devons donner une condition assurant l'existence du moment d'ordre quatre de  $(X_t)$  ce qui, en conséquence, requiert des Lemmes techniques. Pour tout  $k \in \mathbb{N}^*$ , considérons les suites

$$v_k^{(a)} = \mathbb{E}(\eta_k^a \theta_k^4 \dots \theta_1^4)$$

où  $a \in \{0, \dots, 4\}$ , et posons

$$V_k = \begin{pmatrix} v_k^{(0)} \\ \vdots \\ v_k^{(4)} \end{pmatrix}. \quad (4.68)$$

Dans la suite des preuves,  $H$  est défini dans (4.6) et  $\{V_0, \dots, V_4\}$  dans (4.5).

**Lemme 4.14.** *Supposons que  $(H_1)$ – $(H_4)$  sont vérifiées. Alors, pour tout  $k \in \mathbb{N}$ ,*

$$V_k = H^k V_0. \quad (4.69)$$

**Preuve.** Nous avons via la filtration  $(\mathcal{F}_t)$  décrite dans (4.8), que pour tout  $k \geq 1$ ,

$$\begin{aligned}
v_k^{(0)} &= \mathbb{E}(\theta_{k-1}^4 \dots \theta_1^4 \mathbb{E}(\theta_k^4 | \mathcal{F}_{k-1})) \\
&= (\theta^4 + 6\theta^2 \tau_2 + \tau_4) v_{k-1}^{(0)} + 4\alpha(\theta^3 + 3\theta \tau_2) v_{k-1}^{(1)} + 6\alpha^2(\theta^2 + \tau_2) v_{k-1}^{(2)} \\
&\quad + 4\alpha^3 \theta v_{k-1}^{(3)} + \alpha^4 v_{k-1}^{(4)}, \\
v_k^{(1)} &= \mathbb{E}(\theta_{k-1}^4 \dots \theta_1^4 \mathbb{E}(\eta_k \theta_k^4 | \mathcal{F}_{k-1})) \\
&= (4\theta^3 \tau_2 + 4\theta \tau_4) v_{k-1}^{(0)} + 4\alpha(3\theta^2 \tau_2 + \tau_4) v_{k-1}^{(1)} + 12\alpha^2 \theta \tau_2 v_{k-1}^{(2)} \\
&\quad + 4\alpha^3 \tau_2 v_{k-1}^{(3)}, \\
v_k^{(2)} &= \mathbb{E}(\theta_{k-1}^4 \dots \theta_1^4 \mathbb{E}(\eta_k^2 \theta_k^4 | \mathcal{F}_{k-1})) \\
&= (\theta^4 \tau_2 + 6\theta^2 \tau_4 + \tau_6) v_{k-1}^{(0)} + 4\alpha(\theta^3 \tau_2 + 3\theta \tau_4) v_{k-1}^{(1)} + 6\alpha^2(\theta^2 \tau_2 + \tau_4) v_{k-1}^{(2)} \\
&\quad + 4\alpha^3 \theta \tau_2 v_{k-1}^{(3)} + \alpha^4 \tau_2 v_{k-1}^{(4)}, \\
v_k^{(3)} &= \mathbb{E}(\theta_{k-1}^4 \dots \theta_1^4 \mathbb{E}(\eta_k^3 \theta_k^4 | \mathcal{F}_{k-1})) \\
&= (4\theta^3 \tau_4 + 4\theta \tau_6) v_{k-1}^{(0)} + 4\alpha(3\theta^2 \tau_4 + \tau_6) v_{k-1}^{(1)} + 12\alpha^2 \theta \tau_4 v_{k-1}^{(2)} \\
&\quad + 4\alpha^3 \tau_4 v_{k-1}^{(3)}, \\
v_k^{(4)} &= \mathbb{E}(\theta_{k-1}^4 \dots \theta_1^4 \mathbb{E}(\eta_k^4 \theta_k^4 | \mathcal{F}_{k-1})) \\
&= (\theta^4 \tau_4 + 6\theta^2 \tau_6 + \tau_8) v_{k-1}^{(0)} + 4\alpha(\theta^3 \tau_4 + 3\theta \tau_6) v_{k-1}^{(1)} + 6\alpha^2(\theta^2 \tau_4 + \tau_6) v_{k-1}^{(2)} \\
&\quad + 4\alpha^3 \theta \tau_4 v_{k-1}^{(3)} + \alpha^4 \tau_4 v_{k-1}^{(4)},
\end{aligned}$$

où la Table 4.1 peut être utilisée afin d'identifier rapidement les coefficients intervenant dans les calculs. Nous obtenons donc la formulation matricielle  $V_k = H V_{k-1}$  où la valeur initiale  $V_1$  est obtenue *via*

$$\begin{aligned}
v_1^{(0)} &= \mathbb{E}(\theta_1^4) = (\theta^4 + 6\theta^2 \tau_2 + \tau_4) + 6\alpha^2 \tau_2 (\theta^2 + \tau_2) + \alpha^4 \tau_4, \\
v_1^{(1)} &= \mathbb{E}(\eta_1 \theta_1^4) = (4\theta^3 \tau_2 + 4\theta \tau_4) + 12\alpha^2 \theta \tau_2^2, \\
v_1^{(2)} &= \mathbb{E}(\eta_1^2 \theta_1^4) = (\theta^4 \tau_2 + 6\theta^2 \tau_4 + \tau_6) + 6\alpha^2 \tau_2 (\theta^2 \tau_2 + \tau_4) + \alpha^4 \tau_2 \tau_4, \\
v_1^{(3)} &= \mathbb{E}(\eta_1^3 \theta_1^4) = (4\theta^3 \tau_4 + 4\theta \tau_6) + 12\alpha^2 \theta \tau_2 \tau_4, \\
v_1^{(4)} &= \mathbb{E}(\eta_1^4 \theta_1^4) = (\theta^4 \tau_4 + 6\theta^2 \tau_6 + \tau_8) + 6\alpha^2 \tau_2 (\theta^2 \tau_4 + \tau_6) + \alpha^4 \tau_4^2.
\end{aligned}$$

Hence,  $V_1 = H V_0$ . □

Pour tout  $1 \leq k < \ell$ , considérons les suites

$$w_{\ell,k}^{(a)} = \mathbb{E}(\eta_\ell^a \theta_\ell^4 \dots \theta_{\ell-k+1}^4 \theta_{\ell-k}^2 \dots \theta_1^2)$$

où  $a \in \{0, \dots, 4\}$ , puis posons

$$W_{\ell,k} = \begin{pmatrix} w_{\ell,k}^{(0)} \\ \vdots \\ w_{\ell,k}^{(4)} \end{pmatrix} G = \begin{pmatrix} \theta^2 + \tau_2 & 2\alpha\theta & \alpha^2 & 0 & 0 \\ 2\theta\tau_2 & 2\alpha\tau_2 & 0 & 0 & 0 \\ \theta^2\tau_2 + \tau_4 & 2\alpha\theta\tau_2 & \alpha^2\tau_2 & 0 & 0 \\ 2\theta\tau_4 & 2\alpha\tau_4 & 0 & 0 & 0 \\ \theta^2\tau_4 + \tau_6 & 2\alpha\theta\tau_4 & \alpha^2\tau_4 & 0 & 0 \end{pmatrix}.$$

Une nouvelle fois,  $G$  peut être obtenue à partir de  $\{V_0, \dots, V_4\}$ ,

$$\begin{cases} G_1 = \theta^2 V_0 + 2\theta V_1 + V_2 \\ G_2 = 2\alpha(\theta V_0 + V_1) \\ G_3 = \alpha^2 V_0 \\ G_4 = 0 \\ G_5 = 0. \end{cases} \quad (4.70)$$

Notons que le block de taille  $3 \times 3$  en haut à gauche de  $G$  est exactement  $M$  définie dans (4.4). Cet argument sera utilisé par la suite afin d'assurer que  $\rho(G) < 1$ .

**Lemme 4.15.** *Supposons que  $(H_1)$ – $(H_4)$  sont vérifiées. Alors, pour tout  $1 \leq k < \ell$ ,*

$$W_{\ell,k} = H^k G^{\ell-k} V_0. \quad (4.71)$$

**Preuve.** Les calculs sont analogues à ceux menés dans les Lemmes 4.12 et 4.14. Nous avons,

$$W_{\ell,k} = H^k U_{\ell-k}$$

où nous étendons la définition de  $U_{k,h}$  dans (4.48) pour  $a \in \{0, \dots, 4\}$ , c'est à dire

$$U_k = \begin{pmatrix} u_k^{(0)} \\ \vdots \\ u_k^{(4)} \end{pmatrix} = \begin{pmatrix} u_{k,0}^{(0)} \\ \vdots \\ u_{k,0}^{(4)} \end{pmatrix} = U_{k,0}.$$

Il reste donc à étudier le comportement des termes  $u_{\ell-k}$  pour  $a = 3$  et  $a = 4$  en faisant appel à la Table 4.1,

$$\begin{aligned} u_{\ell-k}^{(3)} &= \mathbb{E}(\theta_{\ell-k-1}^2 \dots \theta_1^2 \mathbb{E}(\eta_{\ell-k}^3 \theta_{\ell-k}^2 | \mathcal{F}_{\ell-k-1})) \\ &= 2\theta \tau_4 u_{\ell-k-1}^{(0)} + 2\alpha \tau_4 u_{\ell-k-1}^{(1)}, \\ u_{\ell-k}^{(4)} &= \mathbb{E}(\theta_{\ell-k-1}^2 \dots \theta_1^2 \mathbb{E}(\eta_{\ell-k}^4 \theta_{\ell-k}^2 | \mathcal{F}_{\ell-k-1})) \\ &= (\theta^2 \tau_4 + \tau_6) u_{\ell-k-1}^{(0)} + 2\alpha \theta \tau_4 u_{\ell-k-1}^{(1)} + \alpha^2 \tau_4 u_{\ell-k-1}^{(2)}. \quad \square \end{aligned}$$

D'où,  $U_{\ell-k} = G U_{\ell-k-1}$ . Nous en déduisons donc que, pour tout  $1 \leq k < \ell$ ,

$$U_{\ell-k} = G^{\ell-k} V_0.$$

**Corollaire 4.16.** *Supposons que  $(H_1)$ – $(H_4)$  sont vérifiées. Alors, le processus  $(X_t)$  possède un moment d'ordre quatre qui, vérifient pour tout  $a \in \{0, \dots, 4\}$ ,*

$$\mathbb{E}(\eta_t^a X_t^4) < \infty.$$

**Preuve.** Posons pour tout  $t \in \mathbb{Z}$  et  $k \geq 1$ ,

$$\Delta_t = \begin{pmatrix} 1 \\ \eta_t \\ \vdots \\ \eta_t^4 \end{pmatrix} P_{t,k} = \prod_{i=0}^{k-1} \theta_{t-i} \quad (4.72)$$

avec  $P_{t,0} = 1$ . Puisque les processus  $(\epsilon_t)$  et  $(\eta_t)$  sont des bruits blancs forts indépendants, nous avons via la représentation causale (4.7) et les notations précédentes,

$$\begin{aligned}
\mathbb{E}(\Delta_t X_t^4) &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{u=0}^{\infty} \sum_{v=0}^{\infty} \mathbb{E}(\Delta_t P_{t,k} P_{t,\ell} P_{t,u} P_{t,v} \epsilon_{t-k} \epsilon_{t-\ell} \epsilon_{t-u} \epsilon_{t-v}) \\
&= \sigma_4 \sum_{k=0}^{\infty} \mathbb{E}(\Delta_t P_{t,k}^4) + 6 \sigma_2^2 \sum_{k=0}^{\infty} \sum_{\ell=k+1}^{\infty} \mathbb{E}(\Delta_t P_{t,k}^2 P_{t,\ell}^2) \\
&= \sigma_4 \sum_{k=0}^{\infty} V_k + 6 \sigma_2^2 \sum_{\ell=1}^{\infty} U_\ell + 6 \sigma_2^2 \sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} W_{\ell,k}.
\end{aligned}$$

En conséquence, les Lemmes 4.14 et 4.15 conjuguée à la stationnarité stricte de  $(\theta_t)$  nous permettent d'achever la preuve, puisque  $\rho(G) = \rho(M) < 1$ .  $\square$

Nous revenons à la preuve du Théorème 4.4 et nous convenons des notations suivantes,

$$\begin{aligned}
M_n^{(1)} &= \sum_{t=1}^n X_{t-1} \left( (1 - 2\alpha\tau_2) \epsilon_t + 2\alpha\theta \eta_t \epsilon_t + 2\alpha\eta_t^2 \epsilon_t \right), \\
M_n^{(2)} &= \sum_{t=1}^n X_{t-1}^2 \left( (1 - 2\alpha\tau_2 + \alpha\theta^2) \eta_t + \alpha\eta_t^3 + 2\alpha\theta(\eta_t^2 - \tau_2) \right), \\
M_n^{(3)} &= 2\alpha^2 \sum_{t=1}^n \eta_{t-1} X_{t-1} \eta_t \epsilon_t, \\
M_n^{(4)} &= \sum_{t=1}^n \eta_{t-1} X_{t-1}^2 \left( 2\alpha^2 \theta \eta_t + 2\alpha^2(\eta_t^2 - \tau_2) \right), \\
M_n^{(5)} &= \alpha^3 \sum_{t=1}^n \eta_{t-1}^2 X_{t-1}^2 \eta_t, \\
M_n^{(6)} &= \alpha \sum_{t=1}^n \eta_t \epsilon_t^2.
\end{aligned}$$

Considérons la filtration  $(\mathcal{F}_n^*)$  donnée dans (4.52), puis posons

$$M_n = \begin{pmatrix} M_n^{(1)} \\ \vdots \\ M_n^{(6)} \end{pmatrix}. \quad (4.73)$$

Sous nos hypothèses,  $M_n$  est une  $(\mathcal{F}_n^*)$ -martingale localement de carrée intégrable. Nous allons utiliser de nouveau le Théorème central limite pour les martingales, comme dans la preuve du Théorème 4.3. Nous devons dans un premier temps, étudier le comportement asymptotique de la variation quadratique prévisible de  $M_n$ . Pour tout  $n \geq 1$ , posons

$$\langle M \rangle_n = \sum_{t=1}^n \mathbb{E} \left( (\Delta M_t) (\Delta M_t)^T \mid \mathcal{F}_{t-1}^* \right), \quad (4.74)$$

avec  $\Delta M_1 = M_1$ . Pour simplifier les calculs, nous introduisons des notations supplémentaires. Les moments d'ordre deux du processus sont définis dans (4.55) et les moments d'ordre quatre sont notés comme suit

$$\mathbb{E} (\Delta_t X_t^4) = \begin{pmatrix} \delta_0 \\ \vdots \\ \delta_4 \end{pmatrix} = \Delta \quad (4.75)$$

où  $\Delta_t$  est donnée dans (4.72). Nous utilisons la stationnarité stricte afin d'établir, via les Corollaires 4.13 et 4.16, que

$$\Delta = (I_5 - H)^{-1} (\sigma_2 R + \sigma_4 V_0) \quad (4.76)$$

où  $R$  est défini à partir de (4.70) par

$$R = 6 \lambda_0 G_1 + 6 \lambda_1 G_2 + 6 \lambda_2 G_3.$$

Nous allons montrer maintenant que le comportement asymptotique de  $\langle M \rangle_n / n$  est entièrement décrit par  $\Lambda$  et  $\Delta$ . Par ergodicité,

$$\frac{1}{n} \sum_{t=1}^n \Delta_t X_t^4 \xrightarrow[n \rightarrow \infty]{p.s.} \Delta. \quad (4.77)$$

Nous revenons à (4.74). Premièrement, il existe des constantes telles que

$$\begin{aligned} \langle M^{(1)}, M^{(2)} \rangle_n &= \sum_{t=1}^n X_{t-1}^3 \mathbb{E} [ (k_{(1)} + k_{(2)} \eta_t + k_{(3)} \eta_t^2) (k_{(4)} \eta_t + k_{(5)} \eta_t^3 \\ &\quad + k_{(6)} (\eta_t^2 - \tau_2)) \epsilon_t ] = 0 \end{aligned}$$

sous nos hypothèses. *Via* des arguments analogues, nous obtenons que

$$\begin{aligned} \langle M^{(1)}, M^{(4)} \rangle_n &= \langle M^{(1)}, M^{(5)} \rangle_n = \langle M^{(1)}, M^{(6)} \rangle_n = \langle M^{(2)}, M^{(3)} \rangle_n \\ &= \langle M^{(3)}, M^{(4)} \rangle_n = \langle M^{(3)}, M^{(5)} \rangle_n = \langle M^{(3)}, M^{(6)} \rangle_n = 0. \end{aligned}$$

Nous nous intéressons maintenant aux contributions non nulles, où nous utilisons les constantes définies dans (4.110) et (4.111). Les calculs étant similaires nous donnons le détail du premier terme

$$\begin{aligned} \langle M^{(1)} \rangle_n &= \sum_{t=1}^n X_{t-1}^2 \mathbb{E} \left( ((1 - 2\alpha \tau_2) \epsilon_t + 2\alpha \theta \eta_t \epsilon_t + 2\alpha \eta_t^2 \epsilon_t)^2 \right) \\ &= \sigma_2 (1 + 4\alpha^2 (\theta^2 \tau_2 - \tau_2^2 + \tau_4)) \sum_{t=1}^n X_{t-1}^2. \end{aligned}$$

Nous obtenons via la présentation de l'ensemble des termes sous forme matricielle que

$$\langle M \rangle_n = K \circ \sum_{t=1}^n \begin{pmatrix} X_t^2 & 0 & \eta_t X_t^2 & 0 & 0 & 0 \\ 0 & X_t^4 & 0 & \eta_t X_t^4 & \eta_t^2 X_t^4 & X_t^2 \\ \eta_t X_t^2 & 0 & \eta_t^2 X_t^2 & 0 & 0 & 0 \\ 0 & \eta_t X_t^4 & 0 & \eta_t^2 X_t^4 & \eta_t^3 X_t^4 & \eta_t X_t^2 \\ 0 & \eta_t^2 X_t^4 & 0 & \eta_t^3 X_t^4 & \eta_t^4 X_t^4 & \eta_t^2 X_t^2 \\ 0 & X_t^2 & 0 & \eta_t X_t^2 & \eta_t^2 X_t^2 & 1 \end{pmatrix} + R_n \quad (4.78)$$

où le produit de Hadamard  $\circ$  est de nouveau utilisé afin d'alléger l'écriture, et où le terme de reste  $R_n$  est constitué de termes isolés de sorte que

$$\frac{R_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (4.79)$$

Nous obtenons ce résultat de la même façon que (4.59) puisque la méthodologie employée pour traiter  $M_n$  est exactement la même. Les combinaisons de (4.57), (4.77), (4.78) et (4.79) conduisent à

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \Gamma \quad (4.80)$$

où  $\Gamma$  est donnée par

$$\Gamma = \begin{pmatrix} \lambda_0 & 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & \delta_0 & 0 & \delta_1 & \delta_2 & \lambda_0 \\ \lambda_1 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & \delta_1 & 0 & \delta_2 & \delta_3 & \lambda_1 \\ 0 & \delta_2 & 0 & \delta_3 & \delta_4 & \lambda_2 \\ 0 & \lambda_0 & 0 & \lambda_1 & \lambda_2 & 1 \end{pmatrix}. \quad (4.81)$$

Dans un second temps, il est nécessaire de vérifier la condition de Lindeberg, c'est à dire pour tout  $\epsilon > 0$ ,

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E} (\|\Delta M_t\|^2 \mathbb{1}_{\{\|\Delta M_t\| \geq \epsilon \sqrt{n}\}} \mid \mathcal{F}_{t-1}^*) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \quad (4.82)$$

dès lors que  $n$  tend vers l'infini. Cette condition découle des Corollaires 4.13 et 4.16, conjuguée aux mêmes arguments invoqués pour obtenir (4.62). Nous déduisons de (4.80) et (4.82) que

$$\frac{M_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, K \circ \Gamma). \quad (4.83)$$

Finalement nous obtenons via une étape très calculatoire que

$$\sqrt{n} (\hat{\theta}_n - \theta^*) = \frac{n}{\sum_{t=1}^n X_{t-1}^2} \frac{\Omega_6^T M_n + r_n}{(1 - 2\alpha\tau_2)\sqrt{n}} \quad (4.84)$$

où  $\Omega_6^T = (1 \ 1 \ 1 \ 1 \ 1 \ 1)$  et  $r_n = o(\sqrt{n})$  p.s. Il ne reste plus qu'à appliquer le Lemme de Slutsky afin de conclure que

$$\sqrt{n} (\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{N}(0, \omega^2)}$$

où

$$\omega^2 = \frac{\Omega_6^T (K \circ \Gamma) \Omega_6}{\lambda_0^2 (1 - 2\alpha\tau_2)^2} \quad (4.85)$$

en utilisant les notations précédentes.

### 4.8.6 Preuve du Théorème 4.5

Soit  $V_n = \sqrt{n} I_6$  une suite satisfaisant les conditions de croissances régulières décrites dans [27]. En utilisant la notation (4.73), nous avons déjà étudié le comportement asymptotique de la variation quadratique prévisible de  $M_n$  dans (4.80) ainsi que la condition de Lindeberg dans (4.82), il ne reste plus qu'à vérifier que

$$\frac{[M]_n - \langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{0} 0 \quad (4.86)$$

où

$$[M]_n = \sum_{t=1}^n (\Delta M_t)(\Delta M_t)^T$$

est la variation totale de  $M_n$ , pour ensuite appliquer le Théorème 2.1 de [27]. Pour être précis avec les hypothèses demandées, nous précisons que (4.82) est également valide pour la convergence presque sûre, par ergodicité. La condition (4.86) est une conséquence immédiate de l'ergodicité des incréments. D'où,

$$\frac{1}{6 \log n} \sum_{t=1}^n \left[ 1 - \left( \frac{t}{t+1} \right)^6 \right] \frac{M_t M_t^T}{t} \xrightarrow[n \rightarrow \infty]{p.s.} K \circ \Gamma$$

et, après simplifications,

$$\frac{1}{\log n} \sum_{t=1}^n \frac{M_t M_t^T}{t^2} \xrightarrow[n \rightarrow \infty]{p.s.} K \circ \Gamma. \quad (4.87)$$

Le terme de reste  $r_n$  dans (4.84) est une longue combinaison des termes isolés, nous détaillons le traitement du terme prépondérant ayant la forme  $\eta_{n-1}^2 X_{n-1}^2 \eta_n$ . Le Corollaire 4.16 implique, pour  $a = 4$  *via* le Théorème ergodique, que

$$\frac{1}{n} \sum_{t=1}^n \eta_{t-1}^4 X_{t-1}^4 \eta_t^2 \xrightarrow[n \rightarrow \infty]{p.s.} \delta_4 \tau_2,$$

ce qui implique

$$\frac{\eta_{n-1}^4 X_{n-1}^4 \eta_n^2}{n} \xrightarrow[n \rightarrow \infty]{p.s.} 0 \text{ et donc } \frac{\eta_{n-1}^4 X_{n-1}^4 \eta_n^2}{n^2} = o(n^{-1}) \text{ p.s.}$$

Il vient que

$$\sum_{t=1}^n \frac{\eta_{t-1}^4 X_{t-1}^4 \eta_t^2}{t^2} = o\left(\sum_{t=1}^n \frac{1}{t}\right) = o(\log n) \text{ p.s.}$$

En traitant les autres termes de reste de manière similaire nous avons

$$\frac{1}{\log n} \sum_{t=1}^n \frac{r_t^2}{t^2} \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad (4.88)$$

Il ne reste plus qu'à combiner ces calculs pour obtenir

$$\begin{aligned} \frac{(1 - 2\alpha\tau_2)^2}{\log n} \sum_{t=1}^n (\widehat{\theta}_t - \theta^*)^2 &= \frac{1}{\log n} \sum_{t=1}^n \frac{\Omega_6^T M_t M_t^T \Omega_6}{S_{t-1}^2} + \frac{1}{\log n} \sum_{t=1}^n \frac{r_t^2}{S_{t-1}^2} \\ &\quad + \frac{2}{\ln n} \sum_{t=1}^n \frac{\Omega_6^T M_t r_t}{S_{t-1}^2} \end{aligned}$$

où

$$S_n = \sum_{t=0}^n X_t^2 \text{ vérifie } \frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \lambda_0. \quad (4.89)$$

L'inégalité de Cauchy-Schwarz permet d'en déduire que le terme croisé est négligeable. Les équations (4.85), (4.87), (4.88) et la remarque précédente nous permettent d'en déduire que,

$$\frac{1}{\log n} \sum_{t=1}^n (\widehat{\theta}_t - \theta^*)^2 \xrightarrow[n \rightarrow \infty]{p.s.} \frac{\Omega_6^T (K \circ \Gamma) \Omega_6}{\lambda_0^2 (1 - 2\alpha\tau_2)^2} = \omega^2$$

qui conclut la première partie de la preuve et découle de l'application du Lemme de Toeplitz dans le terme de droite de la décomposition. L'estimation de la vitesse de convergence de  $\widehat{\theta}_n$  est plus facile à obtenir. Nous rappelons que  $M_n$  est une  $(\mathcal{F}_n^*)$ -martingale vectorielle dont les incréments sont strictement stationnaire et ergodique. D'où,

$$N_n = \Omega_6^T M_n \quad (4.90)$$

est une  $(\mathcal{F}_n^*)$ -martingale scalaire s'intégrant au même ordre que, et nos hypothèses garantissent que  $\mathbb{E}((\Delta N_1)^2) = \Omega_6^T (K \circ \Gamma) \Omega_6 < \infty$ . Le Théorème de [93] nous permet d'en déduire que

$$\limsup_{n \rightarrow +\infty} \frac{N_n}{\sqrt{2n \log \log n}} = \sqrt{\Omega_6^T (K \circ \Gamma) \Omega_6} \text{ p.s.} \quad (4.91)$$

et

$$\liminf_{n \rightarrow +\infty} \frac{N_n}{\sqrt{2n \log \log n}} = -\sqrt{\Omega_6^T (K \circ \Gamma) \Omega_6} \text{ p.s.} \quad (4.92)$$

en remplaçant  $N_n$  par  $-N_n$ . Grâce à la décomposition (4.84),

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \sqrt{\frac{n}{2 \log \log n}} (\widehat{\theta}_n - \theta^*) &= \frac{1}{\lambda_0 (1 - 2\alpha\tau_2)} \limsup_{n \rightarrow +\infty} \frac{N_n + r_n}{\sqrt{2n \log \log n}} \\ &= \omega \text{ a.s.} \end{aligned}$$

en utilisant (4.91) et le fait que  $r_n = o(\sqrt{n})$  p.s. Le résultat symétrique s'obtient via (4.92) et la preuve est complète.

#### 4.8.7 Preuve du Théorème 4.7

Nous allons démontrer ce résultat en deux étapes. Premièrement, nous allons identifier une matrice de covariance  $\Sigma$  telle que

$$\sqrt{n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\vartheta}_n - \vartheta^* \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma) \quad (4.93)$$

où  $\widehat{\theta}_n$  et  $\widehat{\vartheta}_n$  sont donnés dans (4.31),  $\theta^* = \rho_X(1)$  est la valeur limite de  $\widehat{\theta}_n$  largement étudiée jusqu'à présent et

$$\vartheta^* = \rho_X(2) = \frac{\theta^2 + \alpha \tau_2 (1 - 2 \alpha \tau_2)}{1 - 2 \alpha \tau_2}.$$

Nous transférerons ce résultat aux nouveaux estimateurs (4.33) *via* la méthode Delta. La première étape est très similaire à celle de la preuve du Théorème 4.4, nous en donnons juste les étapes clés. Le développement de  $\widehat{\vartheta}_n$  va nous conduire à nouvelle contribution étant une  $(\mathcal{F}_n^*)$ -martingale scalaire que nous définissons

$$\begin{aligned} L_n = & \alpha \sum_{t=1}^n X_{t-1} \eta_t \epsilon_t + \sum_{t=1}^n X_{t-1}^2 (\alpha \theta \eta_t + \alpha (\eta_t^2 - \tau_2)) \\ & + \alpha^2 \sum_{t=1}^n \eta_{t-1} X_{t-1}^2 \eta_t + \sum_{t=2}^n X_{t-2} \epsilon_t + \sum_{t=2}^n X_{t-2} \epsilon_{t-1} \eta_t \\ & + \theta \sum_{t=2}^n X_{t-2}^2 \eta_t + \sum_{t=2}^n X_{t-2}^2 \eta_{t-1} \eta_t + \alpha \sum_{t=2}^n \eta_{t-2} X_{t-2}^2 \eta_t \end{aligned} \quad (4.94)$$

s'obtenant *via* un développement calculatoire  $\sum_{t=2}^n X_{t-2} X_t$ . L'erreur statistique  $\widehat{\vartheta}_n - \vartheta^*$  admet donc le développement

$$(\widehat{\vartheta}_n - \vartheta^*) S_{n-2} = \theta^* \Omega_6^T M_n + L_n + s_n$$

où  $M_n$  est donnée dans (4.73),  $S_n$  dans (4.89),  $\Omega_6^T = (1 \ 1 \ 1 \ 1 \ 1 \ 1)$  et  $s_n$  est constitué de termes isolés, dont chacun est un  $o(\sqrt{n})$  p.s. dès que le processus possède des moments d'ordre quatre, *i.e.* sous  $(H_4)$ . En combinant cela avec (4.84), nous avons,

$$\sqrt{n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\vartheta}_n - \vartheta^* \end{pmatrix} = \frac{A_n}{\sqrt{n}} \begin{pmatrix} M_n \\ L_n \end{pmatrix} + T_n \quad (4.95)$$

où

$$A_n = \begin{pmatrix} \frac{n}{S_{n-1}} & \frac{\Omega_6^T}{1-2\alpha\tau_2} & 0 \\ \frac{n}{S_{n-2}} & \frac{\theta \Omega_6^T}{1-2\alpha\tau_2} & \frac{n}{S_{n-2}} \end{pmatrix} \xrightarrow[n \rightarrow \infty]{p.s.} A = \begin{pmatrix} \frac{\Omega_6^T}{\lambda_0(1-2\alpha\tau_2)} & 0 \\ \frac{\theta \Omega_6^T}{\lambda_0(1-2\alpha\tau_2)} & \frac{1}{\lambda_0} \end{pmatrix} \quad (4.96)$$

est une matrice de taille  $2 \times 7$  et  $T_n = o(1)$  p.s. Nous devons comme dans les autres preuves relatives à la normalité asymptotique étudier le comportement asymptotique de la variation quadratique prévisible de notre nouvelle  $(\mathcal{F}_n^*)$ -martingale. Le crochet  $\langle M \rangle_n$  est déjà traité dans (4.80). Pour le terme croisé, nous avons besoin d'une notation supplémentaire. Soit

$$\mu_{a,b,c,p,q} = \mathbb{E} (\eta_{t-1}^a \eta_t^b \epsilon_t^c X_{t-1}^p X_t^q) \quad (4.97)$$

et notons que  $\mu_{0,b,0,0,2} = [\Lambda]_{b+1}$  est décrit dans (4.55) pour  $b \in \{0, 1, 2\}$  et que  $\mu_{0,b,0,0,4} = [\Delta]_{b+1}$  est décrit dans (4.75) pour  $b \in \{0, \dots, 4\}$ . D'où, nous obtenons *via* les arguments usuels reposant sur l'ergodicité et le fait que les termes de reste sont négligeables, que

$$\frac{\langle M, L \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} (L \circ \Upsilon) \Omega_6 \quad (4.98)$$

où  $L$  est définie dans (4.113) et  $\Upsilon$  donné par

$$\Upsilon = \begin{pmatrix} \theta^* \lambda_0 & \lambda_0 & 0 & 0 & 0 & 0 \\ \delta_0 & \delta_1 & \mu_{0,0,0,2,2} & \mu_{0,1,0,2,2} & \mu_{1,0,0,2,2} & \mu_{0,0,1,1,2} \\ \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ \delta_1 & \delta_2 & \mu_{0,1,0,2,2} & \mu_{0,2,0,2,2} & \mu_{1,1,0,2,2} & \mu_{0,1,1,1,2} \\ \delta_2 & \delta_3 & \mu_{0,2,0,2,2} & \mu_{0,3,0,2,2} & \mu_{1,2,0,2,2} & \mu_{0,2,1,1,2} \\ \lambda_0 & \lambda_1 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (4.99)$$

Nous avons donc,

$$\begin{aligned} \frac{\langle L \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \ell &= m_{(1)} \lambda_0 + m_{(2)} \delta_0 + m_{(3)} \delta_1 + m_{(4)} \delta_2 + \theta m_{(5)} \mu_{0,0,0,2,2} \\ &\quad + \alpha m_{(5)} \mu_{1,0,0,2,2} + (1 + \alpha) m_{(5)} \mu_{0,1,0,2,2} + m_{(5)} \mu_{0,0,1,1,2} \\ &\quad + m_{(6)} \mu_{0,2,0,2,2} + \alpha m_{(6)} \mu_{1,1,0,2,2} + m_{(6)} \mu_{0,1,1,1,2} \end{aligned} \quad (4.100)$$

où les constantes sont décrites dans (4.114). Cette dernière convergence conjuguée à (4.98), (4.80) et les notations usuelles, impliquent

$$\frac{1}{n} \left\langle \begin{pmatrix} M \\ L \end{pmatrix} \right\rangle_n \xrightarrow[n \rightarrow \infty]{p.s.} \Sigma_{\text{ML}} = \begin{pmatrix} K \circ \Gamma & (L \circ \Upsilon) \Omega_6 \\ \Omega_6^T (L \circ \Upsilon)^T & \ell \end{pmatrix}. \quad (4.101)$$

La condition de Lindeberg est vérifiée et le Lemme de Slutsky est appliqué sur la relation (4.95), en prenant en compte la normalité asymptotique de notre martingale et la remarque succédant (4.95), nous pouvons identifier  $\Sigma$  définie dans (4.93) à

$$\Sigma = A \Sigma_{\text{ML}} A^T \quad (4.102)$$

où  $A$  est donnée (4.96). Cela achève la première partie de la preuve.

**Remarque 4.8.1** *Notons que, en dépit de la structure complexe de  $\Sigma$ , elle dépend uniquement des paramètres et peut être explicitée. Chaque terme de la forme  $\mu_{a,b,c,p,q}$  dans  $\Sigma_{\text{ML}}$  existe sous nos hypothèses, en exploitant les moments d'ordre quatre du processus. Nous pouvons calculer chacun de ces termes via la même méthode des Lemmes techniques.*

Considérons la fonction  $f$  définie dans (4.32) où la matrice Jacobienne est

$$\nabla f(x, y) = \begin{pmatrix} \frac{(1-2y)(1+2x^2)}{(1-2x^2)^2} & \frac{-2x}{1-2x^2} \\ \frac{-2x(1-2y)}{(1-2x^2)^2} & \frac{1}{1-2x^2} \end{pmatrix}.$$

Le nouveau couple d'estimateurs (4.33) satisfait donc

$$\sqrt{n} \begin{pmatrix} \tilde{\theta}_n - \theta \\ \tilde{\gamma}_n - \gamma \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( \cdot, \nabla^T f(\theta^*, \vartheta^*) \Sigma \nabla f(\theta^*, \vartheta^*) \right)$$

en vertu de la méthode Delta, où le cas pathologique  $\theta^* = \pm \frac{1}{\sqrt{2}}$  est exclue de notre étude.

Nous bifurquons maintenant sur les démonstrations des résultats énoncés au sein de la Section 4.5. Comme nous pouvons nous y attendre, les démonstrations sont très calculatoires et nous avons donc décidé de créer une Section 4.9 dédiée à l'énumération de nombreuses décompositions martingales afin de clarifier les preuves.

### 4.8.8 Preuve de la Proposition 4.8

Nous allons proposer une démonstration similaire à celle du Théorème 4.1 dans le cas vectoriel. Les hypothèses formulées sur  $(\eta_{k,t})$  impliquent que le processus  $(C_\theta + N_{t-1}D_\alpha + N_t)$  est strictement stationnaire et ergodique. Nous déduisons de la condition (4.42) qu'il existe  $\delta$  et une variable aléatoire  $k_0$  telle que, pour tout  $k \geq k_0$ ,

$$\frac{1}{k} \sum_{\ell=0}^{k-1} \log \|C_\theta + N_{t-\ell-1}D_\alpha + N_{t-\ell}\| < \delta \text{ p.s.}$$

Considérons pour  $n \geq 1$

$$\Phi_{n,t} = E_t + \sum_{k=1}^n E_{t-k} \prod_{\ell=0}^{k-1} (C_\theta + N_{t-\ell-1}D_\alpha + N_{t-\ell}).$$

D'où par l'inégalité triangulaire, pour  $n$  assez grand,

$$\begin{aligned} \|\Phi_{n,t}\| &\leq \|E_t\| + \sum_{k=1}^n \|E_{t-k}\| \prod_{\ell=0}^{k-1} \|C_\theta + N_{t-\ell-1}D_\alpha + N_{t-\ell}\| \\ &\leq |\epsilon_t| + \sum_{k=1}^{k_0-1} |\epsilon_{t-k}| \prod_{\ell=0}^{k-1} \|C_\theta + N_{t-\ell-1}D_\alpha + N_{t-\ell}\| + \sum_{k=k_0}^n |\epsilon_{t-k}| \exp(\delta k) \end{aligned}$$

et le Lemme 2.2 de [10] assure que le membre de droite de la dernière inégalité converge sous la condition (4.42). D'où,

$$\limsup_{n \rightarrow +\infty} \|\Phi_{t,n}\| < +\infty \text{ p.s.}$$

donc la quantité définie dans 4.44 est finie presque sûrement. Le fin de la preuve est similaire à celle du Théorème 4.1.

Soit  $\mathcal{F}_n^{(2)}$  la filtration générée par  $\mathcal{F}_0^{(2)} = \sigma(\Psi_0, \eta_{1,0}, \dots, \eta_{p,0})$  et, pour  $n \geq 1$ ,

$\mathcal{F}_n^{(2)} = \sigma(\Psi_0, \eta_{1,0}, \dots, \eta_{p,0}, (\epsilon_1, \eta_{1,1}, \dots, \eta_{p,1}), \dots, ((\epsilon_n, \eta_{1,n}, \dots, \eta_{p,n}))$  Pour éviter un nombre colossal de notations, dans tout ce qui suit,  $M_n, \delta_n$  et  $R_n$  sont des termes génériques non nécessairement identiques d'une ligne à l'autre désignant  $\mathcal{F}_n^{(2)}$ -martingales, différences de  $\mathcal{F}_n^{(2)}$ -martingales et termes de reste, respectivement. Considérons,

$$M_\theta = \begin{pmatrix} \theta_1 & 0 & \dots & \dots & 0 \\ \theta_2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \theta_p & \dots & \dots & \dots & \theta_1 \end{pmatrix} + \begin{pmatrix} 0 & \theta_2 & \dots & \dots & \theta_p \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \theta_p & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

ainsi que

$$M_{\alpha,\beta} = \begin{pmatrix} \frac{1}{1-2\beta_1} & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix} \left( M_\theta + \begin{pmatrix} 0 & \beta_2 & 0 & \dots & 0 \\ \beta_1 & 0 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \beta_1 & 0 \end{pmatrix} \right)$$

où  $\beta_1 = \alpha_1\tau_{1,2}$  et  $\beta_2 = \alpha_2\tau_{2,2} + \dots + \alpha_p\tau_{p,2}$ . Posons ensuite  $U_0$  comme étant la première colonne de  $M_{\alpha,\beta}$  et  $K$  la matrice constituée des colonnes restantes à laquelle nous rajoutons une colonne constituée uniquement de 0 à droite. L'ensemble des cas pathologiques que nous notons  $\Theta^*$  est formé par la réunion des ensembles suivants  $\{2\beta_1 = 1\}$ ,  $\{\det(I_p - K) = 0\}$ ,  $\{s_0 = 1\}$  ou  $\{U^*(I_p - K)^{-1}U_0 = 1 - s_0\}$ .

#### 4.8.9 Preuve du Théorème 4.10

**Proposition 4.17.** *Supposons que  $\Theta = \Theta_4 \setminus \Theta^*$ . Alors, nous avons la décomposition*

$$\sum_{t=1}^n \Phi_{t-1}Y_t - n\bar{L}_1 = (I_p - K)^{-1}U_0 \left( \sum_{t=1}^n Y_t^2 - n\bar{\ell}_0 \right) + M_n + R_n$$

où  $M_n$  est une  $\mathcal{F}_n^{(2)}$ -martingale et  $\|R_n\| = o(\sqrt{n})$  p.s.

**Preuve.** La proposition découle des Lemmes 4.21 et 4.22. Elles permettent de montrer dès que  $\Theta \cap \Theta^* = \emptyset$ , qu'il existe une décomposition de la forme

$$\frac{1}{n} \sum_{t=1}^n \Phi_{t-1}Y_t = (I_p - K)^{-1}U_0 \frac{1}{n} \sum_{t=1}^n Y_t^2 + \frac{M_n}{n} + \frac{R_n}{n}.$$

Le théorème ergodique nous permet d'en déduire que  $\bar{L}_1 = (I_p - K)^{-1}U_0\bar{\ell}_0$  qui donne la décomposition voulue puisque  $\|R_n\| = o(\sqrt{n})$  p.s.  $\square$

**Proposition 4.18.** *Supposons que  $\Theta = \Theta_4 \setminus \Theta^*$ . Alors, nous avons la décomposition*

$$\sum_{t=1}^n Y_t^2 - n\bar{\ell}_0 = M_n + R_n$$

où  $M_n$  est une  $\mathcal{F}_n^{(2)}$ -martingale et  $R_n = o(\sqrt{n})$  p.s.

**Preuve.** Cela est une conséquence de Lemme 4.25. Dès que  $\Theta \cap \Theta^* = \emptyset$ , nous avons via ce Lemme et sous ces notations que :

$$\sum_{t=1}^n Y_t^2 - n\bar{\ell}_0 = \frac{U^*}{1 - s_0} (\Phi_{t-1}Y_t - n\bar{L}_1) + M_n + R_n$$

Cette décomposition conjuguée à la proposition 4.17 nous permet d'obtenir

$$\left( \sum_{t=1}^n Y_t^2 - n\bar{\ell}_0 \right) \left( 1 - \frac{U^*(I_p - K)^{-1}}{1 - s_0} \right) = \frac{U^*M_n + R_n}{1 - s_0}$$

qui achève cette preuve où nous rappelons que  $R_n$  est terme de reste.  $\square$

Nous revenons maintenant à la preuve du Théorème 4.10. Le point clé consisté à noter qu'il existe une matrice  $\bar{A}$  telle que

$$\sum_{t=1}^n \Phi_{t-1} Y_t - S_{n-1} \theta^* = \bar{A} \sum_{t=1}^n \Phi_{t-1} Y_t - \theta^* \sum_{t=1}^n Y_t^2 + R_n.$$

En divisant par  $n$  et en passant à la limite, nous avons via le Théorème ergodique et la décomposition ci-dessus que  $0 = \bar{L}_1 - \bar{\Lambda}_0 \theta^* = \bar{A} \bar{L}_1 - \theta^* \bar{\ell}_0$ . D'où,

$$\sum_{t=1}^n \Phi_{t-1} Y_t - S_{n-1} \theta^* = \bar{A} \left( \sum_{t=1}^n \Phi_{t-1} Y_t - n \bar{L}_1 \right) - \theta^* \left( \sum_{t=1}^n Y_t^2 - n \bar{\ell}_0 \right) + R_n.$$

La combinaison des Propositions 4.17 et 4.18 nous permet d'en déduire la décomposition suivante

$$\sum_{t=1}^n \Phi_{t-1} Y_t - S_{n-1} \theta^* = M_n + R_n \quad (4.103)$$

où  $M_n$  et  $R_n$  sont des termes génériques dont la signification coïncident avec les preuves précédentes. Comme cela est évoqué dans la Section 4.9, un élément générique  $m_n$  de  $M_n$  se présente

$$m_n = \sum_{t=1}^n Y_{t-d_1}^{a_1} Y_{t-d_2}^{a_2} \eta_{k,t-d_3}^{a_3} \eta_{\ell,t-d_4}^{a_4} \nu_t \quad (4.104)$$

où  $d_i > 0$ ,  $a_i \in \{0, 1\}$  et  $\nu_t$  est une variable aléatoire centrée dont la loi ne dépend pas de  $t$ , indépendante de  $\mathcal{F}_{t-1}^{(2)}$ . Dans ce cas précis la variation quadratique prévisible de ce terme s'écrit

$$\langle m \rangle_n = v_\nu \sum_{t=1}^n Y_{t-d_1}^{a_1} Y_{t-d_2}^{a_2} \eta_{k,t-d_3}^{a_3} \eta_{\ell,t-d_4}^{a_4}$$

où  $v_\nu = \mathbb{E}(\nu_t^2) < \infty$  en vertu de nos hypothèses. La fin de la démonstration est similaire à celles des Théorèmes 4.3 et 4.4. Nous avons dans un premier temps grâce à l'ergodicité et aux hypothèses de moments qu'il existe une matrice  $\bar{L}$  telle que

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \bar{L}.$$

( $Y_t$ ) étant strictement stationnaire et ergodique, la condition de Lindeberg se vérifie composante par composante, sur l'espace paramétrique  $\Theta_4$ , via les mêmes arguments utilisés dans la preuve des Théorèmes 4.3 et 4.4. D'où, pour tout  $\epsilon > 0$ ,

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E} \left( \|\Delta M_t \Delta^2 \mathbb{1}_{\{\|\Delta M_t \Delta \geq \epsilon \sqrt{n}\}} \mid \mathcal{F}_{t-1}^{(2)} \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Nous déduisons donc du Théorème central limite pour les martingales que

$$\frac{M_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \bar{L}).$$

Il ne reste plus qu'à appliquer le Lemme de Slutsky pour achever la preuve.

### 4.8.10 Preuve du Théorème 4.11

Nous revenons à notre  $\mathcal{F}_n^{(2)}$ -martingale  $m_n$  donnée dans (4.104) considérée comme un terme générique de  $M_n$ . En vertu de la loi du logarithme itérée de Hartman-Wintner pour les martingales formulée dans [93], et grâce à notre hypothèse sur  $\Phi_0$  qui garantit que les incréments de  $m_n$  sont strictement stationnaire et ergodique,

$$\limsup_{n \rightarrow +\infty} \frac{m_n}{\sqrt{2n \log \log n}} = v_m \text{ and } \liminf_{n \rightarrow +\infty} \frac{m_n}{\sqrt{2n \log \log n}} = -v_m \text{ p.s.}$$

où  $v_m = \mathbb{E}((\Delta m_1)^2) < +\infty$ , conséquence du fait que nous nous restreignons à  $\Theta_4 \subset \Theta$ . La limite inférieure est obtenue en remplaçant  $m_n$  by  $-m_n$ , qui est une martingale avec les mêmes propriétés. En exploitant ces estimations et généralisant à  $M_n$ , nous pouvons en déduire que

$$\limsup_{n \rightarrow +\infty} \frac{\|M_n\|}{\sqrt{2n \log \log n}} < +\infty \text{ p.s.} \quad (4.105)$$

car la décomposition (4.103) implique  $\|\hat{\theta}_n - \theta^*\| \leq \|S_{n-1}^{-1}\| \|M_n + R_n\|$ , ce qui permet de conclure la preuve *via* (4.105). Nous rappelons que  $n S_{n-1}^{-1}$  converge *p.s.* et que  $\|R_n\| = o(\sqrt{n})$ , qui de fait implique  $o(\sqrt{n \log \log n})$  *p.s.*

## 4.9 Décompositions Martingales intervenant à la Section 4.5

Dans toute cette section  $M_n, \delta_n$  et  $R_n$  sont des termes génériques non nécessairement identiques d'une ligne à l'autre désignant  $\mathcal{F}_n^{(2)}$ -martingales, différences de  $\mathcal{F}_n^{(2)}$ -martingales et termes de reste, respectivement.

**Lemme 4.19.** *Pour tout  $t$  nous avons les décompositions*

- Pour  $k \geq 1$  et  $\ell \gg 2$ ,  $Y_t Y_{t-k} \eta_{\ell,t} = \tau_{\ell,2} Y_{t-\ell} Y_{t-k} + \delta_t$ .
- Pour  $k \geq 1$  et  $\ell \gg 2$ ,  $Y_t Y_{t-k} \eta_{1,t} \eta_{\ell,t} = \delta_t$ .
- $Y_t Y_{t-1} \eta_{1,t}^2 = \tau_{1,2} (Y_{t-1} \langle \theta, \Phi_{t-1} \rangle + \alpha_1 Y_{t-1}^2 \eta_{1,t-1} + \sum_{k=2}^p \alpha_k \tau_{k,2} Y_{t-k} Y_{t-k-1}) + \delta_t$ .
- Pour  $i \geq 2$ ,  $Y_t Y_{t-i} \eta_{i,t-1} \eta_{1,t} = \tau_{1,2} \tau_{i,2} Y_{t-i} Y_{t-i-1} + \delta_t$ .
- $Y_t Y_{t-1} \eta_{1,t-1} \eta_{1,t} = \tau_{1,2} Y_{t-1}^2 \eta_{1,t-1} + \delta_t$ .

**Lemme 4.20.** *Nous avons la décomposition*

$$\sum_{t=1}^n Y_t^2 \eta_{1,t} = \frac{2\tau_{1,2}}{1-2\beta_1} \left( \sum_{k=1}^p \theta_k \sum_{t=1}^n Y_t Y_{t-k+1} + \beta_2 \sum_{t=1}^n Y_t Y_{t-1} \right) + M_n + R_n$$

où  $\beta_1 = \alpha_1 \tau_{1,2}$  et  $\beta_2 = \alpha_2 \tau_{2,2} + \dots + \alpha_p \tau_{p,2}$ .

**Preuve.** En développant  $(Y_t)$  via la relation (4.40) puis en utilisant le Lemme 4.20, nous avons pour tout  $t$  et après simplifications

$$Y_t^2 \eta_{1,t} = 2\tau_{1,2} \left( \sum_{k=1}^p \theta_k Y_t Y_{t-k} + \sum_{k=2}^p \alpha_k \tau_{k,2} Y_{t-k} Y_{t-k-1} + \alpha_1 Y_{t-1}^2 \eta_{1,t-1} \right) + \delta_t.$$

Il ne reste plus qu'à sommer par rapport à  $t$  pour obtenir la décomposition voulue. Nous précisons que  $\{2\beta_1 = 1\} \subset \Theta^*$ .  $\square$

**Lemme 4.21.** *Pour  $i \geq 2$ , nous avons la décomposition*

$$\sum_{t=1}^n Y_t Y_{t-i} = \sum_{k=1}^p \theta_k \sum_{t=1}^n Y_t Y_{t-|i-k|} + \beta_1 \sum_{t=1}^n Y_t Y_{t-|i-2|} + M_n + R_n$$

où  $\beta_1 = \alpha_1 \tau_{1,2}$ .

**Preuve.** De façon analogue, nous développons  $(Y_t)$  via 4.40) puis nous utilisons le Lemme 4.20 pour obtenir, pour tout  $t$ ,

$$Y_t Y_{t-i} = \sum_{k=1}^p \theta_k Y_{t-i} Y_{t-k} + \alpha_1 \tau_{1,2} Y_{t-2} Y_{t-i} + \delta_t.$$

Le résultat découle en sommant par rapport à  $t$ .  $\square$

**Lemme 4.22.** *Nous avons la décomposition*

$$\sum_{t=1}^n Y_t Y_{t-1} = \frac{1}{1-2\beta_1} \left( \sum_{k=1}^p \theta_k \sum_{t=1}^n Y_t Y_{t-k+1} + \beta_2 \sum_{t=1}^n Y_t Y_{t-1} \right) + M_n + R_n$$

où  $\beta_1 = \alpha_1 \tau_{1,2}$  et  $\beta_2 = \alpha_2 \tau_{2,2} + \dots + \alpha_p \tau_{p,2}$ .

**Preuve.** En développant  $(Y_t)$  puis en utilisant le Lemme 4.19, il vient, pour tout  $t$ ,

$$Y_t Y_{t-1} = \sum_{k=1}^p \theta_k Y_{t-1} Y_{t-k} + \alpha_1 Y_{t-1}^2 \eta_{1,t-1} + \sum_{k=2}^p \alpha_k \tau_{k,2} Y_{t-k} Y_{t-k-1} + \delta_t.$$

Nous utilisons ensuite le Lemme 4.20 puis nous sommons sur  $t$  pour achever la démonstration.  $\square$

**Lemme 4.23.** *Nous avons les décompositions suivantes :*

- Pour  $i \geq 2$ ,  $Y_t Y_{t-i} \eta_{1,t}^2 = \tau_{1,2} (Y_{t-i} \langle \theta, \Phi_{t-1} \rangle + \alpha_1 \tau_{1,2} Y_{t-2} Y_{t-i}) + \delta_t$ .
- Pour  $i \geq 2$ ,  $y_t Y_{t-i} \eta_{1,t}^2 \eta_{i,t} = \tau_{1,2} \tau_{i,2} Y_{t-i}^2 + \delta_t$ .
- $Y_t Y_{t-1} \eta_{1,t}^3 = \tau_{1,4} Y_{t-1}^2 + \delta_t$ .
- Pour  $i \geq 2$ ,  $Y_t Y_{t-i} \eta_{i,t-1} \eta_{1,t}^2 = \tau_{1,2} \tau_{i,2} (\alpha_i Y_{t-i}^2 + \theta_1 Y_{t-i} Y_{t-i-1}) + \delta_t$ .
- $Y_t Y_{t-1} = \eta_{1,t-1} \eta_{1,t}^2 = \tau_{1,2} (\tau_{1,2} Y_{t-2} \sum_{k=2}^p \theta_k Y_{t-k} + \theta_1 Y_{t-1}^2 \eta_{1,t-1} + \alpha_1 Y_{t-1}^2 \eta_{1,t-1}^2) + \delta_t$ .
- Pour  $i \geq 2$ ,  $Y_t Y_{t-i} \eta_{i,t-1} = \tau_{i,2} (\theta_1 Y_{t-i} Y_{t-i-1} + \alpha_i Y_{t-i}^2) + \delta_t$ .
- $Y_t Y_{t-1} \eta_{1,t-1} = \theta_1 Y_{t-1}^2 \eta_{1,t-1} + \tau_{1,2} Y_{t-2} \sum_{k=2}^p \theta_k Y_{t-k} + \alpha_1 Y_{t-1}^2 \eta_{1,t-1}^2 + \delta_t$ .

**Lemme 4.24.** *Il existe des constantes  $\gamma_0, \gamma_1, \dots, \gamma_p$  telles que nous avons la décomposition suivante :*

$$\sum_{t=1}^n Y_t^2 \eta_{1,t}^2 = \frac{1}{1-\alpha_1 \beta_1} \left( \sum_{k=1}^p \gamma_k \sum_{t=1}^n Y_t Y_{t-k+1} + \sigma_2 \tau_{1,2} n \right) + M_n + R_n$$

où  $\beta_1 = \alpha_1 \tau_{1,2}$ .

**Preuve.** Nous obtenons que

$$Y_t^2 \eta_{1,t}^2 = Y_t \eta_{1,t} \left( \sum_{k=1}^p \theta_k Y_{t-k} + \sum_{k=1}^p Y_{t-k} \eta_{k,t} + \sum_{k=1}^p \alpha_k Y_{t-k} \eta_{k,t-1} \right) + \eta_{1,t}^2 \epsilon_t^2 + \delta_t.$$

Nous devons ensuite poursuivre les étapes calculatoires en utilisant les Lemmes 4.19 et 4.23. Le Lemme 4.20 nous donne ensuite le résultat. Nous précisons que  $\{\alpha_1 \beta_1 = 1\} \subset \Theta^*$ .  $\square$

**Lemme 4.25.** *Il existe un vecteur déterministe  $U = (u_1, \dots, u_{p-1}, 0)$  telle que nous avons la décomposition suivante :*

$$\sum_{t=1}^n Y_t^2 = \frac{U^*}{1-s_0} \sum_{t=1}^n \Phi_{t-1} Y_t + \left( \bar{\ell}_0 - \frac{U^* \bar{L}_1}{1-s_0} \right) n + M_n + R_n$$

$$\text{où } s_0 = \left( 1 + \alpha_1 \theta_2 + \frac{2\alpha_1 \theta_1^2}{1-2\beta_1} + \frac{\sigma_2}{1-\alpha_1 \beta_1} \right) \tau_{1,2} + \sum_{k=2}^p (1 + \alpha_k)^2 \tau_{k,2}.$$

**Preuve.** Nous commençons par développer  $(Y_t)$  comme dans les autres preuves pour obtenir

$$Y_t^2 = Y_t \left( \sum_{k=1}^p \theta_k Y_{t-k} + \sum_{k=1}^p Y_{t-k} \eta_{k,t} + \sum_{k=1}^p \alpha_k Y_{t-k} \eta_{k,t-1} \right) + \epsilon_t^2 + \delta_t.$$

Nous utilisons ensuite les Lemmes 4.19 et 4.23 afin de simplifier cette somme. Nous achevons ensuite la preuve en utilisant les Lemmes 4.20 et 4.24.  $\square$

## 4.10 Intégrabilité du modèle introduit à la Section 4.5

### 4.10.1 Moments d'ordre deux

Notre objectif est de reformuler la condition  $\mathbb{E}(\eta_{k,t}^a X_t^2) < +\infty$  for  $a \in \{0, 1, 2\}$  sur les paramètres dirigeant le processus. Nous utiliserons à cet égard de manière répétitive la relation  $vec(AXB) = (B^T \otimes A)vec(X)$  et  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ . L'écriture vectorielle (4.41) du processus conjuguée à sa  $\mathcal{F}_t^{(1)}$ -mesurabilité induite par (4.44), permet d'obtenir pour tout  $t$ , en développant et en prenant l'espérance,

$$\begin{aligned} \mathbb{E}(vec(\Phi_t \Phi_t^*)) &= (C_\theta \otimes C_\theta + \Gamma_0) \mathbb{E}(vec(\Phi_{t-1} \Phi_{t-1}^*)) \\ &\quad + (I_p \otimes C_\theta) \mathbb{E}(((N_{t-1} D_\alpha) \otimes I_p) vec(\Phi_{t-1} \Phi_{t-1}^*)) \\ &\quad + (C_\theta \otimes I_p) \mathbb{E}((I_p \otimes (N_{t-1} D_\alpha)) vec(\Phi_{t-1} \Phi_{t-1}^*)) \\ &\quad + \mathbb{E}(((N_{t-1} D_\alpha) \otimes (N_{t-1} D_\alpha)) \vec{\Phi}_{t-1} \Phi_{t-1}^*) + \vec{\Sigma}_0 \end{aligned}$$

où  $\Gamma_0 = \mathbb{E}(N_0 \otimes N_0)$  and  $\Sigma_0 = \mathbb{E}(E_0 E_0^T)$ . De manière compacte,

$$U_t = (C_\theta \otimes C_\theta + \Gamma_0) U_{t-1} + (I_p \otimes C_\theta) V_{1,t-1} + (C_\theta \otimes I_p) V_{2,t-1} + W_{t-1} + vec(\Sigma_0)$$

où  $U_t, V_{1,t}, V_{2,t}$  et  $W_t$  sont des matrices facilement identifiables via la relation ci-dessus. En travaillant sur les autres composantes et en les plaçant dans  $\Omega_t$ , nous obtenons le système linéaire

$$\Omega_t = A_2 \Omega_{t-1} + B_2 \quad (4.106)$$

où  $B_2^T = ((\vec{\Sigma}_0), 0, 0, \Gamma_{\alpha\alpha}(\vec{\Sigma}_0))$ ,

$$A_2 = \begin{pmatrix} C_\theta \otimes C_\theta + \Gamma_0 & I_p \otimes C_\theta & C_\theta \otimes I_p & I_{p^2} \\ G_\alpha \otimes C_\theta + \Gamma_{\alpha c} & \Gamma_\alpha & G_\alpha \otimes I_p & 0 \\ C_\theta \otimes G_\alpha + \Gamma'_{\alpha c} & I_p \otimes G_\alpha & \Gamma'_\alpha & 0 \\ \Gamma_{\alpha\alpha} (C_\theta \otimes C_\theta) + \Lambda_{\alpha\alpha} & \Gamma_{\alpha\alpha} (I_p \otimes C_\theta) & \Gamma_{\alpha\alpha} (C_\theta \otimes I_p) & \Gamma_{\alpha\alpha} \end{pmatrix}$$

où, nous notons,  $G_\alpha = \mathbb{E}(N_0 D_\alpha N_0)$ ,  $\Gamma_\alpha = \mathbb{E}((N_0 D_\alpha) \otimes N_0)$ ,  $\Gamma'_\alpha = \mathbb{E}(N_0 \otimes (N_0 D_\alpha))$ ,  $\Gamma_{\alpha c} = \mathbb{E}((N_0 D_\alpha C_\theta) \otimes N_0)$ ,  $\Gamma'_{\alpha c} = \mathbb{E}(N_0 \otimes (N_0 D_\alpha C_\theta))$ ,  $\Gamma_{\alpha\alpha} = \mathbb{E}((N_0 D_\alpha) \otimes (N_0 D_\alpha))$  and  $\Lambda_{\alpha\alpha} = \mathbb{E}((N_0 D_\alpha N_0) \otimes (N_0 D_\alpha N_0))$ . Grâce à la relation (4.106), nous devons retenir la condition  $\rho(A_2) < 1$  pour assurer les moments d'ordre deux, comme cela est fait dans le Corollaire 2.1 de [79]. Nous précisons que les moments d'ordre quatre de  $(\eta_{k,t})$  et d'ordre deux de  $(\epsilon_t)$  sont requis.

**Remarque 4.10.1** *Let  $\Omega = (I_{4p^2} - A_2)^{-1} B_2$  la solution de (4.106) ayant pour colonnes  $U, V_1, V_2$  and  $W$ . Alors, par ergodicité*

$$\frac{1}{n} \sum_{t=1}^n \Phi_{t-1} \Phi_{t-1}^* \xrightarrow[n \rightarrow \infty]{p.s.} \text{vec}^{-1}(U) \text{ et } \frac{1}{n} \sum_{t=1}^n \Phi_{t-1} \Phi_{t-1}^* D_\alpha N_{t-1} \xrightarrow[n \rightarrow \infty]{p.s.} \text{vec}^{-1}(V_1).$$

*D'où, nous déduisons de (4.40) et (4.45) que  $\theta^* = \theta + \Delta^{-1} \lambda$ , où  $\Delta = \text{vec}^{-1}(U)$  et  $\lambda$  est la première colonne de  $\text{vec}^{-1}(V_1)$ . Bien qu'elle n'admette pas de forme explicite, cette expression de  $\theta^*$  dépend uniquement des paramètres.*

## 4.10.2 Moments d'ordre quatre

Pour reformuler la condition  $\mathbb{E}(\eta_{k,t}^a X_t^4) < +\infty$  pour  $a \in \{0, \dots, 4\}$  uniquement sur sous la forme de contraintes sur les paramètres, la même méthode peut être utilisée. Cela va cependant augmenter de façon considérable le nombre de calculs à traiter. Nous donnons juste les étapes clés. Au lieu de travailler sur  $\mathbb{E}(\text{vec}(\Phi_t \Phi_t^*))$  comme cela est fait pour l'étude des moments d'ordre deux, nous commencer par développer le terme  $\mathbb{E}(\text{vec}(\text{vec}(\Phi_t \Phi_t^*) \text{vec}^*(\Phi_t \Phi_t^*)))$  qui donnera, après des étapes calculatoires, un autre système de la forme

$$\Pi_t = A_4 \Pi_{t-1} + B_4. \quad (4.107)$$

Nous précisons que l'existence des moments d'ordre 8 de  $(\eta_{k,t})$  sont requis et apparaissent donc comme coefficients dans  $A_4$ . Finalement, le condition d'existence des moments d'ordre quatre est  $\rho(A_4) < 1$ , comme cela est fait dans le Corollaire 4.16 pour  $p = 1$ .

## 4.11 Notations additionnelles

Cette section est entièrement dédiée à l'énumération des nombreuses constantes invoquées dans ce chapitre, pour plus de clarté. La première d'entre elle est donnée par

$$\begin{cases} \bar{k}_{(1)} &= (1 + \alpha \theta)^2 \tau_2 + \alpha^2 (\tau_4 - \tau_2^2) \\ \bar{k}_{(1-2)} &= \alpha^2 (1 + \alpha \theta) \tau_2 \\ \bar{k}_{(2)} &= \alpha^4 \tau_2 \\ \bar{k}_{(3)} &= (1 + \alpha^2 \tau_2) \sigma_2 \end{cases} \quad (4.108)$$

et sert à construire la matrice

$$\bar{K} = \begin{pmatrix} \bar{k}_{(1)} & \bar{k}_{(1-2)} & 0 \\ \bar{k}_{(1-2)} & \bar{k}_{(2)} & 0 \\ 0 & 0 & \bar{k}_{(3)} \end{pmatrix}. \quad (4.109)$$

Nous définissons également

$$\begin{cases} k_{(1)} &= \sigma_2 (1 + 4\alpha^2 (\theta^2 \tau_2 - \tau_2^2 + \tau_4)) \\ k_{(1-3)} &= 4\alpha^3 \theta \tau_2 \sigma_2 \\ k_{(2)} &= (1 - 2\alpha \tau_2 + \alpha \theta^2) (2\alpha \tau_4 + \tau_2 (1 - 2\alpha \tau_2 + \alpha \theta^2)) \\ &\quad + \alpha^2 (\tau_6 + 4\theta^2 (\tau_4 - \tau_2^2)) \\ k_{(2-4)} &= 2\alpha^2 \theta \tau_2 (1 + \alpha \theta^2 - 4\alpha \tau_2) + 6\alpha^3 \theta \tau_4 \\ k_{(2-5)} &= \alpha^3 (\alpha \tau_4 + \tau_2 (1 - 2\alpha \tau_2 + \alpha \theta^2)) \\ k_{(2-6)} &= \alpha \sigma_2 (\alpha \tau_4 + \tau_2 (1 - 2\alpha \tau_2 + \alpha \theta^2)) \\ k_{(3)} &= 4\alpha^4 \tau_2 \sigma_2 \\ k_{(4)} &= 4\alpha^4 (\theta^2 \tau_2 - \tau_2^2 + \tau_4) \\ k_{(4-5)} &= 2\alpha^5 \theta \tau_2 \\ k_{(4-6)} &= 2\alpha^3 \theta \tau_2 \sigma_2 \\ k_{(5)} &= \alpha^6 \tau_2 \\ k_{(5-6)} &= \alpha^4 \tau_2 \sigma_2 \\ k_{(6)} &= \alpha^2 \tau_2 \sigma_4 \end{cases} \quad (4.110)$$

que nous présentons sous forme matricielle

$$K = \begin{pmatrix} k_{(1)} & 0 & k_{(1-3)} & 0 & 0 & 0 \\ 0 & k_{(2)} & 0 & k_{(2-4)} & k_{(2-5)} & k_{(2-6)} \\ k_{(1-3)} & 0 & k_{(3)} & 0 & 0 & 0 \\ 0 & k_{(2-4)} & 0 & k_{(4)} & k_{(4-5)} & k_{(4-6)} \\ 0 & k_{(2-5)} & 0 & k_{(4-5)} & k_{(5)} & k_{(5-6)} \\ 0 & k_{(2-6)} & 0 & k_{(4-6)} & k_{(5-6)} & k_{(6)} \end{pmatrix}. \quad (4.111)$$

De plus, nous devons considérer

$$\left\{ \begin{array}{l} \ell'_{(1)} = \sigma_2 \\ \ell_{(1)} = 2\alpha^2 \theta \tau_2 \sigma_2 \\ \ell'_{(2)} = \alpha \theta (\tau_2 (1 - 2\alpha \tau_2 + \alpha \theta^2) - \alpha (2\tau_2^2 - 3\tau_4)) \\ \ell_{(2)} = \alpha \tau_4 + \tau_2 (1 - 2\alpha \tau_2 + \alpha \theta^2) \\ \ell_{(3)} = 2\alpha^3 \tau_2 \sigma_2 \\ \ell'_{(4)} = 2\alpha^3 (\theta^2 \tau_2 - \tau_2^2 + \tau_4) \\ \ell_{(4)} = 2\alpha^2 \theta \tau_2 \\ \ell_{(5)} = \alpha^4 \tau_2 \\ \ell_{(6)} = \alpha \tau_2 \sigma_2 (1 + \alpha) \end{array} \right. \quad (4.112)$$

qui sous forme matricielle conduit à

$$L = \begin{pmatrix} \ell'_{(1)} & \ell_{(1)} & 0 & 0 & 0 & 0 \\ \ell'_{(2)} & \alpha^2 \ell_{(2)} & \theta \ell_{(2)} & \ell_{(2)} & \alpha \ell_{(2)} & \ell_{(2)} \\ \ell_{(3)} & 0 & 0 & 0 & 0 & 0 \\ \ell'_{(4)} & \alpha^2 \ell_{(4)} & \theta \ell_{(4)} & \ell_{(4)} & \alpha \ell_{(4)} & \ell_{(4)} \\ \alpha \theta \ell_{(5)} & \alpha^2 \ell_{(5)} & \theta \ell_{(5)} & \ell_{(5)} & \alpha \ell_{(5)} & \ell_{(5)} \\ \theta \ell_{(6)} & \alpha \ell_{(6)} & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (4.113)$$

Nous terminons enfin par les constantes suivantes,

$$\left\{ \begin{array}{l} m_{(1)} = \sigma_2 (1 + \tau_2 (1 + \alpha^2)) \\ m_{(2)} = \theta^2 (1 + \alpha^2) \tau_2 + (1 - \alpha^2) \tau_2^2 + \alpha^2 \tau_4 \\ m_{(3)} = 2\alpha \theta (1 + \alpha^2) \tau_2 \\ m_{(4)} = \alpha^2 (1 + \alpha^2) \tau_2 \\ m_{(5)} = 2\alpha \theta \tau_2 \\ m_{(6)} = 2\alpha^2 \tau_2. \end{array} \right. \quad (4.114)$$

# Chapitre 5

## Procédure One-Step pour un bruit Gaussien fractionnaire

### 5.1 Historique et motivation

Les résultats présentés dans ce chapitre sont résumés dans [26] et sont le fruit d'un travail mené avec mon directeur de thèse et d'une collaboration avec Eirini Votsi que je tiens à remercier.

Nous nous intéressons dans ce chapitre à l'expérience statistique générée par un bruit Gaussien fractionnaire (fGn) pour le schéma haute-fréquence. Notre objectif est de présenter un nouvel estimateur de l'exposant de Hurst et de la volatilité puis d'étudier leurs propriétés asymptotiques.

Dans une expérience statistique grand échantillon où le pas entre deux observations est constant et l'horizon temps vers l'infini à mesure que la taille de l'échantillon grandit, l'estimation jointe du paramètre de Hurst et de la volatilité fût très étudiée ces dernières décennies. La propriété LAN a été obtenue dans [31] pour un processus Gaussien stationnaire ce qui, en particulier, permet d'obtenir la propriété LAN pour le (fGn). Nous rappelons que l'obtention de la propriété LAN permet de définir une notion d'efficacité asymptotique pour les estimateurs comme cela est décrit dans [48, 60, 69]. L'estimateur du maximum de vraisemblance est naturellement asymptotiquement efficace dans cette expérience statistique comme l'attestent les travaux de Dalhaus [33, 34]. L'estimation par la méthode de Whittle (étant une approximation du maximum de vraisemblance) fût étudiée dans le même contexte et conduit à un estimateur asymptotiquement efficace (voir [43]).

Contrairement au cas grand échantillon, la propriété LAN ne peut être obtenue dans une expérience haute fréquence, où l'horizon est fixé et le pas entre deux observations consécutives tend vers 0. Dans ce cas, l'estimation conjointe du paramètre de Hurst et de la volatilité conduit à un vecteur gaussien limite dégénéré ce qui ne permet pas d'extraire une région de confiance du couple. Récemment, ce problème fut résolu en dans [23] en utilisant des matrices non diagonales comme vitesse. Cette méthode permet d'obtenir d'une part, la propriété LAN dans le cadre d'un fGn pour des observations haute-fréquence et d'autre part, l'estimateur

du maximum de vraisemblance est asymptotiquement efficace. Ce phénomène fût également observé pour le processus symétrique stable où le paramètre d'échelle et de stabilité sont estimés de manière conjointe. La propriété LAN pour ce processus fût obtenue récemment dans [25] en utilisant une méthode similaire à celle présentée dans [23].

L'estimateur de Whittle est également asymptotiquement efficace dans le cas haute-fréquence pour le fGn comme cela est démontré dans [44]. D'autres estimateurs comme les variations quadratiques généralisées [61] et les estimateurs basés sur des méthodes d'ondelettes [42] furent étudiés. Malheureusement, ces estimateurs ne sont pas en générale asymptotiquement efficaces.

Bien que le MLE soit asymptotiquement efficace dans le schéma haute-fréquence, il ne s'exprime pas sous une forme explicite et nous devons donc optimiser numériquement la fonction de vraisemblance afin d'obtenir une estimation. Cette dernière étape s'avère être extrêmement coûteuse sur le plan computationnel et demande donc par conséquent du temps. L'objectif de ce chapitre est bâtir un estimateur moins onéreux sur le plan computationnel et de qualité comparable au MLE. Nous allons donc à cet égard introduire un estimateur du paramètre de Hurst et de la volatilité par la procédure one-step telle que décrite dans [66, 68]. Cette dernière consiste dans un premier temps à faire appel à un estimateur initial des paramètres précités asymptotiquement normal mais non asymptotiquement efficace. Dans un second temps, nous effectuons une itération dite Fisher-scoring que nous présenterons dans une autre section. Pour appliquer la procédure one-step, nous allons considérer comme estimateur initial du paramètre de Hurst et de la volatilité un estimateur basé sur la méthode des variations quadratiques généralisées (QVG). Cette méthode d'estimation est présentée et étudiée dans [61] et fût utilisée récemment dans [75] dans le cadre d'un mouvement Brownien fractionnaire linéaire stable.

L'estimateur induit par la procédure one-step comporte des avantages vis à vis du MLE et des QVG en matière de temps de calculs et de variance asymptotique. Le coût computationnel est bien moins onéreux par rapport au MLE et il possède la même variance asymptotique. D'un autre côté, il est optimal en terme de variance ce qui n'est pas le cas d'un estimateur par QGV. Nous démontrons par la suite que l'estimateur induit par la procédure one-step lorsque nous choisissons un estimateur initial via la méthode des QGV est asymptotiquement normal et possède la même variance que le MLE. Nous illustrons ce résultat par des simulations en comparant notre estimateur one-step au MLE ainsi qu'à l'estimateur initial.

Le chapitre s'articule de la manière suivante : la Section 5.2 décrit le bruit gaussien fractionnaire dans le schéma haute-fréquence et nous rappelons les propriétés asymptotiques du MLE. Dans la Section 5.3, nous présentons l'estimateur induit par la procédure one-step et nous montrons qu'il est asymptotiquement efficace. Nous illustrons les performances de cet estimateur pour des échantillons de taille modérée via des simulations Monte-Carlo.

## 5.2 Bruit gaussien fractionnaire dans le schéma haute-fréquence

Le modèle considéré est le bruit Gaussien fractionnaire i.e.,

$$X_i = \sigma \left( B_H \left( t_i^{(n)} \right) - B_H \left( t_{i-1}^{(n)} \right) \right), \quad i = 1, \dots, n, \quad (5.1)$$

où  $(B_H(t), t \geq 0)$  est un mouvement Brownien fractionnaire et  $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = T$  est un grille régulière, avec  $t_i^{(n)} = \frac{iT}{n} = i\Delta_n$  pour  $0 \leq i \leq n$  et  $\Delta_n = \frac{T}{n}$ . Dans le schéma haute-fréquence (échantillon de remplissage), l'horizon  $T$  est fixé et le pas  $\Delta_n = t_1^{(n)} - t_0^{(n)} \rightarrow 0$  dès lors que la taille de notre échantillon  $n \rightarrow \infty$ . Nous considérons comme cela est mentionné dans l'introduction l'estimation jointe du paramètre de Hurst  $H$  et de la volatilité  $\sigma$ .

La vitesse optimale pour l'estimation du paramètre  $\theta = (H, \sigma)^* \in \Theta \subset (0, 1) \times \mathbb{R}_*^+$ , où  $*$  désigne la transposition, fut obtenue récemment dans [23]. Nous rappelons par ailleurs que la borne inférieure du risque induit par un estimateur est donnée pour tout  $C > 0$  par

$$\liminf_{n \rightarrow \infty} \sup_{|\phi_n(\theta_0)^{-1}(\bar{\theta}_n - \theta)| \leq C} E_{\theta}^n \left[ c \left( \phi_n(\theta_0)^{-1}(\bar{\theta}_n - \theta) \right) \right] \geq \int_{\mathbb{R}^2} c \left( \mathcal{I}_{\phi}(\theta_0)^{-1/2} z \right) \Phi_2(z) dz,$$

pour tout estimateur  $(\bar{\vartheta}_n, n \geq 1)$ , une fonction coût, une suite de matrice  $\phi_n$  et une information de Fisher  $\mathcal{I}_{\phi}(\vartheta_0)$  qui sera décrite ultérieurement. Ici  $\Phi_2$  est la densité d'un vecteur gaussien bidimensionnel standard.

### 5.2.1 Rappels sur le fGn

Le bruit gaussien  $(X_i, i \geq 1)$  fractionnaire défini dans (5.1) est un processus Gaussien stationnaire dont la fonction d'autocovariance est

$$r(k) = \mathbf{E} X_i X_{i+k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} f_{\Delta_n}(x) dx,$$

pour tout  $i, k \in \mathbb{N}$ , avec

$$f_{\Delta_n}(x) = \sigma^2 C_H \Delta_n^{2H} 2(1 - \cos(x)) \sum_{k \in \mathbb{Z}} |x + 2k\pi|^{-2H-1} = \sigma^2 \Delta_n^{2H} f_H(x)$$

où  $C_H = \frac{\Gamma(2H+1) \sin(\pi H)}{2\pi}$ . La fonction d'autocovariance possède l'expression analytique suivante (voir see [89, Section 7.2.3.] )

$$r(k) = \frac{\sigma^2}{2} \Delta_n^{2H} \left( |k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H} \right).$$

En conséquence, pour les observations  $X^{(n)} = (X_1, \dots, X_n)$ , la fonction de vraisemblance en  $\theta = (H, \sigma)^*$  s'écrit de manière explicite

$$\mathcal{L}(\theta, X^{(n)}) = (2\pi)^{-\frac{n}{2}} (\det \Sigma_n)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \langle X^{(n)}, \Sigma_n^{-1} X^{(n)} \rangle \right)$$

où  $\Sigma_n = (r(i-j))_{i,j}$ . Posons

$$A_n = \sqrt{n} \left\{ \frac{1}{n\sigma^2\Delta_n^{2H}} \langle X_n, T_n(H)^{-1} X_n \rangle - 1 \right\}$$

et

$$B_n = \frac{1}{\sqrt{n}} \left\{ \frac{1}{2} \partial_H \log (\det T_n(H)) + \frac{1}{2\sigma^2\Delta_n^{2H}} \langle X_n, \partial_H \{T_n(H)^{-1}\} X_n \rangle \right\},$$

où  $T_n(H) = \sigma^{-2}\Delta_n^{-2H}\Sigma_n$ . Nous rappelons que  $(A_n, B_n)$  converge vers  $(A, B)$  en loi sous  $P_{(H,\sigma)}^n$  vers un vecteur gaussien  $(A, B)$  dont la fonction de covariance est donnée par

$$J(H) = \begin{pmatrix} 2 & -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda \\ -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda & \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda \end{pmatrix}.$$

La fonction de log-vraisemblance est notée  $\ell_n(\vartheta) = \log \mathcal{L}(\vartheta, X^{(n)})$  et le score s'écrit donc

$$\nabla \ell_n(\theta) = \begin{pmatrix} \partial_H \ell_n(\theta) \\ \partial_\sigma \ell_n(\theta) \end{pmatrix} = \begin{pmatrix} A_n \sqrt{n} \log \Delta_n - B_n \sqrt{n} \\ A_n \sqrt{n} \sigma^{-1} \end{pmatrix}. \quad (5.2)$$

Nous allons donc dans la section suivante présenter la vitesse optimale telle que décrite dans [23] qui permet de faire converger le score en loi vers un vecteur gaussien non singulier ce qui n'est pas le cas dans [63] où la propriété LAN obtenue est singulière.

## 5.2.2 Propriétés asymptotiques du MLE

L'estimateur du maximum de vraisemblance  $(\hat{\theta}_n, n \geq 1)$  est défini par

$$\mathcal{L}(\hat{\theta}_n, X^{(n)}) = \sup_{\theta \in \Theta} \mathcal{L}(\theta, X^{(n)}), \quad n \geq 1.$$

Cet estimateur est asymptotiquement normal et asymptotiquement efficace pour différentes vitesses non diagonales comme cela est démontré dans [23]. Pour être précis, lorsque  $n \rightarrow \infty$ ,

$$\phi_n(\theta)^{-1} (\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, \mathcal{I}_\varphi^{-1}(\vartheta)) \quad (5.3)$$

en loi sous  $P_{(H,\sigma)}^n$  où

$$\phi_n(\theta) = \frac{1}{\sqrt{n}} \begin{pmatrix} \alpha_n & \bar{\alpha}_n \\ \beta_n & \bar{\beta}_n \end{pmatrix} \quad (5.4)$$

vérifie les propriétés suivantes

1. le réel  $\omega$  défini par  $\omega = n \det(\varphi_n(\vartheta)) = \alpha_n \bar{\beta}_n - \bar{\alpha}_n \beta_n$  vérifie  $\omega \neq 0$  et ne dépend pas de  $n$ ;
2.  $\alpha_n \rightarrow \alpha$  pour  $\alpha \in \mathbb{R}$ ;
3.  $\bar{\alpha}_n \rightarrow \bar{\alpha}$  pour  $\bar{\alpha} \in \mathbb{R}$ ;
4.  $\gamma_n := \alpha_n \log \Delta_n + \beta_n \sigma^{-1} \rightarrow \gamma$  pour  $\gamma \in \mathbb{R}$ ;

5.  $\bar{\gamma}_n := \bar{\alpha}_n \log \Delta_n + \bar{\beta}_n \sigma^{-1} \rightarrow \bar{\gamma}$  pour  $\bar{\gamma} \in \mathbb{R}$ ;
6.  $\alpha \bar{\gamma} - \bar{\alpha} \gamma \neq 0$ ,

et l'information de Fisher se présente de la manière suivante

$$\mathcal{I}_\phi(\theta) = \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} J(H) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix}. \quad (5.5)$$

En utilisant [23], beaucoup d'exemples peuvent être utilisés. Nous en donnons quelques uns.

1. Une matrice symétrique définie positive  $\varphi_n$  donnée par

$$\phi_n(\theta) = \frac{1}{\sqrt{n}} \begin{pmatrix} -\frac{2}{\log \Delta_n} & \sigma \\ \sigma & -\sigma^2 \log \Delta_n \end{pmatrix},$$

où  $\alpha = 0$ ,  $\bar{\alpha} = \sigma$ ,  $\gamma = -1$ ,  $\bar{\gamma} = 0$  et  $\omega = \sigma^2$ .

2. Deux autres exemples peuvent également être employés. Le premier est donné par

$$\phi_n(\theta) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ -\sigma \log \Delta_n & 1 \end{pmatrix} \quad \text{d'où} \quad \phi_n^{-1}(\theta) = \sqrt{n} \begin{pmatrix} 1 & 0 \\ \sigma \log \Delta_n & 1 \end{pmatrix}, \quad (5.6)$$

qui implique  $\alpha = 1$ ,  $\bar{\alpha} = 0$ ,  $\gamma = 0$ ,  $\bar{\gamma} = \sigma^{-1}$  et  $\omega = 1$ . Le second exemple est

$$\phi_n(\theta) = \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{1}{\log \Delta_n} & 1 \\ 0 & -\sigma \log \Delta_n \end{pmatrix} \quad \text{avec} \quad \phi_n^{-1}(\theta) = -\frac{\sqrt{n}}{\sigma} \begin{pmatrix} -\sigma \log \Delta_n & -1 \\ 0 & \frac{1}{\log \Delta_n} \end{pmatrix}, \quad (5.7)$$

impliquant  $\alpha = 0$ ,  $\bar{\alpha} = 1$ ,  $\gamma = 1$ ,  $\bar{\gamma} = 0$  et  $\omega = -\sigma$ .

Nous rappelons que tous les exemples présentés invoquent à chaque fois des matrices non diagonales et que ces dernières dépendent de  $\sigma$ . D'autres exemples sont formulés dans [23].

### 5.3 Efficacité asymptotique de l'estimateur one-step MLE

La procédure numérique permettant d'évaluer le MLE nécessite le calcul de l'inverse  $T_n(H)^{-1}$  ainsi qu'un algorithme d'optimisation numérique. Cela a pour conséquence que l'évaluation du MLE nécessite un temps de calcul assez important. Il est donc important de proposer des estimateurs moins onéreux en temps de calcul présentant des propriétés asymptotiques similaires à celles du MLE. Nous proposons dans cette optique un estimateur one-step MLE (OSMLE) qui est asymptotiquement efficace mais beaucoup plus rapide à évaluer par rapport au MLE introduit dans [23] et à l'estimateur de Whittle [44]. L'évaluation du OSMLE nécessite une seule itération d'un Fisher scoring ainsi qu'un estimateur initial convenablement choisi.

Considérons la suite de matrices  $\phi_n(\theta)$  définie par (5.4). L'information de Fisher correspondante est décrite par  $\mathcal{I}_\phi(\theta)$  définie par (5.5). Nous supposons que nous disposons d'un estimateur initial asymptotiquement normal à même vitesse que celle du MLE mais non asymptotiquement efficace. Cet estimateur initial que nous notons  $\tilde{\theta}_n = (\tilde{H}_n, \tilde{\sigma}_n)^*$  vérifie donc, lorsque  $n \rightarrow \infty$ ,

$$\phi_n(\theta)^{-1} \left( \tilde{\theta}_n - \theta \right) \longrightarrow \mathcal{N}(0, \Gamma_\varphi(\vartheta))$$

en loi sous  $P_{(H,\sigma)}^n$  pour une matrice symétrique définie positive  $\Gamma_\phi(\theta)$  dont l'inverse est bien entendu supérieur à  $\mathcal{I}_\phi(\theta)^{-1}$  au sens des matrices définies positives. Nous considérons alors l'estimateur  $\bar{\theta}_n = (\bar{H}_n, \bar{\sigma}_n)^*$  induit par la procédure one-step défini par :

$$\bar{\theta}_n = \tilde{\theta}_n + \left( \varphi_n^{-*}(\tilde{\theta}_n) \mathcal{I}_\varphi(\tilde{\theta}_n) \varphi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla \ell_n(\tilde{\theta}_n). \quad (5.8)$$

Nous présentons dans la section suivante un estimateur initial ( $\tilde{\theta}_n, n \geq 1$ ) que nous utiliserons dans la procédure one-step.

### 5.3.1 Estimateur initial via la méthode des variations quadratiques généralisées

Nous considérons les variations quadratiques généralisées suivantes

$$V_{n,1} = \frac{1}{n-1} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \quad \text{et} \quad V_{n,2} = \frac{1}{n-3} \sum_{i=1}^{n-3} (X_{i+3} + X_{i+2} - X_{i+1} - X_i)^2.$$

Des calculs nous conduisent à

$$\begin{aligned} \mathbb{E}(V_{n,1}) &= \frac{1}{n-1} \sum_{i=1}^{n-1} \mathbb{E}((X_{i+1} - X_i)^2) = \mathbb{E}((X_2 - X_1)^2) \\ &= 2(r(0) - r(1)) = \sigma^2 \Delta_n^{2H} (4 - 2^{2H}). \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}(V_{n,2}) &= \frac{1}{n-3} \sum_{i=1}^{n-3} \mathbb{E}((X_{i+3} + X_{i+2} - X_{i+1} - X_i)^2) \\ &= \sigma^2 \Delta_n^{2H} 2^{2H} (4 - 2^{2H}). \end{aligned}$$

Nous déduisons des calculs précédents un estimateur initial ( $\tilde{\theta}_n, n \geq 1$ ) en combinant les deux variations quadratiques généralisées, où  $\tilde{\theta}_n = (\tilde{H}_n, \tilde{\sigma}_n)^*$  est défini par

$$\tilde{H}_n = \frac{1}{2} \log_2 \frac{V_{n,2}}{V_{n,1}}$$

et

$$\tilde{\sigma}_n = \sqrt{\frac{V_{n,1}}{\Delta_n^{2\tilde{H}_n}} \cdot \frac{1}{4 - 2^{2\tilde{H}_n}}}.$$

Le résultat suivant garantit que l'estimateur initial ( $\tilde{\theta}_n, n \geq 1$ ) est asymptotiquement normal à vitesse identique du MLE.

**Théorème 5.1.** *L'estimateur initial basé sur les variations quadratiques généralisées ( $\tilde{\theta}_n, n \geq 1$ ) est asymptotiquement normal, i.e. lorsque  $n \rightarrow +\infty$ ,*

$$\phi_n(\theta)^{-1} \left( \tilde{\theta}_n - \theta \right) \longrightarrow \mathcal{N}(0, \Gamma_\phi(\theta)),$$

*en loi sous  $P_{(H,\sigma)}^n$  où  $\Gamma_\phi(\theta)$  est une matrice symétrique définie positive définie dans (5.11) qui dépend de  $\sigma$  et  $H$ .*

**Preuve.** Voir Section 4. □

Nous précisons que dans, la preuve précédente, l'estimateur initial vérifie un théorème central limite pour tout  $H \in (0, 1)$ . La variation quadratique généralisée standard  $\tilde{V}_{n,1} = \frac{1}{n} \sum_{i=1}^n X_i^2$  vérifie un théorème centrale limite seulement pour  $H < \frac{3}{4}$  comme cela est décrit dans [61] et présente donc un comportement différent. Les résultats précédents peuvent s'étendre à d'autres estimateurs dont la vitesse est optimale (par exemple les variations quadratiques généralisées à un ordre supérieur [61] ou même des estimateurs par ondelettes [42]).

### 5.3.2 Propriétés asymptotiques de l'estimateur one-step MLE

Nous présentons l'un des résultats principaux de ce chapitre relatif à l'estimateur one-step MLE,  $\bar{\theta}_n = (\bar{H}_n, \bar{\sigma}_n)^*$ , défini dans (5.8).

**Théorème 5.2.** *L'estimateur one-step MLE ( $\bar{\theta}_n, n \geq 1$ ) est asymptotiquement normal, i.e., lorsque  $n \rightarrow \infty$ ,*

$$\phi_n(\theta)^{-1} \left( \bar{\theta}_n - \theta \right) \longrightarrow \mathcal{N} \left( 0, \mathcal{I}_\phi^{-1}(\theta) \right).$$

*en loi sous  $P_{(H,\sigma)}^n$  et est asymptotiquement efficace (au sens de Fisher, i.e. avec la même vitesse et surtout le même variance que le MLE).*

**Preuve.** Voir Section 5.5.2. □

### 5.3.3 Simulations

L'objectif de cette section est de comparer numériquement les performances des estimateurs présentés dans les sections antérieures (à savoir le QVG, MLE et OSMLE) pour des échantillons de taille modérée via des simulations Monte-Carlo. L'estimateur one-step MLE présente des avantages par rapport au MLE et au QGV en terme de temps de calcul et de variance asymptotique, respectivement. Premièrement, le temps de calcul est moins élevé par rapport au MLE et il possède la même variance asymptotique. En particulier, l'estimation par le MLE nécessite une procédure d'optimisation numérique itérative (comme la procédure de Fisher scoring présentée dans (5.8)) et un critère de convergence. Par opposition au MLE, le OSMLE qui partage la même variance asymptotique que le MLE nécessite seulement une seule itération de la procédure de Fisher scoring sur une condition initiale convenablement choisie ce qui réduit considérablement le temps de calcul. Deuxièmement, le OSMLE est asymptotiquement efficace et est donc plus performant que le QGV en terme de variance asymptotique.

Dans cette expérience statistique, le score tel que décrit dans (5.2) n'a pas de forme explicite et s'exprime seulement via les quantités  $\partial_H \log \det(T_n(H))$  et  $\partial_H \{T_n(H)^{-1}\}$ . Dans la procédure numérique, les égalités suivantes sont utilisées

$$\partial_H \{T_n(H)^{-1}\} = -T_n(H)^{-1} \partial_H \{T_n(H)\} T_n(H)^{-1},$$

$$\partial_H^2 \{T_H^{-1}\} = -T_H^{-1} \partial \partial_H \{T_H\} T_H^{-1},$$

et

$$\partial_H^2 \log \det(T_n(H)) = \text{Tr}(T_H^{-1} \partial_H \{T_H\}).$$

L'information de Fisher régulière est évaluée numériquement en utilisant la procédure décrite dans [44].

### Erreur statistique de $H$

Afin d'évaluer les erreurs statistiques renormalisées  $\sqrt{n}(\tilde{H}_n - H)$  (QGV),  $\sqrt{n}(\hat{H}_n - H)$  (MLE) et  $\sqrt{n}(\bar{H}_n - H)$  (OSMLE), nous considérons la matrice

$$\mathcal{I}_\vartheta = \begin{pmatrix} \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda & \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda \\ \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda & \frac{2}{\sigma^2} \end{pmatrix}. \quad (5.9)$$

La variance asymptotique du QGV donnée dans (5.11) est calculée via les termes

$$a_\vartheta = -\frac{1}{2 \log 2\sigma^2(4 - 2^{2H})}, \quad b_\vartheta = \frac{1}{2 \log 2\sigma^2 2^{2H}(4 - 2^{2H})},$$

$$c_\vartheta = \frac{4 - 2 \cdot 2^{2H}}{2\sigma(4 - 2^{2H})^2} \quad \text{et} \quad d_\vartheta = \frac{1}{2\sigma(4 - 2^{2H})^2}.$$

Des simulations Monte-Carlo (avec 10000 répétitions) sont réalisées afin de mettre en exergue le comportement des estimateurs (QGV, MLE and OSMLE) pour le paramètre  $H$  sur des échantillons de taille  $n = 256$ . Les résultats sont donnés dans la Figure 5.1 pour  $\sigma = 0.5$  et deux valeurs de  $H$  différentes :  $H = 0.2$  (courte-mémoire) et  $H = 0.8 > \frac{3}{4}$  (longue-mémoire). Les simulations sont réalisées sous  $R$  [98].

L'estimateur QVG est clairement non-optimal en terme de variance asymptotique pour les différents valeurs de  $H$ . La variance théorique donnée dans (5.11) et la variance asymptotique minimale donnée dans (5.9) sont superposées. Nous observons que le MLE a un comportement proche de sa loi limite pour un échantillon de taille modérée. Pour des valeurs de  $H$  grande le MLE présente un léger biais qui s'atténue lorsque nous augmentons la taille de l'échantillon. Le biais qui caractérise nettement le OSMLE pour un échantillon de taille modérée, provient du biais existants sur l'estimateur QGV initial de  $\sigma$  (voir Section 5.3.3) et s'atténue lorsque la taille de l'échantillon augmente. Pour être précis, le biais se transfère sur toute les composantes du OSMLE à cause de la vitesse non diagonale.

## Erreur statistique de $\sigma$

Nous considérons dans cet exemple la matrice  $\varphi_n(\vartheta)$  définie dans (5.7) avec l'information de Fisher régulière correspondante

$$\mathcal{I}_\varphi(\vartheta) = \begin{pmatrix} 2 & \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda \\ \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda & \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda \end{pmatrix}.$$

La variance asymptotique du QGV s'obtient via les termes suivants

$$a_\vartheta = \frac{4 - 2 \cdot 2^{2H}}{2\sigma^2(4 - 2^{2H})^2}, \quad b_\vartheta = \frac{1}{2\sigma^2(4 - 2^{2H})^2},$$

$$c_\vartheta = -\frac{1}{2 \log 2\sigma^2(4 - 2^{2H})} \quad \text{et} \quad d_\vartheta = \frac{1}{2 \log 2\sigma^2 2^{2H}(4 - 2^{2H})}.$$

Par le biais de simulations Monte-Carlo, nous générons 10000 trajectoires du fGn haute-fréquence pour  $n = 256$ ,  $\sigma = 0.5$  et deux valeurs du paramètres de Hurst  $H$  :  $H = 0.2$  et  $H = 0.8$ . La Figure 5.1 affiche la distribution des trois estimateurs de  $\sigma$ , pour des trajectoires générées avec  $H = 0.2$  puis  $H = 0.8$ .

Nous observons de nouveau que l'estimateur QGV de  $\sigma$  n'est pas optimal en terme de variance pour les valeurs de  $H$  considérées et présente un biais bien visible pour un échantillon de taille modérée  $n = 256$  (qui disparaît lorsque  $n$  augmente). Nous observons là encore que le MLE a un comportement proche de sa loi limite pour un échantillon de taille modérée. Pour des valeurs de  $H$  grande le MLE présente de nouveau un léger biais qui s'atténue lorsque nous augmentons la taille de l'échantillon. Le biais qui caractérise nettement le OSMLE pour un échantillon de taille modérée, provient dans ce cas encore du biais existants sur l'estimateur QGV initial de  $\sigma$  (see Section 5.3.3) et s'atténue lorsque la taille de l'échantillon augmente.

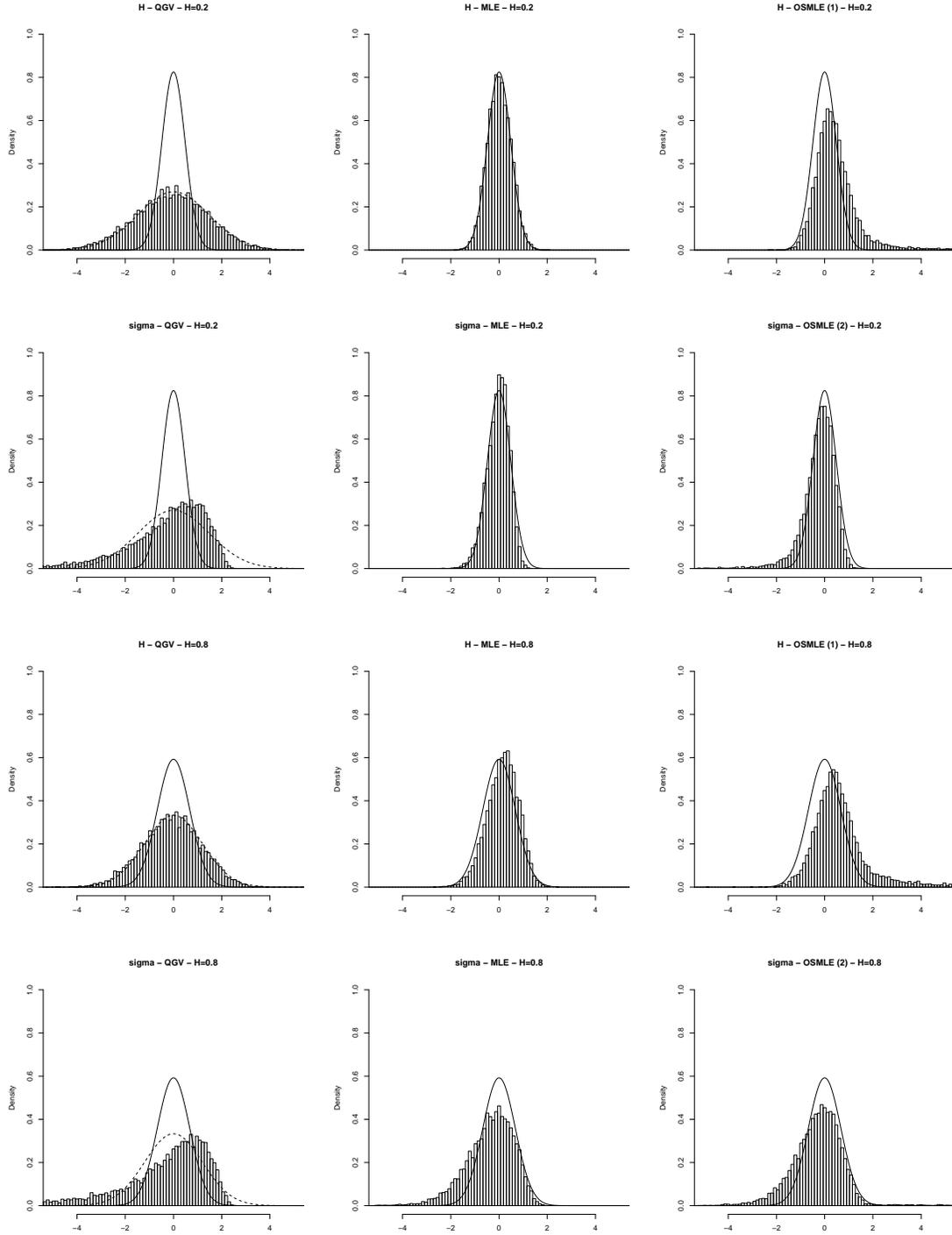


FIGURE 5.1 – Les erreurs statistiques normalisées  $\sqrt{n}(\tilde{H}_n - H)$  et  $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\tilde{\sigma}_n - \sigma)$  (estimateur QGV à gauche),  $\sqrt{n}(\hat{H}_n - H)$  et  $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\hat{\sigma}_n - \sigma)$  (MLE au milieu) et  $\sqrt{n}(\bar{H}_n - H)$  and  $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\bar{\sigma}_n - \sigma)$  (OSMLE à droite) pour  $n = 256$ . Les deux premières lignes présentent 10000 simulations Monte-Carlo dans le cas courte mémoire  $H = 0.2$  et les deux dernières dans le cas longue mémoire  $H = 0.8$ . Nous faisons apparaître la loi induite par la variance optimale (ligne pleine). La loi limite induite par l'estimateur QGV est également représentée (ligne en pointillée).

## Temps de Calcul

Les temps de calcul nécessaires afin de générer des échantillons d'erreurs statistiques sont affichés dans la Table 5.1. Pour  $H = 0.2$  (respectivement  $H = 0.8$ ), la procédure Monte-Carlo pour le MLE (Colonne 2) requiert approximativement 5.6 (respectivement 3.4) fois le temps de calcul pour le OSMLE (Colonne 1). Cela montre que le OSMLE est beaucoup plus avantageux que le MLE en terme de temps de calcul.

	OSMLE	MLE
$H = 0.2$	6280.627	35466.128
$H = 0.8$	5311.633	18231.750

TABLE 5.1 – Temps de calcul (en secondes) pour générer via une procédure Monte-Carlo 10000 réalisations du OSMLE et MLE pour un échantillon de taille  $n = 256$ .

## 5.4 Conclusion et perspectives

Nous venons d'établir dans ce chapitre un nouvel estimateur du paramètre de Hurst et de la variance via la procédure one-step. Ce nouvel estimateur est de qualité comparable au MLE comme l'attestent les simulations. Il serait intéressant d'étudier d'autres estimateurs initiaux afin de corriger le biais présent dans l'estimation de la variance. Il serait également intéressant d'utiliser d'autres variations quadratiques généralisées afin de les comparer numériquement sur des échantillons de taille modérée afin d'espérer résoudre ce problème de biais.

## 5.5 Preuve des résultats principaux

### 5.5.1 Preuve du Théorème 5.1

Posons dans un premier temps

$$f(u, v) = \begin{pmatrix} h(u, v) \\ s(u, v) \end{pmatrix},$$

où  $h(u, v) = \frac{1}{2} \log_2 \left( \frac{v}{u} \right)$  et  $s(u, v) = \sqrt{\frac{u}{w(u, v)}}$ . Nous avons donc

$$\frac{\partial}{\partial u} f(u, v) = \begin{pmatrix} \frac{\partial}{\partial u} h(u, v) \\ -s(u, v) \log \Delta_n \frac{\partial}{\partial u} h(u, v) + \frac{s(u, v) 2^{2h(u, v)} \log 2}{(4 - 2^{2h(u, v)})} \frac{\partial}{\partial u} h(u, v) + \frac{1}{2s(u, v) \Delta_n^{2h(u, v)} (4 - 2^{2h(u, v)})} \end{pmatrix}$$

et

$$\frac{\partial}{\partial v} f(u, v) = \begin{pmatrix} \frac{\partial}{\partial v} h(u, v) \\ -s(u, v) \log \Delta_n \frac{\partial}{\partial v} h(u, v) + \frac{s(u, v) 2^{2h(u, v)} \log 2}{(4 - 2^{2h(u, v)})} \frac{\partial}{\partial v} h(u, v) \end{pmatrix}.$$

Fixons dans un second temps  $u^* = \sigma^2 \Delta_n^{2H} (4 - 2^{2H})$  et  $v^* = \sigma^2 \Delta_n^{2H} 2^{2H} (4 - 2^{2H})$  de sorte que

$$f(u^*, v^*) = \begin{pmatrix} H \\ \sigma \end{pmatrix}.$$

Une développement de Taylor de  $f$  au voisinage de  $(u^*, v^*)$  nous conduit à

$$f(u, v) - f(u^*, v^*) = \frac{\partial}{\partial u} f(u^*, v^*) (u - u^*) + \frac{\partial}{\partial v} f(u^*, v^*) (v - v^*) + o\left(\left\|\begin{pmatrix} u - u^* \\ v - v^* \end{pmatrix}\right\|\right).$$

D'où, pour  $\varphi_n(\vartheta)$  tel que décrit dans (5.6), nous avons

$$\begin{aligned} \phi_n(\theta)^{-1} (f(u, v) - f(u^*, v^*)) &= \frac{\sqrt{n}}{\omega} \begin{pmatrix} \bar{\beta}_n & -\bar{\alpha}_n \\ -\beta_n & \alpha_n \end{pmatrix} (f(u, v) - f(u^*, v^*)) \\ &= \begin{pmatrix} a_{n,\theta} \frac{\sqrt{n}}{\Delta_n^{2H}} (u - u^*) + b_{n,\theta} \frac{\sqrt{n}}{\Delta_n^{2H}} (v - v^*) \\ c_{n,\theta} \frac{\sqrt{n}}{\Delta_n^{2H}} (u - u^*) + d_{n,\theta} \frac{\sqrt{n}}{\Delta_n^{2H}} (v - v^*) \end{pmatrix} \\ &\quad + o\left(\phi_n(\theta)^{-1} \left\|\begin{pmatrix} u - u^* \\ v - v^* \end{pmatrix}\right\|\right). \end{aligned} \quad (5.10)$$

Où nous avons posé

$$\begin{aligned} a_{n,\theta} &= \frac{1}{\omega} \left[ \left( \bar{\beta}_n + \bar{\alpha}_n \sigma \log \Delta_n - \bar{\alpha}_n \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( -\frac{1}{2 \log 2 \sigma^2 (4 - 2^{2H})} \right) - \bar{\alpha}_n \frac{1}{2\sigma(4 - 2^{2H})} \right], \\ b_{n,\theta} &= \frac{1}{\omega} \left( \bar{\beta}_n + \bar{\alpha}_n \sigma \log \Delta_n - \bar{\alpha}_n \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( \frac{1}{2 \log 2 \sigma^2 2^{2H} (4 - 2^{2H})} \right), \\ c_{n,\theta} &= \frac{1}{\omega} \left[ \left( -\beta_n - \alpha_n \sigma \log \Delta_n + \alpha_n \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( -\frac{1}{2 \log 2 \sigma^2 (4 - 2^{2H})} \right) + \alpha_n \frac{1}{2\sigma(4 - 2^{2H})} \right], \end{aligned}$$

et

$$d_{n,\theta} = \frac{1}{\omega} \left( -\beta_n - \alpha_n \sigma \log \Delta_n + \alpha_n \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( \frac{1}{2 \log 2 \sigma^2 2^{2H} (4 - 2^{2H})} \right).$$

Nous avons grâce aux propriétés de la suite  $\phi_n(\theta)$ , que  $a_{n,\theta}$ ,  $b_{n,\theta}$ ,  $c_{n,\theta}$  et  $d_{n,\theta}$  convergent, lorsque  $n \rightarrow \infty$ , vers

$$\begin{aligned} a_\theta &= \frac{1}{\omega} \left[ \left( \bar{\gamma} \sigma - \bar{\alpha} \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( -\frac{1}{2 \log 2 \sigma^2 (4 - 2^{2H})} \right) - \bar{\alpha} \frac{1}{2\sigma(4 - 2^{2H})} \right], \\ b_\theta &= \frac{1}{\omega} \left( \bar{\gamma} \sigma - \bar{\alpha} \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( \frac{1}{2 \log 2 \sigma^2 2^{2H} (4 - 2^{2H})} \right), \\ c_\theta &= \frac{1}{\omega} \left[ \left( -\gamma \sigma + \alpha \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( -\frac{1}{2 \log 2 \sigma^2 (4 - 2^{2H})} \right) + \alpha \frac{1}{2\sigma(4 - 2^{2H})} \right], \end{aligned}$$

et

$$d_\theta = \frac{1}{\omega} \left( -\gamma \sigma + \alpha \frac{\sigma 2^{2H} \log 2}{4 - 2^{2H}} \right) \cdot \left( \frac{1}{2 \log 2 \sigma^2 2^{2H} (4 - 2^{2H})} \right),$$

respectivement. En injectant en  $u = V_{n,1}$  et  $v = V_{n,2}$  dans (5.10), nous obtenons

$$\phi_n(\theta)^{-1} (\tilde{\theta}_n - \theta) = \begin{pmatrix} a_\theta & b_\theta \\ c_\theta & d_\theta \end{pmatrix} \frac{\sqrt{n}}{\Delta_n^{2H}} \begin{pmatrix} V_{n,1} - \mathbb{E}_\theta V_{n,1} \\ V_{n,2} - \mathbb{E}_\theta V_{n,2} \end{pmatrix} + o_{P_\theta^n}(1)$$

comme dans le [99, Théorème 3.1].

Pour tout  $H \in (0, 1)$ , les variations quadratiques généralisées sont asymptotiquement normale (voir [61, Théorème 2 (ii)]), d'où, lorsque  $n \rightarrow \infty$

$$\frac{\sqrt{n}}{\Delta_n^{2H}} \begin{pmatrix} V_{n,1} - \mathbb{E}_\theta V_{n,1} \\ V_{n,2} - \mathbb{E}_\theta V_{n,2} \end{pmatrix} \longrightarrow \mathcal{N}(0, \mathbb{E}(\theta) \Gamma_1(H) \mathbb{E}(\theta))$$

en loi sous  $P_{(H,\sigma)}^n$  avec

$$\mathbb{E}(\theta) = \sigma^2(4 - 2^{2H}) \begin{pmatrix} 1 & 0 \\ 0 & 2^{2H} \end{pmatrix}$$

et

$$\Gamma_1(H) = \begin{pmatrix} \gamma(\mu, \mu) & \gamma(\mu, \nu) \\ \gamma(\nu, \mu) & \gamma(\nu, \nu) \end{pmatrix},$$

où  $\mu = (1, -2, 1)$ ,  $\nu = (1, 0, -2, 0, 1)$  et

$$\gamma(\mu, \nu) = \frac{2 \sum_{j \in \mathbb{Z}} \left( \sum_{k, \ell} \mu_k \nu_\ell |j + k - \ell|^{2H} \right)^2}{\left( \sum_{k, \ell} \mu_k \mu_\ell |k - \ell|^{2H} \right) \left( \sum_{k, \ell} \nu_k \nu_\ell |k - \ell|^{2H} \right)}.$$

Il ne reste plus qu'à appliquer la méthode Delta (telle que décrite dans [99]) afin d'en déduire que notre estimateur initial est bien asymptotiquement normal avec pour matrice de covariance limite

$$\Gamma_\phi(\theta) = \begin{pmatrix} a_\theta & b_\theta \\ c_\theta & d_\theta \end{pmatrix} E(\vartheta) \Gamma_1(H) E(\theta) \begin{pmatrix} a_\theta & b_\theta \\ c_\theta & d_\theta \end{pmatrix}^*. \quad (5.11)$$

### 5.5.2 Preuve du Théorème 5.2

Premièrement, via (5.8), nous avons que

$$\phi_n(\theta)^{-1} (\bar{\theta}_n - \theta) = \phi_n(\theta)^{-1} (\tilde{\theta}_n - \theta) + \phi_n(\theta)^{-1} \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla \ell_n(\tilde{\theta}_n).$$

Soit  $\theta^* \in B(\theta, |\tilde{\theta}_n - \theta|)$ , où  $B(\theta, r)$  est une boule ouverte de rayon  $r$  centrée en  $\theta$ . Un déve-

loppement de Taylor de  $\nabla \ell_n(\tilde{\theta}_n)$  au voisinage de  $\theta$  nous conduit à

$$\begin{aligned}
\phi_n(\theta)^{-1} (\bar{\theta}_n - \theta) &= \phi_n(\theta)^{-1} (\tilde{\theta}_n - \theta) \\
&\quad + \phi_n(\theta)^{-1} \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \left( \nabla \ell_n(\theta) + \nabla^2 \ell_n(\theta^*) (\tilde{\theta}_n - \theta) \right) \\
&= \phi_n(\theta)^{-1} \left( I_2 + \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla^2 \ell_n(\theta^*) \right) (\tilde{\theta}_n - \theta) \\
&\quad + \phi_n(\theta)^{-1} \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla \ell_n(\theta) \\
&= \phi_n(\theta)^{-1} \left( I_2 + \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \phi_n(\theta)^{-*} \phi_n(\theta)^* \nabla^2 \ell_n(\theta^*) \phi_n(\theta) \right. \\
&\quad \left. \times \phi_n(\theta)^{-1} \right) (\tilde{\theta}_n - \theta) + \phi_n(\theta)^{-1} \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla \ell_n(\theta) \\
&= \phi_n(\theta)^{-1} \left( I_2 - \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \phi_n(\theta)^{-*} \mathcal{I}_n(\theta^*) \phi_n(\theta)^{-1} \right) \\
&\quad \times (\tilde{\theta}_n - \theta) + \phi_n(\theta)^{-1} \left( \phi_n(\tilde{\theta}_n)^{-*} \mathcal{I}_\phi(\tilde{\theta}_n) \phi_n(\tilde{\theta}_n)^{-1} \right)^{-1} \nabla \ell_n(\theta). \tag{5.12}
\end{aligned}$$

Nous devons à présent s'assurer que les conditions de Sweeting (formulées dans [96] puis utilisées dans [25] dans une même optique) sont vérifiées, cela impliquera que le tout premier terme de l'équation (5.12) tend vers 0 en probabilités et que le second converge en loi vers un vecteur gaussien centré de matrice de covariance  $\mathcal{I}_\phi^{-1}(\theta)$ . Les conditions de Sweeting sont données par

*C1.* La convergence uniforme

$$\mathcal{I}_n(\theta) = -\phi_n(\theta) \nabla^2 \ell_n(\theta) \phi_n(\theta) \xrightarrow{u} \mathcal{I}_\phi(\theta).$$

Où, nous notons  $\xrightarrow{u}$  la convergence uniforme standard par rapport à  $\theta$  pour tout compact appartenant à  $(0, 1) \times (0, \infty)$ .

*C2.* Contrôle en norme de la matrice  $\phi_n$  de la manière suivante : pour tout  $c > 0$

$$\sup_{\theta' \in \mathfrak{N}_n(c; \theta)} |\phi_n(\theta')^{-1} \phi_n(\theta) - I_2| \xrightarrow{u} 0,$$

où

$$\mathfrak{N}_n(c; \theta) := \{\theta' \in (0, 1) \times (0, \infty); |\phi_n(\theta)^{-1}(\theta' - \theta)| \leq c\}$$

est un voisinage de  $\theta$ , et  $I_k$  désigne la matrice identité de taille  $k \times k$ .

*C3.* Une convergence uniforme de la hessienne  $\nabla_n^2(\theta)$ , donnée par,

$$\sup_{\theta^1, \theta^2 \in \mathfrak{N}_n(c; \theta)} |\phi_n(\theta)^\top \{\nabla^2 \ell_n(\theta^1, \theta^2) - \nabla^2 \ell_n(\theta)\} \phi_n(\theta)| \xrightarrow{u} 0$$

pour tout  $c > 0$ , où

$$\nabla^2 \ell_n(\theta^1, \theta^2) := \begin{pmatrix} \partial_H^2 \ell_n(\theta^1) & \partial_H \partial_\sigma \ell_n(\theta^1) \\ \partial_H \partial_\sigma \ell_n(\theta^2) & \partial_\sigma^2 \ell_n(\theta^2) \end{pmatrix}.$$

Nous allons maintenant établir une par une les conditions de Sweeting.

C1. Cette première condition est démontrée dans le [23, Théorème 3.1].

C2. **Preuve.** Pour des fonctions positives  $a_n(\theta)$  et  $b_n(\theta)$ , nous notons  $a_n(\theta) \lesssim_u b_n(\theta)$  lorsqu'il existe une constante  $C$  telle que  $\sup_{\theta \in K} |a_n(\theta)/b_n(\theta)| \leq C$  pour tout compact  $K$  appartenant à  $(0, 1) \times (0, \infty)$  et  $n$  suffisamment grand. De plus, afin d'alléger les notations, nous écrirons que  $a_n(\theta, \theta') = o_{u,c}(1)$  lorsque  $\sup_{\theta' \in \mathfrak{N}_n(c;\theta)} |a_n(\theta, \theta')| \rightarrow_u 0$ .

Premièrement, nous notons  $\omega(\theta) = \alpha_{n,\theta}\bar{\beta}_{n,\theta} - \beta_{n,\theta}\bar{\alpha}_{n,\theta}$ ,  $\delta_{n,\theta} = \alpha_{n,\theta}\bar{\gamma}_{n,\theta} - \bar{\alpha}_{n,\theta}\bar{\gamma}_{n,\theta}$  et  $\delta(\theta) = \alpha_\theta\bar{\gamma}_\theta - \bar{\alpha}_\theta\bar{\gamma}_\theta$ . Deuxièmement, puisque

$$\phi_n(\theta)^{-1} = \frac{\sqrt{n}}{\omega(\theta)} \begin{pmatrix} \bar{\beta}_{n,\theta} & -\bar{\alpha}_{n,\theta} \\ -\beta_{n,\theta} & \bar{\alpha}_{n,\theta} \end{pmatrix},$$

nous obtenons

$$\phi_n(\theta')^{-1}\phi_n(\theta) = \frac{1}{\omega(\theta')} \begin{pmatrix} \bar{\beta}_{n,\theta'}\alpha_{n,\theta} - \bar{\alpha}_{n,\theta'}\beta_{n,\theta} & \bar{\beta}_{n,\theta'}\bar{\alpha}_{n,\theta} - \bar{\alpha}_{n,\theta'}\beta_{n,\theta} \\ \alpha_{n,\theta'}\beta_{n,\theta} - \beta_{n,\theta'}\alpha_{n,\theta} & \alpha_{n,\theta'}\bar{\beta}_{n,\theta} - \beta_{n,\theta'}\bar{\alpha}_{n,\theta} \end{pmatrix}.$$

Comme  $\beta_{n,\theta} = \sigma\gamma_{n,\theta} - \alpha_{n,\theta}\sigma \log(\Delta_n)$  et  $\bar{\beta}_{n,\theta} = \sigma\bar{\gamma}_{n,\theta} - \bar{\alpha}_{n,\theta}\sigma \log(\Delta_n)$ , nous obtenons via une étape calculatoire que l'élément (1, 2) de la matrice précitée s'écrit :

$$\frac{\bar{\alpha}_{n,\theta}\bar{\alpha}_{n,\theta'} \log(\Delta_n)(\sigma - \sigma')}{\sigma'\delta(\theta')} + o_{u,c}(1).$$

Nous avons de plus que

$$\phi_n(\theta')^{-1}(\theta' - \theta) = A_n(\theta)b_n(\theta, \theta'),$$

où

$$A_n(\theta) = \frac{1}{\sigma\delta_{n,\theta}} \begin{pmatrix} \bar{\gamma}_{n,\theta} & -\bar{\alpha}_{n,\theta} \\ -\gamma_{n,\theta} & \alpha_{n,\theta} \end{pmatrix}$$

et

$$b_n(\theta, \theta') = \begin{pmatrix} \sqrt{n}\sigma(H - H') \\ \sqrt{n}\sigma \log(\Delta_n)(H' - H) + \sqrt{n}(\sigma' - \sigma) \end{pmatrix}.$$

Nous déduisons de la continuité en  $\theta$  que

$$A_n(\theta)^* A_n(\theta) = \frac{1}{\sigma^2\delta^2(\theta)} \begin{pmatrix} \bar{\gamma}_\theta^2 + \gamma_\theta^2 & -(\alpha_\theta\gamma_\theta + \bar{\alpha}_\theta\bar{\gamma}_\theta) \\ \text{sym.} & \alpha_\theta^2 + \bar{\alpha}_\theta^2 \end{pmatrix} + o_{u,c}(1).$$

La plus petite valeur propre de cette matrice est

$$\lambda_\theta^{\min} = \frac{1}{2\sigma^2\delta^2(\theta)} (f(\theta) - \sqrt{f(\theta)^2 - 4g(\theta)}),$$

où  $f(\theta) = \bar{\gamma}_\theta^2 + \gamma_\theta^2 + \bar{\alpha}_\theta^2 + \alpha_\theta^2$  et  $g(\theta) = (\alpha_\theta\gamma_\theta + \bar{\alpha}_\theta\bar{\gamma}_\theta)^2$ . Nous supposons de plus que  $\inf_{\theta \in K} |\delta(\theta)| > 0$ , pour tout compact  $K \subset (0, 1) \times (0, \infty)$ . Donc puisque  $|\lambda_\theta^{\min}|^{-1} \lesssim_u 1$ , nous avons

$$\sup_{\theta' \in \mathfrak{N}_n(c;\theta)} |b_n(\theta, \theta')| \lesssim_u 1,$$

qui implique

$$\sup_{\theta' \in \mathfrak{N}_n(c; \theta)} \left| \frac{\sqrt{n}}{\log(\Delta_n)} (\sigma' - \sigma) \right| \lesssim_u 1,$$

et

$$\sup_{\theta' \in \mathfrak{N}_n(c; \theta)} |\log(\Delta_n)(\sigma' - \sigma)| \rightarrow_u 0.$$

En traitant les autres termes de la matrice  $\phi_n(\theta')^{-1} \phi_n(\theta)$  de manière similaire, nous obtenons que

$$\sup_{\theta' \in \mathfrak{N}_n(c; \theta)} |\phi_n(\theta')^{-1} \phi_n(\theta) - I_2| \rightarrow_u 0. \quad \square$$

**C3. Preuve.** Nous avons via l'inégalité des accroissements finis que

$$\begin{aligned} & \sup_{\theta^1, \theta^2 \in \mathfrak{N}_n(c; \theta)} |\phi_n(\theta)^\top \{\nabla^2 \ell_n(\theta^1, \theta^2) - \nabla^2 \ell_n(\theta)\} \phi_n(\theta)| \\ & \lesssim_u \sup_{\theta^1, \theta^2, \theta' \in \mathfrak{N}_n(c; \theta)} |\phi_n(\theta)^\top \{\partial_\theta^3 \ell_n(\theta^1, \theta^2)[\theta' - \theta]\} \phi_n(\theta)|. \end{aligned} \quad \square$$

La décomposition de  $-\phi_n^\top \partial^3 \ell_n \phi_n$  décrite au sein du [23, Théorème 3.1] nous donne le résultat souhaité.

Les conditions de Sweeting formulées dans (5.12) sont donc vérifiées et nous conduisent à

$$\varphi_n(\vartheta)^{-1} (\bar{\vartheta}_n - \vartheta) \rightarrow \mathcal{N}(0, \mathcal{I}_\varphi^{-1}(\vartheta)).$$

en loi sous  $P_{(H, \sigma)^n}$  lorsque  $n \rightarrow \infty$ .

# Bibliographie

- [1] J. Anděl. Autoregressive series with random parameters. *Math. Operationsforsch. Statist.*, 7-5 :735–741, 1976.
- [2] S. Asmussen and H. Albrecher. *Ruin probabilities*. Chapman & Hall, 2010.
- [3] A. Aït-Sahalia and J. Jacod. Fisher’s information for discretely sampled lévy processes. *Econometrica*, 76(4) :727–761, 2008.
- [4] A. Aue and L. Horváth. Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Stat. Sinica.*, 21 :973–999, 2011.
- [5] A. Aue, L. Horváth, and J. Steinebach. Estimation in random coefficient autoregressive models. *J. Time. Ser. Anal.*, 27-1 :61–76, 2006.
- [6] J. Beran. *Statistics for long-memory processes*. Chapman & Hall, 1994.
- [7] B. Bercu. Central limit theorem and law of iterated logarithm for least squares algorithms in adaptive tracking. *SIAM Journal on Control and Optimization*, 36 :910 – 928, 1998.
- [8] B. Bercu, P. Cénac, and G. Fayolle. On the almost sure central limit theorem for vector martingales : Convergence of moments and statistical applications. *Journal of Applied Probability*, 46(1) :151–169, 2009.
- [9] B. Bercu and F. Proïa. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *ESAIM : Probability and Statistics*, 17 :500–530, 2013.
- [10] I. Berkes, L. Horváth, and P. Kokoszka. Garch processes : structure and estimation. *Bernoulli*, 9(2) :201–227, 04 2003.
- [11] I. Berkes, L. Horváth, and S. Ling. Estimation in nonstationary random coefficient autoregressive models. *J. Time. Ser. Anal.*, 30-4 :395–416, 2009.
- [12] I. Berkes, W. Liu, and W. Wu. Komlòs - Major - Tusnády approximation under dependence. *The Annals of Probability*, 42(2) :794–817, 2014.
- [13] C. Berzin and J. Léon. Estimation in models driven by fractional brownian motion. *Annales de l’Institut Henri Poincaré - Probabilités et Statistiques*, 44(2) :191–213, 2008.
- [14] P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, 12 :788–792, 1961.
- [15] P. Billingsley. *Convergence of Probability Measures, 2nd Edition*. Wiley Series in Probability and Statistics, 1999.

- [16] G. E. P. Box, G.M. Jenkins, and G.C. Reinsel. *Time Series Analysis : Forecasting and Control*. Holden-Day, 1976.
- [17] G. E. P. Box and D. A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332) :1509–1526, 1970.
- [18] A. Brandt. The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients. *Adv. Appl. Probab.*, 18 :211–220, 1986.
- [19] T. Breusch. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31) :334–55, 1978.
- [20] J. Brockwell and R. A. Davis. *Time series : Theory and methods*. 1991.
- [21] A. Brouste, C. Cai, and M. Kleptsyna. Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise. *Mathematical Methods of Statistics*, 23(2) :103–115, Apr 2014.
- [22] A. Brouste, C. Cai, M. Soltane, and L. Wang. Testing for the change of the mean-reverting parameter of an autoregressive model with stationary gaussian noise. *Statistical Inference for Stochastic Process*, 23 :301–318, 2020.
- [23] A. Brouste and M. Fukasawa. Local asymptotic normality property for fractional gaussian noise under high-frequency observations. *The Annals of Statistics*, 46(5) :2045–2061, 2018.
- [24] A. Brouste and M. Kleptsyna. Kalman type filter under stationary noises. *Systems and Control Letters*, 61(12) :1229–1234, 2012.
- [25] A. Brouste and H. Masuda. Efficient estimation of stable lévy process with symmetric jumps. *Statistical Inference for Stochastic Processes*, 21 :289–307, 2018.
- [26] A. Brouste, M. Soltane, and E. Votsi. One-step estimation for the fractional Gaussian noise model at high-frequency. To appear in *ESAIM Probability and Statistics*.
- [27] F. Chaabane and F. Maaouia. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Stat.*, 4 :137–189, 2000.
- [28] S. Choi, W. J. Hall, and A. Schick. Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.*, 24(2) :841–861, 04 1996.
- [29] D. Chong. Crossing probabilities for a square root boundary by a Bessel process. *Comm. Statist. Theory Methods*, 10 :2197–2213, 1981.
- [30] T. Chong. Structural change in AR(1) models. *Econometric Theory*, 17 :87–155, 2001.
- [31] S. Cohen, F. Gamboa, C. Lacaux, and J-M. Loubes. LAN property for some fractional type Brownian motion. *ALEA : Latin American Journal of Probability and Mathematical Statistics*, 10(1) :91–106, 2013.
- [32] H. Cramer and M. Leadbetter. *Stationary and related stochastic processes*. New York, Wiley edition, 1967.
- [33] R. Dahlhaus. Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 17(4) :1749–1766, 1989.
- [34] R. Dahlhaus. Correction efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 34(2) :1045–1047, 2006.

- [35] R. A. Davis, D. Huang, and Y. Yi-Ching. Testing for a change in the parameter values and order of an autoregressive model. *The Annals of Statistics*, 23(1) :282–304, 02 1995.
- [36] H. Douglas, D. Timmer, and J. Pignatiello. The development and evaluation of cusum-based control charts for an ar(1) process. *IIE Transaction in quality and reliability*, 30 :525–534, 1998.
- [37] H. Douglas, D. Timmer, and J. Pignatiello. Change point estimates for the parameters of an ar(1) process. *Quality and Reliability engineering international*, 19 :355–369, 2003.
- [38] R. Drogin. An invariance principle for martingales. *Annals of Mathematical Statistics*, 43(2) :602–620, 1972.
- [39] M. Dufflo. *Random iterative models*, volume 34 of *Applications of Mathematics*, New York. Springer-Verlag, Berlin, 1997.
- [40] J. Durbin. The fitting of time series models. *The fitting of time-series models. Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 28 (3). pp. 233-244. ISSN 03731138, 28, 01 1960.
- [41] E. Eberlein. On strong invariance principles under dependence assumptions. *The Annals of Probability*, 14 :260–270, 1986.
- [42] P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*.
- [43] R. Fox and M. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*, 14(2) :517–532, 1986.
- [44] M. Fukasawa and T. Takabatake. Asymptotically efficient estimators for self-similar stationary gaussian noises under high frequency observations. *Bernoulli*, 25(3) :1870–1900, 2019.
- [45] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6) :933–949, 2018.
- [46] Z. Griliches. A note on serial correlation bias in estimates of distributed lags. *Econometrica*, 29(1) :65–73, 1961.
- [47] X. Guyon and J. Léon. Convergence en loi des h-variations d'un processus gaussien stationnaire sur  $\mathbb{R}$ . *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, B, 25(3) :265–282, 1989.
- [48] J. Hájek. Local asymptotic minimax and admissibility in estimation. *Proc. 6th Berkeley Symp. Math. Statist. Prob.*, 1 :175–194, 1972.
- [49] P. Hall. Martingale invariance principle. *The Annals of Probability*, 5(6) :875–887, 1977.
- [50] P. Hall. On the rate of convergence of normal extremes. *Journal of Applied Probability*, 16(2) :433–439, 1979.
- [51] W. J. Hall and D. J. Mathiason. On large-sample estimation and testing in parametric models. *International Statistical Review / Revue Internationale de Statistique*, 58(1) :77–97, 1990.

- [52] I. Helland. Central limit theorems for martingales with discrete or continuous time. *Scandinavian Journal of Statistics*, 8(2) :79–94, 1982.
- [53] D. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57(1) :1–17, 1970.
- [54] D. Hinkley. Time-ordered classification. *Biometrika*, 59(3) :509–523, 1972.
- [55] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, New-York, 1985.
- [56] L. Horváth. The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of Statistics*, 21 :671–680, 1992.
- [57] L.J. Hosking. Fractional differencing. *Biometrika*, 68(1) :165–176, 1981.
- [58] S. Y. Hwang and I. V. Basawa. Explosive random-coefficient ar(1) processes and related asymptotics for least-squares estimation. *J. Time. Ser. Anal.*, 26-6 :807–824, 2005.
- [59] S. Y. Hwang, I. V. Basawa, and T. Y. Kim. Least squares estimation for critical random coefficient first-order autoregressive processes. *Stat. Probab. Lett.*, 76 :310–317, 2006.
- [60] I. A. Ibragimov and R. Z. K. Has minskii . Statistical estimation : Asymptotic theory. *Springer-Verlag*, 16, 1981.
- [61] J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Annales de l’I.H.P. section B*, 33(4) :407–436, 1997.
- [62] U. Jürgens. The estimation of a random coefficient AR(1) process under moment conditions. *Statist. Hefte.*, 26 :237–249, 1985.
- [63] R. Kawai. Fisher information for fractional brownian motion under high-frequency discrete sampling. *Communications in Statistics - Theory and Methods*, 42(9) :1628–1636, 2013.
- [64] A. Koubková. First-order autoregressive processes with time-dependent random parameters. *Kybernetika.*, 18-5 :408–414, 1982.
- [65] J. Kuelbs and W. Philipp. Almost sure invariance principles for partial sums of mixing b-valued random variables. *The Annals of Probability*, 8 :1003–1036, 1980.
- [66] Y. Kutoyants and A. Motrunich. On multi-step mle-process for markov sequences. 19 :705–724, 2016.
- [67] T.L Lai and C.Z Wei. Asymptotic properties of projections with applications to stochastic regression problems. *Journal of Multivariate Analysis*, 12(3) :346 – 370, 1982.
- [68] L. Le Cam. On the asymptotic theory of estimation and testing hypothesis. *Proceedings of the 3rd Berkeley Symposium*, 1 :355–368, 1956.
- [69] L. Le Cam. Limits of experiments. *Proceedings of the 6th Berkeley Symposium*, 1 :245–261, 1972.
- [70] L. Le Cam and G. Lo Yang. Locally asymptotically normal families. *Springer US*, 978-1-4684-0377-0, pages 52–98, 1990.
- [71] S. Lee, J. Ha, O. Na, and S. Na. The cusum test for parameter change in time series models. *Scandinavian Journal of Statistics*, 30 :781–796, 2003.

- [72] J. Li and Z.C. Hu. Toeplitz lemma, complete convergence, and complete moment convergence. *Communications in Statistics - Theory and Methods*, 46(4) :1731–1743, 2017.
- [73] R. Liptser and A. Shiryaev. Statistics of random processes : I, ii general theory. *2nd Edition, Springer, New York.*, Volumes 5 and 6, 2001.
- [74] E. Malinvaud. Estimation et prévision dans les modèles économiques autoregressifs. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 29(2) :1–32, 1961.
- [75] Otryakhin D. Mazur, S. and M. Podolskij. Estimation of the linear fractional stable motion. *Bernoulli*, 26(1) :226–252, 2020.
- [76] D. McLeish. Dependent central limit theorem and invariance principle. *The Annals of Probability*, 2(4) :620–628, 1974.
- [77] D. McLeish. An extended martingale invariance principle. *The Annals of Probability*, 6(1) :144–150, 1978.
- [78] D. F. Nicholls and B. G. Quinn. The estimation of multivariate random coefficient autoregressive models. *J. Multivar. Anal.*, 11 :544–555, 1981.
- [79] D. F. Nicholls and B. G. Quinn. Multiple autoregressive models with random coefficients. *J. Multivar. Anal.*, 11 :185–198, 1981.
- [80] D. F. Nicholls and B. G. Quinn. *Random Coefficient Autoregressive Models : An Introduction*, volume 11 of *Lecture Notes in Statistics*. Springer-Verlag, New-York, 1982.
- [81] D. Picard. Testing and estimation change-points in time series. *Advances in Applied Probability*, 17(4) :841–867, 1985.
- [82] F. Proïa. Further results on the h-test of Durbin for stable autoregressive processes. *Journal of Multivariate Analysis*, 118 :77 – 101, 2013.
- [83] F. Proïa and M. Soltane. Comments on the presence of serial correlation in the random coefficients of an autoregressive process. To appear in *Statistics and Probability Letters*.
- [84] F. Proïa and M. Soltane. A test of correlation in the random coefficients of an autoregressive process. *Mathematical Methods of Statistics*, 27(2) :119–144, 2018.
- [85] Fred L. Ramsey. Characterization of the partial autocorrelation function. *The Annals of Statistics*, 2(6) :1296–1301, 11 1974.
- [86] P. Robinson. Log-periodogram regression of time series with long-range dependence. *The Annals of Statistics*, 23(3) :1048–1072, 1995.
- [87] P. M. Robinson. Statistical inference for a random coefficient autoregressive model. *Scand. J. Stat.*, 5-3 :163–168, 1978.
- [88] H. Rootzen. On the functional central limit theorem for martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 38 :199–210, 1977.
- [89] G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random Processes*. Chapman & hall edition, 2000.
- [90] A. Schick.  $\sqrt{n}$ -consistent estimation in a random coefficient autoregressive model. *Austral. J. Statist.*, 38-2 :155–160, 1996.

- [91] M. Soltane. Asymptotic efficiency in autoregressive processes driven by stationary gaussian noise. *In Revision*.
- [92] W. Stout. On the functional central limit theorem for martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 38 :199–210, 1977.
- [93] W. F. Stout. The Hartman-Wintner law of the iterated logarithm for martingales. *Ann. Math. Stat.*, 41-6 :2158–2160, 1970.
- [94] W. F. Stout. *Almost sure convergence*, volume 24 of *Probability and Mathematical Statistics*. Academic Press, New-York-London, 1974.
- [95] V. Strassen. Almost sure behavior of sums of independent random variables and martingales. *Proc. Fifth Berkeley Sympos. math. Statist. Probab.*, 2 :315–343, 1965.
- [96] T. Sweeting. Uniform asymptotic normality of the maximum likelihood estimator. *The Annals of Statistics*, 8(6) :1375–1381, 1980.
- [97] M. Taniguchi and Y. Kakizawa. *Asymptotic Theory of Statistical Inference for Time Series*. Springer Series in Statistics. Springer, New-York, 2000.
- [98] R Core Team. R core team (2016) a language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, 2016.
- [99] van der Vaart A. *Asymptotic Statistics*. Cambridge University Press, 1998.
- [100] A. Wald. Asymptotically most powerful tests of statistical hypotheses. *Ann. Math. Statist.*, 12(1) :1–19, 03 1941.

---

**Titre : Statistique asymptotique de certaines séries chronologiques à mémoire**

**Mots clés :** Modélisation chronologique, Normalité locale asymptotique, Modèle autorégressif, Processus longue mémoire, Bruit gaussien fractionnaire

**Résumé :** Cette thèse est dévolue à la statistique inférentielle asymptotique de différents modèles chronologiques dirigés par un bruit comportant de la mémoire. Dans ces modèles, l'estimateur des moindres carrés n'est pas consistant et nous considérons d'autres estimateurs. Nous commençons par étudier les propriétés asymptotiques presque-sûres de l'estimateur du maximum de vraisemblance du coefficient d'autorégression dans un processus autorégressif dirigé par un bruit gaussien stationnaire. Nous présentons ensuite une procédure statistique afin de

détecter un changement de régime au sein de ce modèle en s'inspirant du cas classique dirigé par un bruit blanc fort. Nous abordons ensuite un modèle autorégressif où les coefficients sont aléatoires et possèdent une courte mémoire. Là encore l'estimateur des moindres carrés n'est pas consistant et nous corrigeons l'estimation afin d'estimer correctement les paramètres du modèle. Pour finir nous étudions un nouvel estimateur joint de l'exposant de Hurst et de la variance dans un bruit gaussien fractionnaire observé à haute fréquence dont les qualités sont comparables au maximum de vraisemblance

---

**Title : Asymptotic statistics from some time series models with long memory**

**Keywords :** Time series analysis, Local asymptotic normality, Autoregressive model,

Long memory process, Fractional Gaussian noise

**Abstract :** This thesis is devoted to asymptotic inference of different chronological models driven by a noise with memory. In these models, the least squares estimator is not consistent and we consider other estimators. We begin by studying the almost-sure asymptotic properties of the maximum likelihood estimator of the autoregressive coefficient in an autoregressive process driven by a stationary Gaussian noise. We then present a statistical procedure in order to detect a change of regime within this model, taking inspiration from the classic case driven by a strong white noise. Then we consider an autoregressive model where the coefficients are random and have a short memory. Here again, the least squares estimator is not

consistent and we correct the previous statistic in order to correctly estimate the parameters of the model. Finally we study a new joint estimator of the Hurst exponent and the variance in a fractional Gaussian noise observed at high frequency whose qualities are comparable to the maximum likelihood estimator