

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITÉ DES SCIENCES ET TECHNOLOGIE DE HOUARI BOUMEDIENNE  
FACULTÉ DES MATHÉMATIQUES



Thèse de Doctorat en Sciences  
Présentée pour l'obtention du grade de Docteur  
EN MATHÉMATIQUES

SPÉCIALITÉ : STATISTIQUES ET PROBABILITÉS

Par : Lynda KERAR

thème

Définition, propriétés et estimation pour des  
processus causaux alpha-stables

Soutenue publiquement, le 11/04/2021, devant le jury composé de :

<b>M<sup>me</sup> :H. SAGGOU</b>	Prof	U.S.T.H.B	Présidente
<b>M<sup>me</sup> :K. DJABALLAH</b>	Prof	U.S.T.H.B	Directrice de thèse
<b>M<sup>me</sup> :M. OURBIH</b>	Prof	C.U Tipaza	Examinatrice
<b>M<sup>r</sup> :K. KHALDI</b>	Prof	U.Boumerdes	Examinateur
<b>M<sup>me</sup> :M. CHIKHI</b>	MCA	U.Constantine1	Examinatrice
<b>M<sup>r</sup> :T. MEDKOUR</b>	MCA	U.S.T.H.B	Examinateur
<b>M<sup>r</sup> :J-M. BARDET</b>	Prof	U.Paris I. Sorbonne. France	Invité

# Table des matières

<b>Résumé</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Remerciements</b>	<b>vii</b>
<b>Introduction</b>	<b>ix</b>
<b>1 Modèles de séries chronologiques</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Modèles causaux affines . . . . .	1
1.2.1 Définition des Modèles causaux affines . . . . .	2
1.2.2 Existence et stationnarité . . . . .	3
1.2.3 Estimation par Quasi-Maximum de Vraisemblance des modèles cau- saux affines . . . . .	6
1.2.4 Comportement asymptotique d'EQMV Gaussien . . . . .	8
1.3 Processus conditionnellement hétéroscédastiques généralisés . . . . .	9
1.3.1 Introduction . . . . .	9
1.3.2 Définitions, représentations des modèles GARCH . . . . .	10
1.3.3 Propriétés des trajectoires . . . . .	13
1.3.4 Etude de stationnarité . . . . .	13

1.3.5	Estimation des modèles GARCH par quasi-maximum de vraisemblance gaussien . . . . .	14
<b>2</b>	<b>EQMV du modèle GARCH avec bruit de Student</b>	<b>19</b>
2.1	Introduction . . . . .	19
2.2	Existence et stationnarité . . . . .	24
2.3	L'EQMV d'un processus GARCH à erreur de loi Student . . . . .	25
2.3.1	Définition de l'estimateur . . . . .	25
2.3.2	Comportement asymptotique des estimateurs . . . . .	27
2.3.3	Preuves des principaux résultats . . . . .	29
<b>3</b>	<b>Processus causaux affines avec bruit de Cauchy</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Présentation des lois stables univariées . . . . .	45
3.2.1	Caractérisation des lois stables univariées . . . . .	46
3.2.2	Processus stochastiques de lois stables . . . . .	51
3.2.3	Loi de Cauchy . . . . .	52
3.3	Les processus causaux affines à erreur de loi de Cauchy . . . . .	53
3.3.1	Existence et stationnarité . . . . .	54
3.3.2	Définition de l'estimateur QMV de loi de Cauchy . . . . .	56
3.3.3	Comportement asymptotique d'EQMV de Cauchy . . . . .	59
	<b>Conclusion</b>	<b>69</b>
	<b>Annexe A : Applications numériques</b>	<b>70</b>
3.4	Application aux données simulées . . . . .	70
3.5	Application aux données réelles . . . . .	72

3.5.1	Application aux données S&P500 . . . . .	73
3.5.2	Application aux données CAC 40 . . . . .	76

# Table des figures

1.1	Simulation d'un processus GARCH(1, 1) de taille 500 avec $\omega = 1, \alpha = 0.2, \beta = 0.7$ et $\eta_t \sim N(0, 1)$ . . . . .	12
1.2	Simulation d'un processus GARCH(1, 1) de taille 500 avec $\omega = 1, \alpha = 0.7, \beta = 0.2$ et $\eta_t \sim N(0, 1)$ . . . . .	12
2.1	Simulation de la densité de la loi de Student . . . . .	20
2.2	Simulation d'un processus GARCH(1, 1) à erreurs de Student . . . . .	22
3.1	Simulation de la densité suivant la loi de Cauchy. . . . .	53
3.2	L'indice boursier SP500. . . . .	73
3.3	Rendement logarithmique de l'indice boursier SP500. . . . .	74
3.4	L'indice boursier CAC 40. . . . .	77
3.5	Rendement logarithmique de l'indice boursier CAC 40. . . . .	77

# Résumé

L'estimateur du maximum de vraisemblance non gaussien est fréquemment utilisé dans les modèles GARCH dans le but de capturer des rendements à queue lourde.

L'estimateur est inconsistant en raison d'une mauvaise spécification de la densité. Pour corriger ce biais, nous proposons la méthode du quasi-maximum de vraisemblance dans le cas d'une loi à queue lourde non gaussienne.

Nous établissons la forte consistance et la normalité asymptotique de l'estimateur du quasi-maximum de vraisemblance (EQMV) pour un processus GARCH avec distribution marginale de Student. Nous donnons d'abord une condition nécessaire et suffisante pour l'existence d'une solution strictement stationnaire pour l'équation GARCH. Des simulations numériques ont élaboré la confirmation de la consistance de l'estimation.

Ensuite nous prouvons la consistance forte des modèles causaux affines dans le cas où le bruit suit une loi de Cauchy

**Mots clés :** Normalité asymptotique, consistance, processus GARCH, processus causaux affines.

# Abstract

In this thesis, We establish the strong consistency and asymptotic normality of the quasi-maximum likelihood estimator (QMLE) for a GARCH process with Student marginal distribution. We first give a necessary and sufficient condition for the existence of a strictly stationary solution for the GARCH equation. Numerical simulations elaborated the confirmation of the consistency of estimate. then We prove the consistency of the quasi-maximum estimator of Cauchy (QMLE) for a general class of causal time series including ARMA, AR (1), GARCH, ARCH (1), ARMA-GARCH, APARCH, ARMA-APARCH, .... We present in particular the advantages (order of moment and robustness) of this estimator with respect to the Gaussian QMLE classic.

**Keywords :** Asymptotic normality, consistency, GARCH process.

# Remerciements

Louange à ALLAH, le miséricordieux, sans Lui rien de tout cela n'aurait pu être, je remercie Allah tout puissant qui m'a donné la force, la volonté et le courage pour terminer ce travail de thèse de doctorat, et je le prie de me guider vers la bonne voie.

D'abord, j'adresse mes remerciements à ma chère directrice de thèse, la Professeur Khe-didja Djaballah, qui m'a dirigée tout au long de cette période de préparation de la thèse. Ainsi, je lui suis très reconnaissante pour son soutien, son assertion, son immédiateté et son luxuriance dans ses réponses. Outre ses qualités humaines, je n'oublierai jamais son rôle majeur en tant que professeur. Je voudrais lui dire merci pour son soutien pendant mes périodes de doutes et ses multiples encouragements répétés. Enfin, être une de ses étudiantes est un grand honneur pour moi.

Je tiens à remercier Madame la Professeur Hafida SAGGOU pour l'honneur qu'elle m'a fait en acceptant d'être présidente de mon jury de la thèse. Je tiens à l'assurer de ma profonde reconnaissance pour l'intérêt qu'elle porte à ce travail.

Je remercie vivement madame Megdouda OURBIH professeur à l'université de Tipaza, monsieur Khaled KHALDI professeur à l'université de Boumerdés et monsieur Tarek MED-KOUR Maître de conférence à l'université des sciences et technologies Houari Boumediène , madame Malika CHIKHI Maître de conférence à l'université de Constantine 1 pour l'intérêt qu'ils ont bien voulu porter à ce travail et d'avoir accepté de le juger.

Je remercie Monsieur le Professeur Jean Marc BARDET pour l'intérêt qu'il a manifesté en participant en qualité de membre invité à ce jury, je le remercie de m'avoir accueilli pour des stages dans son laboratoire SAMM à l'université Sorbonne Panthéon durant la

préparation de la thèse.

Je suis infiniment gré à mademoiselle Yasmina MESSACI et madame Sara BENDJEDDOU pour leurs aides précieuses et leurs soutiens scientifiques. J'adresse toute ma gratitude à tous mes amies et à toutes les personnes qui m'ont aidé dans la réalisation de ce travail.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute mon affection et remerciement à ma famille, en particulier, à ma chère maman et à mon cher papa, votre présence et vos encouragements sont pour moi les piliers fondateurs de ce que je suis et de ce que je fais, vous m'avez permis d'effectuer mes recherches dans les meilleures conditions possibles.

A mes frères Ali, Mounir et Salim, à ma belle-sœur Zahra.  
A titre plus personnel, je remercie chaleureusement mon mari, Younes, pour la grande patience, l'encouragement et l'aide dans la rédaction de la thèse. Je tiens à le remercier surtout pour son soutien moral ininterrompu et ses nombreux conseils.  
Je remercie mon beau père qui m'a encouragé à achever ce travail, mes affections à ma belle-mère à qui je souhaite la guérison inchallah.

A ma chère maman et cher papa  
A mon cher mari et notre cher fils Yacine

# Introduction

En modélisation statistique, les lois à variance finie, en particulier la loi normale, sont largement utilisées pour étudier de nombreux phénomènes physiques et des données de nature variée. Les résultats obtenus à partir d'une telle modélisation sont généralement satisfaisants. Par exemple, l'analyse classique des séries chronologiques traite principalement de l'analyse statistique des processus stationnaires et, en particulier, des processus linéaires où les innovations à valeurs réelles sont indépendantes et identiquement distribuées de moyenne nulle et de variance finie. Malheureusement, une telle modélisation devient inefficace lorsque l'on désire étudier des phénomènes présentant de nombreuses valeurs extrêmes, qui ne peuvent être considérées comme des valeurs aberrantes. Les lois stables non-gaussiennes sont une alternative toute naturelle car elles sont une généralisation de la loi normale et prennent en compte des queues lourdes. C'est ainsi que s'est développée l'analyse statistique des processus dans lesquels les innovations sont distribuées à queue lourde tel que des lois  $\alpha$ -stables (loi de student, loi de Cauchy...).

Les séries chronologiques sont des données mesurées sur des intervalles de temps réguliers. Les données macroéconomique et financière sont, généralement déclarées mensuellement, hebdomadairement, quotidiennement et intrajournalièrement. On observe les données aux dates de 1 à  $T$ . Nous considérons les observations  $x_1, \dots, x_T$ , comme réalisations des variables aléatoires  $X_1, \dots, X_T : (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow \mathbb{R}$ ,  $\omega \in \Omega$  tel que  $x_t = X_t(\omega)$ . L'importance des considérations de risque et d'incertitude dans la théorie des séries chronologiques a nécessité le développement de techniques permettant de modéliser la variance et la covariance variant dans le temps. Depuis le papier d'Engle (1982), les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) ont été privilégiés par les économétriciens pour la modélisation des séries financières en temps discret. Plusieurs raisons expliquent le succès de ces modèles. L'un d'eux est certainement qu'ils peuvent être appliqués à des séries de données avec des

queues lourdes. En 1986, Bollerslev a proposé une généralisation de ces modèles ARCH aux modèles autorégressifs conditionnellement hétéroscédastiques généralisés (GARCH), dont la variance conditionnelle au présent dépend de son passé et du passé de processus. Il a étudié les conditions de stationnarité et de la structure des autocorrélations pour cette classe de modèles. L'article de Francq et Zakoian (2004) amène que les modèles GARCH sont basés sur la variance conditionnelle modélisée comme fonction affine du carré des innovations passées. Cette spécification permet de saisir deux caractéristiques importantes de la série financière-succession des périodes de calme et périodes de turbulences, leptokurticité des distributions marginales tout en étant assez simple pour permettre une étude approfondie des propriétés probabilistes et statistiques du modèle. En 2007, Doukhan et Wintenberger ont écrit sous forme générale la classe des processus causaux affines ( $X_t = M_t \varepsilon_t + f_t$ ), qui contient à la fois les processus ARMA ou AR (1), ainsi que les processus GARCH ou ARCH (1), APARCH, ARMA-GARCH et de nombreux autres processus. Concernant l'inférence statistique, plusieurs méthodes d'estimation paramétriques ont été utilisées pour l'estimation des paramètres de ces modèles, parmi lesquelles, on trouve :

La méthode des moindres carrés (MC) : cette méthode consiste à trouver les valeurs des paramètres de modèle qui minimisent la somme de carrés des erreurs  $\sum_{t=1}^n \varepsilon_t^2$  cette méthode a été utilisé par Box et Jenkins (1970) pour les modèles ARMA, Engel (1982) pour les modèles ARCH.

La méthode du maximum de vraisemblance (MV) : cette méthode est la plus utilisée dans l'estimation des paramètres de modèles de séries chronologiques, Box et Jenkins (1970) l'ont utilisée pour l'estimation des modèles ARMA, Engel l'a utilisée pour les modèles ARCH, Bolerselv (1986) pour les modèles GARCH, Bardet et Wintenberger (2009) pour la classe générale des processus causaux affines avec bruits Gaussiens, Bardet, Boularouk et Djaballah (2017) pour la classe générale des processus causaux affine avec bruit de Laplace.

La méthode du maximum de vraisemblance nécessite la connaissance de la loi de probabilité des bruits blancs. Pour l'inférence statistique dans les modèles gaussiens conditionnels, les estimateurs du quasi- maximum de vraisemblance (EQMV) sont assez simples à calculer et ont de bonnes propriétés asymptotiques. Cependant, les processus bruit blanc non

gaussiens sont fréquemment utilisés dans des situations réelles telle que les séries financières. Bollerslev (1986) a étudié les propriétés de l'estimateur quasi-maximum de vraisemblance (EQMV) et des statistiques de test associées dans des modèles dynamiques qui paramètrent conjointement les moyennes conditionnelles et les covariances conditionnelles, lorsqu'une log-vraisemblance normale est maximisée mais l'hypothèse de normalité est ignorée. Francq et Zakoïan (2012, 2013) ont établi les propriétés asymptotiques de l'EQMV standard du vecteur de paramètres, ils ont montré que, si l'ordonnée à l'origine ne peut pas être estimée de manière consistante, l'EQMV des paramètres est consistant et asymptotiquement normal avec ou sans la stationnarité stricte. Des résultats asymptotiques pour les modèles GARCH stationnaires  $(p, q)$  ont été établis pour la première fois dans des conditions faibles par Berkes, Horvath et Kokoszka (2003) et Berkes et Horvath (2004). Ils ont établi la consistance et la normalité asymptotique de l'EQMV dans une famille des modèles GARCH  $(p, q)$  sous des hypothèses faibles sur les paramètres et la distribution du bruit sous-jacente. Straumann et Mikosch (2006) ont fourni une théorie générale de l'EQMV dans les modèles de séries chronologiques hétéroscédastiques et l'ont appliqué à AGARCH  $(p, q)$ . Dans l'article de Bardet et Wintenberger (2009), la forte consistance et la normalité asymptotique de l'estimateur du quasi-maximum vraisemblance est donnée pour une classe générale des processus causaux multidimensionnels. Bardet, Boularouk et Djaballah (2017) ont prouvé la consistance et la normalité asymptotique de l'estimateur du quasi maximum de vraisemblance laplacien (EQMV) pour la même classe générale de séries chronologiques causales. Dans ce travail nous étudierons les propriétés asymptotiques de l'estimateur du quasi-vraisemblance pour la classe des processus GARCH avec un bruit de distribution de Student et la classe des modèles causaux affines à erreur de loi de Cauchy.

La présente thèse comporte trois chapitres théoriques et en annexe des résultats numériques. Dans le premier chapitre nous énonçons certaines notions mathématiques qui s'avèrent nécessaires à la compréhension des chapitres suivants. Dans le deuxième chapitre nous présentons la contribution de la thèse, nous étudions les propriétés asymptotiques de l'estimateur de quasi-vraisemblance pour la classe des processus autoregressive conditionnelement heteroscedastique générale (GARCH) avec bruit de Student au lieu de bruit Gaussien étudié dans la littérature. Tout d'abord, nous définissons la fonction de vraisemblance et l'estimateur de quasi-vraisemblance pour les processus GARCH avec bruit Student ( $t$ -GARCH), puis

nous définissons un espace de stationnarité pour notre classe. Nous prouvons la consistance de l'estimateur du quasi maximum vraisemblance pour la classe des processus t-GARCH, sous condition d'existence des moments de premier ordre qui est une condition plus faible que celle exigée pour le cas des processus avec bruit Gaussien. Comme nous prouvons la normalité asymptotique de notre estimateur, sous condition d'existence des moments d'ordre deux, un résultat qui se trouve une deuxième fois plus performant que celui des processus avec bruits Gaussiens. Par la suite dans le troisième chapitre, nous étudions l'inférence statistique pour le modèle général causal affine à innovations de loi de Cauchy, alors qu'en annexe et dans le but de confirmer les résultats asymptotiques obtenus et de comparer l'EQMV-Student avec l'EQMV-Gauss, nous avons effectué des expériences types Monte-Carlo sur plusieurs modèles de séries chronologiques de différentes tailles et de lois de probabilité du bruit et nous avons effectué également une analyse des données d'une série financière réel (Stock *S&P500*).

# Chapitre 1

## Modèles de séries chronologiques

### 1.1 Introduction

La modélisation représentant une étape cruciale dans l'étude des séries chronologiques. Elle a connu une grande évolution durant ces dernières années et plusieurs modèles de représentation ont été proposés. Ici dans ce chapitre nous allons présenter la classe des modèles causaux affines ARMA ou AR (1), ainsi que les processus GARCH ou ARCH (1), APARCH, ARMA-GARCH et de nombreux autres processus. Nous allons donner l'espace de stationnarité de cette classe et les propriétés asymptotiques de l'estimateur de quasi-vraisemblance avec bruits gaussien, ensuite nous allons traiter un cas particulier de cette classe qui est le modèle autoregressive conditionnellement généralisé (GARCH).

### 1.2 Modèles causaux affines

La sélection des modèles est un outil important pour les statisticiens et tous ceux qui traitent les données. Cette question a reçu une attention considérable dans la littérature récente. Dans cette section nous étudions le problème de sélection des modèles dans une grande classe de modèles de séries chronologiques causales, qui comprend à la fois les processus ARMA ou AR (1), ainsi que les processus GARCH ou ARCH (1), APARCH, ARMA-GARCH et de nombreux autres processus. Nous allons définir la classe générale de séries chronologiques appelées processus causaux affines. Cette classe peut être définie comme suit.

### 1.2.1 Définition des Modèles causaux affines

Soit  $\mathbb{R}^\infty$  l'espace de suite de nombres réels avec un nombre fini de non nul, si  $M, f : \mathbb{R}^\infty \rightarrow \mathbb{R}$  sont deux fonctions mesurables, alors une classe causale affine  $AC(M, f)$  est :

Class  $AC(M, f)$  : Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  appartient à  $AC(M; f)$  s'il satisfait :

$$X_t = M_{\theta_0}(X_{t1}, X_{t2}, \dots)\xi_t + f_{\theta_0}(X_{t1}, X_{t2}, \dots) \text{ pour tout } t \in \mathbb{Z} \quad (1.1)$$

Avec :

- $\theta_0 \in \Theta \subset \mathbb{R}^d, d \in \mathbb{N}^*$ , le vecteur des paramètres inconnus, également appelé le vrai paramètre ;
- $(\xi_t)_{t \in \mathbb{Z}}$  une suite de variables aléatoires indépendantes centrées identiquement distribuées (VAIID) satisfaisant  $E(|\xi_0|^r) < \infty$  avec  $r \geq 1$  et  $E[|\xi_0|] = 1$ .

Si  $r \geq 2$ , notons  $\sigma_t^2 = var(\xi_0)$ ;

- $(\theta, (x_n)_{n \in \mathbb{N}}) \rightarrow M_\theta((x_n)_{n \in \mathbb{N}}) \in (0, \infty)$  et  $(\theta, (x_n)_{n \in \mathbb{N}}) \rightarrow f_\theta((x_n)_{n \in \mathbb{N}}) \in \mathbb{R}$  deux applications connues.

#### Exemples

Si  $M((X_{t-i})_{i \in \mathbb{N}^*}) = \sigma_t$  et  $f((X_{t-i})_{i \in \mathbb{N}^*}) = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p}$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est le processus  $AR(p)$ ;

Si  $M((X_{t-i})_{i \in \mathbb{N}^*}) = \sqrt{a_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$  et  $f((X_{t-i})_{i \in \mathbb{N}^*}) = 0$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est le processus  $ARCH(p)$ .

Si  $M((X_{t-i})_{i \in \mathbb{N}^*}) = \sqrt{\omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2}$  et  $f((X_{t-i})_{i \in \mathbb{N}^*}) = 0$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est le processus  $GARCH(p, q)$ .

De nombreux modèles de séries chronologiques classiques tels que  $ARMA(p; q)$  -  $GARCH(p; q)$  appartiennent à  $AC(M; f)$ . L'existence de solutions stationnaires et ergodiques de cette classe a été étudiée dans Doukhan et Wintenberger(2007).

Dans ce dernier, il a été prouvé que les séries les plus célèbres utilisées en économétrie,

comme les processus ARMA, AR (1), GARCH, ARCH (1), TARCH, ARMA-GARCH peuvent s'écrire comme une solution causale et stationnaire de l'équation (1.1).

Dans Bardet et Wintenberger (2009), il a été établi que, sous plusieurs conditions sur  $M_\theta$ ,  $f_\theta$  et si  $E[|\zeta_0|^r]$  avec  $r \geq 2$ , l'estimateur habituel du quasi-maximum de vraisemblance Gaussien (EQMV) de  $\theta$  est fortement consistant et quand  $r \geq 4$ , il est asymptotiquement normal. Cet estimateur a été défini par Weiss (1986) pour les processus ARCH, et l'étude asymptotique de cet estimateur a été obtenue par Lumsdaine (1996) pour le processus GARCH (1, 1), Berkes et al. (2003) pour le processus GARCH (p, q), Francq et Zakoian (2004) pour les processus ARMA-GARCH, Mikosch et Straumann (2006) pour les modèles hétéroscédastiques généraux, et Robinson et Zaffaroni (2006) pour les processus ARCH (1). Ces résultats adressés aux processus satisfaisant presque partout l'équation (1.1) ainsi que sa généralisation multivariée, fournissent un cadre général et unifié pour les propriétés asymptotiques de l'EQMV Gaussien.

### 1.2.2 Existence et stationnarité

Il existe deux approches pour montrer l'existence et la stationnarité du processus.

- Bougerol et Picard (1992) ont donné des conditions de type Lyapunov pour l'existence d'une solution stationnaire des processus GARCH, qui sont des cas particuliers de (1.1).
- Doukhan et Wintenberger (2007) qui ont prouvé l'existence d'une solution à (1.1) sous une condition de type Lipschitz. Cette condition d'existence donne également la finitude des moments d'un certain ordre en termes de fonctions d'Orlicz, afin de pouvoir travailler avec des moments plus généraux que des moments de puissance. Cette preuve a été établie par Doukhan et Wintenberger (2007) dans leur théorème (3.1) qui donne l'existence d'une solution à (1.1) et que le  $\Phi$ ème moment de cette solution est fini. Nous allons définir cette dernière afin de l'utiliser dans les chapitres suivants.

#### Définitions et hypothèses

**Définition 1.1.** *Espaces Orlicz : Les espaces d'Orlicz sont des généralisations des espaces  $L^m$  classiques. Soit  $\Phi$  une fonction d'Orlicz, c'est-à-dire : définie sur  $\mathbb{R}^+$ , convexe, croissante*

et satisfaisante  $\Phi(0) = 0$ . Pour toute variable aléatoire  $X$  à valeurs dans  $E$ , la norme  $\|X\|_{\Phi}$  est définie par l'équation

$$\|X\|_{\Phi} = \inf \left\{ u > 0, \text{ avec } E\Phi \left( \frac{\|X\|}{u} \right) \leq 1 \right\}.$$

L'espace d'Orlicz  $\mathbb{L}^{\Phi}$  est donné par

$$\mathbb{L}^{\Phi} = \{X \text{ variables aléatoires à valeur dans } E \text{ telles que } \|X\|_{\Phi} < \infty\}$$

Il s'agit d'un espace de Banach muni de la norme  $\|\cdot\|_{\Phi}$ .

Selon Doukhan et Wintenberger (2007) nous avons les hypothèses suivantes :

### Hypothèses

Soit  $(f_{\theta})_{\theta \in \Theta}$  et  $(M_{\theta})_{\theta \in \Theta}$  deux familles de fonctions connues telles que  $\theta \in \Theta$ , où  $f_{\theta}$  et  $M_{\theta}$  prend leurs valeur dans un espace de Banach. Supposons qu'il existe une fonction d'Orlicz  $\Phi$  telle que pour tout  $x, y$  dans  $E^{\infty}$  pour  $\Psi = f_{\theta}$  ou  $M_{\theta}$ , on définit

$$\|\Psi(x; \zeta_0) - \Psi(y; \zeta_0)\|_{\Phi} \leq \sum_{j=1}^{\infty} a_j \|x_j - y_j\|, \quad (1.2)$$

où  $(a_j)_{j \geq 1}$  est une suite de nombres réels non négatifs tels que

$$a = \sum_{j=1}^{\infty} a_j < 1, \quad (1.3)$$

et

$$\mu_{\Phi} = \|\Psi(0, 0, \dots, \zeta_0)\|_{\Phi} < \infty. \quad (1.4)$$

La propriété de Lipschitz de  $\Psi$  et l'hypothèse de moment (1.4) induisent que

$\|\Psi(c; \zeta_0)\|_{\Phi} < \infty$  pour toute constante  $c \in E^{(\infty)}$ . Nous choisissons  $c = (0; 0; \dots)$  dans la condition (1.4) pour plus de commodité.

L'application du théorème (3.1) de Doukhan et Wintenberger (2007) sur les modèles causaux affines générale et sous les conditions Lipschitzienne donne :

$$\|\zeta_0\|_{\Phi} \sum_{i=1}^{\infty} \text{Lip } M_i + \sum_{i=1}^{\infty} \text{Lip } f_i < 1, \quad (1.5)$$

Bardet et Wintenberger (2009), Bardet, Kamila et Kengne (2019), ont adapté les conditions de stationnarité du modèle générale affine au cas gaussien selon l'approche de Doukhan et Wintenberger (2007), les resultats dans ce cas sont donnés ci dessous.

### Stationnarité dans le cas gaussien

#### Notations

Dans la suite, nous considérerons un sous-ensemble  $\Theta$  de  $\mathbb{R}^d$  ( $d \in \mathbb{N}$ ) . Nous utiliserons les normes suivantes :

- Le symbole  $\|\cdot\|$  désigne la norme euclidienne habituelle sur  $\mathbb{R}^{\nu}$ , avec  $\nu \geq 1$  ;
- Si  $X$  est  $\mathbb{R}^{\nu}$  variables aléatoires admet un moment d'ordre  $r$  avec  $r \geq 1$ , nous posons  $\|X\|_r = (E \|X\|^r)^{1/r}$  ;
- Pour tout ensemble  $\Theta \in \mathbb{R}^d$  et pour toute fonction  $g : \Theta \rightarrow \mathbb{R}^d$ ,  $d \geq 1$ , nous définissons  $\|g\|_{\Theta} = \sup_{\theta \in \Theta} \{\|g(\theta)\|\}$ .

Soit  $(f_{\theta})_{\theta \in \Theta}$  et  $(M_{\theta})_{\theta \in \Theta}$  deux familles de fonctions connues telles que  $\theta \in \Theta$ , les deux fonctions prennent des valeurs réelles définies sur  $\mathbb{R}^{\infty}$ .

Nous commenons par donner une condition sur  $f$  et  $M$  assurant l'existence du moment d'ordre  $r$ , de stationnarité et d'ergodicité des séries chronologiques appartenant à  $AC(M_{\theta}; f_{\theta})$ . Cette condition, obtenue initialement par Doukhan et Wintenberger (2007), s'écrit en termes de coefficients de Lipschitz de ces deux fonctions. Par conséquent, pour  $\Psi = f_{\theta}$  ou  $M_{\theta}$ , nous définissons l'hypothèse :

## Hypothèses

Supposons que  $\|\Psi_\theta(0)\|_\Theta < \infty$  et qu'il existe une suite de nombres réels non négatifs  $(\alpha_k(\Psi_\theta, \Theta))_{k \geq 1}$  tel que  $\sum_{k=1}^{\infty} \alpha_k(\Psi_\theta, \Theta) < \infty$  satisfait :

$$\|\Psi_\theta(x) - \Psi_\theta(y)\|_\Theta \leq \sum_{k=1}^{\infty} \alpha_k(\Psi_\theta, \Theta) |x_k - y_k| \text{ pour tout } x, y \in \mathbb{R}^\infty.$$

Afin d'assurer l'existence d'une solution stationnaire d'ordre  $r$  ( $r \geq 1$ ) pour (1.1), définissons l'ensemble

$$\Theta(r) = \left\{ \theta \in \mathbb{R}^d, A(f_\theta, \{\theta\}) \text{ et } A(M_\theta, \{\theta\}) \text{ on a } \sum_{k=1}^{\infty} \alpha_k(f_\theta, \{\theta\}) + \|\zeta_0\|_r \sum_{k=1}^{\infty} \alpha_k(M_\theta, \{\theta\}) < 1 \right\}. \quad (1.6)$$

Alors, pour tout  $\theta \in \Theta(r)$ , il existe une solution stationnaire et ergodique avec un moment d'ordre  $r$  appartenant à  $AC(M_\theta; f_\theta)$ , à partir de Bardet et Wintenberger (2009), nous avons la proposition ci dessous.

**Proposition 1.1.** (Bardet et Wintenberger (2009)) : Si  $\theta_0 \in \Theta(r)$  pour un certain  $r \geq 1$ , alors il existe une unique solution causale de (1.1)  $(X_t)$  indépendant de  $(\zeta_i)_{i>t}$  pour  $t \in \mathbb{Z}$ , qui est stationnaire, ergodique et satisfait  $E \|X\|_0^r < \infty$ .

### 1.2.3 Estimation par Quasi-Maximum de Vraisemblance des modèles causaux affines

Dans la suite, une famille de modèles  $AC(M_\theta; f_\theta)$  avec  $\theta \in \Theta \subset \mathbb{R}^d$ , où  $\theta \rightarrow M_\theta$  et  $\theta \rightarrow f_\theta$  sont deux fonctions fixes, nous allons considérer les estimateurs du quasi-maximum de vraisemblance gaussiens (EQMV). Cette approche comme estimation semi-paramétrique a été introduite pour les processus GARCH (p; q) dans [Jeantheau (1998)] où sa consistance a été également prouvée, et la normalité asymptotique de cet estimateur a été établie dans [Berkes, Horváth et Kokoszka (2003)] et [Francq, et Zakoïan (2004)]. Ces résultats ont été étendus aux processus causaux affines gaussien dans [Bardet et Wintenberger (2009)], ensuite une extension à l'EQMV laplacienne de ces processus a été donnée par [Bardet, Boularouk,

et Djaballah (2017)].

Le QMV gaussien est dérivé de la log-vraisemblance conditionnelle (par rapport à la filtration  $\sigma\{(X_t)_{t \leq 0}\}$ ) de  $(X_1; \dots; X_n)$  lorsque  $(\zeta_t)$  est supposé être un bruit blanc gaussien.

Par définition, si  $X \in \mathcal{AC}(M_\theta; f_\theta)$  alors l'espérance de  $X_{t+1}$  conditionnellement à son propre passé  $X_t, X_{t-1}, \dots$  est égale à  $f_\theta(X_{t-1}, X_{t-2}, \dots) = f_t(\theta)$  et sa variance conditionnelle à  $M_\theta(X_{t-1}, X_{t-2}, \dots)^2 = M_t(\theta)$ . La méthode de Quasi Vraisemblance (QV) consiste à considérer la vraisemblance dans le cas où les innovations  $\varepsilon_t$  sont gaussiennes :

$$L_n(T, \theta) = -\frac{1}{2} \sum_{t \in T} \ell_t(\theta) \text{ avec } \ell_t(\theta) = \frac{(X_t - f_t(\theta))^2}{M_t^2(\theta)} + \log(M_t^2(\theta)).$$

Comme seules  $X_1, \dots, X_n$  sont observées, la fonction  $L_n$  ne peut pas être calculée car elle dépend des valeurs passées inconnues  $X_{-j}$  pour  $j > 0$ . On approche  $L_n$  par la QV  $\hat{L}_n$  définie comme

$$\hat{L}_n(T, \theta) = -\frac{1}{2} \sum_{t \in T} \hat{\ell}_t(\theta) \text{ avec } \hat{\ell}_t(\theta) = \frac{(X_t - \hat{f}_t(\theta))^2}{\hat{M}_t^2(\theta)} + \log(\hat{M}_t^2(\theta)),$$

avec  $\hat{f}_t(\theta) = f_\theta(X_{t-1}, \dots, X_1, u)$  et  $\hat{M}_t^2(\theta) = M_\theta^2(X_{t-1}, \dots, X_1, u)$  pour toute suite de valeurs initiales  $u = (u_n)$  avec un nombre fini de valeurs non nulles.

Un estimateur quasi-maximum de vraisemblance (EQMV) est défini par :

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \log \hat{\ell}_n(X_1, \dots, X_n, \theta).$$

Bardet et Wintenberger ont étudié les propriétés asymptotiques dans le cas gaussien, ils ont prouvés la forte consistance de l'estimateur et sa normalité asymptotique. Leurs résultats sont donnés ci dessous.

### Hypothèses nécessaires pour la convergence de l'EQMV

Cela nécessite quelques hypothèses supplémentaires sur  $\Theta$  et les fonctions  $f$  et  $M$  :

- Hypothèse H1 (Compacité) :  $\Theta$  est un ensemble compact.

- Hypothèse H2 (Borne inférieure de la variance conditionnelle) : il existe une constante  $\underline{M} > 0$  tel que  $\forall \theta \in \Theta$ , alors  $M_\theta(x) > \underline{M}$  pour tout  $x \in \mathbb{R}^{\mathbb{N}}$ .
- Hypothèse H3 (Identifiabilité) : Les fonctions  $M_\theta$  et  $f_\theta$  sont telles que : pour tout  $\theta_1, \theta_2 \in \Theta$ , alors  $M_{\theta_1} = M_{\theta_2}$  et  $f_{\theta_1} = f_{\theta_2}$  implique que  $\theta_1 = \theta_2$ .
- Hypothèse H4 : Une des familles  $(\partial f_{\theta_0}^t / \partial \theta_i)_{1 \leq i \leq d}$  ou  $(\partial M_{\theta_0}^t / \partial \theta_i)_{1 \leq i \leq d}$  est p.s linéairement indépendante, où

$$\frac{\partial f_{\theta_0}^t}{\partial \theta} = \frac{\partial f_\theta}{\partial \theta}(X_{t-1}, \dots) \text{ et } \frac{\partial M_{\theta_0}^t}{\partial \theta} = \frac{\partial M_\theta}{\partial \theta}(X_{t-1}, \dots)$$

### 1.2.4 Comportement asymptotique d'EQMV Gaussien

La quasi vraisemblance est une approximation de la vraisemblance obtenue en remplaçant  $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots)$  et  $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots)$  par  $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots, \mu)$  et  $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots, \mu)$  où  $\mu = (u_n)_{n \geq 0}$  est égale à zéro. Le lemme suivant, déjà prouvé dans Bardet et Wintenberger (2009) nous permet d'avoir une estimation de l'erreur d'approximation, il représente une étape cruciale dans la preuve de la consistance de l'EQMV.

**Lemme 1.1.** *Supposons que  $\theta_0 \in \Theta(r)$  pour  $r \geq 1$  et que  $X$  est la solution causale et stationnaire de (1.1). Si  $(A_0(\Psi, \Theta))$  est satisfaite, alors  $\Psi_\theta^t \in L^r(C(\Theta, \mathbb{R}^d))$  et il existe une constante  $C > 0$  indépendante de  $t$  tel que :*

$$E \left[ \left\| \hat{\Psi}_t^2 - \Psi_t^2 \right\|_{\Theta}^r \right] \leq CE [\|X_0\|^r] \left( \left\| \sum_{j \geq t} \alpha_j(\Psi) \right\|_{\Theta} \right)^r \text{ pour tout } t \in \mathbb{N}^* .$$

#### Consistance forte

**Théorème 1.1.** *(Bardet et wintenberger(2009)) Supposons que les hypothèses H1, H2 et H3 sont vérifiées et que  $\theta_0 \in \Theta(r)$  avec  $r \geq 1$ . Soit  $X$  une solution stationnaire de (1.1).*

*Si  $(A_0(f, \Theta))$  et  $(A_0(M, \Theta))$  tel que :*

$$\alpha_j^0(f, \Theta) + \alpha_j^0(M, \Theta) = O(j^{-\ell}) \text{ pour certain } \ell > 3/2$$

alors,

l'EQMV  $\hat{\theta}_n$  converge fortement, c'est-à-dire

$$\hat{\theta}_n \rightarrow \theta_0 \text{ p.s.n } \rightarrow \infty$$

## Normalit asymptotique

**Théorème 1.2.** (Bardet et wintenberger(2009)) Supposons que  $\theta_0 \in \Theta(4) \cap \overset{\circ}{\Theta}$  l'intérieur de  $\Theta$ , et soit  $X$  la solution stationnaire de (1.1). Sous les hypothèses du théorème 1 et  $H4$ , si, pour  $i = 1, 2$ ,  $(A_i(f, \Theta))$  et  $(A_i(M, \Theta))$  tel que

$$\alpha_j^{(1)}(f, \Theta) + \alpha_j^{(1)}(M, \Theta) = O(j^{-\ell}) \text{ pour certain } \ell > 3/2$$

alors le EQMV  $\hat{\theta}_n$  est asymptotiquement normal, i.e.

$$\sqrt{n} \left( \hat{\theta}_n - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{D} \mathcal{N} \left( 0, F(\theta_0)^{-1} G(\theta_0) F(\theta_0)^{-1} \right),$$

où

$$(F(\theta_0))_{ij} = E \left[ \partial^2 q_0(\theta_0) / \partial \theta_i \partial \theta_j \right] \text{ et } (G(\theta_0))_{ij} = E \left[ \frac{\partial q_i(\theta_0)}{\partial \theta_i} \frac{\partial q_i(\theta_0)}{\partial \theta_j} \right]$$

Pour plus de détails voir Bardet et Wintenberger (2009).

## 1.3 Processus conditionnellement hétéroscédastiques généralisés

### 1.3.1 Introduction

Il existe des séries temporelles, particulièrement dans le domaine de la finance (par exemple les variations des prix du pétrole, des indices boursiers *S&P500*) dont la modélisation des différentes caractéristiques est difficile à obtenir par des modèles linéaires, car ces modèles

sont incapables de capter toutes les asymétries cycliques et les variations instantanées, alors un grand intérêt est accordé aux spécifications non linéaires, ces modèles introduisent une distinction significative entre les phases d'expansion et les phases de récession, ils sont alors suffisamment flexibles et permettent de tenir compte des différentes spécifications et des relations correspondantes à chaque phase.

Les modèles ARCH (autorégressifs conditionnellement hétéroscédastiques) ont été introduits par Engle (1982) et leur extension GARCH (ARCH généralisés) est due à Bollerslev (1986). Leur caractérisation repose essentiellement sur le concept de variance conditionnelle. Dans ces modèles, celle-ci s'écrit comme une fonction affine des valeurs passées du carré de la série. Cette spécification particulière se révèle très fructueuse car elle permet une étude complète des propriétés des solutions tout en étant assez générale. Les modèles GARCH sont en effet susceptibles de capter les propriétés caractéristiques de certaines séries telle que les séries financières.

### 1.3.2 Définitions, représentations des modèles GARCH

Un processus  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est dit processus conditionnelement hétéroscédastique généralisé, s'il est solution de l'équation (1.1) avec les valeurs

$M((X_{t-i})_{i \in \mathbb{N}^*}) = \sqrt{\omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2}$  et  $f((X_{t-i})_{i \in \mathbb{N}^*}) = 0$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est le processus GARCH( $p, q$ ).

Donc l'écriture du modèle GARCH( $p, q$ ) est donnée dans la définition ci-dessous.

**Définition 1.2. (Processus GARCH ( $p, q$ ))** Soit  $(\eta_t)$  une suite iid. On dit que  $(X_t)$  est un processus GARCH ( $p, q$ ) s'il vérifie

$$\begin{cases} X_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z}, \end{cases} \quad (1.7)$$

où  $\omega$ ,  $\alpha_i$  et  $\beta_j$  sont les paramètres du modèle, avec  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  et  $\omega > 0$ .

Notons que cette écriture fait dépendre  $\sigma_t$  des valeurs passées de  $X_t^2$  et de ses propres valeurs passées. avec  $\sigma_t^2$  est mesurable par rapport à la tribu  $\sigma(X_u, u < t)$ . Il est bien sûr nécessaire d'assurer la positivité de  $\sigma_t$ , on impose généralement que le paramètre  $\theta$  du modèle vérifie

$\theta = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q) \in \mathbb{R}_+^* \times \mathbb{R}_+^{p+q}$  tel que  $p \geq 0; q \geq 0; \omega > 0; \alpha_i \geq 0;$   
 $i = 1, \dots, p; \beta_j \geq 0;$  où  $\alpha_i$  et  $\beta_j$  sont des constantes non négatives et  $\omega$  est une constante (strictement) positive.  $j = 1, \dots, q;$  avec  $p$  et  $q$  représente l'ordre du processus *GARCH*.

En remplaçant  $X_{t-i}$  par  $\sigma_{t-i}\eta_{t-i}$  dans (1.7), nous obtenons

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \sigma_{t-i}^2 \eta_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (1.8)$$

que l'on peut écrire :

$$\sigma_t^2 = \omega + \sum_{i=1}^r a_i(\eta_{t-i}) \sigma_{t-i}^2, \quad (1.9)$$

où  $a_i(z) = \alpha_i z^2 + \beta_i, i = 1, \dots, r$ . Cette représentation montre que dans le cas d'un GARCH, le processus de volatilité vérifie une équation autorégressive, mais avec coefficients aléatoires.

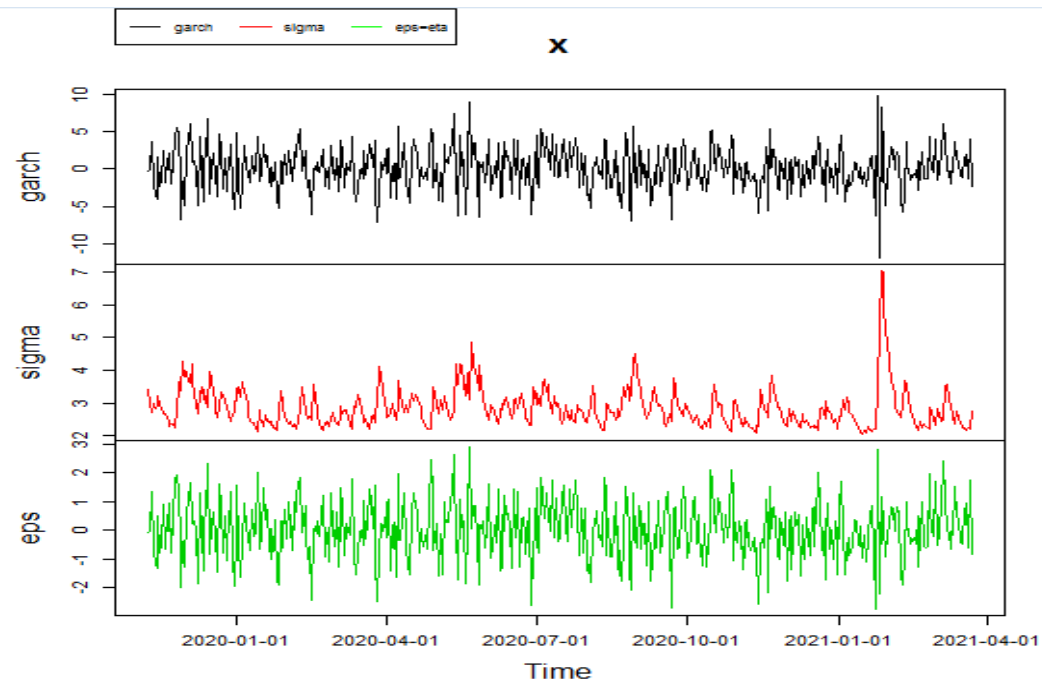


FIGURE 1.1 – Simulation d’un processus GARCH(1, 1) de taille 500 avec  $\omega = 1, \alpha = 0.2, \beta = 0.7$  et  $\eta_t \sim N(0, 1)$ .

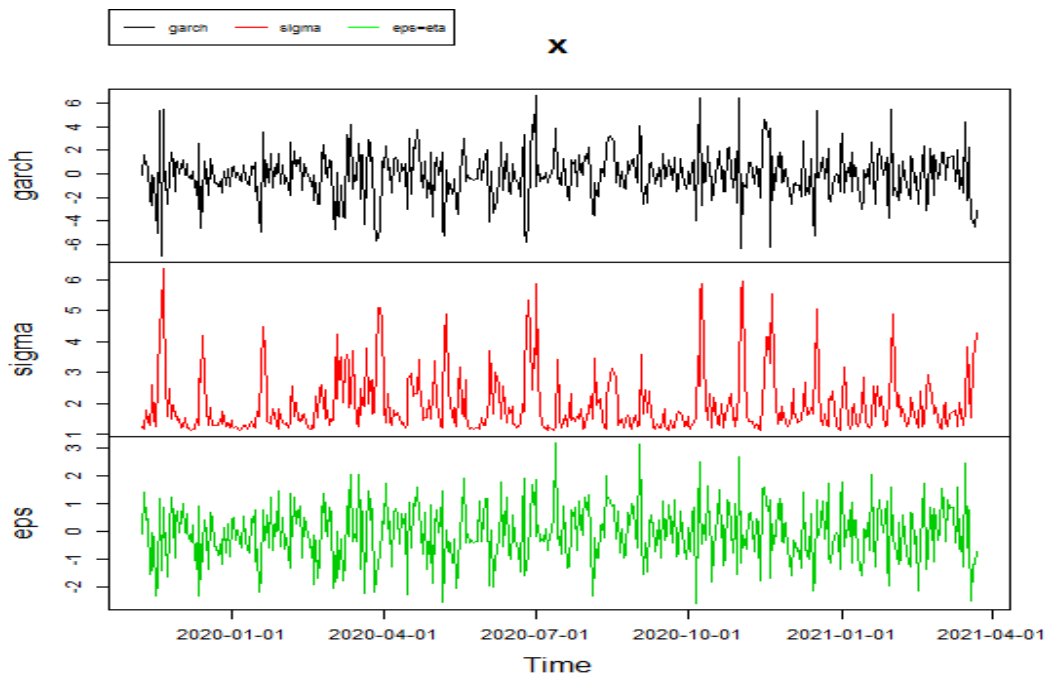


FIGURE 1.2 – Simulation d’un processus GARCH(1, 1) de taille 500 avec  $\omega = 1, \alpha = 0.7, \beta = 0.2$  et  $\eta_t \sim N(0, 1)$ .

De la (figure1.2), avec  $\alpha = 0,7$  et  $\beta = 0,2$ , nous remarquons la présence de valeurs absolue plus élevées. Les deux processus sont également différents en termes de persistance des chocs : lorsque  $\beta$  s'approche de 1, le choc sur la volatilité a un effet persistant. En revanche, lorsque la valeur de  $\alpha$  est importante, des variations de volatilité s'exhibent brusquement en réponse aux chocs.

### 1.3.3 Propriétés des trajectoires

Par rapport aux modèles usuels de séries temporelles (ARMA), cette structure permet au bruit d'avoir un ordre de grandeur fonction des variables passées. Ainsi, vont se succéder des périodes à forte volatilité (grandes valeurs en module des  $X_{t-i}$  et donc de  $\sigma_t^2$ ) et d'autres où les fluctuations sont de plus faibles amplitudes. Les simulations des figures précédentes (figures 1.1 et 1.2) mettent bien en évidence cette propriété dite de regroupement de la volatilité (volatility clustering).

### 1.3.4 Etude de stationnarité

Il existe deux approches pour l'étude de la stationnarité des modèles GARCH.

- Bougerol et Picard (1992) ont donné des conditions de type Lyapunov pour l'existence d'une solution stationnaire aux équations de récurrence stochastiques (SRE), qui sont des cas particuliers de (1.1), cette approche a été traitée pour les modèles conditionnellement hétéroscédastiques (GARCH) par Francq et Zakoian (2010) .

- Doukhan et Wintenberger (2007) ont défini un espace de stationnarité sous des conditions lipchitziennes.

Dans ce qui suit nous allons souligner sur la deuxième approche afin de dériver les conditions de stationnarité du modèle GARCH( $p, q$ ) et ceci dans le cas de gauss.

D'abord nous définissons le processus GARCH comme solution de système d'équation :

$$\begin{cases} X_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases}$$

On peut représenter  $\sigma_t$  sous la forme causale  $\sigma_t^2 = \sum_{j=0}^{\infty} b_j X_{t-j}^2$ , d'où l'équation (1.7) admet la représentation (1.1) avec  $f_{\theta}(X_{t-1}, X_{t-2}, \dots) = 0$  et  $M_{\theta}^t = \sigma_t$ .

Pour un sous-ensemble compact  $\Theta$  de  $\mathbb{R}^d$ , la fonction  $\sigma_t^2(\theta)$  satisfait  $\|\sigma_t^2(\theta)\|_{\Theta} < \infty$ , et il existe une séquence  $(\alpha_j)_j$  de nombres non négatifs tels que  $\forall x, y \in \mathbb{R}^{\infty}$ ,

$$|\sigma_t^2(x) - \sigma_t^2(y)| \leq \sum_{j=1}^{\infty} \alpha_j^{(0)} |x_j - y_j|, \quad (1.10)$$

avec  $\sum_{j=1}^{\infty} \alpha_j^{(0)} < \infty$ .

Ces inégalités de type Lipschitz sur  $\sigma_t^2(\theta)$ , pour le processus (1.7) sont essentielles pour établir la proposition 1.1 et la forte consistance de l'EQMV.

Selon Doukhan et Wintenberger (2007), la condition d'existence d'une solution stationnaire des processus causaux affines (1.1) donnée par (1.5), se réduit pour le modèle GARCH( $p, q$ ) à :

$$\Theta_1(r) = \left\{ \theta \in \mathbb{R}^d / (E[|\eta_0|^r])^{1/r} \sum_{j=1}^{\infty} \alpha_j(\sigma_t, \{\theta\}) < 1 \right\}.$$

Si  $\theta \in \Theta_1$ , alors le processus (1.7) a une solution stationnaire.

### 1.3.5 Estimation des modèles GARCH par quasi-maximum de vraisemblance gaussien

Dans cette partie nous étudions la méthode du maximum de vraisemblance conditionnelle (à des valeurs initiales). Nous présentons une procédure itérative de calcul de la log-vraisemblance gaussienne, conditionnellement à des valeurs initiales fixes ou aléatoires. Cette vraisemblance est écrite comme si la loi des variables  $\eta_t$  était normale centrée réduite (on parle de pseudo ou quasi-vraisemblance), mais cette hypothèse n'est pas nécessaire pour la

convergence forte de l'estimateur. Elle a évidemment un effet sur la variance de la loi normale asymptotique de l'estimateur.

L'estimation par *QMV* gaussienne constituant la méthode de référence, nous détaillons maintenant sa définition et ses propriétés asymptotiques. Le critère à minimiser construit en supposant que le processus des innovations suit une loi gaussienne.

### Quasi-vraisemblance conditionnelle

On supposera que les observations  $X_1, \dots, X_n$  constituent une réalisation (de longueur  $n$ ) d'un processus GARCH  $(p, q)$ , solution strictement stationnaire non anticipative du modèle (1.7)

Généralement pour écrire la vraisemblance du modèle, il faut spécifier une distribution particulière pour les variables iid  $\eta_t$ . Nous considérons généralement la quasi-vraisemblance gaussienne, i.e. la vraisemblance obtenue à partir d'une loi normale centrée réduite pour les  $\eta_t$ . Nous ne ferons cependant pas l'hypothèse que cette loi constitue la vraie distribution du processus iid.

La spécification d'une distribution gaussienne pour les variables  $\eta_t$  ne permet pas d'en déduire simplement la loi de l'échantillon. Nous travaillons avec la vraisemblance de  $X_1, \dots, X_n$  conditionnellement à certaines valeurs initiales.

La vraisemblance conditionnelle gaussienne  $L_n(\theta)$  s'écrit

$$L_n(\theta) = L_n(\theta; X_1, \dots, X_n) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\hat{\sigma}_t^2}} \exp\left(-\frac{X_t^2}{2\hat{\sigma}_t^2}\right), \quad (1.11)$$

où les  $\hat{\sigma}_t^2$  sont définis récursivement, pour  $t \geq 1$ , par

$$\hat{\sigma}_t^2 = \hat{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \hat{\sigma}_{t-j}^2. \quad (1.12)$$

Un estimateur du QMV de  $\theta$  est défini comme toute quantité  $\hat{\theta}_n$  vérifiant presque sûrement

$$L_n(\hat{\theta}_n) = \sup_{\theta \in \Theta} L_n(\theta). \quad (1.13)$$

En prenant le logarithme, que maximiser la vraisemblance revient à minimiser ce critère par rapport à  $\theta$

Nous définissons respectivement la vraisemblance de Gauss et le quasivraisemblance de Gauss par :

$$L_n(\theta) = L_n(\theta; n, x_{n-1}, \dots) = n^{-1} \sum_{t=1}^n \ell_t, \text{ où } \ell_t = \ell_t(\theta) = \frac{X_t^2}{\sigma_t^2} + \log \sigma_t^2.$$

$$\hat{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \hat{\ell}_t, \text{ où } \hat{\ell}_t = \hat{\ell}_t(\theta) = \frac{X_t^2}{\hat{\sigma}_t^2} + \log \hat{\sigma}_t^2, \quad (1.14)$$

et  $\hat{\sigma}_t^2$  est défini par (1.12). Un estimateur du quasi-maximum de vraisemblance est donc une solution mesurable de l'équation

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \hat{L}_n(\theta). \quad (1.15)$$

Nous obtenons les équations de vraisemblance en annulant la dérivée par rapport à  $\theta$  du critère  $\hat{L}_n(\theta)$ .

### Propriétés asymptotiques de l'estimateur du QMV

Pour la convergence, les hypothèses suivantes sont faites, Nous allons reprendre les hypothèses proposées par Bardet et Wintenberger (2009) et nous les adaptons au processus GARCH( $p, q$ ), nous obtenons :

- Hypothèse  $\hat{H}1$  (Compacité) :  $\Theta$  est un ensemble compact.
- Hypothèse  $\hat{H}2$  (Borne inférieure de la variance conditionnelle) : il existe une constante  $\underline{\sigma} > 0$  tel que  $\forall \theta \in \Theta$ , alors  $\sigma_\theta(x) > \underline{\sigma}$  pour tout  $x \in \mathbb{R}^N$ .
- Hypothèse  $\hat{H}3$  (Identifiabilité) : La fonction  $\sigma_\theta$  est telle que : pour tout  $\theta_1, \theta_2 \in \Theta$ ,

alors  $\sigma_{\theta_1} = \sigma_{\theta_2}$  implique que  $\theta_1 = \theta_2$ .

•Hypothèse  $\hat{H}4$  : Une famille  $(\partial\sigma_{\theta_0}^t/\partial\theta_i)_{1 \leq i \leq d}$  est *p.s* linéairement indépendante, où

$$\frac{\partial\sigma_{\theta_0}^t}{\partial\theta} = \frac{\partial\sigma_{\theta}}{\partial\theta}(X_{t-1}, \dots)$$

### Consistance forte

Selon Bardet et Wintenberger (2009), nous adaptons les théorèmes (1.1) et (1.2) au processus GARCH( $p, q$ ) nous obtenons :

**Proposition 1.2.** *Supposons que les hypothèses  $\hat{H}1$ ,  $\hat{H}2$  et  $\hat{H}3$  sont vérifiées et que  $\theta_0 \in \Theta(r)$  avec  $r \geq 1$ . Soit  $X$  une solution stationnaire de (1.7).*

*Si  $(A_0(\sigma, \Theta))$  tel que :*

$$\alpha_j^0(\sigma, \Theta) = O(j^{-\ell}) \text{ pour certain } \ell > 3/2,$$

*alors,*

*la suite de l'EQMV  $\hat{\theta}_n$  converge fortement, c'est-à-dire*

$$\hat{\theta}_n \rightarrow \theta_0 \text{ p.s.n } \rightarrow \infty$$

### Normalit asymptotique

**Proposition 1.3.** *Supposons que  $\theta_0 \in \Theta(4) \cap \overset{\circ}{\Theta}$  l'intérieur de  $\Theta$ , et soit  $X$  la solution stationnaire de (1.7). Sous les hypothèses  $\hat{H}1$ ,  $\hat{H}2$ ,  $\hat{H}3$  et  $\hat{H}4$  si, pour  $i = 1, 2$  et  $(A_i(\sigma, \Theta))$  tel que*

$$\alpha_j^{(1)}(\sigma, \Theta) = O(j^{-\ell}) \text{ pour certain } \ell > 3/2,$$

*alors l'EQMV  $\hat{\theta}_n$  est asymptotiquement normal.*

$$\sqrt{n} \left( \hat{\theta}_n - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{D} \mathcal{N} \left( 0, F(\theta_0)^{-1} G(\theta_0) F(\theta_0)^{-1} \right),$$

où

$$(F(\theta_0))_{ij} = E \left[ \partial^2 q_0(\theta_0) / \partial \theta_i \partial \theta_j \right] \text{ et } (G(\theta_0))_{ij} = E \left[ \frac{\partial q_t(\theta_0)}{\partial \theta_i} \frac{\partial q_t(\theta_0)}{\partial \theta_j} \right]$$

# Chapitre 2

## Estimation du quasi-maximum de vraisemblance du modèle GARCH avec bruit de loi de Student

### 2.1 Introduction

Dans ce chapitre nous analysons l'estimateur du quasi- maximum de vraisemblance (EQMV) dans le cas d'un modèle GARCH [prédéfini dans l'équation 1.7], nous supposons que les innovations ont une distribution qui suit une loi de Student. La distribution  $t$  de Student est une loi symétrique avec des queues plus épaisses que la loi normale. Des queues lourdes sont couramment observées dans les données économiques et financières, en particulier les données à haute fréquence. Le modèle GARCH  $(p, q)$  avec les innovations de Student a attiré l'attention dans la littérature et est considéré comme un modèle réaliste pour les données réelles. Si les queues de la distribution de l'innovation sont lourdes, un estimateur du quasi-maximum de vraisemblance basé sur la densité de Student est plus efficace que l'estimateur de quasi-vraisemblance basé sur la densité normale. En général, une mauvaise spécification de la distribution de l'innovation conduit à des estimations du maximum de vraisemblance non consistant.

La distribution de Student est une loi de probabilité, impliquant le rapport entre une variable suivant une distribution normale standard et la racine carrée d'une variable de loi de  $\chi^2$ . Soit  $Z$  une variable aléatoire avec la distribution normale standard et  $U$  une variable indépendante distribuée selon la loi de  $\chi^2(\nu)$ ; où  $\nu$  est le nombre de degrés de liberté. Par

définition, la variable  $T = \frac{Z}{\sqrt{U/\nu}}$  est une distribution de la loi de Student avec  $\nu$  degrés de liberté. La densité de  $T$ , noté  $f_T$ , est donnée par :

$$f_T(t) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \nu > 0.$$

où  $\Gamma$  est la fonction gamma. La densité  $f$  associée à la variable  $T$  est symétrique, centrée sur 0 et en forme de cloche (voir figure ci-dessous).

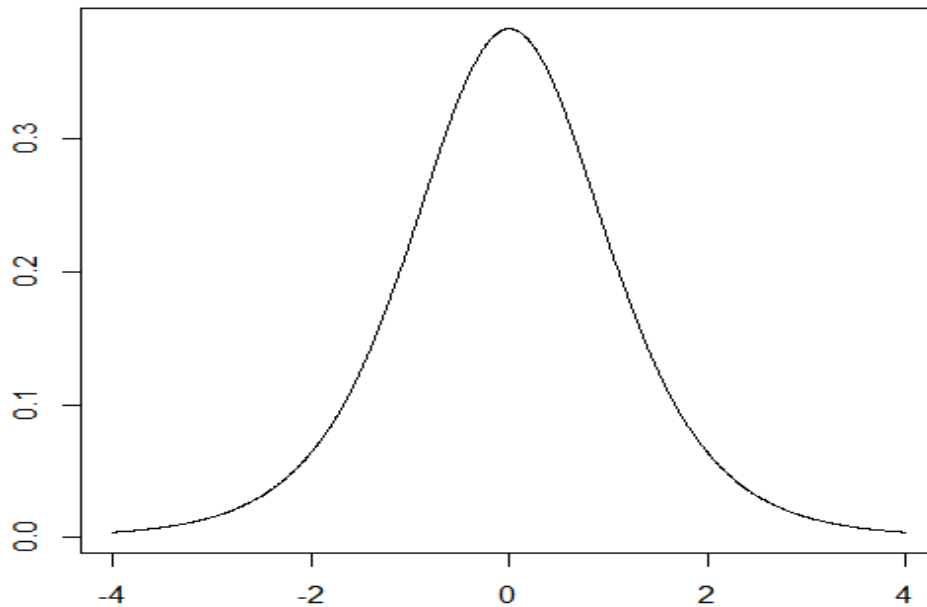


FIGURE 2.1 – Simulation de la densité de la loi de Student

Lorsque nous connaissons la distribution des innovations  $\eta_t$ , l'estimateur de maximum de vraisemblance (conditionnelle) serait utilisé.

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $X$  qui est une solution de (1.7) où  $\theta \in \Theta \subset \mathbb{R}^d$  est inconnue. La forme exacte de la fonction log-vraisemblance de (1.7) dépend de la forme paramétrique de la distribution de l'innovation. Si  $\eta_t$  a une distribution de Student standardisée avec  $\nu > 2$  degrés de liberté, alors la fonction log-vraisemblance est

$$\begin{aligned} \log L &= n \log \left( \frac{\Gamma(\nu + 1)}{\sqrt{\pi(\nu - 2)}\Gamma(\frac{\nu}{2})} \right) - \frac{1}{2} \sum_{t=1}^N \log \sigma_t^2 \\ &\quad - \frac{\nu + 1}{2} \sum_{t=1}^N \log \left[ 1 + \frac{X_t^2}{\sigma_t^2 \times (\nu - 2)} \right]. \end{aligned} \quad (2.1)$$

L'estimateur considéré est un cas particulier de l'estimateur du quasi-maximum de vraisemblance général proposé par Berkes et Horvath (2004). En fait, leur estimateur est assez général et couvre plusieurs lois de probabilité : EQMV gaussien, EQMV Laplace, EQMV Student, EQMV gaussien généralisé, EQMV Gamma ...

L'approche de Doukhan et Wintenberger (2007) que nous utilisons offre une alternative simple pour dériver l'existence d'une solution  $X$  à partir du modèle (1.7), que celle basée sur l'exposant de Lyapunov, introduite par Bougerol et Picard (1992). Doukhan et Wintenberger prouvent l'existence d'une solution strictement stationnaire faiblement dépendante de l'équation  $X_t = F(X_{t-1}, X_{t-2}, X_{t-3}, \dots, \xi_t)$  sous une hypothèse de type Lipschitz spécifique sur  $F$ . Bougerol a donné des conditions de type Lyapunov pour l'existence d'une solution stationnaire aux équations de récurrence stochastiques ; qui sont des cas particuliers de processus  $p$ -Markov  $X_t = F(X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-p}, \xi_t)$ . L'application de la méthode (basée sur l'approche de Doukhan et Wintenberger) à un modèle plus général que GARCH tel que ARCH ( $\infty$ ), AR ( $\infty$ )-ARCH ( $\infty$ ) est plus étendue. Cela dépasse le cadre de Berkes et Horvath (2004) et celui de Francq et Zakoïan (2013) qui ne s'applique qu'au modèle GARCH.

Nous fournissons également les preuves de la consistance et de la normalité asymptotique de l'EQMV avec le bruit distribué selon la loi de Student. Cela n'a été discuté, heuristiquement, que par certains auteurs.

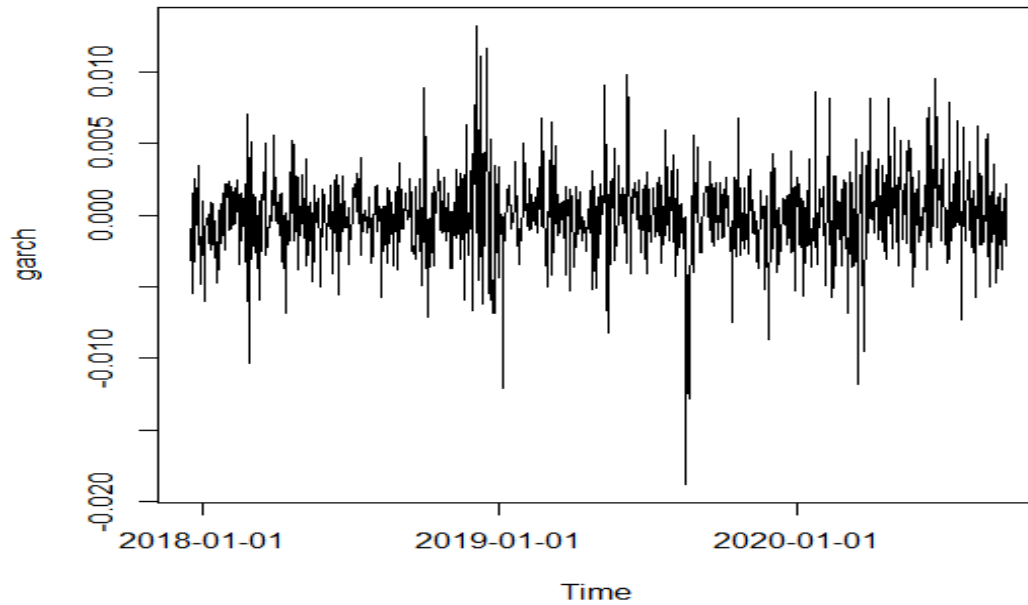


FIGURE 2.2 – Simulation d'un processus GARCH(1, 1) à erreurs de Student

### Notations et hypothèses :

- Le symbole  $\|\cdot\|$  désigne la norme euclidienne habituelle d'un vecteur appartenant à  $\mathbb{R}^d$  ou d'une matrice de dimension à  $m \times p$ .  $\|A\| = \sup_{\|Y\| \leq 1} \{\|AY\| \mid Y \in \mathbb{R}^p\}$ . - Pour une fonction mesurable à valeurs vectorielles  $g$  définie sur un ensemble  $U$ ,

$$\|g\|_{\Theta} = \sup_{\theta \in \Theta} \|g(\theta)\|$$

où  $\Theta$  désigne un sous-ensemble de  $\mathbb{R}^d$ , et  $C(\Theta, \mathbb{R})$  désigne l'espace des fonctions continues sur  $\Theta$  muni de la norme uniforme  $\|\cdot\|_{\Theta}$ .

- Si  $h : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction borélienne. Alors  $h$  est une fonction Lipschitzienne (*Liph*) si

$$Liph = \sup_{\substack{x, y \in \mathbb{R} \\ x \neq y}} \frac{|h(x) - h(y)|}{|x - y|} < \infty.$$

**Hypothèses** Pour  $\theta \in \Theta$ ,  $\sigma_t^2(\theta) : \mathbb{R}^\infty \rightarrow ]0, \infty[$  est une fonction borélienne. L'existence d'une solution à (1.7) sera prouvée sous une condition de type Lipschitz.

**H1** : Pour un sous-ensemble compact  $\Theta$  de  $\mathbb{R}^d$ , la fonction  $\sigma_t^2(\theta)$  satisfait  $\|\sigma_t^2(\theta)\|_\Theta < \infty$ , et il existe une séquence  $(\alpha_j)_j$  de nombres non négatifs tels que  $\forall x, y \in \mathbb{R}^\infty$ .

$$|\sigma_t^2(x) - \sigma_t^2(y)| \leq \sum_{j=1}^{\infty} \alpha_j^{(0)} |x_j - y_j| \quad (2.2)$$

$$\left\| \frac{\partial \sigma_t^2(x)}{\partial x} - \frac{\partial \sigma_t^2(y)}{\partial y} \right\| \leq \sum_{j=1}^{\infty} \alpha_j^{(1)} |x_j - y_j| \quad (2.3)$$

$$\text{avec } \sum_{j=1}^{\infty} \alpha_j^{(0)} < \infty \text{ et } \sum_{j=1}^{\infty} \alpha_j^{(1)} < \infty.,$$

Ces inégalités de type Lipschitz sur  $\sigma_t^2(\theta)$ , pour le processus (1.7) sont essentielles pour établir la proposition ci-dessous et la forte consistance de l'EQMV.

**H2** : La classe de fonctions  $\{\sigma_t^2(\theta) | \theta \in \Theta\}$  est uniformément bornée inférieurement, c'est-à-dire qu'il existe une constante  $\inf_{\theta \in \Theta} (\sigma_t^2(\theta)) \geq \sigma$  pour certains  $\sigma > 0$  pour tous  $\theta \in \Theta$ .

La même condition d'identification que dans Jeantheau (1998) sera utilisée :

**H3** :  $\sigma_t^2(\theta) = \sigma_t^2(\theta_0)$  *p.s.*  $\Rightarrow \theta = \theta_0 \quad \forall \theta \in \Theta$  et certaines  $t \in \mathbb{Z}$ .

**H4** : Le vrai paramètre  $\theta_0$  se trouve à l'intérieur de l'ensemble compact  $\Theta$ .

**H5** :  $\sigma_t^2(\theta)$  est deux fois continment différentiable sur  $\Theta$ .

$$\mathbf{H6} : E \left( \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \right)^2 < \infty, E \left( \frac{1}{(\sigma_t^2(\theta_0))^2} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right) < \infty.$$

L'exigence que  $\theta_0$  soit à l'intérieur de  $\Theta$  permet la différenciation de  $L_n$  dans un voisinage ouvert de  $\theta_0$ .

## 2.2 Existence et stationnarité

La stationnarité est une condition nécessaire à la consistance et à la normalité asymptotique de l'estimateur du paramètre, tout comme l'existence des moments du processus. Dans la proposition ci-dessous, nous prouvons l'existence d'une solution stationnaire du modèle (1.7) sous certaines restrictions sur le paramètre  $\theta_0$ . Pour assurer une solution stationnaire de (1.7), pour  $r \geq 2$ , définissons l'ensemble

$$\Theta(r) = \left\{ \theta \in \mathbb{R}^d, (2.2) \text{ on a, } (E |\eta_0|^r)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)} < 1 \right\}.$$

Il est clair que  $\theta \in \Theta(r)$  seulement si  $E |\eta_0|^r < \infty$ . Ensuite, sur la base des travaux de Doukhan et Wintenberger (2007), nous obtenons ce qui suit.

**Proposition 2.1.** Si  $\theta_0 \in \Theta(r)$  pour certains  $r \geq 2$ , alors il existe une solution unique et causale ( $X_t$  est indépendante de  $(\eta_k)_{k>t}$  pour  $t \in \mathbb{Z}$ )  $X$  de (1.7), qui est stationnaire, ergodique et satisfait  $E \|X_0\|^r < \infty$ .

La proposition (2.1) relie le moment d'ordre  $r$  des innovations  $(\eta_t)$  à un moment  $r$  de  $X_0$ . Cependant, il est connu que la consistance de l'EQMV peut être obtenue, généralement, pour  $r > 2$ . Nous appliquons le résultat de Doukhan et Wintenberger (2007) qui fournit les conditions de l'existence d'une solution stationnaire d'une équation de type

$$X_t = F(X_{t-1}, X_{t-2}, \dots; \eta_t) \text{ p.s pour tout } t \in \mathbb{Z}. \quad (2.4)$$

Si  $E(\eta_0)^r < \infty$  et  $F$  satisfait, pour  $x = (x_i)_{i \geq 1}, y = (y_i)_{i \geq 1} \in \mathbb{R}^\infty$  :

- $E \|F(0; \eta_0)\|^r < \infty$ ;
- $(E \|F(x; \eta_0) - F(y; \eta_0)\|^r)^{1/r} \leq \sum_{j \geq 1} a_j \|x_j - y_j\|$  avec  $\sum_{j \geq 1} a_j < 1$ .

L'existence d'une solution stationnaire causale unique  $X$  de (2.4), telle que  $E[\|X_0\|^r] < \infty$ , est prouvée dans Doukhan et Wintenberger (2007). Nous identifions  $F$  à partir de (1.7) comme suit :

$$F(X_{t-1}, X_{t-2}, \dots; \xi_t) = \sigma_t \eta_t.$$

Evidemment,  $E \|F(0; \eta_0)\|^r < \infty$  car  $E \|\eta_0\|^r < \infty$ , et nous avons

$$\begin{aligned} (E \|F(x, \eta_0) - F(y, \eta_0)\|^r)^{1/r} &\leq (E \|(\sigma_t(x) - \sigma_t(y))\eta_0\|^r)^{1/r} \\ &\leq (E \|\eta_0\|^r)^{1/r} |(\sigma_t(x) - \sigma_t(y))| \\ &\leq (E \|\eta_0\|^r)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)} |x_j - y_j|. \end{aligned}$$

Avec  $\sum_{j=1}^{\infty} \alpha_j^{(0)} < \infty$ .

Hypothèses supplémentaires requises pour la convergence de l'EQMV.

## 2.3 L'EQMV d'un processus GARCH à erreur de loi Student

### 2.3.1 Définition de l'estimateur

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $X$  qui est une solution *p.s.* de (1.7) où  $\theta \in \Theta \subset \mathbb{R}$  est inconnu. Pour estimer  $\theta$ , nous considérons la log-vraisemblance de  $(X_1, \dots, X_n)$  conditionnellement à  $(X_0, X_{-1}, \dots)$ . Si  $\eta_t$  a une distribution de Student standardisée avec  $\nu > 2$  (par rapport à la mesure de Lebesgue), alors, à partir de la définition causale affine de  $X$ , cette log-vraisemblance (2.1) peut être écrite :

$$\begin{aligned} L_n(\theta) &= \log l_n(X_1, \dots, X_n, \theta) = n \log \left( \frac{\Gamma(\nu + 1)}{\sqrt{\pi(\nu - 2)} \Gamma(\frac{\nu}{2})} \right) \\ &\quad - \frac{1}{2} \sum \log \sigma_t^2 - \frac{\nu + 1}{2} \sum \log \left[ 1 + \frac{X_t^2}{\sigma_t^2 \times (\nu - 2)} \right]. \end{aligned}$$

où  $\sigma_t = \sigma(X_{t-1}, X_{t-2}, \dots)$ . Cependant,  $\sigma_t$  n'est généralement pas calculable car  $X_0, X_{-1}, \dots$  sont inconnus. Pour estimer  $\theta$ , nous considérons la log-vraisemblance de  $(X_1, \dots, X_n)$  conditionnellement à  $(X_0, X_{-1}, \dots)$ .

En pratique, nous n'observons que  $X_1, \dots, X_n$  et le logarithme de la fonction quasi-maximum ne peut pas être calculé à partir de nos données. Ainsi, une quasi-log-vraisemblance est considérée au lieu de la log-vraisemblance et elle est définie par :

$$\begin{aligned} \hat{L}_n(\theta) &= \log \hat{l}_n(X_1, \dots, X_n, \theta) = n \log \left( \frac{\Gamma(\nu + 1)}{\sqrt{\pi(\nu - 2)} \Gamma(\frac{\nu}{2})} \right) \\ &\quad - \frac{1}{2} \sum \log \hat{\sigma}_t^2 - \frac{\nu + 1}{2} \sum \log \left[ 1 + \frac{X_t^2}{\hat{\sigma}_t^2 \times (\nu - 2)} \right]. \end{aligned}$$

Avec  $\hat{\sigma}_t = \sigma(X_{t-1}, X_{t-2}, \dots, X_1, u)$ , où  $u = (u_n)_{n \in \mathbb{N}}$  est une suite finie de zéro. Le choix de  $(u_n)_{n \in \mathbb{N}}$  n'influe pas le comportement asymptotique de  $L_n$ , et  $(u_n)$  pourrait typiquement être choisi comme une suite égale à zéro. Il est démontré que le choix des valeurs initiales n'a pas d'importance pour les propriétés asymptotiques de l'EQMV dans Bardet et Wintenberger (2009). Cependant, cela peut être important d'un point de vue pratique. Nous définissons la vraisemblance de Student par  $L_n(\theta) = - \sum_{t=1}^n q_t(\theta)$ , avec

$$q_t(\theta) = \log \sigma_t^2 + (\nu + 1) \log \left[ 1 + \frac{X_t^2}{\sigma_t^2 \times (\nu - 2)} \right].$$

On remplace donc  $L_n(\theta)$  par

$$\hat{L}_n(\theta) = - \sum_{t=1}^n \hat{q}_t(\theta)$$

Avec

$$\hat{q}_t(\theta) = \log \hat{\sigma}_t^2 + (\nu + 1) \log \left[ 1 + \frac{X_t^2}{\hat{\sigma}_t^2 \times (\nu - 2)} \right].$$

Un estimateur du quasi-maximum de vraisemblance (EQMV) est défini par :

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \log \hat{L}_n(X_1, \dots, X_n, \theta).$$

### 2.3.2 Comportement asymptotique des estimateurs

En général, la fonction de vraisemblance n'est pas calculable depuis  $X_0, X_{-1}, \dots$  qui sont inconnus, c'est pourquoi nous utilisons la quasi-vraisemblance obtenue en remplaçant le passé du processus par des zéros. Les propriétés asymptotiques de la quasi-vraisemblance ont été étudiées dans plusieurs articles.

#### Consistance forte

Nous rappelons d'abord le lemme déjà prouvée dans Bardet et Wintenberger (2009). Il s'agit d'une étape cruciale dans la preuve de la consistance de l'estimateur.

**Lemme 2.1.** *Supposons que  $\theta_0 \in \Theta(r)$  pour  $\nu > 2$  et  $X$  est la solution stationnaire de l'équation (1.7). Soit  $\Theta$  un ensemble compact de  $\mathbb{R}^d$ . Si H1 est vrai, alors*

$\forall \theta \in \Theta \quad \sigma_t^2 \in L^r(C(\Theta, \mathbb{R}^d))$  et

$$E \left[ \left\| \hat{\sigma}_t^2 - \sigma_t^2 \right\|_{\Theta}^r \right] \leq E \left[ \|X_0\|^r \right] \left( \left\| \sum_{j=t+1}^{\infty} \alpha_j^{(0)} \right\|_{\Theta} \right)^r.$$

pour tout  $t \in \mathbb{N}^*$ .

Nous sommes maintenant en mesure de dénoncer notre premier résultat. Une forte consistance d'une suite de l'EQMV de Student pour une solution d'équation (1.7) est donnée dans le théorème ci-dessous.

**Théorème 2.1.** *(Djaballah et Kerar (2019)) Supposons que  $\theta_0 \in \Theta(2) \cap \Theta$ . Soit  $X$  la solution stationnaire de (1.7). Si  $\theta_0 \in \Theta$  est un ensemble compact de  $\mathbb{R}$  tel que les hypothèses H1, H2 et H3 sont satisfaites avec*

$$\alpha_j^{(0)} = O(j^{-\ell}).$$

pour certains  $\ell > \frac{3}{2}$ , alors l'estimateur du quasi-maximum de vraisemblance de Student  $\hat{\theta}_n$  converge fortement, c'est-à-dire

$$\hat{\theta}_n \rightarrow \theta_0 \text{ p.s.n } \rightarrow \infty.$$

## Normalité asymptotique de l'EQMV

**Proposition 2.2.** *Nous avons*

$$E \left( -(\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \right) = -1.$$

et

$$E \left( -(\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \right)^2 = 2.$$

En suivant l'approche classique, nous établirons la normalité asymptotique de  $\hat{\theta}_n$  au moyen du Développement de Taylor de  $\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta}$  avec  $L_n(\theta) = -\sum_{t=1}^n q_t(\theta)$ , où

$$q_t(\theta) = \log \sigma_t^2 + (\nu + 1) \log \left[ 1 + \frac{X_t^2}{\sigma_t^2 \times (\nu - 2)} \right].$$

On en déduit que  $\frac{\partial q_t(\theta)}{\partial \theta}$  et  $\frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'}$  sont des suites de variables aléatoires stationnaires et ergodiques avec des valeurs dans  $C(\Theta, \mathbb{R}^d)$  et  $C(\Theta, \mathbb{R}^{d \times d})$ , respectivement.

Nous sommes maintenant prêts à énoncer notre résultat sur la normalité asymptotique.

**Théorème 2.2.** (Djaballah et Kerar (2019)) *Soit  $(X_t)$  un processus stationnaire de forme (1.7) et  $\theta_0$  un vrai vecteur des paramètres. Sous les hypothèses H1-H6, l'EQMV  $\hat{\theta}_n$  est asymptotiquement normal, c'est-à-dire,*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow^D N(0, - \left( \frac{\nu - 1}{\nu + 1} \right) J_0^{-1}).$$

où  $J_0$  est défini par  $\left( E \left( \frac{\partial q_t(\theta_0)}{\partial \theta_i} \frac{\partial q_t(\theta_0)}{\partial \theta_j} \right) \right)_{ij}$ .

Nous suivons une preuve qui est similaire au Theorem2 dans Bardet et Wintenberger (2009). De la proposition 1, nous avons  $\hat{\theta}_n \rightarrow \theta_0$  p.s  $n \rightarrow \infty$ . Comme  $\theta_0 \in \dot{\Theta}$ , un Développement de Taylor de  $\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta_i} \in \mathbb{R}$  implique

$$\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta_i} = \frac{\partial L_n(\theta_0)}{\partial \theta_i} + \frac{\partial^2 L_n(\bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} (\hat{\theta}_n - \theta_0).$$

Pour  $n$  suffisamment grand pour que les  $\bar{\theta}_{n,i} \in \Theta$ , qui sont compris entre  $\hat{\theta}_n$  et  $\theta_0$ , pour tout  $1 \leq i \leq d$ . Nous appliquons le théorème central limite pour les suites de différence de martingale ergodique stationnaire à variance finie.

### Remarque

Aucune hypothèse technique sur la distribution de  $\eta_t$  n'est requise, à part l'existence d'un moment de quatrième ordre. Cette hypothèse est clairement nécessaire à l'existence de la variance du vecteur  $\frac{\partial q_t(\theta_0)}{\partial \theta}$ . Notez également que cette hypothèse n'implique pas l'existence d'un moment de second ordre pour les processus observés.

### 2.3.3 Preuves des principaux résultats

Les preuves des théorèmes 2.1 et 2.2 sont normalisées et découlent des arguments similaires utilisés pour montrer la forte consistance et la normalité asymptotique de l'EQMV pour les modèles GARCH standard dans Francq et Zakoian (2004), Berkes et Horváth (2004) et Straumann (2005) et Bardet et Wintenberger (2009). L'objectif principal ici est d'énoncer les hypothèses de base et de déterminer la distribution asymptotique de l'EQMV pour le modèle GARCH avec le bruit distribué selon la loi de Student. Puisqu'il existe plusieurs similitudes entre le modèle GARCH avec bruit distribué selon la loi normal et le modèle GARCH avec bruit distribué avec la loi de Student, certaines étapes de la preuve dans le cas d'un modèle GARCH avec erreur de Student sont similaires à celle du GARCH avec la loi normal. Ainsi, nous donnons des preuves que lorsque cela semble pertinent et nous référons à Bardet et Wintenberger (2009) pour les détails.

**Preuve du théorème 2.1**

La consistance de  $\hat{\theta}_n$  est démontrée en utilisant la loi des grands nombres. Une loi uniforme (en  $\theta$ ) de grands nombres sur  $(q_t)_{t \in \mathbb{N}^*}$  est établie. Ensuite, il est prouvé que  $L(\theta) := -E(q_t(\theta))/2$  a un maximum unique dans  $\theta_0$ . nous montrons d'abord que  $L_n/n \rightarrow L$ , où la convergence presque sûre est uniforme, sur un ensemble compact qui contient le vrai paramètre. Dans un deuxième lemme, nous prouvons que la fonction limite  $L$  est uniquement maximisée au vrai paramètre  $\theta_0$ ; cela implique la consistance forte du EQMV (notez qu'il n'y a pas de maximisation par rapport à  $\nu$  dans le cas d'un EQMV).

**Lemme 2.2.** *Si  $\theta_0 \in \Theta$ , un ensemble compact de  $\mathbb{R}^d$  tel que les hypothèses H1, H2 et H3 sont vérifiées. Une loi uniforme (en  $\theta$ ) de grands nombres sur  $(\hat{q}_t)_{t \in \mathbb{N}^*}$  est établie*

*Démonstration.* En utilisant la proposition (2.1), avec  $q_t = G(X_t, X_{t-1}, \dots)$ , on en déduit que  $(q_t)_{t \in \mathbb{Z}}$  est une suite stationnaire et ergodique. D'après Straumann et Mikosch (2005), nous savons que si  $(\zeta_t)_{t \in \mathbb{Z}}$  est une suite stationnaire et ergodique d'éléments aléatoires avec des valeurs en  $C(\Theta, \mathbb{R})$ , alors la loi des grands nombres uniforme (en  $\theta \in \Theta$ ) est impliquée par  $E \|\zeta_0\|_{\Theta} < \infty$ .

Par conséquent,  $(q_t)_{t \in \mathbb{Z}}$  satisfait une loi forte des grands nombres uniforme (en  $\theta \in \Theta$ ) dès que  $E[\sup_{\theta} |q_t(\theta)|] < \infty$ . Tout d'abord, notons que  $\sigma_t^2 \geq \sigma$  pour tout  $t \in \mathbb{Z}$ , à partir de l'inégalité du  $\log(x) \leq x - 1$ , pour tout  $x \in ]0, \infty[$  et Lemme 1, pour tout  $t \in \mathbb{Z}$ , on peut écrire

$$\begin{aligned} |q_t(\theta)| &= \left| \log \sigma_t^2 + (\nu + 1) \log \left[ 1 + \frac{1}{(\nu - 2)} \times \frac{X_t^2}{\sigma_t^2} \right] \right| \\ &\leq \frac{\nu + 1}{\nu - 2} \times \frac{X_t^2}{\sigma} + \left| \log \sigma + \frac{\sigma_t^2}{\sigma} - 1 \right| \text{ for all } \theta \in \Theta \end{aligned} \quad (2.5)$$

$\Rightarrow$

$$\sup_{\theta \in \Theta} |q_t(\theta)| \leq \frac{\nu + 1}{\nu - 2} \times \frac{X_t^2}{\sigma} + |\log \sigma| + \frac{\|\sigma_t^2\|_{\Theta}}{\sigma}$$

Mais  $\forall t \in \mathbb{Z}$ ,  $E[X_0^2] < \infty$  ( $\nu > 2$ ) et  $E \|\sigma_t^2\|_{\Theta} < \infty$ . Par conséquent, le côté droit de (2.5) a

un premier moment fini et donc

$$E \left[ \sup_{\theta \in \Theta} |q_t(\theta)| \right] < \infty.$$

Ainsi, la loi forte uniforme des grands nombres pour  $(q_t(\theta))$  résulte :

$$\left\| \frac{L_n(\theta)}{n} - L(\theta) \right\|_{\Theta} \rightarrow 0 \text{ p.s. } n \rightarrow \infty \text{ Avec}$$

$$L(\theta) = -\frac{1}{2} E [q_t(\theta)]$$

■

**Lemme 2.3.** .Si  $\theta_0 \in \Theta$ , un ensemble compact de  $\mathbb{R}^d$  tel que les hypothèses H1, H2 et H3 se vérifient, alors nous avons

$$\frac{1}{n} \left\| \hat{L}_n(\theta) - L_n(\theta) \right\|_{\Theta} \rightarrow 0 \text{ p.s. } n \rightarrow \infty$$

*Démonstration.* Nous montrons que  $\frac{1}{n} \left\| \hat{L}_n(\theta) - L_n(\theta) \right\|_{\Theta} \rightarrow 0$  p.s.  $n \rightarrow \infty$ . En effet, pour tout  $\theta \in \Theta$  et  $t \in \mathbb{N}^*$

$$\begin{aligned} |\hat{q}_t(\theta) - q_t(\theta)| &= \left| \log \hat{\sigma}_t^2 + (\nu + 1) \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{X_t^2}{\hat{\sigma}_t^2} \right) \right. \\ &\quad \left. - \log \sigma_t^2 + (\nu + 1) \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{X_t^2}{\sigma_t^2} \right) \right| \\ &\leq \left| \log \hat{\sigma}_t^2 - \log \sigma_t^2 \right| \\ &\quad + \left| (\nu + 1) \left[ \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{X_t^2}{\hat{\sigma}_t^2} \right) - \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{X_t^2}{\sigma_t^2} \right) \right] \right| \\ &\leq \left| \hat{\sigma}_t^2 - \sigma_t^2 \right| + \frac{(\nu + 1)}{(\nu - 2)} \left| \frac{X_t^2}{\hat{\sigma}_t^2} - \frac{X_t^2}{\sigma_t^2} \right| \\ &\leq \left| \hat{\sigma}_t^2 - \sigma_t^2 \right| + \frac{(\nu + 1)}{(\nu - 2)} \left| \frac{1}{\hat{\sigma}_t^2} - \frac{1}{\sigma_t^2} \right| X_t^2 \end{aligned}$$

Par conséquent, nous avons :

$$|\hat{q}_t(\theta) - q_t(\theta)| \leq \|\hat{\sigma}_t^2 - \sigma_t^2\|_{\Theta} + \frac{X_t^2}{(\nu-2)} \left\| \frac{1}{\hat{\sigma}_t^2} - \frac{1}{\sigma_t^2} \right\|_{\Theta}$$

Nous pouvons écrire,

$$\left\| \frac{1}{\hat{\sigma}_t^2} - \frac{1}{\sigma_t^2} \right\|_{\Theta} \leq \left\| \frac{1}{\hat{\sigma}_t^2} \right\|_{\Theta} \|\hat{\sigma}_t^2 - \sigma_t^2\|_{\Theta} \left\| \frac{1}{\sigma_t^2} \right\|_{\Theta},$$

et

$$\left\| \frac{1}{\sigma_t^2} \right\|_{\Theta} \leq \sigma^{-1}.$$

Donc

$$\sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)| \leq C \left[ \|\hat{\sigma}_t^2 - \sigma_t^2\|_{\Theta} \left( \frac{1}{\sigma^2} + \frac{(\nu+1)}{(\nu-2)} X_t^2 \right) \right].$$

Des inégalités de Hölder et de Minkowski, nous pouvons écrire :

$$E[ (|XY|^{2/3}) ] \leq E[|X|^{2p/3}]^{1/p} (E[|Y|^{2q/3}])^{1/q}$$

.Si  $p = 3$  et  $q = 3/2 \Rightarrow$

$$E[ (|XY|^{2/3}) ] \leq E[|X|^2]^{1/3} (E[|Y|])^{2/3},$$

par conséquent

$$\begin{aligned} E \left[ \left( \sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)| \right)^{2/3} \right] &\leq C (E [\|\hat{\sigma}_t^2 - \sigma_t^2\|_{\Theta}])^{2/3} \times \left( E \left[ \frac{1}{\sigma^2} + \frac{(\nu+1)}{(\nu-2)} X_t^2 \right] \right)^{1/3} \\ &\leq C' \left( \sum_{j=t+1}^{\infty} \alpha_j^{(0)} \right)^{2/3}, \end{aligned}$$

pour tout  $\theta \in \Theta$ , ou  $C' > 0$ , étant donné que  $E[X_t^2] < \infty$  et le lemme ci-dessus.

Considérons,  $n \in \mathbb{N}^*$ ,  $S_n = \sum_{t=1}^n \frac{1}{t} \sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)|$ . Appliquer le lemme de Krone-

cker, si  $\lim_{n \rightarrow \infty} S_n < \infty$  p.s, alors  $\frac{1}{n} \left\| \hat{L}_n(\theta) - L_n(\theta) \right\|_{\Theta} \rightarrow 0$  p.s,  $n \rightarrow \infty$ . voir Bardet et Wintenberger (2009) pour le reste de la preuve. ■

**Lemme 2.4.** *si  $\theta_0 \in \Theta$ , un ensemble compact de  $\mathbb{R}^d$  tel que les hypothèses H1, H2 and H3 sont vérifiant alors  $L(\theta) = -\frac{1}{2}E(q_t(\theta))$  a un maximum unique dans  $\theta_0$ .*

*Démonstration.* Pour  $\theta \in \Theta$ ,

$$\begin{aligned} L(\theta) &= -\frac{1}{2}E(q_t(\theta)) \\ &= -\frac{1}{2}E \left[ \log \sigma_t^2 + (\nu + 1) \log \left( 1 + \frac{X_t^2}{\sigma_t^2 \times (\nu - 2)} \right) \right]. \end{aligned}$$

Par conséquent,

$$\begin{aligned} &L(\theta) - L(\theta_0) \\ &= -\frac{1}{2}E \left[ \log \sigma_t^2(\theta) + (\nu + 1) \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 \right) \right] \\ &\quad + \frac{1}{2}E \left[ \log \sigma_t^2(\theta_0) + (\nu + 1) \log \left( 1 + \frac{1}{(\nu - 2)} \eta_t^2 \right) \right] \\ &\leq -E (\log \sigma_t^2(\theta) - \log \sigma_t^2(\theta_0)) \\ &\quad - (\nu + 1)E \left[ \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 \right) - \log \left( 1 + \frac{1}{(\nu - 2)} \eta_t^2 \right) \right] \\ &\leq -E \left[ \log \frac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)} + (\nu + 1) \left\{ \log \left( 1 + \frac{1}{(\nu - 2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 \right) - \log \left( 1 + \frac{1}{(\nu - 2)} \eta_t^2 \right) \right\} \right] \\ &\leq -E \left[ -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \frac{(\nu + 1)}{(\nu - 2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 - \frac{(\nu + 1)}{(\nu - 2)} \eta_t^2 \right]. \end{aligned}$$

Nous utilisons  $\log x \leq x - 1$

$$\begin{aligned}
& E \left[ -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \frac{(\nu+1)}{(\nu-2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 - \frac{(\nu+1)}{(\nu-2)} \eta_t^2 \right] \\
&= E \left[ -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \frac{(\nu+1)}{(\nu-2)} \times \left( \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1 \right) \eta_t^2 \right] \\
&= E \left( -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right) + E \left( \frac{(\nu+1)}{(\nu-2)} \times \left( \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1 \right) \right) E \eta_t^2 \\
&= E \left( -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right) + E \left( \frac{\nu}{\nu-2} \times \frac{(\nu+1)}{(\nu-2)} \times \left( \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1 \right) \right) \\
&= E \left( -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \frac{\nu(\nu+1)}{(\nu-2)^2} \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - \frac{\nu(\nu+1)}{(\nu-2)^2} \right),
\end{aligned}$$

$$\delta = \frac{\nu}{\nu-2} \times \frac{(\nu+1)}{(\nu-2)} \geq 1 \Rightarrow \log \delta \geq 0 \text{ Donc}$$

$$\begin{aligned}
& E \left[ -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \frac{(\nu+1)}{(\nu-2)} \times \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \eta_t^2 - \frac{(\nu+1)}{(\nu-2)} \eta_t^2 \right] \\
&= E \left( -\log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - \delta \right) \\
&= E \left( -\log \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right) - \delta + \log \delta \\
&\leq E \left( -\log \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right).
\end{aligned}$$

D'où

$$\begin{aligned}
L(\theta) - L(\theta_0) &\leq -E \left( -\log \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} + \delta \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} \right) \\
&\leq 0.
\end{aligned}$$

Cela implique  $L(\theta) - L(\theta_0) < 0$  presque sûrement pour tout  $\theta \in \Theta$ ,  $\theta \neq \theta_0$ . Un supremum de  $L(\theta)$  n'est donc atteint que pour  $\theta = \theta_0$  qui est le maximum unique. ■

Ces trois lemmes (2.2),(2.3) et (2.4) conduisent à la forte consistance de  $\hat{\theta}_n$ .

**Preuve du théorème 2.2**

Une fois la consistance de l'estimateur établie, l'étape suivante consiste à déterminer sa distribution asymptotique. Bien que la méthodologie soit plutôt standard, la validation nécessaire de tous les arguments formels est assez technique. La preuve du théorème 2.2 est basée sur un développement de Taylor de  $\frac{\partial L_n(\theta)}{\partial \theta}$  à  $\theta_0$ , qui est donnée par

$$\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta_i} = \frac{\partial L_n(\theta_0)}{\partial \theta_i} + \frac{\partial^2 L_n(\bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} (\hat{\theta}_n - \theta_0),$$

pour  $n$  suffisamment grand pour que les  $\bar{\theta}_{n,i} \in \Theta$ , qui sont compris entre  $\hat{\theta}_n$  et  $\theta_0$ , pour tout  $1 \leq i \leq d$ .

La preuve du théorème 2.2 est également basée sur plusieurs lemmes : Lemme (2.5), (2.6) et (2.7), très similaire à ceux de Bardet et Wintenberger (2009).

**Lemme 2.5.** *Soit  $\theta_0 \in \Theta(r)$  ( $r \geq 4$ ) et supposons que  $\theta_0 \in \dot{\Theta}$ , un ensemble compact de  $\mathbb{R}^d$  tel que H1-H6 soit vérifiant. Alors,*

$$\frac{1}{\sqrt{n}} \frac{\partial L_n(\theta_0)}{\partial \theta} \rightarrow N_d(0, J_0),$$

ou  $J_0 = (J_0)_{1 \leq i, j \leq d}$  est fini et son expression est donnée par (2.7).

*Démonstration.*

$$L_n(\theta_0) = - \sum_{t=1}^n q_t(\theta_0),$$

avec

$$q_t(\theta_0) = \log \sigma_t^2(\theta_0) + (\nu + 1) \log \left( 1 + \frac{X_t^2}{\sigma_t^2(\theta_0) \times (\nu - 2)} \right),$$

pour tout  $1 \leq i \leq d$ ,  $\frac{\partial L_n(\theta)}{\partial \theta_i} = -\frac{1}{2} \sum_{t=1}^n \frac{\partial q_t(\theta)}{\partial \theta_i}$ . Les calculs donnent les relations

$$\frac{\partial q_t(\theta)}{\partial \theta_k} = \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_k} + \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{\partial}{\partial \theta_k} \left( \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)} \right),$$

le deuxième terme de la somme est égal à

$$\begin{aligned} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{\partial}{\partial \theta_k} \left( \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)} \right) &= \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{X_t^2}{(\nu-2)} \left[ -\frac{\partial \sigma_t^2(\theta)}{\partial \theta_k} \frac{1}{(\sigma_t^2(\theta))^2} \right] \\ &= \frac{(\nu+1)}{1 + \frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)} \left[ -\frac{\partial \sigma_t^2(\theta)}{\partial \theta_k} \frac{1}{(\sigma_t^2(\theta))^2} \right]. \end{aligned}$$

Notons  $F_t = \sigma(X_t, X_{t-1}, \dots)$ , prouvons que  $(\frac{\partial q_t(\theta_0)}{\partial \theta}, F_t)_{t \in \mathbb{Z}}$  est un processus de différence de martingale. En effet, pour tout  $t \in \mathbb{Z}$ ,

$$E \left( \frac{\partial q_t(\theta_0)}{\partial \theta_k} \middle| F_t \right) = \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} + E \left( \frac{(\nu+1)}{1 + \frac{\sigma_t^2(\theta_0)\eta_t^2}{(\nu-2)\sigma_t^2(\theta_0)}} \frac{\sigma_t^2(\theta_0)\eta_t^2}{(\nu-2)} \left[ -\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \frac{1}{(\sigma_t^2(\theta_0))^2} \right] \middle| F_t \right),$$

alors

$$\begin{aligned} E \left( \frac{\partial q_t(\theta_0)}{\partial \theta_k} \middle| F_t \right) &= \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} + E \left( (\nu+1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \left[ -\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \frac{1}{\sigma_t^2(\theta_0)} \right] \middle| F_t \right) \\ &= \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} - \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \frac{1}{\sigma_t^2(\theta_0)} E \left( (\nu+1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \right). \end{aligned}$$

Nous concluons en notant que le deuxième terme de la somme est égal à

$$-\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \frac{1}{(\sigma_t^2(\theta_0))^2}.$$

Compte tenu de la proposition précédente. Afin d'appliquer le théorème central limite pour les différences de martingales, nous devons prouver

$$E\left[\left\|\frac{\partial q_t(\theta_0)}{\partial \theta}\right\|_{\Theta}^2\right] < \infty.$$

Nous avons

$$\frac{\partial q_t(\theta_0)}{\partial \theta_k} = \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} + (\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \left[ -\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k} \frac{1}{\sigma_t^2(\theta_0)} \right].$$

$\Rightarrow$

$$\begin{aligned} E\left(\frac{\partial q_t(\theta_0)}{\partial \theta_k}\right)^2 &= E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 + E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 (\nu + 1)^2 E\left(\frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}}\right)^2 \\ &\quad - 2E\left((\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}}\right) E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 \\ &= E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 (\nu + 1)^2 E\left(\frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}}\right)^2 - E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 \\ &\leq E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 E\left((\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}}\right)^2 \\ &\leq E\left(\frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_k}\right)^2 < \infty \end{aligned}$$

Par conséquent, puisque  $r \geq 4$ , les conditions de moment pour le TCL sont remplies

$$E\left(\left\|\frac{\partial q_t(\theta_0)}{\partial \theta}\right\|^2\right) = \sum_{k=1}^d E\left(\frac{\partial q_t(\theta)}{\partial \theta_k}\right)^2 < \infty.$$

Nous calculons la matrice de covariance asymptotique de  $\frac{\partial q_t(\theta_0)}{\partial \theta}$ .

$$\begin{aligned} \frac{\partial q_t(\theta_0)}{\partial \theta_i} \times \frac{\partial q_t(\theta_0)}{\partial \theta_j} &= \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \\ &\quad - (\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \left( \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \frac{1}{\sigma_t^2(\theta_0)} \right) \\ &\quad - (\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \left( \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{1}{\sigma_t^2(\theta_0)} \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right) \\ &\quad + \left( (\nu + 1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \right)^2 \left( \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{1}{\sigma_t^2(\theta_0)} \right) \left( \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \frac{1}{\sigma_t^2(\theta_0)} \right), \end{aligned}$$

Par conséquent

$$\begin{aligned} E \left[ \frac{\partial q_t(\theta_0)}{\partial \theta_i} \frac{\partial q_t(\theta_0)}{\partial \theta_j} \right] &= E \left( \frac{1}{(\sigma_t^2(\theta_0))^2} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right) - 2E \left( \frac{1}{(\sigma_t^2(\theta_0))^2} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right) \\ &\quad + 2E \left( \frac{1}{(\sigma_t^2(\theta_0))^2} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right) \end{aligned} \quad (2.6)$$

$\Rightarrow$

$$\begin{aligned} (J_0)_{ij} &= E \left[ \frac{\partial q_t(\theta)}{\partial \theta_i} \frac{\partial q_t(\theta)}{\partial \theta_j} \right] \\ &= E \left( \frac{1}{(\sigma_t^2(\theta_0))^2} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_j} \right). \end{aligned} \quad (2.7)$$

■

**Lemme 2.6.** *Sous les hypothèses du lemme précédent*

$$\frac{1}{n} \frac{\partial^2 L_n(\theta)}{\partial \theta \partial \theta'} \rightarrow \frac{1}{n} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} = -\frac{1}{2} E \left( \frac{\partial^2 q_0(\theta)}{\partial \theta \partial \theta'} \right) \quad p.s \quad n \rightarrow \infty.$$

*Démonstration.* La convergence uniforme est prouvée en montrant l'équicontinuité presque sûre de  $\frac{1}{n} \sum_{t=1}^n v_t$ . Soit  $(v_t)$  une suite de variables aléatoires stationnaire et ergodique avec des

valeurs dans  $C(K, \mathbb{R}^d)$ . Alors l'uniforme SLLN (loi forte uniforme des grands nombres) est impliquée par  $E \|v_0\|_K < \infty$ . La seconde dérivée du processus est stationnaire et ergodique (c'est une fonction mesurable de  $X_t, X_{t-1}, \dots$ ). Par conséquent, il satisfait à une loi uniforme des grands nombres (ULLN) si son premier moment uniforme est borné. Il faut montrer que :

$$E \left\| \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j} \right\|_{\Theta} < \infty. \text{ Nous calculons } \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j}.$$

Comme

$$\frac{\partial q_t(\theta)}{\partial \theta_j} = \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} + \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{\partial}{\partial \theta_j} \left( \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)} \right),$$

alors

$$\begin{aligned} \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j} &= (\sigma_t^2(\theta))^2 \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_i} \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_j} - \sigma_t^2(\theta) \frac{\partial^2 (\sigma_t^2(\theta))^{-1}}{\partial \theta_i \partial \theta_j} \\ &\quad - (\nu+1) \left( \frac{\frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)}}{1 + \frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)}} \right)^2 \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_i} \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_j} \\ &\quad + (\nu+1) \frac{\frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)}}{1 + \frac{\sigma_t^2(\theta)\eta_t^2}{(\nu-2)}} \frac{\partial^2 (\sigma_t^2(\theta))^{-1}}{\partial \theta_i \partial \theta_j} \end{aligned} \quad (2.8)$$

$\Rightarrow$

$$\begin{aligned} (B_0)_{ij} &= E \left( \frac{\partial^2 q_t(\theta_0)}{\partial \theta_i \partial \theta_j} \right)_{ij} = \left(1 - \frac{2}{\nu+1}\right) E \left[ (\sigma_t^2(\theta_0))^2 \frac{\partial (\sigma_t^2(\theta_0))^{-1}}{\partial \theta_i} \frac{\partial (\sigma_t^2(\theta_0))^{-1}}{\partial \theta_j} \right] \\ &= \left(1 - \frac{2}{\nu+1}\right) E \left[ (\sigma_t^2(\theta_0))^2 \left( -\frac{\partial \sigma_t^2(\theta)}{\partial \theta_i} \frac{1}{(\sigma_t^2(\theta))^2} \right) \left( -\frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \frac{1}{(\sigma_t^2(\theta))^2} \right) \right] \\ &= \left(1 - \frac{2}{\nu+1}\right) E \left( \frac{1}{(\sigma_t^2(\theta))^2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_i} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right), \end{aligned} \quad (2.9)$$

$1 - \frac{2}{\nu+1} = \frac{\nu-1}{\nu+1} > 0$  car  $\nu > 2$ .  $B_0$  est une matrice définie positive. En effet, pour tout  $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ ,

$$y' B_0 y = \frac{\nu - 1}{\nu + 1} E \left[ \frac{1}{(\sigma_t^2(\theta))^2} \left( \sum_{1 \leq i \leq d} \left[ \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta_i} y_i \right] \right)^2 \right].$$

Ce terme n'est pas négatif. Alors,  $B_0$  est une matrice inversible, et il existe  $n$  suffisamment grand pour que  $B_n$  soit une matrice inversible. ■

**Lemme 2.7.**

$$E \left( \frac{1}{\sqrt{n}} \left\| \frac{\partial L_n}{\partial \theta} - \frac{\partial \hat{L}_n}{\partial \theta} \right\|_{\Theta} \right) \rightarrow 0 \text{ si } n \rightarrow \infty.$$

*Démonstration.*  $\left\| \frac{\partial q_t(\theta)}{\partial \theta_i} - \frac{\partial \hat{q}_t(\theta)}{\partial \theta_i} \right\|_{\Theta}$

$$\begin{aligned} \frac{\partial \hat{q}_t(\theta)}{\partial \theta_j} - \frac{\partial q_t(\theta)}{\partial \theta_j} &= \frac{1}{\hat{\sigma}_t^2(\theta)} \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} + \frac{(\nu + 1)}{1 + \frac{X_t^2}{(\nu - 2)\hat{\sigma}_t^2(\theta)}} \frac{\partial}{\partial \theta_j} \left( \frac{X_t^2}{(\nu - 2)\hat{\sigma}_t^2(\theta)} \right) - \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \\ &\quad + \frac{(\nu + 1)}{1 + \frac{X_t^2}{(\nu - 2)\sigma_t^2(\theta)}} \frac{\partial}{\partial \theta_j} \left( \frac{X_t^2}{(\nu - 2)\sigma_t^2(\theta)} \right) \\ &= \left( \frac{1}{\hat{\sigma}_t^2(\theta)} \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right) + \frac{(\nu + 1)}{1 + \frac{X_t^2}{(\nu - 2)\hat{\sigma}_t^2(\theta)}} \frac{\partial}{\partial \theta_j} \left( \frac{X_t^2}{(\nu - 2)\hat{\sigma}_t^2(\theta)} \right) \\ &\quad - \frac{(\nu + 1)}{1 + \frac{X_t^2}{(\nu - 2)\sigma_t^2(\theta)}} \frac{\partial}{\partial \theta_j} \left( \frac{X_t^2}{(\nu - 2)\sigma_t^2(\theta)} \right). \end{aligned}$$

Cela implique

$$\begin{aligned}
 \frac{\partial \hat{q}_t(\theta)}{\partial \theta_j} - \frac{\partial q_t(\theta)}{\partial \theta_j} &\leq \frac{1}{\sigma} \left( \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right) \\
 &+ \left( \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\hat{\sigma}_t^2(\theta)}} \frac{X_t^2}{(\nu-2)} \frac{\partial (\hat{\sigma}_t^2(\theta))^{-1}}{\partial \theta_j} - \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{X_t^2}{(\nu-2)} \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_j} \right) \\
 &\leq \frac{1}{\sigma} \left( \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right) \\
 &+ \frac{X_t^2}{(\nu-2)} \left( \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\hat{\sigma}_t^2(\theta)}} \frac{\partial (\hat{\sigma}_t^2(\theta))^{-1}}{\partial \theta_j} - \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma_t^2(\theta)}} \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_j} \right) \\
 &\leq \frac{1}{\sigma} \left( \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right) + \frac{X_t^2}{(\nu-2)} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma}} \left( \frac{\partial (\hat{\sigma}_t^2(\theta))^{-1}}{\partial \theta_j} - \frac{\partial (\sigma_t^2(\theta))^{-1}}{\partial \theta_j} \right)
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial \hat{q}_t(\theta)}{\partial \theta_j} - \frac{\partial q_t(\theta)}{\partial \theta_j} &\leq \frac{1}{\sigma} \times \left( \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right) \\
 &+ \frac{X_t^2}{(\nu-2)} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma}} \times \frac{1}{\sigma^2} \left( \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right).
 \end{aligned}$$

Puisque  $\sigma_t^2 \geq \sigma$  pour tout  $t \in \mathbb{Z}$  et  $\hat{\sigma}_t^2 \geq \sigma$  pour tout  $t \in \mathbb{Z}$ .

$\Rightarrow$

$$\begin{aligned}
 \left\| \frac{\partial \hat{q}_t(\theta)}{\partial \theta_j} - \frac{\partial q_t(\theta)}{\partial \theta_j} \right\|_{\Theta} &\leq \frac{1}{\sigma} \left\| \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right\|_{\Theta} \\
 &+ \left( \frac{X_t^2}{(\nu-2)} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma}} \right) \times \frac{1}{\sigma^2} \left\| \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right\|_{\Theta}
 \end{aligned}$$

$$\begin{aligned}
E \left\| \frac{\partial \hat{q}_t(\theta)}{\partial \theta_j} - \frac{\partial q_t(\theta)}{\partial \theta_j} \right\|_{\Theta} &\leq \frac{1}{\sigma} E \left\| \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right\|_{\Theta} \\
&+ E \left( \frac{X_t^2}{(\nu-2)} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma}} \right) \times \frac{1}{\sigma^2} E \left\| \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right\|_{\Theta},
\end{aligned}$$

mais

$$\left\| \frac{\partial \hat{\sigma}_t^2(\theta)}{\partial \theta_j} - \frac{\partial \sigma_t^2(\theta)}{\partial \theta_j} \right\|_{\Theta} \leq \sum_{j \geq 1} \alpha_j^{(1)}(\Theta) < \infty,$$

et

$$E \left( \frac{X_t^2}{(\nu-2)} \frac{(\nu+1)}{1 + \frac{X_t^2}{(\nu-2)\sigma}} \right) = 1.$$

■

*Démonstration.* de la proposition (2.2)

$$f(t) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

et

$$\begin{aligned} f' &= \frac{\partial \left( \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \right)}{\partial t} \\ &= -\frac{2}{\sqrt{\pi}} \frac{t}{\nu^{\frac{3}{2}}} \frac{\Gamma(\frac{1}{2}\nu + \frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \times \frac{\frac{1}{2}\nu + \frac{1}{2}}{\left(\frac{1}{\nu}(t^2 + \nu)\right)^{\frac{1}{2}\nu + \frac{3}{2}}}, \end{aligned}$$

donc

$$\frac{f'}{f} = \frac{-\frac{2}{\sqrt{\pi}} \frac{t}{\nu^{\frac{3}{2}}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu+1}{2} \left(\frac{t^2}{\nu} + 1\right)^{-\frac{\nu+3}{2}}}{\frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}} = -t \frac{\nu+1}{t^2 + \nu}$$

$\Rightarrow$

$$\frac{f'}{f} = -(\nu+1) \frac{t}{\nu + t^2}$$

$$\begin{aligned} E \left( -(\nu+1) \frac{\frac{\eta_t^2}{(\nu-2)}}{1 + \frac{\eta_t^2}{(\nu-2)}} \right) &= \int x f'(x) dx = x f(x) \Big|_{-\infty}^{\infty} - \int f(x) dx \\ &= - \int f(x) dx \\ &= -1 \end{aligned}$$

$$\begin{aligned}
E\left(\eta_0^2 \frac{f''}{f}\right) &= \int x^2 f''(x) dx = x^2 f'(x)|_{-\infty}^{\infty} - 2 \int x f'(x) dx \\
&= -2 \int x f'(x) dx \\
&= 2
\end{aligned}$$

$$\begin{aligned}
E\left(-(\nu+1) \frac{\frac{\eta_i^2}{(\nu-2)}}{1 + \frac{\eta_i^2}{(\nu-2)}}\right)^2 &= \int \left(x \frac{f'(x)}{f(x)}\right)^2 f(x) dx \\
&= I.
\end{aligned}$$

Nous calculons  $I$

$$I = \int \left(x \frac{f'(x)}{f(x)}\right)^2 f(x) dx = \int \frac{f'(x)}{f(x)} x^2 f'(x) dx.$$

L'intégration par partie est réalisée. On a

$$\begin{aligned}
I &= x^2 \frac{f'(x)}{f(x)} f(x)|_{-\infty}^{\infty} - \int 2x \frac{f'(x)}{f(x)} f(x) - \int x^2 \frac{f'(x)^2}{f(x)^2} f(x) + \int x^2 \frac{f''(x) f(x)}{f(x)^2} f(x) \\
&= x^2 f'(x)|_{-\infty}^{\infty} - 2 \int x f'(x) - \int x^2 \frac{f'(x)^2}{f(x)} + \int x^2 f''(x)
\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}
2I &= x^2 f'(x)|_{-\infty}^{\infty} - 2 \int x f'(x) + \int x^2 f''(x) \\
&= 4.
\end{aligned}$$

on en déduit  $I = 2$ . ■

# Chapitre 3

## Processus causaux affines avec bruit de loi de Cauchy

### 3.1 Introduction

Les séries financières étant réputées avoir des queues de distribution très lourdes, il semble pertinent d'autoriser des erreurs ayant une loi stable non nécessairement gaussienne.

Ces lois sont de plus en plus utilisées dans des modélisations des processus temporelles. D'ailleurs, plusieurs travaux consacrent une partie à l'analyse des séries chronologiques  $\alpha$ -stables, à l'instar de la modélisation des processus linéaires *ARMA* qui utilisent des lois  $\alpha$ -stables ( thèse d'Estampes(2003)), et de celle des processus non linéaires *GARCH* innovations  $\alpha$ -stable (thèse de Lepage(2012)).

### 3.2 Présentation des lois stables univariées

Dans les années 60, Mandelbrot étudie les fluctuations boursières, pour lesquelles il était tout-à-fait clair que le modèle gaussien ne convenait pas. Il s'appuie alors sur les lois de Pareto pour mettre en évidence un nouveau modèle de variation des prix, appelé "lois  $\alpha$  stables". Le paramètre  $\alpha$ , compris entre 0 et 2, représente l'exposant caractéristique des lois stables et lorsque celui-ci est strictement inférieur à 2, la variance de la loi stable est infinie.

Les lois stables sont une famille de lois qui présentent un grand intérêt dans la modélisation de nombreux problèmes physiques. La loi la plus connue de cette famille est la loi normale.

La caractéristique la plus importante de ces lois est leur index de stabilité  $\alpha$ . Encore appelé exposant caractéristique, ce paramètre compris entre 0 et 2 indique la vitesse de décroissance des queues de distributions. La loi normale, par exemple, correspond à un index de stabilité  $\alpha = 2$ . C'est la seule loi stable qui ne soit pas à queue lourde.

La loi normale est la première des lois stables ayant fait son apparition et a trouvé immédiatement des applications en théorie des erreurs, puis en physique statistique et dans les autres sciences. La deuxième loi stable qui a paru en physique théorique a été le profil de dispersion de Lorentz d'une raie spectrale (Lorentz, 1906) connu en théorie des probabilités comme la loi de Cauchy, à savoir, la distribution symétrique stable avec les paramètres  $\alpha = 1$  et  $\beta = 0$ , où  $\beta$  est un paramètre d'asymétrie. La distribution stable de paramètres  $\alpha = 1/2$  et  $\beta = 1$ , appelée distribution de Lévy. Ce sont là les seules lois ayant une expression explicite de la densité.

Nous présentons plus en détail la famille des distributions  $\alpha$ -stables. Le seul fait que les lois stables ont une queue de type lourde ou bien asymptotiquement parétienne (pour faire référence à la loi de Pareto) ne suffit pas pour justifier leur importance. La raison profonde provient de la probabilité de stabilité qui affirme que toute combinaison linéaire de variable aléatoire réel  $\alpha$ -stables est aussi de loi  $\alpha$ -stable.

### 3.2.1 Caractérisation des lois stables univariées

L'intérêt des mathématiciens pour les sommes de variables aléatoires indépendantes se voit clairement dans des résultats bien connus de la théorie des probabilités, tels que la loi des grands nombres ou le théorème centrale limite. C'est aussi l'étude de ces sommes qui a permis la caractérisation des lois dites  $\alpha$ -stables (ou simplement stables). L'écriture  $\stackrel{d}{=}$  indique une égalité en distribution.

**Définition 3.1.** *Une variable aléatoire  $X$  est dite stable si  $a, b$  deux nombres réels positifs et  $X_1, X_2$  deux variables aléatoires indépendantes de même loi que  $X$ . Il existe  $c \in \mathbb{R}^+$  et  $d \in \mathbb{R}$  tel que :*

$$aX_1 + bX_2 \stackrel{d}{=} cX + d. \quad (3.1)$$

Cette définition montre que la famille des lois stables est préservée par convolution, d'où cette notion de stabilité. Nous pouvons utiliser une autre définition des variables aléatoires stables, équivalente à la première.

**Définition 3.2.** *Une variable aléatoire  $X$  est dite stable si, pour tout entier non nul  $n$ , il existe des constantes  $a_n > 0$  et  $b_n$  telles que*

$$\sum_{i=1}^n X_i \stackrel{d}{=} a_n X + b_n. \tag{3.2}$$

Lorsque  $b_n = 0$ , nous parlons de distribution strictement stable.

Les constantes de normalisation sont de la forme  $a_n = n^{1/\alpha}$  avec  $0 < \alpha < 2$ . La démonstration est détaillée dans Feller (1971, pages 166 – 167). La constante  $\alpha$  est appelée index de stabilité ou exposant caractéristique de  $X$ , la variable  $X$  est alors dite  $\alpha$ -stable.

Cette définition montre qu'une variable aléatoire stable possède un domaine d'attraction. Cela veut dire qu'il existe une suite de variables aléatoires indépendantes et identiquement distribuées *i.i.d.*  $\{X_i\}$   $i \in \mathbb{N}$ , une suite de réels positifs  $\{a_n\}$  et une suite de réels  $\{b_n\}$ , telles que

$$\frac{1}{a_n} \left( \sum_{i=1}^n X_i - b_n \right) \xrightarrow[n \rightarrow \infty]{d} X.$$

Les lois stables sont les seules lois qui peuvent être obtenues comme limites de sommes normalisées de variables aléatoires indépendantes et identiquement distribuées, et par conséquent les seules ayant un domaine d'attraction.

**Remarque**

Le terme «stable» se réfère à la propriété que la somme des variables aléatoires indépendantes identiquement distribuées ont la même distribution que l'originale.

### La fonction caractéristique des lois alpha stables

En général, les variables aléatoires réelles continues sont définies par leur densité  $f$ . Les lois stables sont définies par leur fonction caractéristique  $\varphi_x$  (qui dépend de 4 paramètres) car, dans le cas des lois alpha stables, la transformée de Fourier inverse ne donne pas une expression simple de la densité  $f$ .

La fonction caractéristique joue un rôle central dans la théorie des lois stables. Dans la pratique, l'avantage principal de l'utilisation des fonctions caractéristiques ressort de cette propriété : la fonction caractéristique d'une somme de variables aléatoires indépendantes est simplement égale au produit des fonctions caractéristiques de chacune des variables aléatoires intervenant dans cette somme. Dans le théorème suivant, qui a aussi valeur de définition, nous reprenons la formulation canonique proposée par Lévy et Khintchine.

**Théorème 3.1.** (*Levy-Khinchin*)

Une variable aléatoire  $X$  a une distribution stable si et seulement s'il existe quatre paramètres uniques :  $0 < \alpha \leq 2, \gamma \geq 0, -1 \leq \beta \leq 1$ , et un réel  $\mu$  tels que la fonction caractéristique de  $X$  s'écrit sous la forme :

$$\varphi_X(t) = Eexp\{itX\} = exp\{i\mu t - \gamma|t|^\alpha[1 + i\beta sign(t)\omega(t, \alpha)]\}, \tag{3.3}$$

où

$$\omega(t, \alpha) \begin{cases} -\tan\frac{\pi\alpha}{2} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi}\ln|t| & \text{si } \alpha = 1 \end{cases},$$

avec  $t$  un réel, et

$$sign(t) \begin{cases} 1 & \text{si } t > 0 \\ 0 & \text{si } t = 0 \\ -1 & \text{si } t < 0 \end{cases},$$

autrement dit :

$$\varphi_X(t) = Eexp\{itX\} = \begin{cases} exp\{i\mu t - \gamma |t| [1 + i\beta sign(t)\frac{2}{\pi} \ln |t|]\} & \text{si } \alpha = 1 \\ exp\{i\mu t - \gamma |t|^\alpha [1 - i\beta sign(t) \tan \frac{\alpha\pi}{2}]\} & \text{si } \alpha \neq 0 \end{cases}. \quad (3.4)$$

L'interprétation des quatre paramètres est donnée par :

- $\alpha$  paramètre principal  $0 < \alpha \leq 2$  : Le paramètre  $\alpha$  est appelé exposant caractéristique ou index de stabilité. Il caractérise les queues de distribution car il détermine la vitesse de décroissance de la queue de distribution, c'est-à-dire que plus  $\alpha$  diminue, plus les queues sont lourdes. C'est pourquoi nous parlons aussi de lois  $\alpha$  stables .

- $\beta$  est le paramètre d'asymétrie,  $-1 \leq \beta \leq 1$  : Lorsque  $\beta > 0$  ( $\beta = 1$ ) la distribution de  $X$  est dite asymétrique (totalement asymétrique) à droite et lorsque  $\beta < 0$  ( $\beta = -1$ ), la distribution de  $X$  est asymétrique (totalement asymétrique) à gauche. Quand  $\beta = 0$ , la distribution de  $X$  est symétrique par rapport à  $\mu$ . Si de plus,  $\mu = 0$ , la loi est dite symétrique alpha-stable, de fonction caractéristique  $\varphi_X(t) = exp\{-\gamma |t|^\alpha\}$ .

- $\mu$  est le paramètre de position : Il caractérise la moyenne de la loi (lorsque  $\alpha > 1$ ).

- $\gamma$  est le paramètre de dispersion ou paramètre d'échelle, il mesure la dispersion de la distribution autour du paramètre de position  $\mu$ .

Par convention,  $X \sim S_\alpha(\mu, \beta, \gamma)$  représentera une loi stable de paramètres  $\alpha, \mu, \beta, \gamma$  et  $S_\alpha S(\gamma)$  une loi symétrique alpha-stable de paramètre  $\gamma$  Enfin, il est assez courant dans la littérature de définir la fonction caractéristique de cette loi par  $\varphi_X(t) = exp\{-\sigma^\alpha |t|^\alpha\}$ .

**Notation**

$X \sim S_\alpha(\mu, \beta, \gamma)$  Ou  $X \stackrel{d}{=} S(\alpha, \mu, \beta, \gamma)$ , Signifie que  $X$  est distribuée selon une loi stable de paramètres  $\alpha, \mu, \beta, \gamma$ . Notons  $S_\alpha S$ , la distribution centré et symétrique de variable aléatoire  $X$ , soit  $\beta = 0$  et  $\mu = 0$ . Une loi  $\alpha$ -stable est dite standard si  $\mu = 0$  et  $\gamma = 1$ . Enfin, reste à noter aussi qu'il est assez courant dans la littérature de remplacer la dispersion  $\gamma$  par  $\sigma^\alpha$  et d'appeler  $\sigma$  paramètre d'échelle.

**Exemple**

Les distributions  $\alpha$  stables les plus connues, et les seules dont nous disposons d'une forme explicite pour les densités, sont les suivantes :

- La distribution gaussienne de densité :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right); \quad -\infty < x < +\infty, \quad (3.5)$$

a pour fonction caractéristique :

$$\varphi_x(t) = \exp\left\{-\frac{\sigma^2}{2}t^2 + i\mu t\right\}. \quad (3.6)$$

L'équation (3.6) correspond à la fonction caractéristique d'une loi  $S_2(\mu, 0, \frac{\sigma}{\sqrt{2}})$

- La distribution de Cauchy généralisée de densité :

$$f(x) = \frac{\gamma}{\pi [(x-\mu)^2 + \gamma^2]}; \quad -\infty < x < +\infty, \quad (3.7)$$

a pour fonction caractéristique :

$$\varphi_X(t) = \exp\{-\gamma|t| + i\mu t\}, \quad (3.8)$$

ce qui correspond à une loi  $S_1(\mu, 0, \gamma)$

- La distribution de Levy a pour densité :

$$f(x) = \left(\frac{\gamma}{2\pi}\right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\left\{-\frac{\gamma}{2(x-\mu)}\right\}; \quad \mu < x < \infty, \quad (3.9)$$

sur  $(\mu, \infty)$ , avec  $\gamma$  le paramètre de dispersion. Sa fonction caractéristique correspond à celle d'une loi  $S_{1/2}(\gamma, 1, \mu)$

### 3.2.2 Processus stochastiques de lois stables

Les processus stables ont joué un rôle important dans la modélisation en finance. Mandelbrot a proposé un modèle pour les prix du coton selon un processus stable à variance infinie. Ce modèle a été fort critiqué à cause de la non existence des seconds moments du processus . Ces processus sont une généralisation des processus gaussiens.

Un vecteur aléatoire  $\underline{X} = (X_1, \dots, X_n)$  est  $\alpha$ -stable si pour tout  $k$  et pour toute famille  $\underline{X}^{(1)}, \dots, \underline{X}^{(k)}$  de même loi que  $\underline{X}$  , il existe  $a_k > 0$  et  $B^{(k)}$  , tels que

$$\underline{X}^{(1)} + \dots + \underline{X}^{(k)} \stackrel{\mathcal{L}}{=} a_k \underline{X} + B^{(k)}.$$

Lorsque  $B^{(k)}$  est le vecteur nul, nous parlons de loi strictement stable. Le paramètre  $a_k$  à le même rôle que dans le cas univarié. Il existe une constante  $\alpha$ ,  $0 < \alpha \leq 2$ , telle que  $a_k = k^{1/\alpha}$

Pour  $\underline{X}$  un vecteur  $\alpha$ -stable, toute combinaison lineaire des composantes de  $\underline{X}$  est une variable aléatoire réelle  $\alpha$ -stable.

Un processus  $\{X(t); t \in T \subset \mathbb{R}\}$  est  $\alpha$ -stable si pour tout  $n \in \mathbb{N}$  et pour tout  $t_1, \dots, t_n \in T$  le vecteur aléatoire  $(X(t_1), \dots, (Xt_n))$  est  $\alpha$ -stable.

**Définition 3.3.** Soit  $\{X(t); t \in T\}$  un processus stochastique. Alors :

1.  $\{X(t); t \in T\}$  est strictement stable si et seulement si toute combinaison linéaire :

$$\sum_{k=1}^n b_k X(t_k), \quad t_1, \dots, t_n \in T, \quad n \in \mathbb{N}, \quad b_1, \dots, b_n \in \mathbb{R},$$

est strictement stable.

2. Si  $\alpha \geq 1$ , alors  $\{X(t); t \in T\}$  est  $\alpha$ -stable si et seulement si toute combinaison linéaire du processus est  $\alpha$ - stable.

#### Remarque

Les processus stables sont une généralisation des processus gaussiens. Il ont joué un rôle important en modélisation.

### 3.2.3 Loi de Cauchy

Parmi les lois  $\alpha$  stables la distribution de Cauchy ( $\alpha = 1$ ), qui représente un cas extrême et sert de contre-exemples pour certains résultats et concepts bien acceptés en statistique. Par exemple, le théorème central limite ne vaut pas pour la distribution limite de la moyenne d'un échantillon aléatoire d'une distribution de Cauchy. En raison de cette nature particulière, certains auteurs considèrent la distribution de Cauchy comme un cas particulier.

La loi de Cauchy, appelée aussi loi de Lorentz, est une loi de probabilité pour les variables aléatoires continues.

Une variable aléatoire  $X$  suit une loi de Cauchy si elle admet une densité  $f_X$  par rapport à la mesure de Lebesgue, dépendant des deux paramètres  $\mu$  et  $\gamma$  ( $\gamma > 0$ ) et définie par :

$$f_X(x) = \frac{1}{\pi\gamma \left[ 1 + \left( \frac{x-\mu}{\gamma} \right)^2 \right]}, \quad x \in \mathbb{R}$$

Cette distribution est symétrique par rapport à  $\mu$  (paramètre de position), le paramètre  $\gamma$  donnant une information sur l'étalement de la fonction (paramètre d'échelle).

La loi de Cauchy possède une queue qui décroît lentement et donne une probabilité assez importante aux valeurs extrêmes. Il s'ensuit que la moyenne et la variance ne sont pas définies, et la médiane est gale à  $\mu$ .

#### Quelques propriétés de la loi Cauchy

- La loi de Cauchy est une loi Lévy-stable, c'est à dire que l'addition de variables de loi de Cauchy donne une somme vérifiant aussi une loi de Cauchy.
- L'inverse d'une variable aléatoire, de loi de Cauchy, suit une loi de Cauchy.
- Le quotient de deux variables aléatoires réelles indépendantes suivant des lois normales standards suit une loi de Cauchy.
- La loi de Cauchy (avec la loi normale et la loi de Lévy) est un cas particulier de lois stables.
- la loi de Cauchy est aussi celle de la loi de Student à un seul degré de liberté.

La fonction de densité d'une loi de Cauchy rappelle celle d'une loi normale, à savoir une forme de cloche, mais avec un étalement plus large (voir figure ci-dessous).

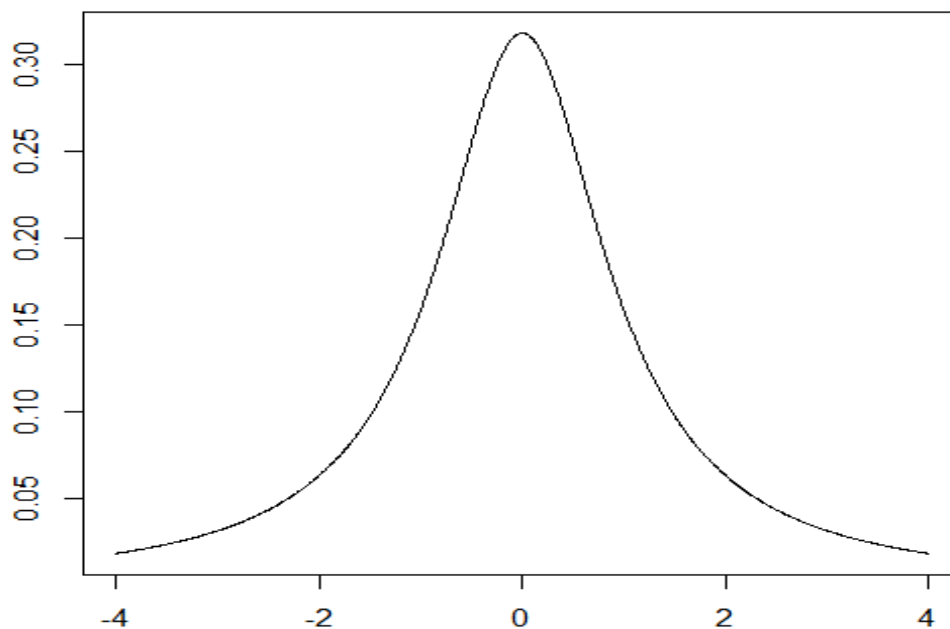


FIGURE 3.1 – Simulation de la densité suivant la loi de Cauchy.

Ce chapitre est consacré à la preuve de l'existence et de la consistance d'un estimateur paramétrique pour une classe générale de séries chronologiques comprenant ARMA, AR(1), GARCH, ARCH(1), ARMA-GARCH, APARCH, ARMA-APARCH, ..., cette classe a été définie et étudiée dans Doukhan et Wintenberger (2007), Bardet et Wintenberger (2009) et Bardet et al.(2012), Bardet, Boularouk et Djaballah (2017).

### 3.3 Les processus causaux affine à erreur de loi de Cauchy

Nous considérons un échantillon observé  $(X_1, \dots, X_n)$  où le processus  $(X_t)_{t \in \mathbb{Z}}$  est une solution de l'équation :

$$X_t = M_{\theta_0}(X_{t-1}, X_{t-2}, \dots)\varepsilon_t + f_{\theta_0}(X_{t-1}, X_{t-2}, \dots), t \in \mathbb{Z} \quad (3.10)$$

tel que :

$\theta_0 \in \Theta \subset \mathbb{R}^d$ , le paramètre inconnu, également appelé le vrai paramètre ;  $(\varepsilon_t)_{t \in \mathbb{Z}}$  une suite de variables aléatoires indépendantes centrées identiquement distribuées.

Si  $M_{\theta_0}(X_{t-1}, X_{t-2}, \dots) = 1$  et  $f_{\theta_0}(X_{t-1}, X_{t-2}, \dots) = \alpha_0 X_{t-1}$  avec  $|\alpha_0| < 1$  alors  $(X_t)$  est un processus AR(1) causal. Dans Doukhan et Wintenberger (2007) et Bardet et Wintenberger (2009), il a été prouvé que les séries les plus célèbres utilisées en économétrie, comme les processus ARMA, AR (1), GARCH, ARCH (1), TAR, ARMA-GARCH peuvent s'écrire comme une solution causale et stationnaire de l'équation (3.10).

La définition de l'EQMV Gaussien est explicitement obtenue dans Bardet et Wintenberger (2009) sous l'hypothèse que  $(\varepsilon_t)$  est une suite Gaussienne. Mais malgré ça, la Quasi-Vraisemblance pourrait être appliquée lorsque la distribution de probabilité de  $(\varepsilon_t)$  est non Gaussienne, comme dans Bardet, Boularouk et Djaballah (2017) qui considère l'EQMV Laplacien.

Dans ce travail, nous nous sommes intéressés à la classe des processus causaux affines dont la loi du bruit est la loi de Cauchy et non pas celle de Gauss utilisé habituellement. Les lois stables non-gaussiennes sont une alternative toute naturelle car elles sont une généralisation de la loi normale et prennent en compte des queues lourdes, telle que la loi de Cauchy. Donc, dans le cas du bruit à queue lourde, un contraste de Cauchy nous donne un meilleur modèle pour représenter le processus stochastique qu'un contraste de Gauss.

Dans un premier temps nous montrons l'existence d'une solution stationnaire de (3.10), ensuite nous prouvons la consistance de l'estimateur du QMV.

### 3.3.1 Existence et stationnarité

Comme dans Doukhan et Wintenberger (2007) et Bardet et Wintenberger(2009), Bardet, Yakoub et Djaballah (2017), l'existence d'une solution stationnaire du modèle (3.10) est prouvée sous des conditions Lipschitziennes sur  $f_{\theta}, M_{\theta}$ , pour l'obtention de l'existence d'une solution causale et ergodique de (3.10). Pour cela nous introduisons le symbole  $\Psi$  pour les

fonctions  $f, M$  et nous notons  $\|\cdot\|_\varphi$  la norme d'Orlicz. Pour  $k = 0, 1, 2$  et un certain sous espace  $\Theta$ , définissons la condition lipchitzienne sur  $\Psi_\theta$ .

**Hypothèse :**

$(A_k(\Psi, \Theta)) : \forall x \in \mathbb{R}^\infty, \theta \in \Theta \rightarrow \Psi_\theta(x) \in C^k(\Theta), \partial_\theta^k \Psi_\theta$  satisfait  $|\partial_\theta^k \Psi_\theta(0)|_\Theta < \infty$  et il existe une suite  $(\partial_\theta^k(\Psi_\theta, \Theta))_j$  de nombres non négatifs telle que  $\forall x, y \in \mathbb{R}^\mathbb{N}$

$$\|\Psi_\theta(x) - \Psi_\theta(y)\|_\varphi \leq \sum_{j=1}^{\infty} \alpha_j(\Psi_\theta) |x_j - y_j| \quad \text{avec} \quad \sum_{j=1}^{\infty} \alpha_j^k(\Psi, \Theta) < 1$$

Afin d'assurer une solution stationnaire de (3.10), définissons l'ensemble :

$$\Theta = \left\{ \begin{array}{l} \theta \in \mathbb{R}, (A_0(f, \{\theta\})) \text{ et } (A_0(M, \{\theta\})) \text{ sont vérifiées,} \\ \sum_{j=1}^{\infty} \alpha_j^{(0)}\{f, \{\theta\}\} + \|\varepsilon_0\|_\varphi \sum_{j=1}^{\infty} \alpha_j M, \{\theta\} < 1 \end{array} \right\}$$

Puis, à partir de l'article de Doukhan et Wintenberger (2007), nous obtenons la proposition ci dessous.

**Proposition 3.1.** *Si  $\theta_0 \in \Theta$ , alors il existe une unique solution causale de (3.10)  $X_t$  indépendante de  $(\varepsilon_i)_{i>t}$  pour  $t \in \mathbb{Z}$  qui est stationnaire, ergodique et satisfait  $\|\varepsilon_0\|_\varphi < \infty$ .*

L'obtention d'une solution causale, stationnaire et ergodique du modèle définit dans l'équation est démontré sous des conditions lipchitziennes sur  $M_\theta$  et  $f_\theta$  comme suit :

$$\|M(x_n) - M(y_n)\| \leq \sum_{j=1}^{\infty} \alpha_j \|x_j - y_j\|,$$

$$\|f(x_n) - f(y_n)\| \leq \sum_{j=1}^{\infty} \alpha_j \|x_j - y_j\|,$$

Afin d'assurer l'existence d'une solution stationnaire il faut définir un espace de stationnarité.

Montrons que  $\|\varepsilon_0\|_\varphi < \infty$  :

$$\|\varepsilon_0\|_\varphi = \inf \left\{ u > 0 \text{ avec } Es\varphi \left( \frac{\|\varepsilon_0\|}{u} \right) \leq 1 \right\}.$$

Nous considérons le modèle affine défini par

$$X_t = M_\theta \varepsilon_t + f_\theta$$

Avec  $\varepsilon_t$  suit une loi de Cauchy de paramètre de position 0 et paramètre d'échelle 1

$$f_\varepsilon(\varepsilon) = \frac{1}{\pi(1 + \varepsilon^2)}, \quad -\infty < \varepsilon < \infty$$

$$\varepsilon_t = \frac{X_t - f_\theta}{M_\theta}.$$

Selon Doukhan et Wintenberger (2007) et en appliquant l'équation (1.5) du chapitre 1 dans le cas de la loi de Cauchy, nous obtenons l'espace de stationnarité suivant :

$$\Theta_2 = \left\{ \left( -\frac{\pi}{4} \tan \frac{\pi}{4} \right) \sum_{i=1}^{\infty} Lip M_i + \sum_{i=1}^{\infty} Lip f_i < 1 \right\}.$$

### 3.3.2 Définition de l'estimateur QMV de loi de Cauchy

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $X$  solution de (3.10) où  $\theta \in \mathbb{R}$  est inconnu. Pour estimer  $\theta$ , nous considérons la log-vraisemblance de  $(X_1, \dots, X_n)$  conditionnellement à  $(X_0, X_{-1}, \dots)$ . Si  $g$  est la densité de probabilité (par rapport à la mesure de Lebesgue) de  $\varepsilon_0$ , alors, à partir de la définition causale affine de  $X$ , la log-vraisemblance peut s'écrire sous la forme :

$$\log(L_\theta(X_1, \dots, X_n)) = \sum_{i=1}^n \log \left( g \left( \frac{X_t - f_\theta^t}{M_\theta^t} \right) \right)$$

où  $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots)$  et  $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots)$ , avec l'hypothèse que  $M_\theta^t > 0$ . Cependant, comme  $X_0, X_{-1}, \dots$  sont inconnus, en général les  $M_\theta^t$  et  $f_\theta^t$  ne sont pas calculables. Pour cela, au lieu d'utiliser directement la log-vraisemblance, nous considérons la quasi-log-vraisemblance qui est définie par :

$$\log(Q(L_\theta(X_1, \dots, X_n))) = \sum_{i=1}^n \log\left(g\left(\frac{X_t - \hat{f}_\theta^t}{\hat{M}_\theta^t}\right)\right),$$

avec  $\hat{f}_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots, X_1, u)$  et  $\hat{M}_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots, X_1, u)$ , où  $u = (u_n)_{n \in \mathbb{N}}$  est une suite finie. Le choix de  $(u_n)_{n \in \mathbb{N}}$  n'influe pas le comportement asymptotique de  $L_n$ , et  $(u_n)$  pourrait être choisie comme une suite égale à zéro. Enfin, l'estimateur de quasi-maximum de vraisemblance (EQMV) est défini par :

$$\hat{\theta}_n = \text{Arg} \max_{\theta \in \Theta} \log(Q L_\theta(X_1, \dots, X_n)).$$

Dans le présent travail, nous considérons la distribution de Cauchy de paramètres  $\mu = 0$  et  $\gamma = 1$ , qui a pour densité :

$$f(x) = \frac{1}{\pi(x^2 + 1)}. \tag{3.11}$$

Nous définissons respectivement la vraisemblance et la quasi-vraisemblance de Cauchy :

De l'équation (3.10) nous obtenons :

$$\varepsilon_t = M_\theta^{-1}(x_t - f_\theta).$$

En faisant un changement de variable

$$(x_t) = f_\varepsilon(M_\theta^{-1}(x_t - f_\theta))M_\theta^{-1}.$$

Nous remplaçons  $f_\varepsilon$  de l'équation (3.11) par son expression, nous obtenons :

$$f_x(x_t) = \frac{1}{\pi [(M_\theta^{-1}(x_t - f_\theta))^2 + 1]} M_\theta^{-1}.$$

Afin d'estimer  $\theta$ , nous considérons la log de vraisemblance de  $(X_1, \dots, X_n)$  conditionnelle à  $(X_0, X_{-1}, \dots)$ . La fonction de vraisemblance s'écrit comme suit :

$$L = \prod_{t=1}^n f_x(x_t) = \prod_{t=1}^n \left( \frac{1}{\pi [(M_\theta^{-1}(x_t - f_\theta))^2 + 1]} M_\theta^{-1} \right).$$

La fonction de log-vraisemblance correspondante est donnée par :

$$\begin{aligned} \log L &= \sum_{t=1}^n (-\log [\pi((M_\theta^{-1}x_t - f_\theta)^2 + 1)] - \log M_\theta) \\ &= \sum_{t=1}^n (-\log \pi - \log [(M_\theta^{-1}(x_t - f_\theta))^2 + 1] - \log M_\theta) \\ &= -n \log \pi - \sum_{t=1}^n \log [(M_\theta^{-1}(x_t - f_\theta))^2 + 1] - n \log M_\theta. \end{aligned}$$

Nous obtenons

$$L_n(\theta) = - \sum_{t=1}^n q_t(\theta),$$

avec

$$q_t(\theta) = \log [(M_\theta^{-1}(x_t - f_\theta))^2 + 1] + \log M_\theta,$$

$$q_t(\theta) = \log \left( \frac{(x_t - f_\theta)^2}{M_\theta^2} + 1 \right) + \log M_\theta.$$

La quasi log-vraisemblance est donnée par :

$$\hat{L}_n(\theta) = - \sum_{t=1}^n \hat{q}_t(\theta),$$

avec

$$\hat{q}_t(\theta) = \log \left[ (\hat{M}_\theta^{-1}(x_t - \hat{f}_\theta))^2 + 1 \right] + \log \hat{M}_\theta.$$

$$\hat{q}_t(\theta) = \log \left( \frac{(x_t - \hat{f}_\theta)^2}{\hat{M}_\theta^2} + 1 \right) + \log \hat{M}_\theta.$$

Pour trouver l'estimateur du maximum de vraisemblance  $\hat{\theta}_n$ , il suffit de maximiser la quasi-maximum de vraisemblance  $\hat{L}_n$  :

$$\hat{\theta}_n = \arg \max \hat{L}_n(\theta).$$

Nous limitons l'ensemble  $\Theta$  de manière à ce qu'une solution  $(X_t)$  de (3.10) stationnaire d'ordre 1 ou 2 existe. Des conditions supplémentaires sont également nécessaires pour assurer la consistance de  $\hat{\theta}_n$ .

### Hypothèses nécessaires pour la convergence d'EQMV-Cauchy

L'EQMV Cauchy pourrait converger et être asymptotiquement Gaussien mais cela nécessite quelques hypothèses supplémentaires sur  $\Theta$  et les fonctions  $f_\theta$  et  $M_\theta$  :

- **A1** (Compacité) :  $\Theta$  est un ensemble compact.
- **A2** (Borne inférieure de la variance conditionnelle) : il existe une constante  $\underline{M} > 0$  tel que  $\forall \theta \in \Theta$ , alors  $M_\theta(x) > \underline{M}$  pour tout  $x \in \mathbb{R}^N$ .
- **A3** (Identifiabilité) : Les fonctions  $f_\theta$  et  $M_\theta$  sont telles que : pour tout  $\theta_1, \theta_2 \in \Theta$ , alors  $M_{\theta_1} = M_{\theta_2}$  et  $f_{\theta_1} = f_{\theta_2}$  implique que  $\theta_1 = \theta_2$

### 3.3.3 Comportement asymptotique d'EQMV de Cauchy

#### Propriétés asymptotiques du quasi vraisemblance

La quasi vraisemblance est une approximation de la vraisemblance obtenue en remplaçant  $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots)$  et  $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots)$  par  $\hat{f}_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots, X_1, u)$  et  $\hat{M}_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots, X_1, u)$ ,  $u = (u_n)_{n>0}$  est égale à zéro.

**Consistance forte**

**Proposition 3.2.** *Supposons que les hypothèses A1, A2 et A3 sont satisfaites et que  $\theta_0 \in \Theta$ . Soit  $X$  une solution stationnaire de (3.10). Si  $(A_0(f, \Theta))$  et  $(A_0(M, \Theta))$  demeurent avec :*

$$\alpha_j^0(f, \Theta) + \alpha_j^0(M, \Theta) = O(j^{-\ell}) \quad \text{pour certain } \ell > 3/2, \quad (3.12)$$

alors, l'EQMV-Cauchy  $(\hat{\theta}_n)_n$  converge fortement, c'est-à-dire

$$\theta_n \xrightarrow[n \rightarrow \infty]{} \theta_0 \quad P_{\theta_0} p.s$$

*Démonstration.* Pour démontrer la consistance il faut passer par la démonstration en deux parties :

i) la première partie : une loi uniforme en  $\Theta$  satisfait la loi forte des grands nombres, alors on aura  $\frac{1}{n} \hat{L}_n(\theta)$  converge vers  $L(\theta) = -E[q_0(\theta)]$ .

ii) la deuxième partie : prouver que  $L(\theta)$  admet un maximum unique en  $\theta_0$

La preuve i :

$$\begin{aligned}
 \left| \log \hat{M}_\theta - \log M_\theta \right| &\leq \log \left( \frac{|\hat{M}_\theta - M_\theta|}{M} + 1 \right) \\
 &\leq \log \left( 1 + \sum_{j=t+1}^{\infty} \left( \frac{\alpha_j(M) x_{t-j}}{M} \right) \right) \\
 &\leq \sum_{j=t+1}^{\infty} \log \left( 1 + \frac{\alpha_j(M) x_{t-j}}{M} \right) \\
 E \left| \log \hat{M}_\theta - \log M_\theta \right| &\leq E \sum_{j=t+1}^{\infty} \log \left( 1 + \frac{\alpha_j(M) X_{t-j}}{M} \right) \\
 &\leq \sum_{j=t+1}^{\infty} E \left( \log \left( 1 + \frac{\alpha_j(M) X_{t-j}}{M} \right) \right) \\
 &\leq \sum_{j=t+1}^{\infty} \int_0^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx \dots (*) \\
 &\leq \sum_{j=t+1}^{\infty} \left[ \int_0^1 \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx + \int_1^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx \right]
 \end{aligned}$$

Nous avons

$$\int_0^1 \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx \leq c\alpha_j$$

et

$$\int_1^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx \leq \int_1^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2} dx.$$

En utilisant un changement de variable nous obtenons :

$$\begin{aligned} \int_1^{+\infty} \frac{\log(1 + \alpha_j x)}{x^2} dx &= \int_{\alpha_j}^{+\infty} \frac{\log(1 + y)}{\frac{1}{\alpha_j^2} y^2} \frac{1}{\alpha_j} dy \\ &= \int_{\alpha_j}^{+\infty} \frac{\log(1 + y)}{\frac{1}{\alpha_j} y^2} dy = \alpha_j \int_{\alpha_j}^{+\infty} \frac{\log(1 + y)}{y^2} dy \\ \int_1^{+\infty} \frac{\log(1 + \alpha_j x)}{x^2} dx &= \alpha_j \int_{\alpha_j}^{+\infty} \frac{\log(1 + y)}{y^2} dy \dots (**) \end{aligned}$$

Nous utilisons l'integration par partie

$$\begin{aligned} \int_{\alpha_j}^{+\infty} \frac{\log(1 + y)}{y^2} dy &= \left[ \frac{-1}{y} \ln(1 + y) \right]_{\alpha_j}^{+\infty} + \int_{\alpha_j}^{+\infty} \frac{1}{y(1 + y)} dy \\ &= 1 + \int_{\alpha_j}^{+\infty} \frac{1}{y(1 + y)} dy \text{ avec } 0 < \alpha_j < 1 \\ &= 1 + \int_{\alpha_j}^{+\infty} \frac{1}{y} - \frac{1}{1 + y} dy \\ &= 1 + [\ln y - \ln(1 + y)]_{\alpha_j}^{+\infty} \\ &= 1 + \left[ \ln \frac{y}{1 + y} \right]_{\alpha_j}^{+\infty} = 1 - \ln \left( \frac{\alpha_j}{1 + \alpha_j} \right) = 1 + \ln \left( \frac{1 + \alpha_j}{\alpha_j} \right) \end{aligned}$$

$$\text{car } \lim_{y \rightarrow +\infty} \left( \ln \frac{y}{1 + y} \right) = \lim_{y \rightarrow +\infty} \left( \ln \frac{y}{y(1 + \frac{1}{y})} \right) = 0$$

$$(**) = \alpha_j \left( 1 + \ln \left( \frac{1 + \alpha_j}{\alpha_j} \right) \right)$$

D'où

$$(*) = \sum_{j=t+1}^{\infty} \int_0^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} dx \leq \sum_{j=t+1}^{\infty} (c\alpha_j + 18\alpha_j + \alpha_j \ln \alpha_j) \text{ avec } \sum_{j=t+1}^{\infty} \alpha_j = 1$$

Donc

$$\sum_{j=t+1}^{\infty} \int_0^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} \leq c\alpha_j + 18\alpha_j + \sum_{j=t+1}^{\infty} \alpha_j \ln \alpha_j$$

$$\sum_{j=t+1}^{\infty} \int_0^{+\infty} \frac{\log \left( 1 + \frac{\alpha_j(M)x}{M} \right)}{x^2 + 1} \leq \sum_{j=t+1}^{\infty} \alpha_j \ln \alpha_j \text{ pour } c \leq -18$$

avec  $\sum \alpha_j < \infty$

Et d'après le lemme 1 de l'article de Bardet et Wintenberger (2009), nous avons :

$$E [\|f_{\theta}\|_{\Theta}^r + \|M_{\theta}\|_{\Theta}^r] < \infty,$$

avec  $0 < r < 1$ , par conséquent :

$$\forall t \in Z, E [\|q_t(\theta)\|_{\Theta}] < \infty$$

Nous montrons que  $\left\| \frac{L_n(\theta)}{n} - L(\theta) \right\| \xrightarrow[n \rightarrow \infty]{p.s.} 0$

$$\begin{aligned}
 |\hat{q}_t(\theta) - q_t(\theta)| &= \left| \left( \log \left( \frac{(x_t - \hat{f}_\theta)^2}{\hat{M}_\theta^2} + 1 \right) + \log \hat{M}_\theta \right) - \left( \log \left( \frac{(x_t - f_\theta)^2}{M_\theta^2} + 1 \right) + \log M_\theta \right) \right| \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \left( \frac{(x_t - \hat{f}_\theta)^2}{\hat{M}_\theta^2} + 1 \right) - \log \left( \frac{(x_t - f_\theta)^2}{M_\theta^2} + 1 \right) \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \frac{\left( \frac{(x_t - \hat{f}_\theta)^2}{\hat{M}_\theta^2} + 1 \right)}{\left( \frac{(x_t - f_\theta)^2}{M_\theta^2} + 1 \right)} \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \frac{\frac{(x_t - \hat{f}_\theta)^2 + \hat{M}_\theta^2}{\hat{M}_\theta^2}}{\frac{(x_t - f_\theta)^2 + M_\theta^2}{M_\theta^2}} \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \left( \frac{(x_t - \hat{f}_\theta)^2 + \hat{M}_\theta^2}{\hat{M}_\theta^2} \frac{M_\theta^2}{(x_t - f_\theta)^2 + M_\theta^2} \right) \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \left( \frac{(x_t - \hat{f}_\theta)^2 M_\theta^2 + \hat{M}_\theta^2 M_\theta^2}{(x_t - f_\theta)^2 \hat{M}_\theta^2 + \hat{M}_\theta^2 M_\theta^2} \right) \\
 &= \left| \log \hat{M}_\theta - \log M_\theta \right| + \log \left( \frac{(x_t - \hat{f}_\theta)^2 M_\theta^2 + \hat{M}_\theta^2 M_\theta^2}{(x_t - f_\theta)^2 \hat{M}_\theta^2 + \hat{M}_\theta^2 M_\theta^2} \right) - \log \left( \frac{(x_t - f_\theta)^2 \hat{M}_\theta^2 + \hat{M}_\theta^2 M_\theta^2}{(x_t - f_\theta)^2 \hat{M}_\theta^2 + \hat{M}_\theta^2 M_\theta^2} \right) \\
 &\leq \left( \sum_{j=t+1}^{\infty} \alpha_j \ln \alpha_j + (x_t - \hat{f}_\theta)^2 M_\theta^2 - (x_t - f_\theta)^2 \hat{M}_\theta^2 \right) \\
 &\leq \sum_{j=t+1}^{\infty} \alpha_j \ln \alpha_j + (x_t - \hat{f}_\theta)^2 (\hat{M}_\theta^2 - M_\theta^2) \\
 &\quad + (2x_t - \hat{f}_\theta - f_\theta) \hat{M}_\theta^2 (\hat{f}_\theta - f_\theta) \\
 &\leq \sum_{j=t+1}^{\infty} \alpha_j \ln \alpha_j + 2 (\|x_t\| + \|\hat{f}_\theta\|) \|\hat{M}_\theta^2 - M_\theta^2\| \\
 &\quad + (2\|x_t\| + \|\hat{f}_\theta\| + \|f_\theta\|) \|\hat{M}_\theta^2\| \|\hat{f}_\theta - f_\theta\|.
 \end{aligned}$$

$$\sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)| \leq \left( \|x_t\| + \|\hat{f}_\theta\| + \|f_\theta\| \right) \times \left( \left\| \hat{M}_\theta^2 - M_\theta^2 \right\| + \left\| \hat{f}_\theta - f_\theta \right\| \right).$$

D'après les inégalités de Hölder et de Minkowski, nous obtenons

$$\begin{aligned} E \left[ |\hat{q}_t(\theta) - q_t(\theta)|^{2/3} \right] &\leq C \left( E \left[ \|X_t\| + \|\hat{f}_\theta\| + \|f_\theta\| \right]^2 \right)^{1/3} \\ &\quad \times \left( E \left[ \left\| \hat{M}_\theta^2 - M_\theta^2 \right\| \right] + E \left[ \left\| \hat{f}_\theta - f_\theta \right\| \right] \right)^{2/3} \\ &\leq C' \left( \sum_{j \geq t} \left[ \alpha_j^{(0)}(f, \Theta) + \alpha_j^{(0)}(M, \Theta) \right] \right)^{2/3} \end{aligned} \quad (3.13)$$

Maintenant, considérons, pour  $n \in \mathbb{N}^*$

$$S_n = \sum_{i=1}^n \frac{1}{i} \sup_{\theta \in \Theta} |\hat{q}_i(\theta) - q_i(\theta)|$$

Appliquer le lemme de Kronecker (voir Feller (1966) page 238), si

$$\lim_{n \rightarrow \infty} S_n < \infty \quad p.s$$

alors

$$\frac{1}{n} \left\| \hat{L}_n - L_n \right\| \xrightarrow[n \rightarrow \infty]{p.s} 0$$

Suivant les arguments de Feller, il reste à montrer que, pour tout  $\varepsilon > 0$ ,

$$\mathbb{P}(\forall n \in \mathbb{N}, \exists m > n \text{ tel que } |S_m - S_n| > \varepsilon) = \mathbb{P}(A) = 0$$

Soit  $\varepsilon > 0$ , et notons

$$A_{m,n} = \{|S_m - S_n| > \varepsilon\}$$

Pour  $m > n$ . nous remarquons que  $A = \bigcap_{n \in \mathbb{N}} \bigcup_{m > n} A_{m,n}$ . pour  $n \in \mathbb{N}^*$ , la suite  $(A_{m,n})_{m > n}$  est évidemment en augmentation, et, si  $A_n = \bigcup_{m > n} A_{m,n}$ . alors  $\lim_{m \rightarrow \infty} \mathbb{P}(A_{m,n}) = \mathbb{P}(A_n)$ . Remarque que  $(A_n)_{n \in \mathbb{N}}$  est une suite décroissante d'ensembles et, par conséquent,

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \mathbb{P}(A_{m,n}) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$$

De l'inégalité de Bienaymé – Chebyshev,

$$\begin{aligned} \mathbb{P}(A_{m,n}) &= \mathbb{P}\left(\sum_{t=n+1}^m \frac{1}{t} \sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)| > \varepsilon\right) \\ &\leq \frac{1}{\varepsilon^{2/3}} \mathbb{E}\left[\left(\sum_{t=n+1}^m \frac{1}{t} \sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)|\right)^{2/3}\right] \\ &\leq \frac{1}{\varepsilon^{2/3}} \sum_{t=n+1}^m \frac{1}{t^{2/3}} \mathbb{E}\left[\sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)|^{2/3}\right] \end{aligned}$$

En utilisant (3.4) et la condition (3.3), depuis  $\ell > 3/2$ , il existe  $C > 0$  tel que

$$\left(\sum_{j=t}^{\infty} \alpha_j^{(0)}(f, \Theta) + \alpha_j^{(0)}(M, \Theta)\right)^{2/3} \leq \frac{C}{t^{2(\ell-1)/3}}$$

Ainsi,  $t^{-2/3} \mathbb{E}\left[\left(\sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)|\right)^{2/3}\right] < C(t^{-2\ell/3})$  pour  $C > 0$ , et

$$\sum_{j=t}^{\infty} \frac{1}{t^{-2/3}} \mathbb{E}\left[\sup_{\theta \in \Theta} |\hat{q}_t(\theta) - q_t(\theta)|^{2/3}\right] < \infty \quad p.s \ell > 3/2$$

Ainsi

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \mathbb{P}(A_{m,n}) \rightarrow 0$$

et

$$\frac{1}{n} \left\| \hat{L}_n - L_n \right\|_{\Theta} \xrightarrow[p.s]{n \rightarrow \infty} 0.$$

■ (ii) Voir la proposition de Jeantheau (1998) Normalité asymptotique

**Théorème 3.2.** *Sous certaines hypothèses la normalité asymptotique de  $\theta_n$  est donnée par :*

$$\sqrt{n}(\theta_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, J^{-1})$$

où  $J$  est une matrice,  $J = E \left[ \frac{\partial \hat{q}_t}{\partial \theta}(\theta_0) \frac{\hat{q}_t}{\partial \theta'}(\theta_0) \right]$ , avec  $\hat{q}_t(\theta) = \log \left[ (\hat{M}_\theta^{-1} (x_t - \hat{f}_\theta))^2 + 1 \right] + \log \hat{M}_\theta$

*Démonstration.* En suivant l'approche classique, nous établirons la normalité asymptotique de  $\hat{\theta}_n$  au moyen du Développement de Taylor de  $\frac{\partial L_n(\hat{\theta}_n)}{\partial \theta}$  avec  $L_n(\theta) = - \sum_{t=1}^n q_t(\theta)$ , où

$$q_t(\theta) = \log \left( \frac{(x_t - f_\theta)^2}{M_\theta^2} + 1 \right) + \log M_\theta$$

Le developpement de Taylor de  $\frac{\partial}{\partial \theta} \left[ L_n(\hat{\theta}) \right]$  est

$$\frac{\partial}{\partial \theta} \left[ L_n(\hat{\theta}) \right] = \frac{\partial}{\partial \theta} L_n(\theta) + \frac{\partial^2}{\partial \theta^2} L_n(\theta) (\hat{\theta} - \theta)$$

Pour démontrer le Théorème, nous décomposons la preuve en 6 étapes:

1. Les dérivées premières du quasi-log-vraisemblance.

nous établissons les dérivées premières du critère c'est à dire calculer

$$\frac{\partial L_n(\theta)}{\partial \theta_i} = - \sum_{t=1}^n \frac{\partial q_t(\theta)}{\partial \theta_i} \quad \text{et} \quad \frac{\partial^2}{\partial \theta^2} L_n(\theta) = - \sum_{t=1}^n \frac{\partial^2 q_t(\theta)}{\partial \theta_i^2}$$

2 Existence des moments de tout ordre du score.

Obtenir l'existence des moments des dérivées premières du critère de log-vraisemblance

3 Normalité asymptotique du vecteur du score.

Le processus  $\{\partial q_t(\theta_0)/\partial \theta\}_t$  est stationnaire et  $\partial q_t(\theta_0)/\partial \theta$  est mesurable par rapport à la tribu  $\mathcal{F}_t$  engendrée par  $\{\eta_u, u \leq t\}$ . Le théorème central limite entraine que

$$\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \left[ - \sum_{t=1}^n q_t(\theta) \right] \xrightarrow{L} N(0, J(\theta))$$

4 Convergence de la matrice  $J$  : Pour montrer la convergence de la matrice  $J$ , il est nécessaire d'avoir les expressions des dérivées du critère à l'ordre 2.

$$\frac{1}{n} - \sum_{t=1}^n \frac{\partial^2 q_t(\theta)}{\partial \theta_i \partial \theta_j} \rightarrow J(i, j) \text{ en probabilité}$$

5 Inversibilité de la matrice  $J$ .

Pour montrer que la matrice  $J$  est inversible on se base sur le calcul de

$$E \left( \frac{\partial^2 q_t(\theta_0)}{\partial \theta_i \partial \theta_j} \right)_{ij}$$

6 Oubli des valeurs initiales.

Montrer que les valeurs initiales n'ont aucun effet sur les dérivées du critère. Ce que nous cherchons à prouver est

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[ \frac{\partial q_n(\theta_0)}{\partial \theta} - \frac{\partial \hat{q}_n(\theta)_0}{\partial \theta} \right] \right\| = 0.$$

■

# Conclusion

Les investisseurs et les gestionnaires des risques sont particulièrement sensibles à la survenance des pertes substantielles. Dans ce contexte, il est primordial de disposer des outils permettant de mieux quantifier les risques liés aux fluctuations des séries financières. Nous avons adopté une approche conditionnelle de la théorie des valeurs extrêmes et plus précisément la famille des lois alpha stables. Cette approche consiste à filtrer les chroniques à travers un des modèles causaux affines à innovations  $\alpha$  stables.

Dans ce travail nous avons étudié les modèles GARCH à innovations de lois  $\alpha$  stables. Nous avons choisi la loi de Student comme loi  $\alpha$  stable, nous avons réussi à prouver la consistance et la normalité asymptotique dans ce cas. Nous avons constaté que la qualité de l'estimateur est satisfaisante, car elle prend en compte l'asymétrie et la lourdeur de la queue.

Par ailleurs, nous avons montré la consistance des estimateurs de quasi vraisemblance pour une classe générale des processus causaux affines, sous l'hypothèse que le bruit est non Gaussien, précisément un bruit de loi de Cauchy. Nous avons montré la consistance de l'estimateur, et à travers des expériences de type Monte-Carlo, ensuite nous avons pu vérifier les résultats théoriques obtenus. D'où les simulations obtenues par  $\alpha = 1$  (loi de Cauchy), nous ont permis de voir que malgré l'absence de tous les moments, les résultats de la simulation sont acceptables.

Le travail que nous avons réalisé dans le chapitre 3 de cette thèse pourrait être complété en étudiant la normalité asymptotique des estimateurs proposés. Ce travail ouvre les portes à la recherche dans le domaine des lois  $\alpha$  stables adaptées aux séries chronologiques. Il sera intéressant de vérifier que le maximum de vraisemblance possède un maximum, moyennant des hypothèses supplémentaires, lorsqu'il s'agit des modèles GARCH à innovations de la loi de Cauchy.

# Annexe A : Applications numériques

## 3.4 Application aux données simulées

Pour illustrer la performance de l'estimateur obtenu dans le chapitre 2, nous avons réalisé l'expérience de Monte-Carlo avec Student-EQMV à l'aide du logiciel R.

Nous avons fourni ici les résultats de simulations pour 2 modèles GARCH, à savoir GARCH (1,1) et GARCH (2,1). Tout d'abord, nous générons des séries temporelles GARCH (1,1) dont la loi est de Student avec 4 degrés de liberté et des coefficients différents. Deuxièmement, nous faisons la même chose avec 6 degrés de liberté. Le modèle de série temporelle satisfaisant (1.7) est considéré comme défini par  $X_t = \sigma_t \eta_t$  où  $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ . Enfin, nous considérons un modèle GARCH (2,1) avec  $ddl = 4$ . Nous avons réalisé une étude de simulation de 1000 itérations avec différentes tailles d'échantillon  $n = 500, 1000, 5000$ . Nous utilisons la méthode du maximum de vraisemblance pour estimer les paramètres considérant, d'abord, la vraisemblance de la loi de Student, puis la vraisemblance de la loi de Gauss. Nous avons obtenu les résultats suivants :

Table 1 Estimation des paramètres du processus GARCH (1,1).

$\alpha_0 = 1e - 6, \alpha_1 = 0.1, \beta_1 = 0.85$				$\alpha_0 = 1e - 4, \alpha_1 = 0.15, \beta_1 = 0.8$	
	$n$	$t$	$\mathcal{N}$	$t$	$\mathcal{N}$
$\hat{\alpha}_0$	500	$6.618998e - 07$	$4.907843e - 08$	0.000124845	$5.255343e - 05$
	1000	$1.155526e - 06$	$7.407231e - 07$	0.0001116504	$7.205254e - 05$
	5000	$1.020559e - 06$	$1.341028e - 06$	0.0001011541	$6.663965e - 05$
$\hat{\alpha}_1$	500	$4.885598e - 02$	$6.039987e - 03$	0.155373548	$9.516783e - 02$
	1000	$1.017001e - 01$	$9.257478e - 02$	0.1524693404	$1.759798e - 01$
	5000	$1.006532e - 01$	$8.101442e - 02$	0.1492774444	$1.501456e - 01$
$\hat{\beta}_1$	500	$3.747688e - 01$	$9.886193e - 01$	0.780794820	$8.773186e - 01$
	1000	$8.394982e - 01$	$8.636183e - 01$	0.7913630367	$8.134969e - 01$
	5000	$8.487412e - 01$	$8.331951e - 01$	0.7996407265	$8.241216e - 01$

Nous répondons exactement au même cadre que dans le précédent, pour un processus GARCH (1,1) avec la loi de Student avec 6 degrés de liberté et des coefficients différents.

Table 2 : Estimation des paramètres des processus GARCH (1,1)

$\alpha_0 = 1e - 4, \alpha_1 = 0.15, \beta_1 = 0.8, ddl = 6$			
	$n$	$t$	$\mathcal{N}$
$\hat{\alpha}_0$	500	0.0001335507	$8.205146e - 05$
	1000	0.0001105052	$5.753943e - 05$
	5000	0.0001023195	0.000152148
$\hat{\alpha}_1$	500	0.1507755553	$1.048567e - 01$
	1000	0.1509206659	$1.095628e - 01$
	5000	0.1506089841	0.193173079
$\hat{\beta}_1$	500	0.7756575934	$8.469935e - 01$
	1000	0.7917576317	$8.623664e - 01$
	5000	0.7980165464	0.747204255

Nous notons que les estimations utilisant la vraisemblance de la distribution de Student à 6 degrés de liberté sont plus proches des vraies valeurs que l'utilisation de la distribution de la loi normale. De plus, si nous calculons le RMSE pour cet exemple, pour  $n = 5000$ , voir les 1000 répliques. On a

Table3 : RMSE

	RMSE (Student)	RMSE (Normal)
$\hat{\alpha}_0$	$1.542479e - 05$	$5.214796e - 05$
$\hat{\alpha}_1$	0.01540112	0.04317308
$\hat{\beta}_1$	0.01840762	0.05279575

Le RMSE dans le cas de la vraisemblance de Student est plus petit que dans le cas de la vraisemblance normale, ce qui confirme que la qualité de l'estimateur dans le cas du modèle de Student est meilleure que dans le cas gaussien dans cette étude.

Nous répondons exactement au même cadre que dans le cas précédent, pour un modèle GARCH (2,1)  $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \beta_1 \sigma_{t-1}^2$  avec  $ddl = 4$ . Premièrement avec  $\alpha_0 = 8e - 6, \alpha_1 = 0.1, \alpha_2 = 0.01, \beta_1 = 0.8$  ensuite avec  $\alpha_0 = 1e - 6, \alpha_1 = 0.005, \alpha_2 = 0.1, \beta_1 = 0.85$ . Nous avons obtenu les résultats suivants :

Table 4 : Estimation des paramètres du processus GARCH (2,1).

	$n$	$t$	$\mathcal{N}$	$t$	$\mathcal{N}$
$\hat{\alpha}_0$	500	$1.232938e - 05$	$8.853293e - 06$	$1.388181e - 06$	$1.737720e - 06$
	1000	$9.734662e - 06$	$2.154225e - 05$	$1.137164e - 06$	$8.552304e - 07$
	5000	$8.329115e - 06$	$9.782854e - 06$	$1.020815e - 06$	$1.315348e - 06$
$\hat{\alpha}_1$	500	$8.387863e - 02$	$7.337239e - 02$	$2.064477e - 02$	$1.000000e - 08$
	1000	$8.979399e - 02$	$5.463446e - 02$	$1.329426e - 02$	$1.000000e - 08$
	5000	$9.618984e - 02$	$2.539513e - 02$	$7.724456e - 03$	$1.000000e - 08$
$\hat{\alpha}_2$	500	$4.539315e - 02$	$7.690743e - 02$	$9.304080e - 02$	$1.518007e - 01$
	1000	$3.245545e - 02$	$9.182182e - 02$	$9.358686e - 02$	$1.150697e - 01$
	5000	$1.672801e - 02$	$8.026546e - 02$	$9.824364e - 02$	$9.521211e - 02$
$\hat{\beta}_1$	500	$7.311759e - 01$	$7.557257e - 01$	$8.215239e - 01$	$7.061804e - 01$
	1000	$7.685140e - 01$	$5.852426e - 01$	$8.414804e - 01$	$8.460193e - 01$
	5000	$7.939337e - 01$	$7.647347e - 01$	$8.484660e - 01$	$8.385998e - 01$

Nous pourrions souligner que l'estimateur basé sur une vraisemblance de loi de Student donne des résultats très convaincants lorsque  $n = 5000$  et pas trop mauvais lorsque  $n = 1000$ . En effet, lorsque la distribution conditionnelle du modèle GARCH (1,1) ou GARCH (2,1) est à innovations de loi de Student, l'estimation QMV calculée sur cette distribution est la meilleure (proche du vrai paramètre) par rapport à l'estimation basée sur la loi normale. Le meilleur modèle GARCH peut produire les meilleures estimations. Les tableaux 1, 2 et 3 montrent que le Student-EQMV fournit une estimation plus précise que le Gaussian-EQMV.

### 3.5 Application aux données réelles

Les rendements financiers ont généralement des queues lourdes, cela conduit souvent à une violation de la normalité conditionnelle de l'erreur d'innovation. Par exemple, Bollerslev et Wooldbridge (1992) ont rapporté que le kurtosis d'échantillon des résidus estimés de EQMV gaussien sur les rendements mensuels de S&P500 est de 4,6, dépassant largement le kurtosis gaussien de 3. Par conséquent, il est intuitivement attrayant de développer un EQMV basé sur des -vraisemblances réduites. Classiquement, pour les processus GARCH rencontrés dans le domaine de la finance, la variable aléatoire  $\eta_t$  peut suivre, au lieu d'une loi gaussienne, une loi de Student à  $v$  degrés de liberté.

### 3.5.1 Application aux données S&P500

Le Standard and Poor's 500 (plus communément appelé le *S&P500*) est un indice boursier de 500 sociétés de premier plan cotées sur le marché boursier américain. Considérons la série de l'indice *S&P500*, observée du 29 juin 2012 au 28 juin 2017 (soit 1257 observations).

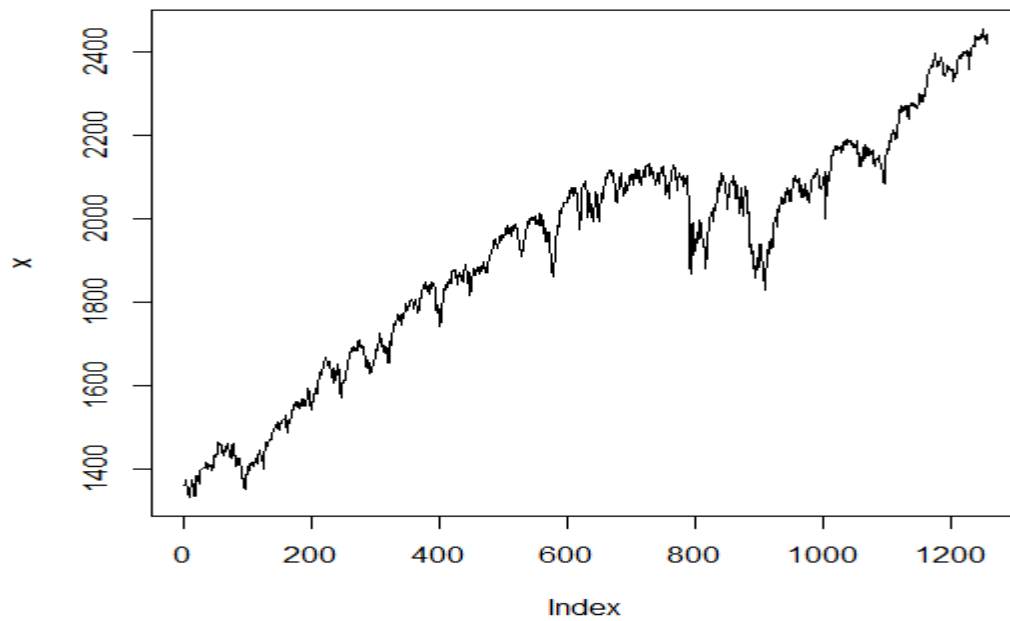


FIGURE 3.2 – L'indice boursier SP500.

Soit  $P_t$  le prix d'un titre au temps  $t$ , on définit donc le rendement au temps  $t$  par

$$R_t = \log\left(\frac{P_t}{P_{t-1}}\right), t = 1, 2, \dots$$

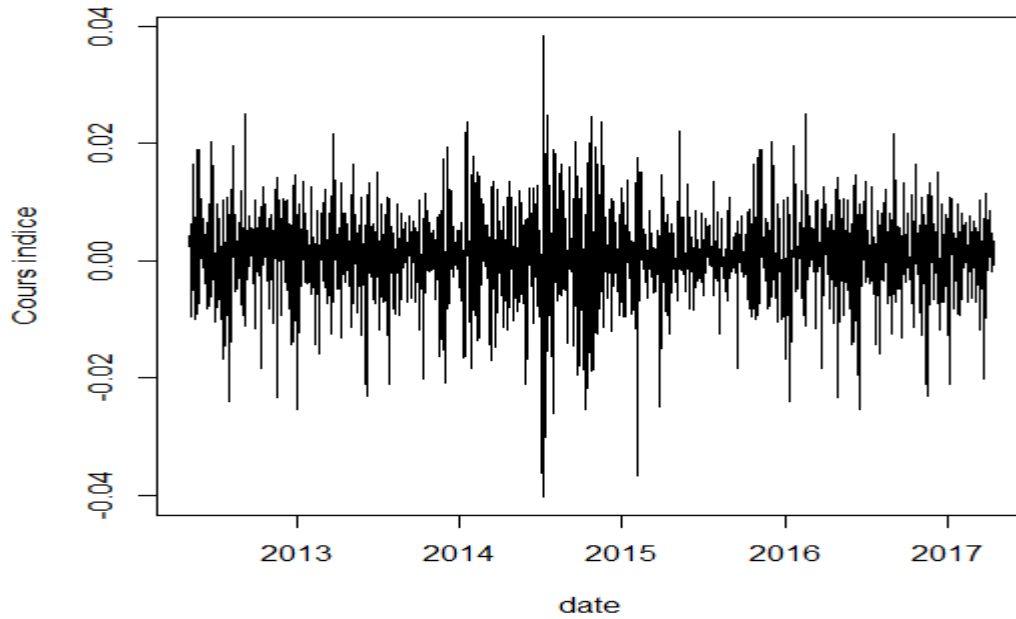


FIGURE 3.3 – Rendement logarithmique de l'indice boursier SP500.

Afin de déterminer le type de modélisation à adopter, l'analyse des moments empiriques de la série de rendements conduit aux conclusions habituelles des études boursières, le kurtosis est différent de 3. Cela signifie que la distribution du modèle n'est pas normale mais plutôt avec des queues épaisses caractérisant une distribution leptokurtique. Cela nous conduit à rejeter l'hypothèse de normalité. La modélisation GARCH de  $(R_t)_t$  revient donc à écrire

$$R_t = \mu + X_t$$

avec

$$X_t = \sigma_t \eta_t$$

où  $\mu$  et  $\sigma_t$  sont respectivement la moyenne et la variance conditionnelle. L'estimation des paramètres des modèles hétéroscédastiques est généralement basée sur la méthode du maximum de vraisemblance. Le modèle choisi est un GARCH (1,1) de paramètres :

$$\hat{\omega} = 7.912453e \times 10^{-6}, \quad \hat{\alpha}_1 = 1.808404 \times 10^{-1}, \quad \hat{\beta}_1 = 6.812782 \times 10^{-1}; \quad \text{et } \hat{\mu} = 7.249024 \times 10^{-4}.$$

Nous remarquons que la somme  $\hat{\alpha}_1 + \hat{\beta}_1 < 1$ ; donc la condition de stationnarité est vérifiée.

D'où, la série des rendements  $(R_t)_t$  admet la représentation suivante :

$$R_t = 7.249024 \times 10^{-4} + \sigma_t \eta_t$$

La variance conditionnelle  $\sigma_t$  est donnée par :

$$\sigma_t^2 = 7.912453 \times 10^{-6} + 1.808404 \times 10^{-1} \times X_{t-1}^2 + 6.812782 \times 10^{-1} \times \sigma_{t-1}^2$$

Pour confirmer que le bruit de la modélisation par le modèle GARCH de la série de prix *S&P500* n'est pas normal, le test Shapiro a été réalisé. L'hypothèse nulle est que la population est normalement distribuée, si la p-value est inférieure au niveau choisi (à 0,05 ici), alors l'hypothèse nulle est rejetée (c'est-à-dire que l'on conclut que les données ne sont pas dérivées d'une distribution normale). Nous notons que la p-value  $< 2.2 \times 10^{-16}$ . Alors la distribution des résidus ne peut pas être gaussienne.

Pour confirmer que le bruit issu de la modélisation par le modèle GARCH des séries *S&P500* est de loi de Student, le test de Kolmogorov-Smirnov a été réalisé.

### Test de Kolmogorov Smirnov sur les résidus

Pour tester l'adéquation d'un échantillon à une loi qui n'est pas nécessairement normale, nous allons utiliser le test de Kolmogorov-Smirnov (KS). Ce test repose sur les propriétés

des fonctions de répartition empiriques. Soit  $X_1, X_2, \dots, X_n$  une suite de  $n$  v.a.i.i.d.. La fonction de répartition empirique de cet échantillon est définie par

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq x\}}(x) = \begin{cases} 0, & x < X_{(1)} \\ \frac{k}{n}, & X_{(k)} \leq x < X_{(k+1)}, k = 1, \dots, n-1 \\ 1, & x \geq X_{(n)} \end{cases}$$

Ce test non paramétrique effectue une comparaison entre la fonction de répartition empirique  $F_n(x)$  et la fonction de répartition théorique  $F(x) = P(X < x)$ . La comparaison entre les deux distributions cumulées, empirique et théorique, s'effectue sur la base de la statistique de Kolmogorov-Smirnov :

$$d = \max |F_n(x) - F(x)|$$

D'après le théorème de Kolmogorov, sous l'hypothèse nulle  $H_0$  d'identité des deux distributions,

$$\{H_0 : F(x) = F_n(x) H_1 : F(x) \neq F_n(x)\}$$

La région critique  $W_\alpha$  de risque  $\alpha$  est déterminée par  $W_\alpha = \{D_n > d_{1-\alpha}(n)\}$  ou  $d_{1-\alpha}(n)$  est le quantile d'ordre  $1 - \alpha$  de la distribution de Kolmogorov-Smirnov pour un échantillon de taille  $n$

Les hypothèses :

$H_0$  : La distribution des résidus suit une loi de Student.

$H_1$  : La distribution des résidus ne suit pas une loi de Student.

La statistique de test calculée est donnée par :  $D_n = 0.023506 < d_{1-\alpha}(n) = d_{1-0.05}(1257) = 0.038$  et la p value=0.2724 > 0.05 Donc on accepte l'hypothèse  $H_0$ , alors la distribution des résidus suit une loi de Student, avec 4 degrés de liberté.

### 3.5.2 Application aux données CAC 40

Le CAC 40 est le principal indice boursier de la place de Paris et voit le jour le 15 juin 1988. L'indice CAC 40 est déterminé à partir des cours des 40 plus importantes capitalisations boursières françaises, correspondant aux 40 entreprises multinationales françaises les plus importantes cotées en continu et en temps réel à la bourse de Paris. Depuis le 1er décembre 2003, le CAC 40 a adopté le système de capitalisation boursière flottante pour s'aligner sur le mode de fonctionnement des grands indices mondiaux. Cela veut dire que, depuis cette date, le nombre de titres disponibles à l'achat sur le marché pour une société est pris en compte dans le calcul de l'indice. Considérons la série de l'indice CAC 40, observée du 18 février 2016 au 17 mars 2021 (soit 1300 observations).

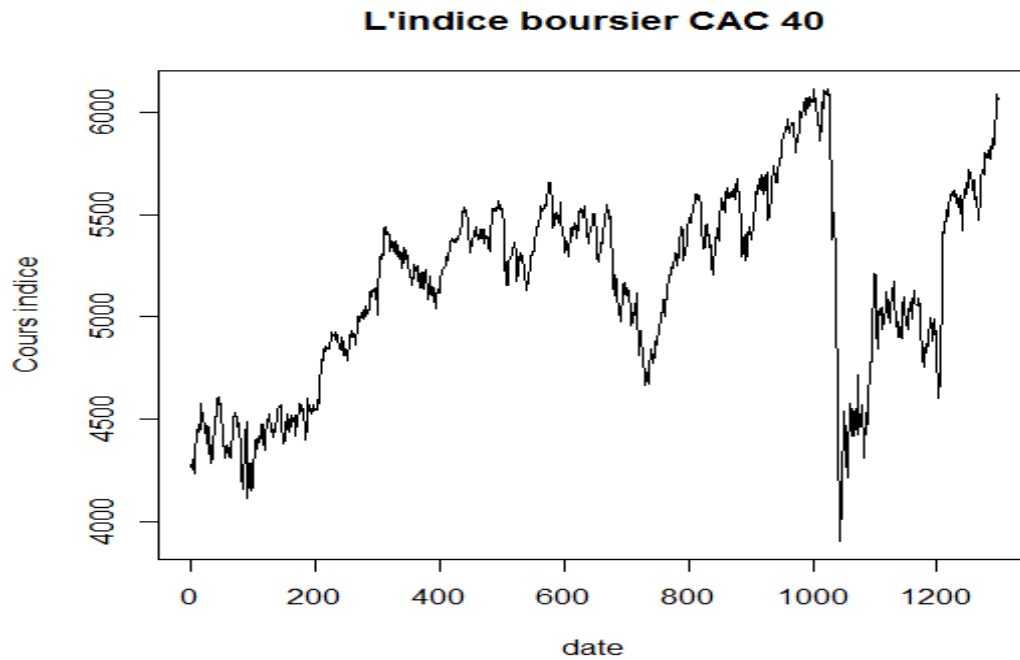


FIGURE 3.4 – L'indice boursier CAC 40.

Soit  $C_t$  la série des rendement logarithmique de l'indice boursier CAC40.

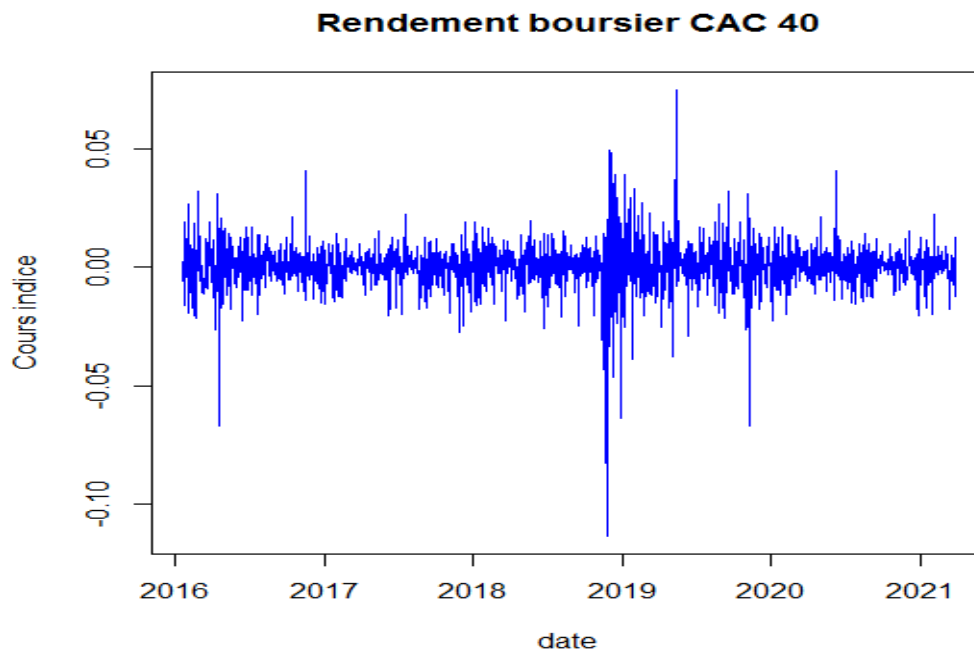


FIGURE 3.5 – Rendement logarithmique de l'indice boursier CAC 40.

A partir des figures (3.4 et 3.5) de la série CAC 40 nous observons un comportement de la variabilité en cluster, c'est-à-dire, les variations de grande amplitude sont suivies par des variations de grande amplitude et les variations de petite amplitude sont suivies par des variations de petite amplitude. Un tel comportement (qualifié d'hétéroscédastique) ne peut certainement pas être expliqué avec les modèles linéaires à variance constante (ou homoscedastique). L'agrégation de ces périodes de variabilité élevée fait partie des caractéristiques des séries financières appelées stylized facts. Un autre phénomène de ce type concerne la distribution probabiliste de ces données est l'apparition de chocs (c'est-à-dire des observations isolées extrêmement élevées en valeur absolue) n'est pas compatible avec une loi normale, mais nécessite plutôt une modélisation avec une distribution plus étalée (heavy tails), d'où la modélisation de la série CAC 40 par le modèle GARCH à erreurs de loi de Student.

Pour modéliser cet indice boursier, nous allons considérer le modèle autoregressif conditionnellement hétéroscédastique généralisé défini par l'équation (1.7) du chapitre 1 avec erreur de loi de Student. Nous allons modéliser la moyenne et la variance conditionnelle de CAC40, conditionnellement à l'information connue ( $F_{t-1}$ ). La modélisation GARCH de  $(C_t)_t$  revient donc à écrire :

$$C_t = \mu + X_t, \quad X_t = \sigma_t \varepsilon_t$$

où  $\mu$  et  $\sigma_t$  sont respectivement la moyenne et la variance conditionnelle.

L'estimation des paramètres des modèles hétéroscédastiques, est généralement fondée sur la méthode du maximum vraisemblance. Le modèle retenu est donc un GARCH(1, 1) de paramètres :

$$\hat{\omega} = 4.240205e - 06, \quad \hat{\alpha}_1 = 1.619948e - 01, \quad \hat{\beta}_1 = 7.973196e - 01$$

et de moyenne conditionnelle  $\hat{\mu} = 6.176000e - 04$

Nous remarquons que la somme  $\hat{\alpha}_1 + \hat{\beta}_1 = (1.619948e - 01) + (7.973196e - 01) = 0.9593144 < 1$ , elle vérifie donc la condition de la stationnarité.

D'où, finalement, la série des rendements  $(C_t)_t$  admet la représentation suivante :

$$C_t = 6.176000e - 04 + \sigma_t \eta_t,$$

la variance conditionnelle  $\sigma_t$  est donnée par :

$$\sigma_t^2 = (6.176000e - 04) + (1.619948e - 01)X_{t-1}^2 + (7.973196e - 01)\sigma_{t-1}^2$$

Pour confirmer que la série des prix du cours *CAC40* est de loi de Student, nous avons effectué le test de Kolmogorov smirnov sur les résidus

Les hypothèses :

$H_0$  : La distribution des résidus suit une loi de Student.

$H_1$  : La distribution des résidus ne suit pas une loi de Student.

A partir du test de Kolmogorov Smirnov nous avons obtenue :

la statistique du test  $D_n = 0.018788 < d_{1-\alpha}(n) = d_{1-0.05}(1300) = 0.0374$  et la *p - valus* = 0.5161 > 0.05, alors nous acceptons l'hypothèse  $H_0$ , donc La distribution des résidus suit une loi de Student.

Nous avons trouvé que les résidus suivent une loi de Student a 3 degré de liberté

Donc la modélisation de *CAC40* nous a donnée un modèle GARCH(1,1) à erreurs de distribution de Student a 3 degré de Liberté

# Bibliographie

- [1] Bardet, J.M. and Wintenberger, O. (2009). Asymptotic normality of the Quasi-Maximum likelihood estimator for multidimensional causal process. *The Annals of Statistics*, **37**, 2730–2759.
- [2] Bardet, JM. Boularouk, Y. and Djaballah,K. (2017). Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes. *Electronic Journal of Statistics Vol. 11*, 452–479.
- [3] Bardet.J.M, Kamila.K et Kengne.W (2019) Consistent model selection criteria and goodness-of-fittest for affine causal processes, *Electronic Journal of Statistics* ISSN : 1935-7524
- [4] Berkes, I. and Horv´ath, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *The Annals of Statistics*, **32**, 633–655.
- [5] Berkes, I., Horv´ath, L. and Kokoszka, P. (2003). GARCH processes : Structure and estimation. *Bernoulli*, **9** 201–227.
- [6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 309–328.
- [7] Bollerslev, T. and Wooldridge, J.M. (1992). Quasi - Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews*, **11**, 143-172.
- [8] Baillie, R. and Bollerslev, T. (1989). Common stochastic trends in a system of exchange rates. *Journal of Monetary Economics*, **44**, 167–81.
- [9] Berkes, I., Horv´ath, L. and Kokoszka, P. (2003), GARCH processes : structure and estimation, *Bernoulli*, 9, 201-227.

- [10] Bougerol, P. and Picard, N. (1992). Stationarity of GA RCH processes and of some nonnegative time series. *Journal of Econometrics* **52**, 115-128.
- [11] Boularouk, Y. and Djaballah, K. (2015). New approximation for ARMA parameters estimate. *Mathematics and Computers in Simulation, 2015, vol. 118, issue C*, 116-122
- [12] Boularouk, Y. and Bardet, J.M. (2020). Generalized Gaussian quasi-maximum likelihood estimation for most common time series. .hal-02902614
- [13] Boussama, F. (1998). Ergodicité, mélange et estimation dans les modèles GARCH. Ph.D. dissertation, Univ. Paris VII
- [14] Boussama, F. (2000). Normalité asymptotique de l'estimateur du pseudo-maximum de vraisemblance d'un modèle GARCH. C. R. Acad. Sci. Paris Ser. I Math. 33181-84. MR1780222
- [15] Box, G.E.P. Jenkins, G.M. (1970), Time Series Analysis; Forecasting and Control, San Francisco : Holden-Day.
- [16] Brandt, A. (1986). The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients. *Advances in Applied Probability*, **18**, 211-220.
- [17] Bougerol, P. and Picard, N. (1992). Stationarity of GA RCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115-128.
- [18] Djaballah, K. and Kerar, I. (2019). Quasi-maximum likelihood estimation of GARCH with student distributed noise. *Communications in Statistics - Simulation and Computation*
- [19] Doukhan, P. and Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications*, **118**, 1997-2013.
- [20] Drost, F.C. and Nijman, T.E. (1993) Temporal aggregation of GARCH processes. *Econometrica* 61, 909-927.
- [21] Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.
- [22] Fan, J., Qi, L. and Xiu, D. (2014). Quasi maximum likelihood estimation of GARCH models with heavy-tailed likelihoods. *Journal of Business and Economic Statistics*, **32**, 178-191.

- [23] FELLER, W. (1966). An Introduction to Probability Theory and Its Applications 2. Wiley, NewYork. MR0210154
- [24] Feller,W (1971). An Introduction to Probability,Theory and Its Applications (livre).
- [25] Francq, C. and Zakoian, J.M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, **10**, 605–63.
- [26] Francq.C et Zakoian.J.M.(2009).Modèles GARCH et à volatilité stochastique.http ://perso.univ-lille3.fr/~cfrancq/Christian-Francq/cours-GARCH-ENSAE/poly09.pdf.
- [27] Francq, C. and Zakoïan, J.M. (2010). GARCH models : Structure, statistical inference and financial applications. *John Wiley*.
- [28] Francq, C. and Zakoïan, J.M. (2012). QML estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory*, **28**, 179-206.
- [29] Francq, C. and Zakoïan, J.M. (2013). Optimal predictions of powers of conditionally heteroskedastic processes. *Journal of the Royal Statistical Society*, **B75**, 345-367.
- [30] Francq, C. and Zakoïan, J.M. and Lepage, G. (2011). Two-stage non Gaussian QML estimation of GARCH Models and testing the efficiency of the Gaussian QMLE. *Journal of Econometrics*, **165**, 246-257.
- [31] Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*, **71**, 285-317.
- [32] Huang, D., Wang, H. and Yao, Q. (2008). Estimating GARCH models : when to use what ? *Econometrics J.* **11**, 27-38.
- [33] Jeanteau.T, (1993), Mod'eles autorégressifs à erreurs conditionnellement hétéroscédastiques. Thèse de l'Université Paris VII.
- [34] Jeanteau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory* **14**, 70–86.
- [35] Kluppelberg, C., Lindner, A. and Maller, R. (2004) A continuous time GARCH process driven by a Lévy process : stationarity and second order behaviour. *Journal of Applied Probability* 41, 601–622.

- [36] Lee, S.W. and Hansen, B. (1994). Asymptotic Theory for the Garch(1,1) Quasi-Maximum Likelihood Estimator. *Econometric Theory*, vol. **10**, issue 01, 29-52.
- [37] Lepage.G,(2012),Inférence statistique des modèles conditionnellement hétéroscédastiques avec innovations stables,contraste non gaussien et volatilité mal spécifiée.Thèse de doctorat.
- [38] Lumsdaine.R.L, (1993), Asymptotic properties of the quasi-maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models, Unpublished manuscript, Princeton University, Dpt of Economics.
- [39] Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1, 1) and covariance stationary GARCH(1, 1) models. *Econometrica* **64**, 575-596.
- [40] Mikosch, T. and Straumann, D. (2006). Stable limits of martingale transforms with application to the estimation of GARCH parameters. *Ann. Statist.* 34 493-522.
- [41] Nelson, D. (1990). Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory* **6**, 318-334.
- [42] Robinson, P.M. and Zaffaroni, P. (2006), Pseudo-maximum likelihood estimation of ARCH(1) models, *Ann. Statist.* 34, 1049-1074.
- [43] Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time Series Models in Econometrics, Finance and Other Fields* (D. R. Cox, D. Hinkley and O. E. Barndorff-Nielsen, eds.) 1-67. Chapman and Hall, London. MR1398214
- [44] Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series : A stochastic recurrence equations approach. *The Annals of Statistics*, **34**, 2449-2495.
- [45] Straumann, D. (2005). Estimation in Conditionally Heteroscedastic Time Series Models. *Lecture Notes in Statistics* **181**. Springer, Berlin.
- [46] Weiss, A.A. (1986), Asymptotic theory for ARCH models estimation and testing. *Econometric Theory*, 2, 107-131.
- [47] Wintenberger.O. (2012) Contributions à la statistique des processus : estimation, prédiction et extrêmes. Université Paris Dauphine - Paris IX, 2012. tel-00757756