

REPUBLIQUE DE CÔTE D'IVOIRE
Union - Discipline – Travail

**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE**



Institut National Polytechnique
Félix HOUPHOUËT-BOIGNY de Yamoussoukro



Ecole Doctorale Polytechnique
N° d'ordre : **126 / 2022**

THESE

Pour l'obtention du grade de

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE FELIX HOUPHOUËT-BOIGNY

Mention : Mathématiques Appliquées

Spécialité : Statistique

Sujet :

**Etude des modèles GARCH hyperboliques univariés et
multivariés stationnaires et leurs applications**

Présentée et soutenue publiquement le 07 Octobre 2022 par

BAMBA Lanciné

JURY

M. OUATTARA Sié	Professeur Titulaire, INP-HB	Président
M. DIOP Aliou	Professeur Titulaire, UGB	Rapporteur
M. YODE ARMEL	Maître de Conférences, UFHB	Rapporteur
M. HILI Ouagnina	Professeur Titulaire, INP-HB	Directeur de thèse
M. N'GUESSAN Assi	Maître de Conférences Habilité, Univ. de Lille	Co-Directeur de thèse
M. MONSAN Vincent	Maître de Conférences, UFHB	Examineur

Etude des modèles GARCH hyperboliques univariées et multivariés stationnaires et leurs applications

Thèse

Lanciné BAMBA

Sous la direction de:

Ouagnina HILI, codirecteur de recherche
Assi N'GUESSAN, codirecteur de recherche

*Je dédie ce présent mémoire à
mon père et à ma mère*

Ne crains pas d'avancer
lentement, crains seulement de
t'arrêter.

Confucius

Table des matières

Résumé	vi
Abstract	vii
Liste des tableaux	viii
Liste des figures	ix
Remerciements	x
Liste des Abréviations et Notations	xiii
Introduction	1
1 Généralités	7
1.1 Séries temporelles	7
1.2 Espérance conditionnelle	11
1.3 Modèles à hétéroscédasticité conditionnelle	12
1.4 Techniques d'approximation de la variance conditionnelle par initialisation	14
1.5 Méthodes de détection des outliers	15
1.6 M-estimateurs et Propriétés Générales	16
1.7 Résultats Fondamentaux	17
2 M-estimateurs des modèles GARCH hyperboliques univariés stationnaires	24
2.1 Processus GARCH hyperbolique univarié	25
2.2 Estimation des paramètres	27
3 Etude comparative des performances des estimateurs du modèle GARCH hyperbolique et applications aux données réelles	58
3.1 Simulations numériques	59
3.2 Applications aux données du S&P 500	79

4 Propriétés asymptotiques de l'estimateur du Quasi-maximum de vraisemblance d'une classe de modèle GARCH hyperbolique multivarié	89
4.1 Présentation du modèle HGARCH multivarié	90
4.2 Condition de Stationnarité pour les modèles CCC-HGARCH	93
4.3 Estimation pour les modèles CCC-HGARCH	96
4.4 Etude de Simulation	116
Conclusion	121
Liste des contributions et activités de recherches	123
Annexe	125
A Boîtes à moustaches des paramètres du modèle HGARCH	125
B Preuves Complémentaires	132
C Article Publié	143
Références bibliographiques	193

Résumé

Dans cette thèse, nous traitons la question de la modélisation des séries temporelles à longue mémoire dans la volatilité. Nous nous intéressons particulièrement aux modèles GARCH hyperboliques univariés et multivariés stationnaires. Pour l'étude du modèle univarié, nous définissons d'abord deux classes d'estimateurs basées sur une fonction bornée, le M-estimateur et le BM-estimateur, afin d'estimer les paramètres du modèle. Ensuite, les propriétés asymptotiques de ces deux classes d'estimateurs à savoir la consistance et la normalité asymptotique sont établies. Puis, à travers des simulations numériques réalisées, nous étudions d'une part le comportement des deux estimateurs proposés et celui de l'estimateur du quasi-maximum de vraisemblance en présence de valeurs aberrantes dans les observations. D'autre part, nous comparons les performances d'estimation et de prévision du M-estimateur et du BM-estimateur à celle de l'estimateur du quasi-maximum de vraisemblance jusqu'alors utilisé pour ces modèles. Enfin, une étude sur des données réelles impliquant les trois estimateurs en compétition est également réalisée. Pour le cas multivarié, nous proposons un modèle GARCH hyperbolique à corrélation conditionnelle constante comme extension du modèle GARCH hyperbolique univarié. Nous donnons également une condition suffisante pour l'existence d'une solution strictement stationnaire de ce modèle multivarié. Par ailleurs, nous établissons, sous certaines conditions, les propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance défini pour ce modèle. De plus, grâce à des simulations numériques, nous illustrons les propriétés de l'estimateur étudié dans le cas d'échantillon de taille finie.

Mots clés : modèle ARCH(∞) - Longue mémoire - volatilité - M-estimateur - BM-estimateur - modèle HGARCH - consistance forte - normalité asymptotique.

Abstract

In this thesis, we address the issue of modeling long memory time series in volatility. We are particularly interested in univariate and multivariate stationary hyperbolic GARCH models. For the study of the univariate model, we first define two classes of estimators based on a bounded function, the M-estimator and the BM-estimator, in order to estimate the model parameters. Then, the asymptotic properties of these two classes of estimators, namely the consistency and the asymptotic normality, are established. Then, through numerical simulations, we study on the one hand the behavior of the two proposed estimators and that of the quasi-maximum likelihood estimator in the presence of outliers in the observations. On the other hand, we compare the estimation and prediction performances of the M-estimator and the BM-estimator with those of the quasi-maximum likelihood estimator used so far for these models. Finally, a study on real data involving the three competing estimators is also performed. For the multivariate case, we propose a hyperbolic GARCH model with constant conditional correlation as an extension of the univariate hyperbolic GARCH model. We also give a sufficient condition for the existence of a strictly stationary solution of this multivariate model. Moreover, we establish, under certain conditions, the asymptotic properties of the quasi-maximum likelihood estimator defined for this model. Moreover, thanks to numerical simulations, we illustrate the properties of the studied estimator in the case of finite sample size.

Keywords: ARCH(∞) model - Long memory - volatility - M-estimator - BM-estimator - HGARCH model - strong consistency - asymptotic normality.

Liste des tableaux

3.1	Résultats pour le modèle HGARCH(1,d,1) normal sans outliers	64
3.2	Résultats du modèle HGARCH(1,d,1) normal avec 5% d'outliers additives et $e = 3$	67
3.3	Résultats du modèle HGARCH(1,d,1) normal avec 5% d'outliers additives et $e = 5$	69
3.4	Résultats du modèle HGARCH(1,d,1) normal avec 10% d'outliers additives et $e = 3$	71
3.5	Résultats du modèle HGARCH(1,d,1) normal avec 10% d'outliers additives et $e = 5$	72
3.6	Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) sans outliers	74
3.7	Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 5% d'outliers additives et $e = 3$	75
3.8	Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 5% d'outliers additives et $e = 5$	76
3.9	Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 10% d'outliers additives et $e = 3$	77
3.10	Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 10% d'outliers additives et $e = 5$	78
3.11	Statistique descriptive des log-rendement du S&P 500	80
3.12	V/S test pour la détection de la longue mémoire	84
3.13	Valeurs des BIC des log-rendement du S&P 500	85
3.14	Paramètres estimés des modèles FIGARCH(1,d,1), HYGARCH(1,d,1) et HGARCH(1,d,1) pour la série des log-rendements journaliers du S&P 500	87
3.15	Résultats de prévision des modèles FIGARCH(1,d,1), HYGARCH(1,d,1) et HGARCH(1,d,1) pour les log-rendements du S&P 500	88
4.1	Résultats d'estimation du modèle CCC-HGARCH(1,d,1) bivarié issu d'une loi de Student à 13 degrés de liberté	118

Liste des figures

3.1	Série brute de l'indice journalier du S&P 500	79
3.2	Série des log-rendements du S&P 500	81
3.3	Boîte à moustache des log-rendements du S&P 500	81
3.4	ACF des log-rendements du S&P 500	83
3.5	ACF du carré des log-rendements $\{r_t^2\}$ du S&P 500	83
4.1	Q-Q plots pour les paramètres estimés du modèle CCC-HGARCH(1,d,1) bivarié pour n=8000	119
4.2	Q-Q plots pour les paramètres estimés du modèle CCC-HGARCH(1,d,1) bivarié pour n=8000	120
A.1	Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 5% of outliers et e=3 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	126
A.2	Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 5% of outliers et e=5 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	127
A.3	Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 10% of outliers et e=5 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	128
A.4	Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 5% of outliers et e=3 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	129
A.5	Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 5% of outliers et e=5 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	130
A.6	Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 10% of outliers et e=5 pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.	131

Remerciements

Ce manuscrit est le résultat de plusieurs années de travail au cours desquelles plusieurs personnes nous ont accompagnés. Je voudrais dans cette partie leur rendre hommage.

Je tiens particulièrement à témoigner toute ma gratitude au Professeur HILI Ouagnina, Directeur de cette thèse et Directeur du l'Unité Mixte de Recherche et d'Innovation Mathématiques et Nouvelles de l'Information, qui a orienté et dirigé nos travaux de recherches vers ce sujet fort intéressant. Je voudrais lui dire un grand merci pour sa patience, ses encouragements, son soutien moral et financier et surtout sa rigueur dans le travail. Il a su nous communiquer son dynamisme et sa passion pour la recherche en suscitant pour nous des collaborations avec des chercheurs dans le monde.

Ma gratitude va ensuite au Professeur N'GUESSAN Assi, co-Directeur de cette thèse pour ses précieux conseils et son aide pour la rédaction du mémoire.

Je voudrais aussi remercier le Professeur OUATTARA Sié, pour avoir accepté de présider le jury de cette thèse.

Mes vifs remerciements sont adressés au Professeur Aliou DIOP, rapporteur externe de cette thèse, pour ses remarques et suggestions très intéressantes qui ont contribué à enrichir le contenu de ce manuscrit.

Je souhaiterais également dire merci au Professeur YODE Armel, rapporteur interne de cette thèse, pour ses conseils avisés qui ont contribué à l'amélioration de la qualité de ce manuscrit.

Je voudrais par suite remercier le Professeur MONSAN Vincent, examinateur de cette thèse, pour sa rigueur et aussi pour la qualité de ses interventions qui ont permis de mieux valoriser les résultats de cette thèse.

Je souhaiterais adresser mes remerciements et ma reconnaissance au Directeur général de l'Institut National Polytechnique Félix Houphouët Boigny (INP-HB) de Yamoussou-

kro, Dr. DIABY Abdoul Kader Moussa pour avoir accepté mon inscription dans son établissement.

Je voudrais également dire merci au Directeur de l'Ecole Doctorale Polytechnique (EDP) de l'INP-HB, Professeur YAO Kouassi Benjamin qui organise son administration afin de contribuer au bien-être de ses doctorants. Je souhaiterais avoir une pensée particulière pour le Professeur SORO Doudjo, Directeur des études à l'EDP, pour son dynamisme et sa disponibilité ainsi qu'à toute son administration.

Je ne saurais oublier les membres de l'équipe Statistique et Probabilité de l'UMRI-MNTI, à savoir Dr. KAMAGATE, Dr N'DRI, Dr SORO, que je remercie, pour leurs encouragements et leurs accompagnements durant tout ce temps. A mes condisciples du laboratoire KANGA Serge, KOUAKOU Geoffroy, KEITA Filamory, BIAN Hubert, KOUASSI Ben Celestin, AMANI Matthias, KADJO Roger, je dis infiniment merci à tous pour tout ce qu'on a partagé ensemble.

Je tiens aussi à remercier les amis et frères de longue date, membres de l'équipe d'Analyse Numérique de l'UMRI-MNTI que sont Dr. N'DRI Cyrille, Dr. ADOU Koffi Achille, Dr. GANON Ardjouma et Dr. EDJA Kouamé Béranger.

Je voudrais également remercier tous les membres du département Mathématiques-Informatique de l'INPHB particulièrement le Directeur Mr SAFFIEDINE et le Professeur TOURE Augustin.

Je voudrais en outre me souvenir de mes maîtres, ceux qui dès le début m'ont communiqué l'amour des Mathématiques. Il s'agit des Professeurs de mon université d'origine, l'Université Nangui Abrogoua. Je fais ainsi référence au Professeur SANGARE Daouda, Professeur GOZO Yoro, Professeur DIAGANA Youssouf, Professeur ASSANE ainsi qu'au Dr. KPATTA Akon et au Dr. ESSOH Modeste. Par suite, je voudrais remercier les condisciples de notre MASTER Mathématiques de l'année 2015-2016, il s'agit entre autres du Dr. DJE Tanoh Jean-Marcel, Dr. NAGACY Pokou, Dr BECHE Eugène Deval et YEKRE Benjamin.

Je tiens à remercier également le Professeur Abdou Kâ DIONGUE, Directeur du Laboratoire d'Etudes et Recherches en Statistiques et Développement (LERSTAD) à l'Université Gaston Berger de Saint-Louis au Sénégal où une partie de ces travaux a été réalisée. Je voudrais à cet effet remercier tout les membres du LERSTAD pour leur accueil chaleureux notamment le Dr. Lamine DIOP, Professeur DEM, Dr. Ali ESSOHAM, Dr MOR Ndongo et aux doctorants SALL Mahamadou, Sema GADO ainsi que

mon regretté ami BADIANE.

Je voudrais aussi dire merci aux condisciples, participant du CIMPA de Lomé 2018 pour la qualité de nos échanges et de nos collaborations. Il s'agit entre autres de Dr. KATCHEPELE, Dr RUTIKANGA, Dr Charline, Tchazimo.

J'ai encore des souvenirs de mes débuts d'années de thèse et de ces personnes qui m'ont apporté leur soutien financier, matériel, psychologique et autres. Je fais allusion à Feu Mr DEDE Dieudonné, ex Directeur de WISDOM EDUCATION, Feue Amandine, ex secrétaire de WISDOM EDUCATION ainsi qu'à mon père spirituel Feu le Pasteur ANOH Anvo Marcel. Je voudrais également rendre un vibrant hommage à mes parents, feu BAMBA Logloua, feu BAMBA Awa et feu HABA KOLKOL, partis sans avoir vu l'aboutissement de cette thèse.

Ma reconnaissance va également à l'endroit de mes communautés, la Mission Pentecoste le Saint-Esprit de Côte d'Ivoire (MPSECI) dirigée par l'Ancien GONDO Emmanuel et le Prophète YAO Michel avec ma promotion les boeufs de labour, ainsi que la MIRAS Mission internationale du Réveil des Ames la Source avec son leader l'Apotro KONE Israel. Je voudrais dire encore un grand merci à mes amis et frères KOFFI Constant, MONEY Stephane, OKOMA Anselme, N'GORAN Konan Prince.

A mes frères qui m'ont reçu et hébergé dans mes temps difficiles à Yamoussoukro, La famille de KONAN Davide, ARDJOUMANE Marcellin et AKA Elysée, merci infiniment pour votre hospitalité et votre bienveillance.

Je remercie également Miss BONY Axelle Meryl et Miss AMOUSSOU Mononde Jocelyne pour leur apport à la correction du présent manuscrit.

Je voudrais également dire merci à DO Ella Amandine pour sa disponibilité et sa générosité car elle nous fait régulièrement à manger.

Liste des Abréviations et Notations

i.i.d. : indépendantes et identiquement distribuées.

$\overset{\circ}{\Theta}$: l'intérieur de Θ .

$\mathbf{E}(X)$: espérance mathématique de X .

$\mathbf{P}(X)$: probabilité de X .

$\text{Tr}(A)$: trace de la matrice A .

$\mathcal{M}_{n \times n}(\mathbf{R})$: l'ensemble des matrices carrées d'ordre n à coefficients réels.

I_N : la matrice identité d'ordre N .

$\mathbf{0}$: la matrice nulle d'ordre N .

$\underline{0}$: le vecteur nul d'ordre N .

$\text{Sp}(A)$: Spectre de A , l'ensemble des valeurs propres de la matrice A .

A' : transposée de la matrice A .

$|A|$: déterminant de la matrice A .

p.s : presque sûrement.

$\text{Cov}_X(X_{t+h}, X_t)$: covariance de X_{t+h} et X_t .

$\stackrel{\text{loi}}{=} :$ égalité en loi.

$C(\Theta, \mathbf{R}^{n \times n})$: ensemble des fonctions continues sur un compact Θ de \mathbf{R}^n à valeurs dans l'espace des matrices carrées d'ordre n à valeurs réelles.

Introduction

L'étude de plusieurs séries temporelles financières telles que les taux d'intérêts futurs, les taux de change et le prix des actifs a révélé que la fonction d'autocovariance du carré des rendements de ces séries décroît très lentement. Comme indiqué dans Ding et al. [23] et Ding et Granger [22], ce phénomène de décroissance lente de la fonction d'autocovariance observé dans ces séries traduit la présence d'une longue mémoire dans la volatilité de celles-ci. Suivant l'approche de Giraitis et al. [33], on dit qu'un processus présente une longue mémoire si sa fonction d'autocovariance n'est pas absolument sommable. Dans le cas contraire, on parle de processus à courte mémoire. Les modèles autorégressifs conditionnellement hétéroscédastiques du type ARCH(p) et GARCH(p, q) respectivement étudiés par Engle [28] et Bollerslev [9] furent les premiers dans la littérature à traiter les problèmes de volatilité dans les séries temporelles financières. Cependant, les modèles ARCH(p) et GARCH(p, q) présentent une courte mémoire. Ils ne sont donc pas adaptés pour la modélisation et la prévision des séries qui présentent la caractéristique de longue mémoire.

Afin d'étudier au mieux le phénomène de longue mémoire dans la volatilité des séries temporelles, Baillie et al. [3] ont proposé les processus FIGARCH (Fractionally Integrated Generalized AutoRegressive Conditional Heteroskedastic). Les travaux menés par Giraitis et al. [34] ont prouvé la présence de longue mémoire dans les modèles FIGARCH. Ce qui les qualifie donc pour la modélisation des séries qui possèdent la caractéristique de longue mémoire. C'est ainsi que Conrad et Haag [17] ont fait cas des contraintes à imposer aux paramètres de la variance conditionnelle du FIGARCH pour que celle-ci soit positive. Cependant, Baillie et al. [3] ont révélé que le modèle FIGARCH possède une variance infinie. Ce qui limite ses applications. Le défaut du modèle FIGARCH est résolu par Davidson [19] qui introduit dans la littérature le modèle GARCH à mémoire hyperbolique appelé modèle GARCH hyperbolique et noté HYGARCH. Le concept de mémoire hyperbolique fut introduit par Davidson [19] pour caractériser les

processus dont les coefficients ont une décroissance hyperbolique. Pour de tels processus, la fonction d'autocovariance décroît de manière hyperbolique c'est-à-dire une décroissance très lente. En d'autres termes, les concepts de mémoire hyperbolique et de longue mémoire traduisent les mêmes faits. Dans ce contexte, les modèles FIGARCH sont désignés dans la littérature (voir Li et al. [46]), comme les tous premiers modèles GARCH à mémoire hyperbolique. Le modèle HYGARCH proposé par Davidson [19] présente donc les propriétés de longue mémoire et de variance finie. Ces propriétés intéressantes du modèle HYGARCH ont motivé Niguez et Rubia [56] à appliquer le HYGARCH à la modélisation du taux de changes entre le Dollar américain et un portefeuille de cinq monnaies internationales à savoir le Deutschmark allemand, le Yen japonais, la Livre sterling britannique, le Franc suisse et la Couronne suédoise. Niguez et Rubia [56] ont mis en exergue le fait que le modèle HYGARCH a de meilleures performances de prévision hors échantillon que le modèle GARCH standard. Leur analyse révèle aussi que la modélisation correcte du paramètre de longue mémoire dans la volatilité est d'une grande importance même si l'on ne s'intéresse qu'à la prévision à court terme de la volatilité. De même, à travers la propriété de longue mémoire, Tang et Shieh [62] comparent les performances des modèles HYGARCH et FIGARCH pour la prédiction des valeurs à risque des prix à la clôture de trois marchés d'indices boursiers à savoir le S&P500, le Nasdaq100 et le Dow Jones. Ils ont également montré que le modèle HYGARCH est plus performant que le FIGARCH. Cependant, certaines critiques ont été émises à l'encontre le modèle HYGARCH. Selon Conrad [18], les conditions qui garantissent la positivité de la variance conditionnelle du modèle HYGARCH sont beaucoup plus restrictives que celles des modèles FIGARCH. De plus, selon Li et al. [46], la variance conditionnelle du modèle HYGARCH a une forme assez complexe. Ces raisons ont motivé Li et al. [46] à proposer un nouveau modèle GARCH hyperbolique (HGARCH) qui a une variance finie, une longue mémoire et une structure plus simple que le HYGARCH.

Pour estimer les paramètres du modèle HGARCH, Li et al. [46], dans leurs travaux ont eu recours à la méthode du quasi-maximum de vraisemblance (QMV). De nombreux auteurs comme Robinson et Zaffaroni [58] et Zakoïan et Francq ([15], P.141-148) ont montré que sous certaines conditions, l'estimateur du quasi-maximum de vraisemblance respectivement pour les modèles ARCH(∞) et GARCH(p, q) est consistant et asymptotiquement normal. Cependant, lorsque des valeurs aberrantes apparaissent dans les données, les méthodes d'estimation telles que le QMV en sont très affectées. Les valeurs aberrantes que nous appellerons outliers dans la suite du travail sont définies par Lu

et al. [49] comme des observations qui s'écartent considérablement de la majorité des données. Pour traiter des données comportant des outliers, deux grandes approches sont présentées dans la littérature, à savoir l'utilisation de méthodes d'estimation robustes sur l'ensemble des observations, ou la détection et la suppression des outliers. Par ailleurs, Fox [30] présente deux principaux types d'outliers pour les séries temporelles : les outliers additives et les outliers d'innovation.

De nombreux auteurs ont proposé des méthodes d'estimation robustes pour traiter les outliers. Ce fut le cas de Huber [38] qui introduit une classe M-estimateur pour l'estimation d'un paramètre de localisation pour les distributions normales contaminées. Huber [39] propose également des estimateurs robustes de régression de type maximum de vraisemblance prenant en compte le fait que le nombre de paramètres du modèle peut augmenter avec la taille de l'échantillon. Plus récemment, Muler et Yohai [54] ont proposé deux classes d'estimateurs robustes pour les modèles ARCH(p) appelées le τ -estimateur et le τ -estimateur filtré. Muler et Yohai [55] ont également généralisé leurs recherches en proposant deux classes d'estimateurs robustes pour les modèles GARCH. La première classe est celle des M-estimateurs qui est obtenue en maximisant une fonction de perte bornée. La seconde classe d'estimateurs robustes de Muler et Yohai [55] appelée M-estimateur bornés et notée BM-estimateur est obtenue par modification de la première classe. Lu et al. [49] ont étendu les travaux de Muler et Yohai [55] aux modèles à erreurs multiplicatives (MEM) avec erreur GARCH et ont proposé deux fonctions de perte pour générer les M-estimateurs et BM-estimateurs présentés dans leurs travaux. De plus, après une étude comparative, Muler et Yohai [55] ainsi que Lu et al. [49] ont montré que les M-estimateurs et BM-estimateurs qu'ils ont proposés sont moins sensibles aux outliers que l'estimateur du QMV pour les modèles GARCH et MEM respectivement. Par ailleurs, lorsque les données contiennent des outliers, les M-estimateurs et BM-estimateurs présentent de meilleures performances de prévision que l'estimateur du QMV pour les modèles GARCH et MEM étudiés par Muler et Yohai [55] et Lu et al. [49]. Les auteurs ont expliqué ces résultats par le fait que la fonction logarithmique utilisée dans l'estimateur QMV n'est pas bornée. Après analyse des résultats des travaux de Muler et Yohai [55] et Lu et al. [49] pour les modèles GARCH et MEM, nous avons décidé pour nos travaux de recherches, d'étendre leur étude aux modèles GARCH hyperboliques.

Dans le cas où l'étude porte sur une série, on parle de cas univarié et les méthodes d'esti-

mation des paramètres des modèles univariés sont abondantes. Cependant, lorsque nous avons deux ou plusieurs séries à observer simultanément, on parle de cas multivarié. Dans ce cas, la typologie des observations n'est plus la même. Pour le cas univarié, les observations sont des scalaires tandis que dans le cas multivarié, les observations sont des vecteurs. Il va s'en dire que les techniques utilisées dans l'étude des séries univariées ne tiendront plus pour le cas multivarié.

Plusieurs auteurs tels que Engle et Kroner [29], Bollerslev [10] ont proposé des processus multivariés pour modéliser la volatilité ou les co-mouvements de plusieurs séries financières. Une variété de spécifications de modèles multivariés a été proposée. Ces différentes spécifications prennent en compte l'augmentation du nombre de paramètres, la condition de stationnarité et la facilité d'exécution numérique des modèles. Comme exemple de spécification des modèles GARCH multivariés, nous pouvons citer le modèle VEC-GARCH proposé par Bollerslev [11] et le modèle BEKK (Baba, Engle, Kraft, Kroner)-GARCH de Engle et Kroner [29]. Les spécifications VEC et BEKK représentent des extensions directes du GARCH univarié au cas multivarié. Silvennoinen et al. [59] ont critiqué ces spécifications parce que le nombre de paramètres à estimer est très élevé et les conditions assurant le fait que la matrice de covariance conditionnelle soit définie positive sont plutôt très restrictives. De plus, l'estimation numérique des paramètres de ces modèles reste assez difficile et très demandée. Une autre critique sur les représentations VEC et BEKK GARCH porte sur l'existence d'une solution strictement stationnaire qui est difficile à vérifier. Bollerslev [10] a proposé un modèle GARCH à corrélation conditionnelle constante (CCC-GARCH) pour lequel la matrice de corrélation est invariante dans le temps. De plus, le nombre de paramètres de la représentation CCC-GARCH est raisonnable car il peut être contrôlé. Les conditions permettant la matrice de covariance conditionnelle d'être définie positive sont assez légères. Toutes ces raisons ont motivé Jeantheau [40] ainsi que He et Terasvirta [36] à généraliser le modèle CCC-GARCH afin de proposer le modèle CCC-GARCH étendu noté ECCC-GARCH. Pour les modèles ECCC-GARCH, les auteurs donnent des conditions permettant l'existence d'une solution stationnaire stricte. Engle [27] a introduit dans la littérature une généralisation intuitive de la représentation CCC-GARCH en proposant la représentation GARCH à corrélation conditionnelle dynamique (DCC-GARCH) avec une matrice de corrélation conditionnelle variable dans le temps. En ce qui concerne le nombre de paramètres du DCC-GARCH, Minovic [52] affirme que ces modèles paramétriques sont très parcimonieux. Le modèle DCC-GARCH est d'une grande utilité pour la mo-

délisation des ensembles de données de grandes dimensions. Cependant la procédure d'estimation des paramètres de ce modèle semble être une tâche considérable. De manière générale, la théorie asymptotique pour l'estimation des paramètres des modèles GARCH multivariés a été établie par de nombreux auteurs. Au nombre desquels, figure Jeantheau [40] qui établit la consistance forte de l'estimateur du quasi-maximum de vraisemblance (QMV) du modèle ECCC-GARCH sous des hypothèses portant sur l'existence de moment. Comte et Libermann [16] ont également étudié la consistance et la normalité asymptotique de l'estimateur QMV du modèle BEKK-GARCH. Les travaux de Jeantheau [40] et de Comte et Libermann [16] sont purement théoriques sans étude de simulations. Francq et Zakoïan [31] de leurs cotés ont abordé l'étude des propriétés asymptotiques de l'estimateur du QMV des modèles asymétriques CCC-GARCH sans aucune hypothèse sur l'existence du moment du processus observé. Une étude de simulation numérique et une application à des données réelles ont été également réalisées dans leur article. Par ailleurs, Bibi [7] a étendu la partie théorique des travaux de Francq et Zakoïan [31] aux modèles CCC-GARCH périodique multivariés. Tous les modèles multivariés cités précédemment sont des extensions des modèles GARCH standard à courte mémoire.

Les modèles GARCH à mémoire longue ont été également étendus au cadre multivarié. De nombreux chercheurs ont étudié ces types de modèles multivariés dans leurs travaux de recherche. Par exemple Teyssière [63] et Kirman et al. [41] ont proposé un modèle FIGARCH multivarié nommé le CCC-FIGARCH pour la modélisation des taux de changes à haute fréquence et des taux de change volatilité. Brunetti et Gilbert [13] ont également introduit le CCC-FIGARCH bivarié pour étudier la volatilité des données du taux de changes du pétrole brut sur le marché américain de New York appelé le NYMEX(New York Mercantile Exchange), et celle sur le marché britannique de Londres appelé l'IPE(International Petroleum Exchange). Pafka et Matyas [57] ont introduit un modèle FIGARCH multivarié diagonal pour étudier la volatilité des taux de changes. Demirel et al. [20] ont également introduit les modèles DCC-ARIMA FIGARCH qu'ils ont appliqué au taux de changes entre le dollar américain et un portefeuille de monnaie de six pays. Concernant la théorie asymptotique sur les modèles GARCH multivariés à longue mémoire, nous pouvons nous référer aux travaux de Bardet et Wintenberger [4]. En effet, les auteurs ont proposé de nombreux théorèmes sur les propriétés asymptotiques de l'estimateur du QMV pour les processus causaux multivariés. Dans le travail que nous vous proposons, nous étendons l'étude menée par Francq et Zakoïan

[31] aux modèles GARCH hyperbolique multivariés. Le modèle GARCH hyperbolique choisi pour notre étude est le HGARCH développé par Li et al. [46].

Notre objectif principal dans ce travail est non seulement de proposer et étudier des méthodes d'estimation robustes permettant de traiter au mieux l'apparition d'outliers dans les modèles GARCH hyperboliques univarié. Mais aussi de généraliser le modèle HGARCH univarié au cas multivarié en proposant un modèle HGARCH multivarié. L'étude de la stationnarité stricte du modèle HGARCH multivarié proposé et celle des propriétés asymptotiques de l'estimateur défini pour les paramètres de ce modèle sera également explorée dans ce travail. Pour ce faire, nous proposerons dans le cas du modèle univarié, deux estimateurs appelés M-estimateur et BM-estimateur pour les modèles GARCH hyperboliques. Le M-estimateur que nous présentons est construit en suivant l'approche de Muler et Yohai [55] tandis que le BM-estimateur de notre travail est motivé par Lu et al. [49]. Nos principaux axes de recherches tourneront dans un premier temps autour de l'étude théorique des propriétés asymptotiques (consistance et normalité asymptotique) des estimateurs que nous proposons pour les modèles univarié et multivarié. Dans un second temps, nous aborderons une étude numérique des propriétés des estimateurs définis dans ce travail pour des échantillons de taille finie dans le but d'illustrer les résultats théoriques. Une application aux données réelles est également conduite.

Cette thèse est articulée autour du plan suivant : le chapitre 1 donne les notions fondamentales et les définitions que nous utiliserons dans la suite du document. Le chapitre 2 contient les résultats théoriques de l'estimation des modèles GARCH hyperboliques univarié par la méthode des M-estimateurs. Le chapitre 3 est consacré à l'étude comparative des performances des deux classes de M-estimateurs définies au chapitre 2 et celles de l'estimateur du quasi-maximum de vraisemblance par le moyen de simulations numériques et d'application aux données réelles. Le chapitre 4 est dédié à l'étude du modèle GARCH hyperbolique multivarié que nous avons introduit dans ce travail. Une partie Annexe qui contient certaines figures issues des simulations réalisées au chapitre 3.

Chapitre 1

Généralités

Le but de ce chapitre est de regrouper les concepts fondamentaux nécessaires à la compréhension des travaux de cette thèse.

1.1 Séries temporelles

1.1.1 Processus stochastiques

Dans ce mémoire, $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité et $\mathcal{T} \subset \mathbb{R}$.

Définition 1.1.1. *Un processus stochastique est une famille de variables aléatoires réelles $(X_t)_{t \in \mathcal{T}}$ définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. Si $\mathcal{T} \subset \mathbb{Z}$, le processus $(X_t)_{t \in \mathcal{T}}$ est appelé série temporelle (ou série chronologique).*

1.1.2 Stationnarité et Ergodicité

Définition 1.1.2. *Soit $(X_t)_{t \in \mathcal{T}}$ un processus tel que la variance de X_t notée $V(X_t)$ soit finie pour tout $t \in \mathcal{T}$, c'est-à-dire $V(X_t) < \infty, \forall t \in \mathcal{T}$. On appelle fonction d'autocovariance de $(X_t)_{t \in \mathcal{T}}$ la fonction*

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = \mathbb{E}[(X_r - \mathbb{E}(X_r))(X_s - \mathbb{E}(X_s))], \quad r, s \in \mathcal{T}.$$

Définition 1.1.3. *La série temporelle $(X_t)_{t \in \mathbb{Z}}$ est dite stationnaire ou stationnaire au second ordre si*

1. $\mathbb{E}(|X_t|^2) < \infty, \forall t \in \mathbb{Z}$,
2. $\mathbb{E}(X_t) = m, \forall t \in \mathbb{Z}, m \in \mathbb{R}$,

$$3. \gamma_X(r, s) = \gamma_X(r + t, s + t), \forall s, r, t \in \mathbb{Z}.$$

Le processus $(X_t)_{t \in \mathbb{Z}}$ est stationnaire au second ordre si sa moyenne, sa variance et sa fonction d'autocovariance sont indépendantes de t , c'est-à-dire

$$\mathbb{E}(X_t) = m, \quad V(X_t) = \sigma^2 \quad \text{et} \quad \text{Cov}(X_t, X_{t+k}) = \gamma(k).$$

En d'autres termes, les moments d'ordre un et deux du processus existent et sont indépendants du temps.

Remarque 1.1.1.

1. Si $(X_t)_{t \in \mathbb{Z}}$ est stationnaire alors $\gamma_X(r, s) = \gamma_X(r - s, 0)$, $\forall r, s, t \in \mathbb{Z}$. Ainsi la fonction d'autocovariance d'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ est $\gamma_X(h) = \gamma_X(h, 0) = \text{Cov}_X(X_{t+h}, X_t)$.
2. La fonction d'autocorrélation du processus $(X_t)_{t \in \mathbb{Z}}$ stationnaire est :

$$\rho_X = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Corr}(X_{t+h}, X_t), \quad \forall t, h \in \mathbb{Z}.$$

Définition 1.1.4. La série temporelle $(X_t)_{t \in \mathbb{Z}}$ est dite strictement stationnaire si

$$(X_{t_1}, \dots, X_{t_n})' \stackrel{\text{loi}}{=} (X_{t_1+h}, \dots, X_{t_n+h})', \quad \forall h \in \mathbb{Z} \quad \text{et} \quad \forall (t_1, \dots, t_n)' \in \mathbb{Z}^n.$$

La stationnarité stricte implique la stationnarité au second ordre. La réciproque est fausse.

Définition 1.1.5. Un processus $\underline{X}_t = (X_{t,1}, X_{t,2}, \dots, X_{t,n})$ n -dimensionnel à temps discret sur $(\Omega, \mathcal{A}, \mathbb{P})$ est dit stationnaire au second ordre ou (faiblement stationnaire) si pour tout $i = 1, \dots, n$, chaque composante $X_{t,i}$ est stationnaire au second ordre et si de plus pour tout couple de composantes, la covariance est stationnaire. Ces définitions sont traduites par les points suivants :

1. $\mathbb{E}(X_{t,i}) = \mu_i$ pour $i = 1, \dots, n$,
2. $\mathbb{E}[(X_{t,i} - \mu_i)(X_{t+k,j} - \mu_j)] = \gamma_{ij}(k) \quad \forall k, t \in \mathbb{Z} \text{ et } i, j = 1, \dots, n.$

Définition 1.1.6 (Voir Straumann ([61], P14)). Soit $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité et (E, \mathcal{F}) un espace mesurable. Considérons une suite de variables aléatoires (X_t) définie sur Ω à valeurs dans E . On définit un espace canonique

$$E^{\mathbb{Z}} = \{(X_t) | X_t \in E \text{ pour tout } t \in \mathbb{Z}\}$$

muni de la σ -algèbre générée par la famille de produit de cylindre définie par

$$\mathcal{F}^{\mathbb{Z}} = \sigma\left(\left\{\prod_{i \in \mathbb{Z}} C_i \mid \exists n \neq 0 : C_i = E \text{ pour } |i| > n \text{ et } C_i \in \mathcal{F} \text{ pour } |i| \leq n\right\}\right)$$

telle que $(E^{\mathbb{Z}}, \mathcal{F}^{\mathbb{Z}})$ soit un espace mesurable. Alors la suite de variables aléatoires (X_t) peut également être considérée comme un élément aléatoire à valeurs dans $E^{\mathbb{Z}}$ et induisant une mesure $\tilde{\mathbb{P}}$ sur $(E^{\mathbb{Z}}, \mathcal{F}^{\mathbb{Z}})$ définie par $\tilde{\mathbb{P}}(C) = \mathbb{P}(X_t \in C)$, pour tout $C \in \mathcal{F}^{\mathbb{Z}}$.

L'ergodicité d'une suite de variables aléatoires est définie en terme d'espace canonique car elle implique une transformation qui préserve la mesure et qui est construite en général sur $(E^{\mathbb{Z}}, \mathcal{F}^{\mathbb{Z}}, \tilde{\mathbb{P}})$. Cette transformation est l'opérateur retard τ à valeurs dans $E^{\mathbb{Z}}$, mesurable et définie par $\tau(u_t) = u_{t-1}, t \in \mathbb{Z}$. De plus, une partie C de $\mathcal{F}^{\mathbb{Z}}$ est dite invariante si $\tau^{-1}(C) = C$. Une suite de variable aléatoire (X_t) est ergodique si pour tout ensemble invariant C , soit $\tilde{\mathbb{P}}(C) = 0$ ou $\tilde{\mathbb{P}}(C) = 1$. Les processus i.i.d sont des exemples standards de processus stationnaire et ergodique.

1.1.3 Convergence presque sûre

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs dans \mathbb{R}^d définies sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 1.1.7. On dit que (X_n) converge presque sûrement vers X et on note

$$X_n \xrightarrow{\text{p.s.}} X \text{ si } \mathbb{P}(\{\omega \in \Omega : \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\}) = 1.$$

1.1.4 Convergence en probabilité

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^d et X une variable aléatoire définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^d . Notons $\|\cdot\|$ une norme quelconque dans \mathbb{R}^d .

Définition 1.1.8. On dit que $(X_n)_{n \in \mathbb{N}}$ converge en probabilité vers X et on note

$$X_n \xrightarrow{\mathbb{P}} X \text{ si } \forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(\{\|X_n - X\| \geq \varepsilon\}) = 0.$$

Théorème 1.1.1. Soit f une application continue sur un intervalle $I \subset \mathbb{R}$ à valeurs réelles. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs dans I .

Si $X_n \xrightarrow{\mathbb{P}} X$ alors $f(X_n) \xrightarrow{\mathbb{P}} f(X)$.

1.1.5 Convergence en loi

Définition 1.1.9. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^d et X une variable aléatoire définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^d .

On dit que la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X et on note $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si pour toute fonction $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ continue et bornée, on a : $\mathbb{E}(g(X_n)) \rightarrow \mathbb{E}(g(X))$.

La convergence presque sûre implique la convergence en probabilité qui elle implique la convergence en loi.

1.1.6 Théorèmes limites

Deux théorèmes limites seront énoncés :

1. La loi forte des grands nombres qui énonce la convergence presque sûre de la moyenne empirique d'une suite de variables aléatoires indépendantes, identiquement distribuées et intégrables.
2. Le théorème central limite qui indique à quelle vitesse cette convergence a lieu sous l'hypothèse supplémentaire que les variables sont de carrés intégrables.

Théorème 1.1.2. (Loi forte des grands nombres de Kolmogorov)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes, identiquement distribuées et intégrables telles que $\mathbb{E}(X_n) = \mu < \infty$. Alors, on a :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\text{p.s.}} \mu.$$

Théorème 1.1.3. (Loi faible des grands nombres)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles vérifiant $\forall n \in \mathbb{N}, \mathbb{E}(X_n) = \mu$
 $V(X_n) = \sigma^2 < \infty$ et $\text{Cov}(X_n, X_m) = 0$ pour tout couple (n, m) tel que $n \neq m$. Alors, on a

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mu.$$

Théorème 1.1.4. (Théorème central limite)

X_1, \dots, X_n une suite de variables aléatoires i.i.d de carrés intégrables telles que $\mathbb{E}(X_1) = \mu < \infty$ et $V(X_1) = \sigma^2 < \infty$. Alors nous avons

$$\sqrt{n} \frac{(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

1.2 Espérance conditionnelle

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, \mathcal{A}_1 une sous tribu de \mathcal{A} c'est-à-dire $\mathcal{A}_1 \subset \mathcal{A}$, et soit $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, alors, il existe une unique variable aléatoire dans $L^1(\Omega, \mathcal{A}_1, \mathbb{P})$, notée $\mathbb{E}(X|\mathcal{A}_1)$, telle que

$$\forall B \subset \mathcal{A}_1, \quad \mathbb{E}(X1_B) = \mathbb{E}\left[\mathbb{E}(X|\mathcal{A}_1)1_B\right]. \quad (1.1)$$

De manière générale, pour toute variable aléatoire Y , \mathcal{A}_1 -mesurable bornée, on a :

$$\mathbb{E}(XY) = \mathbb{E}\left[\mathbb{E}(X|\mathcal{A}_1)Y\right]. \quad (1.2)$$

En particulier, si la tribu \mathcal{A}_1 est engendrée par une variable aléatoire Z , on notera

$$\mathbb{E}(X|\mathcal{A}_1) = \mathbb{E}(X|\sigma(Z)) = \mathbb{E}(X|Z). \quad (1.3)$$

1.2.1 Propriétés de l'espérance conditionnelle

Dans ce paragraphe, nous donnons quelques propriétés de l'espérance conditionnelle.

- a). Si $X \geq 0$ alors $\mathbb{E}(X|\mathcal{A}_1) \geq 0$.
- b). Si X est \mathcal{A}_1 -mesurable, alors on a $\mathbb{E}(X|\mathcal{A}_1) = X$, p.s.
- c). L'application $X \mapsto \mathbb{E}(X|\mathcal{A}_1)$ est linéaire.
- d). Nous avons $\mathbb{E}\left[\mathbb{E}(X|\mathcal{A}_1)\right] = \mathbb{E}(X)$.
- e). $|\mathbb{E}(X|\mathcal{A}_1)| \leq \mathbb{E}(|X||\mathcal{A}_1)$ p.s, ce qui conduit à $\mathbb{E}\left[|\mathbb{E}(X|\mathcal{A}_1)|\right] \leq \mathbb{E}(|X|)$.
- f). Si $X \geq Y$ p.s alors $\mathbb{E}(X|\mathcal{A}_1) \geq \mathbb{E}(Y|\mathcal{A}_1)$ p.s.
- g). Si g est une fonction positive et convexe alors on a : $\mathbb{E}\left(g(X)|\mathcal{A}_1\right) \geq g\left(\mathbb{E}(X|\mathcal{A}_1)\right)$ p.s.
- h). Soit X une variable aléatoire réelle et Z une variable aléatoire \mathcal{A}_1 -mesurable. Si X et Z sont positives ou si X et XZ sont intégrables alors : $\mathbb{E}(ZX|\mathcal{A}_1) = Z\mathbb{E}(X|\mathcal{A}_1)$ p.s.
- i). Soient \mathcal{A}_1 et \mathcal{A}_2 deux sous-tribus de \mathcal{A} telles que $\mathcal{A}_1 \subset \mathcal{A}_2$. Alors, pour toute variable aléatoire X positive ou intégrable, on a : $\mathbb{E}\left[\mathbb{E}(X|\mathcal{A}_2)|\mathcal{A}_1\right] = \mathbb{E}(X|\mathcal{A}_1)$ p.s.

Théorème 1.2.1. *Deux sous-tribus \mathcal{A}_1 et \mathcal{A}_2 sont indépendantes si et seulement si pour toute variable aléatoire Y , \mathcal{A}_2 -mesurable positive (ou pour toute variable aléatoire*

$Y \in L^1(\Omega, \mathcal{A}_2, \mathbb{P})$, ou encore pour toute variable aléatoire Y de la forme $Y = 1_B$, avec $B \in \mathcal{A}_2$, on a

$$\mathbb{E}(Y|\mathcal{A}_1) = \mathbb{E}(Y).$$

Remarque 1.2.1. Soient X et Y deux variables aléatoires réelles. Etant donnée que les variables aléatoires mesurables par rapport à la tribu engendrée par X notée $(\sigma(X))$ sont les fonctions de X , d'après le théorème précédent X et Y sont indépendantes si et seulement si pour toute fonction g borélienne telle que $\mathbb{E}(|g(X)|) < \infty$, on a

$$\mathbb{E}[g(X)|Y] = \mathbb{E}(g(X)).$$

En particulier, si X est intégrable, on a

$$\mathbb{E}[X|Y] = \mathbb{E}(X).$$

Définition 1.2.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(\mathcal{A}_n)_{n \in \mathbb{N}}$ une suite croissante de sous-tribus de \mathcal{A} (appelée filtration). Le processus X_n est une différence de martingale si :

1. X_n est \mathcal{A}_n -mesurable pour chaque n .
2. $\mathbb{E}(|X_n|) < \infty$ pour chaque n .
3. $\mathbb{E}(X_n|\mathcal{A}_{n-1}) = 0$

Si la troisième condition est remplacée par la condition suivante $\mathbb{E}(X_n|\mathcal{A}_{n-1}) = X_{n-1}$, alors on dit que X_n est une martingale.

1.3 Modèles à hétéroscédasticité conditionnelle

Soit p_t le cours d'un titre à l'instant t , on appelle rendement simple de p_t , le processus r_t défini par

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}}.$$

La volatilité d'un actif ou d'un titre est la variabilité ou la variance de ses rendements sur une période de temps donnée. Selon Aragon [2], la volatilité d'un titre n'est rien d'autre que l'écart type de son rendement. Lorsque la variance du processus est constante au cours du temps, on parle de processus homoscédastique. Au contraire, on parle d'hétéroscédasticité lorsque la variance du processus change au cours du temps.

1.3.1 Présentation des modèles ARCH

Les modèles de type ARCH (AutoRegressive Conditional Hétéroscedasticity) permettent de modéliser des séries chronologiques qui présentent une dépendance dans les données, une volatilité (où variance ou variabilité) qui change au fil du temps et une distribution à queue lourde et asymétrique c'est-à-dire non gaussienne. En d'autres termes, la variance du processus conditionnellement à son passé dépend du temps (hétéroscédasticité conditionnelle).

Définition 1.3.1. *Modèle ARCH(p)*

On dit que le processus X_t suit un processus ARCH(p) s'il est défini par :

$$\begin{cases} X_t = \varepsilon_t \sqrt{h_t} \\ h_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \end{cases}$$

où $\alpha_0 > 0$, pour $i = 1, \dots, p$, $\alpha_i \geq 0$ et ε_t est un processus i.i.d, centré, de variance σ^2 noté $\varepsilon_t \sim i.i.d(0, \sigma^2)$.

Ce processus peut aussi s'écrire de la manière suivante

$$X_t^2 = \varepsilon_t^2 [\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2].$$

1.3.2 Processus GARCH(p, q)

Un processus (X_t) est un processus GARCH(p, q) s'il s'écrit de la manière suivante :

$$\begin{cases} X_t = \varepsilon_t \sqrt{h_t} \\ h_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \end{cases} \quad (1.4)$$

avec $\alpha_0 > 0$, pour $i = 1, \dots, p$, $\alpha_i \geq 0$, pour $j = 1, \dots, q$, $\beta_j \geq 0$ et $\varepsilon_t \sim i.i.d(0, 1)$.

1.3.3 Processus FIGARCH et HYGARCH

Un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus GARCH fractionnaire intégré (FIGARCH(p,d,q)) s'il s'écrit de la manière suivante

$$\begin{cases} X_t = \varepsilon_t \sqrt{h_t} \\ h_t = \alpha_0 + \beta(L)h_t + \{1 - \beta(L) - \delta(L)(1 - L)^d\}X_t^2 \end{cases}$$

Un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus GARCH hyperbolique (HYGARCH(p,d,q)) s'il a la structure suivante :

$$\begin{cases} X_t = h_t^{1/2} \varepsilon_t \\ h_t = \gamma + \beta(L)h_t + \{1 - \beta(L) - \delta(L)[1 - \phi + \phi(1 - L)^d]\} X_t^2 \\ 0 < \phi < 1, \end{cases} \quad (1.5)$$

où $\varepsilon_t \sim i.i.d(0; 1)$ et $0 \leq d \leq 1$. Avec $\delta(L) = \sum_{j=1}^q \delta_j L^j$ et $\beta(L) = \sum_{i=1}^p \beta_i L^i$ deux polynômes en L de degrés respectifs q et p pour lesquels avec $\alpha_0, \gamma > 0$, pour $i = 1, \dots, p$, $\alpha_i \geq 0$, pour $j = 1, \dots, q$, $\beta_j \geq 0$.

1.4 Techniques d'approximation de la variance conditionnelle par initialisation

Le choix des initialisations dans la variance conditionnelle des modèles GARCH en général est une étape incontournable pour l'étude numérique du comportement des estimateurs. A cet effet, il existe différentes initialisations proposées dans la littérature. Nous pouvons citer celles définies dans Straumann ([61], page 74) ainsi que celles données dans Zakoïan et Francq ([15], page 142 et 294).

En considérant un échantillon x_1, x_2, \dots, x_n issu d'un processus GARCH(p,q) défini en (1.4), la log-vraisemblance conditionnelle gaussienne a la forme suivante :

$$\log f_\theta(x_1, \dots, x_n | x_0, \dots, x_{-q+1}, h_0, \dots, h_{-p+1}) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \left(\frac{x_t^2}{h_t(\theta)} + \log h_t(\theta) \right).$$

Puisque les valeurs de x_0, \dots, x_{-q+1} et celles de h_0, \dots, h_{-p+1} ne sont pas disponibles dans la pratique, la log-vraisemblance conditionnelle gaussienne ne peut donc être évaluée numériquement sans certaines initialisations sur les valeurs h_0, \dots, h_{-p+1} et x_0, \dots, x_{-q+1} . Ainsi Straumann [61] propose les initialisations suivantes :

$$X_t = 0 \quad \text{et} \quad \tilde{h}_t(\theta) = \frac{\alpha_0}{1 - \beta(1)} \quad \text{pour} \quad t \leq 0 \quad (1.6)$$

avec

$$\tilde{h}_t(\theta) = \alpha_0 + \sum_{i=1}^{\min(q, t-1)} \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \tilde{h}_{t-j} \quad \text{pour} \quad t \geq 1. \quad (1.7)$$

Zakoïan et Francq [15] ont également proposé les initialisations suivantes :

$$X_t = \tilde{h}_t(\theta) = 0 \quad \text{pour tout } t \leq 0 \quad (1.8)$$

ou encore

$$X_t = \tilde{h}_t(\theta) = \alpha_0 \quad \text{pour } t \leq 0 \quad (1.9)$$

De plus, Straumann [61] ainsi que Zakoïan et Francq [15] ont prouvé que ces valeurs initiales sont sans impact sur le comportement asymptotique de l'estimateur. La fonction $\tilde{h}_t(\theta)$ définie en (1.7), avec les différentes initialisations présentées dans cette partie peut être interprétée comme un prédicteur de la variance conditionnelle à l'instant t .

1.5 Méthodes de détection des outliers

1.5.1 La boîte à moustache ou boxplot

Cette méthode graphique ou visuelle, présentée dans Vinutha et al. [65], consiste à déterminer d'abord le premier et le troisième quartile notés respectivement (Q_1 et Q_3). Ensuite, l'écart interquartile I est calculé par la formule ($I = Q_3 - Q_1$). Enfin, deux seuils (R_1 et R_2) représentant les moustaches sont déterminés par les formules suivantes : $R_1 = Q_1 - 1.5I$ et $R_2 = Q_3 + 1.5I$. Toutes les valeurs inférieures à R_1 et supérieures à R_2 constituent l'ensemble de valeurs qui dévient de la majorité des données.

1.5.2 La méthode de Hampel

Cette méthode consiste à considérer comme outliers toutes valeurs en dehors de l'intervalle formé par la médiane, plus ou moins k ($k=3$ généralement) fois la déviation absolue de la médiane

$$I = [\text{médiane} - k * \text{mad}; \text{médiane} + k * \text{mad}]$$

avec mad : la déviation absolue de la médiane et $\text{mad} = \text{médiane}(|y_i - \tilde{y}|)$, \tilde{y} représente la médiane de la série y .

1.5.3 Le test statistique de Grubbs

Le test statistique de Grubbs (voir Grubbs [35]) permet de tester si une valeur est un outlier, la valeur maximale ou minimale de la série précisément. L'hypothèse nulle de ce test spécifie que la valeur considérée n'est pas un outlier, tandis que son hypothèse alternative indique que la valeur considérée est un outlier. La règle de décision est la

suivante : Si la p-value de la statistique de test est inférieure au niveau de significativité de 5% en général, la valeur considérée est une outlier. Pour réaliser ce test sur le logiciel R, on applique la fonction `grubbs.test()` du package "outliers".

1.5.4 Le test basé sur les travaux de Chen et Liu [14]

Cette technique utilise la fonction "locate.outliers" du package "tsoutliers" du logiciel R. Cette fonction, basée sur les travaux de Chen et Liu [14], permet de localiser les outliers dans des données de séries temporelles. Elle applique les statistiques de Student pour la signification des outliers à chaque instant et sélectionne celles qui sont significatives par rapport à d'une valeur critique.

1.6 M-estimateurs et Propriétés Générales

1.6.1 M-estimateurs et Robustesse d'un estimateur

Soient X une variable aléatoire réelle de loi P , définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et X_1, X_2, \dots, X_n un n-échantillon issu de X , on suppose que P appartient à la famille de lois $\{\mathbb{P}_\theta, \theta \in \Theta\}$. De manière générale, trouver un estimateur $\hat{\theta}_n$ de θ revient à maximiser en θ la fonction objective définie par

$$M_n(\theta) = \frac{1}{n} \sum_{i=1}^n \rho(x, \theta),$$

où ρ est une fonction mesurable en x . Un estimateur $\hat{\theta}_n$ qui maximise $M_n(\theta)$ sur Θ est appelé M-estimateur. Si de plus, ρ est dérivable (où admet des dérivées partielles) par rapport à θ alors $\hat{\theta}_n$ vérifie

$$\sum_{i=1}^n \psi(x, \hat{\theta}_n) = 0,$$

avec $\psi x, \hat{\theta}_n = \frac{\partial \rho(x, \theta)}{\partial \theta}$ pour $\theta = \hat{\theta}_n$. Le choix de la fonction ρ est donc déterminant pour identifier les différents types de M-estimateur. Il existe des exemples classiques de M-estimateurs. Nous pouvons citer entre autres :

1. l'estimateur du maximum de vraisemblance défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \sum_{i=1}^n -\log f(X_i, \theta)$$

pour lequel la fonction $\rho(x, \theta) = -\log f(X, \theta)$ avec $f(X, \theta)$ la vraisemblance du modèle.

2. L'estimateur de Huber [38] avec une fonction ρ définie par :

$$\rho(x) = \begin{cases} \frac{1}{2}x^2 & \text{si } |x| \leq k \\ k|x| - \frac{1}{2}k^2 & \text{si } |x| > k, \end{cases}$$

Définition 1.6.1. *Un estimateur est dit robuste lorsqu'il a la capacité de ne pas être perturbé par des petites modifications dans les paramètres du modèle choisi au départ pour l'estimation ou dans les données. Il présente toujours de bonnes propriétés même après une perturbation du modèle de départ.*

Dans ce présent manuscrit, il sera question de proposer et d'étudier deux classes de M-estimateurs robustes à partir d'une fonction ρ judicieusement choisie.

1.6.2 Propriétés Générales

Définition 1.6.2. *Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles avec $(v_n \neq 0$ pour tout $n \in \mathbb{N})$. On écrit que*

— $u_n = o(v_n)$ lorsque le quotient $\frac{u_n}{v_n}$ tend vers 0 quand n tend vers $+\infty$.

— $u_n = O(v_n)$ lorsque le quotient $\frac{u_n}{v_n}$ reste borné quand n tend vers $+\infty$.

Définition 1.6.3. *Une matrice carrée A d'ordre n symétrique est dite définie positive si elle vérifie l'une des propriétés équivalentes suivantes :*

1. *Pour tout vecteur non nul $x \in \mathbb{R}^n$, on a $x'Ax > 0$.*

2. *Toutes les valeurs propres de A sont strictement positives, c'est-à-dire $Sp(A) \in \mathbb{R}_+^*$.*

1.7 Résultats Fondamentaux

Théorème de convergence monotone (corollaire du théorème de convergence monotone) dans (Loeve [48], P125). Si X_n est une suite de variables aléatoires positives alors :

$$\mathbb{E}\left(\sum_n X_n\right) = \sum_n \mathbb{E}(X_n).$$

Critère d'intégrabilité dans (Loeve [48], P121) X est intégrable $\iff \mathbb{E}(|X|) < \infty \implies X < \infty$ p.s.

Lemme de Kronecker (Loeve [48], P250). Si la série de terme général x_n converge vers un nombre s fini et pour toute suite de réels positifs (b_n) divergeant vers l'infini, on a :

$$\lim_{n \rightarrow +\infty} \frac{1}{b_n} \sum_{k=1}^n b_k x_k \rightarrow 0.$$

Théorème 3.5.8 dans Stout ([60], P.182-183). Soit X_t une suite de variables aléatoires stationnaire et ergodique et soit $\rho : \mathbb{R}_\infty \rightarrow \mathbb{R}$ une fonction mesurable. Alors, la suite de variables aléatoires Y_t définie par $Y_t = \rho(X_t, X_{t+1}, \dots)$ est également strictement stationnaire et ergodique.

Théorème ergodique de Birkhoff (Straumann [61], P17.) Pour une suite X_t de variables aléatoires réelles, si la suite X_t est strictement stationnaire et ergodique avec $\mathbb{E}|X_0| < \infty$ alors

$$\frac{1}{n} \sum_{t=1}^n X_t \rightarrow \mathbb{E}(X_0) \quad p.s., \quad n \rightarrow \infty. \quad (1.10)$$

Théorème Central Limite des martingales (Corollaire) : Si le processus X_t est une différence de martingale par rapport à sa filtration naturelle, strictement stationnaire et ergodique, de moyenne zéro et de variance $v = \mathbb{E}(X_t^2) < \infty$, alors

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n X_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, v). \quad (1.11)$$

Critère de Cramer-Wold Soit $(X_n)_{n \in \mathbb{N}}$, une suite de vecteurs aléatoires dans \mathbb{R}^k ($k \geq 2$), alors pour tout vecteur non nul b de \mathbb{R}^k si

$$b'X_n \xrightarrow{\mathcal{L}} b'X \quad \text{alors} \quad X_n \xrightarrow{\mathcal{L}} X. \quad (1.12)$$

Théorème de convergence uniforme ou loi forte uniforme des grands nombres (Straumann [61], P19.) Soit X_t une suite de variables aléatoires définie sur $C(\Theta, \mathbb{R}^n)$. Alors, la loi forte uniforme des grands nombres est induite par la condition

$$\mathbb{E}(\|X_t\|_\Theta) < \infty.$$

Théorème 2.1.1. dans Stout [60], P.11 Soit $(T_n)_{n \leq 1}$ une suite de variables aléatoires telles que

$$\sum_{n=1}^{\infty} \mathbb{P}(|T_n| > \epsilon) < \infty,$$

pour chaque $\epsilon > 0$, alors

$$T_n \rightarrow 0 \quad \text{p.s.} \quad \text{quand} \quad n \rightarrow \infty.$$

Inégalité d' Hadamard. Soit $A = (a_{ij})$ une matrice carrée d'ordre n dont les vecteurs colonnes sont X_1, \dots, X_n . Pour $j = 1, \dots, n$, on note par $\|X_j\| = (\sum_{i=1}^n |a_{ij}|^2)^{1/2}$, alors on a :

$$|\det(A)| \leq \|X_1\| \dots \|X_n\|.$$

Par transformation, nous pouvons obtenir la relation suivante :

$$\begin{aligned} |\det(A)| &\leq \|X_1\| \dots \|X_n\| \\ &\leq \prod_{j=1}^n \|X_j\| = \prod_{j=1}^n \left(\left(\sum_{i=1}^n |a_{ij}|^2 \right)^{1/2} \right) \\ &\leq \prod_{j=1}^n \left(\left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2} \right) = \prod_{j=1}^n \|A\| \\ |\det(A)| &\leq \|A\|^n \end{aligned} \tag{1.13}$$

Technique de différentiation de matrice dans Horn et Johnsen [37] et Magnus et Neudecker [51]. Soient f une fonction définie sur un espace matriciel et à valeurs réelles et M une matrice dépendant d'un paramètre θ alors on a

$$\frac{\partial f(M)}{\partial \theta} = \sum_{i,j} \frac{\partial f(M)}{\partial m_{ij}} \frac{\partial m_{ij}}{\partial \theta} = \text{tr} \left(\frac{\partial f(M)}{\partial M'} \frac{\partial M}{\partial \theta} \right).$$

De plus, si M est inversible, alors pour tous vecteur y et matrices C et B avec les dimensions appropriées. Nous avons

$$\begin{aligned} \frac{\partial y' M y}{\partial M'} &= y y', & \frac{\partial \text{tr}(C' M' B M')}{\partial M'} &= C' M B' + B' M C', & \frac{\partial \log |M|}{\partial M'} &= M^{-1}, \\ \frac{\partial M^{-1}}{\partial \theta} &= -M^{-1} \frac{\partial M}{\partial \theta} M^{-1}, & \frac{\partial \log |C M^{-1} B|}{\partial M'} &= -M^{-1} B C M^{-1}, & \frac{\partial \log |C M B|}{\partial M'} &= B C. \end{aligned}$$

Théorème 7.4.1 dans Doukhan [25] P 167-168. Soit B un espace de Banach, (E, \mathcal{E}) un espace mesurable et soit $(\psi_n)_{n \in \mathbb{Z}}$ une suite i.i.d d'éléments de E . Supposons que $M : B^{\mathbb{N}^*} \times E \rightarrow B$ vérifie pour $p \geq 1$ et $x_0 \in B^{\mathbb{N}^*}$,

$$\|M(x_0, \psi_0)\|_p < \infty.$$

Supposons également qu'il existe une suite a_j , ($j \geq 1$) tels que pour tous $x = (x_j)_{j \geq 1}$, $y = (y_j)_{j \geq 1} \in B^{\mathbb{N}^*}$:

$$\begin{aligned} \|M(y, \psi_0) - M(x, \psi_0)\|_p &\leq \sum_{j \geq 1} a_j \|x_j - y_j\|, \\ a = \sum_{j \geq 1} a_j &< 1. \end{aligned} \tag{1.14}$$

Alors, il existe une solution strictement stationnaire dans \mathbb{L}^p pour l'équation :

$$X_t = M(X_{t-1}, X_{t-2}, X_{t-3}, \dots, \psi_t), \quad \forall t \in \mathbb{Z}. \tag{1.15}$$

L'unicité de la solution est vérifiée si $X_t = H(\psi_t, \psi_{t-1}, \dots)$, pour $H \in \mathbb{L}^p(E^{\mathbb{N}}, \mathcal{E}^{\otimes \mathbb{N}}, \mu^{\otimes \mathbb{N}})$, avec μ la distribution de ψ_0 .

Proposition 1 dans Bardet et Wintenberger [4]. Soit Ψ , un symbole générique désignant les fonctions f , M ou H et Θ un sous ensemble compact de \mathbb{R}^d . La fonction $\Psi_\theta : (\mathbb{R}^m)^\infty \rightarrow V$ est une fonction borélienne à valeurs dans un espace vectoriel de dimension finie V .

Supposons que la fonction Ψ_θ vérifie la condition $\|\Psi_\theta(0)\|_\Theta < \infty$ et qu'il existe une suite $(\alpha_j(\Psi, \Theta))_j$ de nombres positifs tels que pour tout $x, y \in (\mathbb{R}^m)^\infty$, on a

$$\begin{aligned} \|\Psi_\theta(x) - \Psi_\theta(y)\| &\leq \sum_{j=1}^{\infty} \alpha_j(\Psi, \Theta) \|x_j - y_j\|, \quad \text{avec} \\ \sum_{j=1}^{\infty} \alpha_j(\Psi, \Theta) &< \infty. \end{aligned}$$

Supposons de plus que

$$\sum_{j=1}^{\infty} \alpha_j(f, \theta) + \mathbb{E}(\|\varepsilon_t\|^r)^{1/r} \sum_{j=1}^{\infty} \alpha_j(M, \theta) < 1, \tag{1.16}$$

alors il existe une unique solution strictement stationnaire, causale, ergodique du modèle

$$X_t = M_{\theta_0}(X_{t-1}, X_{t-2}, \dots) \cdot \varepsilon_t + f_{\theta_0}(X_{t-1}, X_{t-2}, \dots) \quad \forall t \in \mathbb{Z}, \tag{1.17}$$

qui vérifie également $\mathbb{E}\|X_0\|^r < \infty$, avec $r \geq 1$. Pour $\Psi = H$ les quantités x, y, x_j, y_j sont respectivement remplacés par $xx', yy', x_jx'_j$ et $y_jy'_j$.

Théorème 4.1.1 dans Amemiya [1]. Notons par y , la suite des observations (y_t) jusqu'à l'instant T . Supposons que les hypothèses suivantes sont vérifiées.

- (i) L'espace des paramètres Θ est un sous-ensemble compact d'un K-espace euclidien (\mathbb{R}^K) et $\theta_0 \in \Theta$.
- (ii) La fonction $L_T(\theta, y)$ est continue en $\theta \in \Theta$ pour tout y et est une fonction mesurable de y pour tout $\theta \in \Theta$.
- (iii) $L_T(\theta)$ converge uniformément en probabilité en $\theta \in \Theta$ vers $L(\theta)$, une fonction déterministe quand T tend vers l'infini.
- (iv) $L(\theta)$ atteint un unique maximum en θ_0 .

Alors la quantité $\hat{\theta}_T$ définie par

$$L_T(\hat{\theta}_T) = \max_{\theta \in \Theta} L_T(\theta) \quad (1.18)$$

converge en probabilité vers θ_0 .

Théorème 4.1.3 dans Amemiya [1]. Supposons que les hypothèses du Théorème 4.1.1 précédent sont vérifiées. Supposons de plus que les hypothèses suivantes sont vérifiées.

- (A1) $\partial^2 L_T / \partial \theta \partial \theta'$ existe et est continue dans un voisinage ouvert, convexe de θ_0 .
- (A2) $(\partial^2 L_T / \partial \theta \partial \theta')_{\theta_T^*}$ converge en probabilité vers une matrice singulière finie $A(\theta_0) = \lim_{T \rightarrow \infty} \mathbb{E}((\partial^2 L_T / \partial \theta \partial \theta')_{\theta_0})$ pour toute suite θ_T^* qui converge en probabilité vers θ_0 quand T tend vers l'infini.
- (A3) $\frac{1}{\sqrt{T}}(\partial L_T / \partial \theta)$ converge en loi vers $\mathcal{N}(0, B(\theta_0))$ avec $B(\theta_0) = \lim \mathbb{E}((\partial L_T / \partial \theta)_{\theta_0} (\partial L_T / \partial \theta')_{\theta_0})$ quand T tend vers l'infini.

Alors, on a

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1}), \quad T \rightarrow \infty. \quad (1.19)$$

1.7.1 Résultat de Straumann([61], p.119)

Le but est de montrer que la matrice $\mathbb{E}(\nabla h_t(\theta) \nabla' h_t(\theta))$ est définie positive pour $\theta = \theta_0$. Cela revient à montrer que $\mathbb{E}(\lambda' \nabla h_t(\theta) \nabla' h_t(\theta) \lambda) > 0$, pour tout vecteur non nul λ de \mathbb{R}^d . Nous avons

$$\begin{aligned} \mathbb{E}\{\lambda' \nabla h_t(\theta) \nabla' h_t(\theta) \lambda\} &= \mathbb{E}\{(\nabla' h_t(\theta) \lambda)' (\nabla' h_t(\theta) \lambda)\} \\ &= \mathbb{E}\{(\nabla' h_t(\theta) \lambda)^2\} \geq 0. \end{aligned} \quad (1.20)$$

Supposons que $\mathbb{E}\{(\nabla' h_t(\theta) \lambda)^2\} = 0$, cela implique que

$$\nabla' h_t(\theta) \lambda = 0 \quad \text{p.s.} \quad (1.21)$$

Soit la fonction : $(\theta, s) \mapsto g_\theta(x, s)$ continument différentiable sur $\Theta \times [0, \infty[$ pour tout $x \in \mathbb{R}^p$, avec Θ un compact de R^d . La variance conditionnelle des processus GARCH y_t présente la forme suivante

$$h_{t+1}(\theta) = g_\theta(y_t, h_t(\theta)). \quad (1.22)$$

Par différentiation des deux membres de (1.22) par rapport à θ , nous obtenons

$$\nabla' h_{t+1}(\theta) = \frac{\partial g_\theta}{\partial \theta'}(y_t, h_t(\theta)) + \frac{\partial g_\theta}{\partial h}(y_t, h_t(\theta)) \nabla' h_t(\theta). \quad (1.23)$$

En composant chaque membre de (1.23) à droite par le vecteur λ , nous obtenons

$$\nabla' h_{t+1}(\theta) \lambda = \frac{\partial g_\theta}{\partial \theta'}(y_t, h_t(\theta)) \lambda + \frac{\partial g_\theta}{\partial h}(y_t, h_t(\theta)) \nabla' h_t(\theta) \lambda. \quad (1.24)$$

A partir de (1.21), la relation (1.23) devient :

$$\nabla' h_{t+1}(\theta) \lambda = \frac{\partial g_\theta}{\partial \theta'}(y_t, h_t(\theta)) \lambda = 0. \quad (1.25)$$

Pour $\theta = \theta_0$ et par stationnarité des processus (h_t) et (∇h_t) pour tout t , la relation (1.25) équivaut aussi à

$$\nabla' h_1(\theta_0) \lambda = \frac{\partial g_\theta}{\partial \theta'}(y_0, h_0(\theta))|_{\theta=\theta_0} \lambda = 0. \quad (1.26)$$

En supposant, de plus que les composantes du vecteur $\frac{\partial g_\theta}{\partial \theta'}(y_0, h_0)|_{\theta=\theta_0}$ sont des variables aléatoires linéairement indépendantes. La relation (1.26) implique que $\lambda = 0$. Ce qui est contraire à l'hypothèse de départ. Donc

$$\mathbb{E}\{\lambda' \nabla h_t(\theta) \nabla' h_t(\theta) \lambda\} > 0.$$

1.7.2 Résultat de Kwan et al.[44]

Considérons la dérivée partielle de la variance conditionnelle du modèle HYGARCH défini en (1.5) par rapport à d , nous obtenons

$$\frac{\partial h_t(\theta)}{\partial d} = \phi \delta(L) (1-L)^d \left(\sum_{j=1}^{\infty} \frac{L^j}{j} \right) y_t^2 + \sum_{k=1}^p \beta_p \frac{\partial h_{t-k}(\theta)}{\partial d} \quad (1.27)$$

et

$$\frac{\partial^2 h_t(\theta)}{\partial d^2} = \phi \delta(L) (1-L)^d \left(\sum_{j=1}^{\infty} \frac{L^j}{j} \right)^2 y_t^2 + \sum_{k=1}^p \beta_p \frac{\partial^2 h_{t-k}(\theta)}{\partial d^2}. \quad (1.28)$$

Etant donné que $(1 - L)^d = 1 - \sum_{j=1}^{\infty} a_j L^j$ avec $\sum_{j=1}^{\infty} a_j = 1$ et $a_j = O(j^{-1-d})$, nous notons

$$(1 - L)^d \left(\sum_{j=1}^{\infty} \frac{L^j}{j} \right) = \sum_{j=1}^{\infty} g_j L^j \quad (1.29)$$

avec

$$g_j = \frac{1}{j} - \left(\frac{a_1}{j-1} + \frac{a_2}{j-2} + \dots + \frac{a_{j-1}}{1} \right) \leq \frac{1}{j} \left(1 - \sum_{i=1}^{j-1} a_i \right) = O(j^{-1-d})$$

ce qui montre que $\sum_{j=1}^{\infty} |g_j| < \infty$. De manière similaire, les coefficients de $(1-L)^d \left(\sum_{j=1}^{\infty} \frac{L^j}{j} \right)^2$ sont également absolument sommable.

Chapitre 2

M-estimateurs des modèles GARCH hyperboliques univariés stationnaires

Ce chapitre traite du problème de l'estimation des paramètres du modèle GARCH hyperbolique univarié. La méthode du quasi-maximum de vraisemblance utilisée jusqu'à présent pour estimer les paramètres du modèle GARCH hyperbolique s'avère être très sensible à la présence de valeurs aberrantes dans les observations. D'où la nécessité de proposer une technique d'estimation plus adaptée au traitement des valeurs aberrantes dans les données. La méthode d'estimation présentée dans cette partie utilise deux classes de M-estimateurs toutes basées sur une fonction critère bornée minutieusement choisie. La première classe d'estimateur est appelée M-estimateur et la seconde BM-estimateur. Le BM-estimateur est une version modifiée du M-estimateur en incluant un mécanisme qui permet de limiter la propagation des effets des valeurs aberrantes dans la variance conditionnelle. Dans ce travail, nous présentons d'abord le modèle GARCH hyperbolique univarié ainsi que la propriété relative à l'existence et à l'unicité d'une solution strictement stationnaire du modèle. Les conditions liées à l'existence de certains moments d'ordre du modèle sont également fournies. Puis, nous présentons les deux classes de M-estimateurs. Enfin, nous établissons les propriétés de consistance et de normalité asymptotique de ces deux classes de M-estimateurs. .

2.1 Processus GARCH hyperbolique univarié

Dans cette partie, nous donnons la description du modèle GARCH hyperbolique univarié ainsi que certaines de ses propriétés.

2.1.1 Présentation du modèle

Dans le but de modéliser au mieux la volatilité de certaines séries temporelles, Li et al. [46] ont proposé un modèle GARCH hyperbolique noté modèle HGARCH(p, d, q). Ce modèle HGARCH(p, d, q) est d'un grand intérêt car il présente un caractère GARCH, une mémoire longue induite par le choix du paramètre d et une variance finie. Le modèle HGARCH(p, d, q) a la forme suivante :

$$\begin{cases} y_t = h_t^{1/2} \varepsilon_t \\ h_t(\theta) = \gamma + \beta(L)h_t + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}y_t^2, \end{cases} \quad (2.1)$$

avec $(\varepsilon_t)_{t \in \mathbb{Z}}$ qui désigne une suite de variables aléatoires i.i.d de moyenne nulle et de variance 1. La variance conditionnelle est $h_t(\theta) = V(y_t | \mathcal{F}_{t-1})$ avec $\mathcal{F}_{t-1} = \sigma(\{\varepsilon_s\}_{s \leq t-1})$ qui représente la filtration générée par le passé de ε_t jusqu'à la date $(t - 1)$. Les paramètres γ et ω sont strictement positifs. L est l'opérateur retard défini par $Lx_t = x_{t-1}$, tels que les polynômes $\beta(L)$ et $\delta(L)$ soient définis par

$$\beta(L) = \sum_{i=1}^p \beta_i L^i \quad \text{et} \quad \delta(L) = 1 - \sum_{j=1}^q \delta_j L^j,$$

où p et q désignent des entiers positifs et pour tous $i = 1, \dots, p$ et $j = 1, \dots, q$, nous avons $\beta_i, \delta_j \geq 0$. Le paramètre d est tel que $0 < d < 1$ avec l'opérateur de différence fractionnel $(1 - L)^d$ défini selon Li et al. [45] par

$$(1 - L)^d = 1 - \sum_{j=1}^{\infty} \pi_j(d) L^j, \quad (2.2)$$

avec

$$\pi_j(d) = \frac{d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} = O(j^{-1-d}) \quad \text{pour } j \geq 0 \quad \text{et} \quad 0 < d < 1,$$

où $\Gamma(\cdot)$ est la fonction Gamma définie par

$$\Gamma(r) = \begin{cases} (r-1)! & \text{si } r \in \mathbb{N}^*, \\ \int_0^{+\infty} \exp(x)x^{r-1} dx & \text{sinon.} \end{cases}$$

Le paramètre d'intérêt θ porté par la variance conditionnelle est un vecteur de paramètres défini par $\theta = (\gamma, \beta, \delta, \omega, d) \in \mathbb{R}^{p+q+3}$. Comme dans Li et al. [46], nous réécrivons la variance conditionnelle (2.1) sous la forme suivante :

$$\begin{aligned} h_t(\theta) &= \frac{\gamma}{1 - \beta(1)} + \omega \left(1 - \frac{\delta(L)}{1 - \beta(L)} (1 - L)^d\right) y_t^2 \\ &= \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{\infty} b_j(\theta) y_{t-j}^2, \end{aligned} \quad (2.3)$$

avec $b_j > 0$ et $b_j(\theta) = O(j^{-1-d})$ pour tout $j \geq 1$. De plus, on a $\sum_{j=1}^{\infty} b_j < \infty$. Cette forme (2.3) confère au modèle HGARCH une structure de modèle ARCH(∞). Remarquons que les conditions liées à l'existence et à l'unicité d'une solution strictement stationnaire et causale des modèles ARCH(∞) ont été étudiées par Douc et al. [24].

2.1.2 Propriétés du modèle HGARCH

En mettant en relation les modèles HGARCH(p, d, q) et ARCH(∞) dans (2.1) et (2.3), Li et al. [46] établissent sur la base du théorème Douc et al. [24], des résultats portant sur l'existence et l'unicité d'une solution strictement stationnaire du modèle HGARCH(p, d, q). Nous rappelons à travers le paragraphe suivant le théorème des auteurs Li et al. [46] relatif à l'existence et l'unicité d'une solution strictement stationnaire du modèle HGARCH(p, d, q) ainsi que les points essentiels de leurs résultats.

Théorème 2.1.1. *Supposons qu'il existe $m \in]0; 1]$ tel que $\mathbb{E}(|\varepsilon_t|^{2m}) \sum_{j=1}^{\infty} b_j^m < 1$. Alors il existe une unique solution strictement stationnaire du modèle ARCH(∞) défini en (2.3) avec $\mathbb{E}(|y_t|^{2m}) < \infty$, et la variance conditionnelle a la forme suivante*

$$h_t(\theta) = \frac{\gamma}{\beta(1)} + \frac{\gamma}{\beta(1)} \sum_{k=1}^{\infty} \sum_{j_1, \dots, j_k \geq 1} b_{j_1} \dots b_{j_k} \varepsilon_{t-j_1}^2 \dots \varepsilon_{t-j_1-j_2-\dots-j_k}^2.$$

Conséquences du Théorème 2.1.1 : Nous faisons un résumé de certains résultats de Li et al. [46] qui pourraient être utile dans la suite du manuscrit.

1. Lorsque $m = 1$, il s'en suit dans le Théorème 2.1.1 que $\mathbb{E}(|\varepsilon_t|^2) \sum_{j=1}^{\infty} b_j = \omega < 1$. Alors il existe toujours une solution strictement stationnaire pour le modèle HGARCH(p, d, q) défini en (2.1) avec $\mathbb{E}(|y_t|^2) < \infty$.
2. Lorsque $\omega \geq 1$, il est impossible qu'une solution stationnaire à variance finie puisse exister. L'existence d'une solution stationnaire dans ce cas dépendra de la distribution du processus ε_t comme pour les modèles GARCH standard (voir

Bougerol et al. [12]).

3. Supposons qu'il existe une solution stationnaire du processus HGARCH noté $(y_t)_t$:

$$\text{Si } \omega < [\mathbb{E}(|\varepsilon_t|^{2m})]^{-\frac{1}{m}} \text{ alors } \mathbb{E}(|y_t|^{2m}) < \infty, \text{ avec } m = 1, 2, 3, 4. \quad (2.4)$$

A partir des propriétés de stationnarité et d'existence des moments d'ordre $n = 1, \dots, 8$ du modèle HGARCH(p, d, q), nous supposons la condition suivante :

(**A**₀) : La distribution de la suite de variables aléatoires $(\varepsilon_t)_t$ admet une densité f absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R} et on a : $f(\cdot) > 0$.

2.2 Estimation des paramètres

Dans cette partie, nous proposons une technique d'estimation des paramètres du modèle HGARCH(p, d, q) en utilisant deux classes de M-estimateurs. Cette technique a été utilisée précédemment par Muler et Yohai [55] ainsi que Lu et al. [49] pour les modèles GARCH et MEM respectivement.

2.2.1 Présentation des M-estimateurs

La première classe d'estimateurs proposée dans ce travail est appelée "M-estimateur" et est définie par minimisation d'une fonction critère bornée convenablement choisie. La seconde classe est appelée "BM-estimateur". Elle est une version modifiée de la première classe puisqu'elle contient un mécanisme qui contrôle la propagation des effets des outliers dans la variance conditionnelle. Nous supposons que y_1, y_2, \dots, y_n est un échantillon de variables aléatoires généré par le modèle HGARCH(p, d, q) défini en (2.1) et

$$\theta_0 = (\gamma_0, \beta_0', \delta_0', \omega_0, d_0)' \in \mathbb{R}^{p+q+3} \quad (2.5)$$

le vrai vecteur de paramètres de ce modèle. Dans la suite du travail, nous supposons que

$$x_t = \log(y_t^2) \quad (2.6)$$

$$w_t = \log(\varepsilon_t^2), \quad (2.7)$$

en appliquant ces transformations à (2.1), nous obtenons

$$x_t = w_t + \log(h_t(\theta_0)). \quad (2.8)$$

En utilisant l'hypothèse (\mathbf{A}_0) dans la partie 2.1.2, nous remarquons que la suite de variables aléatoires (w_t) admet une densité que nous noterons $g(\cdot)$. Cette fonction g est définie par Lu et al. [49] et Muler et Yohai [55] de la manière suivante : si la densité f de ε_t est symétrique par rapport à zéro, alors

$$g(w) = f(e^{(w/2)})e^{(w/2)}. \quad (2.9)$$

Dans le cas où f correspond à la densité de la loi normale centrée réduite $\mathcal{N}(0, 1)$, alors $g = g_0$ et

$$g_0(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(e^w - w)}. \quad (2.10)$$

Pour les modèles HGARCH, la première classe de M-estimateurs que nous proposons appelée "M-estimateur" est inspirée des travaux de Muler et Yohai [55], tandis que la deuxième classe nommée "BM-estimateur" est construite comme dans Lu et al. [49]. La construction des deux classes de M-estimateurs pour les modèles HGARCH se déroule comme suit. Soit

$$\theta = (\gamma, \beta', \delta', \omega, d)' \in \mathbb{R}^{p+q+3}, \quad (2.11)$$

le vecteur de paramètres d'intérêt. Nous définissons la suite de fonction suivante qu'on utilisera pour la construction de la première classe de M-estimateurs. Ainsi, nous avons :

$$M_n(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \rho(x_t - \log h_t(\theta)), \quad (2.12)$$

avec x_t défini en (2.6), $h_t(\theta)$ la variance conditionnelle du processus (y_t) et ρ la fonction supposée bornée. Cependant, par construction dans (2.3), la variance conditionnelle $h_t(\theta)$ dépend de tout le passé du processus y_t . Il se trouve que dans la pratique, les données $(y_i, i \leq 0)$ c'est-à-dire $y_0^2, \dots, y_{1-q}^2, \dots$ sont non observables. Ainsi, la fonction $M_n(\theta)$ définie en (2.12) est une fonction théorique qui ne peut être évaluée numériquement sans émettre des conditions préalables sur les valeurs $(y_i, i \leq 0)$ non observables. En suivant les mêmes techniques que Straumann [61] (voir section 1.4), nous proposons les valeurs initiales suivantes pour les données non observables. Nous supposons

$$y_t = 0 \quad \text{pour } t \leq 0. \quad (2.13)$$

Dans ce cas, une initialisation de la variance conditionnelle est donnée par

$$\tilde{h}_t(\theta) = \frac{\gamma}{1 - \beta(1)} \quad \text{pour } t \leq 0 \quad (2.14)$$

En utilisant (2.13) et (2.14), on obtient

$$\tilde{h}_t(\theta) = \gamma + \beta(L)\tilde{h}_t(\theta) + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}y_t^2, \quad (2.15)$$

$$= \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{t-1} b_j(\theta)y_{t-j}^2 \quad \text{pour } t \geq 1, \quad (2.16)$$

où $\tilde{h}_t(\theta)$ est la variance conditionnelle de (y_t) calculée avec ces valeurs initiales. Ainsi, $\tilde{h}_t(\theta)$ défini en (2.13) représente le premier prédicteur de la variance conditionnelle dans ce travail. La première classe de M-estimateurs que nous proposons donc pour les modèles HGARCH(p, d, q) est définie par minimisation de la fonction

$$\tilde{M}_n(\theta) = \frac{1}{n - p} \sum_{t=p+1}^n \rho(x_t - \log \tilde{h}_t(\theta)). \quad (2.17)$$

Ainsi, nous définissons le M-estimateur du modèle HGARCH(p, d, q) par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{M}_n(\theta), \quad (2.18)$$

avec Θ qui représente l'espace des paramètres, un ensemble compact de \mathbb{R}^{p+q+3} . Par analogie avec la construction du M-estimateur, nous remarquons que l'estimateur du quasi-maximum de vraisemblance du modèle HGARCH(p, d, q) se définit par minimisation de la fonction

$$M_{0,n}(\theta) = \frac{1}{n - p} \sum_{t=p+1}^n \rho_0(x_t - \log \tilde{h}_t(\theta)), \quad (2.19)$$

avec

$$\rho_0 = -\log(g_0), \quad (2.20)$$

g_0 est donné dans (2.10) et x_t est défini en (2.6). Cependant, il paraît plus qu'important de s'intéresser à la variance conditionnelle $h_t(\theta)$ définie en (2.1). En effet, lorsqu'un outlier apparaît dans une série à l'instant t , le calcul des observations $h_{t+u}(u > 0)$ s'en trouve affecté puisque $h_{t+u}(u > 0)$ est lié à son passé. Ainsi, tout comme l'estimateur du quasi-maximum de vraisemblance, les variances conditionnelles $h_t(\theta)$ et $\tilde{h}_t(\theta)$ s'avèrent être également très sensibles aux outliers. Il va s'en dire que cette sensibilité aura un impact sur la robustesse du M-estimateur. Ce qui est problématique. Afin d'analyser au mieux ce phénomène, nous proposons une deuxième classe d'estimateurs appelée le BM-estimateur similaire à l'estimateur présenté dans Lu et al. [49]. Le BM-estimateur est une version modifiée du M-estimateur défini en (2.18) car il contient un mécanisme qui limite la propagation des effets des outliers sur $\tilde{h}_t(\theta)$, le premier prédicteur de la variance conditionnelle. En d'autres termes, le BM-estimateur est construit à partir

d'un second prédicteur de la variance conditionnelle. L'expression de ce prédicteur se présente comme suit :

$$\bar{h}_{t,l}(\theta) = \gamma + \beta(L)\bar{h}_{t,l}(\theta) + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}\bar{h}_{t,l}(\theta)s_l\left(\frac{y_t^2}{\bar{h}_{t,l}(\theta)}\right), \quad (2.21)$$

avec

$$s_l(u) = \begin{cases} u & \text{si } u \leq l \\ l & \text{si } u > l, \end{cases} \quad (2.22)$$

et $y_t = 0$ pour $t \leq 0$. Par suite, la fonction $M_n(\theta)$ définie (2.12) peut être modifiée comme suit

$$\bar{M}_{nl}(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \rho(x_t - \log \bar{h}_{t,l}(\theta)). \quad (2.23)$$

L'estimateur issu de $\bar{M}_{nl}(\theta)$ est le BM-estimateur et il est défini par

$$\hat{\theta}_n^B = \arg \min_{\theta \in \Theta} \bar{M}_{nl}(\theta), \quad (2.24)$$

Remarquons que plus l est grand, plus $\tilde{h}_t(\theta)$ et $\bar{h}_{t,l}(\theta)$ sont de plus en plus proches. Ce qui entrainerait l'équivalence de $\hat{\theta}_n$ et $\hat{\theta}_n^B$.

2.2.2 Résultats Asymptotiques

Dans ce paragraphe, nous établissons les propriétés asymptotiques (consistance et normalité asymptotique) du M-estimateur ($\hat{\theta}_n$) du modèle HGARCH(p, d, q). Pour ce faire, la supposition des hypothèses suivantes est nécessaire.

- (A₁) Le processus y_t est une solution strictement stationnaire et ergodique du modèle HGARCH(p, d, q) défini en (2.1). Le polynôme $\beta(L)$ est tel que $\sum_{j=1}^p \beta_j < 1$, $j = 1, \dots, p$. Les polynômes $\beta(L)$ et $\delta(L)$ n'ont aucune racines communes.
- (A₂) Θ est un ensemble compact de \mathbb{R}^{p+q+3} , $\theta_0 \in \overset{\circ}{\Theta}$ est le vrai paramètre et $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$.
- (A₃) ρ est continue, bornée, croissante et ses trois premières dérivées sont chacune continues et bornées.
- (A₄) Considérons $J(u) = \mathbf{E}(\rho(w_t - u))$, avec w_t défini en (2.7). Alors, il existe une fonction ρ qui respecte (A₃) telle que $J(u)$ atteint son minimum $J(u_0)$ en un unique u_0 .

(A₅) Soit ψ la dérivée première de ρ . Nous supposons les relations suivantes :
 $\mathbf{E}(\psi(x_t - \log h_t(\theta_0))) = 0$, $\mathbf{E}(\psi^2(x_t - \log h_t(\theta_0))) < \infty$ et $\mathbf{E}(\psi'(x_t - \log \tilde{h}_t(\theta_0))) > 0$.

(A₆) Pour tout $j \geq 1$, la quantité $b_j(\theta)$ définie en (2.3) a ses $k^{\text{ième}}$ dérivées continues sur \mathbb{R}^{p+q+2} telle que pour K positif, on a

$$\left| \frac{\partial^k b_j(\theta)}{\partial \theta_{i_1} \dots \partial \theta_{i_k}} \right| \leq K b_j^{1-\eta}(\theta) \quad \text{pour tout } \eta > 0,$$

θ_i est la i -ème composante du vecteur θ .

(A₇) Soit $(\theta, s) \mapsto g_\theta(x, s)$ une fonction continue et différentiable sur $\Theta \times [0, \infty[$ telle que $h_{t+1}(\theta) = g_\theta(y_t, h_t)$.

Les composantes du vecteur $\frac{\partial g_\theta}{\partial \theta'}(y_0, h_0)|_{\theta=\theta_0}$ sont des variables aléatoires linéairement indépendantes.

Remarque 2.2.1. *Les hypothèses (A₁) et (A₂) sont des hypothèses standards. La condition $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$ issue de (2.4) est suffisante pour garantir l'existence du moment d'ordre 4 du processus, condition cruciale pour cette étude. De plus, la condition sur les polynômes $\beta(L)$ et $\delta(L)$ permet d'assurer l'identifiabilité du modèle.*

Les hypothèses (A₃) et (A₄) sont nécessaires pour prouver la consistance de l'estimateur. Ces hypothèses seront utilisées pour montrer l'existence et l'unicité de la solution en laquelle $\mathbf{E}(\rho(x_t - \log h_t(\theta)))$ atteint son minimum.

L'hypothèse (A₆) est la même que l'hypothèse F(l) dans Robinson et Zaffaroni [58]. L'hypothèse (A₇) est similaire à la condition N4 dans Straumann ([61], p.85-86;116). Elles permettent avec les hypothèses techniques (A₅) d'établir les lemmes utilisés pour la preuve de la normalité asymptotique.

Comme dans Muler et Yohai [55], w_t est intégrable et la quantité $J(u)$ en (A₄) est finie pour tout u si la dérivée première de ρ est finie.

Pour satisfaire empiriquement les théorèmes suivants, la série temporelle doit être un modèle GARCH à mémoire longue, strictement stationnaire, ergodique et de moment d'ordre 4 fini.

Notations

Pour une meilleure compréhension de cette partie, nous adoptons les notations suivantes.

Soit $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ une matrice carrée d'ordre n . Nous supposons que

$$\|A\| = \left(\text{tr}(AA') \right)^{1/2} = \left(\sum_{1 \leq i, j \leq n} |a_{i,j}|^2 \right)^{1/2}. \quad (2.25)$$

Pour $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, nous notons par $\|u\| = \sum_{i=1}^n |u_i|$. Par les propriétés de la norme matricielle, nous avons pour tout vecteur colonne $u, v \in \mathbb{R}^n$,

$$\|uv'\| \leq \|u\| \|v\|, \quad (2.26)$$

De plus, pour toute fonction H continue sur un compact de \mathbb{R}^n à valeurs dans l'espace des matrices carrées d'ordre n à valeurs réelles c'est-à-dire $H \in C(\Theta, \mathbb{R}^{n \times n})$, on définit par

$$\|H\|_{\Theta} = \sup_{\theta \in \Theta} \|H(\theta)\|. \quad (2.27)$$

Étant donnée une fonction différentiable $f(u_1, \dots, u_n)$, nous notons $\nabla f = \left(\frac{\partial f}{\partial u_1}, \dots, \frac{\partial f}{\partial u_n} \right)$ et $\nabla^2 f$ la matrice dont les coefficients (i, j) sont $\frac{\partial^2 f}{\partial u_i \partial u_j}$.

Le théorème 2.2.1 montre que sous certaines conditions, la fonction $M(\theta)$ atteint son minimum en une unique valeur θ_0 définie en (2.5).

Théorème 2.2.1. *Supposons que les hypothèses (\mathbf{A}_0) , (\mathbf{A}_4) et (\mathbf{A}_7) sont satisfaites. Soit $x_t = \log(y_t^2)$ et définissons pour tout $\theta \in \mathbb{R}^{p+q+3}$*

$$M(\theta) = \mathbf{E}(\rho(x_t - \log h_t(\theta))). \quad (2.28)$$

Alors $M(\theta)$ atteint son minimum en un unique vecteur θ lorsque les paramètres qui composent ce vecteur θ valent respectivement $\gamma = e^{u_0} \gamma_0$, pour tous $1 \leq i \leq p$, $\beta_i = \beta_{0i}$, $1 \leq j \leq q$, $\delta_j = \delta_{0j}$, $\omega = e^{u_0} \omega_0$ et $d = d_0$, c'est-à-dire $\theta = (e^{u_0} \gamma_0, \beta'_0, \delta'_0, e^{u_0} \omega_0, d_0)'$

Remarque 2.2.2. *Supposons que la densité $g(w)$ de la suite de variables aléatoires $w_t = \log(\varepsilon_t^2)$, définie en (2.9), soit continue, positive et unimodale pour tout w . De plus, si nous considérons $\rho = m(-\log(g))$, avec m une fonction monotone alors l'hypothèse (\mathbf{A}_4) est vérifiée avec $u_0 = 0$.*

Nous donnons dans le paragraphe suivant la preuve du Théorème 2.2.1.

Preuve du Théorème 2.2.1 Soit θ_0 le vrai paramètre du modèle HGARCH(p, d, q) défini en (2.5). En utilisant la relation (2.8), nous obtenons

$$\begin{aligned} M(\theta) &= \mathbf{E}(\rho(x_t - \log h_t(\theta))) = \mathbf{E}(\rho(\log(y_t^2) - \log h_t(\theta))) \\ &= \mathbf{E}(\rho(w_t + \log h_t(\theta_0) - \log h_t(\theta))) \\ &= \mathbf{E}(\rho(w_t - \log(\frac{h_t(\theta)}{h_t(\theta_0)}))). \end{aligned}$$

A partir des relations (2.6) – (2.7), nous obtenons $w_t = x_t - \log h_t(\theta_0) = \log(\varepsilon_t^2)$ une suite de variables aléatoires i.i.d de densité g .

Par l'hypothèse (\mathbf{A}_4), la quantité $\mathbf{E}(\rho(w_t - u))$ a un unique minimum en u_0 . Alors le minimum de $M(\theta) = \mathbf{E}(\rho(w_t - \log(\frac{h_t(\theta)}{h_t(\theta_0)})))$ noté $M(\ddot{\theta})$ est atteint lorsque $\log(\frac{h_t(\ddot{\theta})}{h_t(\theta_0)}) = u_0$. Ce qui signifie que $M(\theta)$ atteint son minimum $M(\ddot{\theta})$ quand

$$h_t(\ddot{\theta}) = e^{u_0} h_t(\theta_0) \quad p.s. \quad (2.29)$$

D'après (2.5) et (2.3), nous obtenons

$$h_t(\ddot{\theta}) = e^{u_0} \frac{\gamma_0}{1 - \beta_0(1)} + e^{u_0} \omega \frac{1 - \beta_0(L) - \delta(L)(1 - L)^{d_0}}{1 - \beta_0(L)} y_t^2. \quad (2.30)$$

Ce qui sous-entend que

$$\ddot{\theta} = (e^{u_0} \gamma_0, \beta_{0_1}, \dots, \beta_{0_p}, \delta_{0_1}, \dots, \delta_{0_q}, e^{u_0} \omega_0, d_0).$$

Montrons l'unicité de $\ddot{\theta}$.

Considérons $\bar{\theta}$, un vecteur de paramètres différent de $\ddot{\theta}$ tel que $M(\bar{\theta})$ soit le minimum de $M(\theta)$. Puisque $M(\bar{\theta})$ est le minimum de $M(\theta)$, alors on a également

$$h_t(\bar{\theta}) = e^{u_0} h_t(\theta_0) \quad p.s. \quad (2.31)$$

Ainsi, nous obtenons

$$h_t(\bar{\theta}) = h_t(\ddot{\theta}) \quad p.s. \quad (2.32)$$

A partir des relations (2.29) et (2.31), il existe U et V , deux voisinages de θ_0 tels que $\ddot{\theta} \in U$ et $\bar{\theta} \in V$. De plus, en utilisant la relation (2.32), nous avons

$$\log(h_t(\bar{\theta})) - \log(h_t(\ddot{\theta})) = 0. \quad (2.33)$$

Par application du théorème de la valeur moyenne combiné à (2.33), il existe $\theta_1 \in U \cap V$ tel que

$$\begin{aligned} \log(h_t(\bar{\theta})) - \log(h_t(\check{\theta})) &= (\bar{\theta} - \check{\theta})' \frac{\nabla h_t(\theta_1)}{h_t(\theta_1)} = 0 \\ &= (\bar{\theta} - \check{\theta})' \nabla h_t(\theta_1) = 0. \end{aligned} \quad (2.34)$$

Ce qui implique également par transposition que

$$\nabla' h_t(\theta_1)(\bar{\theta} - \check{\theta}) = 0 \quad (2.35)$$

En utilisant le resultat dans Straumann [61] (section 1.7.1 du chapitre 1) et d'après l'hypothèse (\mathbf{A}_7) , la relation (2.35) conduit à

$$(\bar{\theta} - \check{\theta}) = 0,$$

d'où l'unicité. De plus, sous les conditions de la Remarque 2.2.2, vérifiées à partir de l'hypothèse (\mathbf{A}_0) , $\check{\theta}$ est égal à θ_0 .

□

2.2.3 Convergence du M-estimateur

Le théorème suivant donne les résultats de la consistance de l'estimateur $\hat{\theta}_n$ défini en (2.18).

Théorème 2.2.2. *Sous les hypothèses (\mathbf{A}_0) - (\mathbf{A}_4) et (\mathbf{A}_7) avec $u_0 = 0$, le M-estimateur $\hat{\theta}_n$, défini en (2.18) converge presque sûrement vers θ_0 quand n tend vers ∞ , c'est-à-dire*

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0 \quad p.s.$$

Pour établir le Théorème 2.2.2, nous aurons besoin de certains lemmes que nous énoncerons au paragraphe suivant. Les preuves de ces lemmes sont inspirées en partie des travaux de Robinson et Zaffaroni [58]. Le lemme 2.2.1 est un recueil de résultats intermédiaires qui permettra d'aborder avec plus d'aisance les démonstrations à venir.

Lemme 2.2.1. *Soit y_t un processus HGARCH(p, d, q) strictement stationnaire et ergodique satisfaisant la relation (2.1). Soit $h_t(\theta)$ la variance conditionnelle du processus y_t telle que définie en (2.3) et $\tilde{h}_t(\theta)$ la variance conditionnelle du processus y_t avec les valeurs initiales telles que définies en (2.16). Supposons que les hypothèses (\mathbf{A}_0) , (\mathbf{A}_2) et (\mathbf{A}_6) sont vérifiées, alors*

1. il existe un réel positif u tel que

$$\sup_{\theta \in \Theta} |h_t(\theta) - \tilde{h}_t(\theta)| \leq u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \quad \text{avec} \quad d_m = \inf_{\theta \in \Theta} d(\theta),$$

2. il existe un voisinage U de θ_0 et $\eta < 1 - \frac{1}{1+\frac{d_m}{2}}$ tels que

$$\begin{aligned} \mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta)\| &< \infty \\ \mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta) \nabla' \log h_t(\theta)\| &\leq \mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|^2 < \infty, \end{aligned}$$

3. il existe un voisinage U de θ_0 , $u, K > 0$ et $\eta < 1 - \frac{1}{1+\frac{d_m}{2}}$ tels que

$$\begin{aligned} \sup_{\theta \in U} \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| &\leq u \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \\ &+ \frac{Ku}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|, \\ \text{avec} \quad \gamma_m &= \inf_{\theta \in \Theta} \gamma, \end{aligned}$$

4. il existe un voisinage U de θ_0 , $u, K > 0$ et $\eta < 1 - \frac{1}{1+\frac{d_m}{2}}$ tels que

$$\begin{aligned} \mathbf{E} \sup_{\theta \in U} \|\nabla^2 h_t(\theta)\| &< \infty, \quad \mathbf{E}(\sup_{\theta \in U} \|\nabla^2 h_t(\theta)\|^2) < \infty \quad \text{et} \\ \sup_{\theta \in U} \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| &\leq K \left(\sum_{k=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-k}^4 \right)^{1/2} \\ &+ u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \sup_{\theta \in U} \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} \right\|. \end{aligned}$$

Preuve

1. En utilisant les expressions de la variance conditionnelle $h_t(\theta)$ et de la variance conditionnelle calculée avec les données initiales $\tilde{h}_t(\theta)$ définies respectivement en (2.3) et (2.16), nous obtenons :

$$0 \leq h_t(\theta) - \tilde{h}_t(\theta) = \sum_{j=1}^{\infty} b_j(\theta) y_{t-j}^2 - \sum_{j=1}^{t-1} b_j(\theta) y_{t-j}^2 = \sum_{j=t}^{\infty} b_j(\theta) y_{t-j}^2.$$

Dans la relation (2.3), nous avons $b_j(\theta) = O(j^{-1-d})$. Ce qui traduit le fait que pour $j \geq 1$, il existe une constante réelle $u > 0$ telle que $b_j(\theta) \leq u j^{-1-d}$. Par suite, nous obtenons,

$$\text{pour tout} \quad \theta \in \Theta, \quad |h_t(\theta) - \tilde{h}_t(\theta)| \leq u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2. \quad (2.36)$$

En effet, en considérant la fonction g_1 définie par

$$\begin{aligned} g_1 : \Theta &\rightarrow]0, 1[\\ \theta &\mapsto g_1(\theta) = d, \end{aligned}$$

nous obtenons une fonction g_1 qui représente la projection suivant la dernière composante (d) de θ défini en (2.11). Par continuité des projections suivant les différentes composantes et sous l'hypothèse (\mathbf{A}_2) , g_1 est bornée et atteint ses bornes. Ce qui justifie l'existence de

$$d_m = \inf_{\theta \in \Theta} g_1(\theta) = \inf_{\theta \in \Theta} d. \quad (2.37)$$

□

2. La démonstration de ce résultat se situe en annexe B
3. Dans cette partie du travail, l'objectif est de montrer que la quantité $\|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\|$ est majorée par le reste d'une série convergente presque sûrement, pour tout $\theta \in U$. Par calcul du gradient, nous obtenons

$$\|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| = \left\| \frac{\nabla h_t(\theta)}{h_t(\theta)} - \frac{\nabla \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\|.$$

En appliquant les mêmes techniques de calcul que dans Berkes et al. [6], nous aboutissons à l'inégalité suivante

$$\begin{aligned} \left\| \frac{\nabla h_t(\theta)}{h_t(\theta)} - \frac{\nabla \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| &= \left\| \frac{\nabla h_t(\theta) \tilde{h}_t(\theta) - \nabla \tilde{h}_t(\theta) h_t(\theta)}{h_t(\theta) \tilde{h}_t(\theta)} \right\| \\ &= \frac{1}{\tilde{h}_t(\theta)} \left\| \frac{\nabla h_t(\theta) \tilde{h}_t(\theta)}{h_t(\theta)} - \nabla \tilde{h}_t(\theta) \right\| \\ &= \frac{1}{\tilde{h}_t(\theta)} \left\| \frac{\nabla h_t(\theta) \tilde{h}_t(\theta)}{h_t(\theta)} - \nabla h_t(\theta) + \nabla h_t(\theta) - \nabla \tilde{h}_t(\theta) \right\| \\ &\leq \frac{1}{\tilde{h}_t(\theta)} \left\| \frac{\nabla h_t(\theta) \tilde{h}_t(\theta)}{h_t(\theta)} - \nabla h_t(\theta) \right\| + \frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| \\ &\leq \frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| + \frac{1}{\tilde{h}_t(\theta)} \left\| \frac{\nabla h_t(\theta) (h_t(\theta) - \tilde{h}_t(\theta))}{h_t(\theta)} \right\| \\ &\leq \frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| \\ &\quad + \frac{1}{\tilde{h}_t(\theta)} \|\nabla \log h_t(\theta)\| |h_t(\theta) - \tilde{h}_t(\theta)|. \end{aligned} \quad (2.38)$$

Traisons d'abord le premier terme du second membre de (2.38), c'est-à-dire $\frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\|$. A partir de (B.6) et avec (B.13), nous avons

$$\nabla h_t(\theta) = \left(\frac{1}{1 - \beta(1)}, \sum_{j=1}^{\infty} b_{j,k}^{(1)}(\theta) y_{t-j}^2 \right) \quad \text{avec } k = 1, \dots, p + q + 2. \quad (2.39)$$

De manière similaire, nous définissons le gradient de $\tilde{h}_t(\theta)$ défini en (2.16) par :

$$\nabla \tilde{h}_t(\theta) = \left(\frac{1}{1 - \beta(1)}, \sum_{j=1}^{t-1} b_{j,k}^{(1)}(\theta) y_{t-j}^2 \right) \quad \text{avec } k = 1, \dots, p + q + 2. \quad (2.40)$$

A partir de (2.39) et (2.40), nous avons pour tout $k = 1, \dots, p + q + 2$,

$$\|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| = \left\| \sum_{j=t}^{\infty} b_{j,k}^{(1)}(\theta) y_{t-j}^2 \right\| \leq \sum_{k=1}^{p+q+2} \sum_{j=t}^{\infty} |b_{j,k}^{(1)}(\theta) y_{t-j}^2|.$$

En utilisant également les mêmes arguments que dans (B.13)–(B.16), nous avons

$$\frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| \leq \frac{Ku(p + q + 2)}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2. \quad (2.41)$$

En ce qui concerne le second terme de (2.38), par le résultat du lemme 2.2.1.(1), nous avons

$$\frac{1}{\tilde{h}_t(\theta)} \|\nabla \log h_t(\theta)\| |h_t(\theta) - \tilde{h}_t(\theta)| \leq \frac{Ku}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|. \quad (2.42)$$

A partir de l'hypothèse (\mathbf{A}_2) qui garantit l'existence de $\mathbf{E}(y_t^4)$, combinée au résultat du lemme 2.2.1.(2) qui garantit l'existence de $\mathbb{E}(\sup_{\theta \in U} \|\nabla \log h_t(\theta)\|^2)$, nous pouvons appliquer l'inégalité de Hölder afin d'obtenir

$$\begin{aligned} \sum_{j=1}^{\infty} j^{-(1+d_m)} \mathbb{E}(y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|) &\leq \sum_{j=1}^{\infty} j^{-(1+d_m)} \mathbb{E}(y_{t-j}^4)^{1/2} \times \\ &\quad \mathbb{E}(\sup_{\theta \in U} \|\nabla \log h_t(\theta)\|^2)^{1/2}. \end{aligned}$$

Par suite, la série précédente converge selon le critère de Riemann car $1 + d_m$ est strictement supérieur à 1.

Ce qui nous permet d'affirmer par le Théorème de convergence monotone et le critère d'intégrabilité (section 1.7 du chapitre 1) que

$$\frac{Ku}{\gamma_m} \sum_{j=1}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\| < \infty \quad \text{p.s.}$$

En appliquant les résultats (2.41)–(2.42) à la relation (2.38), nous montrons le résultat.

4. La démonstration de ce résultat se situe en annexe B □

Les différents lemmes qui suivront sont inspirés des travaux de Muler et Yohai [55] pour les modèles GARCH. Contrairement à Muler et Yohai [55] qui travaillent avec la variance conditionnelle exprimée sous la forme de série de Volterra, dans notre étude, nous travaillons avec la variance conditionnelle exprimée sous forme de somme de série telle que définie en (2.3) comme dans Robinson et zaffaroni [58]. De plus, les différents lemmes que nous proposons, prennent en compte le caractère de longue mémoire propre au modèle GARCH hyperbolique.

Le lemme suivant est établi pour montrer que sous certaines conditions, la fonction $M_n(\theta)$ théorique peut être approximée par la fonction $\tilde{M}_n(\theta)$ qu'on peut évaluer numériquement.

Lemme 2.2.2. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_3) sont vérifiées, alors $\tilde{M}_n(\theta)$ définie en (2.17) est une bonne approximation de $M_n(\theta)$ définie en (2.12). Ce qui se traduit par*

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \xrightarrow{n \rightarrow \infty} 0 \quad p.s.$$

Preuve En utilisant (2.12) et (2.17), nous obtenons

$$|\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{1}{n-p} \sum_{t=p+1}^n |\rho(x_t - \log \tilde{h}_t(\theta)) - \rho(x_t - \log h_t(\theta))|. \quad (2.43)$$

D'après l'hypothèse (\mathbf{A}_3) , la fonction ρ a sa dérivée première bornée, donc elle est lipschitzienne. Cela entraîne qu'il existe un réel $K > 0$ tel que

$$|\rho(x_t - \log \tilde{h}_t(\theta)) - \rho(x_t - \log h_t(\theta))| \leq K |\log h_t(\theta) - \log \tilde{h}_t(\theta)|. \quad (2.44)$$

En appliquant au deuxième membre de (2.44) le résultat suivant

$$\left| \log \frac{x}{y} \right| \leq \frac{|x - y|}{\min(x, y)}, \quad x, y > 0,$$

nous obtenons

$$\begin{aligned} |\log h_t(\theta) - \log \tilde{h}_t(\theta)| &\leq \frac{|\tilde{h}_t(\theta) - h_t(\theta)|}{\min(\tilde{h}_t(\theta), h_t(\theta))} \quad \text{avec} & (2.45) \\ \frac{1}{\min(\tilde{h}_t(\theta), h_t(\theta))} &< \frac{1}{\gamma}. \end{aligned}$$

En substituant (2.45) dans (2.44), il en ressort que

$$|\rho(x_t - \log \tilde{h}_t(\theta)) - \rho(x_t - \log h_t(\theta))| \leq K \frac{1}{\gamma} |h_t(\theta) - \tilde{h}_t(\theta)|.$$

Par suite (2.43) devient

$$|\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{K}{\gamma} \frac{1}{n-p} \sum_{t=p+1}^n |h_t(\theta) - \tilde{h}_t(\theta)|. \quad (2.46)$$

En appliquant au second membre de (2.46) le lemme 2.2.1.(1), nous avons

$$|\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{Ku}{(n-p)\gamma_m} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2. \quad (2.47)$$

En passant au sup, nous obtenons

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{Ku}{(n-p)\gamma_m} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2. \quad (2.48)$$

A ce stade du travail, nous nous intéressons au second membre de (2.48). L'objectif est de montrer que cette quantité tend vers 0 quand n tend vers l'infini. Pour ce faire, nous considérons la série des espérances suivante :

$$\begin{aligned} \sum_{t \geq 1} \frac{1}{t} \mathbb{E} \left(\sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \right) &= \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-1-d_m} \mathbb{E}(y_{t-j}^2) \\ &= \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}} \frac{1}{j^{\frac{d_m}{2}}} \mathbb{E}(y_{t-j}^2) \\ &\leq \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}} \frac{1}{t^{\frac{d_m}{2}}} \mathbb{E}(y_{t-j}^2) \\ &\leq \sum_{t \geq 1} \frac{1}{t^{1+\frac{d_m}{2}}} \sum_{j=1}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}} \mathbb{E}(y_{t-j}^2). \end{aligned}$$

Comme dans Koulikov [42], en utilisant l'hypothèse (\mathbf{A}_1) qui assure la stricte stationnarité du processus y_t , il existe un $K > 0$ tel que

$$\sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-1-d_m} \mathbb{E}(y_{t-j}^2) \leq K \sum_{t \geq 1} \frac{1}{t^{1+\frac{d_m}{2}}} \sum_{j=1}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}}.$$

En remarquant que la série $\sum_{t \geq 1} \frac{1}{t^{1+\frac{d_m}{2}}} \sum_{j=1}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}}$ converge au sens de Riemann, par le Théorème de convergence monotone et le critère d'intégrabilité (section 1.7 du chapitre 1), nous avons

$$\sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 < \infty \quad p.s. \quad (2.49)$$

Par suite, nous appliquons le lemme de Kronecker (section 1.7 du chapitre 1) à (2.49) et obtenons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 = 0 \quad p.s. \quad (2.50)$$

Ce qui conduit au fait que, d'après (2.48),

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \xrightarrow{n \rightarrow \infty} 0, \quad p.s.$$

□

Le lemme 2.2.3 suivant a pour but de montrer que la suite de fonction $M_n(\theta)$ converge uniformément presque sûrement vers $M(\theta)$ quand n tend vers l'infini.

Lemme 2.2.3. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_3) sont vérifiées, alors nous avons*

$$\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow{n \rightarrow \infty} 0, \quad p.s. \quad \text{avec} \quad M(\theta) = \mathbf{E}(\rho(x_t - \log h_t(\theta))).$$

Preuve D'après l'hypothèse (\mathbf{A}_3) , la fonction ρ est bornée. Cela conduit à

$$\mathbf{E}(|\rho(x_t - \log h_t(\theta))|) < \infty. \quad (2.51)$$

De plus, d'après l'hypothèse (\mathbf{A}_1) , x_t défini en (2.6) est la composée d'un processus strictement stationnaire ergodique par une application mesurable. Ce qui fait également de x_t un processus strictement stationnaire ergodique d'après le Théorème 3.5.8 dans Stout [60] (section 1.7 du chapitre 1). Pour ces mêmes raisons, $\rho(x_t - \log h_t(\theta))$ est aussi un processus strictement stationnaire et ergodique. Par ailleurs, à partir de (2.51), nous appliquons le Théorème ergodique dans Straumann [61] (section 1.7 du chapitre 1) pour aboutir à

$$M_n(\theta) \xrightarrow{n \rightarrow \infty} \mathbf{E}(\rho(x_t - \log h_t(\theta))) \quad p.s. \quad (2.52)$$

Pour prouver le lemme 2.2.3, il suffit de montrer que la convergence dans (2.52) est uniforme.

L'objectif de la partie suivante est de montrer que la suite de fonction $M_n(\theta)$ est équi-continue. Pour cela, en utilisant (2.12), nous avons pour tous θ_1 et θ_2 éléments de Θ , avec $\theta_1 \neq \theta_2$,

$$|M_n(\theta_1) - M_n(\theta_2)| \leq \frac{1}{n-p} \sum_{t=p+1}^n |\log h_t(\theta_2) - \log h_t(\theta_1)|. \quad (2.53)$$

Par suite, en multipliant les deux termes de (2.53) par $\frac{1}{\|\theta_1 - \theta_2\|}$, nous obtenons

$$\sup_{\theta_1, \theta_2 \in \Theta} |M_n(\theta_1) - M_n(\theta_2)| \frac{1}{\|\theta_1 - \theta_2\|} \leq \frac{1}{n-p} \sum_{t=p+1}^n \nu_t, \quad (2.54)$$

avec

$$\nu_t = \sup_{\theta_1, \theta_2 \in \Theta} \frac{1}{\|\theta_1 - \theta_2\|} |\log h_t(\theta_2) - \log h_t(\theta_1)|.$$

Pour la suite du travail, nous emploierons des techniques similaires à celles utilisées dans Berkes et al. [6]. Ainsi, par le théorème de la valeur moyenne, il existe un ψ élément de Θ vérifiant

$$\|\psi - \theta_1\| \leq \|\theta_1 - \theta_2\| \quad \text{et} \quad \|\psi - \theta_2\| \leq \|\theta_1 - \theta_2\|,$$

tel que

$$|\log h_t(\theta_2) - \log h_t(\theta_1)| \leq \|\theta_1 - \theta_2\| \sup_{\psi \in \Theta} \|\nabla \log h_t(\psi)\|, \quad \theta_1, \theta_2 \in \Theta.$$

Cela entraîne que

$$\frac{1}{\|\theta_1 - \theta_2\|} |\log h_t(\theta_2) - \log h_t(\theta_1)| \leq \sup_{\psi \in \Theta} \|\nabla \log h_t(\psi)\|, \quad \theta_1, \theta_2 \in \Theta. \quad (2.55)$$

Par (2.55) et le lemme 2.2.1(2), nous aboutissons à

$$\mathbf{E}(\nu_t) < \infty \quad \text{c'est-à-dire} \quad \nu_t < \infty \quad p.s. \quad (2.56)$$

De plus, par le Théorème 3.5.8 dans Stout [60] (section 1.7 du chapitre 1), ν_t est un processus strictement stationnaire et ergodique. En utilisant le Théorème ergodique dans Straumann [61] (section 1.7 du chapitre 1), nous obtenons

$$\frac{1}{n-p} \sum_{t=p+1}^n \nu_t \rightarrow \mathbf{E}(\nu_t), \quad p.s., \quad n \rightarrow \infty.$$

Par ailleurs, en se référant à (2.56), nous arrivons à

$$\frac{1}{n-p} \sum_{t=p+1}^n \nu_t = O(1), \quad p.s., \quad n \rightarrow \infty. \quad (2.57)$$

En utilisant ce résultat dans (2.54), nous pouvons dire que la suite de fonction $M_n(\theta)$ est équicontinue presque sûrement. D'après l'hypothèse (\mathbf{A}_1) par laquelle Θ est un compact et (2.52), nous avons M_n une suite de fonction équicontinue qui converge presque sûrement sur un compact. Ce qui justifie la convergence uniforme de la suite M_n qui se traduit par :

$$\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow{n \rightarrow \infty} 0, \quad p.s. \quad \square$$

Nous pouvons à présent passer à la preuve du Théorème 2.2.2.

Preuve du Théorème 2.2.2

Nous avons par l'inégalité triangulaire, la relation suivante

$$|\tilde{M}_n(\theta) - M(\theta)| \leq |\tilde{M}_n(\theta) - M_n(\theta)| + |M_n(\theta) - M(\theta)|$$

Par passage au sup, nous obtenons

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \leq \sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| + \sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)|$$

A partir des Lemme 2.2.2 et Lemme 2.2.3, nous parvenons à montrer que

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0, \quad p.s. \quad (2.58)$$

Ce qui traduit le fait que

$$\mathbf{P}\left(\left\{\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0\right\}\right) = 1. \quad (2.59)$$

L'objectif de cette démonstration est d'aboutir à

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0, \quad p.s.,$$

c'est-à-dire

$$\mathbf{P}(\{\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0\}) = 1.$$

Pour ce faire, il suffit de montrer l'inclusion suivante :

$$\left\{\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0\right\} \subset \{\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0\}.$$

Notons $B = \left\{\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0\right\}$, l'objectif est de montrer que

$$B \subset \{\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0\}.$$

Procédons par l'absurde en supposant que $B \not\subset \{\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0\}$, c'est-à-dire il existe une sous suite $\hat{\theta}_{n_k}$ de B telle que

$$\hat{\theta}_{n_k} \xrightarrow[n \rightarrow \infty]{} \theta_1 \quad \text{avec} \quad \theta_1 \neq \theta_0. \quad (2.60)$$

La fonction M définie en (2.2.3) est continue sur Θ . En effet, \tilde{M}_n définie en (2.17) est une suite de fonction continue sur Θ car le fonction ρ est continue. D'après (2.58), la suite de fonction \tilde{M}_n converge uniformément sur Θ presque sûrement vers la fonction

M . Ainsi M est la limite uniforme de fonction continue donc elle est continue.

D'après le Théorème 2.2.1, M atteint son minimum en une unique valeur θ_0 et l'hypothèse $\theta_1 \neq \theta_0$ dans (2.60), nous avons

$$M(\theta_1) > M(\theta_0) \quad \text{soit} \quad M(\theta_1) - M(\theta_0) > 0. \quad (2.61)$$

La relation (2.61) permet de dire qu'il existe un voisinage V de θ_1 noté $V(\theta_1)$ et $\mu > 0$ tels que pour tout $\theta \in V(\theta_1)$, la relation (2.61) donne

$$M(\theta) > M(\theta_0) + \mu. \quad (2.62)$$

La définition de (2.60) conduit à , pour tout $\eta > 0$, il existe un rang $k_0 > 0$ tel que pour tout $k \geq k_0$, on ait $\|\hat{\theta}_{n_k} - \theta_1\|_{\Theta} < \eta$. En particulier pour tout $k \geq k_0$, nous avons également $\|\hat{\theta}_{n_k} - \theta_1\|_{\Theta} < \mu$. Ce qui entraîne que

$$\hat{\theta}_{n_k} \in V(\theta_1). \quad (2.63)$$

Par suite, puisque la sous suite $\hat{\theta}_{n_k} \in B$, nous avons

$$\sup_{\theta \in \Theta} |\tilde{M}_{n_k}(\theta) - M(\theta)| \xrightarrow{n \rightarrow \infty} 0,$$

qui se traduit par, pour tout $\eta > 0$, il existe un rang $k_0 > 0$ tel que pour tout $k \geq k_0$, on ait $\sup_{\theta \in \Theta} |\tilde{M}_{n_k}(\theta) - M(\theta)| < \eta$. En particulier, nous pouvons obtenir également, pour tout $k \geq k_0$

$$\sup_{\theta \in \Theta} |\tilde{M}_{n_k}(\theta) - M(\theta)| < \frac{\mu}{2} \quad (2.64)$$

En outre, nous avons l'égalité suivante :

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) = \tilde{M}_{n_k}(\hat{\theta}_{n_k}) - M(\hat{\theta}_{n_k}) + M(\hat{\theta}_{n_k}). \quad (2.65)$$

D'après (2.63) et (2.64), nous avons pour tout $k \geq k_0$ et pour $\hat{\theta}_{n_k} \in V(\theta_1)$

$$-\frac{\mu}{2} < \tilde{M}_{n_k}(\hat{\theta}_{n_k}) - M(\hat{\theta}_{n_k}) < \frac{\mu}{2}. \quad (2.66)$$

En substituant (2.66) dans (2.65), la relation (2.65) devient

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) > -\frac{\mu}{2} + M(\hat{\theta}_{n_k}). \quad (2.67)$$

Comme pour tout $k \geq k_0$, $\hat{\theta}_{n_k} \in V(\theta_1)$ alors d'après (2.62), nous obtenons

$$M(\hat{\theta}_{n_k}) > M(\theta_0) + \mu. \quad (2.68)$$

En remplaçant (2.68) dans (2.67), nous aboutissons à

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) > M(\theta_0) + \frac{\mu}{2}. \quad (2.69)$$

Par définition $\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{M}_n(\theta)$, ainsi nous notons

$$\hat{\theta}_{n_k} = \arg \min_{\theta \in \Theta} \tilde{M}_{n_k}(\theta).$$

Cela conduit à

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) \leq \tilde{M}_{n_k}(\theta) \quad \forall \theta \in \Theta,$$

en particulier

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) \leq \tilde{M}_{n_k}(\theta_0). \quad (2.70)$$

En utilisant la relation (2.64), nous obtenons pour tout $k \geq k_0$

$$\begin{aligned} |\tilde{M}_{n_k}(\theta_0) - M(\theta_0)| &< \frac{\mu}{2} \\ \tilde{M}_{n_k}(\theta_0) &< M(\theta_0) + \frac{\mu}{2}. \end{aligned} \quad (2.71)$$

En remplaçant (2.71) dans (2.70), nous aboutissons à

$$\tilde{M}_{n_k}(\hat{\theta}_{n_k}) < M(\theta_0) + \frac{\mu}{2},$$

ce qui est contraire à (2.69). Par conséquent $B \subset \{\hat{\theta}_n \xrightarrow{n \rightarrow \infty} \theta_0\}$.

En passant aux probabilités de ces événements avec (2.81), nous obtenons bien

$$\mathbf{P}(\{\hat{\theta}_n \xrightarrow{n \rightarrow \infty} \theta_0\}) = 1.$$

□

2.2.4 Normalité asymptotique du M-estimateur

Nous énonçons par le théorème suivant un résultat qui porte sur la distribution asymptotique du M-estimateur étudié dans ce travail.

Théorème 2.2.3. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_7) sont satisfaites avec $u_0 = 0$, alors*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}[0, A(\theta_0)^{-1}B(\theta_0)A(\theta_0)^{-1}], \quad n \rightarrow \infty,$$

avec

$$\begin{aligned} A(\theta_0) &= \mathbf{E}_{\theta_0}(\nabla^2 \rho(x_t - \log h_t(\theta))), & B(\theta_0) &= \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0)))D_0 \\ \text{et } D_0 &= \mathbf{E}[\nabla \log h_t(\theta_0)(\nabla \log h_t(\theta_0))']. \end{aligned}$$

Dans la partie suivante, nous donnerons sous forme de lemme, les résultats intermédiaires qui nous seront utiles pour la preuve du Théorème 2.2.3.

Lemme 2.2.4. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_7) sont vérifiées, alors*

$$\frac{1}{\sqrt{n-p}} \sum_{t=p+1}^n \nabla \rho(x_t - \log h_t(\theta_0)) \xrightarrow{\mathcal{L}} \mathbf{N}(0, \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0)))D_0),$$

avec $D_0 = \mathbf{E}[\nabla \log h_t(\theta_0)(\nabla \log h_t(\theta_0))']$.

Preuve Cette preuve a pour but de montrer que $b'\nabla \rho(x_t - \log h_t(\theta_0))$ avec b un vecteur non nul de \mathbb{R}^{p+q+3} est une différence de martingale, strictement stationnaire et ergodique. Elle s'inspire des travaux de Muler et Yohai [55]. En utilisant (2.12) et l'hypothèse (\mathbf{A}_3) , nous définissons par l'équation suivante le gradient de la fonction M_n en θ_0 , ainsi nous avons

$$\nabla M_n(\theta_0) = \frac{1}{n-p} \sum_{t=p+1}^n \nabla \rho(x_t - \log h_t(\theta_0)),$$

avec

$$\nabla \rho(x_t - \log h_t(\theta_0)) = \nabla \log h_t(\theta_0) \psi(x_t - \log h_t(\theta_0)). \quad (2.72)$$

D'après l'hypothèse (\mathbf{A}_1) , la suite de variables aléatoires $b'\nabla \rho(x_t - \log h_t(\theta_0))$ est la composée d'un processus strictement stationnaire ergodique par une application continue. Ce qui fait que $b'\nabla \rho(x_t - \log h_t(\theta_0))$ est également un processus strictement stationnaire ergodique d'après le Théorème 3.5.8 dans Stout [60] (section 1.7 du chapitre 1).

Afin de prouver que $b'\nabla \rho(x_t - \log h_t(\theta_0))$ est une différence de martingale, il suffit de montrer que

$$\mathbf{E}(b'\nabla \rho(x_t - \log h_t(\theta_0)) | \mathcal{F}_{t-1}) = 0.$$

En effet, pour tout vecteur b non nul, élément de \mathbb{R}^{p+q+3} , la relation (2.72) combinée à (2.8), nous permet d'obtenir

$$\mathbf{E}(b'\nabla \rho(x_t - \log h_t(\theta_0)) | \mathcal{F}_{t-1}) = \mathbf{E}(\psi(w_t) b' \nabla \log h_t(\theta_0) | \mathcal{F}_{t-1}). \quad (2.73)$$

Par construction en (2.1), nous avons $h_t(\theta)$ qui est une fonction \mathcal{F}_{t-1} -mesurable. Ce qui entraîne par application des propriétés de l'espérance conditionnelle (section 1.2.1

chapitre 1) avec (2.7) que

$$\begin{aligned}\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))|\mathcal{F}_{t-1}) &= b'\nabla \log h_t(\theta_0)\mathbf{E}(\psi(w_t)|\mathcal{F}_{t-1}), \\ &= b'\nabla \log h_t(\theta_0)\mathbf{E}(\psi(\log \varepsilon_t^2)|\mathcal{F}_{t-1}).\end{aligned}$$

Par suite, puisque par définition en (2.1), le processus ε_t est indépendant de \mathcal{F}_{t-1} , nous obtenons

$$\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))|\mathcal{F}_{t-1}) = b'\nabla \log h_t(\theta_0)\mathbf{E}(\psi(w_t)).$$

De plus, en utilisant l'hypothèse (\mathbf{A}_5) dans laquelle $\mathbf{E}(\psi(w_t)) = 0$ avec $w_t = x_t - \log h_t(\theta_0)$, nous aboutissons finalement à

$$\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))|\mathcal{F}_{t-1}) = 0. \quad (2.74)$$

Ce qui nous permet de conclure que $b'\nabla\rho(x_t - \log h_t(\theta_0))$ est bien une différence de martingale. De plus, d'après (2.74) nous avons

$$\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))) = \mathbf{E}(\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))|\mathcal{F}_{t-1})) = 0.$$

Ainsi, le processus $b'\nabla\rho(x_t - \log h_t(\theta_0))$ est une différence de martingale strictement stationnaire, ergodique et centrée.

A ce stade, nous voulons appliquer le théorème central limite des martingales (section 1.7 du chapitre 1) au processus $b'\nabla\rho(x_t - \log h_t(\theta_0))$ afin de déterminer sa distribution limite. Pour ce faire, il faudrait s'assurer que :

$$\mathbf{E}([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2) < \infty.$$

En effet, par application de l'espérance conditionnelle, nous obtenons

$$\mathbf{E}([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2) = \mathbf{E}(\mathbf{E}(b'\nabla\rho(x_t - \log h_t(\theta_0))^2|\mathcal{F}_{t-1})). \quad (2.75)$$

En explicitant la quantité élevée au carré dans le second membre de (2.75), nous avons

$$\mathbf{E}([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2) = \mathbf{E}(\mathbf{E}[b'\nabla\rho(x_t - \log h_t(\theta_0))\nabla'\rho(x_t - \log h_t(\theta_0))b|\mathcal{F}_{t-1}]).$$

En appliquant (2.72) à l'équation précédente, nous aboutissons à

$$\mathbf{E}([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2) = \mathbf{E}(\mathbf{E}[\psi^2(w_t)b'\nabla \log h_t(\theta_0)\nabla' \log h_t(\theta_0)b|\mathcal{F}_{t-1}])).$$

Étant donné que $h_t(\theta)$ définie en (2.1) est une fonction \mathcal{F}_{t-1} - mesurable, nous obtenons

$$\mathbf{E}([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2) = \mathbf{E}(b'\nabla \log h_t(\theta_0)\nabla' \log h_t(\theta_0)b.\mathbf{E}[\psi^2(w_t)|\mathcal{F}_{t-1}])).$$

Le processus ε_t étant indépendant de \mathcal{F}_{t-1} , avec (2.7), il en ressort que :

$$\begin{aligned} \mathbf{E}\left([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2\right) &= \mathbf{E}\left(b'\nabla\log h_t(\theta_0)\nabla'\log h_t(\theta_0)b\mathbf{E}\left(\psi^2(w_t)\right)\right) \\ &= \mathbf{E}\left(b'\nabla\log h_t(\theta_0)\nabla'\log h_t(\theta_0)b\right)\mathbf{E}\left[\psi^2(w_t)\right] \end{aligned} \quad (2.76)$$

Par ailleurs, en utilisant le résultat Straumann [61] (section 1.7.1 chapitre 1) et l'hypothèse (\mathbf{A}_7) , nous montrons que $\mathbf{E}\left(b'\nabla\log h_t(\theta_0)\nabla'\log h_t(\theta_0)b\right) > 0$ c'est-à-dire D_0 est une matrice définie positive. Par suite, en utilisant l'hypothèse (\mathbf{A}_5) , nous pouvons affirmer d'après (2.76) que

$$\mathbf{E}\left([b'\nabla\rho(x_t - \log h_t(\theta_0))]^2\right) < \infty. \quad (2.77)$$

Ainsi, par application du théorème central limite des martingales (section 1.7 du chapitre 1), nous obtenons

$$\frac{1}{\sqrt{n-p}} \sum_{t=1+p}^n b'\nabla\rho(x_t - \log h_t(\theta_0)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbf{E}(\psi^2(w_t))b'D_0b\right). \quad (2.78)$$

En appliquant le critère de Cramer-Wold (section 1.7 du chapitre 1) à (2.78), nous obtenons le résultat escompté. \square

Lemme 2.2.5. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_7) sont vérifiées. Alors, il existe un voisinage U de θ_0 tel que*

1.

$$\sup_{\theta \in U} \|\nabla^2 M_n(\theta) - \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta)))\| \xrightarrow[n \rightarrow \infty]{} 0, \quad p.s.$$

2. $A(\theta_0)$ est une matrice définie positive donnée par

$$A(\theta_0) = \mathbf{E}_{\theta_0}(\nabla^2 \rho(x_t - \log h_t(\theta))) = \mathbf{E}(\psi'(w_t))D_0.$$

Preuve Le but de la preuve est de montrer que la suite de matrice $\nabla^2 M_n(\theta)$ converge uniformément presque sûrement vers une certaine quantité et que cette quantité en question, évaluée en θ_0 , est une matrice définie positive.

1. En utilisant (2.12) et (2.72), nous définissons la dérivée seconde de $M_n(\theta)$ par rapport au vecteur des paramètres. En réalité, $\nabla^2 M_n(\theta)$ représente la matrice hessienne de $M_n(\theta)$. Ainsi, nous avons

$$\nabla^2 M_n(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \nabla^2 \rho(x_t - \log h_t(\theta)),$$

avec

$$\begin{aligned} \nabla^2 \rho(x_t - \log h_t(\theta)) &= \left[\psi(x_t - \log h_t(\theta)) + \psi'(x_t - \log h_t(\theta)) \right] \nabla \log h_t(\theta) \times \\ &\quad (\nabla \log h_t(\theta))' - \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 h_t(\theta)}{h_t(\theta)}. \end{aligned} \quad (2.79)$$

En passant à la norme, nous obtenons

$$\begin{aligned} \|\nabla^2 \rho(x_t - \log h_t(\theta))\| &\leq \left\| \left[\psi(x_t - \log h_t(\theta)) + \psi'(x_t - \log h_t(\theta)) \right] \nabla \log h_t(\theta) \right. \\ &\quad \left. \times (\nabla \log h_t(\theta))' + \left\| \frac{1}{h_t(\theta)} \psi(x_t - \log h_t(\theta)) \nabla^2 h_t(\theta) \right\| \right\|. \end{aligned}$$

Ce qui conduit à

$$\begin{aligned} \|\nabla^2 \rho(x_t - \log h_t(\theta))\| &\leq |(\psi + \psi')(x_t - \log h_t(\theta))| \|\nabla \log h_t(\theta) (\nabla \log h_t(\theta))'\| \\ &\quad + |\psi(x_t - \log h_t(\theta))| \left\| \frac{1}{h_t(\theta)} \nabla^2 h_t(\theta) \right\|. \end{aligned} \quad (2.80)$$

D'après l'hypothèse (\mathbf{A}_3) , les fonctions ψ et ψ' , dérivées première et seconde respectives de la fonction ρ sont continues et bornées. D'où, il existe deux réel positifs K et K' tels que (2.80) donne

$$\|\nabla^2 \rho(x_t - \log h_t(\theta))\| \leq K \|\nabla \log h_t(\theta) (\nabla \log h_t(\theta))'\| + K' \left\| \frac{1}{h_t(\theta)} \nabla^2 h_t(\theta) \right\|. \quad (2.81)$$

De plus, en utilisant les résultats du lemme 2.2.1(2-4) qui traduisent que

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla \log h_t(\theta) (\nabla \log h_t(\theta))'\|) < \infty \quad \text{et} \quad \mathbf{E}(\sup_{\theta \in U} \|\nabla^2 h_t(\theta)\|) < \infty,$$

nous pouvons conclure à partir de (2.81) que

$$\mathbf{E}(\sup_{\theta \in \Theta} \|\nabla^2 \rho(x_t - \log h_t(\theta))\|) < \infty, \quad (2.82)$$

c'est-à-dire

$$\mathbf{E}(\|\nabla^2 \rho(x_t - \log h_t(\theta))\|_{\Theta}) < \infty. \quad (2.83)$$

La condition (2.83) permet d'utiliser la loi forte uniforme des grands nombres (section 1.1.6 du chapitre 1) à la quantité $\nabla^2 \rho(x_t - \log h_t(\theta))$. Ainsi, nous obtenons

$$\sup_{\theta \in \Theta} \left\| \frac{1}{n-p} \sum_{t=p+1}^n \nabla^2 \rho(x_t - \log h_t(\theta)) - \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta))) \right\| \xrightarrow[n \rightarrow \infty]{} 0, \quad p.s.,$$

d'où le résultat.

2. Dans ce paragraphe, l'objectif sera de montrer que $\mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta_0)))$ est une matrice définie positive. Considérons b un vecteur non nul de \mathbb{R}^{p+q+3} , en procédant comme dans (2.75)-(2.76), nous obtenons

$$\mathbf{E}_{\theta_0}(b' \nabla^2 \rho(x_t - \log h_t(\theta)) b) = \mathbf{E}(\psi'(x_t - \log h_t(\theta_0))) \mathbf{E}[b' \nabla \log h_t(\theta_0) (\nabla \log h_t(\theta_0))' b].$$

Ce qui nous amène à

$$\mathbf{E}_{\theta_0}(b' \nabla^2 \rho(x_t - \log h_t(\theta)) b) = \mathbf{E}(\psi'(x_t - \log h_t(\theta_0))) b' D_0 b$$

D'après l'hypothèse (\mathbf{A}_5) avec la condition $\mathbf{E}(\psi'(x_t - \log h_t(\theta_0))) > 0$ combinée au fait que D_0 soit une matrice définie positive, nous pouvons conclure que

$$\mathbf{E}_{\theta_0}(b' \nabla^2 \rho(x_t - \log h_t(\theta)) b) > 0.$$

Par conséquent, $A(\theta_0)$ est une matrice définie positive. \square

Lemme 2.2.6. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_7) sont vérifiées, alors*

$$\frac{1}{n-p} \sum_{t=p+1}^n \|\nabla \rho(x_t - \log h_t(\theta_0)) - \nabla \rho(x_t - \log \tilde{h}_t(\theta_0))\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s.$$

Preuve A partir de (2.72), nous avons

$$\|\nabla M_n(\theta_0) - \nabla \tilde{M}_n(\theta_0)\| \leq \frac{1}{n-p} \sum_{t=p+1}^n \|\nabla \rho(x_t - \log h_t(\theta_0)) - \nabla \rho(x_t - \log \tilde{h}_t(\theta_0))\|,$$

ce qui conduit à

$$\begin{aligned} \|\nabla M_n(\theta_0) - \nabla \tilde{M}_n(\theta_0)\| &\leq \frac{1}{n-p} \sum_{t=1}^n \|\psi(x_t - \log h_t(\theta_0)) \nabla \log h_t(\theta_0) \\ &\quad - \psi(x_t - \log \tilde{h}_t(\theta_0)) \nabla \log \tilde{h}_t(\theta_0)\|. \end{aligned} \quad (2.84)$$

En ajoutant et en retranchant la même quantité au second terme de (2.84), nous obtenons

$$\begin{aligned} \|\nabla M_n(\theta_0) - \nabla \tilde{M}_n(\theta_0)\| &\leq \frac{1}{n-p} \sum_{t=1}^n \|\psi(x_t - \log h_t(\theta_0)) \nabla \log h_t(\theta_0) \\ &\quad - \psi(x_t - \log h_t(\theta_0)) \nabla \log \tilde{h}_t(\theta_0) \\ &\quad + \psi(x_t - \log h_t(\theta_0)) \nabla \log \tilde{h}_t(\theta_0) \\ &\quad - \psi(x_t - \log \tilde{h}_t(\theta_0)) \nabla \log \tilde{h}_t(\theta_0)\|. \end{aligned} \quad (2.85)$$

Par soucis de concision, nous adoptons les notations suivantes pour la relation (2.85)

$$\|\nabla M_n(\theta_0) - \nabla \tilde{M}_n(\theta_0)\| \leq u_n + v_n,$$

avec

$$\begin{aligned} u_n &= \frac{1}{n-p} \sum_{t=p+1}^n \|\psi(x_t - \log h_t(\theta_0))(\nabla \log h_t(\theta_0) - \nabla \log \tilde{h}_t(\theta_0))\|, \\ v_n &= \frac{1}{n-p} \sum_{t=p+1}^n \|(\psi(x_t - \log h_t(\theta_0)) - \psi(x_t - \log \tilde{h}_t(\theta_0)))\nabla \log \tilde{h}_t(\theta_0)\|. \end{aligned}$$

L'objectif du paragraphe suivant est de montrer que les suites u_n et v_n convergent presque sûrement vers 0 quand n tend vers l'infini. Nous traitons dans un premier temps la convergence presque sûre de u_n vers 0. La suite u_n peut se réécrire de la manière suivante

$$u_n = \frac{1}{n-p} \sum_{t=p+1}^n |\psi(x_t - \log h_t(\theta_0))| \|\nabla \log h_t(\theta_0) - \nabla \log \tilde{h}_t(\theta_0)\|. \quad (2.86)$$

En appliquant le lemme 2.2.1(3) au second membre de (2.86), nous aboutissons à

$$\begin{aligned} u_n &\leq \frac{u}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} |\psi(x_t - \log h_t(\theta_0))| y_{t-j}^2 + \\ &\quad \frac{Ku}{\gamma_m} \frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\| |\psi(x_t - \log h_t(\theta_0))|. \end{aligned}$$

D'après l'hypothèse (\mathbf{A}_3) la fonction ψ est bornée, il existe un réel K_3 positif tel que la relation précédente donne

$$\begin{aligned} u_n &\leq \frac{uK_3}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 + \\ &\quad \frac{K.K_3u}{\gamma_m} \frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|. \end{aligned} \quad (2.87)$$

Nous nous intéressons au second membre de (2.87) et nous souhaitons montrer d'abord que son premier terme converge vers 0 presque sûrement quand n tend vers l'infini.

En effet, en utilisant les mêmes techniques que dans (B.14) – (B.16), nous obtenons

$$\begin{aligned} \sum_{t \geq 1} \frac{1}{t} \mathbf{E} \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 &= \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} \mathbf{E} y_{t-j}^2, \\ &\leq K \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} \frac{1}{t^{\frac{d_m}{2}(1-\eta)}} \frac{1}{j^{(1+\frac{d_m}{2})(1-\eta)}}, \\ &\leq K \sum_{t \geq 1} \frac{1}{t^{1+\frac{d_m}{2}(1-\eta)}} \sum_{j=1}^{\infty} \frac{1}{j^{(1+\frac{d_m}{2})(1-\eta)}}. \end{aligned} \quad (2.88)$$

D'après le critère de convergence des séries de Riemann, le second terme de (2.88) converge, lorsque $\eta < 1 - \frac{1}{1+\frac{d_m}{2}}$. De plus, par le Théorème de convergence monotone et

le critère d'intégrabilité (section 1.7 du chapitre 1), nous déduisons que

$$\sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 < \infty \quad \text{p.s.} \quad (2.89)$$

A partir du résultat (2.89), nous pouvons appliquer le lemme de Kronecker (section 1.7 du chapitre 1) comme dans la preuve du lemme 2.2.2 afin d'aboutir à

$$\frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (2.90)$$

Nous venons ainsi de montrer que le premier terme du second membre de (2.87) converge vers 0 presque sûrement quand n tend vers l'infini. A présent, intéressons nous au second terme du second membre de (2.87). Nous commençons d'abord par étudier la convergence de la série des espérances suivante. D'après le Théorème de convergence monotone (section 1.7 du chapitre 1), nous avons

$$\sum_{t \geq 1} \frac{1}{t} \mathbf{E} \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta_0)\| = \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_m)} \mathbf{E}(y_{t-j}^2 \|\nabla \log h_t(\theta_0)\|).$$

En procédant comme dans (2.42) dans le lemme 2.2.1(3), nous aboutissons à

$$\sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_m)} \mathbf{E}(y_{t-j}^2 \|\nabla \log h_t(\theta_0)\|) \leq K \sum_{t \geq 1} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_m)}.$$

En utilisant les mêmes arguments que dans (2.48)-(2.50) au Lemme 2.2.2, nous parvenons à

$$\frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta_0)\| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (2.91)$$

Ainsi, par les relations (2.90) et (2.91), nous obtenons

$$u_n \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (2.92)$$

Nous montrons également la convergence presque sûre de v_n vers 0 quand n tend vers l'infini. La suite v_n peut se réécrire de la manière suivante

$$v_n \leq \frac{1}{n-p} \sum_{t=p+1}^n |(\psi(x_t - \log h_t(\theta_0)) - \psi(x_t - \log \tilde{h}_t(\theta_0)))| \|\nabla \log \tilde{h}_t(\theta_0)\|.$$

D'après l'hypothèse (\mathbf{A}_3), il existe un réel K_4 positif tel que

$$v_n \leq \frac{K_4}{n-p} \sum_{t=p+1}^n \|\nabla \log \tilde{h}_t(\theta_0)\|.$$

En utilisant les résultats (B.5) et (B.17), nous obtenons

$$v_n \leq \frac{K_4 u}{\gamma_m} \frac{1}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2.$$

Par ailleurs, par les mêmes arguments que dans (2.87)-(2.90), nous obtenons

$$v_n \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s. \quad (2.93)$$

Par conséquent, en combinant (2.92) et (2.93), nous aboutissons au résultat du lemme. \square

Lemme 2.2.7. *Supposons que les hypothèses (\mathbf{A}_1) - (\mathbf{A}_7) sont vérifiées, alors il existe un voisinage U de θ_0 tel que*

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|\nabla^2 \rho(x_t - \log h_t(\theta)) - \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta))\| \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s.$$

Preuve Pour faciliter la compréhension de cette partie, nous procédons à quelques notations. Ainsi, nous posons

$$\nu(x_t - \log h_t(\theta)) = \psi(x_t - \log h_t(\theta)) + \psi'(x_t - \log h_t(\theta)), \quad (2.94)$$

et

$$P_t(\theta) = \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta), \quad (2.95)$$

également

$$a_t = \nabla^2 \rho(x_t - \log h_t(\theta)) - \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta)). \quad (2.96)$$

En utilisant les relations (2.79) et (2.94), nous avons

$$\begin{aligned} a_t &= \nu(x_t - \log h_t(\theta)) \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nu(x_t - \log \tilde{h}_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta) \\ &\quad - \frac{1}{h_t(\theta)} \psi(x_t - \log h_t(\theta)) \nabla^2 h_t(\theta) + \frac{1}{\tilde{h}_t(\theta)} \psi(x_t - \log \tilde{h}_t(\theta)) \nabla^2 \tilde{h}_t(\theta). \end{aligned} \quad (2.97)$$

A partir des relations (2.94)-(2.96), l'expression (2.97) se réécrit comme suit

$$a_t = L_t(\theta) - M_t(\theta) \quad (2.98)$$

avec

$$L_t(\theta) = \nu(x_t - \log h_t(\theta)) \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nu(x_t - \log \tilde{h}_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta),$$

et

$$M_t(\theta) = \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \psi(x_t - \log \tilde{h}_t(\theta)) \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)}. \quad (2.99)$$

L'objectif dans ce paragraphe sera de montrer les convergences suivantes

$$\sup_{\theta \in \tilde{U}} \frac{1}{n-p} \sum_{t=1}^n \|L_t(\theta)\| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.} \quad \text{et} \quad \sup_{\theta \in \tilde{U}} \frac{1}{n-p} \sum_{t=1}^n \|M_t(\theta)\| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.} \quad (2.100)$$

Traisons premièrement la convergence de la première quantité de (2.100) définie par $L_t(\theta)$. Pour cela, en ajoutant et retranchant une même quantité, nous réécrivons $L_t(\theta)$ de la façon suivante

$$\begin{aligned} L_t(\theta) &= \nu(x_t - \log h_t(\theta)) \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nu(x_t - \log h_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta) \\ &\quad + \nu(x_t - \log h_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta) - \nu(x_t - \log \tilde{h}_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta), \end{aligned}$$

ce qui conduit d'après (2.95) à

$$\begin{aligned} L_t(\theta) &= \nu(x_t - \log h_t(\theta)) P_t(\theta) + \\ &\quad [\nu(x_t - \log h_t(\theta)) - \nu(x_t - \log \tilde{h}_t(\theta))] \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta). \end{aligned} \quad (2.101)$$

Intéressons nous particulièrement au cas de $P_t(\theta)$ dans (2.101). En utilisant (2.95), nous avons

$$P_t(\theta) = \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta).$$

En ajoutant et retranchant à $P_t(\theta)$ la même quantité, nous obtenons

$$\begin{aligned} P_t(\theta) &= \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \nabla' \log h_t(\theta) + \nabla \log \tilde{h}_t(\theta) \nabla' \log h_t(\theta) \\ &\quad - \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta) \\ &= \left(\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \right) \nabla' \log h_t(\theta) \\ &\quad + \nabla \log \tilde{h}_t(\theta) \left(\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \right)'. \end{aligned} \quad (2.102)$$

En passant à la norme dans (2.102), nous obtenons

$$\|P_t(\theta)\| \leq \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right). \quad (2.103)$$

Dans le paragraphe suivant, nous étudierons de près le premier terme de $L_t(\theta)$ défini en (2.101). Ce qui reviendrait à étudier la convergence de

$$s_n = \frac{1}{n-p} \sum_{t=p+1}^n |\nu(x_t - \log h_t(\theta))| \|P_t(\theta)\|.$$

En utilisant (2.103), nous avons

$$s_n \leq \frac{1}{n-p} \sum_{t=p+1}^n |\nu(x_t - \log h_t(\theta))| \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| \times \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right).$$

Par application du lemme 2.2.1(3), nous obtenons

$$s_n \leq \frac{u}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 |\nu(x_t - \log h_t(\theta))| \times \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right) + \frac{K_3}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta)\| \times |\nu(x_t - \log h_t(\theta))| \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right).$$

D'après l'hypothèse (\mathbf{A}_3), il existe un réel K_5 positif tel que

$$s_n \leq \frac{K_5 u}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right) \quad (2.104) \\ + \frac{K_3 K_5}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta)\| \times \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right).$$

En appliquant à (2.104), les mêmes arguments que dans les preuves des lemmes 2.2.6 et 2.2.1(2), nous aboutissons à

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n |\nu(x_t - \log h_t(\theta))| \|P_t(\theta)\| \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s. \quad (2.105)$$

Considérons à présent le second terme de $L_t(\theta)$ défini en (2.101). A partir de (2.94), nous obtenons

$$b_t = |\nu(x_t - \log h_t(\theta)) - \nu(x_t - \log \tilde{h}_t(\theta))|, \quad (2.106) \\ \leq |\psi(x_t - \log h_t(\theta)) - \psi(x_t - \log \tilde{h}_t(\theta))| + |\psi'(x_t - \log h_t(\theta)) - \psi'(x_t - \log \tilde{h}_t(\theta))|.$$

La norme du second terme dans (2.101) a la forme suivante

$$b_t \|\nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta)\|. \quad (2.107)$$

D'après l'hypothèse (\mathbf{A}_3), b_t défini en (2.106) est borné alors il existe un réel K_6 positif tel qu'à partir de (2.107), nous obtenons

$$\frac{1}{n-p} \sum_{t=p+1}^n b_t \|\nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta)\| \leq \frac{K_6}{n-p} \sum_{t=p+1}^n \|\nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta)\|.$$

Par suite, en appliquant les résultats issus de la preuve du lemme 2.2.1(2), ainsi que les mêmes arguments dans le lemme 2.2.6, nous aboutissons à

$$\sup_{\theta \in \tilde{U}} \frac{1}{n-p} \sum_{t=p+1}^n b_t \|\nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s. \quad (2.108)$$

A partir des relations (2.101), (2.105) et (2.108), nous arrivons à

$$\sup_{\theta \in \tilde{U}} \frac{1}{n-p} \sum_{t=p+1}^n \|L_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s. \quad (2.109)$$

Dans la partie suivante, nous montrons la convergence de la seconde quantité définie en (2.100) c'est-à-dire

$$\frac{1}{n-p} \sum_{t=p+1}^n \|M_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s.$$

D'après (2.99), nous pouvons réécrire la suite de variables aléatoires $M_t(\theta)$ de la manière suivante

$$\begin{aligned} M_t(\theta) &= \frac{1}{h_t(\theta)} \psi(x_t - \log h_t(\theta)) \nabla^2 h_t(\theta) - \frac{1}{\tilde{h}_t(\theta)} \psi(x_t - \log \tilde{h}_t(\theta)) \nabla^2 \tilde{h}_t(\theta), \\ &= \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \\ &\quad + \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} - \psi(x_t - \log \tilde{h}_t(\theta)) \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)}, \\ &= [\psi(x_t - \log \tilde{h}_t(\theta)) - \psi(x_t - \log h_t(\theta))] \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \\ &\quad + \psi(x_t - \log h_t(\theta)) \left[\frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} - \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} \right]. \end{aligned}$$

Par passage à la norme, nous obtenons le résultat suivant

$$\begin{aligned} \|M_t(\theta)\| &\leq \|[\psi(x_t - \log h_t(\theta)) - \psi(x_t - \log \tilde{h}_t(\theta))]\| \left\| \frac{\nabla^2 \tilde{h}_t(\theta)}{h_t(\theta)} \right\| + \\ &\quad \|\psi(x_t - \log \tilde{h}_t(\theta))\| \left\| \left[\frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right] \right\|. \end{aligned}$$

De plus, d'après l'hypothèse (\mathbf{A}_3) , il existe un réel K_7 positif tel que nous ayons

$$\|M_t(\theta)\| \leq K_7 \left\| \frac{\nabla^2 \tilde{h}_t(\theta)}{h_t(\theta)} \right\| + K_7 \left\| \left[\frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right] \right\|.$$

Par suite, en appliquant les résultats du Lemme 2.2.1.(4) et les mêmes arguments que dans la preuve du lemme 2.2.6, nous obtenons

$$\sup_{\theta \in \tilde{U}} \frac{1}{n-p} \sum_{t=p+1}^n \|M_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s. \quad (2.110)$$

A partir de tous les résultats intermédiaires précédents, nous passons à présent à la preuve du lemme 2.2.7. D'après (2.98), nous avons

$$\|a_t\| \leq \|L_t(\theta)\| + \|M_t(\theta)\|.$$

Par suite, nous aboutissons à

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|a_t\| \leq \sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|L_t(\theta)\| + \sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|M_t(\theta)\|.$$

Par les relations (2.109) et (2.110), nous obtenons le résultat souhaité. \square

En combinant les lemmes 2.2.4 à 2.2.7, nous proposons dans le paragraphe suivant la preuve du théorème 2.2.3.

Preuve du Théorème 2.2.3

A partir du lemme 2.2.4 et du lemme 2.2.6, nous obtenons

$$\frac{1}{\sqrt{n-p}} \sum_{t=1+p}^n \nabla \rho(x_t - \log \tilde{h}_t(\theta_0)) \xrightarrow{L} \mathcal{N}\left(0, \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0)))D_0\right).$$

De même, en combinant le lemme 2.2.5(1) et le lemme 2.2.7, nous concluons qu'il existe un voisinage U de θ_0 tel que

$$\sup_{\theta \in U} \left\| \frac{1}{n-p} \sum_{t=p+1}^n \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta)) - \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta))) \right\| \xrightarrow{n \rightarrow \infty} 0 \quad p.s.$$

De plus, d'après le lemme 2.2.5(2), la matrice $A(\theta_0) = \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta_0)))$ est non singulière. En considérant que le Théorème 2.2.2 est vérifié, les hypothèses du Théorème 4.1.3 de Amémiya (section 1.7 du chapitre 1) sont vérifiées, d'où le résultat. \square

2.2.5 Cas de BM-estimateur

Il est à noter que les propriétés asymptotiques du M-estimateur défini en (2.18) et du BM-estimateur défini en (2.24) sont exactement les mêmes. Ce qui sous-entend que les résultats de la consistance et de la normalité asymptotique obtenus dans ce chapitre pour le M-estimateur sont aussi valables pour le BM-estimateur. En effet, en analysant la construction de ces deux estimateurs, nous nous apercevons qu'ils présentent la même structure de construction. La différence entre eux ne repose uniquement que sur la manière de traiter les observations dans l'échantillon. Le M-estimateur utilise un prédicteur de la variance conditionnelle non observée qui est obtenu à partir de

toutes les observations de l'échantillon sans qu'aucunes d'elles ne soient modifiées. Par contre, le BM-estimateur utilise un prédicteur de la variance conditionnelle qui est obtenu après quelques modifications de certaines observations dans l'échantillon (les valeurs aberrantes notamment). Cette différence d'ordre technique n'est donc pas de nature à remettre en question le fait que les résultats asymptotiques du M-estimateur s'appliquent au BM-estimateur.

Nous faisons remarquer que Lu et al. [49] dans leurs travaux ont également avancé ces mêmes arguments pour expliquer la différence entre leurs estimateurs.

Après avoir établi théoriquement les propriétés asymptotiques du M-estimateur et du BM-estimateur dans ce chapitre, il paraît important d'évaluer numériquement certaines de ces propriétés sur des échantillons de taille finie. De plus, une analyse comparative des performances de ces estimateurs serait d'une grande utilité.

Chapitre 3

Etude comparative des performances des estimateurs du modèle GARCH hyperbolique et applications aux données réelles

Ce chapitre consiste essentiellement à étudier les propriétés à échantillon fini du M-estimateur, du BM-estimateur et de l'estimateur du quasi maximum de vraisemblance à travers des simulations numériques. Une étude comparative des performances d'estimations et de prévisions des estimateurs en compétition est également réalisée d'une part sur des échantillons non contaminés par des outliers, d'autre part sur des échantillons contaminés par des outliers. A cet effet, selon Grubbs [35] et Fox [30], une outlier est une observation qui semble s'écarter de façon marquée (considérable) des autres observations de l'échantillon dans lequel elle se trouve. Elle peut être le résultat d'une erreur d'échantillonnage, d'une erreur de mesures issues d'instruments deffectueux, d'une erreur de calcul ou d'enregistrement d'une valeur numérique. Elle peut également être une manifestation extrême de la variabilité aléatoire inhérente aux données. Les outliers ne sont pas des valeurs prévisibles (qu'on peut estimer) contrairement aux valeurs extrêmes. En effet, les valeurs extrêmes d'une série statistique sont les valeurs minimum et maximum de cette série. Ces valeurs sont régies par des lois. De plus, il existe une théorie, appelée la théorie des valeurs extrêmes qui s'intéresse à l'étude du comportement des valeurs extrêmes issues d'un échantillon de variables aléatoires (voir Longin [47]).

L'étude comparative des performances des estimateurs présentée dans ce chapitre s'intéressera principalement aux observations comportant des outliers. Une étude sur les données réelles sera également abordée.

3.1 Simulations numériques

3.1.1 Cadre des simulations

Dans ce paragraphe, nous conduisons des études de simulations afin d'étudier le comportement des estimateurs définis au chapitre 2 pour les modèles GARCH hyperboliques. Il s'agit de l'estimateur du quasi maximum de vraisemblance (QMV), du M-estimateur (M_m) défini en (2.18) et du BM-estimateur défini en (2.24). Les propriétés asymptotiques de l'estimateur du QMV ont été étudiées par Li et al. [46]. Puisque le M-estimateur et le BM-estimateur sont basés sur une fonction ρ définie en (\mathbf{A}_3) dans le chapitre 2, il est nécessaire de choisir une fonction qui vérifie toutes les propriétés définies en (\mathbf{A}_3) pour construire numériquement ces estimateurs. Pour ce faire, nous utilisons la fonction proposée par Muler et Yohai [55] et reprise également par Lu et al. [49]. La fonction ρ considérée dans cette étude est donc définie par $\rho = m(-\log g_0)$, avec g_0 donné dans (2.10) et

$$m(x) = \begin{cases} x & \text{si } x \leq 4 \\ R(x) & \text{si } 4 < x \leq 4.30 \\ 4.15 & \text{si } x > 4.30, \end{cases}$$

où

$$R(x) = 0.5(x^4 - a^4) - \frac{2}{3}(2a + b)(x^3 - a^3) + \frac{(a^2 + 2ab)(x^2 - a^2)}{(b - a)^3} - \frac{(2ba^2(x - a))}{(b - a)^3} - \frac{(x - a)^3}{3(b - a)^2} + x,$$

avec $a = 4$, $b = 4.3$.

Ces choix de $R(x)$, a et b dans Muler et Yohai [55] assurent que m est continue. Par ailleurs, dans le cas particulier du BM-estimateur, il nous faut choisir une valeur pour le seuil l défini en (2.22). D'après les notations adoptées en (2.21) et (2.22), nous avons

$$s_l(u) = s_l\left(\frac{y_t^2}{\bar{h}_{t,l}(\theta)}\right), \quad \text{avec } u = \frac{y_t^2}{\bar{h}_{t,l}(\theta)}. \quad (3.1)$$

Pourtant, selon (2.1), le rapport du carré du processus par la variance conditionnelle donne le carré du bruit, en d'autres termes, on a

$$\frac{y_t^2}{\bar{h}_{t,l}(\theta)} = \varepsilon_t^2.$$

Ce qui entraîne d'après (3.1) que

$$u = \varepsilon_t^2. \quad (3.2)$$

Comme expliqué en (2.22), lorsque la valeur de l est suffisamment grande, $s_l(u)$ se réduit à u . Par conséquent, nous pouvons dire que

$$\mathbf{P}\{s_l(u) = u\} = \mathbf{P}\{u \leq l\},$$

ce qui donne d'après (3.2) que

$$\mathbf{P}\{s_l(\varepsilon_t^2) = \varepsilon_t^2\} = \mathbf{P}\{\varepsilon_t^2 \leq l\}.$$

Afin de trouver une valeur à l , nous réalisons plusieurs essais et retenons le cas pour lequel il y'a une proportion de 2,5% d'outliers dans la série observée, c'est-à-dire

$$\mathbf{P}\{\varepsilon_t^2 \leq l\} = 0,975. \quad (3.3)$$

En d'autres termes, l est choisi suffisamment grand de sorte que 97,5% des observations de la série soient conservées. En considérant que ε_t suit une loi normale centrée réduite, ε_t^2 suit une loi de Khi-deux à un degré de liberté. Par conséquent, l est le quantile d'ordre 0.975 d'un Khi-deux à un degré de liberté. D'après la table de Khi-deux, la valeur de l cherchée est 5.02. Ainsi, pour la construction du BM-estimateur, nous prendrons

$$l = 5. \quad (3.4)$$

3.1.2 Choix du modèle

Pour réaliser cette étude numérique, nous générons les observations par le modèle HGARCH(1,d,1) défini comme suit

$$\begin{cases} y_t = h_t^{1/2} \varepsilon_t \\ h_t(\theta) = \gamma + \beta h_{t-1} + \omega \{1 - \beta L - (1 - \delta L)(1 - L)^d\} y_{t-1}^2. \end{cases} \quad (3.5)$$

Pour la programmation sous le logiciel libre R, nous utilisons l'expression de la variance conditionnelle suivante

$$h_t(\theta) = \gamma + \beta h_{t-1} + \omega \{\pi_1(d) - \beta + \delta\} y_{t-1}^2 + \omega \sum_{j=2}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) y_{t-j}^2. \quad (3.6)$$

En effet, d'après (2.2), nous avons

$$h_t(\theta) = \gamma + \beta h_{t-1} + \omega \left\{ 1 - \beta L - (1 - \delta L) \left(1 - \sum_{j=1}^{\infty} \pi_j(d) L^j \right) \right\} y_t^2.$$

En développant les termes entre parenthèses, nous obtenons

$$\begin{aligned} h_t(\theta) &= \gamma + \beta h_{t-1} + \omega \left\{ 1 - \beta L - (1 - \delta L) + \sum_{j=1}^{\infty} (1 - \delta L) \pi_j(d) L^j \right\} y_t^2 \\ &= \gamma + \beta h_{t-1} + \omega \left\{ 1 - \beta L - (1 - \delta L) + \sum_{j=1}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) L^j \right\} y_t^2 \\ &= \gamma + \beta h_{t-1} + \omega \left\{ 1 - \beta L - 1 + \delta L \right\} y_t^2 + \omega \left\{ \sum_{j=1}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) L^j \right\} y_t^2 \\ &= \gamma + \beta h_{t-1} + \omega \left\{ -\beta L + \delta L \right\} y_t^2 + \omega \sum_{j=1}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) y_{t-j}^2 \\ &= \gamma + \beta h_{t-1} + \omega \left\{ \pi_1(d) - \beta + \delta \right\} y_{t-1}^2 + \omega \sum_{j=2}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) y_{t-j}^2. \end{aligned}$$

L'expression (3.6) de la variance conditionnelle a également été utilisée par Li et al. [46] pour leurs études numériques. Pour simuler le processus y_t défini en (3.5), nous avons considéré deux distributions pour le bruit ε_t . Ces distributions sont les suivantes :

1. ε_t suit la loi normale centrée réduite.
2. ε_t suit la loi de Student à sept degrés de liberté.

Egalement, pour chacune des distributions précédentes, nous examinons les trois cas suivants.

1^{er} cas Le processus y_t est généré sans outliers.

2^{ème} cas Le processus y_t est généré avec une proportion de 5% d'outliers.

3^{ème} cas Le processus y_t est généré avec une proportion de 10% d'outliers.

En ce qui concerne les outliers, il s'agit exclusivement d'outliers additives. Le principe d'obtention des séries contaminées se fait en deux étapes. D'abord, nous générons des observations issues du modèle HGARCH et qui ne contiennent pas d'outliers. Ensuite, nous sélectionnons dans la série obtenue, un certain nombre d'observations selon le pourcentage d'outliers qu'on veut avoir dans la série. A ces observations sélectionnées, nous ajoutons une certaine quantité afin d'obtenir les outliers. Ainsi, la série y_t^* contaminée par la présence d'outliers additives que nous avons considérée est définie comme dans Muler et Yohai [55] et Lu et al. [49] par l'expression suivante

$$y_t^* = \begin{cases} y_t + e h_t^{1/2} & \text{si } t = t_j, \quad 1 \leq j \leq r = kn/100, \\ y_t & \text{sinon,} \end{cases} \quad (3.7)$$

avec y_t , la série non contaminée générée par le modèle HGARCH. k est le pourcentage de contamination qui correspond à la proportion d'outliers qu'on a insérée dans les observations. Les instants t_1, \dots, t_r sont choisis à intervalle régulier et représentent les instants où les outliers apparaissent dans la série et n représente la taille de l'échantillon. Après plusieurs essais ($e = 1, \dots, 5$), nous ne présenterons que deux valeurs de e pour cette étude, à savoir $e = 3$ et $e = 5$. Les résultats présentés avec ces deux valeurs de e résument l'ensemble des études de simulations réalisées. Nous explorons ces deux cas pour évaluer l'influence de ce paramètre sur les résultats.

Nous considérons pour cette étude de simulation trois échantillons de taille respective $n = 1000, 2000$ et 4000 . Les trois estimateurs en compétition seront évalués par rapport à leurs performances d'estimations et de prévisions. Par conséquent, dans chaque échantillon considéré, 80% des observations sont utilisées pour l'estimation et les 20% restant pour la prévision hors échantillon. Au niveau de la prévision, nous retenons quatre horizons de prévisions qui sont $l = 1, 5, 10$ et 20 . En effet, comme dans Li et al. [46], nous avons considéré l'horizon 1 et l'horizon 5 comme horizons à court terme tandis que les horizons 10 et 20 sont considérés comme des horizons à long terme. De plus, l'expression des différents horizons de prévisions de la variance conditionnelle du modèle HGARCH(1,d,1) est donnée par les relations suivantes, pour l'horizon $l = 1$, nous avons

$$\hat{h}_{t,\theta}(1) = \gamma + \beta h_{t-1} + \omega \{\pi_1(d) - \beta + \delta\} y_{t-1}^2 + \omega \sum_{j=2}^{\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) y_{t-j}^2. \quad (3.8)$$

Pour l'horizon $l \geq 2$, nous avons

$$\begin{aligned} \hat{h}_{t,\theta}(l) &= \gamma + \{\omega(\pi_1(d) + \delta) + (1 - \omega)\beta\} \hat{h}_{t,\theta}(l-1) \\ &\quad + \omega \sum_{j=2}^{l-1} (\pi_j(d) - \delta \pi_{j-1}(d)) \hat{h}_{t,\theta}(j-1) + \omega \sum_{j=l}^{+\infty} (\pi_j(d) - \delta \pi_{j-1}(d)) y_{t+l-j}^2, \end{aligned}$$

avec \hat{h}_t qui désigne la valeur prédite de h_t . Par ailleurs, pour chaque taille d'échantillon considérée, nous réalisons $N = 1000$ réplifications. A chacune des réplifications, les 2000 premières observations sont supprimées afin d'atténuer dans les séries générées, l'effet des valeurs initiales que nous avons imposées dans la construction des prédicteurs de la variance conditionnelle en (2.16) et (2.21). En d'autres termes, pour chacune des 1000 réplifications, des séries de $2000 + n$ observations seront générées chaque fois. Le vrai vecteur de paramètres que nous avons utilisé pour générer les observations dans cette étude de simulation est donné par

$$\theta_0 = (\gamma_0, \beta_0, \delta_0, \omega_0, d_0) = (0.1, 0.4, 0.2, 0.5, 0.6).$$

Pour comparer les performances d'estimations du M-estimateur, du BM-estimateur et de l'estimateur du quasi-maximum de vraisemblance, nous utiliserons, comme dans Kwan et al. [43], les indices du RMSE (Root Mean Square Error) et du MAE (Mean Absolute Error). Pour la comparaison des performances de prévisions, nous utilisons uniquement le RMSE. Ces indices sont définis par

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_0)^2} \quad \text{et} \quad MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_0|, \quad (3.9)$$

où \hat{y}_i désigne la valeur prédite de y_0 et N le nombre de réplifications. Les résultats des simulations sont consignés dans les tableaux 3.1 - 3.10.

3.1.3 Cas du modèle HGARCH généré par la loi normale centrée réduite

Dans cette partie, nous présentons sous forme de tableaux les résultats des simulations réalisées avec un modèle HGARCH(1,d,1) pour lequel le bruit ε_t suit une loi normale centrée réduite.

Le tableau **3.1** contient les valeurs du RMSE et du MAE pour les parties estimation et prévision des trois estimateurs en compétition pour le modèle HGARCH(1,d,1) généré par une loi normale centrée réduite sans outliers. Il s'agit de l'estimateur du Quasi Maximum de vraisemblance (QMV), le M-estimateur (M_m) et le M-estimateur borné (BM). A partir du tableau **3.1**, nous constatons que les valeurs du RMSE et du MAE pour la partie estimation des paramètres du modèle de toutes les méthodes d'estimation diminuent et tendent vers zéro quand la taille de l'échantillon augmente. Cette décroissance des indices du RMSE et du MAE quand n devient grand traduit la convergence de chacun des estimateurs. Les résultats consignés au tableau **3.1** illustrent bien le résultat théorique de convergence des M et BM estimateurs que nous avons établis au chapitre 2. Il en est de même pour l'estimateur du QMV.

TABLEAU 3.1 – Résultats pour le modèle HGARCH(1,d,1) normal sans outliers

Résultats d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.04361152	0.21071733	0.18531022	0.14781896	0.28407642
<i>RMSE</i>	M_m	0.04637462	0.21347503	0.19956722	0.16168414	0.28187780
	BM	0.0444476	0.2055457	0.2033027	0.1589536	0.2755895
1000	QMV	0.0342135	0.1742121	0.1428745	0.1112259	0.2486029
<i>MAE</i>	M_m	0.0360574	0.1763166	0.1502495	0.1220998	0.2471377
	BM	0.0345099	0.1695302	0.1517171	0.1193955	0.2400837
2000	QMV	0.03429911	0.18467751	0.13484239	0.10624846	0.23219205
<i>RMSE</i>	M_m	0.03572537	0.18539711	0.14022858	0.12291937	0.23774955
	BM	0.03485032	0.17928535	0.13608130	0.12163424	0.22657021
2000	QMV	0.0280271	0.1578663	0.1094460	0.0805565	0.1964174
<i>MAE</i>	M_m	0.0287165	0.1564427	0.1115852	0.0908726	0.2020973
	BM	0.0277052	0.1504490	0.1079263	0.0899925	0.1902408
4000	QMV	0.02655261	0.14789295	0.11457942	0.07717198	0.17047913
<i>RMSE</i>	M_m	0.02742995	0.15315357	0.11539842	0.08854180	0.18248984
	BM	0.02589646	0.14443635	0.11133267	0.09524189	0.17441386
4000	QMV	0.0220159	0.1267519	0.0912920	0.0595648	0.1382659
<i>MAE</i>	M_m	0.0225451	0.1315196	0.0915822	0.0672422	0.1499602
	BM	0.0208985	0.1228190	0.0875777	0.0733259	0.1417378

RMSE pour les prévisions

	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QMV	0.4969 (0.534) ¹	0.5804 (0.631) ¹	0.5045 (0.534) ¹	0.6378 (0.871) ¹
M_m	0.4958 (0.000) ²	0.5815 (0.055) ²	0.5068 (0.234) ²	0.6421 (0.608) ²
BM	0.4948 (0.000) ³	0.5811(0.069) ³	0.5099 (0.231) ³	0.6436 (0.662) ³
n=400				
QMV	0.5671 (0.149) ¹	0.5065 (0.187) ¹	0.5339 (0.417) ¹	0.5214(0.149) ¹
M_m	0.5682 (0.0000) ²	0.5060 (0.011) ²	0.5335 (0.079) ²	0.5218(0.346) ²
BM	0.5689 (0.0000) ³	0.5067(0.008) ³	0.5374 (0.077) ³	0.5293(0.294) ³
n=800				
QMV	0.4918 (0.139) ¹	0.5391 (0.142) ¹	0.4867 (0.297) ¹	0.4610(0.139) ¹
M_m	0.4961 (0.000) ²	0.5385 (0.015) ²	0.4863 (0.073) ²	0.4616(0.151) ²
BM	0.4980 (0.000) ³	0.5385 (0.003) ³	0.4905 (0.027) ³	0.4685(0.080) ³

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

De plus, nous remarquons que pour la partie estimation des paramètres, l'estimateur du QMV, le M-estimateur et le BM-estimateur ont des résultats qui sont presque similaires. Cependant, selon les résultats du tableau **3.1**, il semblerait que l'estimateur du QMV présente les plus petites valeurs de RMSE et de MAE. Nous pouvons donc dire que pour l'estimation des paramètres dans un modèle HGARCH(1,d,1) normal sans outliers, l'estimateur du QMV est un peu plus meilleur que les autres. En ce qui concerne les résultats de la partie prévision, nous constatons que les valeurs du RMSE obtenues pour les trois estimateurs sont pratiquement similaires. Ce qui permet de dire que les trois estimateurs ont presque les mêmes performances de prévision. De plus, nous remarquons que l'estimateur QMV présente en général les plus petites valeurs. Par conséquent, il paraît évident que l'estimateur QMV est sensiblement le meilleur pour la prévision. Les valeurs données entre parenthèses représentent les p-values du test de Diebold-Mariano expliqué dans Diebold et Mariano [21]. Ce test nous permet d'évaluer la qualité de chaque prévision. En d'autres termes, il permet d'évaluer la précision de prévision entre deux méthodes. L'hypothèse nulle de ce test est que les deux méthodes ont la même précision de prévision. L'hypothèse alternative est que les deux méthodes ont des niveaux de précision différents. Dans le cas du tableau **3.1**, nous constatons que pour tous les horizons de prévisions, les valeurs dans (...) ¹ sont supérieures à 0.05. L'hypothèse nulle est toujours acceptée dans ces différents cas.

Par conséquent, nous pouvons conclure qu'il n'ya pas de différence significative entre les précisions de prévision du QMV et du M_m. Pour les horizons 1 et 5, les valeurs dans (...) ² et (...) ³ sont pratiquement inférieures à 0.05. Ce qui veut dire que l'hypothèse nulle est rejetée. Nous en déduisons que pour les horizons 1 et 5, il existe une différence significative entre les précisions prévision des différents estimateurs. Cette différence de précision peut-être due au choix spécifique des observations dans l'échantillon. Pour la prévision à long terme, l'hypothèse nulle est toujours acceptée. Donc les prévisions ont la même précision pour les horizons 10 et 20.

Dans le tableau **3.2**, nous reportons les valeurs du RMSE et du MAE pour les parties estimation et prévision des trois estimateurs en compétition pour le modèle HGARCH(1,d,1) normal contenant 5% d'outliers. Dans ce cas, le paramètre e défini dans (3.7) pour générer les outliers est égal à 3. Pour la partie estimation, nous remarquons que les valeurs des RMSE et MAE des différents estimateurs ne sont pas proches les unes des autres. De plus, seuls les RMSE et MAE du BM-estimateur diminuent quand la taille de l'échantillon augmente. Ce n'est pas le cas des paramètres δ et d pour l'estimateur du QMV et le paramètre d du M-estimateur. Ce tableau illustre bien la propriété de convergence du BM-estimateur. Donc l'ajout d'outliers a engendré des perturbations dans le comportement des trois différents estimateurs. En outre, le BM-estimateur présente les plus petites valeurs de MAE et de RMSE. Nous pouvons donc dire que le BM-estimateur est meilleur pour l'estimation dans le cas où il y'a 5% d'outliers. Pour les résultats des prévisions, nous observons que les résultats obtenus pour les M et BM estimateurs sont pratiquement similaires. Cependant, le M-estimateur se présente comme ayant les plus petits RMSE pour les horizons 5, 10 et 20. Par conséquent, nous pouvons dire qu'il est le meilleur en termes de prévision. L'estimateur QMV présente les valeurs de RMSE les plus élevées. Il est important de noter l'apparition de grandes valeurs de RMSE à l'horizon 1. Ces valeurs sont dues à l'apparition d'une outlier à cet horizon pendant l'insertion des outliers additives dans les données. En effet, les techniques d'estimation que nous explorons dans notre travail ne sont pas adaptées à la prévision des outliers. Etant donné que la prévision des outliers n'est pas l'objet de notre étude, et que nous n'avons pas dans ce travail les outils nécessaires pour ce type d'estimation, il serait donc surréaliste de comparer nos méthodes d'estimation pour cet horizon. Enfin, avec les p-values du test de Diebold-Mariano, il semblerait que toutes les méthodes de prévision ont significativement les mêmes précisions. Nous pouvons donc déduire que les résultats de prévision obtenus seront indépendants des données dans l'échantillon.

TABLEAU 3.2 – Résultats du modèle HGARCH(1,d,1) normal avec 5% d'outliers additives et $e = 3$.

Résultat d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.1824507	0.2389641	0.5594601	0.3168024	0.3022920
<i>RMSE</i>	M_m	0.1225048	0.2466524	0.3924226	0.3304879	0.3205431
	BM	0.08825207	0.24250316	0.32239109	0.23873986	0.29328744
1000	QMV	0.1567677	0.1955715	0.4668623	0.2907232	0.2566926
<i>MAE</i>	M_m	0.1003156	0.2031034	0.2885874	0.3035852	0.2755282
	BM	0.0664480	0.2000983	0.2395116	0.1964150	0.2524568
2000	QMV	0.1647756	0.2190157	0.5680752	0.2982690	0.3040075
<i>RMSE</i>	M_m	0.1168934	0.2334681	0.3385703	0.3147113	0.3279849
	BM	0.07689233	0.22260976	0.25345363	0.20714702	0.28940060
2000	QMV	0.1433223	0.1766261	0.4749507	0.2791704	0.2573172
<i>MAE</i>	M_m	0.0995834	0.1901209	0.2424024	0.2918625	0.2822235
	BM	0.0599613	0.1828199	0.1897606	0.1697019	0.2452585
4000	QMV	0.1519731	0.2069074	0.5819863	0.2892006	0.3113580
<i>RMSE</i>	M_m	0.1097818	0.2080384	0.2808735	0.3010942	0.3278539
	BM	0.06359509	0.17851374	0.19118510	0.17245534	0.26633820
4000	QMV	0.1310870	0.1698337	0.4978103	0.2755681	0.2651011
<i>MAE</i>	M_m	0.0977251	0.1664175	0.1958090	0.2809762	0.2831434
	BM	0.0516965	0.1459989	0.1416375	0.1428029	0.2216879

RMSE pour la prévision				
	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QMV	3.9332(0.986) ¹	0.6026 (0.984) ¹	0.6199 (0.986) ¹	0.6294 (0.998) ¹
M_m	3.9954(0.714) ²	0.5813 (0.602) ²	0.6040 (0.706) ²	0.5037 (0.745) ²
BM	3.9861(0.925) ³	0.5832 (0.920) ³	0.6069 (0.929) ³	0.5509 (0.988) ³
n=400				
QMV	3.9175(0.548) ¹	0.5083 (0.567) ¹	0.5292 (0.595) ¹	0.5727(0.892) ¹
M_m	3.9748(0.369) ²	0.4807 (0.394) ²	0.4957 (0.434) ²	0.5247(0.810) ¹
BM	3.9610(0.632) ³	0.4837 (0.647) ³	0.5029 (0.670) ³	0.5335(0.919) ³
n=800				
QMV	3.7245(0.580) ¹	0.7621 (0.613) ¹	0.5159 (0.641) ¹	0.5016(0.883) ¹
M_m	3.7788(0.479) ²	0.7485 (0.514) ²	0.4814 (0.554) ²	0.4506(0.767) ²
BM	3.7705(0.6348) ³	0.7493 (0.664) ³	0.4900 (0.687) ³	0.4579(0.910) ³

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

Dans le tableau **3.3**, nous rapportons les résultats du RMSE et du MAE pour les parties estimations et les prévisions des trois estimateurs pour un modèle HGARCH (1,d,1) normal contenant 5% d'outliers avec $e = 5$.

Pour la partie estimation, il est nécessaire de constater que les valeurs des RMSE et MAE sont éloignées les unes des autres. Ce qui traduit une nette différence entre les trois estimateurs en compétition. De plus, seuls les RMSE et MAE du BM-estimateur diminuent quand la taille de l'échantillon augmente. Ce n'est pas le cas du paramètre d pour le M-estimateur. Par ailleurs, il semble que tous les paramètres du QMV soient perturbés par la présence de ce type d'outliers additives. Ce tableau illustre la propriété de convergence du BM-estimateur et montre la grande sensibilité du QMV aux outliers. En outre, le BM-estimateur présente les plus petites valeurs du RMSE et du MAE. Le BM-estimateur est donc le meilleur pour l'estimation des paramètres. En ce qui concerne les résultats de prévision, nous notons que les valeurs pour les BM et M estimateurs sont assez proches les unes des autres. Ce qui laisserait à penser que ces deux estimateurs ont approximativement les mêmes performances de prévisions. Néanmoins, le M-estimateur apparaît comme l'estimateur qui a les plus petits RMSE. Il peut donc être considéré comme le meilleur pour la prévision pour les horizons à courts et longs termes. Le QMV a indiscutablement la pire des performances. La présence de grandes valeurs de RMSE à l'horizon 1 s'explique par l'apparition d'une outlier à cet horizon pendant l'insertion des outliers additives dans les données. Avec les p-values du test de Diehold-Mariano, nous pouvons dire que toutes les méthodes de prévisions ont significativement les mêmes précisions. Nous pouvons donc déduire que les résultats de prévision obtenus seront indépendants des données dans l'échantillon.

Après avoir étudié différentes valeurs de e , nous pouvons conclure que pour le modèle normal HGARCH(1,d,1), l'ajout d'une proportion de 5% d'outliers à engendré des perturbations dans le comportement des estimateurs. Ainsi, le QMV s'est révélé être très affecté par la présence de ces outliers surtout au niveau de la prévision. Le BM-estimateur est plus résistant aux outliers que tous les autres estimateurs. Nous dirons que pour la partie estimation, il est l'estimateur le plus robuste. C'est donc le meilleur pour l'estimation des paramètres. Le M-estimateur quant à lui sera privilégié pour la prévision.

TABLEAU 3.3 – Résultats du modèle HGARCH(1,d,1) normal avec 5% d'outliers additives et $e = 5$.

Résultat d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.4518902	0.2519396	0.5201002	0.4221391	0.2643575
<i>RMSE</i>	M_m	0.1005877	0.2705183	0.2636832	0.3620779	0.3086049
	BM	0.05153823	0.22127280	0.22321691	0.17737535	0.25853110
1000	QMV	0.4259810	0.2367324	0.4359100	0.3940847	0.2472110
<i>MAE</i>	M_m	0.0823139	0.2320896	0.1902356	0.3398584	0.2604685
	BM	0.0406182	0.1861499	0.1691550	0.1378854	0.2162802
2000	QMV	0.4584906	0.2494533	0.5231717	0.4321780	0.2648688
<i>RMSE</i>	M_m	0.09058128	0.26341472	0.20000807	0.35350364	0.30909291
	BM	0.03900253	0.18963719	0.16686540	0.13309793	0.22634063
2000	QMV	0.4323235	0.2366489	0.4417509	0.4128225	0.2504236
<i>MAE</i>	M_m	0.0741354	0.2257466	0.1443882	0.3326238	0.2586171
	BM	0.0320692	0.1590653	0.1288769	0.1050128	0.1838603
4000	QMV	0.4671310	0.2512758	0.5348963	0.4458519	0.2543854
<i>RMSE</i>	M_m	0.07613541	0.23680213	0.14095148	0.34634290	0.30902588
	BM	0.02974118	0.15332327	0.13119966	0.09548084	0.18348735
4000	QMV	0.4394668	0.2386203	0.4598530	0.4346389	0.2422358
<i>MAE</i>	M_m	0.0634537	0.2032611	0.1096697	0.3315560	0.2616048
	BM	0.0248348	0.1291357	0.1034101	0.0768272	0.1497912

RMSE pour la prévision

	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QMV	9.0302(0.107) ¹	0.7066 (0.388) ¹	1.4168 (0.502) ¹	52.318 (0.768) ¹
M_m	9.3083(0.243) ²	0.4326 (0.286) ²	0.5263 (0.352) ²	0.5948 (0.699) ²
BM	9.2814(0.106) ³	0.4399 (0.388) ³	0.5342 (0.502) ³	0.6008 (0.768) ³
n=400				
QMV	9.6258(0.071) ¹	0.7098 (0.370) ¹	0.9192 (0.493) ¹	10.400(0.766) ¹
M_m	9.8883(0.196) ²	0.5033 (0.232) ²	0.5172 (0.276) ²	0.5669(0.734) ²
BM	9.848(0.070) ³	0.5021 (0.369) ³	0.5237 (0.492) ³	0.5726(0.763) ³
n=800				
QMV	9.4842(0.868) ¹	0.7063 (0.823) ¹	0.7582 (0.843) ¹	1.3591102(0.981) ¹
M_m	9.7395(0.205) ²	0.5254 (0.238) ²	0.5114 (0.297) ²	0.4798(0.711) ²
BM	9.7061(0.995) ³	0.5277 (0.994) ³	0.5168 (0.994) ³	0.4928(0.993) ³

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

Dans le tableau **3.4**, nous résumons les valeurs du RMSE et du MAE pour le modèle HGARCH (1,d,1) normal avec 10% d'outliers pour $e = 3$. Pour la partie estimation, au fur et à mesure que la taille de l'échantillon augmente, on observe la décroissance du RMSE uniquement pour les paramètres γ et β du QMV, les paramètres δ et d du M-estimateur et le paramètre d du BM-estimateur. Toutes les méthodes d'estimations sont perturbées par ce type et cette proportion d'outliers. Cependant, à partir des deux indices MAE et RMSE, le M-estimateur sera désigné comme le meilleur car il contrôle mieux le paramètre de longue mémoire (d) qui caractérise les propriétés du modèle HGARCH. Pour les résultats de prévision, les valeurs du RMSE du M-estimateur et du BM-estimateur sont quelque peu similaires. Cependant, un avantage sera donné au M-estimateur qui a le plus petit RMSE à tous les horizons à l'exception de l'horizon 1. De plus, les p-values du test de Mariano-Diebold montrent que toutes les prévisions ont la même précision.

Dans le tableau **3.5**, nous résumons les valeurs du RMSE et du MAE des trois estimateurs en compétition pour le modèle HGARCH(1,d,1) normal avec 10% d'outliers pour $e = 5$. Pour la partie estimation, les valeurs du RMSE et du MAE des trois estimateurs sont éloignées les unes des autres. De plus, nous constatons que tous les paramètres du BM-estimateur diminuent à mesure que la taille de l'échantillon augmente. Ce qui n'est pas le cas pour le M-estimateur et pour le QMV. Il y'a donc la propriété de convergence du BM-estimateur qui est vérifiée par les résultats du tableau **3.5**. Par ailleurs, le BM-estimateur a également les valeurs des RMSE et MAE les plus faibles. Il se comporte donc le mieux et est préférable aux autres estimateurs pour la partie estimation. D'après les résultats de la partie prévision, les p-values du test de Diebold-Mariano montrent que les prévisions ont généralement la même précision. Notons, à partir des valeurs du RMSE que le QMV a les pires performances. Le M-estimateur et le BM-estimateur ont les valeurs les plus faibles et assez proches. Cependant, le M-estimateur a les plus petits RMSE en général pour les prévisions aux horizons 5, 10 et 20. Il peut donc être considéré comme le plus performant pour la prévisions.

TABLEAU 3.4 – Résultats du modèle HGARCH(1,d,1) normal avec 10% d'outliers additives et $e = 3$.

Résultats d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.3490146	0.2645680	0.5756572	0.4169015	0.2973865
<i>RMSE</i>	M_m	0.2630635	0.2919219	0.4506256	0.4437230	0.3121641
	BM	0.2516656	0.2941109	0.4914643	0.4202650	0.3145411
1000	QMV	0.3207306	0.2222750	0.4904890	0.3991049	0.2690081
<i>MAE</i>	M_m	0.2403757	0.2642968	0.3517488	0.4258935	0.2860985
	BM	0.2283504	0.2641756	0.3907854	0.3991946	0.2879458
2000	QMV	0.3465734	0.2480457	0.6116746	0.4243433	0.3001224
<i>RMSE</i>	M_m	0.2827183	0.2887196	0.4275370	0.4604038	0.3015558
	BM	0.2690417	0.2859154	0.4947990	0.4425642	0.3099522
2000	QMV	0.3202454	0.2033926	0.5335604	0.4088358	0.2750252
<i>MAE</i>	M_m	0.2672020	0.2675340	0.3333156	0.4455246	0.2813127
	BM	0.2512289	0.2587626	0.3930219	0.4269325	0.2906585
4000	QMV	0.3395143	0.2293033	0.6480537	0.4319478	0.2995103
<i>RMSE</i>	M_m	0.2910831	0.2933086	0.3962053	0.4697166	0.2903102
	BM	0.2788411	0.2885405	0.5036131	0.4525467	0.2968972
4000	QMV	0.3174852	0.1787691	0.5842657	0.4189693	0.2723853
<i>MAE</i>	M_m	0.2779050	0.2787670	0.3092967	0.4581473	0.2731245
	BM	0.2647884	0.2636224	0.4062010	0.4394239	0.2775671
RMSE pour la prévision						
	horizon 1	horizon 5	horizon 10	horizon 20		
n=200						
QMV	3.9046(0.511) ¹	0.5719 (0.332) ¹	0.6023 (0.620) ¹	0.6019(0.869) ¹		
M_m	3.9903(0.846) ²	0.5072 (0.751) ²	0.5204 (0.761) ²	0.5190 (0.708) ²		
BM	3.9904(0.471) ³	0.5102 (0.298) ³	0.5285 (0.592) ³	0.5358 (0.858) ³		
n=400						
QMV	4.2126(0.521) ¹	0.5985(0.348) ¹	0.6157 (0.521) ¹	0.6286(0.851) ¹		
M_m	4.2808 (0.892) ²	0.5245 (0.850) ²	0.5455(0.854) ²	0.5576(0.942) ²		
BM	4.2801(0.497) ³	0.5262 (0.322) ³	0.5510 (0.622) ³	0.5672(0.844) ³		
n=800						
QMV	3.9106(0.530) ¹	0.6528 (0.359) ¹	0.5724 (0.640) ¹	0.7397(0.869) ¹		
M_m	3.9874(0.885) ²	0.5903 (0.872) ²	0.4894 (0.702) ²	0.6821(0.982) ²		
BM	3.9906(0.524) ³	0.5906 (0.357) ³	0.4900 (0.635) ³	0.6843(0.871) ³		

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

TABLEAU 3.5 – Résultats du modèle HGARCH(1,d,1) normal avec 10% d'outliers additives et $e = 5$.

Résultats d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.5501989	0.2822957	0.2388864	0.4177522	0.4067638
<i>RMSE</i>	M_m	0.1045551	0.2685897	0.2263825	0.3950856	0.3133597
	BM	0.05622594	0.22811116	0.21373243	0.19377377	0.26376515
1000	QMV	0.4693620	0.2301027	0.1820533	0.3945965	0.3870595
<i>MAE</i>	M_m	0.0834447	0.2359551	0.1728466	0.3745398	0.2647210
	BM	0.0447219	0.1951067	0.1655397	0.1538771	0.2211813
2000	QMV	0.4952832	0.3023534	0.2140564	0.4032478	0.4204120
<i>RMSE</i>	M_m	0.08620839	0.26333737	0.16675999	0.39160136	0.28710061
	BM	0.04286325	0.20270592	0.14565109	0.14326776	0.22103315
2000	QMV	0.4023883	0.2512016	0.1632025	0.3771393	0.3999923
<i>MAE</i>	M_m	0.06777219	0.2328813	0.1248515	0.3777633	0.2328527
	BM	0.0367684	0.1748798	0.1161040	0.1136829	0.1815490
4000	QMV	0.4048013	0.3344157	0.1878658	0.3741157	0.4475092
<i>RMSE</i>	M_m	0.06420363	0.23992399	0.11429608	0.39244864	0.26899302
	BM	0.03558152	0.16938805	0.10684028	0.10766980	0.17592334
4000	QMV	0.305757	5 0.2903262	0.1519310	0.3469654	0.4290392
<i>MAE</i>	M_m	0.0510370	0.2081071	0.0904891	0.3847413	0.2148922
	BM	0.0315424	0.1463628	0.0856704	0.0907294	0.1417067

RMSE pour la prévision				
	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QML	9.9947(0.334) ¹	0.9447 (0.128) ¹	0.9868 (0.448) ¹	1.1100 (0.795) ¹
M_m	10.344(0.143) ²	0.5785 (0.035) ²	0.5637 (0.288) ²	0.4775 (0.561) ²
BM	10.321(0.362) ³	0.5905(0.149) ³	0.5731 (0.473) ³	0.4996 (0.812) ³
n=400				
QML	9.1181(0.293) ¹	0.8742 (0.112) ¹	0.9473 (0.427) ¹	0.9670(0.744) ¹
M_m	9.4576(0.140) ²	0.5053 (0.035) ²	0.5574 (0.279) ²	0.4865 (0.440) ²
BM	9.4380(0.318) ³	0.5108 (0.131) ³	0.5552(0.450) ³	0.4978 (0.765) ³
n=800				
QML	9.1177(0.285) ¹	0.8564(0.114) ¹	0.9395 (0.429) ¹	0.9687 (0.708) ¹
M_m	9.4037 (0.154) ²	0.5221 (0.063) ²	0.5952(0.342) ²	0.5054(0.154) ²
BM	9.3951(0.308) ³	0.5270 (0.129) ³	0.5957 (0.447) ³	0.5156(0.738) ³

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

Après avoir étudié le comportement des estimateurs en compétition en présence d'outliers générés par différentes valeurs de e , nous pouvons affirmer que pour le modèle HGARCH(1,d,1) normal, l'ajout d'une proportion de 10% d'outliers à entraîner des variations dans le comportement de tous les estimateurs. Le BM-estimateur par ses performances est donc le meilleur pour l'estimation des paramètres. Le M-estimateur quant à lui, sera privilégié pour la prévision.

Nous considérons maintenant que le processus ε_t suit la loi de Student avec sept degrés (t_7) de liberté.

3.1.4 Cas du modèle HGARCH issu d'une loi de Student à 7 degrés de liberté (Student's t_7)

Les tableaux suivants résument les résultats des simulations réalisées avec un modèle HGARCH(1,d,1) pour lequel le bruit ε_t suit une loi de Student à sept degrés de liberté.

Les tableaux **3.6** - **3.10** contiennent les résultats d'estimations et de prévisions de l'estimateur du quasi-maximum de vraisemblance (QMV), du M-estimateur (M_m) et du BM-estimateur (BM) pour le modèle HGARCH(1,d,1) issu d'une loi de Student à sept degrés de liberté. De manière générale, nous pouvons remarquer des valeurs assez élevées des RMSE de la partie prévision. Ces valeurs révèlent des prévisions qui ne sont pas très intéressantes pour la distribution de Student. Les analyses et interprétations des résultats et de ces tableaux pour les parties estimation et prévision sont similaires à celles du modèle HGARCH(1,d,1) généré par une loi normale. Ce qui traduit le fait que les performances des trois estimateurs en compétition ne dépendent pas de la distribution du bruit.

Nous rappelons de plus que les figures A.1 et A.2 en annexe présentent les boxplots des paramètres du HGARCH (1,d,1) normal pour les différentes méthodes d'estimation avec 5% d'outliers. Ces figures montrent que pour un échantillon de taille finie, la distribution des observations issues du BM-estimateur et du M-estimateur est plus symétrique que celle des observations du QMV. Les variances des paramètres pour le BM-estimateur sont également inférieures à celles du M-estimateur et du QMV. Les boxplots, pour la présence de 10% d'outliers et ceux représentés pour un HGARCH (1,d,1) issu d'une loi de Student sont consignés en annexe et fournissent des résultats pratiquement similaires à ceux de la loi normale.

TABLEAU 3.6 – Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) sans outliers

Résultats d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.09183176	0.23980786	0.17317487	0.20881758	0.28760310
<i>RMSE</i>	M_m	0.07066585	0.23744129	0.17829237	0.16498721	0.27705762
	BM	0.07015049	0.22272819	0.17078276	0.22687306	0.26764230
1000	QMV	0.0693544	0.2019195	0.1410889	0.1730103	0.2523373
<i>MAE</i>	M_m	0.0542833	0.2013467	0.1401859	0.1282793	0.2391913
	BM	0.05298725	0.18790208	0.13553592	0.19398930	0.22888948
2000	QMV	0.06879443	0.22685563	0.14799033	0.20351628	0.24512754
<i>RMSE</i>	M_m	0.05465952	0.22021419	0.14525310	0.13959519	0.23359104
	BM	0.05081348	0.20909129	0.14098730	0.22429108	0.22973904
2000	QMV	0.0539783	0.1943295	0.1256835	0.1763391	0.2040543
<i>MAE</i>	M_m	0.0432261	0.1887166	0.1206872	0.1102202	0.1932087
	BM	0.0402513	0.1776479	0.1171477	0.2004755	0.1894803
4000	QMV	0.06271285	0.19250059	0.13029092	0.19252975	0.18103072
<i>RMSE</i>	M_m	0.04714152	0.18944588	0.13263131	0.11835210	0.17697064
	BM	0.04130632	0.17764559	0.13117694	0.21894204	0.17141760
4000	QMV	0.0503974	0.1617177	0.10883393	0.17647090	0.14252188
<i>MAE</i>	M_m	0.0382459	0.1596679	0.1101325	0.1001134	0.1376027
	BM	0.0334107	0.1481438	0.1092638	0.2062580	0.13406586

RMSE pour la prévision

	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QMV	1.4417(0.079) ¹	1.5008 (0.388) ¹	2.0214 (0.500) ¹	3.9055 (0.728) ¹
M_m	1.4039(0.046) ²	1.5181 (0.354) ²	2.0190 (0.481) ²	3.7083 (0.724) ²
BM	1.4347(0.040) ³	1.5474(0.346) ³	2.0252 (0.475) ³	4.9411 (0.722) ³
n=400				
QMV	1.6033(0.004) ¹	1.9133 (0.236) ¹	1.7101 (0.384) ¹	3.1987(0.632) ¹
M_m	1.5526 (0.000) ²	1.8538 (0.163) ²	1.6257(0.339) ²	2.1238(0.212) ²
BM	1.6250(0.009) ³	1.8666 (0.260) ³	1.6416 (0.392) ³	2.3427(0.674) ³
n=800				
QMV	1.9835(0.104) ¹	2.7104 (0.280) ¹	3.0961 (0.447) ¹	1.6243(0.473) ¹
M_m	1.9306(0.055) ²	2.7036 (0.094) ²	3.0899 (0.187) ²	1.67049(0.758) ²
BM	1.9047(0.827) ³	2.7020 (0.929) ³	3.1001 (0.949) ³	1.625(0.947) ³ 1

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

TABLEAU 3.7 – Résultats pour le modèle HGARCH(1,d,1) issu d’une loi de Student à 7 degrés de liberté (Student’s t_7) avec 5% d’outliers additives et $e = 3$.

Résultats d’estimation						
n		γ	β	δ	ω	d
1000	QMV	0.2232684	0.2433654	0.2671698	0.2210327	0.2967160
<i>RMSE</i>	M_m	0.1740031	0.2557992	0.2585510	0.2280809	0.3082918
	BM	0.1383870	0.2353445	0.2590069	0.2230483	0.2678836
1000	QMV	0.1924815	0.2009800	0.1942890	0.1751197	0.2565371
<i>MAE</i>	M_m	0.1403866	0.2169373	0.1874856	0.1822322	0.2664562
	BM	0.1031449	0.1962265	0.1877156	0.1771517	0.2281657
2000	QMV	0.2119749	0.2238317	0.2018728	0.1905110	0.2622385
<i>RMSE</i>	M_m	0.1612009	0.2490745	0.2023851	0.1830490	0.2956904
	BM	0.1122019	0.2300487	0.1922464	0.1909857	0.2535307
2000	QMV	0.1895571	0.1843541	0.1570523	0.1496991	0.2202929
<i>MAE</i>	M_m	0.1370239	0.2132900	0.1515922	0.1438609	0.2518536
	BM	0.0885643	0.1935508	0.1514619	0.1503139	0.2133295
4000	QMV	0.2040775	0.1984260	0.1788925	0.1685548	0.2155104
<i>RMSE</i>	M_m	0.1472668	0.2232871	0.1630939	0.1544880	0.2670434
	BM	0.0979746	0.2008141	0.1596105	0.1756070	0.2147209
4000	QMV	0.1908384	0.1646101	0.1424864	0.1348654	0.1740652
<i>MAE</i>	M_m	0.1303340	0.1890230	0.1263907	0.1188400	0.2257310
	BM	0.0808824	0.1690851	0.1273080	0.1430466	0.1747903
Résultat pour la prévision						
	horizon 1	horizon 5	horizon 10	horizon 20		
n=200						
QMV	10.524(0.121) ¹	2.0534 (0.456) ¹	2.0960 (0.612) ¹	3.5490 (0.831) ¹		
M_m	10.764(0.001) ²	2.0183 (0.100) ²	1.9765 (0.199) ²	2.4867 (0.246) ²		
BM	11.007(0.004) ³	2.0930 (0.063) ³	2.1170 (0.074) ³	3.7757 (0.017) ³		
n=400						
QMV	7.6929(0.782) ¹	2.4893 (0.688) ¹	2.9023 (0.780) ¹	2.2268(0.825) ¹		
M_m	7.8378(0.767) ²	2.5061 (0.639) ²	2.8753 (0.706) ²	1.9196(0.942) ³		
BM	7.8162(0.799) ³	2.5041 (0.743) ³	2.8886 (0.814) ³	2.0495(0.876) ³		
n=800						
QMV	11.306(0.825) ¹	1.6450 (0.852) ¹	1.7670 (0.887) ¹	2.1912(0.979) ¹		
M_m	11.5541(0.243) ²	1.5575 (0.473) ²	1.5962 (0.623) ²	1.9012(0.898) ²		
BM	11.410(0.344) ³	1.5844 (0.330) ³	1.7198 (0.403) ³	2.1591(0.882) ³		

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

TABLEAU 3.8 – Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 5% d'outliers additives et $e = 5$.

Résultat d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.5648307	0.2499760	0.5356514	0.2686789	0.3028719
<i>RMSE</i>	M_m	0.1521896	0.2985108	0.2123590	0.2748437	0.3057135
	BM	0.07160839	0.21086305	0.18447847	0.17551215	0.24892119
1000	QMV	0.5165514	0.2111579	0.4531793	0.2220389	0.2586513
<i>MAE</i>	M_m	0.1208892	0.2630725	0.1542754	0.2491959	0.2612923
	BM	0.0517626	0.1744124	0.1451344	0.1378828	0.2093185
2000	QMV	0.5638567	0.2559949	0.5365817	0.2427489	0.3114296
<i>RMSE</i>	M_m	0.1326667	0.2951217	0.1584158	0.2523152	0.3046324
	BM	0.05040214	0.19104861	0.15281403	0.15767507	0.21891329
2000	QMV	0.5223559	0.2201130	0.4617962	0.1972170	0.2711683
<i>MAE</i>	M_m	0.1069465	0.2616240	0.1195407	0.2315005	0.2635410
	BM	0.0375038	0.1561707	0.1213493	0.1253809	0.17851517
4000	QMV	0.5360030	0.2449803	0.5214461	0.2142798	0.3337353
<i>RMSE</i>	M_m	0.1117087	0.2877303	0.1437599	0.2397087	0.2706203
	BM	0.03662889	0.15565170	0.11949934	0.13285513	0.16512541
4000	QMV	0.5022896	0.2099410	0.4543578	0.1726850	0.2965178
<i>MAE</i>	M_m	0.0921130	0.2570777	0.1114871	0.2268937	0.2323010
	BM	0.0279544	0.1276121	0.0980505	0.1136700	0.1294532
RMSE pour la prévision						
	horizon 1	horizon 5	horizon 10	horizon 20		
n=200						
QMV	19.7853(0.115) ¹	1.7169 (0.390) ¹	2.9207 (0.504) ¹	84.904 (0.744) ¹		
M_m	20.320(0.204) ²	1.1987 (0.239) ²	1.5002 (0.279) ²	2.0512 (0.787) ²		
BM	20.159(0.115) ³	1.2299 (0.390) ³	1.4853 (0.504) ³	2.0723 (0.744) ³		
n=400						
QMV	23.1851(0.084) ¹	2.8283 (0.378) ¹	3.4013 (0.498) ¹	20.5635(0.744) ¹		
M_m	23.8281(0.168) ²	2.6667 (0.224) ²	3.0344 (0.393) ²	2.4068(0.168) ²		
BM	23.663809(0.081) ³	2.68182 (0.3758) ³	3.0569 (0.496) ³	2.1355(0.739) ³		
n=800						
QMV	149.31(0.604) ¹	1.6908 (0.687) ¹	4.2449 (0.772) ¹	8.6097(0.928) ¹		
M_m	153.59(0.378) ²	1.4169 (0.462) ²	1.7551 (0.568) ²	2.0146(0.778) ²		
BM	150.57(0.771) ³	2.4810 (0.856) ³	2.4128 (0.903) ³	3.2987(0.973) ³		

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

TABLEAU 3.9 – Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 10% d'outliers additives et $e = 3$.

Résultats d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.4169460	0.3011928	0.3134388	0.2586651	0.3063438
<i>RMSE</i>	M_m	0.3609773	0.3102409	0.3444614	0.3071906	0.3333049
	BM	0.3340910	0.3042965	0.3711966	0.2880218	0.3035411
1000	QMV	0.3661369	0.2610087	0.2267887	0.2114776	0.2655122
<i>MAE</i>	M_m	0.3019093	0.2706237	0.2495066	0.2660483	0.2899952
	BM	0.2771095	0.2636659	0.2746938	0.2454367	0.2610387
2000	QMV	0.3930570	0.2951317	0.2241389	0.2276060	0.3226785
<i>RMSE</i>	M_m	0.3418088	0.3291794	0.2572199	0.2657475	0.3341221
	BM	0.3167903	0.3106951	0.2843631	0.2572686	0.3087179
2000	QMV	0.3557297	0.2589721	0.1682103	0.1804710	0.2938298
<i>MAE</i>	M_m	0.2801584	0.2975768	0.1897024	0.2260184	0.2934769
	BM	0.2672879	0.2751862	0.1989421	0.2107103	0.2696851
4000	QMV	0.3685170	0.2855625	0.1699708	0.1920961	0.3269855
<i>RMSE</i>	M_m	0.3096062	0.3378079	0.1963209	0.2261551	0.3477793
	BM	0.2824697	0.3114089	0.2060521	0.2226867	0.3293762
4000	QMV	0.3437049	0.2540167	0.1341495	0.1549976	0.3089090
<i>MAE</i>	M_m	0.2507114	0.3095731	0.1445923	0.1905236	0.3039645
	BM	0.2435649	0.2811718	0.1434312	0.1776148	0.2950128
RMSE pour la prévision						
	horizon 1	horizon 5	horizon 10	horizon 20		
n=200						
QMV	7.9437(0.007) ¹	2.3832 (0.129) ¹	2.2573 (0.268) ¹	8.0016 (0.776) ¹		
M_m	8.4704(0.516) ²	2.0927 (0.543) ²	1.4930 (0.271) ²	7.5822 (0.932) ²		
BM	8.3580(0.002) ³	2.1473(0.103) ³	1.6014 (0.249) ³	7.6327 (0.745) ³		
n=400						
QMV	8.9515(0.503) ¹	2.4550 (0.406) ¹	3.2298 (0.503) ¹	2.1013(0.893) ¹		
M_m	9.2804(0.412) ²	2.4418 (0.244) ²	3.1948 (0.412) ²	1.8368 (0.795) ²		
BM	9.2363(0.554) ³	2.4460 (0.4868) ³	3.1899 (0.554) ³	1.9145(0.913) ³		
n=800						
QMV	8.1438(0.644) ¹	2.0040 (0.493) ¹	2.0690 (0.644) ¹	2.2298(0.929) ¹		
M_m	8.3414(0.413) ²	1.9390 (0.204) ²	1.9678 (0.413) ²	2.0503(0.830) ²		
BM	8.2981(0.790) ³	1.9316 (0.708) ³	1.9924 (0.790) ³	2.0941(0.963) ³		

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

TABLEAU 3.10 – Résultats pour le modèle HGARCH(1,d,1) issu d'une loi de Student à 7 degrés de liberté (Student's t_7) avec 10% d'outliers additives et $e = 5$.

Résultat d'estimation						
n		γ	β	δ	ω	d
1000	QMV	0.6860862	0.3554409	0.3890850	0.2438352	0.3200711
<i>RMSE</i>	M_m	0.1999738	0.3024286	0.2161062	0.3468562	0.3238086
	BM	0.09285297	0.22338258	0.19531870	0.16254759	0.25989716
1000	QMV	0.5052916	0.3204261	0.3180892	0.2026061	0.2640737
<i>MAE</i>	M_m	0.1407247	0.2678568	0.1627060	0.3265580	0.2790750
	BM	0.0578108	0.1891003	0.1521024	0.1212633	0.2194328
2000	QMV	0.6635013	0.3786370	0.3438742	0.2097120	0.2765339
<i>RMSE</i>	M_m	0.1281592	0.3177094	0.1558212	0.3321473	0.2976933
	BM	0.05196373	0.20082194	0.13942675	0.11535523	0.22756159
2000	QMV	0.4481679	0.3505733	0.2886819	0.1708698	0.2237653
<i>MAE</i>	M_m	0.0898310	0.2870026	0.1236213	0.3165158	0.2564725
	BM	0.0375878	0.1691607	0.1161633	0.0874029	0.1890520
4000	QMV	0.6921490	0.3924226	0.3213152	0.1795246	0.2494398
<i>RMSE</i>	M_m	0.09497276	0.32971093	0.13290190	0.33440822	0.26983722
	BM	0.03520161	0.18135530	0.11295034	0.08119154	0.19342388
4000	QMV	0.445390	0.371323	0.2802407	0.1472661	0.1974306
<i>MAE</i>	M_m	0.0668671	0.3007275	0.1077316	0.3241814	0.2302379
	BM	0.0286096	0.1533838	0.0923323	0.0638807	0.15688908

RMSE pour la prévision				
	horizon 1	horizon 5	horizon 10	horizon 20
n=200				
QMV	18.541(0.092) ¹	2.9111 (0.357) ¹	3.6467 (0.507) ¹	8.1302 (0.841) ¹
M_m	19.076(0.067) ²	2.5303 (0.008) ²	2.7537 (0.170) ²	2.7898 (0.499) ²
BM	18.979(0.072) ³	2.5066 (0.336) ³	2.7527 (0.489) ³	2.7908 (0.824) ³
n=400				
QMV	18.530(0.838) ¹	2.5925 (0.718) ¹	2.4935 (0.838) ¹	3.7704(0.965) ¹
M_m	18.957(0.126) ²	2.2340 (0.029) ²	1.7104 (0.126) ²	1.4589(0.369) ²
BM	18.919(0.985) ³	2.2260 (0.975) ³	1.6567 (0.985) ³	1.4905(0.997) ³
n=800				
QMV	18.625(0.351) ¹	1.9258 (0.152) ¹	3.1682 (0.351) ¹	2.5572(0.727) ¹
M_m	19.0333(0.232) ²	1.4154 (0.183) ²	2.7993 (0.232) ²	1.5179(0.477) ²
BM	19.085(0.382) ³	1.4163 (0.176) ³	2.7881 (0.382) ³	1.5448(0.786) ³

Notes : (...) ¹ représente la p-value du test de Diebold-Mariano (DM-test) pour le QMV et le M_m, (...) ² représente la p-value du test de Diebold-Mariano (DM-test) pour le M_m et le BM, (...) ³ représente la p-value du test de Diebold-Mariano (DM-test) pour le BM et le QMV.

3.2 Applications aux données du S&P 500

Dans cette partie, nous présentons les résultats d'application de notre étude théorique à des données réelles. Notre objectif est de comparer les performances des M et BM-estimateurs proposées dans cette étude à celles de l'estimateur du quasi-maximum de vraisemblance concernant la prévision de la volatilité sur un exemple de données réelles.

3.2.1 Description et analyse des données

Nous nous intéressons à l'indice journalier du Standard and Poor's 500 (S&P 500). La période d'observation s'étend du 01 Janvier 1990 au 31 Décembre 2007. Ce qui fait un total de 4536 valeurs observées pour le S&P 500. Les études et analyses sur ces données ont été réalisées sur le logiciel R. Nous présentons à travers la figure 3.1 la série brute des observations.

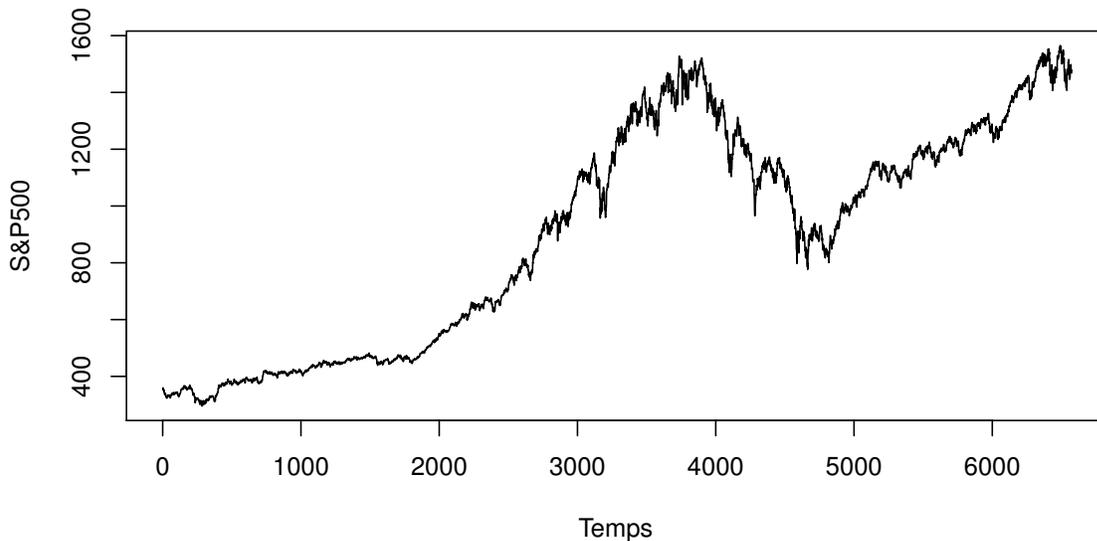


FIGURE 3.1 – Série brute de l'indice journalier du S&P 500

De plus, en désignant par y_t l'observation du S&P 500 à la date t , le log-rendement r_t à la date t est défini par

$$r_t = \ln y_t - \ln y_{t-1}. \quad (3.10)$$

Dans le but de recueillir les premières impressions concernant les données étudiées, nous procédons, dans la partie suivante, à l'analyse des données des log-rendements du S&P

500 comme dans Lux et al. [50]. Le tableau **3.11** présente les éléments de statistiques descriptives de la série des log-rendements journaliers du S&P 500.

TABLEAU 3.11 – Statistique descriptive des log-rendement du S&P 500

Moyenne	Ecart type	Skewness	Kurtosis	Min	Max	JBT	Arch-T
0.0003	0.0099	-0.1243	3.7597	-0.0711	0.0557	2687.6 (0.0000)	519.76 (0.0000)

Notes : JBT désigne le test de Jarque-Bera, Arch-T désigne le test de l'effet ARCH, les p -values sont données entre parenthèses.

Ces éléments de statistiques descriptives fournissent un certain nombre d'indicateurs permettant d'expliquer la distribution de la série. Nous remarquons dans cette série que la moyenne est petite et proche de zéro. Ce qui traduit le fait qu'en moyenne, les données tournent autour de 0. En outre, le skewness ou coefficient d'asymétrie évalué à -0.1243 est différent de zéro (la valeur du skewness d'une loi normale). Ce coefficient négatif traduit une distribution asymétrique, donc une queue de distribution de la série des log-rendements étalée vers la gauche. Le kurtosis ou coefficient d'aplatissement est de 3.7597 supérieur au kurtosis de la loi normale évalué à 3. Ce qui traduit le fait que la distribution des log-rendements est plus élevée que celle de la loi normale et retombe assez brutalement par rapport à la distribution de la loi normale. On dira de la distribution des $\{r_t\}$ qu'elle est à queue lourde ou leptokurtique. Ces indicateurs nous donnent de suspecter que la distribution de la série des log-rendements $\{r_t\}$ n'est pas normale. Pour vérifier l'hypothèse de non-normalité de la série des log-rendements, nous appliquons le test de normalité de Jarque-Bera. L'hypothèse nulle de ce test est la normalité de la distribution. Le résultat fourni au tableau **3.11** révèle une p -value inférieure à 0.05, donc un rejet de l'hypothèse nulle. Par conséquent, nous pouvons affirmer que la distribution des log-rendements $\{r_t\}$ n'est pas normale. Nous observons à la figure **3.2** une période de forte volatilité entre les instants 1 et 500 suivie d'une période de faible volatilité entre 500 et 1800 encore suivie d'une forte volatilité de 1800 à 3250 etc... En d'autres termes, nous pouvons déduire de la tendance de ce graphique, une alternance de forte et faible variabilité. Cette alternance de variabilité traduit le phénomène de clustering de la volatilité qui est une caractéristique de l'hétéroscédasticité conditionnelle. La figure **3.2**, nous permet donc de suspecter la présence d'hétéroscédasticité conditionnelle dans la série.

Nous présentons par ailleurs à la figure **3.2**, la série des log-rendements journalier $\{r_t\}$ et à la figure **3.3** la boîte à moustache (boxplot) de $\{r_t\}$ pour l'indice S&P 500.

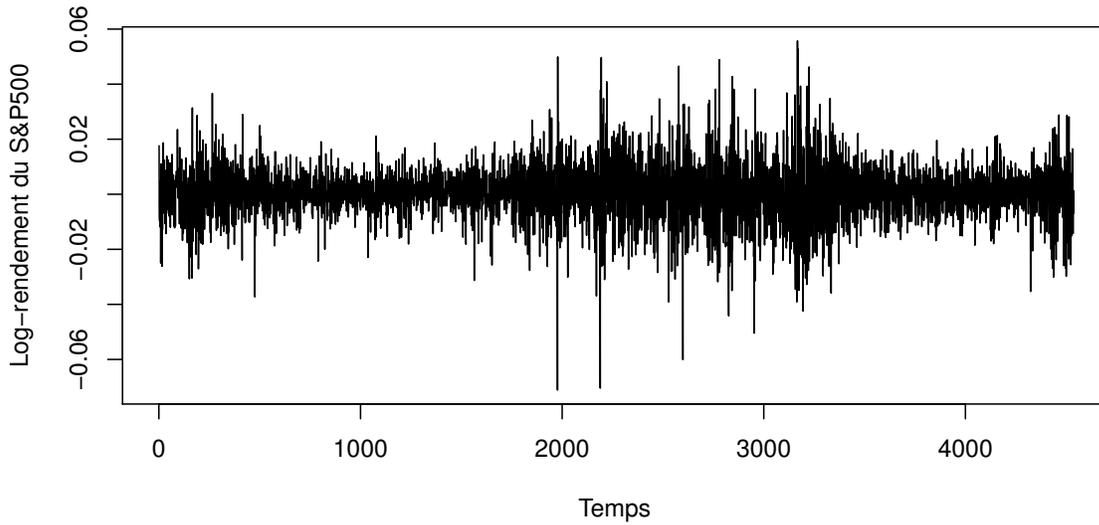


FIGURE 3.2 – Série des log-rendements du S&P 500

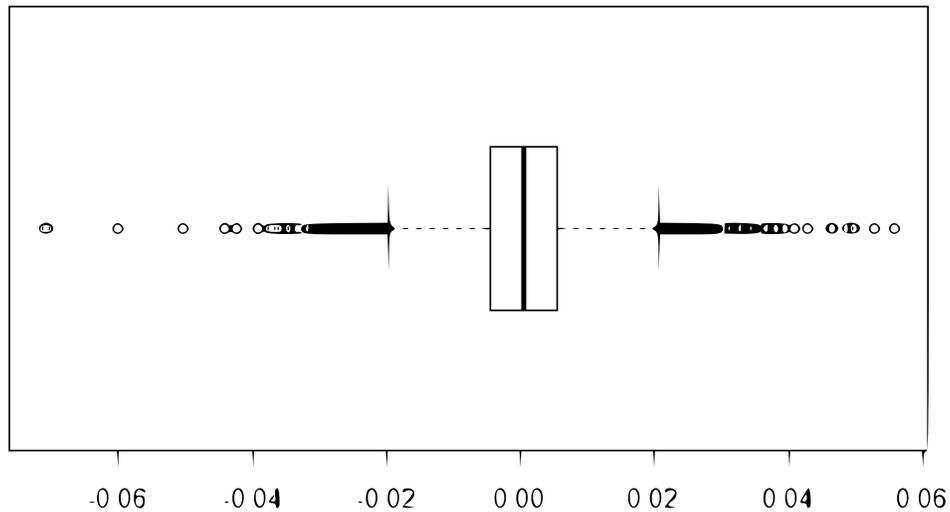


FIGURE 3.3 – Boîte à moustache des log-rendements du S&P 500

Par cette remarque, nous procédons comme dans Vee et al. [64] afin de vérifier la présence ou l'absence de l'hétéroscédasticité conditionnelle dans la série à travers un test disponible à cet effet. Le tableau 3.11 fournit également la valeur de la p-value du test ARCH que nous avons de plus réalisé sur les données de $\{r_t\}$. Ces valeurs ont été obtenues en utilisant la fonction "ArchTest" définie dans le package FinTS du logiciel R. En effet, le test de Multiplicateur de Lagrange ou test d'effet ARCH permet de tester

- H_0 : Absence d'hétéroscédasticité conditionnelle contre
- H_1 : présence d'hétéroscédasticité conditionnelle.

Au regard des valeurs obtenues du logiciel, au seuil de 5%, la valeur de la statistique de test évaluée à 519.76 est supérieure à la valeur critique évaluée à 21.03. En effet, cette valeur critique correspond au quantile d'ordre $\alpha = 5\%$ d'un Khi-deux à 12 degrés de libertés. Ainsi, au seuil de 5%, nous rejetons l'hypothèse nulle. En d'autres termes, il y'a une présence d'hétéroscédasticité conditionnelle dans la série $\{r_t\}$ au seuil de 5%. Ce résultat confirme donc les remarques soulevées dans la figure **3.2**. Par conséquent, pour l'ajustement de ces données, il serait intéressant de penser aux modèles à hétéroscédasticité conditionnelle de type ARCH, GARCH...

Nous réalisons également deux tests sur la série $\{r_t\}$ pour étudier la stationnarité de cette série. Il s'agit des tests de Dickey-Fuller Augmenté (ADF-test) et de racine unitaire de Phillips Perron (PP-test) contenu dans le package "tseries" de R. Ces deux tests ont les mêmes hypothèses à savoir

- H_0 : non stationnarité de la série contre
- H_1 : stationnarité de la série

Les résultats sortis du logiciel montrent que la série est stationnaire. En effet, pour le ADF-test, la p-value est de 0.01 inférieur au niveau de significativité de 5% avec une statistique de test d'une valeur de -16.355 . Ce qui nous donne de rejeter l'hypothèse nulle donc d'accepter l'hypothèse de stationnarité de la série au seuil de 5%. Le PP-test conduit également à la même conclusion. Ce qui traduit le fait qu'un processus stationnaire serait adapté à ces données.

Dans le paragraphe suivant, nous présentons des figures qui nous permettrons d'étudier la présence de la longue mémoire dans les données. A cet effet, nous présentons dans la figure 3.4, le graphe de la fonction d'autocovariance de la série des log-rendements du S&P 500.

Nous remarquons dans la figure 3.4 que toutes les variations sont à l'intérieur de l'intervalle de confiance et varient autour de zéro. Par contre, dans le cas de la figure 3.5 qui montre le graphe de la fonction d'autocovariance de la série $\{r_t^2\}$, tous les pics sont à l'extérieur de l'intervalle de confiance.

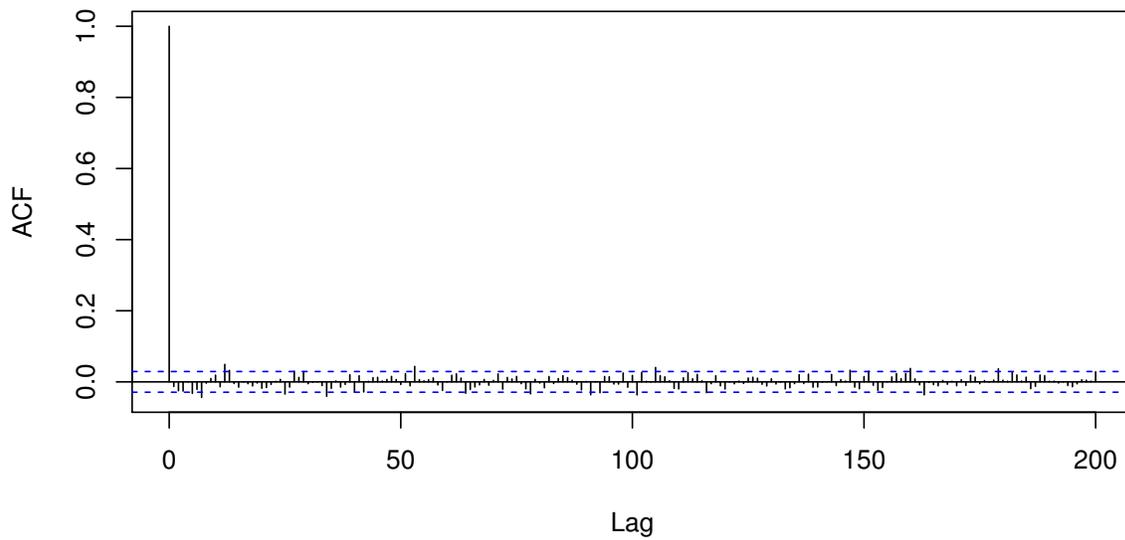


FIGURE 3.4 – ACF des log-rendements du S&P 500

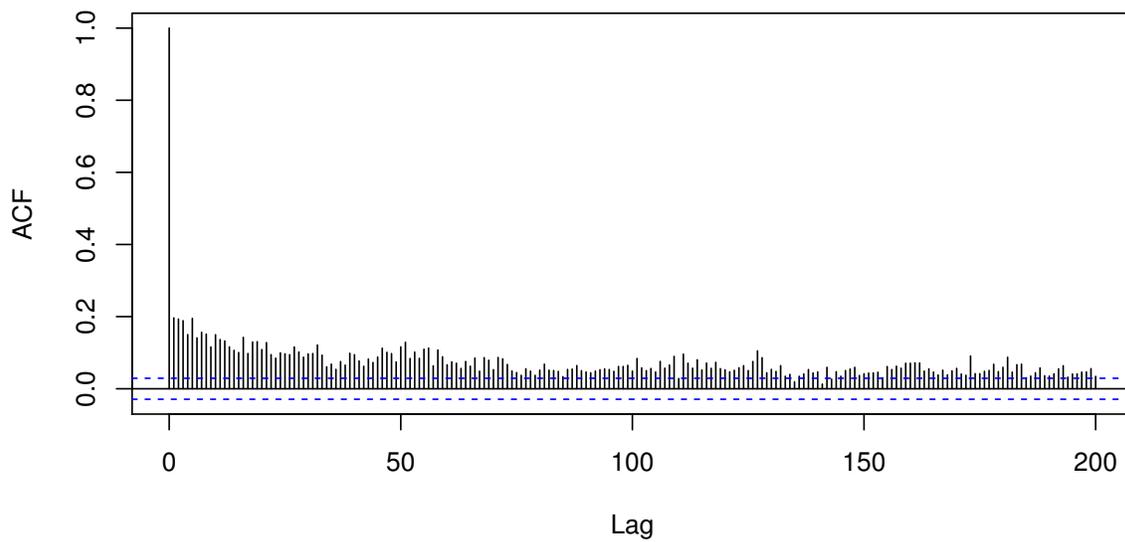


FIGURE 3.5 – ACF du carré des log-rendements $\{r_t^2\}$ du S&P 500

De plus, nous observons dans le cas de la figure 3.5, une sorte de décroissance de la

représentation de l'ACF. Selon l'étude menée par Mootamri [53], cette décroissance pourrait nous faire penser à la présence de longue mémoire dans la variance du processus. Pour s'en assurer, nous testerons la présence de longue mémoire dans les séries $\{r_t\}$ et $\{r_t^2\}$. Nous appliquons donc le "rescaled variance (V/S)" test défini dans Giratis et al. [32]. Ce test présente les hypothèses suivantes

- H_0 : mémoire courte contre
- H_1 : présence de mémoire longue.

Le test a été réalisé au niveau de significativité de 5% pour lequel la valeur critique est de 0.1869. Comme dans Kwan et al. [44], nous avons effectué une série de tests sous différentes valeurs de la bande passante q^* . Les résultats de ces tests sont consignés dans le tableau 3.12 suivant

TABLEAU 3.12 – V/S test pour la détection de la longue mémoire

$q^*=$	1	5	10	15	20	25	30
$\{r_t\}$	0.113	0.12273	0.13649	0.13693	0.13781	0.14151	0.14393
$\{r_t^2\}$	4.6551	2.8725	2.028	1.6101	1.3541	1.1781	1.0504

D'après le tableau **3.12**, toutes les valeurs des statistiques de tests pour la série des log-rendements sont inférieures à la valeur critique de 0.1869. Ceci est vérifié pour différentes valeurs de la bande passante q^* . Donc dans le cas de $\{r_t\}$, l'hypothèse nulle du test est acceptée, c'est-à-dire la série $\{r_t\}$ des log-rendements présente une mémoire courte au seuil de 5%. En ce qui concerne la série du carré des log-rendements $\{r_t^2\}$, pour différentes valeurs de la bande passante q^* , les valeurs des statistiques de tests sont toutes supérieures à la valeur critique de 0.1869. Ce qui traduit le rejet de l'hypothèse nulle. Ainsi, au seuil de 5%, la présence de longue mémoire est toujours suggérée dans la série des $\{r_t^2\}$. Par suite, nous pouvons dire que dans la série $\{r_t\}$, la longue mémoire est présente uniquement dans la variance du processus et non dans la série observée. Ce qui implique, comme dans Kwan et al. [44], qu'un modèle GARCH hyperbolique pur serait plus adapté pour l'ajustement de ces données. Après avoir analysé les données des log-rendements de la série soumise à notre étude, il en ressort trois caractéristiques majeures. Ces caractéristiques sont

- la stationnarité de la série,
- la présence d'effet ARCH dans la série,
- la présence de la longue mémoire uniquement dans la volatilité de la série.

3.2.2 Sélection du meilleur modèle pour la série des log-rendements du S&P500

Dans ce paragraphe, nous le meilleur modèle permettant d'ajuster la série des log-rendements du S&P500.

Au vu des caractéristiques précédentes, nous choisissons les modèles GARCH à mémoire hyperbolique (HYGARCH et HGARCH) pour ajuster les données. Pour un travail plus englobant, nous travaillerons également avec le modèle FIGARCH.

Pour avoir une idée des valeurs de p et q à utiliser dans nos modèles, nous proposons plusieurs modèles différents en faisant varier p et q . Puis, pour chacun des modèles proposés, nous déterminons les valeurs du critère d'information bayésien (BIC) correspondantes, comme dans Bisaglia [8]. Cette détermination du BIC se fera uniquement sur les modèle $HGARCH(p, d, q)$ pour différentes valeurs de p et q .

Le modèle ayant le plus petit BIC sera retenu et les valeurs de p et q correspondant à ce modèle seront également utilisées pour les autres modèles. Le calcul du BIC se fait de la manière suivante

$$BIC = -2 \log(L(\theta)) + k \log n,$$

avec n qui définit la taille de l'échantillon, k le nombre de paramètres à estimer et $L(\theta)$ la vraisemblance estimée en la valeur de l'estimation du quasi maximum de vraisemblance obtenu.

Le tableau 3.13 suivant donne les valeurs des BIC pour les différents modèles GARCH hyperbolique HGARCH(p,d,q).

TABLEAU 3.13 – Valeurs des BIC des log-rendement du S&P 500

HGARCH(p,d,q)	$p = 1$	$p = 2$	$p = 3$
$q = 1$	45.7778	54.07257	62.36357
$q = 2$	54.06925	62.36332	70.65738

D'après le tableau 3.13, la plus petite valeur du BIC des modèles $HGARCH(p, d, q)$ vaut 45.7778. Elle est obtenue pour $p = 1$ et $q = 1$. Dans la suite du travail, nous considérerons donc le modèle $HGARCH(1, d, 1)$ pour la série des log-rendements du S&P500.

3.2.3 Evaluation de différents estimateurs pour l'estimation et la prévision de la volatilité

Dans cette partie, nous comparons les performances de prévision de la volatilité des trois estimateurs utilisés dans les simulations sur les modèles sélectionnés pour la série $\{r_t\}$.

L'analyse de la trajectoire de la série des log-rendements à la figure 3.2 et la boîte à moustache de la série des log-rendements représentée à la figure 3.3 indiquent clairement qu'il y'a certaines valeurs qui s'écartent de la majorité des données délimitée par les moustaches de la boîte. Par ailleurs, en utilisant la méthode du quartile exploitée dans Vinutha et al. [65], nous évaluons approximativement à 237, le nombre d'observations qui semblent dévier de la majorité des données de la série.

Dans le but de détecter avec plus de précision, le nombre d'outliers dans la série des log-rendement $\{r_t\}$, nous utilisons la fonction "locate.outliers" du package "tsoutliers" du logiciel R. Cette fonction, basée sur les travaux de Chen et Liu [14], permet de localiser les outliers dans des données de séries temporelles. Elle applique les statistiques de Student pour la signification des outliers à chaque instant et sélectionne celles qui sont significatives compte tenu d'une valeur critique. Ainsi, en appliquant la fonction "locate.outliers", le nombre d'outliers dans la série $\{r_t\}$ est de 189 soit 4.16% des observations.

Pour réaliser notre étude, les 4000 premières observations de $\{r_t\}$ sont utilisées pour estimer les paramètres des modèles et les 536 observations restantes seront utilisées pour la prévision hors échantillon. En utilisant également la fonction "locate.outliers", nous trouvons 153 outliers dans les 4000 observations de l'échantillon d'estimation, soit 3.82%. Cette information rend favorable l'utilisation du BM-estimateur et du M-estimateur sur ce type de données.

Dans le paragraphe suivant, nous estimons les paramètres des modèles FIGARCH(1,d,1), HYGARCH(1,d,1) et HGARCH(1,d,1) par l'estimateur du quasi-maximum de vraisemblance (QMV), le M-estimateur (M_m) et le BM-estimateur (BM). Les résultats de ces estimations sont donnés dans le tableau **3.14** ci-dessous :

TABLEAU 3.14 – Paramètres estimés des modèles FIGARCH(1,d,1), HYGARCH(1,d,1) et HGARCH(1,d,1) pour la série des log-rendements journaliers du S&P 500

		Estimation					
		γ	β	δ	ω	ϕ	d
FIGARCH	QMV	2.937e-06 (1.642e-09)	0.572055 (27.7340)	0.157408 (9.1042)	-	-	0.462010 (22.9565)
	M_m	9.108e-07 (1.6794e-09)	0.564157 (31.1347)	0.202621 (15.6541)	-	-	0.379969 (21.5788)
	BM	1.792e-06 (2.072e-09)	0.666034 (34.3124)	0.116032 (7.93528)	-	-	0.537644 (29.5184)
HYGARCH	QMV	2.999e-06 (2.512e-09)	0.573681 (27.7576)	0.157355 (9.09769)	0.997883 (48.2533)	-	0.464701 (23.1209)
	M_m	2.107e-06 (3.575e-09)	0.592762 (24.6349)	0.188374 (9.55433)	0.953799 (43.7134)	-	0.438298 (19.2540)
	BM	1.854e-06 (2.157e-09)	0.671324 (33.7348)	0.115029 (7.82908)	0.996961 (30.5291)	-	0.545684 (29.4939)
HGARCH	QMV	3.338e-06 (2.098e-09)	0.583783 (28.2041)	0.154509 (9.09820)	0.970660 (95.5484)	-	0.477609 (24.0711)
	M_m	2.019e-06 (3.344e-09)	0.605768 (27.5468)	0.184486 (10.9450)	0.910084 (91.2994)	-	0.436593 (20.3980)
	BM	1.887e-06 (2.059e-09)	0.676744 (34.0157)	0.112767 (7.75440)	0.986514 (85.1323)	-	0.551590 (30.1549)

Notes : Le nombre entre parenthèse représente la valeur de la statistique de Student "t-statistic" pour les paramètres estimés.

Le tableau **3.14** présente les valeurs des paramètres estimés obtenues par application des trois méthodes d'estimation sur les modèles FIGARCH, HYGARCH et HGARCH. Pour chaque modèle, nous notons une différence entre les valeurs obtenues par chacune des trois méthodes d'estimation. Cette différence peut être expliquée par la présence des outliers dans les données.

Le tableau **3.15** montre les résultats du RMSE (Root Mean Square Error) et du MAE (Mean Absolute Error) concernant les prévisions de la volatilité en utilisant les trois méthodes d'estimation.

TABLEAU 3.15 – Résultats de prévision des modèles FIGARCH(1,d,1), HYGARCH(1,d,1) et HGARCH(1,d,1) pour les log-rendements du S&P 500

Methods	BM	M_m	QMV
FIGARCH			
RMSE	4.076e+59	5.001e+27	1.997e+36
MAE	4.742e+58	8.255e+26	2.918e+35
HYGARCH			
RMSE	15.52145e-5	14.85304e-5	15.85435e-5
MAE	12.19993e-5	10.95313e-5	12.76183e-5
HGARCH			
RMSE	15.49579e-5	14.97837e-5	15.70795e-5
MAE	12.15469e-5	11.21154e-5	12.52305e-5

Pour les prévisions, nous remarquons que pour les modèles HYGARCH et HGARCH, les valeurs du RMSE et du MAE pour le M-estimateur (M_m) sont plus faibles que celles du BM-estimateur et du QMV. Nous pouvons donc conclure que pour la prévision hors échantillon des modèles HYGARCH et HGARCH, le M-estimateur a la meilleure performance et le QMV la pire. Le M-estimateur est donc préférable pour la prévision des modèles HYGARCH et HGARCH. Cela confirme les résultats obtenus dans les études de simulations effectuées dans ce travail.

De plus, nous notons que le modèle HGARCH affiche de plus petites valeurs de RMSE et MAE par rapport au HYGARCH pour le BM estimateur et le QMV. Par conséquent, le modèle HGARCH a donc une bonne précision de prévision que le modèle HYGARCH pour le BM estimateur et le QMV. Cependant, en observant le comportement du M-estimateur, le modèle HYGARCH est le meilleur. Pour toutes les méthodes d'estimation, nous constatons que les modèles HYGARCH et HGARCH ont de bien meilleures performances que le modèle FIGARCH pour nos données. Cette remarque s'explique certainement par l'existence de variance finie des données qui est une caractéristique que ne possède pas le modèle FIGARCH.

Chapitre 4

Propriétés asymptotiques de l'estimateur du Quasi-maximum de vraisemblance d'une classe de modèle GARCH hyperbolique multivarié

Dans ce chapitre, nous proposons un modèle HGARCH multivarié appelé le modèle GARCH hyperbolique à corrélation conditionnelle constante noté (CCC-HGARCH). Ce modèle présente de nombreux avantages. Il permet de contrôler le nombre de paramètres à estimer, ce qui le rend parcimonieux. De plus, la propriété relative à la stationnarité stricte du CCC-HGARCH s'obtient sous certaines conditions relativement simple. Ce modèle offre également une flexibilité dans les applications et pour l'estimation numérique des paramètres. Le principal objectif de ce travail est d'étudier la stationnarité du modèle et les propriétés asymptotiques (consistance et normalité asymptotique) de l'estimateur du quasi-maximum de vraisemblance utilisé pour l'estimation des paramètres du modèle proposé. En effet, bien qu'étant très sensible à la présence de valeurs aberrantes dans les observations d'après le chapitre 3, l'estimateur du quasi-maximum de vraisemblance demeure très populaire en raison de ses bonnes propriétés asymptotiques et de sa facilité d'exécution numérique. Notre travail dans cette partie sera donc purement théorique, soutenu par des simulations numériques. Les résultats obtenus dans cette partie constitueront une référence pour des perspectives de recherches concernant l'étude d'autres méthodes d'estimations pour les modèles

GARCH hyperbolique multivariés.

4.1 Présentation du modèle HGARCH multivarié

Dans cette partie, nous présentons le modèle HGARCH (p, d, q) multivarié que nous étudierons par la suite. Ce modèle est conçu par extension du HGARCH (p, d, q) univarié.

4.1.1 Modèle GARCH hyperbolique univarié

Le modèle HGARCH (p, d, q) univarié introduit par Li et al.[46], étudié au chapitre 2, a la paramétrisation suivante :

$$y_t = h_t^{1/2} \epsilon_t \quad (4.1)$$

avec

$$h_t = \alpha + \beta(L)h_t + \omega(1 - \beta(L) - \delta(L)(1 - L)^d)y_t^2. \quad (4.2)$$

où ϵ_t est un processus i.i.d de moyenne 0 et de variance 1. Les paramètres du modèle sont tels que $\gamma, \omega > 0$ et $\delta_i, \beta_j \geq 0$ pour $i = 1, \dots, p$ et $j = 1, \dots, q$, avec $0 < d < 1$.

La paramétrisation ARCH(∞) du modèle HGARCH est donnée par

$$\begin{aligned} h_t(\theta) &= \frac{\alpha}{1 - \beta(1)} + \omega \left(1 - \frac{(1 - \delta(L))(1 - L)^d}{1 - \beta(L)}\right) y_t^2, \\ &= \frac{\alpha}{1 - \beta(1)} + \sum_{j=1}^{\infty} b_j(\theta) y_{t-j}^2, \end{aligned} \quad (4.3)$$

où les b_j sont positifs et on a $b_j(\theta) = O(j^{-1-d})$ pour tous $j \geq 1$.

La variance conditionnelle (4.2) peut se réécrire comme suit :

$$\begin{aligned} h_t &= \alpha + \beta(L)h_t + \omega(1 - \beta(L) - \delta(L)(1 - L)^d)y_t^2 \\ &= \alpha + \sum_{j=1}^p \beta_j h_{t-j} + \omega \left\{1 - \sum_{j=1}^p \beta_j L^j - (1 - \sum_{i=1}^q \delta_i L^i)(1 - L)^d\right\} y_t^2 \end{aligned} \quad (4.4)$$

$$h_t = \alpha + \sum_{j=1}^p \beta_j h_{t-j} + (1 - \beta(L) - \delta(L)(1 - L)^d)\omega y_t^2 \quad (4.5)$$

$$\left(1 - \sum_{j=1}^p \beta_j L^j\right) h_t = \alpha + \left((1 - \beta(L))\omega - (\delta(L)(1 - L)^d)\omega\right) y_t^2.$$

4.1.2 Le modèle HGARCH à corrélation conditionnelle constante

Soit $\underline{r}_t \in \mathbb{R}^N$, le vecteur représentant l'extension du modèle (4.1)-(4.2) au cas multivarié et \mathcal{I}_{t-1} l'ensemble des informations générées par le passé de \underline{r}_t jusqu'à la date $t - 1$. De plus, nous avons $\mathbb{E}(\underline{r}_t | \mathcal{I}_{t-1}) = 0$ et $Var(\underline{r}_t | \mathcal{I}_{t-1}) = \mathbb{E}(\underline{r}_t \underline{r}_t' | \mathcal{I}_{t-1}) = H_t$. En nous inspirant des travaux de Bauwens et al. [5] et Silvennoinen et al. [59], nous proposons une spécification particulière pour les modèles GARCH hyperboliques multivariés. Il s'agit du modèle GARCH hyperbolique à corrélation conditionnelle constante appelé modèle CCC-HGARCH. Il se présente comme suit

$$\underline{r}_t = H_t^{1/2}(\theta)\underline{\varepsilon}_t \quad (4.6)$$

$$H_t = D_t P D_t, \quad D_t = \text{diag}(\sqrt{h_{11t}}, \sqrt{h_{22t}}, \dots, \sqrt{h_{NNt}}) = \text{diag}(\underline{h}_t^{1/2}) \quad (4.7)$$

$$P = (\rho_{ij})_{1 \leq i, j \leq N} \quad \text{avec} \quad \rho_{ii} = 1 \quad \text{et} \quad \rho_{ij} = \rho_{ji} \quad i \neq j, \quad -1 \leq \rho_{ij} \leq 1, \quad (4.8)$$

avec

$$\underline{h}_t = \underline{\alpha} + \sum_{j'=1}^p \mathbf{B}_{j'} \underline{h}_{t-j'} + \left[(I_N - \sum_{j'=1}^p \mathbf{B}_{j'} L^{j'}) - (I_N - \sum_{i'=1}^q \Delta_{i'} L^{i'}) (I - L)^d \right] \underline{\Omega} \underline{\zeta}_t. \quad (4.9)$$

Le processus $\underline{\varepsilon}_t$ est une suite de vecteur i.i.d de \mathbb{R}^N pour laquelle $\mathbb{E}(\underline{\varepsilon}_t) = 0$, $Var(\underline{\varepsilon}_t) = \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t') = I_N$.

Nous avons $\underline{h}_t = (h_{11t}, h_{22t}, \dots, h_{NNt})$, où le processus h_{iit} pour $i = 1, \dots, N$ correspond à la variance conditionnelle dans le cas univarié définie en (4.4) – (4.5). Comme dans He et Terasvirta [36], nous notons que \underline{h}_t défini en (4.9) désigne le vecteur de dimension N dont les composantes sont les carrés des éléments de D_t .

Les matrices $\underline{\Omega}$ et $(I - L)^d$ sont des matrices diagonales définies par $\underline{\Omega} = \text{diag}(\omega_i)$, avec $0 < \omega_i < 1$ pour $i = 1, \dots, N$ et $(I - L)^d = \text{diag}((1 - L)^{d_i})$ avec $0 < d_i < 1$ pour $i = 1, \dots, N$, le vecteur $\underline{\zeta}_t = (r_{1t}^2, r_{2t}^2, \dots, r_{Nt}^2)'$. $\underline{\alpha}$ est un vecteur de taille N dont les coefficients α_i , pour $i = 1, \dots, N$, sont tous strictement positifs. Les matrices $\mathbf{B}_{j'}$ pour $j' = 1, \dots, p$ et $\Delta_{i'}$ pour $i' = 1, \dots, q$ sont des matrices diagonales d'ordre N dont les coefficients sont tous positifs.

Les coefficients de h_{iit} pour $i = 1, \dots, N$ vérifient les conditions en (4.3) telles que chaque élément de \underline{h}_t soit strictement positif pour chaque t . De plus, $\underline{\varepsilon}_t$ admet une densité positive absolument continue par rapport à la mesure de Lebesgue. $H_t = (h_{ijt})_{1 \leq i, j \leq N}$ est la matrice de covariance conditionnelle du processus \underline{r}_t . Par ailleurs, la matrice

de corrélation conditionnelle P est symétrique et invariante dans le temps. Tout ceci entraîne que

$$H_t = (h_{ijt}) = \begin{cases} (h_{iit}) & \text{si } i = j, \quad i, j = 1, \dots, N, \\ (h_{iit}^{1/2} h_{jjt}^{1/2} \rho_{ij}), & \text{si } i \neq j. \end{cases} \quad (4.10)$$

Ainsi, la matrice de covariance conditionnelle H_t est toujours symétrique définie positive lorsque P est définie positive. L'espace des paramètres Θ est un sous-ensemble compact de

$$]0, +\infty[^N \times [0, +\infty[^{(p+q)N} \times]0, 1[^N \times]-1, 1[^{N(N-1)/2} \times]0, 1[^N.$$

Un élément de Θ a la forme suivante

$$\theta = (\gamma_1, \dots, \gamma_N, \beta_{ii}, \delta_{ii}, \omega_{ii}, \rho_{ij}, d_{ii}) \quad \text{avec } 1 \leq i, j \leq N \quad \text{et } i \neq j. \quad (4.11)$$

4.1.3 Notations et définitions

I_N désigne la matrice identité d'ordre N . $|A|$ (resp. $\text{tr}(A)$) désigne le déterminant (resp. la trace) de la matrice A . Soit A une matrice appartenant à $\mathcal{M}_{n \times n}(\mathbb{R})$. On suppose que

$$\|A\| = \left(\text{tr}(AA') \right)^{1/2} = \left(\sum_{1 \leq i, j \leq n} |a_{i,j}|^2 \right)^{1/2}. \quad (4.12)$$

Par la définition dans (4.12) de la norme d'une matrice, il s'en suit l'inégalité suivante (voir Comte et Liberman [16], Francq et Zakoïan [31] et Magnus et Newdecker [51])

$$|\text{tr}(AB)| \leq \|A\| \|B\| \quad (4.13)$$

Pour un vecteur $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, on désigne par $\|u\| = (\sum_{i=1}^n |u_i|^2)^{1/2}$. Etant donné une fonction différentiable $f(u_1, \dots, u_n)$, on désigne par $\nabla f = (\frac{\partial f}{\partial u_1}, \dots, \frac{\partial f}{\partial u_n})$ le gradient de f et par $\nabla^2 f$ une matrice carré d'ordre n dont l'élément (i, j) est $\frac{\partial^2 f}{\partial u_i \partial u_j}$. De plus, pour toute fonction H continue sur un compact de \mathbb{R}^n à valeurs dans l'espace des matrices carrées d'ordre n à valeurs réelles c'est-à-dire $H \in C(\Theta, \mathbb{R}^{n \times n})$, on définit par

$$\|H\|_{\Theta} = \sup_{\theta \in \Theta} \|H(\theta)\|. \quad (4.14)$$

4.2 Condition de Stationnarité pour les modèles CCC-HGARCH

Dans ce paragraphe, nous étudions la stationnarité stricte du modèle CCC-HGARCH. Cette étude nous permet d'aboutir à une condition suffisante qui garantit l'existence d'une solution strictement stationnaire du processus CCC-HGARCH défini en (4.6)-(4.9). L'idée dans cette partie est de généraliser le Théorème 2.1.1 de Douc et al. [24]. Cependant, contrairement au cas univarié, obtenir une écriture de la matrice de covariance du CCC-HGARCH sous forme de série de Volterra n'est pas évident.

Pour cette étude, nous nous inspirons du travail de Bardet et Wintenberger [4] qui ont exploité un résultat de Doukhan [26] et Doukhan et Wintenberger [26] (voir Théorème 7.4.1, section 1.7 du chapitre 1) pour étudier les conditions d'existence d'une solution strictement stationnaire pour les processus causaux multidimensionnels définis par :

$$X_t = M_{\theta_0}(X_{t-1}, X_{t-2}, \dots) \cdot \varepsilon_t + f_{\theta_0}(X_{t-1}, X_{t-2}, \dots) \quad \forall t \in \mathbb{Z}, \quad (4.15)$$

(voir Proposition 1 dans Bardet et Wintenberger [4], section 1.7 du chapitre 1).

Dans un premier temps, nous faisons un travail préliminaire qui consiste à redéfinir certaines formules utiles pour la démonstration de nos résultats. Ainsi, en partant du fait que P soit une matrice symétrique définie positive, il existe alors une matrice $P^{1/2}$ telle que $P = P^{1/2}P^{1/2}$. Nous pouvons ainsi réécrire la matrice de covariance conditionnelle H_t définie en (4.7) par

$$\begin{aligned} H_t &= D_t P^{1/2} P^{1/2} D_t \\ &= D_t P^{1/2} (D_t' P^{1/2})' \\ &= D_t P^{1/2} (D_t P^{1/2})'. \end{aligned} \quad (4.16)$$

Le choix de P , définie positive est suffisant pour assurer que H_t est aussi symétrique définie positive. Cela nous permet également d'affirmer qu'il existe une matrice $H_t^{1/2}$ telle que

$$H_t = H_t^{1/2} H_t^{1/2}.$$

Par identification dans (4.16), nous avons

$$H_t^{1/2} = D_t P^{1/2}. \quad (4.17)$$

En utilisant (4.17), le processus r_t défini dans (4.1) se réécrit comme

$$\begin{aligned} r_t &= D_t P^{1/2} \underline{\varepsilon}_t, \\ r_t &= D_t \tilde{\varepsilon}_t \quad \text{avec} \quad \tilde{\varepsilon}_t = P^{1/2} \underline{\varepsilon}_t. \end{aligned} \quad (4.18)$$

Ainsi, le processus ζ_t défini dans (4.9) peut être écrit comme

$$\zeta_t = D_t^2 \underline{\eta}_t \quad \text{avec} \quad \underline{\eta}_t = (\tilde{\varepsilon}_{1t}^2, \tilde{\varepsilon}_{2t}^2, \dots, \tilde{\varepsilon}_{Nt}^2)'. \quad (4.19)$$

Les processus $\tilde{\varepsilon}_t$ et $\underline{\eta}_t$ sont également i.i.d car composés de $\underline{\varepsilon}_t$ par des applications mesurables. Nous établissons dans un second temps le résultat suivant

Théorème 4.2.1. *Supposons que la condition suivante*

$$\mathbb{E}(\|\underline{\eta}_t\|) \sum_{j=1}^{\infty} \|B_j(\theta)\|_{\Theta} < 1 \quad \text{c'est-à-dire} \quad \mathbb{E}(\|\underline{\varepsilon}_t\|^2) \sum_{j=1}^{\infty} \|B_j(\theta)\|_{\Theta} < 1$$

soit vérifiée avec

$$B_j(\theta) = \begin{pmatrix} b_{1j}(\theta) & 0 & 0 & \dots \\ 0 & b_{2j}(\theta) & 0 & \dots \\ & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & b_{Nj}(\theta) \end{pmatrix}.$$

Alors, il existe une solution strictement stationnaire du modèle CCC-HGARCH défini en (4.6) - (4.7) tel que $\mathbb{E}(\|r_t\|^2) < \infty$.

Preuve. Nous remarquons que les processus r_t et ζ_t définis en (4.18) et (4.19) appartiennent à la famille des processus définis par (4.15) et par Doukhan [25] (voir Théorème 7.4.1, section 1.7 du chapitre 1). En effet, par identification dans l'équation (1.15), nous aboutissons au lien suivant avec (4.19) :

$$\zeta_t = M(\zeta_{t-1}, \zeta_{t-2}, \zeta_{t-3}, \dots, \eta_t) = D^2(\zeta_{t-1}, \zeta_{t-2}, \zeta_{t-3}, \dots) \cdot \eta_t \quad (4.20)$$

tel que

$$D^2(\zeta_{t-1}, \zeta_{t-2}, \zeta_{t-3}, \dots) = \begin{pmatrix} a_1 + \sum_{j=1}^{\infty} b_{1j} r_{1t-j}^2 & 0 & 0 & \dots \\ 0 & a_2 + \sum_{j=1}^{\infty} b_{2j} r_{2t-j}^2 & 0 & \dots \\ & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & a_N + \sum_{j=1}^{\infty} b_{Nj} r_{Nt-j}^2 \end{pmatrix}.$$

Cette matrice se décompose également comme suit :

$$D^2(\zeta_{t-1}, \zeta_{t-2}, \zeta_{t-3}, \dots) = \text{diag}(\underline{a}) + \begin{pmatrix} \sum_{j=1}^{\infty} b_{1j} L^j & 0 & 0 & \dots \\ 0 & \sum_{j=1}^{\infty} b_{2j} L^j & 0 & \dots \\ & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sum_{j=1}^{\infty} b_{Nj} L^j \end{pmatrix} \text{diag}(\zeta_t)$$

avec $diag(\underline{a})$ la matrice carrée diagonale dont les coefficients sont les coordonnées du vecteur $\underline{a} = (a_1, a_2, \dots, a_N)'$.

Par ailleurs, d'après (4.20), nous obtenons

$$\begin{aligned} \|M(\underline{0}, \underline{\eta}_t)\| &= \|D^2(\underline{0}) \cdot \underline{\eta}_t\| \\ &\leq \|D^2(\underline{0})\| \cdot \|\underline{\eta}_t\| \\ \mathbb{E}(\|M(\underline{0}, \underline{\eta}_t)\|) &\leq K \mathbb{E}(\|\underline{\eta}_t\|), \end{aligned}$$

par suite,

$$\mathbb{E}(\|M(\underline{0}, \underline{\eta}_t)\|) < \infty \quad \text{si} \quad \mathbb{E}(\|\underline{\eta}_t\|) < \infty \quad \text{c'est-à-dire} \quad \text{si} \quad \mathbb{E}(\|\underline{\varepsilon}_t\|^2) < \infty. \quad (4.21)$$

De plus, notons par $u = (u_i)_{i \geq 1}$ et $v = (v_i)_{i \geq 1}$ deux familles de vecteurs de $(\mathbb{R}^N)^\infty$. Nous parvenons à, pour tout $r = 1, 2, 3, 4$,

$$\begin{aligned} (\mathbb{E}\|M_{\theta_0}(u, \underline{\eta}_t) - M_{\theta_0}(v, \underline{\eta}_t)\|^r)^{1/r} &= (\mathbb{E}\|D_{\theta_0}^2(u) \cdot \underline{\eta}_t - D_{\theta_0}^2(v) \cdot \underline{\eta}_t\|^r)^{1/r} \\ &\leq (\mathbb{E}\|D_{\theta_0}^2(u) - D_{\theta_0}^2(v)\|^r \|\underline{\eta}_t\|^r)^{1/r} \\ &\leq (\mathbb{E}\|\underline{\eta}_t\|^r)^{1/r} \|D_{\theta_0}^2(u) - D_{\theta_0}^2(v)\| \\ &\leq (\mathbb{E}\|\underline{\eta}_t\|^r)^{1/r} \left\| \sum_{j=1}^{\infty} B_j(\theta_0) [diag(u_j) - diag(v_j)] \right\| \\ &\leq (\mathbb{E}\|\underline{\eta}_t\|^r)^{1/r} \sum_{j=1}^{\infty} \|B_j(\theta_0)\|_{\Theta} \|u_j - v_j\| \quad (4.22) \end{aligned}$$

avec

$$B_j(\theta_0) = \begin{pmatrix} b_{1j}(\theta_0) & 0 & 0 & \dots \\ 0 & b_{2j}(\theta_0) & 0 & \dots \\ & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & b_{Nj}(\theta_0) \end{pmatrix}.$$

Par ailleurs, d'après (4.3), nous obtenons

$$\begin{aligned} \|B_j(\theta_0)\|_{\Theta} &\leq K \cdot j^{(-1-d^*)^2} \quad \text{avec} \quad d^* = \min_{1 \leq i \leq N} d_i, \\ &= O(j^{-1-d^*}), \end{aligned} \quad (4.23)$$

par application du critère de convergence de Riemann, la série $\sum_{j=1}^{\infty} \|B_j(\theta_0)\|_{\Theta}$ converge.

Finalement, à partir de (4.21) et (4.22), la Proposition 1 dans Bardet et Wintenberger [4] et le Théorème 7.4.1 dans Doukhan [25] (voir section 1.7 du chapitre 1) stipulent

qu'il existe une solution strictement stationnaire dans \mathbb{L}^1 satisfaisant l'équation (4.20) si

$$\mathbb{E}(\|\eta_t\|) \sum_{j=1}^{\infty} \|B_j(\theta)\|_{\Theta} < 1. \quad (4.24)$$

Soit $\tilde{\zeta}_t$ cette solution strictement stationnaire, alors le processus $r_t = g(\tilde{\zeta}_t)$, avec g mesurable, est solution strictement stationnaire de l'équation (4.18).

Dans ce travail, nous supposons que les degrés p et q des polynômes matriciels sont des entiers fixés, connus d'avance. Ils peuvent être déterminés par application des critères de l'AIC et du BIC.

4.3 Estimation pour les modèles CCC-HGARCH

Dans cette partie, nous étudions les propriétés asymptotiques de l'estimateur du Quasi-Maximum de vraisemblance (QMV) pour les modèles CCC-HGARCH.

4.3.1 Présentation et Propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance

Soit r_1, r_2, \dots, r_n un n -échantillon généré par le modèle CCC-HGARCH défini dans (4.6)-(4.7). La quasi-vraisemblance gaussienne de CCC-HGARCH s'écrit :

$$L(\theta, r_1, r_2, \dots, r_n) = \prod_{t=1}^n \frac{1}{(2\pi)^{N/2} |H_{t,\theta}|^{1/2}} \exp\left(-\frac{1}{2} r_t' H_{t,\theta}^{-1} r_t\right).$$

La log-vraisemblance donne

$$\begin{aligned} L_n(\theta) &= -\frac{1}{2n} \sum_{t=1}^n l_t(\theta), \quad \text{avec} \\ l_t(\theta) &= \log |H_{t,\theta}| + r_t' H_{t,\theta}^{-1} r_t \end{aligned} \quad (4.25)$$

Dans la pratique, les valeurs de $r_0, \dots, r_{1-q}, \dots$ sont non observables. De manière similaire aux relations (2.13)-(2.14) au chapitre 2, nous faisons les initialisations suivantes :

$$r_t = \underline{0} \quad \text{et} \quad \underline{h}_t = \underline{\alpha} + \sum_{j=1}^p \mathbf{B}_j \underline{h}_{t-j}, \quad \text{pour} \quad t \leq 0. \quad (4.26)$$

Pour $t \geq 1$, nous définissons récursivement la matrice de covariance conditionnelle \tilde{H}_t définie avec les valeurs initiales de (4.26) par :

$$\begin{aligned} \tilde{H}_t &= \tilde{D}_t P \tilde{D}_t, \quad \tilde{D}_t = \text{diag}(\sqrt{\tilde{h}_{11t}}, \sqrt{\tilde{h}_{22t}}, \dots, \sqrt{\tilde{h}_{NNt}}) = \text{diag}(\tilde{h}_t^{1/2}) \\ \tilde{h}_t &= \underline{\gamma} + \sum_{j=1}^p \mathbf{B}_j \tilde{h}_{t-j} + \left[(I_N - \sum_{j=1}^p \mathbf{B}_{j'} L^{j'}) - (I_N - \sum_{i'=1}^q \Delta_{i'} L^{i'}) (I - L)^d \right] \underline{\Omega}_{\tilde{\zeta}_t} \end{aligned} \quad (4.27)$$

La log-vraisemblance définie avec \tilde{H}_t est

$$\begin{aligned}\tilde{L}_n(\theta) &= \frac{1}{2n} \sum_{t=1}^n \tilde{l}_t(\theta), \quad \text{où} \\ \tilde{l}_t(\theta) &= \log |\tilde{H}_{t,\theta}| + \underline{r}'_t \tilde{H}_{t,\theta}^{-1} \underline{r}_t.\end{aligned}\tag{4.29}$$

Par suite, l'estimateur du quasi-maximum de vraisemblance (QMV) pour le modèle CCC-HGARCH est défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{L}_n(\theta).\tag{4.30}$$

Afin d'établir les propriétés asymptotiques de l'estimateur du QMV du modèle CCC-HGARCH, nous donnons les hypothèses suivantes :

- (**A**₀) Le vrai paramètre $\theta_0 \in \overset{\circ}{\Theta}$ où $\overset{\circ}{\Theta}$ désigne l'intérieur de Θ , avec Θ un sous-ensemble compact de \mathbb{R}^e pour $e = (p + q + 3)N + N(N - 1)/2$. De plus, nous avons $\theta_0 = (\gamma_{0i}, \beta_{0ii}, \delta_{0ii}, \omega_{0ii}, \rho_{0ij}, d_{0ij})$ avec $1 \leq i, j \leq N$ et $i \neq j$.
- (**A**₁) Le processus $\underline{r}_t \in \mathbb{R}^N$ est une solution strictement stationnaire et ergodique du modèle CCC-HGARCH défini dans (4.6) et (4.7). De plus, $Sp(\mathbf{B}_j) < 1$, pour $j = 1, \dots, p$.
- (**A**₂) La matrice de corrélation $P(\theta)$ est symétrique définie positive pour tout $\theta \in \Theta$.
- (**A**₃) Les composantes de ε_t sont indépendantes. De plus, leur carré a une distribution non dégénérée.
- (**A**₄) Il existe une constante positive c_1 telle que pour tout $t \geq 1$ et pour tout $\theta \in \Theta$, $|H_t| \geq c_1$.
- (**A**₅) Les fonctions H_t de paramètre θ sont continues et nous avons pour tout $\theta \in \Theta$, $H_t(\theta) = H_t(\theta_0) \implies \theta = \theta_0$.
- (**A**₆) Nous supposons les relations suivantes $\mathbb{E}(\|\underline{\varepsilon}_t \underline{\varepsilon}'_t\|^2) < \infty$ et $\mathbb{E}(\|\underline{r}_t\|^4) < \infty$.

Les hypothèses (**A**₀) – (**A**₁) sont standards. L'hypothèse (**A**₁) est aussi utilisée pour assurer l'existence du moment d'ordre deux du processus. De plus, la seconde condition de (**A**₁) permet d'écrire la variance conditionnelle univariée définie dans (4.3) sous la forme $ARCH(\infty)$. L'hypothèse (**A**₂) implique que $H_{t,\theta}$ est une matrice définie positive. Les hypothèses (**A**₄) – (**A**₅) sont cruciales pour la consistance. L'hypothèse (**A**₆) contient des arguments techniques utilisés pour étudier l'existence de la matrice des dérivées secondes.

4.3.2 Consistance de l'estimateur

Nous pouvons donc établir à partir du théorème suivant le résultat de la consistance forte de l'estimateur.

Théorème 4.3.1. *Sous les hypothèses (\mathbf{A}_0) - (\mathbf{A}_5) l'estimateur du QMV, noté $\hat{\theta}_n$, défini dans (4.30) converge presque sûrement vers θ_0 , c'est-à-dire,*

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0 \quad p.s.$$

Avant de donner la preuve du théorème, nous présentons d'abord sous forme de lemmes, les résultats intermédiaires obtenus. Ainsi, nous énonçons le lemme suivant :

Lemme 4.3.1. *Soit r_t un processus strictement stationnaire et ergodique du CCC-HGARCH satisfaisant (4.1). Soit $h_t(\theta)$ défini comme dans (4.9) et $\tilde{h}_t(\theta)$ défini comme dans (4.28). Sous les hypothèses (\mathbf{A}_0) - (\mathbf{A}_2) , il existe aussi des réels positifs K et u tels que,*

1.

$$\sup_{\theta \in \Theta} \|H_t(\theta) - \tilde{H}_t(\theta)\| \leq K \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|r_{t-j}\|^2 \right)^{1/2} \text{ avec } d_m = \inf_{\theta \in \Theta} d^*(\theta) \text{ et}$$

$$d^* = \min_{1 \leq i \leq N} d_i,$$

2.

$$\sup_{\theta \in \Theta} \|H_t^{-1}(\theta)\| = O(1), \quad \sup_{\theta \in \Theta} \|\tilde{H}_t^{-1}(\theta)\| = O(1), \quad \text{et} \quad \mathbb{E}(\sup_{\theta \in \Theta} \|H_t\|) < \infty.$$

Preuve

1. Les éléments diagonaux des matrices $H_t(\theta)$ et $\tilde{H}_t(\theta)$ notés respectivement h_{iit} et \tilde{h}_{iit} sont les variances conditionnelles d'un modèle HGARCH univarié. Ainsi, h_{iit} peut être défini comme la variance conditionnelle d'un modèle ARCH(∞) (voir (2.3)). Nous obtenons donc pour $i = 1, \dots, N$,

$$h_{iit} = \alpha + \sum_{j=1}^{\infty} b_{ij} r_{it}^2, \quad \tilde{h}_{iit} = \alpha + \sum_{j=1}^{t-1} b_{ij} r_{it}^2, \quad \text{avec } b_{ij} = O(j^{-1-d_i}) \text{ et } b_{ij} > 0.$$

La différence des variances conditionnelles univariées conduit à

$$|h_{iit} - \tilde{h}_{iit}| \leq \sum_{j=t}^{+\infty} b_{ij} r_{it}^2. \quad (4.31)$$

Dans le cas multivarié, la différence des matrices de variances conditionnelles donne

$$H_t(\theta) - \tilde{H}_t(\theta) = D_t P D_t - \tilde{D}_t P \tilde{D}_t = (D_t - \tilde{D}_t) P D_t - \tilde{D}_t P (\tilde{D}_t - D_t).$$

En passant à la norme, nous avons

$$\begin{aligned} \|(D_t - \tilde{D}_t) P D_t - \tilde{D}_t P (\tilde{D}_t - D_t)\| &\leq \|D_t - \tilde{D}_t\| \|P\| [\|D_t\| + \|\tilde{D}_t\|] \\ &\leq 2 \|P\| \|D_t\| \|D_t - \tilde{D}_t\|. \end{aligned}$$

En utilisant les définitions (4.7) et (4.27), nous obtenons

$$\|D_t - \tilde{D}_t\| = \left(\sum_{i=1}^N (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2 \right)^{1/2} \leq \left(\sum_{i=1}^N |h_{iit} - \tilde{h}_{iit}| \right)^{1/2}. \quad (4.32)$$

En effet, nous avons

$$\begin{aligned} (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2 &= h_{iit} + \tilde{h}_{iit} - 2(h_{iit} \tilde{h}_{iit})^{1/2} \quad \text{et} \\ \left((h_{iit} - \tilde{h}_{iit})^{1/2} \right)^2 &= |h_{iit} - \tilde{h}_{iit}| = h_{iit} - \tilde{h}_{iit}. \end{aligned}$$

En faisant la différence, nous obtenons

$$\begin{aligned} \left((h_{iit} - \tilde{h}_{iit})^{1/2} \right)^2 - (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2 &= h_{iit} - \tilde{h}_{iit} - h_{iit} - \tilde{h}_{iit} + 2(h_{iit} \tilde{h}_{iit})^{1/2} \\ &= -2 \left[\tilde{h}_{iit} - (h_{iit} \tilde{h}_{iit})^{1/2} \right]. \end{aligned} \quad (4.33)$$

Cependant, d'après (2.3) et (2.16), nous avons

$$h_{iit} = \tilde{h}_{iit} + \mu_t \quad \text{avec} \quad \mu_t = \sum_{j=t}^{\infty} b_j(\theta) y_{t-j}^2 > 0. \quad (4.34)$$

En substituant (4.34) dans (4.33), nous avons

$$\begin{aligned} \left[(h_{iit} - \tilde{h}_{iit})^{1/2} \right]^2 - (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2 &= -2 \left(\tilde{h}_{iit} - ((\tilde{h}_{iit} + \mu_t) \tilde{h}_{iit})^{1/2} \right) \\ &= -2 \left[\tilde{h}_{iit} - \left((\tilde{h}_{iit} + \mu_t) \tilde{h}_{iit} \right)^{1/2} \right] \\ &= -2 \left[\tilde{h}_{iit} - \left(\tilde{h}_{iit}^2 + \mu_t \tilde{h}_{iit} \right)^{1/2} \right]. \end{aligned} \quad (4.35)$$

Par construction, les variances conditionnelles h_t et \tilde{h}_t définies en (2.3) et (2.16) sont strictement positives. Ainsi, (4.35) devient

$$\begin{aligned} \left[(h_{iit} - \tilde{h}_{iit})^{1/2} \right]^2 - (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2 &= -2 \left[\tilde{h}_{iit} - \tilde{h}_{iit} \left(1 + \frac{\mu_t}{\tilde{h}_{iit}} \right)^{1/2} \right] \\ &= -2 \tilde{h}_{iit} \left[1 - \left(1 + \frac{\mu_t}{\tilde{h}_{iit}} \right)^{1/2} \right]. \end{aligned} \quad (4.36)$$

D'après (4.34), μ_t est positif, ce qui entraîne que

$$\left(1 + \frac{\mu_t}{\tilde{h}_{iit}}\right)^{1/2} > 1 \quad \text{puis} \quad 1 - \left(1 + \frac{\mu_t}{\tilde{h}_{iit}}\right)^{1/2} < 0.$$

Par suite, le second membre de (4.36) est positif. Il s'en suit que

$$\left[(h_{iit} - \tilde{h}_{iit})^{1/2}\right]^2 > (h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2})^2. \quad (4.37)$$

En utilisant (4.34), nous aboutissons à

$$(h_{iit} - \tilde{h}_{iit})^{1/2} > h_{iit}^{1/2} - \tilde{h}_{iit}^{1/2}. \quad (4.38)$$

La relation (4.32) est ainsi justifiée et entraîne par le Lemme 2.2.1.(1) que

$$\|D_t - \tilde{D}_t\| \leq \left(\sum_{i=1}^N |h_{iit} - \tilde{h}_{iit}|\right)^{1/2} \leq u \left(\sum_{i=1}^N \sum_{j=t}^{+\infty} j^{-1-d_i} r_{i,t-j}^2\right)^{1/2} \quad (4.39)$$

$$\leq u \left(\sum_{j=t}^{+\infty} j^{-1-d^*} \|r_{t-j}\|^2\right)^{1/2}, \quad \text{avec} \quad d^* = \min_{1 \leq i \leq N} d_i. \quad (4.40)$$

Par suite, nous obtenons

$$\sup_{\theta \in \Theta} \|H_t(\theta) - \tilde{H}_t(\theta)\| \leq 2u \|P\| \|D_t\| \left(\sum_{j=t}^{+\infty} j^{-1-d^*} \|r_{t-j}\|^2\right)^{1/2}$$

$$\sup_{\theta \in \Theta} \|H_t(\theta) - \tilde{H}_t(\theta)\| \leq K \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|r_{t-j}\|^2\right)^{1/2}, \quad \text{avec} \quad d_m = \inf_{\theta \in \Theta} d^*(\theta).$$

2. Nous débutons la preuve du second point de ce lemme en utilisant la relation (4.27) afin d'aboutir à

$$\sup_{\theta \in \Theta} \|\tilde{H}_t^{-1}(\theta)\| \leq \sup_{\theta \in \Theta} \|\tilde{D}_t^{-1}\|^2 \|P^{-1}\|. \quad (4.41)$$

Par (4.27), nous parvenons également à

$$\|\tilde{D}_t^{-1}\|^2 = \sum_{i=1}^N \frac{1}{\tilde{h}_{iit}} \leq \sum_{i=1}^N \frac{1}{\min_{1 \leq i \leq N} \alpha_i(\theta)} \leq \frac{N}{\alpha^*(\theta)} \quad \text{avec} \quad \alpha^* = \min_{1 \leq i \leq N} \alpha_i.$$

Par passage au sup, nous obtenons

$$\sup_{\theta \in \Theta} \|\tilde{D}_t^{-1}\|^2 \leq \sup_{\theta \in \Theta} \frac{N}{\alpha^*(\theta)} \leq \frac{N}{\inf_{\theta \in \Theta} \alpha^*(\theta)} \leq u. \quad (4.42)$$

De plus, il ressort de l'hypothèse (\mathbf{A}_2) que

$$\|P^{-1}\| = \frac{1}{|\lambda_{\min}|}, \quad \text{et} \quad \|P\| = |\lambda_{\max}| \quad (4.43)$$

où λ_{min} et λ_{max} désignent respectivement la plus petite et la plus grande valeur propre associées à la matrice P . Par (4.42)-(4.43), nous pouvons conclure qu'il existe un réel positif u tel que

$$\sup_{\theta \in \Theta} \|\tilde{H}_t^{-1}(\theta)\| \leq u.$$

En utilisant les mêmes arguments, nous parvenons également à

$$\sup_{\theta \in \Theta} \|H_t^{-1}(\theta)\| \leq u. \quad (4.44)$$

En ce qui concerne l'existence du sup de la norme de la matrice $H_t(\theta)$, nous notons que

$$\begin{aligned} \|H_t\| &= \|D_t P D_t\| \leq \|D_t\|^2 \|P\| \leq \lambda_{max} \|D_t\|^2 \quad \text{avec} \\ \|D_t\|^2 &= \sum_{i=1}^N h_{iit}. \end{aligned} \quad (4.45)$$

Cela entraîne d'après (4.40) que

$$\|D_t\|^2 \leq u \sum_{j=1}^{+\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2.$$

Par suite, nous avons

$$\sup_{\theta \in \Theta} \|H_t\| \leq \lambda_{max} \sum_{j=1}^{+\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2, \quad (4.46)$$

En passant aux espérances, et par application du Théorème de convergence monotone (section 1.7 du chapitre 1), nous obtenons

$$\mathbb{E}(\sup_{\theta \in \Theta} \|H_t\|) \leq \lambda_{max} \sum_{j=1}^{+\infty} j^{-1-d^*} \mathbb{E}(\|\underline{r}_{t-j}\|^2),$$

par application de l'hypothèse (\mathbf{A}_1) , nous obtenons le résultat. □

Lemme 4.3.2. *Supposons que les hypothèses (\mathbf{A}_0) - (\mathbf{A}_6) sont satisfaites, alors nous avons*

$$\sup_{\theta \in \Theta} |L_n(\theta) - \tilde{L}_n(\theta)| \rightarrow 0 \quad p.s., \quad n \rightarrow \infty.$$

Preuve De (4.25) et (4.29), nous obtenons

$$\begin{aligned} \sup_{\theta \in \Theta} |L_n(\theta) - \tilde{L}_n(\theta)| &\leq \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |\log |H_t| - \log |\tilde{H}_t|| \\ &\quad + \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |r'_t(H_t^{-1} - \tilde{H}_t^{-1})r_t|. \end{aligned} \quad (4.47)$$

Nous notons que

$$\begin{aligned} \tilde{H}_t^{-1}(H_t - \tilde{H}_t)H_t^{-1} &= [\tilde{H}_t^{-1}H_t - \tilde{H}_t^{-1}\tilde{H}_t]H_t^{-1} = \tilde{H}_t^{-1}H_tH_t^{-1} - \tilde{H}_t^{-1}\tilde{H}_tH_t^{-1} \\ &= \tilde{H}_t^{-1} - H_t^{-1}. \end{aligned} \quad (4.48)$$

En utilisant (4.48), la seconde somme du second membre de (4.47) peut être écrite comme

$$\begin{aligned} |r'_t(H_t^{-1} - \tilde{H}_t^{-1})r_t| &= |r'_t\tilde{H}_t^{-1}(H_t - \tilde{H}_t)H_t^{-1}r_t| \\ &= |tr(\tilde{H}_t^{-1}(H_t - \tilde{H}_t)H_t^{-1}r_t r'_t)|. \end{aligned}$$

En utilisant la relation de la trace d'une matrice définie en (4.13), nous arrivons à

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |r'_t(H_t^{-1} - \tilde{H}_t^{-1})r_t| &\leq \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |tr(\tilde{H}_t^{-1}(H_t - \tilde{H}_t)H_t^{-1}r_t r'_t)|, \\ &\leq \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} \|\tilde{H}_t^{-1}\| \|H_t - \tilde{H}_t\| \|H_t^{-1}\| \|r_t r'_t\|. \end{aligned} \quad (4.49)$$

En utilisant le Lemme 4.3.1, nous aboutissons à

$$\frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |r'_t(H_t^{-1} - \tilde{H}_t^{-1})r_t| \leq \frac{Ku}{n} \sum_{t=1}^n \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|r_{t-j}\|^2 \right)^{1/2} \|r_t\|^2. \quad (4.50)$$

Nous souhaitons montrer que le second membre de (4.50) tend vers 0, quand n tend vers l'infini. Le but est de montrer la convergence de la série suivante

$$c_t = \sum_{t \geq 1} \frac{1}{t} \mathbb{E} \left(\left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|r_{t-j}\|^2 \right)^{1/2} \|r_t\|^2 \right).$$

En utilisant l'hypothèse (\mathbf{A}_6) pour laquelle $\mathbb{E}(\|r_t\|^4) < \infty$, nous appliquons l'inégalité de Cauchy pour obtenir

$$\begin{aligned} c_t &\leq \sum_{t \geq 1} \frac{1}{t} \left(\mathbb{E} \sum_{j=t}^{+\infty} j^{-1-d_m} \|r_{t-j}\|^2 \right)^{1/2} \left(\mathbb{E}(\|r_t\|^4) \right)^{1/2} \\ &\leq \sum_{t \geq 1} \frac{1}{t} \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \mathbb{E} \|r_{t-j}\|^2 \right)^{1/2} \left(\mathbb{E}(\|r_t\|^4) \right)^{1/2}. \end{aligned} \quad (4.51)$$

Par les hypothèses (\mathbf{A}_1) et (\mathbf{A}_6) , et en utilisant les mêmes arguments comme dans le chapitre 2 (Lemme 2.2.2), la série définie dans (4.51) converge selon le critère de Riemann. Par le Lemme de Kronecker (section 1.7 du chapitre 1), nous déduisons que

$$\frac{1}{n} \sum_{t=1}^n \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|\underline{r}_t\|^2 \rightarrow 0 \quad \text{p.s, } n \rightarrow \infty. \quad (4.52)$$

Ainsi, nous avons

$$\frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |\underline{r}'_t (H_t^{-1} - \tilde{H}_t^{-1}) \underline{r}_t| \rightarrow 0 \quad \text{p.s, } n \rightarrow \infty. \quad (4.53)$$

Nous traitons à présent la première somme de (4.47). Par définition dans (4.10), les déterminants des matrices H_t et \tilde{H}_t sont strictement positifs, ainsi nous obtenons

$$\log |H_t| - \log |\tilde{H}_t| = \log \frac{|H_t|}{|\tilde{H}_t|} = \log |H_t \tilde{H}_t^{-1}|. \quad (4.54)$$

De plus, nous remarquons que

$$\begin{aligned} H_t \tilde{H}_t^{-1} &= I_N + H_t \tilde{H}_t^{-1} - I_N \\ &= I_N + H_t \tilde{H}_t^{-1} - \tilde{H}_t \tilde{H}_t^{-1} \\ &= I_N + (H_t - \tilde{H}_t) \tilde{H}_t^{-1}. \end{aligned}$$

Par application du déterminant et de l'inégalité de Hadamard (section 1.7 du chapitre 1), nous obtenons

$$|H_t \tilde{H}_t^{-1}| = |I_N + (H_t - \tilde{H}_t) \tilde{H}_t^{-1}| \leq \|I_N + (H_t - \tilde{H}_t) \tilde{H}_t^{-1}\|^N.$$

Par suite, nous aboutissons à

$$\begin{aligned} \log |H_t \tilde{H}_t^{-1}| &\leq N \log \|I_N + (H_t - \tilde{H}_t) \tilde{H}_t^{-1}\| \\ &\leq N \|H_t - \tilde{H}_t\| \|\tilde{H}_t^{-1}\|. \end{aligned} \quad (4.55)$$

En utilisant le Lemme 4.3.1, nous avons

$$\log |H_t \tilde{H}_t^{-1}| \leq N \sup_{\theta \in \Theta} \|\tilde{H}_t^{-1}(\theta)\| K \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|\underline{r}_{t-j}\|^2 \right)^{1/2}.$$

Ainsi, la première somme de (4.47) devient

$$\frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |\log |H_t| - \log |\tilde{H}_t|| \leq NK \frac{1}{n} \sum_{t=1}^n \left(\sum_{j=t}^{+\infty} j^{-1-d_m} \|\underline{r}_{t-j}\|^2 \right)^{1/2}.$$

En utilisant les mêmes raisons que dans (4.49) – (4.53), nous avons

$$\frac{1}{n} \sum_{t=1}^n \sup_{\theta \in \Theta} |\log |H_t| - \log |\tilde{H}_t|| \rightarrow 0 \quad \text{p.s.}, \quad n \rightarrow \infty. \quad (4.56)$$

En combinant (4.53) et (4.56), nous obtenons le résultat attendu. \square

Lemme 4.3.3. *Supposons que les hypothèses (\mathbf{A}_0) - (\mathbf{A}_5) sont satisfaites, alors nous avons*

$$\sup_{\theta \in \Theta} |L_n(\theta) - \mathbb{E}_{\theta_0}(l_t(\theta))| \rightarrow 0 \quad \text{p.s.} \quad n \rightarrow \infty,$$

où $l_t(\theta) = \log |H_t| + \underline{r}'_t H_t^{-1} \underline{r}_t$.

Preuve Dans un premier temps, nous montrerons l'existence de $\mathbb{E}(\sup_{\theta \in \Theta} \|l_t(\theta)\|)$. Soit λ_i , $i = 1, \dots, N$, les valeurs propres de H_t à t fixé. Puisque la matrice H_t est symétrique définie positive, d'après (\mathbf{A}_2) et (4.10), alors toutes ses valeurs propres λ_i sont positives et nous avons

$$\log |H_t| = \sum_i^N \log \lambda_i \leq \sum_i^N \lambda_i = \text{Tr}(H_t) \quad (4.57)$$

$$\log |H_t| \leq \sum_i^N (H_t)_{ii}. \quad (4.58)$$

Comme les éléments $(H_t)_{ii}$ de la matrice H_t correspondent aux variances conditionnelles univariées des composantes du processus \underline{r}_t , alors nous obtenons

$$\log |H_t| \leq \sum_i^N h_{iit} \quad (4.59)$$

En utilisant les mêmes techniques comme dans (4.45) – (4.46), nous parvenons à :

$$\sup_{\theta \in \Theta} \log |H_t| \leq u \sum_{j=1}^{+\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2.$$

En appliquant l'espérance et le Théorème de convergence monotone (section 1.7 du chapitre 1), nous avons

$$\mathbb{E}(\sup_{\theta \in \Theta} \log |H_t|) \leq u \sum_{j=1}^{+\infty} j^{-1-d^*} \mathbb{E}(\|\underline{r}_{t-j}\|^2).$$

De plus, par l'hypothèse de stationnarité du processus en (\mathbf{A}_1) , nous obtenons par application du critère de convergence de la série de Riemann que

$$\mathbb{E}(\sup_{\theta \in \Theta} \log |H_t(\theta)|) < \infty. \quad (4.60)$$

Par ailleurs, en partant de la relation suivante

$$|\log |H_t|| = \max(\log |H_t|, 0) + \max(-\log |H_t|, 0),$$

pour $\max(\log |H_t|, 0) = \log |H_t|$, nous sommes dans le cas de (4.60).

Pour $\max(-\log |H_t|, 0) = -\log |H_t|$, nous avons

$$\sup_{\theta \in \Theta} (-\log |H_t(\theta)|) = \inf_{\theta \in \Theta} (\log |H_t(\theta)|) \leq \sup_{\theta \in \Theta} \log |H_t(\theta)|. \quad (4.61)$$

Nous obtenons d'après (4.60)-(4.61)

$$\mathbb{E}(\sup_{\theta \in \Theta} |\log |H_t||) < \infty. \quad (4.62)$$

Nous avons en outre

$$\begin{aligned} |r'_t H_{t,\theta}^{-1} r_t| &= |Tr(r'_t H_{t,\theta}^{-1} r_t)| = \|Tr(H_{t,\theta}^{-1} r_t r'_t)\| \leq \|r_t r'_t\| \|H_{t,\theta}^{-1}\| \\ &\leq \|r_t\|^2 \|H_{t,\theta}^{-1}\|. \end{aligned}$$

A partir du Lemme 4.3.1, nous avons

$$\begin{aligned} \sup_{\theta \in \Theta} |r'_t H_{t,\theta}^{-1} r_t| &\leq \|r_t\|^2 \sup_{\theta \in \Theta} \|H_{t,\theta}^{-1}(\theta)\| \\ \mathbb{E}(\sup_{\theta \in \Theta} |r'_t H_{t,\theta}^{-1} r_t|) &\leq \mathbb{E}(\|r_t\|^2). \end{aligned}$$

Par l'hypothèse (\mathbf{A}_1) , nous obtenons

$$\mathbb{E}(\sup_{\theta \in \Theta} |r'_t H_{t,\theta}^{-1} r_t|) < \infty. \quad (4.63)$$

De (4.62) et (4.63), nous parvenons finalement à

$$\mathbb{E}(\sup_{\theta \in \Theta} \|l_t(\theta)\|) \leq \mathbb{E}(\sup_{\theta \in \Theta} |\log |H_t||) + \mathbb{E}(\sup_{\theta \in \Theta} |r'_t H_{t,\theta}^{-1} r_t|) < \infty. \quad (4.64)$$

Ce qui met fin à la première partie.

Pour la seconde partie, nous notons que la suite $l_t(\theta)$ définie dans ce lemme est une fonction du processus r_t . Ainsi, nous avons $l_t(\theta) = f(r_t, r_{t-1}, \dots)$ pour f mesurable. A partir de l'hypothèse (\mathbf{A}_1) et du Théorème 3.5.8 dans Stout [60] (section 1.7), la suite $l_t(\theta)$ est strictement stationnaire et ergodique. De plus, en utilisant le théorème ergodique de Birkhoff dans Straumann [61] (section 1.7), nous obtenons $L_n(\theta)$ converge presque sûrement vers $\mathbb{E}(l_t(\theta))$. Par ailleurs, d'après le Théorème de convergence uniforme dans Straumann [61] (section 1.7), la relation (4.64) suffit à appliquer la loi forte uniforme

des grands nombres à suite $l_t(\theta)$. La loi forte uniforme des grands nombres traduit le fait que

$$\left\| \frac{1}{n} \sum_{t=1}^n l_t(\theta) - \mathbf{E}_{\theta_0}(l_t(\theta)) \right\|_{\Theta} \xrightarrow{p.s} 0 \quad n \rightarrow \infty.$$

Ce qui se traduit par

$$\sup_{\theta \in \Theta} |L_n(\theta) - \mathbf{E}_{\theta_0}(l_t(\theta))| \rightarrow 0 \quad p.s \quad n \rightarrow \infty.$$

Ainsi, via le théorème de convergence uniforme (section 1.7), nous prouvons la convergence uniforme. \square

Preuve du Théorème 4.3.1. A partir des Lemmes 4.3.2 et 4.3.3, nous obtenons

$$\sup_{\theta \in \Theta} |\tilde{L}_n(\theta) - L(\theta)| \rightarrow 0 \quad p.s \quad n \rightarrow \infty, \quad (4.65)$$

c'est-à-dire $\tilde{L}_n(\theta)$ converge uniformément presque sûrement vers $L(\theta)$. Par ailleurs, la suite de fonctions $\tilde{L}_n(\theta)$ définie en (4.29) est continue en θ car elle est composée de somme et produit de fonctions élémentaires continue en θ . Pour les mêmes raisons, elle est aussi une fonction mesurable des observations. Par suite, l'hypothèse (\mathbf{A}_0) montre que l'ensemble Θ est compact. Pour satisfaire aux conditions du Théorème 4.1.1 dans Amemiya [1] (section 1.7), il nous reste à montrer que $L(\theta)$ a un unique maximum en θ_0 .

D'après le Lemme (4.3.3), $l_t(\theta_0) = \log |H_t(\theta_0)| + \underline{r}'_t H_t^{-1}(\theta_0) \underline{r}_t$. Nous avons en θ_0 , à partir de (4.6)

$$\begin{aligned} \mathbb{E}(\underline{r}'_t H_t^{-1}(\theta_0) \underline{r}_t) &= \mathbb{E}(Tr(H_t^{-1}(\theta_0) \underline{r}_t \underline{r}'_t)) \\ &= \mathbb{E}\left(Tr\left[H_t^{-1}(\theta_0) H_t^{1/2}(\theta_0) \underline{\varepsilon}_t \underline{\varepsilon}'_t (H_t^{1/2}(\theta_0))'\right]\right). \end{aligned} \quad (4.66)$$

En passant aux espérances conditionnelles et par l'indépendance de (ε_t) , nous obtenons

$$\begin{aligned} \mathbb{E}(\underline{r}'_t H_t^{-1}(\theta_0) \underline{r}_t) &= Tr\left(\mathbb{E}\left[\mathbb{E}\left(H_t^{-1}(\theta_0) H_t^{1/2}(\theta_0) \underline{\varepsilon}_t \underline{\varepsilon}'_t (H_t^{1/2}(\theta_0))' \middle| \mathcal{I}_{t-1}\right)\right]\right) \\ &= Tr\left(\mathbb{E}\left[H_t^{-1}(\theta_0) H_t^{1/2}(\theta_0) \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}'_t | \mathcal{I}_{t-1}) (H_t^{1/2}(\theta_0))'\right]\right) \\ &= Tr\left(\mathbb{E}\left[H_t^{-1}(\theta_0) H_t^{1/2}(\theta_0) \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}'_t) (H_t^{1/2}(\theta_0))'\right]\right) \\ &= Tr\left(\mathbb{E}\left[H_t^{-1}(\theta_0) H_t^{1/2}(\theta_0) I_N (H_t^{1/2}(\theta_0))'\right]\right) \\ &= \mathbb{E}\left(Tr(H_t^{-1}(\theta_0) H_t(\theta_0))\right) \\ &= \mathbb{E}(Tr(I_N)) = N. \end{aligned} \quad (4.67)$$

En utilisant les relations (4.60) et (4.67), nous avons

$$\mathbb{E}(l_t(\theta_0) = \log |H_t(\theta_0)| + \underline{r}'_t H_t^{-1}(\theta_0) \underline{r}_t) < \infty. \quad (4.68)$$

De plus, pour tout θ appartenant à Θ , nous obtenons

$$\mathbb{E}(l_t(\theta)) - \mathbb{E}(l_t(\theta_0)) = -\mathbb{E}\left(\log |H_t(\theta_0)| - \log |H_t(\theta)|\right) + \mathbb{E}(\underline{r}'_t H_t^{-1}(\theta) \underline{r}_t) - \mathbb{E}(\underline{r}'_t H_t^{-1}(\theta_0) \underline{r}_t). \quad (4.69)$$

En appliquant (4.54) et (4.67) à (4.69), nous parvenons à

$$\mathbb{E}(l_t(\theta)) - \mathbb{E}(l_t(\theta_0)) = -\mathbb{E}(\log |H_t(\theta_0) H_t^{-1}(\theta)|) + \mathbb{E}(\underline{r}'_t H_t^{-1}(\theta) \underline{r}_t) - N. \quad (4.70)$$

Nous pouvons remarquer que

$$\mathbb{E}(\underline{r}'_t H_t^{-1}(\theta) \underline{r}_t) = \mathbb{E}(\text{Tr}(H_t^{-1}(\theta) \underline{r}_t \underline{r}'_t)) = \mathbb{E}\left(\text{Tr}\left[H_t^{-1}(\theta) H_t^{1/2}(\theta_0) \underline{\varepsilon}_t \underline{\varepsilon}'_t (H_t^{1/2}(\theta_0))'\right]\right).$$

En utilisant les mêmes arguments que dans (4.66)-(4.69), nous avons

$$\mathbb{E}(\underline{r}'_t H_t^{-1}(\theta) \underline{r}_t) = \mathbb{E}(\text{Tr}(H_t^{-1}(\theta) H_t(\theta_0))) = \mathbb{E}(\text{Tr}(H_t(\theta_0) H_t^{-1}(\theta))).$$

Ainsi, la relation (4.70) devient

$$\mathbb{E}(l_t(\theta)) - \mathbb{E}(l_t(\theta_0)) = -\mathbb{E}(\log |H_t(\theta_0) H_t^{-1}(\theta)|) + \mathbb{E}(\text{Tr}(H_t(\theta_0) H_t^{-1}(\theta))) - N. \quad (4.71)$$

Etant donné que $H_t(\theta_0)$ et $H_t^{-1}(\theta)$ sont des matrices définies positives, alors les N valeurs propres de $H_t(\theta_0) H_t^{-1}(\theta)$ sont strictement positives. En notant λ_i ces valeurs, nous obtenons par (4.71)

$$\begin{aligned} \mathbb{E}(l_t(\theta)) - \mathbb{E}(l_t(\theta_0)) &= -\mathbb{E}(\log(\prod_{i=1}^N \lambda_i)) + \mathbb{E}(\sum_{i=1}^N \lambda_i) - N \\ &= \mathbb{E}\left(\sum_{i=1}^N (-\log \lambda_i + \lambda_i - 1)\right) \geq 0, \end{aligned} \quad (4.72)$$

en effet, (4.72) se justifie par la relation $\log x \leq x - 1$ pour tout x strictement positif avec égalité pour $x = 1$. Ce qui sous-entend que dans (4.72), l'égalité surviendra si et seulement si pour tout $i = 1, \dots, N$, $\lambda_i = 1$. Un tel cas se produira seulement lorsque $H_t(\theta_0) = H_t(\theta)$. Cependant d'après l'hypothèse (\mathbf{A}_5), l'égalité $H_t(\theta_0) = H_t(\theta)$ est vérifiée uniquement lorsque $\theta_0 = \theta$.

Par conséquent pour tout $\theta \in \Theta$ avec $\theta_0 \neq \theta$, nous avons

$$\mathbb{E}(l_t(\theta)) - \mathbb{E}(l_t(\theta_0)) > 0. \quad (4.73)$$

Nous montrons ainsi que $L(\theta)$ a un unique minimum en θ_0 .

Finalement, toutes les hypothèses du Théorème 4.1.1 dans Amemiya [1] (section 1.7) sont vérifiées, nous obtenons donc le résultat désiré. \square

4.3.3 Normalité asymptotique de l'estimateur

Le théorème suivant nous donne le résultat de la distribution asymptotique de l'estimateur $\hat{\theta}_n$.

Théorème 4.3.2. *Supposons que les hypothèses (\mathbf{A}_0) - (\mathbf{A}_6) sont satisfaites. Alors, nous avons*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}(\theta_0)B(\theta_0)A^{-1}(\theta_0)) \quad n \rightarrow \infty,$$

$$\text{avec } A(\theta_0) = \mathbb{E}(\nabla^2 l_t(\theta_0)) \quad \text{et} \quad B(\theta_0) = \mathbb{E}(\nabla l_t(\theta_0)(\nabla l_t(\theta_0))').$$

Soit $s = 1, \dots, (p + q + 3)N$ tel que θ_s désigne uniquement les paramètres du modèle GARCH hyperbolique présents dans θ . Et $u = (p + q + 3)N + 1, \dots, e$ tel que θ_u désigne les paramètres de la matrice de corrélation P présents dans θ . En d'autres termes, le vecteur de paramètres θ se réécrit de la manière suivante $\theta = (\theta_s, \theta_u)$. Nous utilisons la technique de différentiation de matrice dans Horn et Johnsen [37] et Magnus et Neudecker [51] (section 4.1.3). Nous donnons les dérivées de $l_t(\theta)$. Ainsi, nous avons

$$\begin{aligned} \frac{\partial l_t(\theta)}{\partial \theta_s} &= -Tr\{(r_t r_t' D_t^{-1} P^{-1} + P^{-1} D_t^{-1} r_t r_t') D_t^{-1} \frac{\partial D_t^{-1}}{\partial \theta_s} D_t^{-1}\} \\ &\quad + 2Tr(D_t^{-1} \frac{\partial D_t^{-1}}{\partial \theta_s}) \end{aligned} \quad (4.74)$$

$$\frac{\partial l_t(\theta)}{\partial \theta_u} = -Tr(P^{-1} D_t^{-1} r_t r_t' D_t^{-1} P^{-1} \frac{\partial P^{-1}}{\partial \theta_u}) + Tr(P^{-1} \frac{\partial P^{-1}}{\partial \theta_u}). \quad (4.75)$$

Pour la suite du travail, nous adopterons les notations suivantes pour les dérivées premières en θ_0

$$D_{0t} = D_t(\theta_0), \quad P_0 = P(\theta_0), \quad \dot{D}_{0ts} = \frac{\partial D_t}{\partial \theta_s}(\theta_0), \quad \dot{P}_{0tu} = \frac{\partial P}{\partial \theta_e}(\theta_0). \quad (4.76)$$

Les dérivées premières définies en (4.74) – (4.75) peuvent être écrites sous la forme

$$\frac{\partial l_t(\theta_0)}{\partial \theta_s} = Tr\{(I_N - P_0^{-1} \tilde{\varepsilon}_t \tilde{\varepsilon}_t') \dot{D}_{0ts} D_{0t}^{-1} + (I_N - \tilde{\varepsilon}_t \tilde{\varepsilon}_t' P_0^{-1}) D_{0t}^{-1} \dot{D}_{0ts}\}, \quad (4.77)$$

$$\frac{\partial l_t(\theta_0)}{\partial \theta_u} = Tr\{(I_N - P_0^{-1} \tilde{\varepsilon}_t \tilde{\varepsilon}_t') P_0^{-1} \dot{P}_{0tu}\}. \quad (4.78)$$

Nous voulons prouver l'existence de $\mathbb{E}(\|\frac{\partial l_t(\theta_0)}{\partial \theta}\|)$.

En utilisant, (4.77) nous obtenons par définition pour $s = 1, \dots, (P + q + 3)N$

$$\left| \frac{\partial l_t(\theta_0)}{\partial \theta_s} \right| \leq |Tr\{(I_N - P_0^{-1} \tilde{\xi}_t \tilde{\xi}_t') \dot{D}_{0ts} D_{0t}^{-1}\}| \quad (4.79)$$

$$+ |Tr\{(I_N - \tilde{\xi}_t \tilde{\xi}_t' P_0^{-1}) D_{0t}^{-1} \dot{D}_{0ts}\}|$$

$$\leq K_2 \|\dot{D}_{0ts} D_{0t}^{-1}\| \quad (4.80)$$

$$\mathbb{E}\left(\left| \frac{\partial l_t(\theta_0)}{\partial \theta_s} \right|\right) \leq K_2 \mathbb{E}(\|\dot{D}_{0ts} D_{0t}^{-1}\|). \quad (4.81)$$

Pour les mêmes raisons, nous avons

$$\mathbb{E}\left(\left| \frac{\partial l_t(\theta_0)}{\partial \theta_u} \right|\right) \leq K_3, \quad (4.82)$$

$$\mathbb{E}\left(\left| \frac{\partial l_t(\theta_0)}{\partial \theta_u} \frac{\partial l_t(\theta_0)}{\partial \theta_f} \right|\right) \leq K_4, \quad f = (p+q+3)N+1, \dots, e, \quad (4.83)$$

$$\mathbb{E}\left(\left| \frac{\partial l_t(\theta_0)}{\partial \theta_s} \frac{\partial l_t(\theta_0)}{\partial \theta_u} \right|\right) \leq K_5 \mathbb{E}(\|\dot{D}_{0ts} D_{0t}^{-1}\|). \quad (4.84)$$

En utilisant l'hypothèse (\mathbf{A}_3) et l'inégalité de Cauchy, nous obtenons aussi pour $k = 1, \dots, (p + q + 3)N$

$$\mathbb{E}\left|\frac{\partial l_t(\theta_0)}{\partial \theta_s} \frac{\partial l_t(\theta_0)}{\partial \theta_k}\right| \leq K_6 \{\mathbb{E}\|\dot{D}_{0ts} D_{0t}^{-1}\|^2 \mathbb{E}\|\dot{D}_{0tk} D_{0t}^{-1}\|^2\}^{1/2}. \quad (4.85)$$

Les dérivées de D_t donnent

$$\dot{D}_{0ts} = \frac{1}{2} D_{0t}^{-1} \text{diag}\left(\frac{\partial h_t}{\partial \theta_s}(\theta_0)\right) = \frac{1}{2} \text{diag}\left(\frac{1}{\sqrt{h_t(i)}} \frac{\partial h_t(i)}{\partial \theta_s}(\theta_0)\right) \quad (4.86)$$

$$\dot{D}_{0ts} D_{0t}^{-1} = \frac{1}{2} \text{diag}\left(\frac{1}{h_t(i)} \frac{\partial h_t(i)}{\partial \theta_s}(\theta_0)\right) \quad \text{pour } i=1, \dots, N. \quad (4.87)$$

Pour $i = 1, \dots, N$, l'existence des dérivées partielles première et seconde de $h_t(i)$ (cas univarié) a été traitée dans le chapitre 2 (Lemme 2.2.1.(2)) sous la condition de l'existence du moment d'ordre quatre du processus. Par suite, nous pouvons dire qu'il existe un voisinage \mathcal{U} de θ_0 tel que

$$\mathbb{E}\left(\sup_{\theta \in \mathcal{U}} \left| \frac{1}{h_t(i)} \frac{\partial h_t(i)}{\partial \theta_s} \right|^2\right) < \infty, \quad (4.88)$$

$$\mathbb{E}\left(\sup_{\theta \in \mathcal{U}} \left| \frac{1}{h_t(i)} \frac{\partial^2 h_t(i)}{\partial \theta_s \partial \theta_j} \right|^2\right) < \infty, \quad \text{pour } 1 \leq s, j \leq (p + q + 3)N. \quad (4.89)$$

Par conséquent, en regroupant (4.85), (4.87) et (4.88)-(4.89), les moments d'ordre 2 du vecteur score et de la hessienne existent.

Lemme 4.3.4. *Supposons les hypothèses (\mathbf{A}_0) - (\mathbf{A}_4) sont vérifiées, alors nous avons*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta}(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B(\theta_0)) \quad n \rightarrow \infty. \quad (4.90)$$

Preuve D'après les relations (4.74)-(4.75), le processus $\{\frac{\partial l_t}{\partial \theta}(\theta_0)\}_t$ est une fonction du passé de r_t donc est \mathcal{I}_{t-1} -mesurable. Par ailleurs, par (4.74)-(4.75), le processus $\{\frac{\partial l_t}{\partial \theta}(\theta_0)\}_t$ est strictement stationnaire et ergodique d'après le Théorème 3.5.8 dans Stout [60]. De plus, l'existence du moment d'ordre 1 du vecteur score de $l_t(\theta)$ montre que le processus $\{\frac{\partial l_t}{\partial \theta}(\theta_0)\}_t$ est intégrable. Par ailleurs, à partir de (4.77) – (4.78) et (\mathbf{A}_0) , pour tout vecteur $\underline{b} \neq \underline{0} \in \mathbb{R}^e$, nous obtenons par linéarité de l'espérance conditionnelle

$$\begin{aligned} \mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= \mathbb{E}\left[\underline{b}' \left(\frac{\partial l_t}{\partial \theta_s}(\theta_0), \frac{\partial l_t}{\partial \theta_u}(\theta_0) \right) | \mathcal{I}_{t-1}\right] \\ &= \mathbb{E}\left(b_s Tr\{(I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') \dot{D}_{0ts} D_{0t}^{-1}\} | \mathcal{I}_{t-1}\right) \\ &\quad + \mathbb{E}\left(b_u Tr\{(I_N - \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t' P_0^{-1}) D_{0t}^{-1} \dot{D}_{0ts}\} | \mathcal{I}_{t-1}\right) \\ &\quad + \mathbb{E}\left(b_u Tr\{(I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') P_0^{-1} \dot{P}_{0tu}\} | \mathcal{I}_{t-1}\right). \end{aligned}$$

En intervertissant trace et espérance, nous obtenons

$$\begin{aligned} \mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= Tr\{\mathbb{E}(b_s (I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') \dot{D}_{0ts} D_{0t}^{-1} | \mathcal{I}_{t-1})\} \\ &\quad + Tr\{\mathbb{E}(b_u (I_N - \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t' P_0^{-1}) D_{0t}^{-1} \dot{D}_{0ts} | \mathcal{I}_{t-1})\} \\ &\quad + Tr\{\mathbb{E}(b_u (I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') P_0^{-1} \dot{P}_{0tu} | \mathcal{I}_{t-1})\}. \end{aligned}$$

D'après (4.87), la quantité $\dot{D}_{0ts} D_{0t}^{-1}$ est fonction des variances conditionnelles univariées du processus r_{it} . En utilisant (4.4), nous constatons que les variances conditionnelles univariées sont des fonctions du passé du processus r_{it} . En d'autres termes, la quantité $\dot{D}_{0ts} D_{0t}^{-1}$ est fonction du passé du processus r_t donc elle est \mathcal{I}_{t-1} -mesurable. Par application des propriétés de l'espérance conditionnelle, nous obtenons

$$\begin{aligned} \mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} \mathbb{E}((I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') | \mathcal{I}_{t-1})\} \\ &\quad + Tr\{b_u \dot{D}_{0ts} D_{0t}^{-1} \mathbb{E}((I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') | \mathcal{I}_{t-1})\} \\ &\quad + Tr\{b_u P_0^{-1} \dot{P}_{0tu} \mathbb{E}((I_N - P_0^{-1} \tilde{\underline{\varepsilon}}_t \tilde{\underline{\varepsilon}}_t') | \mathcal{I}_{t-1})\}. \end{aligned}$$

En remplaçant le processus $\tilde{\underline{\varepsilon}}_t$ défini en (4.18), nous aboutissons à

$$\begin{aligned} \mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1} P_0^{1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t' | \mathcal{I}_{t-1})) P_0^{1/2}\} \\ &\quad + Tr\{b_u \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1} P_0^{1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t' | \mathcal{I}_{t-1})) P_0^{1/2}\} \\ &\quad + Tr\{b_u P_0^{-1} \dot{P}_{0tu} (I_N - P_0^{-1} P_0^{1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t' | \mathcal{I}_{t-1})) P_0^{1/2}\}. \end{aligned}$$

Par construction du modèle en (4.6)-(4.9), le processus $\underline{\varepsilon}_t$ est indépendant du processus r_t , ce qui nous conduit à

$$\begin{aligned}\mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t') P_0^{1/2})\} \\ &+ Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t') P_0^{1/2})\} \\ &+ Tr\{b_u P_0^{-1} \dot{P}_{0tu} (I_N - P_0^{-1/2} \mathbb{E}(\underline{\varepsilon}_t \underline{\varepsilon}_t') P_0^{1/2})\}.\end{aligned}$$

Puisque le processus $\underline{\varepsilon}_t$ est centré de variance la matrice identité, nous obtenons finalement

$$\begin{aligned}\mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) &= Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1/2} I_N P_0^{1/2})\} \\ &+ Tr\{b_s \dot{D}_{0ts} D_{0t}^{-1} (I_N - P_0^{-1/2} I_N P_0^{1/2})\} \\ &+ Tr\{b_u P_0^{-1} \dot{P}_{0tu} (I_N - P_0^{-1/2} I_N P_0^{1/2})\}.\end{aligned}$$

Par suite, nous avons

$$\mathbb{E}(\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0) | \mathcal{I}_{t-1}) = 0.$$

Ce qui nous permet de dire que $\{\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0)\}_t$ est une différence de martingales strictement stationnaire et ergodique. Par ailleurs, en utilisant les relations (4.85) et (4.88), nous obtenons que la matrice $B(\theta_0)$ existe et $\underline{b}' \frac{\partial l_t}{\partial \theta}(\theta_0)$ est de carré intégrable. Par le Théorème central limite pour les différences de martingales (section 1.7), nous obtenons le résultat. \square

Dans le paragraphe suivant, nous abordons l'étude portant sur les dérivées partielles secondes. Pour tous $s, j = 1, \dots, (p + q + 3)N$ et $u, v = 1 + (p + q + 3)N, \dots, e$, nous obtenons les dérivées partielles secondes de la manière suivante :

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \theta_s \partial \theta_j}(\theta_0) &= Tr(a_1 + a_2 + a_3) \\ \frac{\partial^2 l_t}{\partial \theta_s \partial \theta_u}(\theta_0) &= Tr(a_4 + a_5) \\ \frac{\partial^2 l_t}{\partial \theta_u \partial \theta_v}(\theta_0) &= Tr(a_6),\end{aligned}\tag{4.91}$$

où

$$\begin{aligned}
a_1 &= D_t^{-1}P^{-1}D_t^{-1}\frac{\partial D_t}{\partial\theta_s}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}\frac{\partial D_t}{\partial\theta_j} + D_t^{-1}\frac{\partial D_t}{\partial\theta_s}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}D_t^{-1}\frac{\partial D_t}{\partial\theta_j} \\
&\quad + D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}D_t^{-1}\frac{\partial D_t}{\partial\theta_s}D_t^{-1}\frac{\partial D_t}{\partial\theta_j} - D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}D_t^{-1}\frac{\partial^2 D_t}{\partial\theta_s\partial\theta_j}, \\
a_2 &= -2D_t^{-1}\frac{\partial D_t}{\partial\theta_s}D_t^{-1}\frac{\partial D_t}{\partial\theta_j} + D_t^{-1}\frac{\partial^2 D_t}{\partial\theta_s\partial\theta_j}, \\
a_4 &= P^{-1}D_t^{-1}\frac{\partial D_t}{\partial\theta_s}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}\frac{\partial P}{\partial\theta_u}, \\
a_6 &= P^{-1}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}\frac{\partial P}{\partial\theta_u}P^{-1}\frac{\partial P}{\partial\theta_v} + P^{-1}\frac{\partial P}{\partial\theta_v}P^{-1}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}\frac{\partial P}{\partial\theta_v} \\
&\quad - P^{-1}D_t^{-1}\underline{r}_t\underline{r}'_tD_t^{-1}P^{-1}\frac{\partial^2 P}{\partial\theta_u\partial\theta_v} - P^{-1}\frac{\partial^2 P}{\partial\theta_u\partial\theta_v} - P^{-1}\frac{\partial P}{\partial\theta_u}P^{-1}\frac{\partial P}{\partial\theta_v}. \tag{4.92}
\end{aligned}$$

Nous obtenons a_3 en permutant $\underline{r}_t\underline{r}'_t$ et P^{-1} dans a_1 . Aussi, en permutant $\underline{r}_t\underline{r}'_t$ et $\frac{\partial D_t}{\partial\theta_s}$ dans a_4 , nous obtenons a_5 .

Lemme 4.3.5. *Supposons que les hypothèses (\mathbf{A}_0) - (\mathbf{A}_5) sont vérifiées, alors il existe un voisinage \mathcal{U} de θ_0 tel que*

$$\sup_{\theta \in \mathcal{U}} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'} - A(\theta_0) \right\| \rightarrow 0, \quad p.s., \quad n \rightarrow \infty, \tag{4.93}$$

où $A(\theta_0) = \mathbb{E}\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta\partial\theta'}\right)$ est une matrice définie positive.

Preuve Nous avons que $\frac{\partial^2 P}{\partial\theta_u\partial\theta_v} = 0$. En utilisant les relations (4.91) – (4.92), nous constatons que le processus des dérivées partielles secondes $\left\{\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right\}_t$ peut être défini comme une fonction mesurable du processus \underline{r}_t . A partir de l'hypothèse (\mathbf{A}_1) et du Théorème 3.5.8 dans Stout [60] (voir section 1.7), $\left\{\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right\}_t$ est un processus strictement stationnaire et ergodique. A partir de (4.91) – (4.92) combinées aux résultats (4.88) – (4.89), nous obtenons

$$\mathbb{E}\left(\sup_{\theta \in \mathcal{U}} \left\| \frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'} \right\| \right) < \infty. \tag{4.94}$$

D'après le Théorème de convergence uniforme dans Straumann [61] (voir section 1.7), la condition (4.94) suffit à appliquer la loi forte uniforme des grands nombres à la suite $\left\{\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right\}_t$. Comme dans la preuve du lemme 4.3.3, la loi forte uniforme des grands nombres donne

$$\sup_{\theta \in \mathcal{U}} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'} - \mathbb{E}\left(\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right) \right\| \rightarrow 0, \quad p.s., \quad n \rightarrow \infty.$$

Ainsi, nous prouvons la convergence uniforme. \square

Lemme 4.3.6. *Supposons que les hypothèses (\mathbf{A}_0) - (\mathbf{A}_6) sont vérifiées. Alors, il existe un voisinage \mathcal{U} de θ_0 tel que*

1.

$$\sup_{\theta \in \mathcal{U}} \left\| \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\partial l_t(\theta)}{\partial \theta} - \frac{\partial \tilde{l}_t(\theta)}{\partial \theta} \right\} \right\| \rightarrow 0 \quad p.s., \quad n \rightarrow \infty.$$

2.

$$\sup_{\theta \in \mathcal{U}} \left\| \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{l}_t(\theta)}{\partial \theta \partial \theta'} \right\} \right\| \rightarrow 0 \quad p.s., \quad n \rightarrow \infty.$$

Preuve

1. A partir de (4.74) et (4.75), nous obtenons

$$\frac{\partial l_t(\theta)}{\partial \theta} - \frac{\partial \tilde{l}_t(\theta)}{\partial \theta} = \text{Tr}(b_{1t} + b_{2t} + b_{3t}) \quad (4.95)$$

où

$$b_{1t} = -D_t^{-1} r_t r_t' D_t^{-1} D_t (D_t^{-1} - \tilde{D}_t^{-1}) P^{-1} D_t^{-1} \frac{\partial D_t}{\partial \theta_s} \quad (4.96)$$

$$b_{2t} = -D_t^{-1} r_t r_t' \tilde{D}_t^{-1} P^{-1} D_t^{-1} \left(\frac{\partial D_t}{\partial \theta_s} - \frac{\partial \tilde{D}_t}{\partial \theta_s} \right) \quad (4.97)$$

pour $s = 1, \dots, (p+q+3)N$. Remarquons que b_{3t} contient des termes de b_{2t} et b_{1t} . En utilisant les mêmes arguments que dans le Lemme 4.3.1.(1), nous avons

$$\sup_{\theta \in \Theta} \|D_t - \tilde{D}_t\| \leq K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|r_{t-j}\|^2 \right)^{1/2}, \quad (4.98)$$

$$\sup_{\theta \in \Theta} \|D_t^{-1}\| = \sup_{\theta \in \Theta} \|\tilde{D}_t^{-1}\| = O(1). \quad (4.99)$$

Nous traitons premièrement b_{1t} .

Par la norme matricielle définie en (4.13), nous obtenons de (4.96)

$$|\text{tr}(b_{1t})| \leq \|D_t^{-1} r_t\| \|r_t' D_t^{-1}\| \|D_t (D_t^{-1} - \tilde{D}_t^{-1})\| \|P^{-1}\| \|D_t^{-1} \frac{\partial D_t}{\partial \theta_i}\|. \quad (4.100)$$

Dans le paragraphe suivant, nous traitons indépendamment les différentes normes définies dans (4.100).

En utilisant la relation suivante :

$$\begin{aligned} D_t (D_t^{-1} - \tilde{D}_t^{-1}) &= D_t (D_t^{-1} - D_t \tilde{D}_t^{-1}) = -(D_t \tilde{D}_t^{-1} - I_N) \\ &= -(D_t \tilde{D}_t^{-1} - \tilde{D}_t \tilde{D}_t^{-1}) = -(D_t - \tilde{D}_t) \tilde{D}_t^{-1}, \end{aligned}$$

combinée à (4.98), nous obtenons

$$\sup_{\theta \in \Theta} \|D_t(D_t^{-1} - \tilde{D}_t^{-1})\| \leq K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|r_{t-j}\|^2 \right)^{1/2}. \quad (4.101)$$

Nous remarquons que de (4.87) et (4.88), nous obtenons

$$\mathbb{E} \sup_{\theta \in \Theta} \|D_t^{-2} \text{diag}\left\{ \frac{\partial \mathbf{h}_t}{\partial \theta_s}(\theta) \right\}\|^2 < \infty.$$

Cela conduit à

$$\mathbb{E} \sup_{\theta \in \Theta} \|D_t^{-1} \frac{\partial D_t}{\partial \theta_s}(\theta)\|^2 < \infty. \quad (4.102)$$

De plus, nous remarquons à partir de (4.99) et de la norme matricielle définie en (4.13) que

$$\begin{aligned} \|D_t^{-1} r_t\| &\leq K \|r_t\| \quad \text{et} \\ \|\tilde{D}_t^{-1} r_t\| &= \|\tilde{D}_t^{-1} D_t D_t^{-1} r_t\| \leq \|\tilde{D}_t^{-1} D_t\| \|D_t^{-1} r_t\| \leq K_4 \|r_t\|. \end{aligned} \quad (4.103)$$

En utilisant les relations (4.101)- (4.103), combinées à (4.43), nous obtenons

$$|tr(b_{1t})| \leq K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|r_{t-j}\|^2 \right)^{1/2} \|r_t\|^2 \|g_{1t}\|, \quad \text{où} \quad (4.104)$$

$$g_{1t} = 2D_t^{-1} \frac{\partial D_t(\theta)}{\partial \theta_s} = 2D_t^{-2} \text{diag}\left\{ \frac{\partial \mathbf{h}_t}{\partial \theta_s}(\theta) \right\} \quad (4.105)$$

A présent, nous traitons le cas b_{2t} .

En utilisant la norme matricielle définie en (4.13) et d'après (4.97), nous avons

$$|tr(b_{2t})| \leq \|D_t^{-1} r_t\| \|r_t' \tilde{D}_t^{-1}\| \|P^{-1}\| \|D_t^{-1} \left(\frac{\partial D_t}{\partial \theta_s} - \frac{\partial \tilde{D}_t}{\partial \theta_s} \right)\|. \quad (4.106)$$

D'après (4.86), nous obtenons

$$\begin{aligned} D_t^{-1} \frac{\partial}{\partial \theta_s} (D_t - \tilde{D}_t) &= \frac{1}{2} D_t (D_t^{-1} - \tilde{D}_t^{-1}) D_t^{-2} \text{diag}\left\{ \frac{\partial \mathbf{h}_t(\theta)}{\partial \theta_s} \right\} \\ &\quad + \frac{1}{2} \tilde{D}_t^{-1} D_t^{-1} \text{diag}\left\{ \frac{\partial \mathbf{h}_t(\theta)}{\partial \theta_s} - \frac{\partial \tilde{\mathbf{h}}_t(\theta)}{\partial \theta_s} \right\}. \end{aligned} \quad (4.107)$$

En passant à la norme, le premier terme du second membre de (4.107) devient

$$\begin{aligned} \|V_t\| &= \|D_t (D_t^{-1} - \tilde{D}_t^{-1}) D_t^{-2} \text{diag}\left\{ \frac{\partial \mathbf{h}_t(\theta)}{\partial \theta_s} \right\}\| \\ &\leq \|D_t (D_t^{-1} - \tilde{D}_t^{-1})\| \|D_t^{-2} \text{diag}\left\{ \frac{\partial \mathbf{h}_t(\theta)}{\partial \theta_s} \right\}\|. \end{aligned}$$

Par les relations (4.101) et (4.105), nous obtenons

$$\sup_{\theta \in \Theta} \|V_t\| \leq 2K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|g_{1t}\|. \quad (4.108)$$

En ce qui concerne le second terme du second membre de (4.107), à partir du cas univarié dans le Lemme 2.2.1.(3), nous avons pour tout $s \in 1, \dots, (p+q+3)N$,

$$\sup_{\theta \in \Theta} \left\| \frac{\partial(\underline{h}_t(\theta) - \tilde{\underline{h}}_t(\theta))}{\partial \theta_s} \right\| \leq K \sum_{j=t}^{\infty} j^{-(1+d^*)(1-\eta)} \|\underline{r}_{t-j}\|^2, \quad \text{avec} \quad (4.109)$$

$$\eta < 1 - \frac{1}{1 + \frac{d_m}{2}}.$$

Par suite, d'après les relations (4.103), (4.108) et (4.109), nous parvenons à

$$\sup_{\theta \in \Theta} \left\| D_t^{-1} \left(\frac{\partial D_t}{\partial \theta_i} - \frac{\partial \tilde{D}_t}{\partial \theta_i} \right) \right\| \leq K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|g_{1t}\|$$

$$+ K \sum_{j=t}^{\infty} j^{-(1+d^*)(1-\eta)} \|\underline{r}_{t-j}\|^2. \quad (4.110)$$

Par les relations (4.103), (4.110) et (4.43), il existe un réel K positif, la relation (4.106) devient :

$$|Tr(b_{2t})| \leq K \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|g_{1t}\|$$

$$+ K \sum_{j=t}^{\infty} j^{-(1+d^*)(1-\eta)} \|\underline{r}_{t-j}\|^2. \quad (4.111)$$

Finalement, en regroupant les relations (4.104) et (4.111), nous obtenons

$$\sup_{\theta \in \mathcal{U}} \left| \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\partial l_t(\theta)}{\partial \theta_s} - \frac{\partial \tilde{l}_t(\theta)}{\partial \theta_s} \right\} \right| \leq \frac{K}{n} \sum_{t=1}^n \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|\underline{r}_t\|^2 \|g_{1t}\|$$

$$+ \frac{K}{n} \sum_{t=1}^n \left(\sum_{j=t}^{\infty} j^{-1-d^*} \|\underline{r}_{t-j}\|^2 \right)^{1/2} \|g_{1t}\|$$

$$+ \frac{K}{n} \sum_{t=1}^n \sum_{j=t}^{\infty} j^{-(1+d^*)(1-\eta)} \|\underline{r}_{t-j}\|^2 \quad (4.112)$$

où g_{1t} définie en (4.105) est, d'après (4.102), une suite de variables aléatoires de carré intégrable.

En utilisant les mêmes arguments comme dans le lemme 4.3.2 et le lemme 2.2.6, nous avons les résultats.

2. En utilisant les mêmes arguments que dans le premier point de ce lemme, nous obtenons le résultat. \square

Preuve du Théorème 4.3.2 En utilisant le Lemme 4.3.4 et le Lemme 4.3.6.(1), nous obtenons

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \tilde{l}_t}{\partial \theta}(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B(\theta_0)) \quad n \rightarrow \infty,$$

où

$$B(\theta_0) = \mathbb{E}(\nabla l_t(\theta_0)(\nabla l_t(\theta_0))').$$

A partir du Lemme 4.3.5 et du Lemme 4.3.6.(2), il existe un voisinage \mathcal{U} de θ_0 tel que

$$\sup_{\theta \in \mathcal{U}} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t(\theta)}{\partial \theta \partial \theta'} - A(\theta_0) \right\| \xrightarrow{p.s} 0 \quad n \rightarrow \infty, \quad (4.113)$$

avec

$$A(\theta_0) = \mathbb{E}\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'}\right).$$

Par ailleurs, d'après (4.91)-(4.92), les dérivées partielles, définies sous forme de trace de matrices carrées, sont continues et existent d'après (4.94).

Toutes ces raisons permettent de vérifier les hypothèses du Théorème 4.1.3. dans Ame-miya [1] (section 1.7). Ainsi, nous aboutissons au résultat. \square

4.4 Etude de Simulation

Dans cette partie, nous conduisons une étude de Monte Carlo afin d'illustrer les propriétés de l'estimateur du quasi-maximum de vraisemblance (QMV) sur des échantillons de tailles finies pour le modèle CCC-HGARCH(1,d,1) bivarié. Quatre tailles d'échantillon sont considérées dans cette étude. Il s'agit respectivement de $n = 1000, 2000, 4000$ et 8000 . Pour chaque échantillon généré, les 2000 premières observations sont supprimées afin à réduire les effets des valeurs initiales. Donc, pour chaque répliation, $2000 + n$ observations seront générées pour chaque échantillon. Pour cette étude, nous ferons $N = 1000$ répliations pour chaque échantillon. A chaque instant t , nous générons la matrice de covariance conditionnelle H_t en deux étapes. D'abord, nous générons les variances conditionnelles univariées qui correspondent aux éléments de la diagonale principale de H_t . En second lieu, en utilisant le résultat des éléments de la diagonale principale, nous calculons les éléments en dehors de la diagonale principale. Pour réussir

le calcul de H_t , nous supposons que les composantes du processus ε_t utilisées pour générer les variances conditionnelles univariées suivent chacune une distribution de Student à 13 degrés de liberté.

Le modèle CCC-HGARCH(1,d,1) bivarié que nous étudions présente donc onze paramètres d'intérêt. Le vecteur des vrais paramètres utilisé pour générer les observations est

$$\begin{aligned}\theta_0 &= (\alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}, \delta_{01}, \delta_{02}, \omega_{01}, \omega_{02}, \rho_0, d_{01}, d_{02}) \\ &= (0.25, 0.35, 0.15, 0.1, 0.2, 0.3, 0.45, 0.4, 0.5, 0.55, 0.6).\end{aligned}$$

Dans ce travail, les paramètres du vecteur θ sont estimés conjointement par la méthode du quasi-maximum de vraisemblance. Par ailleurs, nous avons utilisé les indices du MAE (mean absolute errors) et du RMSE (root mean squared errors) comme outils pour évaluer la consistance de l'estimateur. Nous rappelons que

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_0)^2} \quad \text{et} \quad MAE = \frac{1}{N} \sum_{i=1}^N |\hat{x}_i - x_0|, \quad (4.114)$$

avec \hat{x}_i désigne la valeur estimée de x_0 et N le nombre de réplifications.

Nous reportons dans le tableau 4.1 les résultats du MAE et RMSE de l'estimateur du Quasi Maximum de vraisemblance (QMV) pour le modèle CCC-HGARCH(1,d,1) bivarié issu d'une loi de Student à 13 degrés de liberté.

TABLEAU 4.1 – Résultats d'estimation du modèle CCC-HGARCH(1,d,1) bivarié issu d'une loi de Student à 13 degrés de liberté

		Estimation										
n		α_1	α_2	β	β_2	δ_1	δ_2	ω_1	ω_2	ρ	d_1	d_2
1000	MAE	0.15658	0.16622	0.03150	0.04417	0.07702	0.06622	0.05689	0.01322	0.23938	0.07722	0.06460
	RMSE	0.16253	0.17157	0.03370	0.04620	0.07726	0.06629	0.06026	0.01527	0.24229	0.07750	0.06466
2000	MAE	0.13207	0.14242	0.02764	0.03950	0.07373	0.06395	0.04623	0.00916	0.21633	0.07550	0.06308
	RMSE	0.13722	0.14716	0.02961	0.04143	0.07391	0.06402	0.04939	0.01095	0.21880	0.07572	0.06315
4000	MAE	0.11475	0.12488	0.01930	0.02903	0.07205	0.06282	0.03830	0.00744	0.20350	0.07495	0.06244
	RMSE	0.11965	0.12964	0.02147	0.03134	0.07220	0.06288	0.04155	0.00929	0.20588	0.07514	0.06250
8000	MAE	0.09558	0.10574	0.01523	0.02425	0.06942	0.06101	0.02933	0.00717	0.18495	0.07284	0.06095
	RMSE	0.09974	0.10985	0.01734	0.02641	0.06954	0.06105	0.03252	0.00916	0.18686	0.07299	0.06099

Nous observons que pour tous les paramètres estimés du modèle, les erreurs sont relativement petites. De plus, pour chacun des paramètres du modèle, les deux indices, MAE et RMSE diminuent et tendent vers zéro quand la taille de l'échantillon augmente. Ce résultat permet donc d'illustrer la convergence de l'estimateur dans ce cas.

Les figures 4.1 et 4.2 présentent les tracés des courbes quantile-quantile (Q-Q plots) pour les paramètres estimés du CCC-HGARCH(1,d,1) bivarié issu d'une loi de Student à 13 degrés de liberté. Ces tracés sont donnés uniquement pour l'échantillon de taille $n = 8000$.

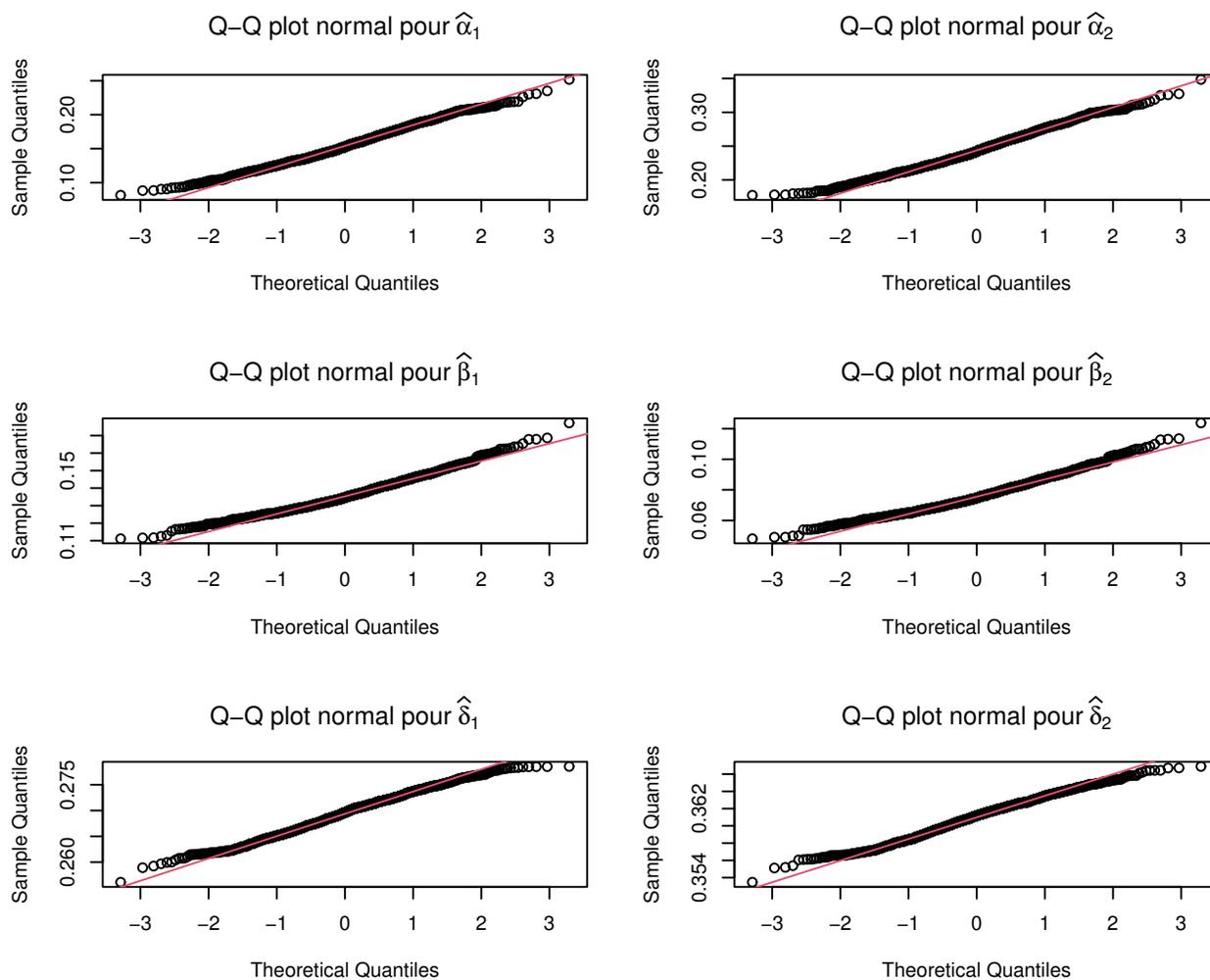


FIGURE 4.1 – Q-Q plots pour les paramètres estimés du modèle CCC-HGARCH(1,d,1) bivarié pour $n=8000$

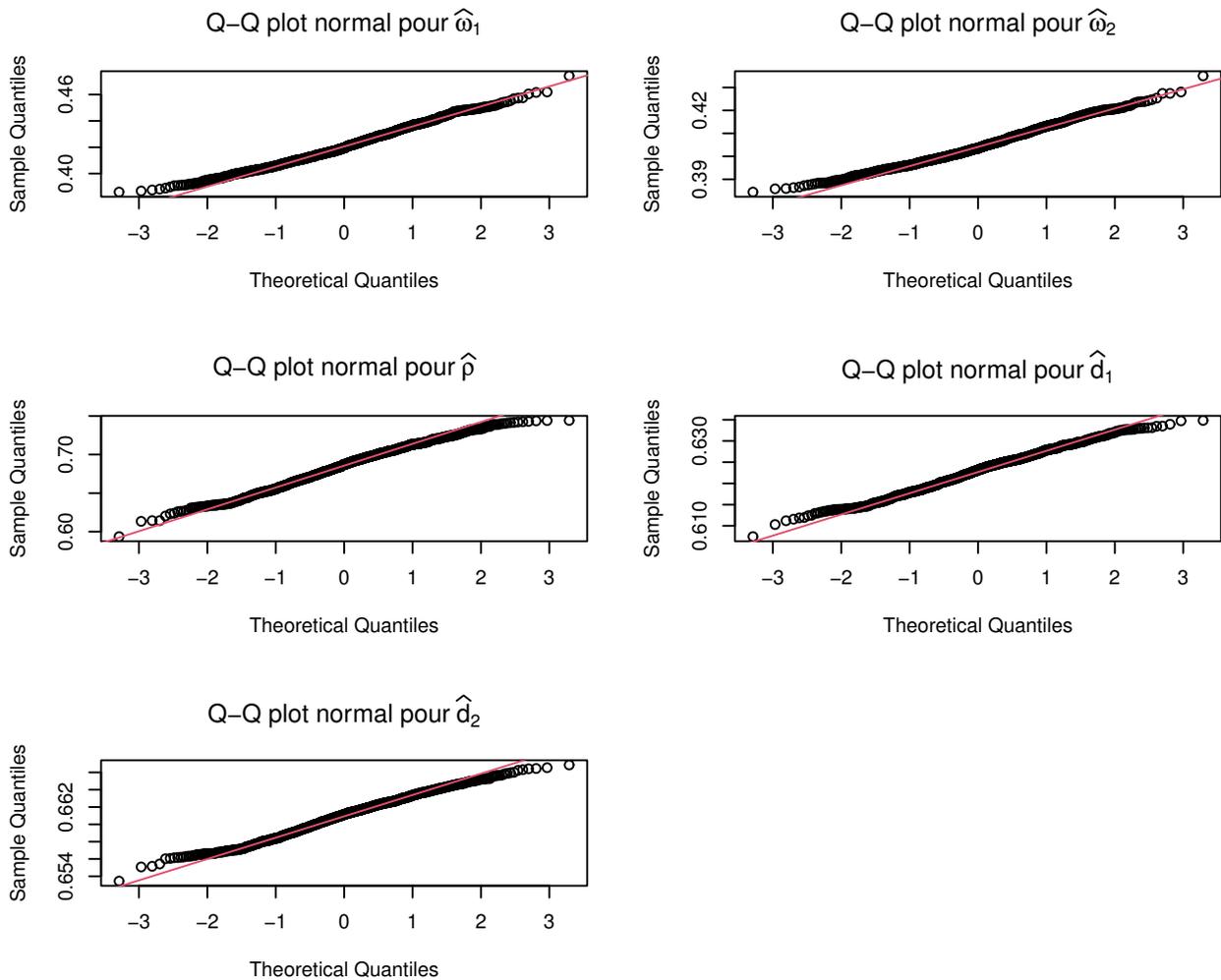


FIGURE 4.2 – Q-Q plots pour les paramètres estimés du modèle CCC-HGARCH(1,d,1) bivarié pour $n=8000$

Nous utilisons ces figures pour comparer la distribution des paramètres estimés à celle de la loi normale. A partir des figures 4.1 et 4.2, nous observons que pour chaque paramètre estimé, tous les points tracés sur les différents graphes sont approximativement alignés sur la même droite. Ainsi, à partir du concept du Q-Q plot, pour les différents graphes de la figure 4.1 et de la figure 4.2, nous pouvons dire que la distribution pour chaque paramètre estimé semble être une distribution normale. Finalement, les résultats obtenus des différentes études numériques semblent bien illustrer nos résultats théoriques.

Conclusion

L'objectif principal de cette thèse est l'estimation des paramètres des processus GARCH hyperboliques univariés et multivariés stationnaires. Pour réaliser cela, nous avons proposé deux approches différentes. La première, appliquée au cas univarié et résumée au chapitre 2, consiste à utiliser deux classes d'estimateurs, nommées M-estimateur et BM-estimateur, définies par une fonction de perte bornée. La seconde approche, présentée au chapitre 4, est relative à l'utilisation de la méthode du quasi-maximum de vraisemblance pour le modèle multivarié. Le choix des estimateurs définis par une fonction bornée est motivé par le fait qu'ils sont pratiquement insensibles à la présence de valeurs aberrantes dans les observations. Ce qui n'est pas le cas de l'estimateur du quasi-maximum de vraisemblance très souvent utilisé dans la littérature pour les modèles GARCH hyperboliques. Nous avons présenté au chapitre 2 les M et BM estimateurs et avons établi leurs propriétés asymptotiques à savoir la consistance et la normalité asymptotique. A travers des expériences de simulations numériques réalisées au chapitre 3, nous avons mis en compétition le M-estimateur, le BM-estimateur et l'estimateur du quasi-maximum de vraisemblance. Nous avons pu constater que dans le cas où les observations ne contiennent pas de valeurs aberrantes (outliers), les trois estimateurs en compétition ont des performances d'estimation et de prévision pratiquement similaires. Cependant, dans le cas où les observations contiennent 5% ou 10% de valeurs aberrantes, le BM-estimateur est le meilleur pour l'estimation, le M-estimateur pour la prévision et l'estimateur du quasi-maximum de vraisemblance a les pires performances d'estimation et de prévision. Une application à des données réelles a donné des résultats allant dans le sens de ceux des simulations. Au chapitre 4, nous avons construit un modèle GARCH hyperbolique multivarié à corrélation conditionnelle constante. Les conditions d'existence d'une solution stationnaire stricte de ce modèle ont été étudiées. Pour l'estimation de ses paramètres, nous avons opté pour la méthode du quasi-maximum de vraisemblance. Sous certaines conditions, nous voyons que cet estimateur est fortement consistant et asymptotiquement normal. Des simulations nu-

mériques réalisées dans le cas du modèle bivarié issu d'une distribution de Student à 13 degrés de liberté ont permis d'illustrer les résultats théoriques obtenus.

Puisque les M et BM estimateurs étudiés aux chapitres 2 et 3 se sont montrés robustes, il serait intéressant d'envisager pour des recherches futures l'estimation des paramètres du modèle GARCH hyperbolique univarié par d'autres méthodes d'estimation robustes. Nous pouvons proposer par exemple la méthode du minimum de distance de Hellinger, tout en faisant une étude comparative des performances de ces estimateurs robustes.

Pour le cas multivariée, une étude rigoureuse portant sur les conditions garantissant l'identifiabilité du modèle CCC-HGARCH pourraient être menée. En outre, nous pouvons étendre l'étude des M et BM estimateurs pour les paramètres du modèle GARCH hyperbolique multivarié défini au chapitre 4. La méthode du minimum de distance de Hellinger pourrait également être développée sur le modèle multivarié. Par ailleurs, l'invariance de la matrice de corrélation conditionnelle est l'un des handicap majeur du modèle multivarié présenté dans ce travail. Une idée intuitive serait de proposer un modèle GARCH hyperbolique à matrice de corrélation conditionnelle dynamique. Pour ce nouveau modèle, différentes méthodes pour l'estimation de ses paramètres peuvent également être étudiées à savoir les méthodes du quasi-maximum de vraisemblance, du M-estimateur, du BM-estimateur et du minimum de distance de Hellinger.

Liste des contributions et activités de recherches

Articles Scientifiques

- "M-estimate for the stationary hyperbolic GARCH models" publié dans METRON, Springer, Août 2021.
- "Asymptotic Properties of Quasi-Maximum Likelihood Estimator for a class of Multivariate Hyperbolic GARCH models" soumis.

Activités de Recherches

- 29 Juin 2022 : Présentation au séminaire de l'équipe Probabilités-Statistique AFRIMATH (Afrique France Réseau International de Mathématiques)
- Du 01 Avril au 31 Juillet 2019 : Séjour de recherches à Université GASTON BERGER de Saint-Louis, Sénégal, financé par le CEA-MITIC (Centre d'Excellence Africain en Mathématiques Informatique et TIC)
- Du 03 au 15 Septembre 2018 Participant à l'école de recherche du CIMPA (Centre Africain de Mathématiques Pures et Appliquées) Lomé 2018
- Depuis Juillet 2018 : Communications annuelles aux doctoriales de l'INP-HB

Annexe

Annexe A

Boîtes à moustaches des paramètres du modèle HGARCH

Nous présentons dans cette partie les boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) univarié avec les différentes méthodes d'estimation.

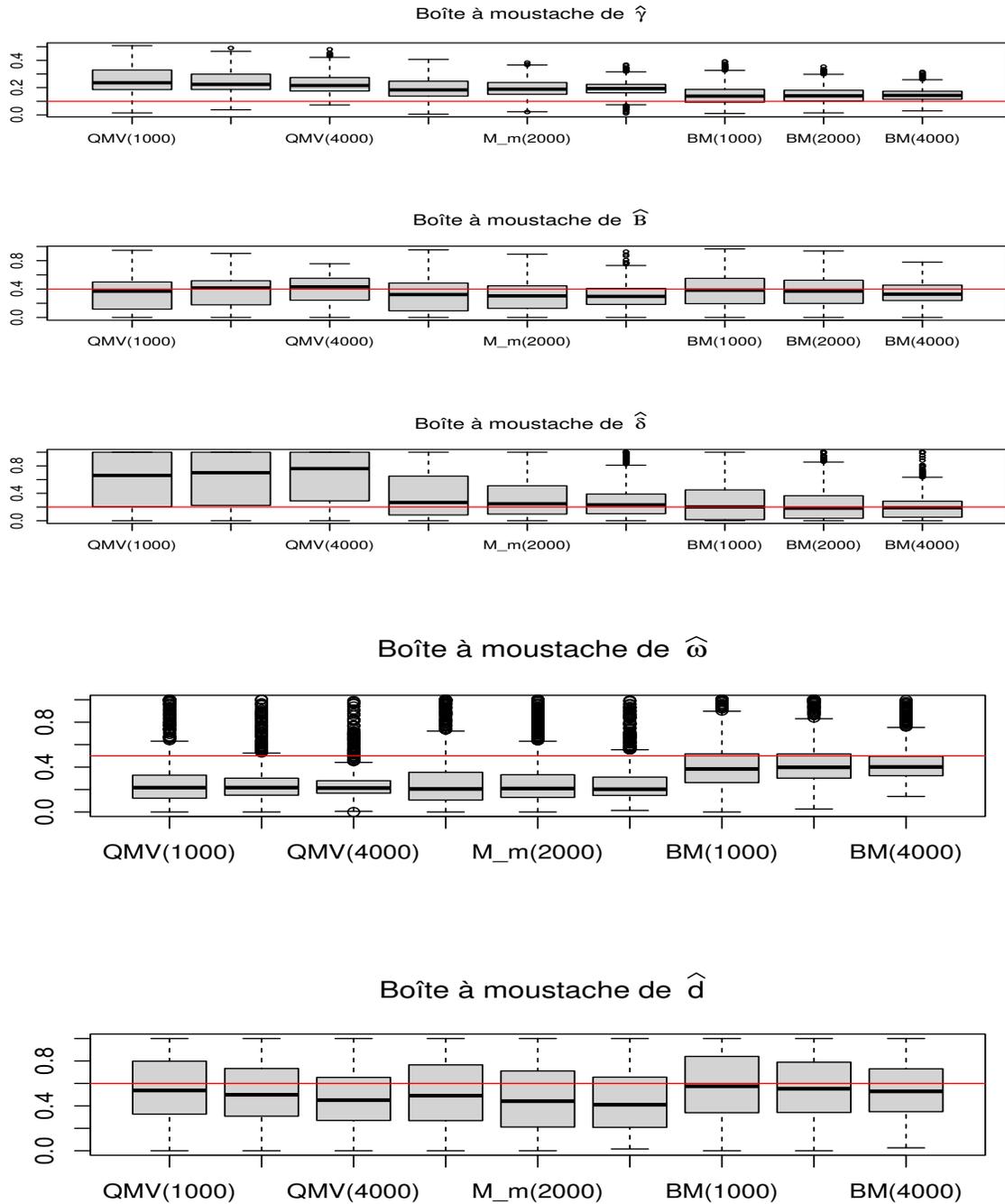


FIGURE A.1 – Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 5% of outliers et $e=3$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

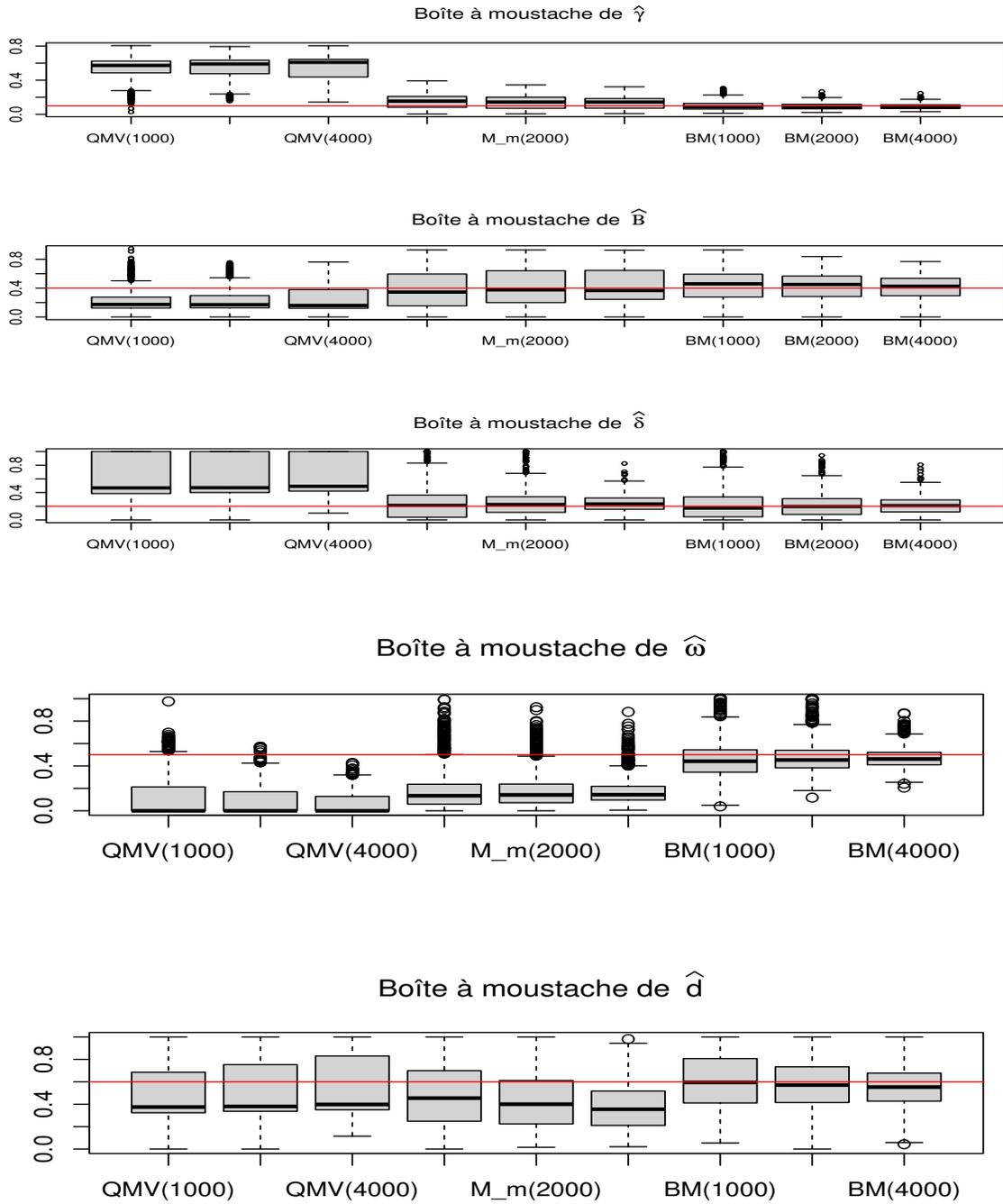


FIGURE A.2 – Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 5% of outliers et $e=5$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

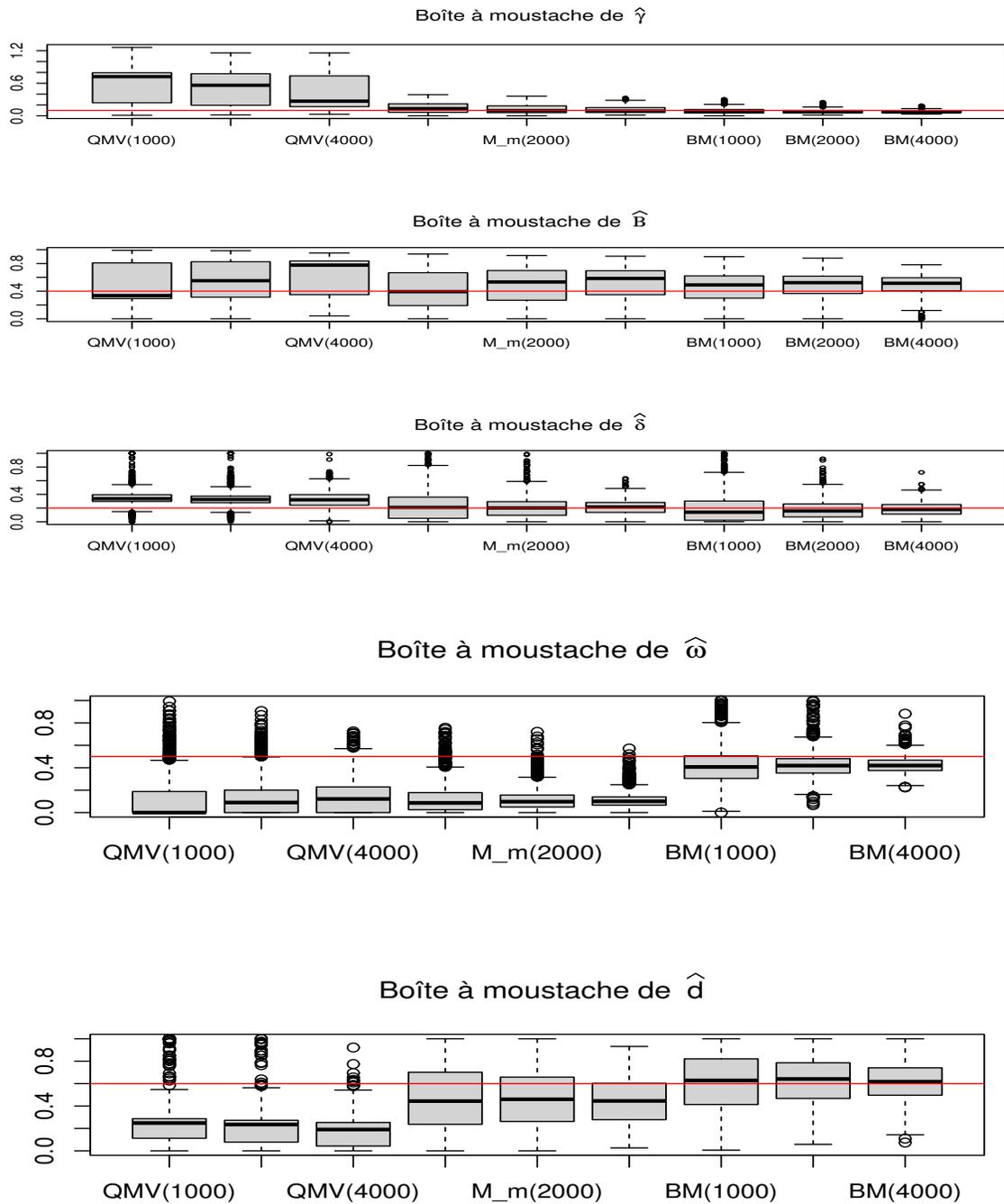


FIGURE A.3 – Boîtes à moustaches des paramètres estimés du modèle HGARCH(1,d,1) Normal avec 10% of outliers et $e=5$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

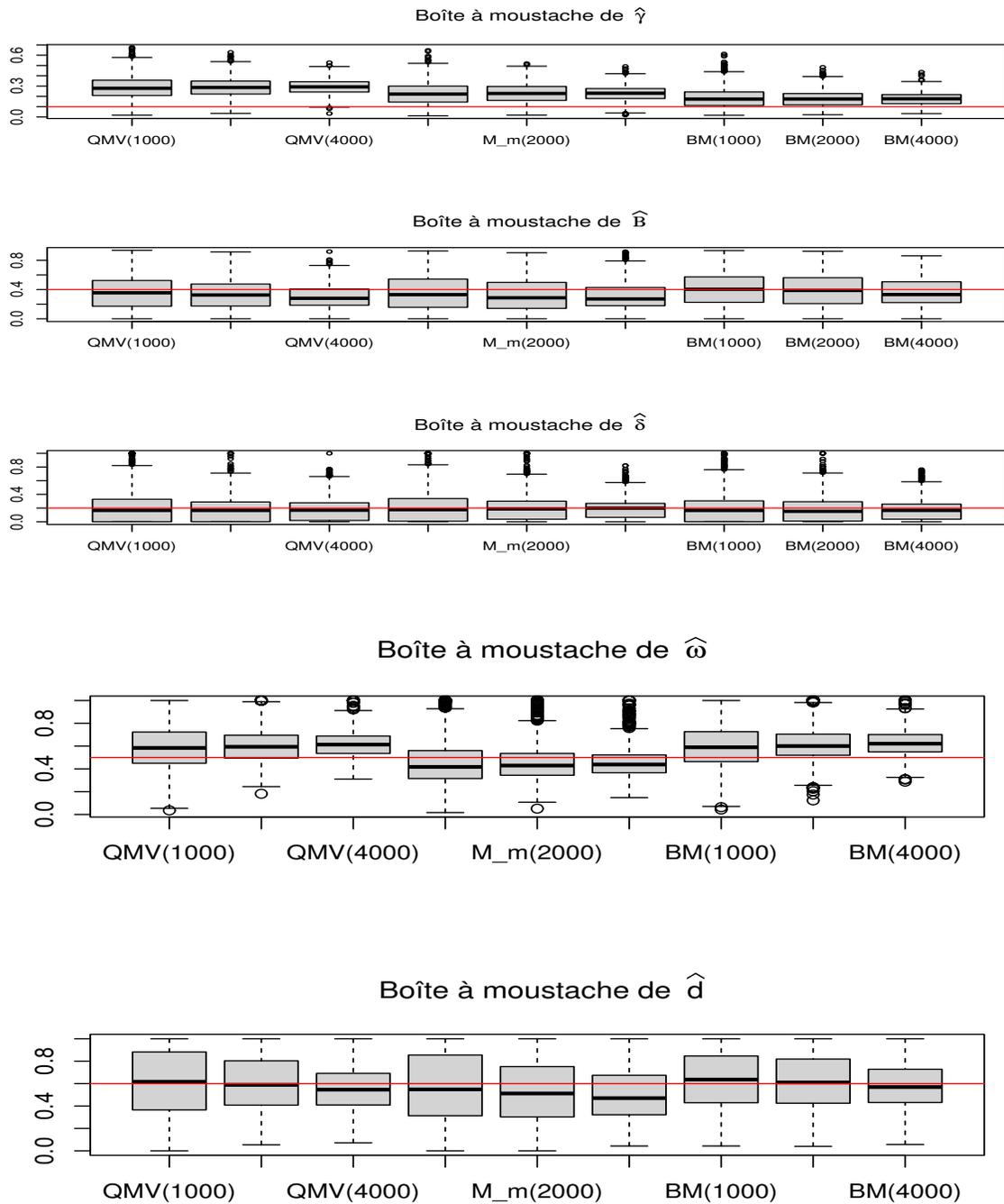


FIGURE A.4 – Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 5% of outliers et $e=3$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

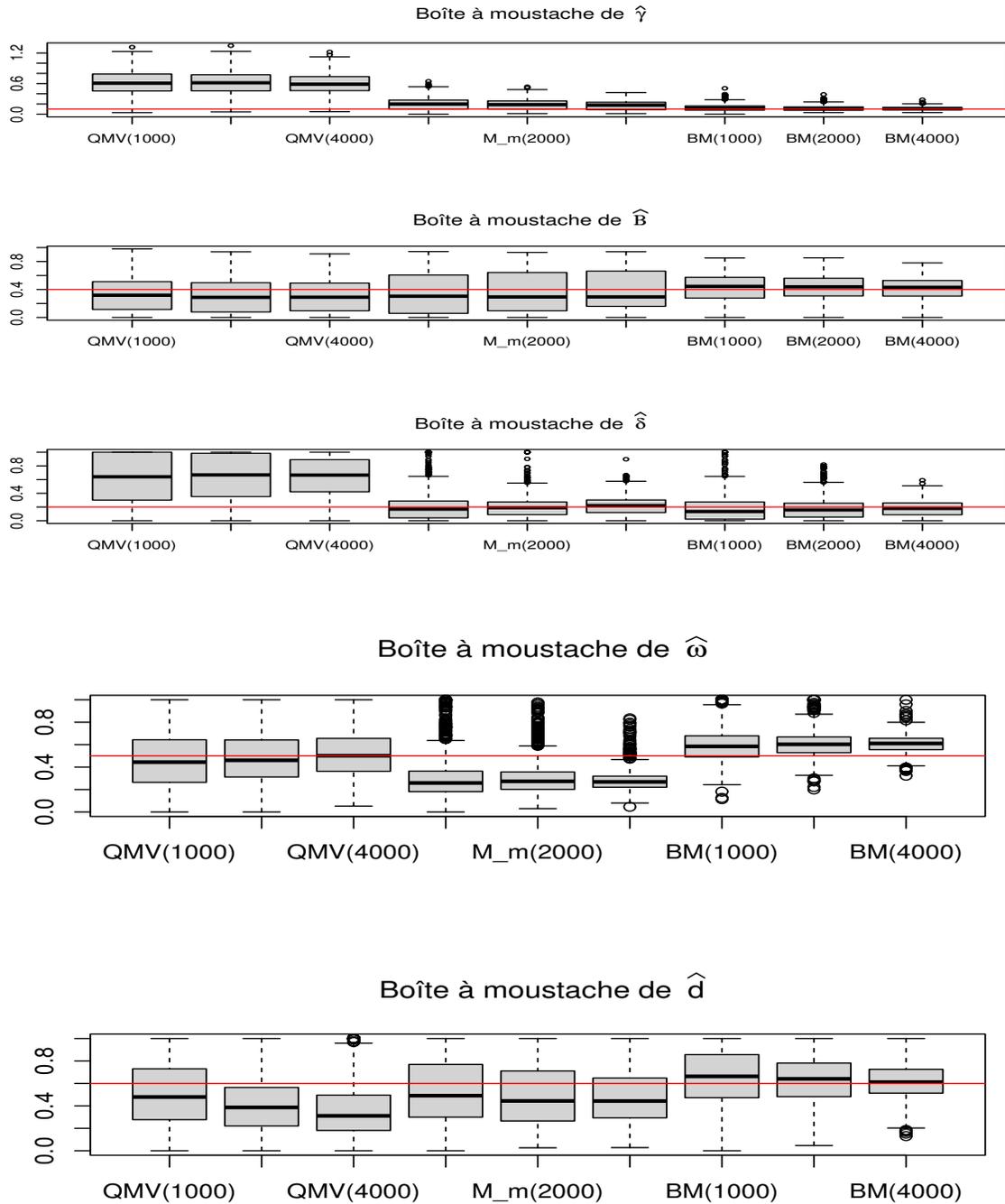


FIGURE A.5 – Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 5% of outliers et $e=5$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

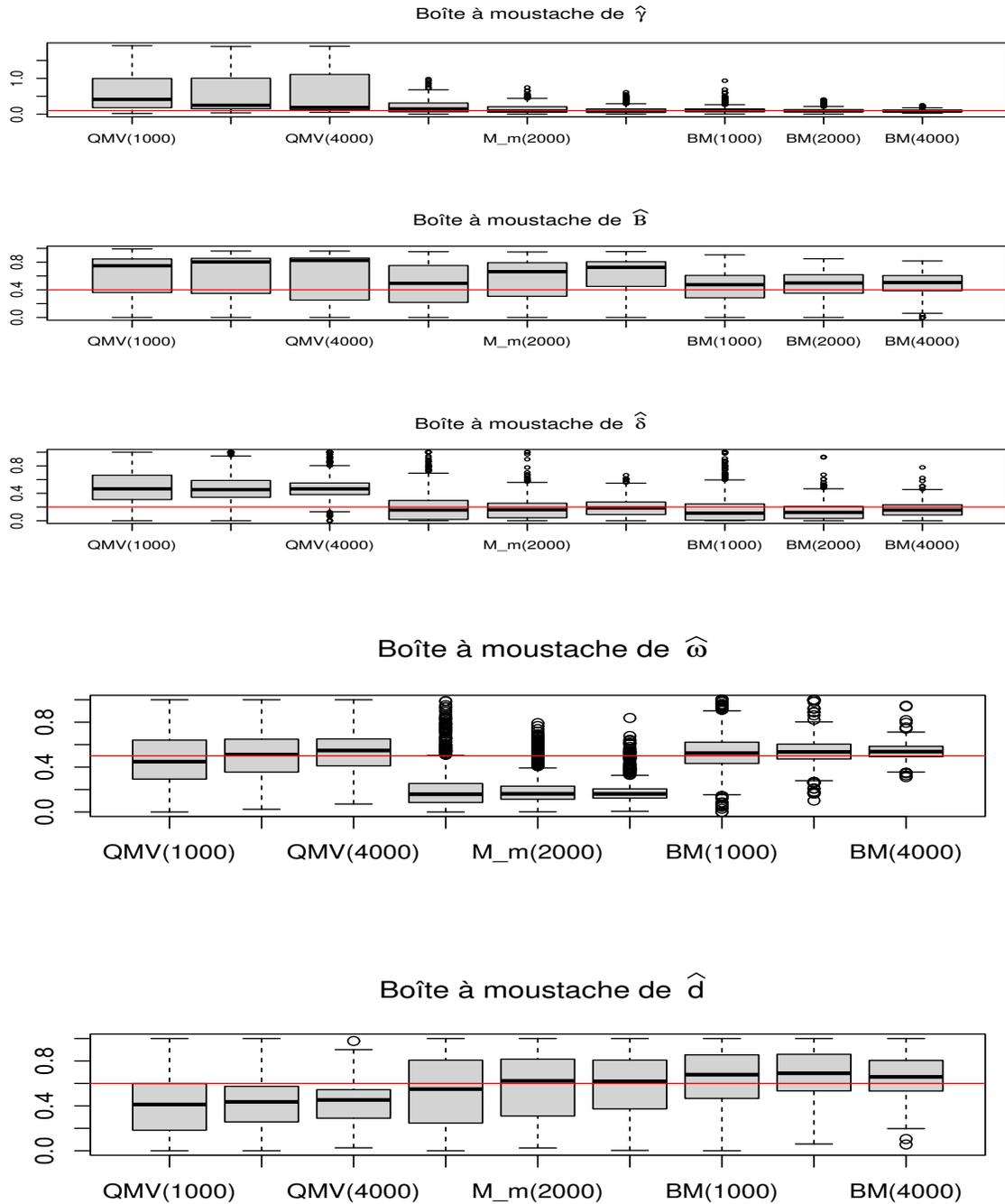


FIGURE A.6 – Boîtes à moustaches des paramètres estimés du modèle Student HGARCH(1,d,1) à sept degré de liberté avec 10% of outliers et $e=5$ pour les trois méthodes d'estimation. La ligne rouge désigne les valeurs du vrai paramètre.

Annexe B

Preuves Complémentaires

Nous présentons dans cette partie les démonstrations des points 2 et 4 du lemme 2.2.1.

Preuve du Lemme 2.2.1.2. L'objectif du point (2) de ce lemme est de montrer que les quantités suivantes $\sup_{\theta \in U} \|\nabla \log h_t(\theta)\|$ et $\sup_{\theta \in U} \|\nabla \log h_t(\theta)\|^2$ existent presque sûrement.

Pour cela, nous proposons une réécriture de la variance conditionnelle $h_t(\theta)$ en (2.1) en fonction des composantes du vecteur θ défini en (2.11). Nous obtenons alors la formulation suivante :

$$h_t(\theta) = \frac{\gamma}{1 - \beta(1)} + \frac{\omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}}{1 - \beta(L)} y_t^2 \quad (\text{B.1})$$

$$h_t(\theta) = \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{\infty} b_j(\theta^*) y_{t-j}^2, \quad \text{où} \quad (\text{B.2})$$

$$\sum_{j=1}^{\infty} b_j(\theta^*) L^j = \frac{\omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}}{1 - \beta(L)} \quad (\text{B.3})$$

avec

$$\theta^* = (\delta', \beta', \omega, d)', \quad (\text{B.4})$$

le vecteur contenant les mêmes paramètres que θ à l'exception de γ . Nous avons d'après (2.1), pour tout θ ,

$$\begin{aligned} h_t(\theta) &\geq \gamma \\ \frac{1}{h_t(\theta)} &\leq \frac{1}{\gamma} \leq \frac{1}{\gamma_m}, \end{aligned}$$

avec $\gamma_m = \inf_{\theta \in \Theta} \gamma(\theta)$ défini comme dans (2.37). Cela nous conduit à l'inégalité suivante :

$$\|\nabla \log h_t(\theta)\| = \left\| \frac{1}{h_t(\theta)} \nabla h_t(\theta) \right\| \leq \frac{1}{\gamma_m} \|\nabla h_t(\theta)\| \quad \text{avec} \quad \gamma_m = \inf_{\theta \in \Theta} \gamma(\theta). \quad (\text{B.5})$$

Par suite, le gradient de $h_t(\theta)$ donne

$$\nabla h_t(\theta) = \left(\frac{1}{1 - \beta(1)}, \frac{\partial h_t(\theta)}{\partial \theta^*} \right). \quad (\text{B.6})$$

Dans l'expression (B.6), la première composante du vecteur représente la dérivée partielle de $h_t(\theta)$ par rapport à γ , et la seconde composante représente les dérivées partielles de $h_t(\theta)$ par rapport aux composantes du vecteur θ^* défini en (B.4).

Nous souhaitons montrer dans le paragraphe suivant, l'existence de toutes les composantes du gradient défini en (B.6).

Pour cela, nous donnons les expressions des dérivées partielles de $h_t(\theta)$ définie en (2.1) par rapport aux composantes de θ^* . Ainsi, nous avons

$$\begin{aligned} \frac{\partial h_t(\theta)}{\partial \delta_i} &= \frac{\omega}{1 - \beta(L)} (1 - L)^d L^i y_t^2 \quad i = 1, \dots, q \\ \frac{\partial h_t(\theta)}{\partial \beta_k} &= \frac{\omega}{(1 - \beta(L)) (1 - \beta(L))} (1 - L)^d L^k y_t^2 \quad k = 1, \dots, p \\ \frac{\partial h_t(\theta)}{\partial \omega} &= 1 - \frac{\delta(L)}{1 - \beta(L)} (1 - L)^d y_t^2 \\ \frac{\partial h_t(\theta)}{\partial d} &= -\omega \frac{\delta(L)}{1 - \beta(L)} (1 - L)^d \log(1 - L) y_t^2. \end{aligned}$$

A partir de (2.2), nous adoptons les notations suivantes

$$\begin{aligned} \phi(L) &= \frac{1}{1 - \beta(L)} (1 - L)^d = \sum_{v=0}^{\infty} \phi_v L^v, \\ \mu(L) &= \frac{\delta(L)}{1 - \beta(L)} = \sum_{u=0}^{\infty} \mu_u L^u, \end{aligned}$$

en nous inspirant des arguments dans Kwan et al. [44] (section 1.7.2 du chapitre 1), les

dérivées partielles se réécrivent ainsi

$$\frac{\partial h_t(\theta)}{\partial \delta_i} = \omega \sum_{v=0}^{\infty} \phi_v L^v L^i y_t^2 = \sum_{v=0}^{\infty} \phi_v y_{t-i-v}^2, \quad i = 1, \dots, q \quad (\text{B.7})$$

$$\begin{aligned} \frac{\partial h_t(\theta)}{\partial \beta_k} &= \omega \sum_{u=0}^{\infty} \mu_u L^u \sum_{v=0}^{\infty} \phi_v L^v y_{t-k}^2 = \sum_{u=0}^{\infty} \sum_{v=0}^{\infty} \mu_u \phi_v y_{t-u-v-k}^2, \quad k = 1, \dots, p, \\ &= \sum_{l=0}^{\infty} \nu_l y_{t-l-k}^2, \quad k = 1, \dots, p, \end{aligned}$$

$$\begin{aligned} \frac{\partial h_t(\theta)}{\partial \omega} &= \left(1 - \sum_{u=0}^{\infty} \mu_u L^u \sum_{j=0}^{\infty} a_j L^j\right) y_t^2 = y_t^2 - \sum_{u=0}^{\infty} \sum_{j=0}^{\infty} \mu_u a_j y_{t-u-j}^2 \\ &= y_t^2 - \sum_{l=0}^{\infty} \rho_l y_{t-l}^2 \end{aligned}$$

$$\frac{\partial h_t(\theta)}{\partial d} = -\omega \sum_{u=1}^{\infty} \mu_u L^u \sum_{j=0}^{\infty} a_j L^j \sum_{j=1}^{\infty} \frac{L^j}{j} y_t^2 = \sum_{j=1}^{\infty} g_j y_{t-j}^2, \quad (\text{B.8})$$

avec les relations suivantes

$$|L| < 1, \quad \sum_{v=0}^{\infty} |\phi_v| < \infty, \quad \sum_{j=1}^{\infty} |g_j| < \infty \quad \text{et} \quad \sum_{u=0}^{\infty} |\mu_u| < \infty \quad (\text{B.9})$$

$$\sum_{u=0}^{\infty} \sum_{v=0}^{\infty} \mu_u \phi_v = \sum_{k=0}^{\infty} \nu_k < \infty, \quad \sum_{u=0}^{\infty} \sum_{j=0}^{\infty} \mu_u a_j = \sum_{l=0}^{\infty} \rho_l < \infty. \quad (\text{B.10})$$

Montrons l'existence des dérivées partielles.

D'après (B.7), nous avons pour tout $i = 1, \dots, q$

$$\sum_{v=0}^{\infty} \mathbb{E} |\phi_v y_{t-i-v}^2| = \sum_{v=0}^{\infty} |\phi_v| \mathbb{E}(y_{t-i-v}^2). \quad (\text{B.11})$$

Par stationnarité du processus (y_t) dans l'hypothèse (\mathbf{A}_1) , la relation (B.11) combinée à (B.9) donne

$$\sum_{v=0}^{\infty} \mathbb{E} |\phi_v y_{t-i-v}^2| \leq K \sum_{v=0}^{\infty} |\phi_v| < \infty.$$

D'après le Théorème de convergence monotone et le critère d'intégrabilité (section 1.7 du chapitre 1), nous avons

$$\frac{\partial h_t(\theta)}{\partial \delta_i} < \infty \quad \text{p.s.} \quad (\text{B.12})$$

De manière similaire, nous montrons que les séries définies en (B.7)-(B.8) convergent presque sûrement.

Dans la suite du travail, nous considérerons la formulation générale suivante pour les dérivées partielles définies en (B.7)-(B.8), pour tout $k = 1, \dots, (p + q + 2)$,

$$\frac{\partial h_t(\theta)}{\partial \theta_k^*} = \sum_{j=1}^{\infty} b_{j,k}^{(1)}(\theta) y_{t-j}^2. \quad (\text{B.13})$$

Avec les $b_{j,k}^{(1)}(\theta)$ qui représentent les coefficients des séries (B.7)-(B.8) issus des dérivées partielles premières de $h_t(\theta)$ par rapport aux $(p + q + 2)$ composantes du vecteur θ^* .

Pour montrer l'existence des dérivées partielles à partir de la relation (B.13), nous procédons de manière similaire à (B.11)-(B.12). Ainsi, nous avons pour tout $k = 1, \dots, (p + q + 2)$,

$$\sum_{j=1}^{\infty} \mathbb{E}(|b_{j,k}^{(1)}(\theta) y_{t-j}^2|) = \sum_{j=1}^{\infty} |b_{j,k}^{(1)}(\theta)| \mathbb{E}(y_{t-j}^2). \quad (\text{B.14})$$

D'après l'hypothèse (\mathbf{A}_6), (B.14) devient

$$\sum_{j=1}^{\infty} \mathbb{E}(|b_{j,k}^{(1)}(\theta) y_{t-j}^2|) \leq K \sum_{j=1}^{\infty} b_j^{(1-\eta)} \mathbf{E}(|y_{t-j}^2|).$$

En appliquant (2.3), la relation (B.14) devient

$$\sum_{j=1}^{\infty} \mathbb{E}(|b_{j,k}^{(1)}(\theta) y_{t-j}^2|) \leq K \sum_{j=1}^{\infty} j^{-(1+d)(1-\eta)} \mathbf{E}(|y_{t-j}^2|).$$

De plus, par l'hypothèse (\mathbf{A}_1) portant sur la stationnarité du processus, nous aboutissons à

$$\sum_{j=1}^{\infty} \mathbb{E}(|b_{j,k}^{(1)}(\theta) y_{t-j}^2|) \leq K_1 \sum_{j=1}^{\infty} j^{-(1+d)(1-\eta)}. \quad (\text{B.15})$$

Ainsi, la série des espérances (B.15) converge selon le critère de Riemann sous la condition $\eta < 1 - \frac{1}{1+d}$, en particulier pour $\eta < 1 - \frac{1}{1+\frac{d_m}{2}}$. En effet, selon le critère de Riemann, la série (B.15) converge pour $(1+d)(1-\eta) > 1$. Ce qui entraîne $\eta < 1 - \frac{1}{1+d}$. Selon (2.37), nous avons

$$\begin{aligned} \frac{d_m}{2} &< d_m < d \\ \frac{1}{1+\frac{d_m}{2}} &> \frac{1}{1+d_m} > \frac{1}{1+d} \\ 1 - \frac{1}{1+\frac{d_m}{2}} &< 1 - \frac{1}{1+d_m} < 1 - \frac{1}{1+d}. \end{aligned}$$

Ce choix particulier de η conduit également à la convergence de la série définie en (B.15). Par suite, en utilisant le Théorème de convergence monotone et le critère d'intégrabilité (section 1.7 du chapitre 1) nous obtenons finalement pour tout $k = 1, \dots, (p + q + 2)$,

$$\begin{aligned} \frac{\partial h_t(\theta)}{\partial \theta_k^*} &< \infty \quad \text{et} \quad \sum_{j=1}^{\infty} j^{-(1+d)(1-\eta)} y_{t-j}^2 < \infty \quad \text{p.s pour,} \\ \eta &< 1 - \frac{1}{1 + \frac{d_m}{2}}. \end{aligned} \quad (\text{B.16})$$

Etant donné que toutes les dérivées partielles qui composent le gradient existent, nous déterminons dans le paragraphe suivant la norme de celui-ci.

En utilisant (B.6), (B.16) et la définition de la norme dans (2.25)–(2.27), nous obtenons pour tout $\theta \in U$,

$$\begin{aligned} \|\nabla h_t(\theta)\| &= \frac{1}{1 - \beta(1)} + \sum_{k=1}^{p+q+2} \left| \frac{\partial h_t(\theta)}{\partial \theta_k^*} \right| \\ &\leq \frac{1}{1 - \beta(1)} + (p + q + 2) \sum_{j=1}^{\infty} j^{-(1+d)(1-\eta)} y_{t-j}^2 \\ &\leq K_2 \sum_{j=1}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \quad \text{avec} \quad \eta < 1 - \frac{1}{1 + \frac{d_m}{2}}, \end{aligned} \quad (\text{B.17})$$

il s'en suit que

$$\sup_{\theta \in U} \|\nabla h_t(\theta)\| \leq K_2 \sum_{j=1}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \quad \text{avec} \quad \eta < 1 - \frac{1}{1 + \frac{d_m}{2}}, \quad (\text{B.18})$$

ainsi

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla h_t(\theta)\|) < \infty. \quad (\text{B.19})$$

En utilisant (B.5) combiné à (B.19), nous aboutissons au premier résultat.

Nous rappelons que la norme définie pour les matrices (voir (2.25) – (2.27)) dans ce travail est une norme matricielle. Pour la démonstration du deuxième résultat du lemme 2.2.1, nous appliquons une propriété des normes matricielles qui sont sous-multiplicatives. Ce qui nous conduit à

$$\begin{aligned} \|\nabla \log h_t(\theta)(\nabla \log h_t(\theta))'\| &\leq \|\nabla \log h_t(\theta)\| \|(\nabla \log h_t(\theta))'\| \\ &\leq \|\nabla \log h_t(\theta)\|^2. \end{aligned} \quad (\text{B.20})$$

Par ailleurs, l'objectif dans cette partie sera de montrer l'existence de la quantité $\mathbf{E}(\sup_{\theta \in U} \|\nabla h_t(\theta)\|^2)$.

En utilisant la relation (B.6) et à partir de la norme choisie pour les vecteurs, nous obtenons

$$\begin{aligned} \|\nabla h_t(\theta)\|^2 &\leq \left(\frac{1}{1-\beta(1)}\right)^2 + \frac{2}{1-\beta(1)} \sum_{k=1}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right| + \sum_{k=1}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right|^2 + \\ &2 \sum_{k_1=1}^{p+q+2} \sum_{k_2=1, k_1 \neq k_2}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_{k_1}^*}\right| \left|\frac{\partial h_t(\theta)}{\partial \theta_{k_2}^*}\right|. \end{aligned} \quad (\text{B.21})$$

De plus, à partir de (B.13), nous obtenons pour tout $k = 1, \dots, (p+q+2)$

$$\left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right|^2 \leq \left(\sum_{j=1}^{\infty} |b_{j,k}^{(1)}(\theta)| y_{t-j}^2\right)^2 \leq \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| |y_{t-j_1}^2 y_{t-j_2}^2|. \quad (\text{B.22})$$

Nous étudions d'abord la convergence de la série des espérances suivante :

$$\sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbf{E}(y_{t-j_1}^2 y_{t-j_2}^2).$$

Par l'inégalité de Hölder, nous avons pour tout $k = 1, \dots, (p+q+2)$

$$\sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbf{E}(y_{t-j_1}^2 y_{t-j_2}^2) \leq \sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbf{E}(y_{t-j_1}^4)^{1/2} \mathbf{E}(y_{t-j_2}^4)^{1/2}.$$

Par ailleurs, l'utilisation de la condition $\omega < \left(\mathbf{E}|\varepsilon_t|^4\right)^{-1/2}$ dans l'hypothèse (\mathbf{A}_2) , nous garantit, d'après la conséquence du Théorème 2.1.1 pour $m = 2$, l'existence du moment d'ordre 4 du processus c'est-à-dire $(\mathbf{E}(y_t^4) < \infty)$. De plus, par l'inégalité de Jensen, pour tout $k = 1, \dots, (p+q+2)$, nous aboutissons à

$$\sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbf{E}(y_{t-j_1}^2 y_{t-j_2}^2) \leq \sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbf{E}(y_{t-j_1}^2) \mathbf{E}(y_{t-j_2}^2).$$

Enfin, en appliquant les mêmes arguments que dans (B.14) – (B.16), la série des espérances précédente est majorée par une série de Riemann qui converge sous la condition $\eta < 1 - \frac{1}{1+\frac{dm}{2}}$. Ainsi, nous obtenons

$$\sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| |y_{t-j_1}^2 y_{t-j_2}^2| < \infty \quad \text{p.s pour tout } k = 1, \dots, (p+q+2).$$

Ce qui entraîne d'après (B.22) et le Théorème de convergence monotone que

$$\left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right|^2 < \infty \quad \text{p.s pour tout } k = 1, \dots, (p+q+2). \quad (\text{B.23})$$

Par un raisonnement similaire à celui dans (B.14) – (B.18), $\|\nabla h_t(\theta)\|^2$ défini en (B.21) est dominée par une série convergente presque sûrement et indépendante de θ . Par suite, nous obtenons

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla h_t(\theta)\|^2) < \infty. \quad (\text{B.24})$$

D’après (B.5) l’inégalité (B.20) devient

$$\|\nabla \log h_t(\theta)(\nabla \log h_t(\theta))'\| \leq \|\nabla \log h_t(\theta)\|^2 \leq \frac{1}{\gamma_m^2} \|\nabla h_t(\theta)\|^2. \quad (\text{B.25})$$

En passant au sup dans (B.25) et d’après (B.24), nous aboutissons au second résultat.

Preuve du Lemme 2.2.1.4. Ce lemme consiste d’abord à montrer l’existence de la norme de la matrice hessienne de $h_t(\theta)$ et ensuite à montrer que le sup de la norme de la différence des matrices hessiennes est dominé. Nous définissons la matrice hessienne par

$$\nabla^2 h_t(\theta) = \left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} \right).$$

En utilisant les mêmes arguments que dans (B.6)-(B.8), nous définissons les dérivées partielles secondes de $\nabla^2 h_t(\theta)$ comme suit :

$$\begin{aligned} \frac{\partial^2 h_t(\theta)}{\partial \gamma^2} &= \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \delta_i} = \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \omega} = \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial d} = 0, \\ \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \beta_j} &= \frac{\partial h_{t-j}(\theta)}{\partial \gamma} + \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \gamma \partial \beta_j} = \frac{1}{1 - \beta_j} \frac{\partial h_{t-j}(\theta)}{\partial \gamma}, \quad j = 1, \dots, p, \\ \frac{\partial^2 h_t(\theta)}{\partial \delta_i^2} &= \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \delta_i^2} = 0, \quad i = 1, \dots, q, \end{aligned} \quad (\text{B.26})$$

$$\begin{aligned} \frac{\partial^2 h_t(\theta)}{\partial \omega \partial \delta_i} &= \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \omega \partial \delta_i} + (1 - L)^d y_{t-i}^2 = \frac{1}{1 - \beta(L)} (1 - L)^d y_{t-i}^2, \\ \frac{\partial^2 h_t(\theta)}{\partial \beta_j \partial \delta_i} &= \frac{\partial h_{t-j}(\theta)}{\partial \delta_i} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \beta_j \partial \delta_i} = \frac{1}{1 - \beta(L)} \frac{\partial h_{t-j}(\theta)}{\partial \delta_i}, \\ \frac{\partial^2 h_t(\theta)}{\partial d \partial \delta_i} &= \omega \times \frac{1}{1 - \beta(L)} \log(1 - L) (1 - L)^d y_{t-i}^2, \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 h_t(\theta)}{\partial \beta_j^2} &= 2 \frac{\partial h_{t-j}(\theta)}{\partial \beta_j} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \beta_j^2} = \frac{2}{1-\beta(L)} \frac{\partial h_{t-j}(\theta)}{\partial \beta_j}, \\
\frac{\partial^2 h_t(\theta)}{\partial \omega \partial \beta_j} &= \frac{\partial h_{t-j}(\theta)}{\partial \omega} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \omega \partial \beta_j} - y_{t-i}^2, \\
\frac{\partial^2 h_t(\theta)}{\partial \omega \partial \beta_j} &= \frac{1}{1-\beta(L)} \frac{\partial h_{t-j}(\theta)}{\partial \omega} - \frac{1}{1-\beta(L)} y_{t-i}^2, \\
\frac{\partial^2 h_t(\theta)}{\partial d \partial \beta_j} &= \frac{1}{1-\beta(L)} \frac{\partial h_{t-j}(\theta)}{\partial d}, \\
\frac{\partial^2 h_t(\theta)}{\partial \omega^2} &= \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \omega^2} = 0, \\
\frac{\partial^2 h_t(\theta)}{\partial d \partial \omega} &= \sum_{j=1}^p \frac{\partial^2 h_{t-j}(\theta)}{\partial d \partial \omega} - \log(1-L)(1-L)^d \delta(L) y_t^2, \\
\frac{\partial^2 h_t(\theta)}{\partial d \partial \omega} &= -\log(1-L)(1-L)^d \frac{\delta(L)}{1-\beta(L)} y_t^2, \\
\frac{\partial^2 h_t(\theta)}{\partial d^2} &= \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial d^2} - \omega \delta(L) (\log(1-L))^2 (1-L)^d y_t^2.
\end{aligned}$$

Nous récrivons les dérivées partielles seconde sous forme de séries en utilisant les résultats et le même principe que dans (B.6) – (B.8). Nous avons alors

$$\begin{aligned}
\frac{\partial^2 h_t(\theta)}{\partial \gamma^2} &= \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \delta_i} = \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \omega} = \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial d} = 0, \\
\frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \beta_j} &= \frac{1}{1-\beta_j} \frac{\partial h_{t-j}(\theta)}{\partial \gamma} = \frac{1}{1-\beta_j} \frac{1}{1-\beta(1)}, \quad j = 1, \dots, p, \\
\frac{\partial^2 h_t(\theta)}{\partial \delta_i^2} &= \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \delta_i^2} = 0, \quad i = 1, \dots, q, \\
\frac{\partial^2 h_t(\theta)}{\partial \omega \partial \delta_i} &= \frac{1}{1-\beta(L)} (1-L)^d y_{t-i}^2 = \sum_{v=0}^{\infty} \phi_v y_{t-i-v}^2, \quad i = 1, \dots, q, \\
\frac{\partial^2 h_t(\theta)}{\partial \beta_j \partial \delta_i} &= \frac{1}{1-\beta(L)} \frac{\partial h_{t-j}(\theta)}{\partial \delta_i} = \frac{1}{1-\beta(L)} \sum_{v=0}^{\infty} \phi_v y_{t-i-v}^2, \quad i = 1, \dots, q,
\end{aligned} \tag{B.27}$$

$$\begin{aligned}
\frac{\partial^2 h_t(\theta)}{\partial d \partial \delta_i} &= \frac{\omega(1-L)^d}{1-\beta(L)} \log(1-L) y_{t-i}^2 = -\omega \sum_{v=0}^{\infty} \phi_v L^v \sum_{j=1}^{\infty} \frac{L^j}{j} y_{t-i}^2, \quad i = 1, \dots, q, \\
\frac{\partial^2 h_t(\theta)}{\partial \beta_j^2} &= \frac{2}{1-\beta(L)} \sum_{k=0}^{\infty} \nu_k L^k y_{t-j}^2, \quad j = 1, \dots, p \\
\frac{\partial^2 h_t(\theta)}{\partial \omega \partial \beta_j} &= \frac{1}{1-\beta(L)} \left(1 - \sum_{l=0}^{\infty} \rho_l L^l\right) y_{t-i}^2 - \frac{1}{1-\beta(L)} y_{t-i}^2, \quad j = 1, \dots, p \\
\frac{\partial^2 h_t(\theta)}{\partial d \partial \beta_j} &= -\omega \frac{1}{1-\beta(L)} \sum_{l=0}^{\infty} \rho_l L^l \sum_{k=1}^{\infty} \frac{L^k}{k} y_{t-j}^2, \quad j = 1, \dots, p \\
\\
\frac{\partial^2 h_t(\theta)}{\partial \omega^2} &= \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial \omega^2} = 0 \\
\frac{\partial^2 h_t(\theta)}{\partial d \partial \omega} &= -\log(1-L)(1-L)^d \frac{\delta(L)}{1-\beta(L)} y_t^2 = -\sum_{k=1}^{\infty} \frac{L^k}{k} \sum_{l=0}^{\infty} a_l L^l \sum_{\mu=0}^{\infty} \mu_u L^u y_t^2 \\
\frac{\partial^2 h_t(\theta)}{\partial d^2} &= \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}(\theta)}{\partial d^2} - \omega \delta(L) (\log(1-L))^2 (1-L)^d y_t^2. \tag{B.28}
\end{aligned}$$

Ainsi, pour $i, j = 1, \dots, p + q + 3$, chaque dérivée partielle seconde (B.27)-(B.28) se réécrit sous forme de série dont la convergence presque sûre se démontre comme dans le cas du lemme 2.2.1.(2). Ainsi, de manière similaire au lemme 2.2.1.(2), nous adoptons également une écriture qui désigne une forme générale des dérivées secondes.

Dans cette forme générale, l'expression $b_k^{(2)}(i, j)$ caractérise les coefficients des différentes séries (B.27)-(B.28) pour $i, j = 1, \dots, (p + q + 3)$ et $k = 1, \dots, (p + q + 3)$. Nous posons donc que

$$\left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'}\right)_{i,j} = 0 \quad \text{ou} \quad \left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'}\right)_{i,j} = \sum_{k=1}^{\infty} b_k^{(2)}(i, j) y_{t-k}^2. \tag{B.29}$$

Ce qui permet de définir le carré de chaque éléments de la matrice hessienne $\nabla^2 h_t(\theta)$ comme suit : pour tout $i, j = 1, \dots, (p + q + 3)$, nous avons

$$\left|\left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'}\right)_{i,j}\right|^2 = \left|\sum_{k=1}^{\infty} b_k^{(2)}(i, j) y_{t-k}^2\right|^2 \leq \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} |b_{k_1}^{(2)}(i, j)| |b_{k_2}^{(2)}(i, j)| |y_{t-k_1}^2 y_{t-k_2}^2|.$$

Par suite, la norme de la matrice $\nabla^2 h_t(\theta)$ obtenue vérifie

$$\|\nabla^2 h_t(\theta)\| \leq \left[\sum_{i,j=1}^{p+q+3} \left(\sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} |b_{k_1}^{(2)}(i, j)| |b_{k_2}^{(2)}(i, j)| |y_{t-k_1}^2 y_{t-k_2}^2| \right) \right]^{1/2}. \tag{B.30}$$

En appliquant l'inégalité $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ pour tous a et b positifs à (B.30), nous obtenons alors

$$\|\nabla^2 h_t(\theta)\| \leq \sum_{i,j=1}^{p+q+3} \left(\sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} |b_{k_1}^{(2)}(i,j)| |b_{k_2}^{(2)}(i,j)| |y_{t-k_1}^2 y_{t-k_2}^2| \right)^{1/2}.$$

En utilisant les mêmes arguments que dans le lemme 2.2.1.(2), nous obtenons le résultat souhaité c'est-à-dire

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla^2 h_t(\theta)\|) < \infty.$$

De plus, suivant (B.30), nous avons

$$\|\nabla^2 h_t(\theta)\|^2 \leq \sum_{i,j=1}^{p+q+3} \left(\sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} |b_{k_1}^{(2)}(i,j)| |b_{k_2}^{(2)}(i,j)| |y_{t-k_1}^2 y_{t-k_2}^2| \right).$$

En utilisant les mêmes arguments que dans le lemme 2.2.1.(2) nous aboutissons à

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla^2 h_t(\theta)\|^2) < \infty. \quad (\text{B.31})$$

Pour montrer le second point du lemme 2.2.1.(4), nous procédons exactement comme dans la preuve du lemme 2.2.1.(3). En faisant la différence des deux matrices, nous obtenons pour tous $i, j = 1, \dots, (p+q+3)$,

$$\left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{h}_t(\theta)}{\partial \theta \partial \theta'} \right)_{ij} = \sum_{k=t}^{\infty} b_k^{(2)}(i,j) y_{t-k}^2.$$

D'où la norme de la différence vérifie

$$\left\| \frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{h}_t(\theta)}{\partial \theta \partial \theta'} \right\| \leq \left[\sum_{i,j=1}^{p+q+3} \left(\sum_{k=t}^{\infty} |b_k^{(2)}(i,j)| y_{t-k}^2 \right)^2 \right]^{1/2}. \quad (\text{B.32})$$

Afin de manipuler au mieux (B.32), nous utilisons des techniques similaires à celles employées dans Robinson et Zaffaroni [58]. Ainsi, nous avons, pour tout $i, j = 1, \dots, (p+q+3)$,

$$\sum_{k=t}^{\infty} |b_k^{(2)}(i,j)| y_{t-k}^2 = \sum_{k=t}^{\infty} |b_k^{(2)}(i,j)|^{1/2} |b_k^{(2)}(i,j)|^{1/2} y_{t-k}^2.$$

D'après l'hypothèse (\mathbf{A}_6) qui garantit l'existence de la série $\sum_{k_1=1}^{\infty} |b_{k_1}^{(2)}(i,j)|$ et par application de l'inégalité de Hölder, nous avons pour tout $i, j = 1, \dots, (p+q+3)$,

$$\left(\sum_{k=t}^{\infty} |b_k^{(2)}(i,j)|^{1/2} |b_k^{(2)}(i,j)|^{1/2} y_{t-k}^2 \right)^2 \leq \sum_{k_1=t}^{\infty} |b_{k_1}^{(2)}(i,j)| \sum_{k=t}^{\infty} |b_k^{(2)}(i,j)| y_{t-k}^4.$$

Ainsi, il existerait un nombre K_2 positif tel que pour tout $i, j = 1, \dots, (p + q + 3)$,

$$\left(\sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^2 \right)^2 \leq K_2 \sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^4. \quad (\text{B.33})$$

De plus, par les relations (B.32) – (B.33), nous avons

$$\|\nabla^2 h_t(\theta) - \nabla^2 \tilde{h}_t(\theta)\| \leq K_2 \sum_{i,j=1}^{p+q+3} \left(\sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^4 \right)^{1/2}.$$

Grâce aux mêmes techniques que (2.38), nous pouvons également avoir l'expression suivante

$$\begin{aligned} \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| &\leq \frac{1}{\tilde{h}_t(\theta)} \left(\|\nabla^2 h_t(\theta) - \nabla^2 \tilde{h}_t(\theta)\| + \right. \\ &\quad \left. \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} \right\| |h_t(\theta) - \tilde{h}_t(\theta)| \right). \end{aligned}$$

En utilisant les mêmes arguments que dans le lemme 2.2.1.(3) avec (B.31), nous obtenons le résultat . □

Annexe C

Article Publié



M-Estimate for the stationary hyperbolic GARCH models

Lanciné Bamba¹ · Ouagnina Hili¹ · Abdou Kâ Diongue² · Assi N'Guessan³

Received: 4 February 2021 / Accepted: 9 July 2021 / Published online: 10 August 2021
© Sapienza Università di Roma 2021

Abstract

In this manuscript, we propose two classes of M-estimates for the hyperbolic GARCH models. The first class called M-estimate is defined by minimizing of a convenient bounded loss function. The second, called BM-estimate is a modified version of the first with a mechanism that limits the propagation of the effect of outliers in the conditional variance. The asymptotic properties of these classes of M-estimates are established. According to the Monte Carlo study, we compare the performance of the M and BM-estimates with that of the quasi-maximum likelihood (QML) estimate. We show that the proposed M and BM-estimates are less affected by outliers than the QML-estimate. Moreover, in the last part, an empirical example indicates that the studied M-estimate is the best for the out-of-sample forecasting.

Keywords ARCH (∞)models · Long memory in volatility · M-estimate · BM-estimate · Outliers

Mathematics Subject Classification 62M10 · 62F12

1 Introduction

Many financial time series such as future interest rates, foreign exchange rates and asset price returns have their autocorrelation function which decays very slowly. These time series are said to present a long memory in volatility as related in Ding et al. [7] and Ding and Granger

Lanciné Bamba
lancine91090@gmail.com

Ouagnina Hili
o_hili@yahoo.fr

Abdou Kâ Diongue
kadiougue@yahoo.fr

Assi N'Guessan
assi.nguessan@polytech-lille.fr

¹ UMRI-Mathématiques et Nouvelles Technologies de l'Information, Institut National Polytechnique Félix Houphouët-Boigny, BP 1093 Yamoussoukro, Côte d'Ivoire

² LERSTAD, Université Gaston Berger de Saint-Louis, BP 234 Saint-Louis, Senegal

³ Laboratoire Paul Painlevé UMR CNRS 8524, Université de Lille, 59655 Villeneuve d'Ascq, France

[6]. The autoregressive conditional heteroscedasticity ARCH(p) and GARCH(p, q) models respectively studied by Engle [8] and Bollerslev [3] were the first in literature used to treat volatility in financial time series. ARCH(p) and GARCH(p, q) models exhibit short memory. They are therefore unsuitable to modelling and forecasting volatility in the financial time series which exhibit long memory. To correct these drawbacks, Baillie et al. [1] proposed the Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic (FIGARCH) processes. The FIGARCH model presents a long memory but have infinite variance which limits its applications. The defect of the FIGARCH model is solved by Davidson [5] which introduced an hyperbolic GARCH (HYGARCH) model. The HYGARCH model presents a long memory and has finite variance. However the conditional variance of HYGARCH model can have a complicated form. Moreover, the memory and the amplitude parameters are both related to the memory of the HYGARCH model. These reasons motivate Li et al. [13] to propose a new hyperbolic GARCH (HGARCH) model with finite variance, long memory and a simpler structure than HYGARCH. To estimate the parameters of HGARCH model, Li et al. [13] used a quasi-maximum likelihood estimate (QMLE) method. Many authors like Robinson and Zaffaroni [17] and Zakoian and Francq [4] p.141–148 have shown that QMLE is consistent and asymptotically normal under certain conditions for ARCH(∞) and GARCH models respectively. However, when outliers appear in data, classical estimation methods such as QMLE are very affected by their presence in the sample. Outliers are defined by Lu et al. [14] as observations that deviate significantly from the majority of data. To deal data with outliers, some researchers use a robust estimation method and others detect and suppress them. Two principal types of outliers have been proposed by Fox [9] for time series called additive outliers and innovation outliers. Many authors have proposed some estimation methods to deal with outliers. Muler and Yohai [15] proposed two classes of robust estimates called τ estimate and τ -filtered estimate for ARCH(p) models. Muler and Yohai [16] proposed two classes of robust estimates for GARCH models. The first, called M-estimate is obtained by maximizing a bounded loss function. The second class of robust estimates of Muler and Yohai [16] called Bounded M-estimate(BM-estimate) is obtained by modification of the first. Lu et al. [14] extended the work of Muler and Yohai [16] to the multiplicative error models (MEM) and proposed two types of loss functions to generate their M-estimates and BM-estimates.

Muler and Yohai [16] and Lu et al. [14] have shown that the M and BM-estimates with bounded loss function are less influenced by outliers than the QML-estimate for the GARCH and the MEM models respectively. Hence, these estimates have better forecasting performance than QMLE when the data are contaminated by outliers for these models. Muler and Yohai [16] and Lu et al. [14] explained their results because the logarithmic function present in the QML-estimate is unbounded. These results have motivated the authors of this work to extend the study of Muler and Yohai [16] and Lu et al. [14] to HGARCH models. Our main objective is to propose some robust estimation methods namely M-estimate and BM-estimate for the hyperbolic GARCH models to deal with presence of outliers in the data. The M-estimate of the present work is motivated by Muler and Yohai [16] and the BM-estimate by Lu et al. [14]. We also study the asymptotic properties of these robust estimates as well as their estimation and forecasting performances when the data are contaminated or not by outliers. In addition, we will compare the M and BM-estimate performance with that of the QML-estimate which is used for HGARCH models until then.

This work is organized as follows. In Sect. 2, we present the HGARCH model and the M and BM-estimate. In Sect. 3, we give the asymptotic property such as consistency and the asymptotic normality for the M-estimates. In Sect. 4, we performed a Monte Carlo simulation to compare the performances of M-estimate, BM-estimate and QML-estimate. In Sect. 5, we

consider a real data analysis. In Sect. 6, we make some conclusions. The proofs of the main results are given in Sect. 7.

2 M-Estimate of the HGARCH models

2.1 Presentation of the HGARCH model

The new hyperbolic GARCH (HGARCH(p, d, q)) model presented by Li et al. [13] can be written as

$$y_t = h_t^{1/2} \varepsilon_t, \quad (1)$$

$$h_t = \gamma + \beta(L)h_t + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}y_t^2, \quad (2)$$

where ε_t is an independent identically distributed random variables sequence with a continuous density f_1 such that ε_t has mean 0 and variance 1. The conditional variance is $h_t = \text{var}(y_t | \mathcal{F}_{t-1})$ where $\mathcal{F}_{t-1} = \sigma(\{\varepsilon_s\}_{s \leq t-1})$ denotes the filtration generated by the past information of ε_t until time $t - 1$. The parameters γ and ω are positive, L is the backshift operator, $\beta(L) = \sum_{i=1}^p \beta_i L^i$, $\delta(L) = 1 - \sum_{j=1}^q \delta_j L^j$ with $\beta_i \geq 0$ for $i = 1, \dots, p$, $\delta_j \geq 0$ for $j = 1, \dots, q$, $0 < d < 1$ and $(1 - L)^d = 1 - \sum_{j=1}^{\infty} \pi_j L^j$ with

$$\pi_j = \frac{d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} = O(j^{-1-d}) \quad \text{as } 0 < d < 1.$$

Li et al. [13] also write this model as the following $ARCH(\infty)$ form

$$h_t = \frac{\gamma}{1 - \beta(1)} + \omega\left(1 - \frac{\delta(L)}{1 - \beta(L)}(1 - L)^d\right)y_t^2 = \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{\infty} b_j y_{t-j}^2 \quad (3)$$

where $b_j = O(j^{-1-d})$, $b_j > 0$ for $j \geq 1$ and $\sum_{j=1}^{\infty} b_j < \infty$.

The stationarity properties of the process y_t are given in Theorem 1 and Theorem 2 of Li et al. [13].

2.2 Robust estimates of HGARCH models

Let y_1, y_2, \dots, y_n be a sample generated by a HGARCH(p, d, q) model defined in (1)–(3) and $\theta_0 = (\gamma_0, \beta'_0, \delta'_0, \omega_0, d_0)' \in \mathbb{R}^{p+q+3}$ the true parameter vector of this model. As well as Muler and Yohai [16], we suppose that

$$x_t = \log(y_t^2), \quad (4)$$

$$w_t = \log(\varepsilon_t^2), \quad (5)$$

This implies that

$$x_t = w_t + \log(h_t(\theta_0)).$$

From Lu et al. [14] and Muler and Yohai [16], we recall that if f_1 , the density function of ε_t defined in Sect. 2.1, is symmetric around 0, then the density function g of w_t equals to

$$g(w) = f_1(e^{(w/2)})e^{(w/2)}. \quad (6)$$

Moreover if f_1 corresponds to the normal $\mathcal{N}(0, 1)$ distribution, then $g = g_0$ and

$$g_0(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(e^w - w)}. \quad (7)$$

Let $\theta = (\gamma, \beta', \delta', \omega, d)' \in \mathbb{R}^{p+q+3}$ be the parameter vector. The function $h_t(\theta)$ defined in (3), is the conditional variance of y_t up to time t knowing all its past. For the rest of the work, we assume that p and q are known. They can also be estimated by AIC or BIC. We define

$$M_n(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \rho(x_t - \log h_t(\theta)). \quad (8)$$

In practice, the observations $y_0^2, \dots, y_{1-q}^2, \dots$ are not available. Following Straumann [19] p.73-75, we set the following initial values for unavailable observations.

$$\begin{aligned} y_t = 0 \quad \text{and} \quad \tilde{h}_t(\theta) &= \frac{\gamma}{1 - \beta(1)} \quad \text{for } t \leq 0, \\ \tilde{h}_t(\theta) &= \gamma + \beta(L)\tilde{h}_t(\theta) + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}y_t^2, \quad t \geq 1 \\ &= \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{t-1} b_j(\theta)y_{t-j}^2 \end{aligned} \quad (9)$$

where $\tilde{h}_t(\theta)$ is the conditional variance of y_t computed with its initial values. Following the same lines as Muler and Yohai [16], we define the QML-estimate of HGARCH(p, d, q) model by minimizing the function

$$M_{0,n}(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \rho_0(x_t - \log \tilde{h}_t(\theta)), \quad (10)$$

where

$$\rho_0 = -\log(g_0), \quad (11)$$

g_0 is given by (7), and x_t is defined by (5).

Muler and Yohai [16] and Lu et al. [14] have shown that for the GARCH and the MEM models respectively, the QML-estimate is very sensitive to the presence of a few outliers in the sample. One reason is that the function ρ_0 is unbounded. In the present work, we propose two estimation methods the M-estimate and the BM-estimate based on bounded function for HGARCH models.

The first estimation method called the M-estimate of HGARCH(p, d, q) model is defined by minimizing the function

$$\tilde{M}_n(\theta) = \frac{1}{n-p} \sum_{t=p+1}^n \rho(x_t - \log \tilde{h}_t(\theta)). \quad (12)$$

Then, the M-estimate of HGARCH(p, d, q) model is defined by

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \tilde{M}_n(\theta) \quad (13)$$

However, the conditional variance $h_t(\theta)$ defined in (3) is also very sensitive to outliers. Indeed, when an outlier appears in the series at time t , the calculation of $h_{t+u}(u > 0)$ is affected because it relies on its past observations. To analyse this issue, we propose as in

Muler and Yohai [16] and Lu et al. [14] the second estimation method of this work called BM-estimate. This BM-estimate is defined as in Lu et al. [14]. It's a modification version of M-estimate, by including a mechanism which limits the propagation of the effect of outliers on the predictors of the conditional variance. For this purpose, we use the following expression for the predictor of the conditional variance.

$$\bar{h}_{t,f}(\theta) = \gamma + \beta(L)\bar{h}_{t,f}(\theta) + \omega\{1 - \beta(L) - \delta(L)(1 - L)^d\}\bar{h}_{t,f}(\theta)s_f\left(\frac{y_t^2}{\bar{h}_{t,f}(\theta)}\right), \tag{14}$$

with $y_t = 0$ for $t \leq 0$ and

$$s_f(u) = \begin{cases} u & \text{if } u \leq f, \\ f & \text{if } u > f. \end{cases} \tag{15}$$

Hence, if f is large, $h_t(\theta)$ and $\bar{h}_{t,f}(\theta)$ are close. The function $M_n(\theta)$ defined in (8) can be modified as

$$\bar{M}_{nf}(\theta) = \frac{1}{n - p} \sum_{t=p+1}^n \rho(x_t - \log \bar{h}_{t,f}(\theta)). \tag{16}$$

The BM-estimate ($\hat{\theta}_n^B$) of HGARCH(p, d, q) model is defined by

$$\hat{\theta}_n^B = \arg \min_{\theta \in \Theta} \bar{M}_{nf}(\theta) \tag{17}$$

Θ is the parameter space, a compact set of \mathbb{R}^{p+q+3} . The function ρ defined in Muler and Yohai [16] is continuous, bounded, monotonic increasing function and has three bounded continuous derivatives.

3 Main results

In this section, we establish the asymptotic properties for the M-estimate ($\hat{\theta}_n$) of HGARCH model. We state the following assumptions.

- (A₀) The process y_t is a strictly stationary and ergodic solution of the HGARCH model defined in (1) and (3).
- (A₁) The polynomial $\beta(L)$ is such that $\sum_{j=1}^p \beta_j < 1, j = 1, \dots, p$.
- (A₂) Θ is a compact set of \mathbb{R}^{p+q+3} , $\theta_0 \in \Theta$ is the real parameter and $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$.
- (A₃) ρ is continuous, bounded, monotonic increasing function and has three bounded continuous derivatives.
- (A₄) Let $J(u) = \mathbf{E}(\rho(w_t - u))$. Then there exists a function ρ that respects (A₃) so that $J(u)$ takes the minimum $J(u_0)$ at a unique value u_0 .
- (A₅) Let ψ be the first-order derivative of ρ . We have $\mathbf{E}(\psi(x_t - \log h_t(\theta_0))) = 0, \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0))) < \infty$ and $\mathbf{E}(\psi'(x_t - \log \tilde{h}_t(\theta_0))) > 0$.
- (A₆) For all $j \geq 1, b_j(\theta)$ has the k-th order derivatives to be continuous on \mathbb{R}^{p+q+2} such that

$$\left| \frac{\partial^k b_j(\theta)}{\partial \theta_{i_1} \dots \partial \theta_{i_k}} \right| \leq K b_j^{1-\eta}(\theta) \quad \text{for all } \eta > 0,$$

θ_i is the i-th element of θ .

(A₇) Let g_θ be a continuous and differentiable function such that $h_{t+1}(\theta) = g_\theta(y_t, h_t)$. The components of the vector $\frac{\partial g_\theta}{\partial \theta'}(y_0, h_0)|_{\theta=\theta_0}$ are linearly independent random variables (see the condition N4 in Straumann [19] p. 85-86;116).

The assumptions (A₀), (A₁) and (A₂) are standard. The condition $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$ is sufficient to ensure the existence of the fourth order moment of the process.

Assumptions (A₃) and (A₄) are crucial for consistency. We used them to show the existence and the uniqueness of the parameter for which $\mathbf{E}(\rho(x_t - \log h_t(\theta)))$ takes the minimum.

The assumption (A₆) is the same as the hypothesis $F(l)$ in Robinson and Zaffaroni [17]. It allows with the technical assumptions (A₅) and (A₇) to establish the lemmas used for the proof of the asymptotic normality.

For satisfying empirically the following theorems, the time series should be a strictly stationary and ergodic long memory GARCH model with finite fourth moment.

Theorem 1 shows that under some assumptions, the function $M(\theta)$ takes the minimum at the unique value θ_0 .

Theorem 1 *Suppose that the assumptions (A₀), (A₄) and (A₇) are satisfied. Let $x_t = \log(y_t^2)$ and define for $\theta \in \mathbb{R}^{p+q+3}$*

$$M(\theta) = \mathbf{E}(\rho(x_t - \log h_t(\theta))). \quad (18)$$

Then

1. $M(\theta)$ is minimized when $\gamma = e^{u_0}\gamma_0$, $\omega = e^{u_0}\omega_0$, for $1 \leq i \leq p$ $\beta_i = \beta_{0i}$, $1 \leq j \leq q$, $\delta_j = \delta_{0j}$, $d = d_0$.
2. Suppose that $w_t = \log(\varepsilon_t^2)$ has a density $g(w)$ that's continuous, positive and unimodal for all w . If we take $\rho = m(-\log(g))$, where m is monotone, (A₄) holds with $u_0 = 0$.

The following theorems give the consistency and the convergence in distribution of $\hat{\theta}_n$.

Theorem 2 *Under the assumptions (A₀) - (A₄) and (A₇) with $u_0 = 0$, the M-estimate $\hat{\theta}_n$ converges almost surely to θ_0 ,*

$$\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta_0 \text{ a.s.}$$

Theorem 3 *If (A₀)-(A₇) are satisfied with $u_0 = 0$, then*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} \mathcal{N}[0, A(\theta_0)^{-1}B(\theta_0)A(\theta_0)^{-1}],$$

with

$$A(\theta_0) = \mathbf{E}_{\theta_0}(\nabla^2 \rho(x_t - \log h_t(\theta))),$$

$$B(\theta_0) = \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0)))D_0 \text{ and } D_0 = \mathbf{E}[\nabla \log h_t(\theta_0)(\nabla \log h_t(\theta_0))']$$

For the same arguments as in Lu et al. [14], we obtain that the consistency and asymptotic normality of the M-estimate $\hat{\theta}_n$ also hold for the BM-estimate $\hat{\theta}_n^B$.

The proofs of Theorems 1-3 are in Appendix 7.

4 Simulation studies

This section conducts a Monte Carlo study to investigate the performance of three estimation methods for the HGARCH model : the QML-estimate (QML) defined in (10), the

M-estimate (M_m) and the BM estimate both based on the loss function ρ defined in (A₃). When ε_t follows the normal distribution $N(0, 1)$, s_f defined in (15) is equal to identity with $P(s_{5.02}(\varepsilon_t^2) = \varepsilon_t^2) = 0.975$. This is what leads to the choice of $f = 5$ for this simulation study. There are two steps in this study. In the first step, ε_t follows a Normal distribution. In the second, ε_t follows a Student-t distribution with seven freedom degree. In each step, we consider three cases. First, ε_t is normal and there is no outliers in the data. Second, ε_t is normal with 5% of additive outliers. Third, ε_t is normal with 10% of additive outliers. The series y_t^* with additive outliers are defined as in Muler and Yohai [16].

$$y_t^* = \begin{cases} y_t + eh_t^{1/2} & \text{if } t = t_j, \quad 1 \leq j \leq r = kn/100 \\ y_t & \text{elsewhere} \end{cases} \tag{19}$$

where y_t is the noncontaminated series observed in HGARCH models with normal ε_t . k is the percentage of contamination, t_1, \dots, t_r represent the time when outliers appear and were chosen equally spaced. After trying, we consider for this study two values for e , $e = 3$ and $e = 5$. We explore both cases of e values to evaluate the influence of this parameter on the result. The HGARCH(1, d, 1) model used in this study is defined in Li et al. [13].

The function ρ is defined as in Muler and Yohai [16] by $\rho = m(-\log g_0)$, with

$$m(x) = \begin{cases} x & \text{if } x \leq 4 \\ R(x) & \text{if } 4 < x \leq 4.30 \\ 4.15 & \text{if } x > 4.30, \end{cases}$$

where

$$R(x) = 0.5(x^4 - a^4) - \frac{2}{3}(2a + b)(x^3 - a^3) + \frac{(a^2 + 2ab)(x^2 - a^2)}{(b - a)^3} - \frac{(2ba^2(x - a))}{(b - a)^3} - \frac{(x - a)^3}{3(b - a)^2} + x,$$

with $a = 4, \quad b = 4.3$.

These choices of $R(x)$, a and b in Muler and Yohai [16] ensure that m is continuous and smooth. We consider three sample sizes for the simulation, 1000, 2000 and 4000 respectively. 80% of observations are used to the estimation and the last 20% for the out-of-sample forecasting. The different forecast horizons are set to 1, 5, 10 and 20. Indeed, as in Li et al. [13], we considered the 1-step ahead and the 5-step ahead as shorter horizons while 10 and 20 ahead forecasts are treated as the long term forecast. There are also $N = 1000$ replications for each sample size. For the generated sequence, we have discarded the first 2000 observations as in Li et al. [13], to mitigate the effect of initialisations. So $2000 + n$ observations will be generated each time. The true parameter vector used to generate the observations is

$$\theta_0 = (\gamma_0, \beta_0, \delta_0, \omega_0, d_0) = (0.1, 0.4, 0.2, 0.5, 0.6).$$

To compare the performance among the M-estimate, the BM-estimate and the QML-estimate, we use the RMSE and the MAE for estimation. For the forecasting comparison, we only use the RMSE.

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad \text{and} \quad MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t|,$$

where \hat{y}_t denotes the predicted value of y_t .

Table 1 contains the RMSE and MAE values for the Quasi Maximum likelihood (QML) estimate, the M-estimate (M_m) and the Bounded M-estimate (BM) for the normal HGARCH(1,d,1) without outliers. From Table 1, the RMSE and MAE for the estimation part of all the estimate methods decrease when the sample size increases.

It looks that for estimation the QML estimate, the M-estimate and the BM-estimate are almost similar but the QML estimate is a little better when there are no outliers in the HGARCH(1,d,1) model. For the forecasting results, all the methods are also almost similar. We note that the QML estimate is little better because it behaves generally the smallest values. From the out-of-sample forecasting results, with the p-values of the Diebold-Mariano test, we generally accept the null hypothesis that assume that the forecasts have the same accuracy taken in pairs for the 10 and 20 step ahead.

In Table 2, we report the RMSE and the MAE values for 5% contaminated sample for the normal HGARCH(1,d,1) model. In this case the parameter e defined in (19) is set to 3. For the estimation part, both the RMSE and MAE of BM-estimate method decrease as the sample size increases. It's not the case for the RMSE and MAE of parameters δ and d for the QML estimate and the parameter d for the M-estimate. In addition, the BM estimate presents the smallest MAE and RMSE values. So, it behaves the best for estimation. From the out-of-sample forecasting results, with the p-values of the Diebold-Mariano test in parenthesis, we conclude that all the forecasts have the same accuracy taken in pairs. We observe that the M-estimate has the smallest RMSE, therefore it performs the best for the 5, 10 and 20 ahead forecast. The highest values of RMSE at 1 ahead forecast are due to the apparition of one outlier at this horizon during the insertion of additive outliers in the generated data. Because the outliers forecasting is not the aims of this study, we cannot compared these estimation methods for the 1 ahead forecast in the rest of the study.

In Table 3, we report the RMSE and MAE for 5% contaminated sample for the normal HGARCH(1,d,1) model with $e = 5$. For estimation, both the RMSE and MAE for BM-estimate method decrease as the sample size increases. It'snt the case of parameter d for the M-estimate. All the parameters of QML-estimate are disturbed by the presence of outliers. Moreover, the BM-estimate has the smallest RMSE and MAE values. It is better for estimation. From the out-of-sample forecasting results, using the same argument as in Table 2 for the forecast accuracy, we note that the BM and M-estimates are close. But, the M-estimate has the smallest RMSE, therefore it is the best for the short and the long horizons forecast excepted the 1-ahead . The highest values of RMSE at 1 ahead forecast are explained by the apparition of large outliers at this time in the generated data.

After studying different values of e , we can conclude that for the normal HGARCH(1,d,1) model with 5% outliers contamination the BM-estimate is more robust than the other estimate methods. It is so the best for estimation and the M-estimate will be prefer for the forecasting. The QML is the worst for estimation and forecasting.

In Table 4, we summarize the RMSE and MAE values for the normal HGARCH(1,d,1) model with 10% of outliers contamination for $e = 3$. For the estimation part, we observe the decreasing of the RMSE only for the parameters γ and β of QML estimate, δ and d for the M-estimate and d for the BM-estimate as the sample size increases. All the estimation methods are disturbed by this type and proportion of outliers. From both indices MAE and RMSE, the M-estimate will be preferred because it better controls the long memory parameter (d), which characterizes the properties of the HGARCH model. From the out-of-sample forecasting results, the p-values of the Mariano-Diebold test show that all the forecasts have the same accuracy taken in pairs. Also, the M-estimate has the smallest RMSE for all the horizons excepted the 1-ahead that correspond to the outlier position.

Table 1 Results for the normal HGARCH(1,d,1) without outliers

<i>n</i>	Estimation					
		γ	β	δ	ω	d
1000	QML	0.04361152	0.21071733	0.18531022	0.14781896	0.28407642
<i>RMSE</i>	M_m	0.04637462	0.21347503	0.19956722	0.16168414	0.28187780
	BM	0.0444476	0.2055457	0.2033027	0.1589536	0.2755895
1000	QML	0.0342135	0.1742121	0.1428745	0.1112259	0.2486029
<i>MAE</i>	M_m	0.0360574	0.1763166	0.1502495	0.1220998	0.2471377
	BM	0.0345099	0.1695302	0.1517171	0.1193955	0.2400837
2000	QML	0.03429911	0.18467751	0.13484239	0.10624846	0.23219205
<i>RMSE</i>	M_m	0.03572537	0.18539711	0.14022858	0.12291937	0.23774955
	BM	0.03485032	0.17928535	0.13608130	0.12163424	0.22657021
2000	QML	0.0280271	0.1578663	0.1094460	0.0805565	0.1964174
<i>MAE</i>	M_m	0.0287165	0.1564427	0.1115852	0.0908726	0.2020973
	BM	0.0277052	0.1504490	0.1079263	0.0899925	0.1902408
4000	QML	0.02655261	0.14789295	0.11457942	0.07717198	0.17047913
<i>RMSE</i>	M_m	0.02742995	0.15315357	0.11539842	0.08854180	0.18248984
	BM	0.02589646	0.14443635	0.11133267	0.09524189	0.17441386
4000	QML	0.0220159	0.1267519	0.0912920	0.0595648	0.1382659
<i>MAE</i>	M_m	0.0225451	0.1315196	0.0915822	0.0672422	0.1499602
	BM	0.0208985	0.1228190	0.0875777	0.0733259	0.1417378
RMSE for Out-of-sample Forecast						
		1-ahead	5-ahead	10-ahead	20-ahead	
n=200						
	QML	0.4969 (0.534) ¹	0.5804 (0.631) ¹	0.5045 (0.534) ¹	0.6378 (0.871) ¹	
	M_m	0.4958 (0.000) ²	0.5815 (0.055) ²	0.5068 (0.234) ²	0.6421 (0.608) ²	
	BM	0.4948 (0.000) ³	0.5811(0.069) ³	0.5099 (0.231) ³	0.6436 (0.662) ³	
n=400						
	QML	0.5671 (0.149) ¹	0.5065 (0.187) ¹	0.5339 (0.417) ¹	0.5214(0.149) ¹	
	M_m	0.5682 (0.0000) ²	0.5060 (0.011) ²	0.5335 (0.079) ²	0.5218(0.346) ²	
	BM	0.5689 (0.0000) ³	0.5067(0.008) ³	0.5374 (0.077) ³	0.5293(0.294) ³	
n=800						
	QML	0.4918 (0.139) ¹	0.5391 (0.142) ¹	0.4867 (0.297) ¹	0.4610(0.139) ¹	
	M_m	0.4961 (0.000) ²	0.5385 (0.015) ²	0.4863 (0.073) ²	0.4616(0.151) ²	
	BM	0.4980 (0.000) ³	0.5385 (0.003) ³	0.4905 (0.027) ³	0.4685(0.080) ³	

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 2 Results for the normal HGARCH(1,d,1) with 5% of additive outliers and $e = 3$

n	Estimation					
		γ	β	δ	ω	d
1000	QML	0.1824507	0.2389641	0.5594601	0.3168024	0.3022920
<i>RMSE</i>	M_m	0.1225048	0.2466524	0.3924226	0.3304879	0.3205431
	BM	0.08825207	0.24250316	0.32239109	0.23873986	0.29328744
1000	QML	0.1567677	0.1955715	0.4668623	0.2907232	0.2566926
<i>MAE</i>	M_m	0.1003156	0.2031034	0.2885874	0.3035852	0.2755282
	BM	0.0664480	0.2000983	0.2395116	0.1964150	0.2524568
2000	QML	0.1647756	0.2190157	0.5680752	0.2982690	0.3040075
<i>RMSE</i>	M_m	0.1168934	0.2334681	0.3385703	0.3147113	0.3279849
	BM	0.07689233	0.22260976	0.25345363	0.20714702	0.28940060
2000	QML	0.1433223	0.1766261	0.4749507	0.2791704	0.2573172
<i>MAE</i>	M_m	0.0995834	0.1901209	0.2424024	0.2918625	0.2822235
	BM	0.0599613	0.1828199	0.1897606	0.1697019	0.2452585
4000	QML	0.1519731	0.2069074	0.5819863	0.2892006	0.3113580
<i>RMSE</i>	M_m	0.1097818	0.2080384	0.2808735	0.3010942	0.3278539
	BM	0.06359509	0.17851374	0.19118510	0.17245534	0.26633820
4000	QML	0.1310870	0.1698337	0.4978103	0.2755681	0.2651011
<i>MAE</i>	M_m	0.0977251	0.1664175	0.1958090	0.2809762	0.2831434
	BM	0.0516965	0.1459989	0.1416375	0.1428029	0.2216879
RMSE for Out-of-sample Forecast						
		1-ahead	5-ahead	10-ahead	20-ahead	
n=200						
	QML	3.9332(0.986) ¹	0.6026 (0.984) ¹	0.6199 (0.986) ¹	0.6294 (0.998) ¹	
	M_m	3.9954(0.714) ²	0.5813 (0.602) ²	0.6040 (0.706) ²	0.5037 (0.745) ²	
	BM	3.9861(0.925) ³	0.5832 (0.920) ³	0.6069 (0.929) ³	0.5509 (0.988) ³	
n=400						
	QML	3.9175(0.548) ¹	0.5083 (0.567) ¹	0.5292 (0.595) ¹	0.5727(0.892) ¹	
	M_m	3.9748(0.369) ²	0.4807 (0.394) ²	0.4957 (0.434) ²	0.5247(0.810) ¹	
	BM	3.9610(0.632) ³	0.4837 (0.647) ³	0.5029 (0.670) ³	0.5335(0.919) ³	
n=800						
	QML	3.7245(0.580) ¹	0.7621 (0.613) ¹	0.5159 (0.641) ¹	0.5016(0.883) ¹	
	M_m	3.7788(0.479) ²	0.7485 (0.514) ²	0.4814 (0.554) ²	0.4506(0.767) ²	
	BM	3.7705(0.6348) ³	0.7493 (0.664) ³	0.4900 (0.687) ³	0.4579(0.910) ³	

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 3 Results for the normal HGARCH(1,d,1) with 5% of additive outliers and $e = 5$

n	Estimation	γ	β	δ	ω	d
1000	QML	0.4518902	0.2519396	0.5201002	0.4221391	0.2643575
	M_m	0.1005877	0.2705183	0.2636832	0.3620779	0.3086049
	BM	0.05153823	0.22127280	0.22321691	0.17737535	0.25853110
1000	QML	0.4259810	0.2367324	0.4359100	0.3940847	0.2472110
	M_m	0.0823139	0.2320896	0.1902356	0.3398584	0.2604685
	BM	0.0406182	0.1861499	0.1691550	0.1378854	0.2162802
2000	QML	0.4584906	0.2494533	0.5231717	0.4321780	0.2648688
	M_m	0.09058128	0.26341472	0.2000807	0.35350364	0.30909291
	BM	0.03900253	0.18963719	0.16686540	0.13309793	0.22634063
2000	QML	0.4323235	0.2366489	0.4417509	0.4128225	0.2504236
	M_m	0.0741354	0.2257466	0.1443882	0.3326238	0.2586171
	BM	0.0320692	0.1590653	0.1288769	0.1050128	0.1838603

Table 3 continued

<i>n</i>	Estimation					
	γ	β	δ	ω	d	
4000	0.4671310	0.2512758	0.5348963	0.4458519	0.2543854	
<i>RMSE</i>	0.07613541	0.23680213	0.14095148	0.34634290	0.30902588	
	0.02974118	0.15332327	0.13119966	0.09548084	0.18348735	
4000	0.4394668	0.2386203	0.4598530	0.4346389	0.2422358	
<i>MAE</i>	0.0634537	0.2032611	0.1096697	0.3315560	0.2616048	
	0.0248348	0.1291357	0.1034101	0.0768272	0.1497912	
RMSE for Out-of-sample Forecast						
<i>n=200</i>						
QML	9.0302(0.107) ¹	0.7066 (0.388) ¹	1.4168 (0.502) ¹	52.318 (0.768) ¹		
M_m	9.3083(0.243) ²	0.4326 (0.286) ²	0.5263 (0.352) ²	0.5948 (0.699) ²		
BM	9.2814(0.106) ³	0.4399 (0.388) ³	0.5342 (0.502) ³	0.6008 (0.768) ³		
<i>n=400</i>						
QML	9.6258(0.071) ¹	0.7098 (0.370) ¹	0.9192 (0.493) ¹	10.400(0.766) ¹		
M_m	9.8883(0.196) ²	0.5033 (0.232) ²	0.5172 (0.276) ²	0.5669(0.734) ²		
BM	9.848(0.070) ^{3,4}	0.5021 (0.369) ³	0.5237 (0.492) ³	0.5726(0.763) ³		
<i>n=800</i>						
QML	9.4842(0.868) ¹	0.7063 (0.823) ¹	0.7582 (0.843) ¹	1.3591102(0.981) ¹		
M_m	9.7395(0.205) ²	0.5254 (0.238) ²	0.5114 (0.297) ²	0.4798(0.711) ²		
BM	9.7061(0.995) ³	0.5277 (0.994) ³	0.5168 (0.994) ³	0.4928(0.993) ³		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 4 Results for the normal HGARCH(1,d,1) with 10% of additive outliers and $e = 3$

n	Estimation	γ	β	δ	ω	d
1000	QML	0.3490146	0.2645680	0.5756572	0.4169015	0.2973865
	M_m	0.2630635	0.2919219	0.4506256	0.4437230	0.3121641
	BM	0.2516656	0.2941109	0.4914643	0.4202650	0.3145411
1000	QML	0.3207306	0.2222750	0.4904890	0.3991049	0.2690081
	M_m	0.2403757	0.2642968	0.3517488	0.4258935	0.2860985
	BM	0.2283504	0.2641756	0.3907854	0.3991946	0.2879458
2000	QML	0.3465734	0.2480457	0.6116746	0.4243433	0.3001224
	M_m	0.2827183	0.2887196	0.4275370	0.4604038	0.3015558
	BM	0.2690417	0.2859154	0.4947990	0.4425642	0.3099522
2000	QML	0.3202454	0.2033926	0.5335604	0.4088358	0.2750252
	M_m	0.2672020	0.2675340	0.3333156	0.4455246	0.2813127
	BM	0.2512289	0.2587626	0.3930219	0.4269325	0.2906585

Table 4 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.3395143	0.2293033	0.6480537	0.4319478	0.2995103
	M_m	0.2910831	0.2933086	0.3962053	0.4697166	0.2903102
	BM	0.2788411	0.2885405	0.5036131	0.4525467	0.2968972
4000	QML	0.3174852	0.1787691	0.5842657	0.4189693	0.2723853
	M_m	0.2779050	0.2787670	0.3092967	0.4581473	0.2731245
4000	QML	0.2647884	0.2636224	0.4062010	0.4394239	0.2775671
	BM					
RMSE for Out-of-sample Forecast						
n=200	1-ahead			10-ahead		20-ahead
	QML	3.9046(0.511) ¹	0.5719 (0.332) ¹	0.6023 (0.620) ¹	0.6019(0.869) ¹	
	M_m	3.9903(0.846) ²	0.5072 (0.751) ²	0.5204 (0.761) ²	0.5190 (0.708) ²	
	BM	3.9904(0.471) ³	0.5102 (0.298) ³	0.5285 (0.592) ³	0.5358 (0.858) ³	
	QML	4.2126(0.521) ¹	0.5985(0.348) ¹	0.6157 (0.521) ¹	0.6286(0.851) ¹	
	M_m	4.2808 (0.892) ²	0.5245 (0.850) ²	0.5455(0.854) ²	0.5576(0.942) ²	
n=800	BM	4.2801(0.497) ³	0.5262 (0.322) ³	0.5510 (0.622) ³	0.5672(0.844) ³	
	QML	3.9106(0.530) ¹	0.6528 (0.359) ¹	0.5724 (0.640) ¹	0.7397(0.869) ¹	
	M_m	3.9874(0.885) ²	0.5903 (0.872) ²	0.4894 (0.702) ²	0.6821(0.982) ²	
BM	3.9906(0.524) ³	0.5906 (0.357) ³	0.4900 (0.635) ³	0.6843(0.871) ³		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

In Table 5, we summarize the RMSE and MAE values for the normal HGARCH(1,d,1) model with 10% of outliers contamination for $e = 5$. For estimation, the RMSE and MAE values for all the parameters of BM-estimate decrease as the sample size increases unlike the M and the QML estimates. The BM-estimate has also the smallest RMSE and MAE values. It behaves more robust and the better. From the out-of-sample forecasting results, the p-values of the Diebold–Mariano test show that the forecasts have generally same accuracy. We note, from the RMSE values that the QML has the worst performance. The M-estimate has the smallest RMSE for the 5, 10 and 20 ahead forecast, therefore it performs the best.

Moreover, the Figures 3 and 4 in Appendix 7.2 present the boxplots of the HGARCH(1,d,1) parameters for the different estimation methods with 5% of outliers contamination. These figures show that for the finite sample the distribution of the BM and the M-estimates are more symmetric than that of the QML-estimate. The variances of the BM-estimate parameters are also less than those of the M-estimate and the QML estimate. The other boxplots, for 10% of outliers contamination and the Student distribution are presented in Appendix 7.2 and provide almost similar results.

We now consider that the ε_t process follows the Student-t distribution with seven freedom degrees (t_7). Table 6 contains the RMSE and MAE results for the Quasi Maximum likelihood (QML) estimate, the M-estimate (M_m) and the Bounded M-estimate (BM) for the Student's t_7 HGARCH(1,d,1) without outliers. From Table 6, the RMSE and MAE for the parameters of all the estimates methods generally decrease when the sample size increases.

It looks that for estimation, the RMSE and MAE results for the BM-estimate are generally smallest. So BM-estimate is the better when there are no outliers in the t_7 HGARCH(1,d,1) model. For the forecasting results, the p-values of Diebold-Mariano test show that the different forecasts have generally the same accuracy taken in pairs. We also note that the M-estimate has the better performance for short and long horizon forecast.

In Table 7, we report the RMSE and MAE results for 5% contaminated sample for the t_7 HGARCH(1,d,1) model with $e = 3$. For the estimation part, the RMSE and MAE of all the estimate methods decrease as the sample size increases. In addition, the BM-estimate has the smallest RMSE and MAE values. It behaves the best for estimation. From the out-of-sample forecasting results, with the p-values of the Diebold- Mariano test we conclude that the forecasts have generally the same accuracy taken in pairs. We also observe that the M-estimate has the smallest RMSE for long and short horizon excepted the 1-ahead forecast. Therefore it performs the best for the short and long forecast horizons. The highest values of RMSE at 1-ahead forecast are due to the apparition of outliers at this horizon in the generated data.

In Table 8, we report the RMSE and MAE values for 5% contaminated sample for the t_7 HGARCH(1,d,1) model with $e = 5$. For estimation, both the RMSE and MAE values for BM and M-estimate methods generally decrease as the sample size increases. It'snt the case for all the parameters of QML-estimate. Moreover, the BM-estimate presents the smallest RMSE and MAE values while QML and M-estimates are very affected by the presence of outliers. The BM-estimate is better for estimation. From the out-of-sample forecasting results, the Diebold-Mariano test indicates that the forecast accuracy for all method are the same taken in pairs. We note that the M-estimate has the smallest values, therefore it is the best for short and long forecast horizon. Also, the QML estimate has the worst performance. The highest values of RMSE at 1 ahead forecast are explained by the apparition of outliers at this time in the generated data.

After studying different values of e , we can conclude that for the estimation of t_7 HGARCH(1,d,1) model with 5% outliers contamination the BM-estimate is more robust than M and QML estimates. But for forecasting, we would prefer the M-estimate.

Table 5 Results for the normal HGARCH(1,d,1) with 10% of additive outliers and $e = 5$

n	Estimation					
		γ	β	δ	ω	d
1000	QML	0.5501989	0.2822957	0.2388864	0.4177522	0.4067638
	M_m	0.1045551	0.2685897	0.2263825	0.3950856	0.3133597
	BM	0.05622594	0.22811116	0.21373243	0.19377377	0.26376515
1000	QML	0.4693620	0.2301027	0.1820533	0.3945965	0.3870595
	M_m	0.0834447	0.2359551	0.1728466	0.3745398	0.2647210
	BM	0.0447219	0.1951067	0.1655397	0.1538771	0.2211813
2000	QML	0.4952832	0.3023534	0.2140564	0.4032478	0.4204120
	M_m	0.08620839	0.26333737	0.16675999	0.39160136	0.28710061
	BM	0.04286325	0.20270592	0.14565109	0.14326776	0.22103315
2000	QML	0.4023883	0.2512016	0.1632025	0.3771393	0.3999923
	M_m	0.06777219	0.2328813	0.1248515	0.3777633	0.2328527
	BM	0.0367684	0.1748798	0.1161040	0.1136829	0.1815490

Table 5 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.4048013	0.3344157	0.1878658	0.3741157	0.4475092
	M_m	0.06420363	0.23992399	0.11429608	0.39244864	0.26899302
	BM	0.03558152	0.16938805	0.10684028	0.10766980	0.17592334
4000	QML	0.305757	5 0.2903262	0.1519310	0.3469654	0.4290392
	M_m	0.0510370	0.2081071	0.0904891	0.3847413	0.2148922
	BM	0.0315424	0.1463628	0.0856704	0.0907294	0.1417067
RMSE for Out-of-sample Forecast						
				10-ahead		20-ahead
n=200	1-ahead					
QML		9.9947(0.334) ¹	0.9447 (0.128) ¹	0.9868 (0.448) ¹		1.1100 (0.795) ¹
M_m		10.344(0.143) ²	0.5785 (0.035) ²	0.5637 (0.288) ²		0.4775 (0.561) ²
BM		10.321(0.362) ³	0.5905(0.149) ³	0.5731 (0.473) ³		0.4996 (0.812) ³
n=400						
QML		9.1181(0.293) ¹	0.8742 (0.112) ¹	0.9473 (0.427) ¹		0.9670(0.744) ¹
M_m		9.4576(0.140) ²	0.5053 (0.035) ²	0.5574 (0.279) ²		0.4865 (0.440) ²
BM		9.4380(0.318) ³	0.5108 (0.131) ³	0.5552(0.450) ³		0.4978 (0.765) ³
n=800						
QML		9.1177(0.285) ¹	0.8564(0.114) ¹	0.9395 (0.429) ¹		0.9687 (0.708) ¹
M_m		9.4037 (0.154) ²	0.5221 (0.063) ²	0.5952(0.342) ²		0.5054(0.154) ²
BM		9.3951(0.308) ³	0.5270 (0.129) ³	0.5957 (0.447) ³		0.5156(0.738) ³

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 6 Results for the student HGARCH(1,d,1) with 7 degree of freedom without outliers

<i>n</i>	Estimation					
		γ	β	δ	ω	<i>d</i>
1000	QML	0.09183176	0.23980786	0.17317487	0.20881758	0.28760310
	M_m	0.07066585	0.23744129	0.17829237	0.16498721	0.27705762
	BM	0.07015049	0.22272819	0.17078276	0.22687306	0.26764230
1000	QML	0.0693544	0.2019195	0.1410889	0.1730103	0.2523373
	M_m	0.0542833	0.2013467	0.1401859	0.1282793	0.2391913
	BM	0.05298725	0.18790208	0.13553592	0.19398930	0.22888948
2000	QML	0.06879443	0.22685563	0.14799033	0.20351628	0.24512754
	M_m	0.05465952	0.22021419	0.14525310	0.13959519	0.23359104
	BM	0.05081348	0.20909129	0.14098730	0.22429108	0.22973904
2000	QML	0.0539783	0.1943295	0.1256835	0.1763391	0.2040543
	M_m	0.0432261	0.1887166	0.1206872	0.1102202	0.1932087
	BM	0.0402513	0.1776479	0.1171477	0.2004755	0.1894803

Table 6 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.06271285	0.19250059	0.13029092	0.19252975	0.18103072
	M_m	0.04714152	0.18944588	0.13263131	0.11835210	0.17697064
	BM	0.04130632	0.17764559	0.13117694	0.21894204	0.17141760
4000	QML	0.0503974	0.1617177	0.10883393	0.17647090	0.14252188
	M_m	0.0382459	0.1596679	0.1101325	0.1001134	0.1376027
	BM	0.0334107	0.1481438	0.1092638	0.2062580	0.13406586
RMSE for Out-of-sample Forecast						
		10-ahead			20-ahead	
	1-ahead	5-ahead				
n=200						
QML	1.4417(0.079) ¹	1.5008(0.388) ¹	2.0214(0.500) ¹	3.9055(0.728) ¹		
M_m	1.4039(0.046) ²	1.5181(0.354) ²	2.0190(0.481) ²	3.7083(0.724) ²		
BM	1.4347(0.040) ³	1.5474(0.346) ³	2.0252(0.475) ³	4.9411(0.722) ³		
n=400						
QML	1.6033(0.004) ¹	1.9133(0.236) ¹	1.7101(0.384) ¹	3.1987(0.632) ¹		
M_m	1.5526(0.000) ²	1.8538(0.163) ²	1.6257(0.339) ²	2.1238(0.212) ²		
BM	1.6250(0.009) ³	1.8666(0.260) ³	1.6416(0.392) ³	2.3427(0.674) ³		
n=800						
QML	1.9835(0.104) ¹	2.7104(0.280) ¹	3.0961(0.447) ¹	1.6243(0.473) ¹		
M_m	1.9306(0.055) ²	2.7036(0.094) ²	3.0899(0.187) ²	1.67049(0.758) ²		
BM	1.9047(0.827) ³	2.7020(0.929) ³	3.1001(0.949) ³	1.625(0.947) ³ 1		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 7 Results for the student HGARCH(1,d,1) with 7 degree of freedom with 5% of additive outliers and $e = 3$

n	Estimation	γ	β	δ	ω	d
1000	QML	0.2232684	0.2433654	0.2671698	0.2210327	0.2967160
	M_m	0.1740031	0.2557992	0.2585510	0.2280809	0.3082918
	BM	0.1383870	0.2353445	0.2590069	0.2230483	0.2678836
1000	QML	0.1924815	0.2009800	0.1942890	0.1751197	0.2565371
	M_m	0.1403866	0.2169373	0.1874856	0.1822322	0.2664562
	BM	0.1031449	0.1962265	0.1877156	0.1771517	0.2281657
2000	QML	0.2119749	0.2238317	0.2018728	0.1905110	0.2622385
	M_m	0.1612009	0.2490745	0.2023851	0.1830490	0.2956904
	BM	0.1122019	0.2300487	0.1922464	0.1909857	0.2535307
2000	QML	0.1895571	0.1843541	0.1570523	0.1496991	0.2202929
	M_m	0.1370239	0.2132900	0.1515922	0.1438609	0.2518536
	BM	0.0885643	0.1935508	0.1514619	0.1503139	0.2133295

Table 7 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.2040775	0.1984260	0.1788925	0.1685548	0.2155104
	M_m	0.1472668	0.2232871	0.1630939	0.1544880	0.2670434
	BM	0.0979746	0.2008141	0.1596105	0.1756070	0.2147209
4000	QML	0.1908384	0.1646101	0.1424864	0.1348654	0.1740652
	M_m	0.1303340	0.1890230	0.1263907	0.1188400	0.2257310
	BM	0.0808824	0.1690851	0.1273080	0.1430466	0.1747903
RMSE for Out-of-sample Forecast						
				10-ahead	20-ahead	
	1-ahead	5-ahead				
n=200						
QML	10.524(0.121) ¹	2.0534 (0.456) ¹		2.0960 (0.612) ¹	3.5490 (0.831) ¹	
M_m	10.764(0.001) ²	2.0183 (0.100) ²		1.9765 (0.199) ²	2.4867 (0.246) ²	
BM	11.007(0.004) ³	2.0930 (0.063) ³		2.1170 (0.074) ³	3.7757 (0.017) ³	
n=400						
QML	7.6929(0.782) ¹	2.4893 (0.688) ¹		2.9023 (0.780) ¹	2.2268(0.825) ¹	
M_m	7.8378(0.767) ²	2.5061 (0.639) ²		2.8753 (0.706) ²	1.9196(0.942) ³	
BM	7.8162(0.799) ³	2.5041 (0.743) ³		2.8886 (0.814) ³	2.0495(0.876) ³	
n=800						
QML	11.306(0.825) ¹	1.6450 (0.852) ¹		1.7670 (0.887) ¹	2.1912(0.979) ¹	
M_m	11.5541(0.243) ²	1.5575 (0.473) ²		1.5962 (0.623) ²	1.9012(0.898) ²	
BM	11.410(0.344) ³	1.5844 (0.330) ³		1.7198 (0.403) ³	2.1591(0.882) ³	

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 8 Results for the student HGARCH(1,d,1) with 7 degree of freedom with 5% of additive outliers and $e = 5$

n	Estimation					
		γ	β	δ	ω	d
1000	QML	0.5648307	0.2499760	0.5356514	0.2686789	0.3028719
	M_m	0.1521896	0.2985108	0.2123590	0.2748437	0.3057135
	BM	0.07160839	0.21086305	0.18447847	0.17551215	0.24892119
1000	QML	0.5165514	0.2111579	0.4531793	0.2220389	0.2586513
	M_m	0.1208892	0.2630725	0.1542754	0.2491959	0.2612923
	BM	0.0517626	0.1744124	0.1451344	0.1378828	0.2093185
2000	QML	0.5638567	0.2559949	0.5365817	0.2427489	0.3114296
	M_m	0.1326667	0.2951217	0.1584158	0.2523152	0.3046324
	BM	0.05040214	0.19104861	0.15281403	0.15767507	0.21891329
2000	QML	0.5223559	0.2201130	0.4617962	0.1972170	0.2711683
	M_m	0.1069465	0.2616240	0.1195407	0.2315005	0.2635410
	BM	0.0375038	0.1561707	0.1213493	0.1253809	0.17851517

Table 8 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.5360030	0.2449803	0.5214461	0.2142798	0.3337353
	M_m	0.1117087	0.2877303	0.1437599	0.2397087	0.2706203
	BM	0.03662889	0.15565170	0.11949934	0.13285513	0.16512541
4000	QML	0.5022896	0.2099410	0.4543578	0.1726850	0.2965178
	M_m	0.0921130	0.2570777	0.1114871	0.2268937	0.2323010
	BM	0.0279544	0.1276121	0.0980505	0.1136700	0.1294532
RMSE for Out-of-sample Forecast						
	1-ahead	5-ahead		10-ahead		20-ahead
n=200						
QML	19.7853(0.115) ¹	1.7169 (0.390) ¹	2.9207 (0.504) ¹	84.904 (0.744) ¹		
M_m	20.320(0.204) ²	1.1987 (0.239) ²	1.5002 (0.279) ²	2.0512 (0.787) ²		
BM	20.159(0.115) ³	1.2299 (0.390) ³	1.4853 (0.504) ³	2.0723 (0.744) ³		
n=400						
QML	23.1851(0.084) ¹	2.8283 (0.378) ¹	3.4013 (0.498) ¹	20.5635(0.744) ¹		
M_m	23.8281(0.168) ²	2.6667 (0.224) ²	3.0344 (0.393) ²	2.4068(0.168) ²		
BM	23.663809(0.081) ³	2.68182 (0.3758) ³	3.0569 (0.496) ³	2.1355(0.739) ³		
n=800						
QML	149.31(0.604) ¹	1.6908 (0.687) ¹	4.2449 (0.772) ¹	8.6097(0.928) ¹		
M_m	153.59(0.378) ²	1.4169 (0.462) ²	1.7551 (0.568) ²	2.0146(0.778) ²		
BM	150.57(0.771) ³	2.4810 (0.856) ³	2.4128 (0.903) ³	3.2987(0.973) ³		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

In Table 9, we summarize the RMSE and MAE values for the Student's t_7 HGARCH(1,d,1) model with 10% of outliers contamination for $e = 3$. For the estimation part, we generally observe the decreasing of the RMSE and MAE values as the sample size increases for all the parameters of estimation methods excepted for the long memory parameter d . For the long memory parameter d and γ , the BM-estimate has the smallest MAE and RMSE values. Excepted these parameters, the QML-estimate seems little better. From the out-of-sample forecasting results, the p-values of the Diebold-Mariano test show that the forecast are generally the same accuracy taken in pairs. The BM and M-estimates are similar but the M-estimate is little better for long and short horizons excepted the 1-ahead. The QML-estimate is the worst.

In Table 10, we summarize the RMSE and the MAE values for the t_7 HGARCH(1,d,1) model with 10% of outliers contamination for $e = 5$. For estimation, the RMSE and MAE values for the parameters of BM-estimate decrease as the sample size increases. It's not the case for the parameter β for QML and the parameters β and ω for the M-estimate. We note that the BM-estimate has always the smallest RMSE and MAE values. It behaves more robust and is the better. From the out-of-sample forecasting results, the p-values of the Diebold-Mariano test show that the forecasts have generally the same accuracy. We note from the RMSE values that the QML has the worst performance for short and long horizon forecast. The BM and M-estimates are close but the BM-estimate will prefer for long and short horizons excepted the 1-ahead.

5 Empirical examples

In this section, we consider the example of the daily log-returns of the Standard and Poor's 500 Index (S&P 500). The observation period is from January 01, 1990 to December 31, 2007. These make a total of 4536 observed values for the S&P 500. We plot in Fig. 1, the daily log-returns $\{r_t\}$ and the boxplot of $\{r_t\}$ for the S&P 500 Index.

Table 11 presents the descriptive statistics for the S&P 500 daily log-returns. We note that the mean is small and close to zero. Also, it can be seen that from the kurtosis, skewness and Jarque-Bera test for normality, the S&P 500 daily log-returns are non-normal. The ARCH test at 5% significance level allows us to reject the null hypothesis of absence of heteroscedasticity. The ARCH test confirms the presence of heteroscedasticity in the data.

Figure 2 shows the autocorrelation functions (ACFs) for the log-returns and the squared log-returns. The ACF of squared log-returns has slow decaying over time. This may be due to long memory in the volatility. We applied the rescaled variance V/S test with bandwidth $q^* = 5$ defined in Giraitis et al. [10] for testing the presence of long memory to the log-returns and the squared log-returns. For the squared log-returns, the test statistic value is 2.8724 yet at the 0.05 significance level, the critical value is 0.1869. We have also done the test for different values of the bandwidth ($q^* = 10, q^* = 15$), and the presence of long memory is always suggested. For the log-returns, the value of the test statistic is 0.12273. These results can testify that the long memory is only present in variance. As stated in Kwan et al. [12], we can apply to $\{r_t\}$ a pure HGARCH model. The smallest Bayesian Information Criteria (BIC) for different HGARCH models is obtained with $p = 1$ and $q = 1$. For example when d is set to 0.4, the smallest $BIC = 45.77$ is obtained by log-likelihood value equals to -2.153746 and 5 parameters that correspond to the HGARCH with $p = 1$ and $q = 1$. We tried this with several value of d , we always come to the same conclusion. So the HGARCH(1,d,1) model

Table 9 Results for the student HGARCH(1,d,1) with 7 degree of freedom with 10% of additive outliers and $e = 3$

n	Estimation	γ	β	δ	ω	d
1000	QML	0.4169460	0.3011928	0.3134388	0.2586651	0.3063438
	M_m	0.3609773	0.3102409	0.3444614	0.3071906	0.3333049
	BM	0.3340910	0.3042965	0.3711966	0.2880218	0.3035411
1000	QML	0.3661369	0.2610087	0.2267887	0.2114776	0.2655122
	M_m	0.3019093	0.2706237	0.2495066	0.2660483	0.2899952
	BM	0.2771095	0.2636659	0.2746938	0.2454367	0.2610387
2000	QML	0.3930570	0.2951317	0.2241389	0.2276060	0.3226785
	M_m	0.3418088	0.3291794	0.2572199	0.2657475	0.3341221
	BM	0.3167903	0.3106951	0.2843631	0.2572686	0.3087179
2000	QML	0.3557297	0.2589721	0.1682103	0.1804710	0.2938298
	M_m	0.2801584	0.2975768	0.1897024	0.2260184	0.2934769
	BM	0.2672879	0.2751862	0.1989421	0.2107103	0.2696851

Table 9 continued

<i>n</i>	Estimation	γ	β	δ	ω	<i>d</i>
4000	QML	0.3685170	0.2855625	0.1699708	0.1920961	0.3269855
	M_m	0.3096062	0.3378079	0.1963209	0.2261551	0.3477793
	BM	0.2824697	0.3114089	0.2060521	0.2226867	0.3293762
4000	QML	0.3437049	0.2540167	0.1341495	0.1549976	0.3089090
	M_m	0.2507114	0.3095731	0.1445923	0.1905236	0.3039645
4000	QML	0.2435649	0.2811718	0.1434312	0.1776148	0.2950128
	BM					
RMSE for Out-of-sample Forecast						
	1-ahead			10-ahead		20-ahead
n=200						
QML	7.9437(0.007) ¹	2.3832 (0.129) ¹	2.2573 (0.268) ¹	8.0016 (0.776) ¹		
M_m	8.4704(0.516) ²	2.0927 (0.543) ²	1.4930 (0.271) ²	7.5822 (0.932) ²		
BM	8.3580(0.002) ³	2.1473(0.103) ³	1.6014 (0.249) ³	7.6327 (0.745) ³		
n=400						
QML	8.9515(0.503) ¹	2.4550 (0.406) ¹	3.2298 (0.503) ¹	2.1013(0.893) ¹		
M_m	9.2804(0.412) ²	2.4418 (0.244) ²	3.1948 (0.412) ²	1.8368 (0.795) ²		
BM	9.2363(0.554) ³	2.4460 (0.4868) ³	3.1899 (0.554) ³	1.9145(0.913) ³		
n=800						
QML	8.1438(0.644) ¹	2.0040 (0.493) ¹	2.0690 (0.644) ¹	2.2298(0.929) ¹		
M_m	8.3414(0.413) ²	1.9390 (0.204) ²	1.9678 (0.413) ²	2.0503(0.830) ²		
BM	8.2981(0.790) ³	1.9316 (0.708) ³	1.9924 (0.790) ³	2.0941(0.963) ³		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

Table 10 Results for the student HGARCH(1,d,1) with 7 degree of freedom with 10% of additive outliers and $e = 5$

n	Estimation	γ	β	δ	ω	d
1000	QML	0.6860862	0.3554409	0.3890850	0.2438352	0.3200711
	M_m	0.1999738	0.3024286	0.2161062	0.3468562	0.3238086
	BM	0.09285297	0.22338258	0.19531870	0.16254759	0.25989716
1000	QML	0.5052916	0.3204261	0.3180892	0.2026061	0.2640737
	M_m	0.1407247	0.2678568	0.1627060	0.3265580	0.2790750
	BM	0.0578108	0.1891003	0.1521024	0.1212633	0.2194328
2000	QML	0.6635013	0.3786370	0.3438742	0.2097120	0.2765339
	M_m	0.1281592	0.3177094	0.1558212	0.3321473	0.2976933
	BM	0.05196373	0.20082194	0.13942675	0.11535523	0.22756159
2000	QML	0.4481679	0.3505733	0.2886819	0.1708698	0.2237653
	M_m	0.0898310	0.2870026	0.1236213	0.3165158	0.2564725
	BM	0.0375878	0.1691607	0.1161633	0.0874029	0.1890520

Table 10 continued

<i>n</i>	Estimation					
	γ	β	δ	ω	d	
4000	0.6921490	0.3924226	0.3213152	0.1795246	0.2494398	
<i>RMSE</i>	0.09497276	0.32971093	0.13290190	0.33440822	0.26983722	
	0.03520161	0.18135530	0.11295034	0.08119154	0.19342388	
4000	0.445390	0.371323	0.2802407	0.1472661	0.1974306	
<i>MAE</i>	0.0668671	0.3007275	0.1077316	0.3241814	0.2302379	
	0.0286096	0.1533838	0.0923323	0.0638807	0.15688908	
RMSE for Out-of-sample Forecast						
	1-ahead		10-ahead		20-ahead	
n=200						
QML	18.541(0.092) ¹	2.9111(0.357) ¹	3.6467(0.507) ¹	8.1302(0.841) ¹		
M_m	19.076(0.067) ²	2.5303(0.008) ²	2.7537(0.170) ²	2.7898(0.499) ²		
BM	18.979(0.072) ³	2.5066(0.336) ³	2.7527(0.489) ³	2.7908(0.824) ³		
n=400						
QML	18.530(0.838) ¹	2.5925(0.718) ¹	2.4935(0.838) ¹	3.7704(0.965) ¹		
M_m	18.957(0.126) ²⁻¹	2.2340(0.029) ²	1.7104(0.126) ²	1.4589(0.369) ²		
BM	18.919(0.985) ³	2.2260(0.975) ³	1.6567(0.985) ³	1.4905(0.997) ³		
n=800						
QML	18.625(0.351) ¹	1.9258(0.152) ¹	3.1682(0.351) ¹	2.5572(0.727) ¹		
M_m	19.0333(0.232) ²	1.4154(0.183) ²	2.7993(0.232) ²	1.5179(0.477) ²		
BM	19.085(0.382) ³	1.4163(0.176) ³	2.7881(0.382) ³	1.5448(0.786) ³		

Notes: The number in bracket (...) ¹ denotes the p-values of Diebold-Mariano (DM) test for QML and M_m methods, (...) ² denotes the DM test p-values for M_m and BM methods, (...) ³ denotes the DM test p-values for BM and QML methods

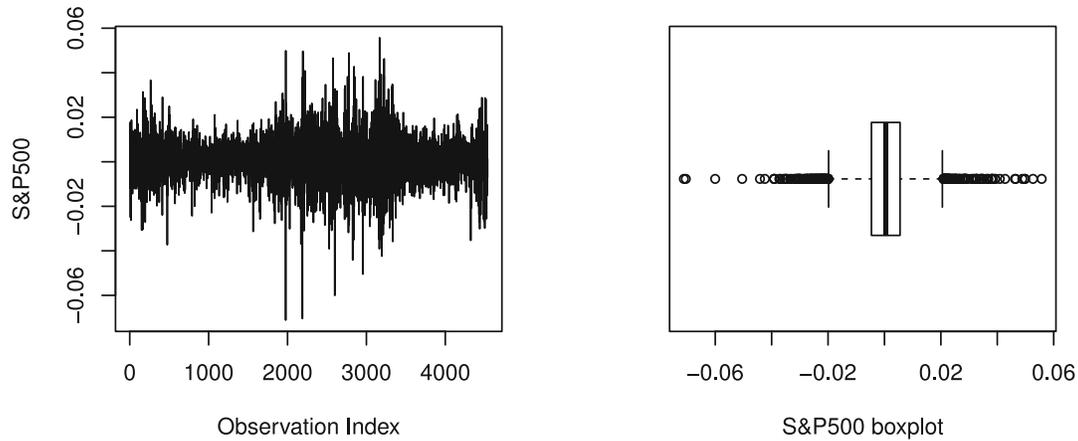


Fig. 1 Plot and boxplot of daily log-returns for S&P 500

Table 11 Descriptive statistics for the S&P 500 daily log-returns

Mean	SD	Skewness	Kurtosis	Min	Max	JBT	Arch-T
0.0003	0.0099	-0.1243	3.7597	-0.0711	0.0557	2687.6 (0.0000)	519.76 (0.0000)

Notes: JBT denotes the Jarque-Bera test, Arch-T denotes ARCH-test, and the p -values are in bracket

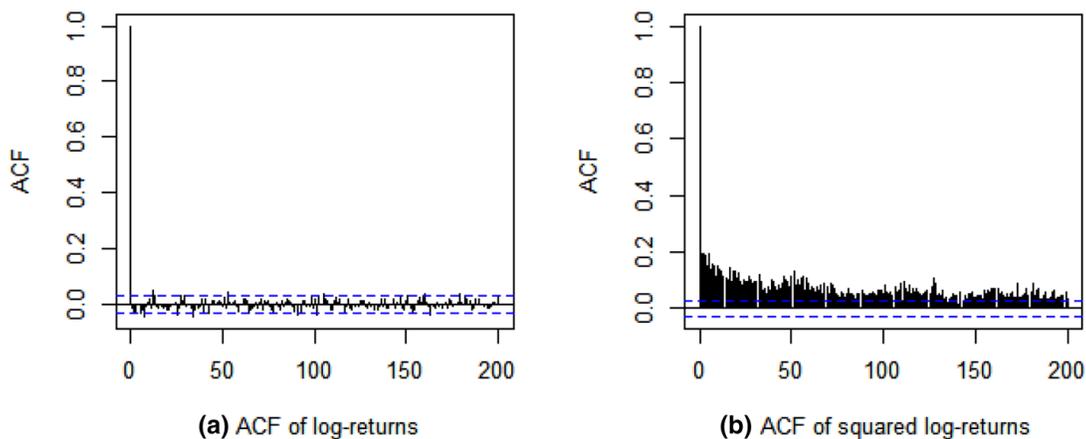


Fig. 2 ACFs of log-returns and squared log-returns for S&P 500

is the best fit for our data. Using the Quartile Method in Vinutha et al. [20], we detect the numbers of outliers in the series. There are 237 outliers in the series.

The first 4000 observations of $\{r_t\}$ are used to estimate the parameters of the model and the remaining 536 observations as out-of-sample forecasting. Using the quartile method, we find 200 outliers in the 4000 sample observations. We compare the performance of the BM-estimate, M-estimate(M_m) and the QML-estimate for the following three hyperbolic GARCH models, FIGARCH, HYGARCH and HGARCH.

Table 12 presents these estimated values for the three estimation methods for FIGARCH, HYGARCH and HGARCH models. We note a difference for the different estimate methods for each model which may be explained by the presence of outliers in data. Table 13 shows the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) results for the out-of-sample forecasting using the three competing estimation methods for FIGARCH,

Table 12 The estimated parameters of the FIGARCH(1,d,1),HYGARCH(1,d,1) and HGARCH(1,d,1) models for the S&P 500 daily log-returns

		Estimation				
		γ	β	δ	ω or ϕ	d
FIGARCH	QML	2.937e-06 (1.642e-09)	0.572055 (27.7340)	0.157408 (9.1042)	–	0.462010 (22.9565)
	M_m	9.108e-07 (1.6794e-09)	0.564157 (31.1347)	0.202621 (15.6541)	–	0.379969 (21.5788)
	BM	1.792e-06 (2.072e-09)	0.666034 (34.3124)	0.116032 (7.93528)	–	0.537644 (29.5184)
HYGARCH	QML	2.999e-06 (2.512e-09)	0.573681 (27.7576)	0.157355 (9.09769)	0.997883 (48.2533)	0.464701 (23.1209)
	M_m	2.107e-06 (3.575e-09)	0.592762 (24.6349)	0.188374 (9.55433)	0.953799 (43.7134)	0.438298 (19.2540)
	BM	1.854e-06 (2.157e-09)	0.671324 (33.7348)	0.115029 (7.82908)	0.996961 (30.5291)	0.545684 (29.4939)
HGARCH	QML	3.338e-06 (2.098e-09)	0.583783 (28.2041)	0.154509 (9.09820)	0.970660 (95.5484)	0.477609 (24.0711)
	M_m	2.019e-06 (3.344e-09)	0.605768 (27.5468)	0.184486 (10.9450)	0.910084 (91.2994)	0.436593 (20.3980)
	BM	1.887e-06 (2.059e-09)	0.676744 (34.0157)	0.112767 (7.75440)	0.986514 (85.1323)	0.551590 (30.1549)

Notes: The number in bracket denotes the t -statistic for the parameter estimation

Table 13 Out-of sample forecasting results of the FIGARCH(1,d,1),HYGARCH(1,d,1) and HGARCH(1,d,1) models for the S&P 500 daily log-returns

Methods	BM	M_m	QMLE
FIGARCH			
RMSE	4.076e+59	5.001e+27	1.997e+36
MAE	4.742e+58	8.255e+26	2.918e+35
HYGARCH			
RMSE	15.52145e-5	14.85304e-5	15.85435e-5
MAE	12.19993e-5	10.95313e-5	12.76183e-5
HGARCH			
RMSE	15.49579e-5	14.97837e-5	15.70795e-5
MAE	12.15469e-5	11.21154e-5	12.52305e-5

HYGARCH and HGARCH models. For the out-of-sample forecasting of HGARCH and HYGARCH models, both RMSE and MAE indices of M_m estimate have lowest values than those of BM and QML estimates. We can conclude that for the out-of-sample forecasting of HGARCH and HYGARCH models, the M-estimate performs the best and the QML the worst. This confirms our simulation study results. Moreover, we note that the HGARCH model has smallest RMSE and MAE values for BM and the QML estimates than the HYGARCH. The HGARCH model has so better performance than the HYGARCH model from forecasting

aspects concerning the BM and the QML estimation. Regarding the M_m estimate, the HYGARCH is better. Both models, HYGARCH and HGARCH have better performance than the FIGARCH model for all the methods. This may be explained by the existence of finite variance of the data which is a limit of FIGARCH models.

6 Conclusion

In this paper, we introduce two types of M-estimates for hyperbolic GARCH (HGARCH) models, the M-estimate and the BM-estimate both based on a bounded ρ function to deal with the presence of outliers in sample. The BM-estimate is a modified version of the M-estimate. According to the Monte Carlo study, we show that for the parameters estimation of the HGARCH(1,d,1) model, the M and BM-estimate are less influenced than QML-estimate when the data are contaminated by outliers. In these conditions, the BM-estimate is more robust than the other methods. Moreover, for the out-of-sample forecast, it has been shown that the M-estimate performs the best when outliers occurring. The results of an empirical example study revealed the same conclusion as those of the simulations studies.

Acknowledgements The first author obtained a grant from the Centre d'Excellence Africain en Mathématiques Informatique et TIC (CEA-MITIC) under the invitation of the third author for staying at University GASTON BERGER (UGB) of Saint-Louis, during four months. The authors are grateful to Professor Aliou DIOP and the members of LERSTAD (Laboratoire d'Etudes et Recherches en Statistique et Développement) of UGB for their contribution to this paper.

7 Appendix

7.1 Appendix 1

Let A be a $\mathcal{M}_{n \times n}(\mathbb{R})$ matrix. We assume that

$$\|A\| = (\text{tr}(AA'))^{1/2} = \left(\sum_{1 \leq i, j \leq n} |a_{i,j}|^2 \right)^{1/2}.$$

For a vector $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, we denote by $\|u\| = \sum_{i=1}^n |u_i|$. We denote by $\nabla f = (\frac{\partial f}{\partial u_1}, \dots, \frac{\partial f}{\partial u_n})$ given a differentiable function $f(u_1, \dots, u_n)$ and $\nabla^2 f$ the $n \times n$ matrix whose (i, j) element is $\frac{\partial^2 f}{\partial u_i \partial u_j}$.

Proof of Theorem 1 1. Let θ_0 be the true parameter. We can write

$$M(\theta) = \mathbf{E}(\rho(x_t - \log h_t(\theta))) = \mathbf{E}(\rho(w_t - \log(\frac{h_t(\theta)}{h_t(\theta_0)}))).$$

From (4)–(5), we have that $w_t = x_t - \log h_t(\theta_0) = \log(\varepsilon_t^2)$ are i.i.d random variable with density g .

From assumption (A4), $M(\theta) = \mathbf{E}(\rho(w_t - \log(\frac{h_t(\theta)}{h_t(\theta_0)})))$ takes its minimum value $M(\ddot{\theta})$ when $\log(\frac{h_t(\ddot{\theta})}{h_t(\theta_0)}) = u_0$ that is

$$h_t(\ddot{\theta}) = e^{u_0} h_t(\theta_0) \quad a.s. \quad (20)$$

Let $\bar{\theta} = (e^{u_0} \gamma_0, \beta_{0_1}, \dots, \beta_{0_p}, \delta_{0_1}, \dots, \delta_{0_q}, e^{u_0} \omega_0, d_0)$ such that

$$h_t(\bar{\theta}) = e^{u_0} h_t(\theta_0), \quad \text{then we obtain } h_t(\bar{\theta}) = h_t(\ddot{\theta}). \quad (21)$$

By (20) and (21), there exist U and V , two neighborhoods of θ_0 such that: $\ddot{\theta} \in U$ and $\bar{\theta} \in V$. From (21), we also have $\log(h_t(\bar{\theta})) - \log(h_t(\ddot{\theta})) = 0$. By the mean value theorem, there exists $\theta' \in U \cap V$ such that

$$|\log(h_t(\bar{\theta})) - \log(h_t(\ddot{\theta}))| = |\bar{\theta} - \ddot{\theta}| \frac{\|\nabla h_t(\theta')\|}{h_t(\theta')} = \|(\bar{\theta} - \ddot{\theta})' \nabla h_t(\theta')\| = 0$$

This implies that

$$(\bar{\theta} - \ddot{\theta})' \nabla h_t(\theta') = 0$$

Similarly to Straumann [19] p.119 and assumption (A7), we have $\bar{\theta} = \ddot{\theta}$.

2. For the proof of the second point, we can see also Theorem 1 of Muler and Yohai [16]. \square

Lemma 1 *Let y_t be a stationary and ergodic HGARCH(p, d, q) process satisfying (1). Let $h_t(\theta)$ be as defined in (3) and $\tilde{h}_t(\theta)$ as defined in (9). Under the assumptions (A0), (A2) and (A6), then, there exists also a positive real u such that,*

1.

$$\sup_{\theta \in \Theta} |h_t(\theta) - \tilde{h}_t(\theta)| \leq u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \quad \text{with } d_m = \min_{\theta \in \Theta} d,$$

2. there exists a neighborhood U of θ_0 and $\eta < 1 - \frac{1}{1 + \frac{d_m}{2}}$ such that

$$\mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta)\| < \infty,$$

$$\mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta) \nabla' \log h_t(\theta)\| \leq \mathbf{E} \sup_{\theta \in U} \|\nabla \log h_t(\theta)\|^2 < \infty,$$

3. there exists a neighborhood U of θ_0 , $u, K > 0$ and $\eta < 1 - \frac{1}{1 + \frac{d_m}{2}}$ such that

$$\begin{aligned} \sup_{\theta \in U} \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| &\leq u \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2 \\ &+ \frac{Ku}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta)\|, \end{aligned}$$

$$\text{with } \gamma_m = \min_{\theta \in \Theta} \gamma,$$

4. there exist a neighborhood U of θ_0 , $u, K > 0$ and $\eta < 1 - \frac{1}{1 + \frac{d_m}{2}}$ such that

$$\begin{aligned} \mathbb{E} \sup_{\theta \in U} \|\nabla^2 h_t(\theta)\| &< \infty \quad \text{and} \\ \sup_{\theta \in U} \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| &\leq K \left(\sum_{k=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-k}^4 \right)^{1/2} \\ &+ u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} \right\|. \end{aligned}$$

Proof 1. As in the same lines of Robinson and Zaffaroni [17], from (3) and (9) we have,

$$0 \leq h_t(\theta) - \tilde{h}_t(\theta) = \sum_{j=1}^{\infty} b_j(\theta)y_{t-j}^2 - \sum_{j=1}^{t-1} b_j(\theta)y_{t-j}^2.$$

By (3), we have $b_j = O(j^{-1-d})$. Then, there exists $u > 0$ such that $b_j \leq uj^{-1-d}$, we obtain

$$\text{for all } \theta \in \Theta, \quad |h_t(\theta) - \tilde{h}_t(\theta)| \leq u \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2. \tag{22}$$

2. We rewrite $h_t(\theta)$ in (3) as

$$h_t(\theta) = \frac{\gamma}{1 - \beta(1)} + \sum_{j=1}^{\infty} b_j(\theta^*)y_t^2 \tag{23}$$

where $\theta^* = (\delta', \beta', \omega, d)'$ contains the parameters of θ without γ . We get

$$\|\nabla \log h_t(\theta)\| = \left\| \frac{1}{h_t(\theta)} \nabla h_t(\theta) \right\| \leq \frac{1}{\gamma_m} \|\nabla h_t(\theta)\| \quad \text{with } \gamma_m = \min_{\theta \in \Theta} \gamma, \tag{24}$$

$$\nabla h_t(\theta) = \left(\frac{1}{1 - \beta(1)}, \frac{\partial h_t(\theta)}{\partial \theta^*} \right). \tag{25}$$

As in Kwan et al. [12], the partial derivatives of $h_t(\theta)$ defined in (23) from the component of θ^* can be written like almost surely convergent series. For example

$$\frac{\partial h_t(\theta)}{\partial \omega} = \left(1 - \frac{\delta(L)}{1 - \beta(L)}\right) (1 - L)^d y_t^2 = y_t^2 - \sum_{u=0}^{\infty} \sum_{j=0}^{\infty} \mu_u \pi_j y_{t-u-j}^2. \tag{26}$$

with

$$\mu(L) = \frac{\delta(L)}{1 - \beta(L)} = \sum_{u=0}^{\infty} \mu_u L^u \quad \text{and} \quad \sum_{u=0}^{\infty} \|\mu_u\| < \infty. \tag{27}$$

Thus, we assume that

$$\text{for all } k = 1, \dots, p + q + 2, \quad \frac{\partial h_t(\theta)}{\partial \theta_k^*} = \sum_{j=1}^{\infty} b_{j,k}^{(1)}(\theta) y_{t-j}^2, \tag{28}$$

where $b_{j,k}^{(1)}(\theta)$ is the coefficient of the first order derivatives of $h_t(\theta)$ by the component of θ^* . By assumptions **(A₀)**, **(A₆)**, the relations (3), (28), and the Beppo Levi's Lemma, for $k = 1, \dots, (p + q + 2)$ and $\theta \in U$, there exists $K > 0$ such that

$$\sum_{j=1}^{\infty} \|b_{j,k}^{(1)}(\theta)\| \mathbb{E}(y_{t-j}^2) \leq K \sum_{j=1}^{\infty} j^{-(1+d_m)(1-\eta)} \mathbb{E}(\|y_{t-j}^2\|) \quad \text{and so,} \tag{29}$$

$$\mathbb{E} \left(\left\| \frac{\partial h_t(\theta)}{\partial \theta_k^*} \right\| \right) < \infty \quad \text{for } \eta < 1 - \frac{1}{1 + \frac{d_m}{2}}. \tag{30}$$

Using (29)–(30), we show that for all $\theta \in U$ the expectation of each partial derivatives is dominated by convergent serie.

It follows that

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla h_t(\theta)\|) < \infty. \quad (31)$$

Using (24), we obtain the first result.

From the matrix norm, we have

$$\|\nabla \log h_t(\theta)(\nabla \log h_t(\theta))'\| \leq \|\nabla \log h_t(\theta)\| \|(\nabla \log h_t(\theta))'\| \leq \|\nabla \log h_t(\theta)\|^2.$$

We want to show the existence of $\mathbf{E}(\sup_{\theta \in U} \|\nabla h_t(\theta)\|^2)$. Using (25), we obtain

$$\begin{aligned} \|\nabla h_t(\theta)\|^2 &\leq \left(\frac{1}{1-\beta(1)}\right)^2 + \frac{2}{1-\beta(1)} \sum_{k=1}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right| + \sum_{k=1}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right|^2 \\ &\quad + 2 \sum_{k_1=1}^{p+q+2} \sum_{k_2=1, k_1 \neq k_2}^{p+q+2} \left|\frac{\partial h_t(\theta)}{\partial \theta_{k_1}^*}\right| \left|\frac{\partial h_t(\theta)}{\partial \theta_{k_2}^*}\right|. \end{aligned} \quad (32)$$

From (28), we obtain

$$\left|\frac{\partial h_t(\theta)}{\partial \theta_k^*}\right|^2 \leq \left(\sum_{j=1}^{\infty} |b_{j,k}^{(1)}(\theta)| y_{t-j}^2\right)^2 \leq \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| |y_{t-j_1}^2 y_{t-j_2}^2|.$$

Using the same arguments as in (28)-(31), combining the Hölder's inequality and condition $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$ of assumption **(A₂)** that ensures $\mathbf{E}y_t^4 < \infty$ due to the Theorem 2 of Li et al. [13] for $m = 2$, we get

$$\sum_{j_1, j_2=1}^{\infty} |b_{j_1,k}^{(1)}(\theta)| |b_{j_2,k}^{(1)}(\theta)| \mathbb{E}(|y_{t-j_1}^2 y_{t-j_2}^2|) < \infty \quad \text{for } \eta < 1 - \frac{1}{1 + \frac{d_m}{2}}.$$

And the result follows.

3. Similarly to Berkes et al. [2], we obtain

$$\begin{aligned} \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| &= \left\| \frac{\nabla h_t(\theta)}{h_t(\theta)} - \frac{\nabla \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\|, \\ \left\| \frac{\nabla h_t(\theta)}{h_t(\theta)} - \frac{\nabla \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| &\leq \frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| \\ &\quad + \frac{1}{\tilde{h}_t(\theta)} \|\nabla \log h_t(\theta)\| |h_t(\theta) - \tilde{h}_t(\theta)|. \end{aligned}$$

From the equations (3), (9), (25) and (28), we have

$$\nabla \tilde{h}_t(\theta) = \left(\frac{1}{1-\beta(1)}, \sum_{j=1}^{t-1} b_{j,k}^{(1)}(\theta) y_{t-j}^2 \right). \quad (33)$$

For $k = 1, \dots, p + q + 2$, we have

$$\|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| = \left\| \sum_{j=t}^{\infty} b_{j,k}^{(1)}(\theta) y_{t-j}^2 \right\| \leq \sum_{k=1}^{p+q+2} \sum_{j=t}^{\infty} |b_{j,k}^{(1)}(\theta) y_{t-j}^2|.$$

Using Lemma 1(1–2) and the same arguments as in (28)–(30), one has

$$\frac{1}{\tilde{h}_t(\theta)} \|\nabla h_t(\theta) - \nabla \tilde{h}_t(\theta)\| \leq \frac{Ku(p+q+2)}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)(1-\eta)} y_{t-j}^2$$

and

$$\frac{1}{\tilde{h}_t(\theta)} \|\nabla \log h_t(\theta)\| |h_t(\theta) - \tilde{h}_t(\theta)| \leq \frac{Ku}{\gamma_m} \sum_{j=t}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta)\|.$$

Hence, from Lemma 1(2) combining the Hölder’s inequality and condition $\omega < (\mathbf{E}|\varepsilon_t|^4)^{-1/2}$ of assumption (A₂) that ensures $\mathbf{E}y_t^4 < \infty$ due to the Theorem 2 of Li et al. [13] for $m = 2$, we obtain that

- $\sum_{j=1}^{\infty} j^{-(1+d_m)} y_{t-j}^2 \|\nabla \log h_t(\theta)\|$ converges almost surely and the result follows.
4. Each second partial derivatives from θ can be rewritten as almost surely convergent series. For $i, j = 1, \dots, (p+q+3)$, let $b_k^{(2)}(i, j)$ be the coefficients of these series. We have

$$\nabla^2 h_t(\theta) = \left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} \right) \quad \text{where} \quad \left(\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} \right)_{i,j} = \sum_{k=1}^{\infty} b_k^{(2)}(i, j) y_{t-k}^2 \quad \text{or} \quad 0.$$

The matrix norm of $\nabla^2 h_t(\theta)$ is such that

$$\|\nabla^2 h_t(\theta)\| \leq \sum_{i,j=1}^{p+q+3} \left(\sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} |b_{k_1}^{(2)}(i, j)| |b_{k_2}^{(2)}(i, j)| |y_{t-k_1}^2 y_{t-k_2}^2| \right)^{1/2}.$$

For the same arguments as in Lemma 1.2 and (24) we get the result. Moreover

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla^2 h_t(\theta)\|) < \infty. \tag{34}$$

For the second point, with the matrix norm, we obtain

$$\left\| \frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{h}_t(\theta)}{\partial \theta \partial \theta'} \right\| \leq \left[\sum_{i,j=1}^{p+q+3} \left(\sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^2 \right)^2 \right]^{1/2}. \tag{35}$$

From the same technical arguments as in Robinson and Zaffaroni [17], Hölder’s inequality and assumption (A₆) that ensures

$\sum_{k=1}^{\infty} |b_{k_1}^{(2)}(i, j)| < \infty$, a.s, we obtain

$$\left(\sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^2 \right)^2 \leq \sum_{k_1=t}^{\infty} |b_{k_1}^{(2)}(i, j)| \sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^4. \tag{36}$$

From (35)–(36), we get

$$\|\nabla^2 h_t(\theta) - \nabla^2 \tilde{h}_t(\theta)\| \leq K_2 \sum_{i,j=1}^{p+q+3} \left(\sum_{k=t}^{\infty} |b_k^{(2)}(i, j)| y_{t-k}^4 \right)^{1/2}. \tag{37}$$

We can also have the following expression

$$\left\| \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right\| \leq \frac{1}{\tilde{h}_t(\theta)} \left(\|\nabla^2 h_t(\theta) - \nabla^2 \tilde{h}_t(\theta)\| \right)$$

$$+\left\|\frac{\nabla^2 h_t(\theta)}{h_t(\theta)}\right\|\|h_t(\theta)-\tilde{h}_t(\theta)\| \Big).$$

Using the same arguments as in Lemma 1.(1;3) and (34), the result follows. \square

Lemma 2 *If the assumptions (A₀)-(A₃) hold, then $\tilde{M}_n(\theta)$ defined in (12) is a good estimate of $M_n(\theta)$ defined in (8),*

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \xrightarrow{n \rightarrow \infty} 0 \quad a.s.$$

Proof From (8) and (12), we have

$$|\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{1}{n-p} \sum_{t=p+1}^n |\rho(x_t - \log \tilde{h}_t(\theta)) - \rho(x_t - \log h_t(\theta))|. \quad (38)$$

Using Lemma 1 and assumption (A₃), ρ is a lipschit function. Then, there exists $K > 0$ such that

$$\begin{aligned} |\rho(x_t - \log \tilde{h}_t(\theta)) - \rho(x_t - \log h_t(\theta))| &\leq K |\log h_t(\theta) - \log \tilde{h}_t(\theta)|, \\ \left| \log \frac{x}{y} \right| &\leq \frac{|x-y|}{\min(x,y)}, \quad x, y > 0, \end{aligned} \quad (39)$$

we obtain

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \leq \frac{Ku}{(n-p)\gamma_m} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2. \quad (40)$$

Using assumption (A₀) and the Beppo Levi's theorem, there exists $K > 0$ such that

$$\begin{aligned} \sum_{t=1}^{\infty} \frac{1}{t} \mathbb{E} \left(\sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 \right) &= \sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} j^{-1-d_m} \mathbb{E}(y_{t-j}^2) \\ &\leq K \sum_{t=1}^{\infty} \frac{1}{t^{1+\frac{d_m}{2}}} \sum_{j=1}^{\infty} \frac{1}{j^{1+\frac{d_m}{2}}} < \infty. \end{aligned}$$

This implies that

$$\sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} j^{-1-d_m} y_{t-j}^2 < \infty \quad a.s.$$

Using Kronecker's Lemma, see also Robinson and Zaffaroni [17],

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M_n(\theta)| \xrightarrow{n \rightarrow \infty} 0, \quad a.s.$$

\square

Lemma 3 *If the assumptions (A₀)-(A₃) hold, then*

$$\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow{n \rightarrow \infty} 0, \quad a.s.,$$

with $M(\theta) = \mathbf{E}(\rho(x_t - \log h_t(\theta)))$.

Proof By (4), we obtain

$$|x_t - \log h_t(\theta)| \leq |\log \epsilon_t^2| + |\log h_t(\theta_0) - \log h_t(\theta)|.$$

Moreover, using (39)-(40), assumption (A₂) and (3), we prove that

$$\begin{aligned} \mathbf{E}(\sup_{\theta \in \Theta} |h_t(\theta)|) &< \infty \quad \text{and} \\ \mathbf{E}(\sup_{\theta \in \Theta} |\log h_t(\theta) - \log h_t(\theta_0)|) &< \infty. \end{aligned}$$

Using the same arguments as in Muler and Yohai [16], we obtain

$$\mathbf{E}(\sup_{\theta \in \Theta} |\rho(x_t - \log h_t(\theta))|) < \infty. \tag{41}$$

By (41) and Theorem 3.5.8 of Stout [18] p.182-183, $\rho(x_t - \log h_t(\theta))$ is an ergodic and stationary process. Using the ergodic theorem in Straumann [19] p.17-18, we obtain

$$M_n(\theta) \xrightarrow[n \rightarrow \infty]{} \mathbf{E}(\rho(x_t - \log h_t(\theta))) \quad a.s. \tag{42}$$

We can show that (42) converges uniformly. From (8) and for all $\theta_1, \theta_2 \in \Theta$, we have

$$\begin{aligned} |M_n(\theta_1) - M_n(\theta_2)| &\leq \frac{1}{n-p} \sum_{t=p+1}^n |\log h_t(\theta_2) - \log h_t(\theta_1)|, \\ \sup_{\theta_1, \theta_2 \in \Theta} |M_n(\theta_1) - M_n(\theta_2)| \frac{1}{\|\theta_1 - \theta_2\|} &\leq \frac{1}{n-p} \sum_{t=p+1}^n v_t, \end{aligned}$$

where

$$v_t = \sup_{\theta_1, \theta_2 \in \Theta} \frac{1}{\|\theta_1 - \theta_2\|} |\log h_t(\theta_2) - \log h_t(\theta_1)|.$$

Using the same technics as in Berkes et al. [2], by the mean value theorem, there exists $\psi \in \Theta$ satisfying $\|\psi - \theta_1\| \leq \|\theta_1 - \theta_2\|$ and $\|\psi - \theta_2\| \leq \|\theta_1 - \theta_2\|$ such that

$$|\log h_t(\theta_2) - \log h_t(\theta_1)| \leq \|\theta_1 - \theta_2\| \sup_{\psi \in \Theta} \|\nabla \log h_t(\psi)\|, \quad \theta_1, \theta_2 \in \Theta. \tag{43}$$

By Lemma 1(2) and (43), we have $\mathbf{E}(|v_t|) < \infty$. Moreover, according to Theorem 3.5.8 of Stout [18] p.182-183, v_t is a stationary and ergodic sequence. Using the ergodic theorem, we obtain

$$\frac{1}{n-p} \sum_{t=p+1}^n v_t = O(1), \quad a.s., \quad n \rightarrow \infty.$$

Hence $M_n(\theta)$ is a.s equicontinuous. Also from (42) with Θ a compact subset, we deduce

$$\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0, \quad a.s.$$

□

Proof of Theorem 2 From Lemma 2 and Lemma 3, we show that

$$\sup_{\theta \in \Theta} |\tilde{M}_n(\theta) - M(\theta)| \xrightarrow[n \rightarrow \infty]{} 0, \quad a.s.$$

Using the same arguments as in Muler and Yohai [16], we get the result. □

Lemma 4 Suppose that assumptions (\mathbf{A}_0) - (\mathbf{A}_7) hold, then

$$\frac{1}{\sqrt{n-p}} \sum_{t=p+1}^n \nabla \rho(x_t - \log h_t(\theta_0)) \xrightarrow{D} \mathbf{N}(0, \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0)))D_0),$$

with $D_0 = \mathbf{E}[\nabla \log h_t(\theta_0)(\nabla \log h_t(\theta_0))']$.

Proof We have

$$\begin{aligned} \nabla M_n(\theta_0) &= \frac{1}{n-p} \sum_{t=p+1}^n \nabla \rho(x_t - \log h_t(\theta_0)), \\ \nabla \rho(x_t - \log h_t(\theta_0)) &= \nabla \log h_t(\theta_0) \psi(x_t - \log h_t(\theta_0)). \end{aligned} \quad (44)$$

We use the same arguments as in Muler and Yohai [16] to prove this Lemma. From assumption (\mathbf{A}_7) and using the same arguments as in Straumann [19] p.118-119, we show that D_0 is positive definite. Also, from assumption (\mathbf{A}_5) , we have $\mathbf{E}(\psi(w_t)) = 0$, with $w_t = x_t - \log h_t(\theta_0)$. Hence, we obtain

$$\begin{aligned} \mathbf{E}(b' \nabla \rho(x_t - \log h_t(\theta_0)) | \mathcal{F}_{t-1}) &= \mathbf{E}(\psi(w_t) b' \nabla \log h_t(\theta_0) | \mathcal{F}_{t-1}), \\ &= b' \nabla \log h_t(\theta_0) \mathbf{E}(\psi(w_t) | \mathcal{F}_{t-1}) = 0. \end{aligned}$$

This implies that for any vector $b \neq \underline{0}$ in \mathbf{R}^{p+q+3} , $b' \nabla \rho(x_t - \log h_t(\theta_0))$ is a centered stationary martingale differences sequence. By the martingale Central Limit Theorem and Standard Cramer-Wold device similarly to Muler and Yohai [16], we have the finally result. \square

Lemma 5 Suppose that assumptions (\mathbf{A}_0) - (\mathbf{A}_7) hold. Then, there exists a neighborhood U of θ_0 such that

1.

$$\sup_{\theta \in U} \|\nabla^2 M_n(\theta) - \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta)))\| \xrightarrow{n \rightarrow \infty} 0, \quad a.s.$$

2. $A(\theta_0)$ is a positive definite matrix given by

$$A(\theta_0) = \mathbf{E}_{\theta_0}(\nabla^2 \rho(x_t - \log h_t(\theta_0))) = \mathbf{E}(\psi'(w_t) D_0).$$

Proof 1. Using (44), we obtain

$$\begin{aligned} \nabla^2 \rho(x_t - \log h_t(\theta)) &= [\psi(x_t - \log h_t(\theta)) + \psi'(x_t - \log h_t(\theta))] \nabla \log h_t(\theta) \\ &\quad \times (\nabla \log h_t(\theta))' - \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 h_t(\theta)}{h_t(\theta)}. \end{aligned} \quad (45)$$

Under (\mathbf{A}_3) , ψ' and ψ are bounded and continuous functions. Moreover, by Lemma 1(2-4), we have

$$\mathbf{E}(\sup_{\theta \in U} \|\nabla^2 \rho(x_t - \log h_t(\theta))\|) < \infty \quad \text{and} \quad \mathbf{E}(\sup_{\theta \in U} \|\nabla^2 M_n(\theta)\|) < \infty. \quad (46)$$

Since $\nabla^2 \rho(x_t - \log h_t(\theta))$ is an ergodic and stationary process, by the ergodic theorem and (46), we have, for all $\theta \in U$,

$$\nabla^2 M_n(\theta) \xrightarrow{n \rightarrow \infty} \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta))) \quad a.s.$$

We follow the same arguments as in the proof of Lemma 3 or Lemma 5.6 of Berkes et al. [2] to obtain the result.

2. Using the same arguments as in Muler and Yohai [16], we have

$$\begin{aligned} \mathbf{E}_{\theta_0}(\nabla^2 \rho(x_t - \log h_t(\theta))) &= \mathbf{E}(\psi'(x_t - \log h_t(\theta_0)))\mathbf{E}[\nabla \log h_t(\theta_0)(\nabla \log h_t(\theta_0))'] \\ &= \mathbf{E}(\psi'(w_t))D_0 = A(\theta_0). \end{aligned}$$

The result is obvious for $\mathbf{E}(\psi'(w_t)) > 0$, this is verified by assumption **(A₅)**. □

Lemma 6 *Suppose that all the assumptions **(A₀)**-**(A₇)** hold. Then*

$$\frac{1}{n-p} \sum_{t=p+1}^n \|\nabla \rho(x_t - \log h_t(\theta_0)) - \nabla \rho(x_t - \log \tilde{h}_t(\theta_0))\| \xrightarrow{n \rightarrow \infty} 0 \text{ a.s.}$$

Proof From (44), we have

$$\frac{1}{n-p} \sum_{t=p+1}^n \|\nabla \rho(x_t - \log h_t(\theta_0)) - \nabla \rho(x_t - \log \tilde{h}_t(\theta_0))\| \leq u_n + v_n,$$

where

$$\begin{aligned} u_n &= \frac{1}{n-p} \sum_{t=p+1}^n \|\psi(x_t - \log h_t(\theta_0))(\nabla \log h_t(\theta_0) - \nabla \log \tilde{h}_t(\theta_0))\|, \\ v_n &= \frac{1}{n-p} \sum_{t=p+1}^n \|\nabla \log \tilde{h}_t(\theta_0)(\psi(x_t - \log h_t(\theta_0)) - \psi(x_t - \log \tilde{h}_t(\theta_0)))\|. \end{aligned}$$

By Lemma 1(3), we obtain

$$\begin{aligned} u_n &\leq \frac{Kv}{\gamma_0(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_0)(1-\eta)} |\psi(x_t - \log h_t(\theta_0))| y_{t-j}^2 \\ &\quad + \frac{Ku}{\gamma_0} \frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_0)} y_{t-j}^2 \|\nabla \log h_t(\theta_0)\| |\psi(x_t - \log h_t(\theta_0))|. \end{aligned}$$

Using the Kronecker’s lemma combined to assumption **(A₃)**, that ensures ψ is bounded. We obtain

$$\frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d_0)(1-\eta)} |\psi(x_t - \log h_t(\theta_0))| y_{t-j}^2 \xrightarrow{n \rightarrow \infty} 0 \text{ a.s.} \tag{47}$$

Indeed, using the same argument as in (29)–(30) we have

$$\begin{aligned} \sum_{t=1}^{\infty} \frac{1}{t} \mathbf{E} \sum_{j=t}^{\infty} j^{-(1+d_0)(1-\eta)} y_{t-j}^2 &= \sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_0)(1-\eta)} \mathbf{E} y_{t-j}^2, \\ &\leq \sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} \frac{1}{t^{\frac{d_0}{2}(1-\eta)}} \frac{1}{j^{(1+\frac{d_0}{2})(1-\eta)}}, \\ &\leq \sum_{t=1}^{\infty} \frac{1}{t^{1+\frac{d_0}{2}(1-\eta)}} \sum_{j=1}^{\infty} \frac{1}{j^{(1+\frac{d_0}{2})(1-\eta)}}. \end{aligned} \tag{48}$$

The second term of (48) converges for $\eta < 1 - \frac{1}{1 + \frac{d_0}{2}}$ and by Beppo Levy's lemma, we obtain

$$\sum_{t=1}^{\infty} \frac{1}{t} \sum_{j=t}^{\infty} j^{-(1+d_0)(1-\eta)} y_{t-j}^2 < \infty \quad \text{a.s.}$$

Using the same technics and Lemma 1(2-3), one has

$$\frac{1}{(n-p)} \sum_{t=p+1}^n \sum_{j=t}^{\infty} \frac{y_{t-j}^2}{j^{1+d_0}} \|\nabla \log h_t(\theta_0)\| |\psi(x_t - \log h_t(\theta_0))| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{a.s.} \quad (49)$$

By (47) and (49), we obtain

$$u_n \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{a.s.} \quad (50)$$

From assumption (A₃) and the relations (38)–(40), we obtain

$$\begin{aligned} v_n &\leq \frac{1}{n-p} \sum_{t=1+p}^n |(\psi(x_t - \log h_t(\theta_0)) - \psi(x_t - \log \tilde{h}_t(\theta_0)))| \|\nabla \log \tilde{h}_t(\theta_0)\|, \\ &\leq \frac{Ku}{\gamma_0} \frac{1}{n-p} \sum_{t=1+p}^n \sum_{j=t}^{\infty} j^{-(1+d_0)} y_{t-j}^2 \|\nabla \log \tilde{h}_t(\theta_0)\|. \end{aligned} \quad (51)$$

From the same lines as u_n combined to (24), (33) and lemma 1(2), we obtain

$$v_n \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{a.s.} \quad (52)$$

From (50) and (52), we have the result. \square

Lemma 7 *Suppose that all the assumptions (A₀)–(A₇) hold. Then, there exists a neighborhood U of θ_0 such that*

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|\nabla^2 \rho(x_t - \log h_t(\theta)) - \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta))\| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{a.s.}$$

Proof We note that

$$v(x_t - \log h_t(\theta)) = \psi(x_t - \log h_t(\theta)) + \psi'(x_t - \log h_t(\theta)), \quad (53)$$

$$P_t(\theta) = \nabla \log h_t(\theta) \nabla' \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta),$$

$$a_t = \nabla^2 \rho(x_t - \log h_t(\theta)) - \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta)). \quad (54)$$

Using (45), (53) and (54), we obtain the following expression

$$a_t = L_t(\theta) - M_t(\theta)$$

where

$$\begin{aligned} L_t(\theta) &= v(x_t - \log h_t(\theta)) \nabla \log h_t(\theta) \nabla' \log h_t(\theta) \\ &\quad - v(x_t - \log \tilde{h}_t(\theta)) \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta), \end{aligned}$$

and

$$M_t(\theta) = \psi(x_t - \log h_t(\theta)) \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \psi(x_t - \log \tilde{h}_t(\theta)) \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)}.$$

We can rewrite

$$L_t(\theta) = v(x_t - \log h_t(\theta))P_t(\theta) + [v(x_t - \log h_t(\theta)) - v(x_t - \log \tilde{h}_t(\theta))] \nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta), \tag{55}$$

we also have

$$\|P_t(\theta)\| \leq \|\nabla \log h_t(\theta) - \nabla \log \tilde{h}_t(\theta)\| \left(\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\| \right).$$

Let

$$s_n = \frac{1}{n-p} \sum_{t=p+1}^n |v(x_t - \log h_t(\theta))| \|P_t(\theta)\|$$

By Lemma 1(3) and (56), we have

$$\begin{aligned} s_n &\leq \frac{K_3}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d)} y_{t-j}^2 \|\nabla \log h_t(\theta)\| |v(x_t - \log h_t(\theta))| \\ &\quad \times (\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\|) + \frac{K_3}{n-p} \sum_{t=p+1}^n \sum_{j=t}^{\infty} j^{-(1+d)(1-\eta)} y_{t-j}^2 \\ &\quad \times |v(x_t - \log h_t(\theta))| (\|\nabla \log \tilde{h}_t(\theta)\| + \|\nabla' \log h_t(\theta)\|). \end{aligned}$$

From assumption (A₃), $v(x_t - \log h_t(\theta))$ is a bounded function. Using the same arguments as in Lemma 6, and Lemma 1(2), we obtain

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n |v(x_t - \log h_t(\theta))| \|P_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad a.s. \tag{56}$$

From (53), let

$$\begin{aligned} b_t &= |v(x_t - \log h_t(\theta)) - v(x_t - \log \tilde{h}_t(\theta))|, \\ &\leq |\psi(x_t - \log h_t(\theta)) - \psi(x_t - \log \tilde{h}_t(\theta))| \\ &\quad + |\psi'(x_t - \log h_t(\theta)) - \psi'(x_t - \log \tilde{h}_t(\theta))|. \end{aligned}$$

Using the similar argument as in the proof of Lemma 6 and Lemma 1(2), we obtain

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n b_t \|\nabla \log \tilde{h}_t(\theta) \nabla' \log \tilde{h}_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad a.s. \tag{57}$$

From (55), (56) and (57), we have

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|L_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad a.s. \tag{58}$$

We also have

$$M_t(\theta) = [\psi(x_t - \log h_t(\theta)) - \psi(x_t - \log \tilde{h}_t(\theta))] \frac{\nabla^2 h_t(\theta)}{h_t(\theta)} \\ + \psi(x_t - \log \tilde{h}_t(\theta)) \left[\frac{\nabla^2 h_t(\theta)}{h_t(\theta)} - \frac{\nabla^2 \tilde{h}_t(\theta)}{\tilde{h}_t(\theta)} \right].$$

From Lemma 1.4 and the same arguments as in the proof of Lemma 6, we obtain

$$\sup_{\theta \in U} \frac{1}{n-p} \sum_{t=p+1}^n \|M_t(\theta)\| \xrightarrow{n \rightarrow \infty} 0 \quad a.s. \quad (59)$$

From (58) and (59), we have the result. \square

Proof of Theorem 3 From Lemma 4 and Lemma 6, we have

$$\frac{1}{\sqrt{n-p}} \sum_{t=1+p}^n \nabla \rho(x_t - \log \tilde{h}_t(\theta_0)) \xrightarrow{D} \mathcal{N}\left(0, \mathbf{E}(\psi^2(x_t - \log h_t(\theta_0))) D_0\right).$$

From Lemma 5(1) and Lemma 7, we conclude that there exists a neighbourhood U of θ_0 such that

$$\sup_{\theta \in U} \left\| \frac{1}{n-p} \sum_{t=p+1}^n \nabla^2 \rho(x_t - \log \tilde{h}_t(\theta)) - \mathbf{E}(\nabla^2 \rho(x_t - \log h_t(\theta))) \right\| \xrightarrow{n \rightarrow \infty} 0 \quad a.s.$$

The remainder of the proof follows the same arguments as in Theorem 3 of Muler and Yohai [16]. \square

7.2 Appendix 2

We present the boxplots for the estimated parameters of the HGARCH(1,d,1) models with the different estimation methods (Figs. 3, 4, 5, 6, 7, 8).

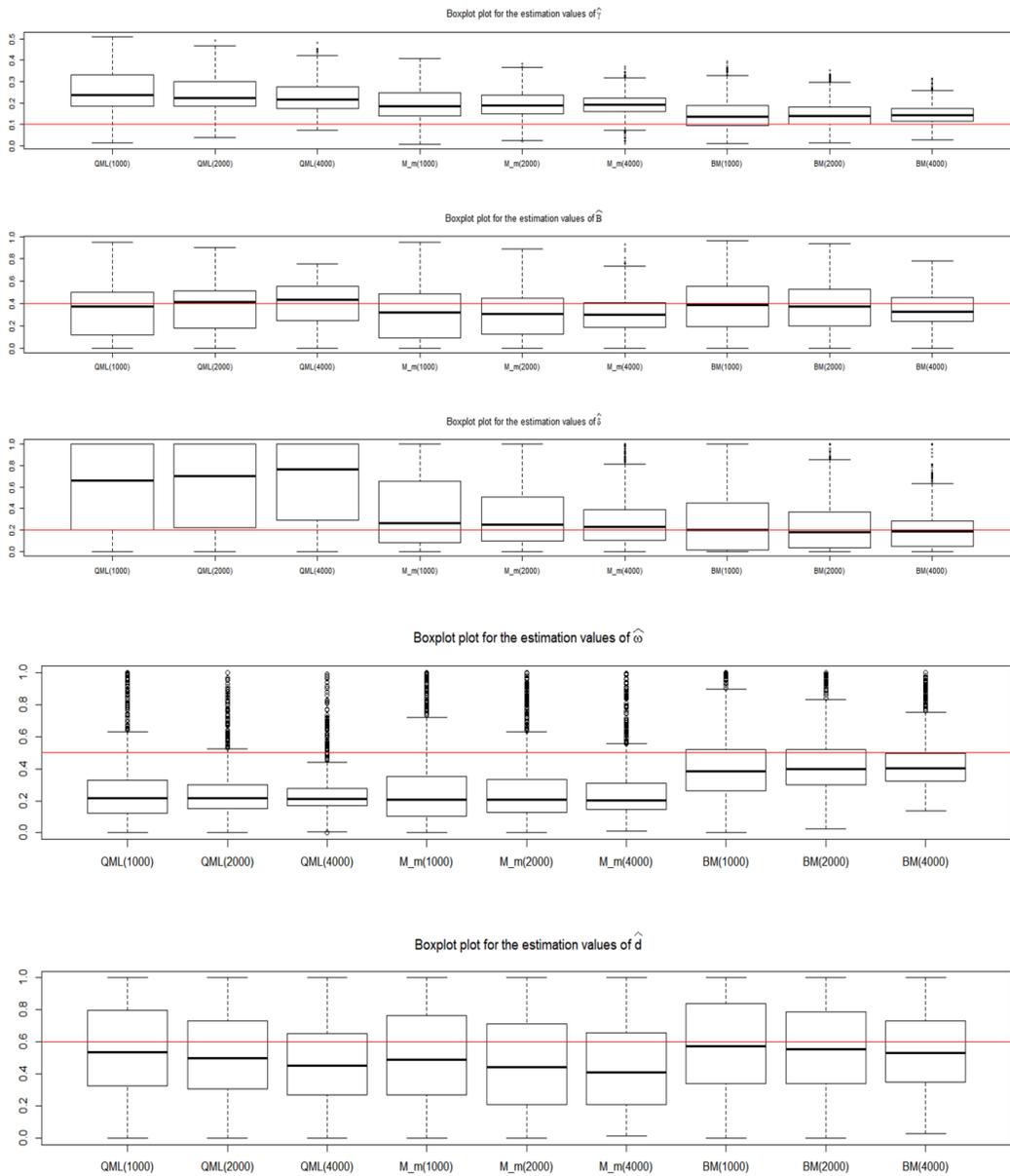


Fig. 3 The boxplots for the different estimation methods for the Normal HGARCH(1,d,1) parameters with 5% of outliers and $\epsilon=3$. The red line denotes the true parameter value

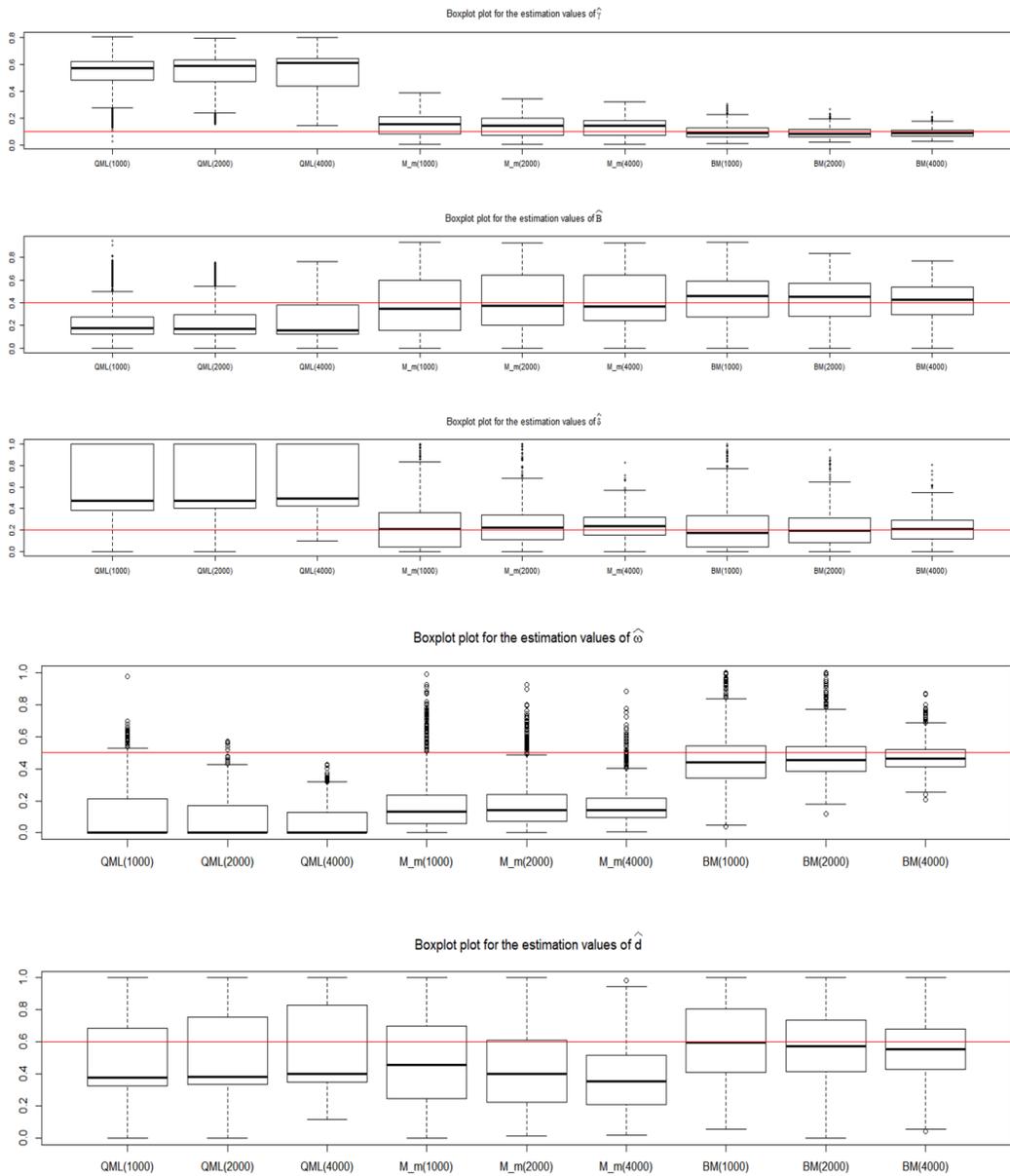


Fig. 4 The boxplots for the different estimation methods for the Normal HGARCH(1,d,1) parameters with 5% of outliers and $\epsilon=5$. The red line denotes the true parameter value

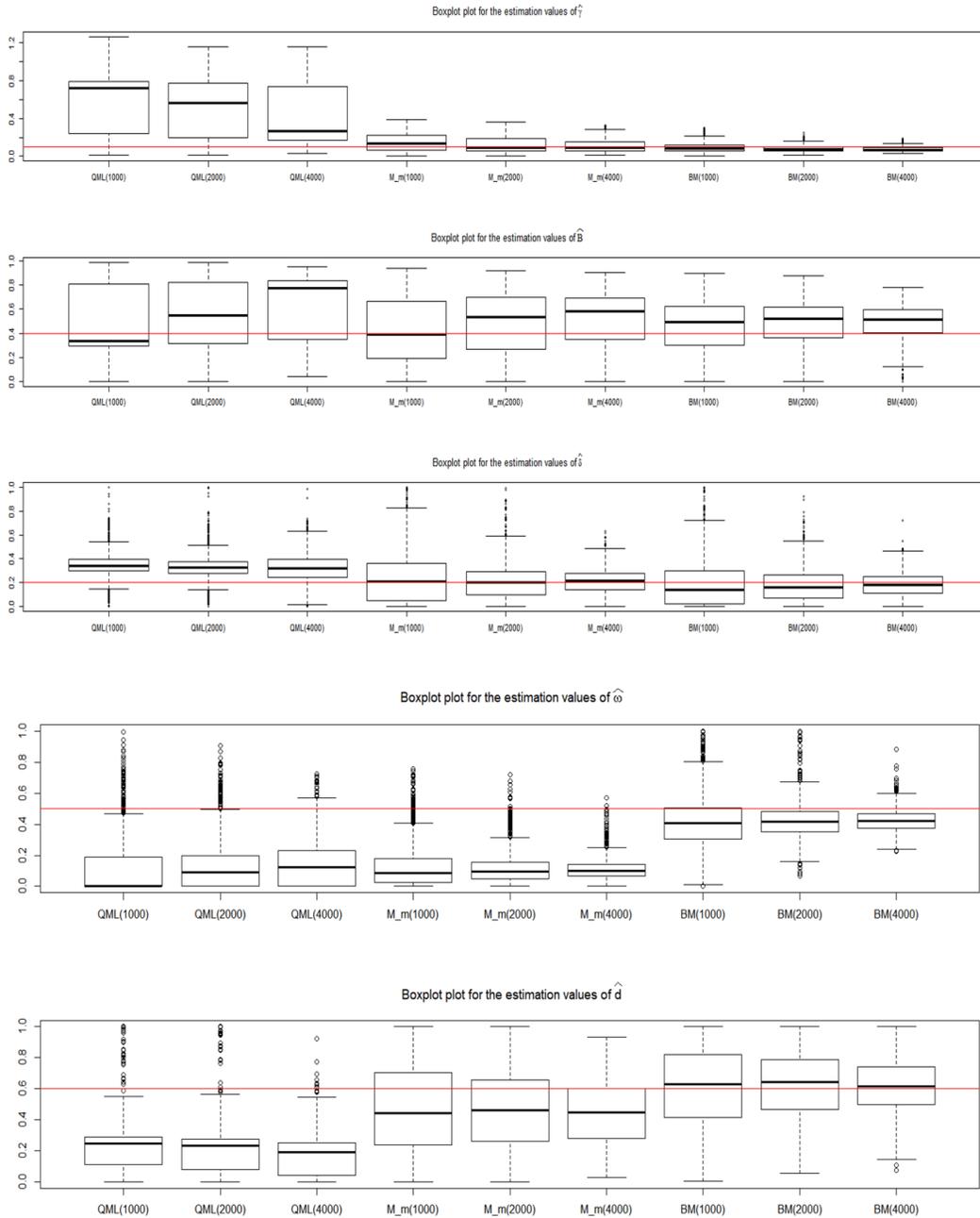


Fig. 5 The boxplots for the different estimation methods for the Normal HGARCH(1,d,1) parameters with 10% of outliers and $\epsilon=5$. The red line denotes the true parameter value

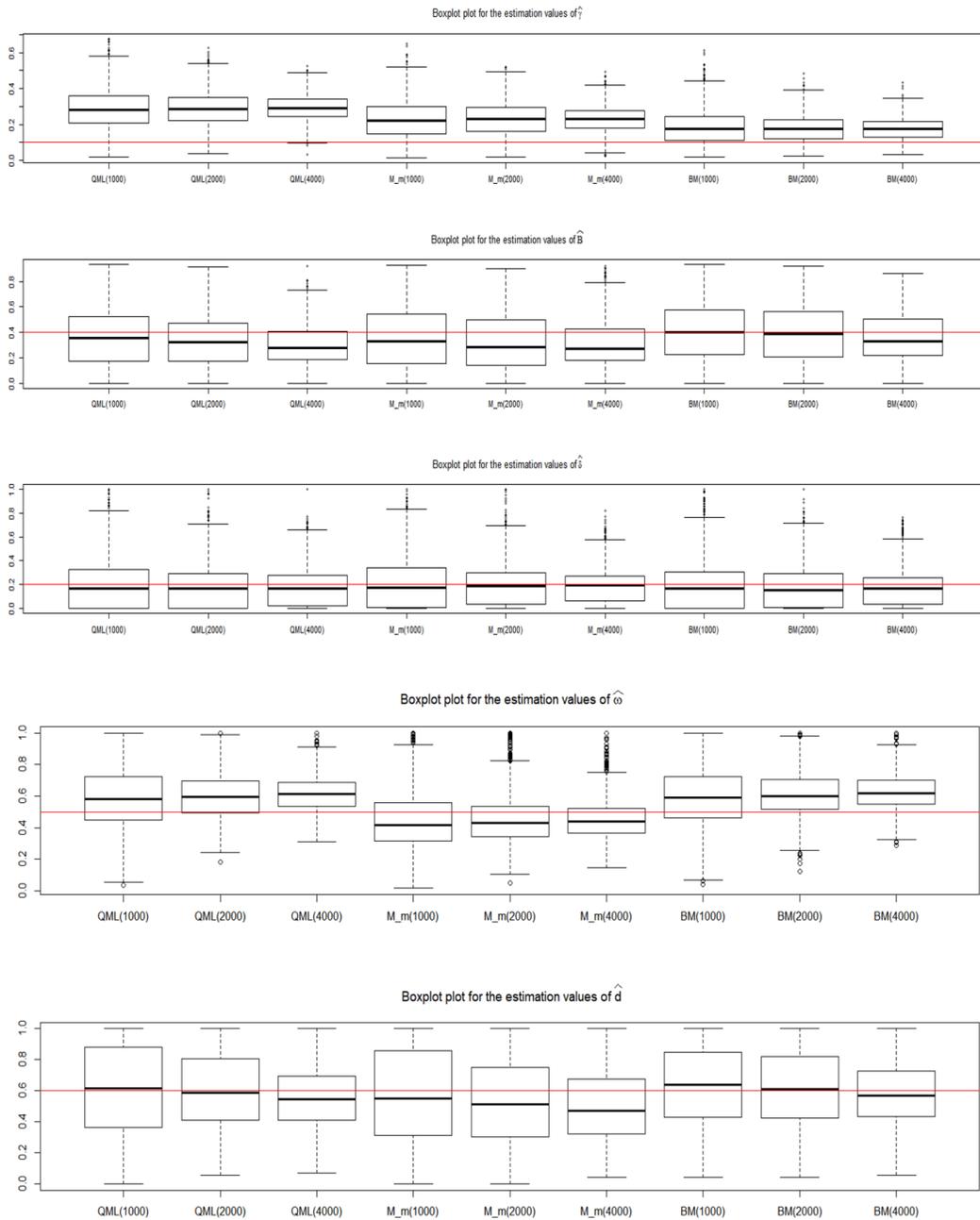


Fig. 6 The boxplots for the different estimation methods for the Student's (t7) HGARCH(1,d,1) parameters with 5% of outliers and $\epsilon=3$. The red line denotes the true parameter value

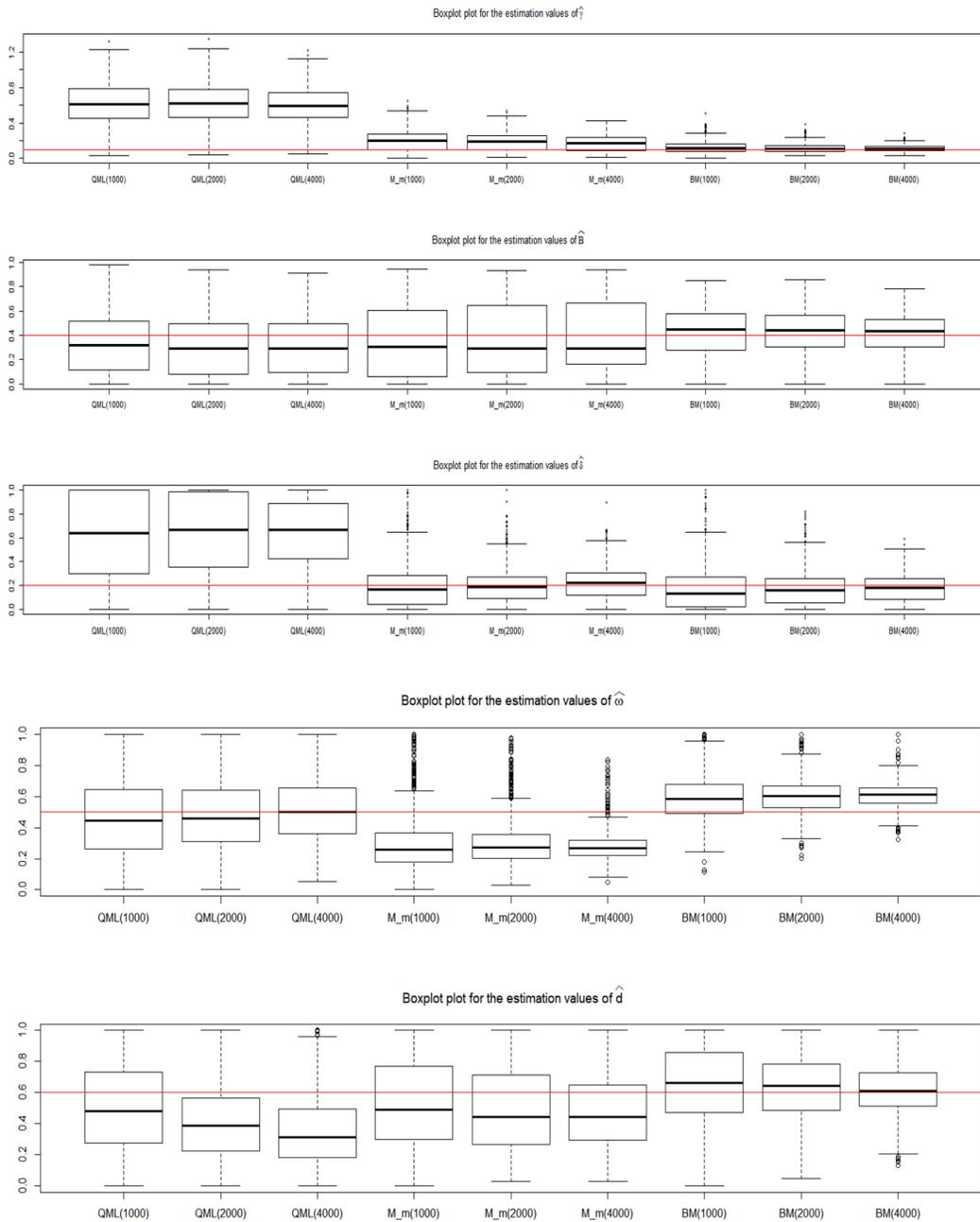


Fig. 7 The boxplots for the different estimation methods for the Student's (t7) HGARCH(1,d,1) parameters with 5% of outliers and $\epsilon=5$. The red line denotes the true parameter value

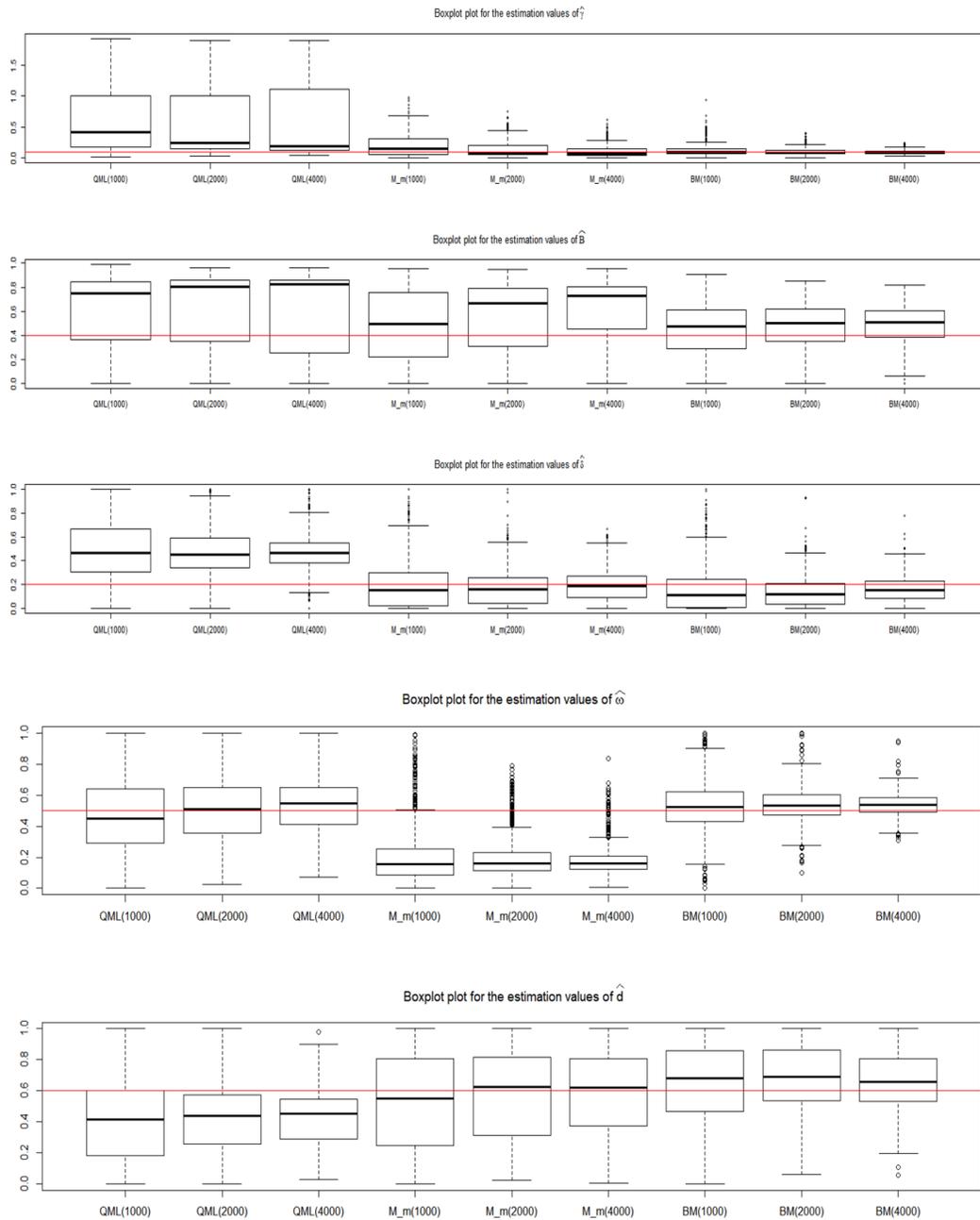


Fig. 8 The boxplots for the different estimation methods for the Student's (t_7) HGARCH(1,d,1) parameters with 10% of outliers and $e=5$. The red line denotes the true parameter value

References

1. Baillie, R.T., Bollerslev, T., Mikkelsen, H.O.: Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econom.* 74(1), 3–30 (1996). [https://doi.org/10.1016/s0304-4076\(95\)01749-6](https://doi.org/10.1016/s0304-4076(95)01749-6)
2. Berkes, I., Horvath, L., Kokoszka, P.: GARCH processes: structure and estimation. *Bernoulli* 9(2), 201–227 (2003). <https://doi.org/10.3150/bj/1068128975>
3. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. *J. Econom.* 31(3), 307–327 (1986). [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
4. Christian, F., Zakoian, J.M.: *GARCH Models : Structure, statistical inference and financial applications*. John Wiley & Sons Ltd (2010)
5. Davidson, J.: Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *J. Business Econom. Statist.* 22(1), 16–29 (2004). <https://doi.org/10.1198/073500103288619359>
6. Ding, Z.X., Granger, C.W.J.: Modeling volatility persistence of speculative returns: A new approach. *J. Econom.* 73, 185–215 (1996)
7. Ding, Z.X., Granger, C.W.J., Engle, R.F.: A long memory property of stock market returns and a new model. *J. Empir. Finance.* 1, 83–106 (1993)
8. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50(4), 987 (1982). <https://doi.org/10.2307/1912773>
9. Fox, A.J.: Outliers in time series. *J. Royal Statist. Soc. Series B (Methodol)* 34(3), 350–363 (1972). <https://doi.org/10.1111/j.2517-6161.1972.tb00912.x>
10. Giraitis, L., Kokoszka, P., Leipus, R., Teyssi re, G.: Rescaled variance and related tests for long memory in volatility and levels. *J. Econom.* 112(2), 265–294 (2003). [https://doi.org/10.1016/s0304-4076\(02\)00197-5](https://doi.org/10.1016/s0304-4076(02)00197-5)
11. Koulikov, D.: Long memory ARCH(∞) models: specification and quasi-maximum likelihood estimation. Working Paper 163, Centre for Analytical Finance, Univ. Aarhus (2003). Available at www.cls.dk/caf/wp/wp-163.pdf and <http://citeseerx.ist.psu.edu/viewdoc/download?10.1.1.194.5011&rep=rep1&type=pdf>
12. Kwan, W., Li, W.K., Li, G.: On the estimation and diagnostic checking of the ARFIMA-HYGARCH model. *Computat. Statist. Data Anal.* 56(11), 3632–3644 (2012). <https://doi.org/10.1016/j.csda.2010.07.010>
13. Li, M., Li, W.K., Li, G.: A new hyperbolic GARCH model. *J. Econom.* 189(2), 428–436 (2015). <https://doi.org/10.1016/j.jeconom.2015.03.034>
14. Lu, W., Wang, Y., Gao, Y.: M-estimates for the multiplicative error model. *J. Statist. Comput. Simul.* 90(1), 1–27 (2019). <https://doi.org/10.1080/00949655.2019.1671387>
15. Muler, N., Yohai, V.J.: Robust estimates for arch processes. *J. Time Ser. Anal.* 23(3), 341–375 (2002). <https://doi.org/10.1111/1467-9892.00268>
16. Muler, N., Yohai, V.J.: Robust estimates for GARCH models. *J. Statist. Plann. Inference* 138(10), 2918–2940 (2008). <https://doi.org/10.1016/j.jspi.2007.11.003>
17. Robinson, P.M., Zaffaroni, P.: Pseudo-maximum likelihood estimation of ARCH(∞) models. *Annal. Statist.* 34(3), 1049–1074 (2006). <https://doi.org/10.1214/009053606000000245>
18. Stout, W.F.: *Almost Sure Convergence*. Academic Press, Inc (1974)
19. Straumann, D.: *Estimation in conditionally heteroscedastic time series models*. Berlin, Springer (2005)
20. Vinutha, H.P., Poornima, B., Sagar, B.M.: Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. In: Satapathy S., Tavares J., Bhateja V., Mohanty J. (eds) *Information and Decision Sciences. Advances in Intelligent Systems and Computing*, vol 701. Springer, Singapore. (2018) https://doi.org/10.1007/978-981-10-7563-6_53

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Références bibliographiques

- [1] Amemiya, T. *Advanced econometrics*. Harvard University Press Cambridge, Massachusetts, 1985.
- [2] Aragon, Y. *Séries temporelles avec R Méthodes et cas*. Springer-Verlag France, 2011.
- [3] Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 1 (sep 1996), 3–30.
- [4] Bardet, J.-M., and Wintenberger, O. Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *The Annals of Statistics* 37, 5B (oct 2009).
- [5] Bauwens, L., Laurent, S., and Rombouts, J. V. K. Multivariate GARCH models : a survey. *Journal of Applied Econometrics* 21, 1 (jan 2006), 79–109.
- [6] Berkes, I., Horvath, L., and Kokoszka, P. GARCH processes : structure and estimation. *Bernoulli* 9, 2 (apr 2003), 201–227.
- [7] Bibi, A. Asymptotic properties of QML estimation of multivariate periodic CCC - GARCH models. *Mathematical Methods of Statistics* 27, 3 (jul 2018), 184–204.
- [8] Bisaglia, L. Model selection for long-memory models. *Quaderni di Statistica* 4 (2002).
- [9] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 3 (apr 1986), 307–327.
- [10] Bollerslev, T. Modelling the coherence in short-run nominal exchange rates : A multivariate generalized arch model. *The Review of Economics and Statistics* 72, 3 (aug 1990), 498.

- [11] Bollerslev, T., Engle, R. F., and Wooldridge, J. M. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96, 1 (feb 1988), 116–131.
- [12] Bougerol, P., and Picard, N. Stationarity of garch processes and of some nonnegative time series. *Journal of Econometrics* 52, 1-2 (apr 1992), 115–127.
- [13] Brunetti, C., and Gilbert, C. L. Bivariate FIGARCH and fractional cointegration. *SSRN Electronic Journal* (2000).
- [14] Chen, C., and Liu, L.-M. Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association* 88, 421 (mar 1993), 284–297.
- [15] Christian, F., and Zakoian, J.-M. *GARCH Models : Structure, Statistical Inference and Financial Applications*. John Wiley & Sons Ltd, 2010.
- [16] Comte, F., and Lieberman, O. Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84, 1 (jan 2003), 61–84.
- [17] Conrad, C. Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics* 4, 3 (may 2006), 413–449.
- [18] Conrad, C. Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics* 157, 2 (aug 2010), 441–457.
- [19] Davidson, J. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business & Economic Statistics* 22, 1 (jan 2004), 16–29.
- [20] Demirel, M., and Unal, G. Applying multivariate-fractionally integrated volatility analysis on emerging market bond portfolios. *Financial Innovation* 6, 1 (dec 2020).
- [21] Diebold, F. X., and Mariano, R. S. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13, 3 (jul 1995), 253.
- [22] Ding, Z., and Granger, C. W. Modeling volatility persistence of speculative returns : A new approach. *Journal of Econometrics* 73, 1 (jul 1996), 185–215.
- [23] Ding, Z., Granger, C. W., and Engle, R. F. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 1 (jun 1993), 83–106.

- [24] Douc, R., Roueff, F., and Soulier, P. On the existence of some arch(∞) processes. *Stochastic Processes and their Applications* 118, 5 (may 2008), 755–761.
- [25] Doukhan, P. *Stochastic Models for Time Series*. Springer International Publishing AG, 2018.
- [26] Doukhan, P., and Wintenberger, O. Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications* 118, 11 (nov 2008), 1997–2013.
- [27] Engle, R. Dynamic conditional correlation. *Journal of Business & Economic Statistics* 20, 3 (jul 2002), 339–350.
- [28] Engle, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50, 4 (jul 1982), 987.
- [29] Engle, R. F., and Kroner, K. F. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 1 (feb 1995), 122–150.
- [30] Fox, A. J. Outliers in time series. *Journal of the Royal Statistical Society : Series B (Methodological)* 34, 3 (jul 1972), 350–363.
- [31] Francq, C., and Zakoïan, J.-M. Qml estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory* 28, 1 (aug 2011), 179–206.
- [32] Giraitis, L., Kokoszka, P., Leipus, R., and Teyssière, G. Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics* 112, 2 (feb 2003), 265–294.
- [33] Giraitis, L., Leipus, R., and Surgailis, D. Recent advances in ARCH modelling. In *Long Memory in Economics*. Springer Berlin Heidelberg, 2007, pp. 3–38.
- [34] Giraitis, L., Surgailis, D., and Škarnulis, A. Stationary integrated ARCH(∞) and AR(∞) processes with finite variance. *Econometric Theory* 34, 6 (oct 2017), 1159–1179.
- [35] Grubbs, F. E. Procedures for detecting outlying observations in samples. *Technometrics* 11, 1 (feb 1969), 1–21.
- [36] He, C., and Teräsvirta, T. An extended constant conditionnal correlation GARCH models and its fourth-moment structure. *Econometric Theory* 20, 05 (oct 2004).
- [37] Horn, R. A., and Johnson, C. R. *Matrix Analysis*. Cambridge University Press, dec 1985.

- [38] Huber, P. J. Robust estimation of a location parameter. *The Annals of Mathematical Statistics* 35, 1 (mar 1964), 73–101.
- [39] Huber, P. J. Robust regression : Asymptotics, conjectures and monte carlo. *The Annals of Statistics* 1, 5 (sep 1973), 799–821.
- [40] Jeantheau, T. Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14, 1 (feb 1998), 70–86.
- [41] Kirman, A., and Teyssière, G. Microeconomic models for long memory in the volatility of financial time series. *Studies in Nonlinear Dynamics & Econometrics* 5, 4 (jan 2002).
- [42] Koulikov, D. Long memory ARCH(∞) models : specification and quasi-maximum likelihood estimation. *Working Paper 163, Centre for Analytical Finance, Univ. Aarhus* (2003).
- [43] Kwan, W., Li, G., and Li, W. K. On the threshold hyperbolic GARCH models. *Statistics and Its Interface* 4, 2 (2011), 159–166.
- [44] Kwan, W., Li, W. K., and Li, G. On the estimation and diagnostic checking of the ARFIMA–HYGARCH model. *Computational Statistics & Data Analysis* 56, 11 (nov 2012), 3632–3644.
- [45] Li, M., Li, G., and Li, W. K. Score tests for hyperbolic GARCH models. *Journal of Business & Economic Statistics* 29, 4 (oct 2011), 579–586.
- [46] Li, M., Li, W. K., and Li, G. A new hyperbolic GARCH model. *Journal of Econometrics* 189, 2 (dec 2015), 428–436.
- [47] Longin, F. La théorie des valeurs extrêmes : présentation et premières applications en finance. *Journal de la Société Française de Statistique* 136, 1 (1995), 77–97.
- [48] Loève, M. *Probability theory I 4th edition*. Springer New York, 1977.
- [49] Lu, W., Wang, Y., and Gao, Y. M-estimates for the multiplicative error model. *Journal of Statistical Computation and Simulation* 90, 1 (oct 2019), 1–27.
- [50] Lux, T., Segnon, M., and Gupta, R. Modeling and forecasting crude oil price volatility :evidence from historical and recent data. Paper, No. 31, Kiel University, FinMaP - Financial Distortions and Macroeconomic Performance,Kiel.

- [51] Magnus, J. R., and Magnus, J. R. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, feb 2019.
- [52] Minovic, J. Z. Modeling multivariate volatility processes : Theory and evidence. *Theoretical and Applied Economics* (2009).
- [53] Mootamri, I. Long memory process in asset returns with multivariate GARCH innovations. *Economics Research International 2011* (sep 2011), 1–15.
- [54] Muler, N., and Yohai, V. J. Robust estimates for arch processes. *Journal of Time Series Analysis 23*, 3 (may 2002), 341–375.
- [55] Muler, N., and Yohai, V. J. Robust estimates for GARCH models. *Journal of Statistical Planning and Inference 138*, 10 (oct 2008), 2918–2940.
- [56] Ñíguez, T.-M., and Rubia, A. Forecasting the conditional covariance matrix of a portfolio under long-run temporal dependence. *Journal of Forecasting 25*, 6 (2006), 439–458.
- [57] Pafka, S., and Matyas, L. Multivariate diagonal FIGARCH : Specification, estimation and application to modelling exchange rates volatility. *SSRN Electronic Journal* (2001).
- [58] Robinson, P. M., and Zaffaroni, P. Pseudo-maximum likelihood estimation of ARCH(∞) models. *The Annals of Statistics 34*, 3 (jun 2006).
- [59] Silvennoinen, A., and Teräsvirta, T. Multivariate GARCH models. In *Handbook of Financial Time Series*. Springer Berlin Heidelberg, 2009, pp. 201–229.
- [60] Stout, W. F. *Almost Sure Convergence*. Academic Press, Inc, 1974.
- [61] Straumann, D. *Estimation in conditionally heteroscedastic time series models*. Berlin Heidelberg Springer, 2005.
- [62] Tang, T.-L., and Shieh, S.-J. Long memory in stock index futures markets : A value-at-risk approach. *Physica A : Statistical Mechanics and its Applications 366* (jul 2006), 437–448.
- [63] Teyssière, G. Modelling exchange rates volatility with multivariate long-memory arch processes, 1998.

- [64] Vee, D. C., Gonpot, P., and Sookia, N. Forecasting volatility of USD/MUR exchange rate using a GARCH (1,1) model with GED and student's-t errors. *University of Mauritius Research Journal* 17, 1 (oct 2011).
- [65] Vinutha, H. P., Poornima, B., and Sagar, B. M. Detection of outliers using inter-quartile range technique from intrusion dataset. In *Advances in Intelligent Systems and Computing*. Springer Singapore, 2018, pp. 511–518.