Stochastic recurrent equations : structure, statistical inference, and financial applications

Baye Matar Kandji

INSTITUT POLYTECHNIQUE DE PARIS

ENSAE

IP PARIS

NNT : 2023IPPAG004

Thèse de doctorat

Stochastic recurrent equations: structure, statistical inference, and financial applications

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École nationale de la statistique et de l'administration économique

École doctorale n°574 École doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 27 juin 2023, par

BAYE MATAR KANDJI

Composition du Jury :

Alessandra Luati Professeur, Imperial College London et Università di Bologna	Président
Jean Marc Bardet Professeur, Université Paris 1 Panthéon-Sorbonne	Rapporteur
Lorenzo Trapani Professeur, University of Leicester	Rapporteur
François Roueff Professeur, Télécom Paris - Institut Polytechnique de Paris	Examinateur
Melanie Schienle Professeur, Karlsruhe Institute of Technology (KIT)	Examinateur
Christian Francq Professeur, Universite de Lille et ENSAE-CREST	Directeur de thèse
Jean-Michel Zakoïan Professeur, Universite de Lille et ENSAE-CREST	Directeur de thèse

Remerciements

Ce manuscrit de thèse est l'aboutissement d'un voyage tout sauf solitaire et je tiens à exprimer ma profonde gratitude à tous ceux qui m'ont soutenu tout au long du parcours qui m'a mené à son achèvement. Sans leurs encouragements, leurs conseils et leur aide indéfectibles, cette réalisation n'aurait pas été possible. Les lignes qui suivent rendent hommage à tous ceux qui ont participé de près ou de loin à cette aventure.

Tout d'abord, je souhaite remercier mes directeurs de thèse, Christian Francq et Jean-Michel Zakoian, pour leur mentorat inestimable et leur soutien continu. Leur expertise, leur patience, leur bienveillance et leur dévouement ont joué un rôle déterminant dans l'orientation de mes recherches et l'amélioration de mes capacités académiques. Je leur suis vraiment reconnaissant pour leurs conseils, leurs commentaires perspicaces et les innombrables heures qu'ils ont consacrées à la révision et à l'amélioration de mon travail.

I would like to express my sincere gratitude to Professors Jean Marc Bardet and Lorenzo Trapani for reviewing this research work. I thank them for the time and dedication they devoted to examining this manuscript. Their insightful comments have greatly contributed to the quality and improvement of this work. I would also like to thank Professors Alessandra Luati, François Roueff, and Melanie Schienle for accepting to be the examiners of my thesis.

J'adresse également mes remerciements au personnel et aux membres du laboratoire de Finance du CREST, qui m'ont offert un environnement universitaire stimulant et les ressources nécessaires pour mener à bien mes recherches. En particulier, j'aimerai remercier Christian Robert pour l'intérêt qu'il a porté à mes recherches et son soutient depuis l'ISFA, ainsi que Jean-David Fermanian, Christian Gourieroux, Caroline Hillairet, Francesco Violante et Peter Tankov. Je suis également redevable à mes collègues doctorants (certains sont devenus docteurs) et post-doctorants: Alicia, Christopher, David, Fallou, Fanny, François, Hugo, Julien, Marcos, Maria, Mariia, Redouane, Théo, Thomas et Yousra. Nos discussions perspicaces et leur volonté de partager leurs connaissances et leurs expériences ont joué un rôle inestimable dans ma compréhension du sujet. Une mention spéciale à Marcos pour nos longues, instructives et passionnantes discussions sur les mathématiques.

Je remercie sincèrement ma famille et mes amis, dont le soutien indéfectible, la compréhension et les encouragements ont été les piliers de ma force. Je suis immensément

reconnaissant envers mes parents pour les sacrifices qu'ils ont consentis pour que je puisse bénéficier des meilleures études. Ce manuscrit de doctorat est autant le vôtre que le mien. Je vous dédie ce travail à tous les deux, avec une profonde gratitude. Je tiens également à exprimer ma plus profonde reconnaissance à ma compagne et meilleure amie, Pheaktra. Ta foi en mes capacités et tes encouragements constants m'ont donné la motivation nécessaire pour persévérer à travers les défis rencontrés tout au long de mes études supérieures et particulièrement au cours de cette thèse.

À tous ceux qui ont joué un rôle dans la réalisation de cette thèse, j'adresse mes remerciements les plus sincères. Mes pensées vont à tous mes professeurs de mathématiques, sans exception, de l'école élémentaire au Sénégal à mes conseillers doctoraux aujourd'hui. Par vos enseignements de qualité, vous avez fait naître en moi une profonde passion pour les mathématiques et vos encouragements ont été ma source constante de force et de motivation. Je pense également à Tom Picard, ami et ancien tuteur de stage chez Nexialog, pour son encadrement formateur lors de ma première expérience de recherche. Je remercie également Christiane Francq d'avoir pris le temps de relire ce manuscrit.

Abstract

We are interested in the theoretical properties of Stochastic Recurrent Equations (SRE) and their applications in finance. These models are widely used in econometrics, including financial econometrics, to explain the dynamics of various processes such as the volatility of financial returns. However, the probability structure and statistical properties of these models are still not well understood, especially when the model is considered in infinite dimensions or driven by non-independent processes. These two features lead to significant difficulties in the theoretical study of these models. In this context, we aim to explore the existence of stationary solutions and the statistical and probabilistic properties of these solutions.

We establish new properties on the trajectory of the stationary solution of SREs, which we use to study the asymptotic properties of the quasi-maximum likelihood estimator (QMLE) of GARCH-type (generalized autoregressive conditional heteroskedasticity) conditional volatility models. In particular, we study the stationarity and statistical inference of semi-strong GARCH(p, q) models where the innovation process is not necessarily independent. We establish the consistency of the QMLE of semi-strong GARCHs without assuming the commonly used condition that the stationary distribution admits a small-order moment. In addition, we are interested in the two-factor volatility GARCH models (GARCH-MIDAS); a long-run, and a short-run volatility. These models were recently introduced by Engle et al. (2013) and have the particularity to admit stationary solutions with heavy-tailed distributions. These models are now widely used but their statistical properties have not received much attention. We show the consistency and asymptotic normality of the QMLE of the GARCH-MIDAS models and provide various test procedures to evaluate the presence of long-run volatility in these models. We also illustrate our results with simulations and applications to real financial data.

Finally, we extend a result of Kesten (1975) on the growth rate of additive sequences to superadditive processes. From this result, we derive generalizations of the contraction property of random matrices to products of stochastic operators. We use these results to establish necessary and sufficient conditions for the existence of stationary solutions of the affine case with positive coefficients of SREs in the space of continuous functions. This class of models includes most conditional volatility models, including functional GARCHs.

Keywords: Inference Without Moments, Iterated Function Systems, Lyapunov exponent, Multiplicative Component GARCH, Residual Bootstrap, Subadditive sequence, Tests on boundary parameters.

Résumé

Nous nous intéressons à l'étude des propriétés théoriques des équations récurrentes stochastiques (SRE) et de leurs applications en finance. Ces modèles sont couramment utilisés en économétrie, y compris en économétrie de la finance, pour styliser la dynamique d'une variété de processus tels que la volatilité des rendements financiers. Cependant, la structure de probabilité ainsi que les propriétés statistiques de ces modèles sont encore mal connues, particulièrement lorsque le modèle est considéré en dimension infinie ou lorsqu'il est généré par un processus non indépendant. Ces deux caractéristiques entraînent de formidables difficultés à l'étude théorique de ces modèles. Dans ces contextes, nous nous intéressons à l'existence de solutions stationnaires, ainsi qu'aux propriétés statistiques et probabilistes de ces solutions.

Nous établissons de nouvelles propriétés sur la trajectoire de la solution stationnaire des SREs que nous exploitons dans l'étude des propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance (QMLE) des modèles de volatilité conditionnelle de type GARCH. En particulier, nous avons étudié la stationnarité et l'inférence statistique des modèles GARCH(p, q) semi-forts dans lesquels le processus d'innovation n'est pas nécessairement indépendant. Nous établissons la consistance du QMLE des GARCH (p, q) semi-forts sans hypothèses d'existence de moment, couramment supposée pour ces modèles, sur la distribution stationnaire. De même, nous nous sommes intéressés aux modèles GARCH à deux facteurs (GARCH-MIDAS); un facteur de volatilité à long terme et un autre à court terme. Ces récents modèles introduits par Engle et al. (2013) ont la particularité d'avoir des solutions stationnaires avec des distributions à queue épaisse. Ces modèles sont maintenant fréquemment utilisés en économétrie, cependant, leurs propriétés statistiques n'ont pas reçu beaucoup d'attention jusqu'à présent. Nous montrons la consistance et la normalité asymptotique du QMLE des modèles GARCH-MIDAS et nous proposons différentes procédures de test pour évaluer la présence de volatilité à long terme dans ces modèles. Nous illustrons nos résultats avec des simulations et des applications sur des données financières réelles.

Enfin, nous étendons le résultat de Kesten (1975) sur le taux de croissance des séquences additives aux processus superadditifs. Nous déduisons de ce résultat des généralisations de la propriété de contraction des matrices aléatoires aux produits d'opérateurs stochastiques. Nous utilisons ces résultats pour établir des conditions nécessaires et suffisantes d'existence de solutions stationnaires du modèle affine à coefficients positifs des SREs dans l'espace des fonctions continues. Cette classe de modèles regroupe la plupart des modèles de volatilité conditionnelle, y compris les GARCH fonctionnels.

Mots clés: Inférence sans moments, Systèmes stochastiques de fonctions itérées, Exposant de Lyapunov, GARCH à composantes multiplicatives, Bootstrap, Processus sous-additif, Tests Statistiques.

Contents

Remerciements i

Abstract iii

Résumé v

List of Figures ix

List of Tables xi

1 Introduction 1
1 Context . . . 2
2 Main results . . . 5

2 Introduction (Français) 13
1 Contexte . . . 14
2 Synthèse des principaux résultats . . . 18

3 IFS and semi-strong GARCH 27
1 Introduction . . . 28
2 Stochastic IFS without moments . . . 29
3 Proof of the main result . . . 31
4 Inference for semi-strong GARCH . . . 36
5 Appendix: Complementary Proofs . . . 41

4 Inference on GARCH-MIDAS models 47
1 Introduction . . . 48
2 Stationarity and existence of moments . . . 49
3 QMLE of GARCH-MIDAS . . . 52
4 Testing the existence of a long-run volatility . . . 55
5 Numerical results . . . 59
6 Conclusion . . . 70
7 Appendix: proofs . . . 70

5 Superadditive processes and fGARCH **81**
1 Introduction 82
2 The growth rate of superadditive processes 83
3 Stationarity of fGARCH models in $\mathcal{C}^0$ 94
4 Perspective 101
5 Appendix: Complementary Proofs 101

Conclusion and perspectives **107**

Bibliography **109**

List of Figures

1.1 Barnsley fern . . . 3
1.2 Decomposition of short- and long-term volatilities for the Euronext 100 stock index . . . 8

2.1 Fougere de Barnsley . . . 15
2.2 Décomposition des volatilités à court et long terme pour l'indice boursier Euronext 100 . . . 20

4.1 Simulation of $r_t = \sqrt{1 + 0.1 r_{t-1}^2}\epsilon_t$ with $\epsilon_t = \sqrt{1 + 0.05\epsilon_{t-1}^2}\eta_t$, $\eta_t \sim \mathcal{N}(0,1)$. . . . 52
4.2 Examples of estimated short and long term volatilities when the GDP is a GARCH-MIDAS (left figure) or a standard GARCH (right figure) . . . 63
4.3 Estimated short and long term volatilities of the GARCH-MIDAS model . 64
4.4 True and estimated volatility estimated by a GARCH-MIDAS and by a standard GARCH . . . 65
4.5 Bootstrap estimate of the distribution of $\widehat{a}_n$ when $a = 0$ (in blue) and observed value of $\widehat{a}_n$ (red vertical line). . . . 68
4.6 GARCH-MIDAS short and long term volatilities for four stock indices from 1990-03-01 to 2021-04-08. . . . 69

List of Tables

4.1 Distribution of the QMLE over 1000 replications 60
4.2 Distribution of the QMLE of a GARCH-MIDAS, when the DGP is a GARCH-MIDAS (first part) and when it is a standard GARCH (second part of the table). In the latter case, the parameter ϑ is undefined (UD). . 62
4.3 Distribution of the QMLE of a GARCH(1,1) when the DGP is the GARCH-MIDAS of Table 4.2 (top panel). 62
4.4 Distribution of the QLIK losses over 1000 replications when the GARCH-MIDAS volatility is estimated by the GARCH-MIDAS model or by the GARCH model. 63
4.5 Empirical relative frequency of rejection of the null that there exists no long-run volatility (*i.e.* $a_0 = 0$) using the score, Wald and LR tests with a fixed value of ϑ, for nominal levels varying from 0.1% to 20%. 66
4.6 Empirical relative frequency of rejection of the null that there exists no long-run volatility (*i.e.* $a_0 = 0$) using the bootstrapped version of the Wald test, for nominal levels varying from 0.1% to 20%. 67
4.7 GARCH-MIDAS fitted on stock returns. The estimated standard deviations are displayed in small font, under the estimated values of the coefficients. The last column gives the bootstrap estimated p-value of the Wald test of $H_0 : a = 0$. 67

Chapter 1

Introduction

Abstract

This thesis focuses on the study of the theoretical properties of stochastic recurrence equations and their applications in financial econometrics. We are interested in the existence of stationary solutions as well as the statistical and probabilistic properties of these solutions. In particular, we will study the stationarity and statistical inference of semi-strong Generalised AutoRegressive Conditional Heteroskedasticity (GARCH) (p, q) models, where the innovation process is not necessarily independent. We will also focus on two-factor GARCH volatility models, one long-run and one short-run volatility. We will illustrate our results with simulations and applications to real financial data. Finally, we will study the contraction property of products of random operators, with an application to the study of the existence of a stationary solution for conditional volatility models in the space of continuous functions (fGARCH).

Contents

1 Context 2
2 Main results 5

1 Context

In recent years, autoregressive processes of the form

$$\boldsymbol{X}_t = \Psi\left(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1}\right) = \boldsymbol{\Psi}_t\left(\boldsymbol{X}_{t-1}\right), \quad t \in \mathbb{Z}, \tag{1.1}$$

where $\Psi(\theta, \cdot)$ is an operator on a Polish space $(F,\ d)$ to itself and $(\boldsymbol{\theta}_t)$ is a strictly stationary and ergodic process, have attracted much attention. This equation, commonly called Iterated Function Systems (IFS), includes a wide range of models. Among them we can mention the Stochastic Recurrence Equation (SRE) introduced by Kesten (1973), where $F = \mathbb{R}^d$, $\boldsymbol{\theta}_t = (\boldsymbol{A}_t, \boldsymbol{B}_t)$ where $\boldsymbol{A}_t$ is a matrix $d \times d$ and $\boldsymbol{B}_t$ is a vector $d \times 1$, and $\Psi(\theta, \cdot)$ is the affine operator such that

$$\boldsymbol{X}_t = \boldsymbol{A}_t \boldsymbol{X}_{t-1} + \boldsymbol{B}_t, \quad t \in \mathbb{Z}. \tag{1.2}$$

SREs include AutoRegressive-Moving-Average (ARMA) and GARCH processes, see Eq. (1.12). A class of bivariate SREs is also often used to draw images of two-dimensional fractals, such as Barnsley's fern, see figure 1.1, Koch's flake or the dragon curve. The construction procedure is as follows: consider a finite sequence of affine transformations of $\mathbb{R}^2$, i.e. $\left(\mathbf{a}_i, \mathbf{b}_i\right), i = 1, \ldots, k$, for some $k \geq 1$, where $(\mathbf{a}_i)$ are matrices 2×2 and $(\mathbf{b}_i)$ are vectors in $\mathbb{R}^2$. Let us assume that $\mathbf{a}_i$ are contracting matrices, i.e. that their respective spectral radii ρ_i are strictly less than 1. Considering positive weights $p_1, \ldots, p_k$, with $p_1 + \cdots + p_k = 1$, we can define the Markov chain $(\boldsymbol{X}_t)$ moving in $\mathbb{R}^2$: starting from x, the chain proceeds by randomly choosing i with probability p_i and moving to $\boldsymbol{X}_{t+1} = a_i \boldsymbol{X}_t + b_i$. The fractal image is obtained by plotting the trajectory $(\boldsymbol{X}_t)$ over the plane from an arbitrarily large n to reduce the effect of initialisation to approximate the stationary distribution. For more details on fractal SREs, see Barnsley (2014). The class of AutoRegressive (AR) models with ARCH errors introduced by Weiss (1984) is an example of processes that verify (1.1) and which are not SREs. A special case of these processes is the AR(1) model with ARCH errors defined by :

$$\boldsymbol{X}_t = \alpha \boldsymbol{X}_{t-1} + \sqrt{\beta + \lambda \boldsymbol{X}_{t-1}^2}\boldsymbol{\eta}_t, \quad t \in \mathbb{Z}, \tag{1.3}$$

where $(\boldsymbol{\eta}_t)$ are independent and identically distributed random variables (iid), $\alpha \in \mathbb{R}$, $\beta > 0$ and $\lambda > 0$. The model (1.3) was considered in Borkovec and Klüppelberg (2001) and Borkovec (2000). Other examples of IFS applications can be found in Buraczewski et al. (2016, Examples 1.0.1 to 1.0.5) and in Diaconis and Freedman (1999).

1.1 Stationarity

Because of their widespread use, the theoretical properties of the class of models (1.1) have also attracted much interest over the last 50 years. In the univariate case of SREs (1.2) with iid innovations $(\boldsymbol{A}_t, \boldsymbol{B}_t) \in \mathbb{R}^2$, the existence of stationary solutions has been studied by, among others, Kesten (1973), Vervaat (1979), Goldie (1991) and Goldie and Maller (2000). The latter paper provides a complete characterisation of the

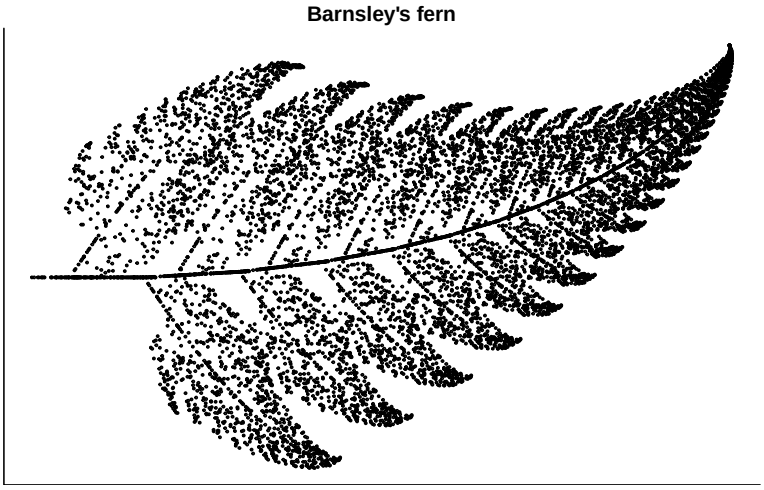

Figure 1.1 – Barnsley fern with 10.000 points with the parameters (see Barnsley (2014, table 3.8.3 and figure 3.8.3)):

$$(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4) = \left(\begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.16 \end{pmatrix}, \begin{pmatrix} 0.85 & 0.04 \\ -0.04 & 0.85 \end{pmatrix}, \begin{pmatrix} 0.20 & -0.26 \\ 0.23 & 0.22 \end{pmatrix}, \begin{pmatrix} -0.15 & 0.28 \\ 0.26 & 0.24 \end{pmatrix} \right)$$
$$(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4) = \left(\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 1.6 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 1.6 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 0.44 \end{pmatrix} \right), \text{ and}$$
$$(p_1, p_2, p_3, p_4) = (0.01, 0.85, 0.07, 0.07)$$

problem by giving a set of necessary and sufficient conditions. In the multivariate iid case, the weakest conditions for the existence of a stationary solution are due to Bougerol and Picard (1992b). Under an irreducibility condition of the model, they show that if $\mathbb{E} \log^+ \boldsymbol{A}_0$ and $\mathbb{E} \log^+ \boldsymbol{B}_0$ are finite, then the equation (1.2) admits a strictly stationary and non-anticipative solution $(\boldsymbol{X}_t)$, i.e. $\boldsymbol{X}_t$ is $\sigma(\boldsymbol{\theta}_k,\ k \leq t)$-measurable if and only if

$$\boldsymbol{\gamma} := \lim_{n \to +\infty} \frac{1}{n} \log(\boldsymbol{\Lambda}_0^{(n)}) = \inf_{n \geq 1} \frac{1}{n} \mathbb{E} \log(\|\boldsymbol{A}_n \boldsymbol{A}_{n-1} \cdots \boldsymbol{A}_1\|) < 0, \tag{1.4}$$

where $\boldsymbol{A}_t^{(0)} = 1$ and $\boldsymbol{A}_t^{(k)} = \|\boldsymbol{A}_{t-1} \cdots \boldsymbol{A}_{t-k+1}\|$ for any integer $k > 0$. The proof of this result relies in part on another property established in the same article. Commonly called the **contraction property of product random matrices**, this property states that if $(\boldsymbol{A}_t)$ is strictly stationary and ergodic, and $\mathbb{E} \log^+ \boldsymbol{A}_0$ is finite, then

$$\boldsymbol{\gamma}(\boldsymbol{A}) < 0 \ \ a.s. \quad \text{if} \quad \boldsymbol{\Lambda}_0^{(n)} \to 0 \ \ a.s. \text{ when } \ n \to \infty. \tag{1.5}$$

This property is commonly used to establish the necessity of (1.4) for the existence of stationary solutions of GARCH-type models.

For the general model (1.1), Elton (1990, Theorem 3) (see also Bougerol (1993, Theorem 3.1)) established a sufficient condition for the existence of stationary solutions. As for the affine model in finite dimensions, this result relates the sign of the Lyapunov coefficient $\boldsymbol{\gamma}(\boldsymbol{\Psi})$, defined below, of the process $(\boldsymbol{\Psi}_t)$ to the existence of stationary solutions.

The contraction of $(\boldsymbol{\Psi}_t)$ occurs when:

$$\boldsymbol{\gamma}(\boldsymbol{\Psi}) := \limsup_{n\to+\infty} \frac{1}{n} \log \boldsymbol{\Lambda}_0^{(n)} < 0, \tag{1.6}$$

where $\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t) = \sup_{x_1,x_2\in F, x_1\neq x_2} \frac{d(\boldsymbol{\Psi}_t(x_1),\boldsymbol{\Psi}_t(x_2))}{d(x_1,x_2)}$, with $\boldsymbol{\Lambda}_t^{(0)} = 1$ and $\boldsymbol{\Lambda}_t^{(k)} = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t \circ \cdots \circ \boldsymbol{\Psi}_{t-k+1})$ for any integer $k > 0$. For the model (1.3), it can be noted that the sufficient condition for the existence of a stationary solution proposed by Borkovec and Klüppelberg (2001, Eq. 3.2) implies (1.6). To our knowledge, there is no result establishing the necessity of equation (1.6) in the general context of (non-affine) IFS or for affine operators $\boldsymbol{\Psi}_t$ in infinite dimension. Functional ARCH/GARCH models in which the volatility process is a curve constitute a new class of models in which $\boldsymbol{\Psi}_t$ is an affine operator in an L^p space or in the space of continuous functions. A functional GARCH process of order $(1,1)$ (fGARCH$(1,1)$) in $\mathcal{C}[0,1]$ is a sequence $(\boldsymbol{r}_t : t \in \mathbb{Z})$ of random elements, where $(\boldsymbol{r}_t(u) : u \in [0,1])$ is a continuous function on $[0,1]$, satisfying:

$$\boldsymbol{r}_t = \boldsymbol{\sigma}_t \boldsymbol{\eta}_t,\ \boldsymbol{\sigma}_t^2 = \delta + \alpha \boldsymbol{r}_{t-1}^2 + \beta \boldsymbol{\sigma}_{t-1}^2 = \delta + \int \boldsymbol{\gamma}_t(\cdot, s)\boldsymbol{\sigma}_{t-1}^2(s)ds = \delta + \boldsymbol{\gamma}_t \boldsymbol{\sigma}_{t-1}, \tag{1.7}$$

where $(\boldsymbol{\eta}_t : t \in \mathbb{Z})$ is a sequence of iid random functions in $\mathcal{C}[0,1]$, δ is a positive function and the integral operators α and β, i.e. $(\alpha x)(u) = \int \alpha(u,s)x(s)ds$ and $(\beta x)(u) = \int \beta(u,s)x(s)ds$ are positive, i.e. they map nonnegative functions to nonnegative function. $\boldsymbol{\gamma}_t(u,s) = \alpha(u,s)\boldsymbol{\eta}_{t-1}^2(s) + \beta(u,s)$ is an element of $\mathcal{C}[0,1]^2$. We can see that $(\boldsymbol{\sigma}_t^2)$ defined in (1.7) verifies (1.1) with $F = \mathcal{C}[0,1]$, $\boldsymbol{\theta}_t = \boldsymbol{\eta}_t$ and $\boldsymbol{\Psi}_t = \delta + \boldsymbol{\gamma}_t(\cdot)$. The fGARCHs were introduced by Hörmann et al. (2013) and Aue et al. (2017) to model the dynamics of the conditional intra-day volatility curve based on the structure of the classical GARCH models. Sufficient conditions for the existence of stationary solutions for different classes of these models have been established in Aue et al. (2017), Hörmann et al. (2013), Kühnert (2020) and in Cerovecki et al. (2019). The latter paper considers the fGARCH(p,q) in $L^2([0,1])$. In this setup, the authors note that one of the main challenges in establishing the necessity of (1.6) is to extend the contraction property of random matrices (1.5) to operator processes $(\boldsymbol{\Psi}_t)$. To our knowledge, this extension has not yet been established.

1.2 Tails

An important property of the marginal distribution of the stationary solution $(\boldsymbol{X}_t)$ of Equation (1.2) comes from Kesten (1973). Under relatively weak assumptions, he establishes in the univariate iid case that this distribution has a power tail law, in the following sense:

$$\mathbb{P}(|\boldsymbol{X}| > x) \sim c x^{-\alpha}, \quad x \to \infty, \tag{1.8}$$

where c and α are strictly positive constants. In particular, this result implies that $\boldsymbol{X}$ has moments of order $s < \alpha$. The work of Goldie (1991) shows that the property (1.8) can be generalized to the iid cases of Model (1.1) under regularity conditions on $\boldsymbol{\Psi}_t$

which are not necessarily restricted to the affine case (1.2). The work of Borkovec and Klüppelberg (2001) on the model (1.3) also gives similar results. More generally, in the iid case, under (1.6) and some regularity assumptions on $\mathbf{\Psi}_0$, it can be shown that the stationary solution of (1.1) admits a small-order moment. This moment property is often used in the statistical inference of models of type (1.1) , such as GARCH type models with iid innovation. See Berkes et al. (2003) or Francq and Zakoian (2004) for the classical GARCH(p, q). However, when $(\boldsymbol{\theta}_t)$ is stationary and ergodic, though not independent, we show on a class of GARCH models with two volatility factors that the strictly stationary solution may admit no-moment.

1.3 Motivation

As we can see, the theoretical properties of Model (1.1) are not well understood when the sequence $(\boldsymbol{\theta}_t)$ is not identically distributed or when the affine model (1.2) is considered in infinite dimensions. Many econometric models that fall under the (1.1) framework belong to one of these two categories. For example, the fGARCH and the semi-strong GARCH models proposed by Lee and Hansen (1994), in which the innovation process follows a non-iid martingale difference process. The multiplicative-component GARCH models (GARCH-MIDAS) of Engle et al. (2013) are another class of IFS that are not generated by an iid sequence. Another example is GARCH-X models, which are IFS driven by a sequence of innovations and covariates that are generally non-iid.

A simple example of GARCH-X is given by the process:

$$\boldsymbol{\epsilon}_t = \boldsymbol{\sigma}_t \boldsymbol{\eta}_t \quad \boldsymbol{\sigma}_t^2 = \omega + \alpha \boldsymbol{\epsilon}_{t-1}^2 + \beta \boldsymbol{\sigma}_{t-1}^2 + \pi' \boldsymbol{x}_{t-1} = \omega + \pi' \boldsymbol{x}_{t-1} + (\alpha \boldsymbol{\eta}_{t-1}^2 + \beta) \boldsymbol{\sigma}_{t-1}^2$$

Here, $\boldsymbol{x}_t$ is a vector of r exogenous covariates, such as daily trading volume, macro-economic data or other market data, and $\boldsymbol{\pi}$ is a vector of parameters associated with these covariates. In these models, the volatility process depends on the covariate process, in addition to the squares of past returns. This additional information allows for better forecasts of the squared returns. These different examples motivate the study of affine SREss in infinite dimensions or IFS models driven by non-iid innovations.

2 Main results

Chapter 3: Exponential control of the trajectories of Iterated Function Systems and application to semi-strong GARCH(p,q) models.

We know, from Elton (1990), that if there exists an element c in F such that:

$$\mathbb{E} \ln^+ d\left(\mathbf{\Lambda}_0(c), c\right) \text{ and } \mathbb{E} \ln^+ \mathbf{\Lambda}_0 \text{ are finite,} \tag{1.9}$$

then the condition (1.6) is sufficient to guarantee the existence of a stationary solution $(\boldsymbol{X}_t)$ to the equation (1.1). If $(\boldsymbol{\theta}_t)$ is iid and $\mathbb{E}(\mathbf{\Lambda}_r)^u$ and $\mathbb{E} d\left(\mathbf{\Psi}_0(c), c\right)^u$ are finite for a certain $r \geq 1$ and $u > 0$, we can show that the marginal distribution of this solution

admits a moment of small order, as follows:

$$\text{there exists } s>0, \quad \text{such that } \mathbb{E}d(\boldsymbol{X},c)^s<\infty. \tag{1.10}$$

When the independence assumption is violated, we show by examples that the property (1.10) is not always true. In this context, the main result of this chapter provides a property on the trajectory of the stationary solution that is weaker than (1.10). The full statement of this result can be found in Theorem 3.2. It states that for any integer t,

$$1.\ \limsup_{n\to\infty}\frac{1}{n}d(\boldsymbol{X}_{t+n},c)=0 \quad \text{and} \quad 2.\ \limsup_{n\to\infty}\frac{1}{n}\ln d(\mathbf{X}_{t-n},c)=0 \ \ a.s. \tag{1.11}$$

This property can be interpreted as an exponential control of the trajectory. We explain why the property in equation (1.10) implies (1.11), but that the converse is not true. Section 3 is entirely devoted to the proof of this result. One of the distinctions made between 1. and 2. in (1.11) is that their proofs use different arguments. The proof of (1.11) uses new methods and relies in part on the subadditive ergodic theorem of Kingman (1973).

Section 4 of the chapter is devoted to the inference of the following GARCH(p,q) models:

$$\begin{aligned}&\boldsymbol{\epsilon}_t=\sqrt{\boldsymbol{h}_t}\boldsymbol{\eta}_t,\\&\boldsymbol{h}_t=\omega_0+\textstyle\sum_{i=1}^q\alpha_{0i}\boldsymbol{\epsilon}_{t-i}^2+\sum_{j=1}^p\beta_{0j}\boldsymbol{h}_{t-j},\quad \forall t\in\mathbb{Z}\end{aligned} \tag{1.12}$$

where $E\boldsymbol{\eta}_t^2=1$, $\omega_0>0,\alpha_{0i}\geqslant 0$ $(i=1,\ldots,q)$ and $\beta_{0j}\geqslant 0$ $(j=1,\ldots,p)$. By rewriting the model in a multivariate autoregressive form, we justify that Model (1.12) is a special case of (1.1), where $\boldsymbol{\theta}_t=\boldsymbol{\eta}_t$. If $(\boldsymbol{\eta}_t)$ is iid, we know by Berkes et al. (2003, Lemma 2.3) (see also (1.10)) that the stationary solution $(\boldsymbol{\epsilon}_t)$ of the (strong) GARCH admits a small order moment. The classical proof of the consistency and asymptotic normality of the Quasi-Maximum Likelihood Estimator (QMLE) of these models is based on this property (see Berkes et al. (2003) and Francq and Zakoian (2004)). Based on the work of Francq and Zakoian (2004), Escanciano (2009) relaxes the iid condition, but keeping the assumption of the existence of moment, and establishes the consistency of the QMLE of the semi-strong GARCH models, i.e. when $(\boldsymbol{\eta}_t)$ verifies:

$$\begin{aligned}&(\boldsymbol{\eta}_t) \text{ is stationary and ergodic,}\\&\boldsymbol{\eta}_t^2 \text{ is non-degenerate and } \mathbb{E}\left[\boldsymbol{\eta}_t^2\mid\mathcal{F}_{t-1}\right]=1 \text{ a.s.,}\end{aligned} \tag{1.13}$$

where $\mathcal{F}_{t-1}$ is the filtration generated by $(\boldsymbol{\epsilon}_{t-1},\boldsymbol{\epsilon}_{t-2},\ldots)$. This assumption is obviously weaker than the iid condition. An explicit construction of such a process which is not reduced to the iid case is proposed in the appendix 5.5 of the chapter. As noted above, we know that the existence of this moment is not guaranteed without the independence assumption and thus in particular under (1.13).

Given that $\boldsymbol{\epsilon}_1,\ldots,\boldsymbol{\epsilon}_n$ is a realisation (of length n) of the semi-strong GARCH process (1.12), for the value $\boldsymbol{\theta}_0=(\omega,\alpha_1,\ldots,\alpha_q,\beta_1,\ldots,\beta_p)^{\mathrm{T}}$ of the parameter belonging to a space of parameters $\boldsymbol{\Theta}\subset]0,+\infty[\times[0,\infty[^{p+q}$. Conditionally on the initial values $\boldsymbol{\epsilon}_0,\ldots,\boldsymbol{\epsilon}_{1-q},\tilde{\boldsymbol{\sigma}}_0^2,\ldots,\tilde{\boldsymbol{\sigma}}_{1-p}^2$, the conditional Gaussian quasi-likelihood is given by

$$L_n(\boldsymbol{\theta}) = L_n\left(\boldsymbol{\theta}; \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n\right) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\boldsymbol{\sigma}}_t^2}} \exp\left(-\frac{\boldsymbol{\epsilon}_t^2}{2\tilde{\boldsymbol{\sigma}}_t^2}\right),$$

where the $\tilde{\boldsymbol{\sigma}}_t^2$ are recursively defined , for $t \geqslant 1$, by $\tilde{\boldsymbol{\sigma}}_t^2 = \tilde{\boldsymbol{\sigma}}_t^2(\boldsymbol{\theta}) = \omega + \sum_{i=1}^{q} \alpha_i \boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{\boldsymbol{\sigma}}_{t-j}^2$. The QMLE of the parameter $\boldsymbol{\theta}_0$ is defined as any measurable solution $\hat{\boldsymbol{\theta}}_n$ of:

$$\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}) \tag{1.14}$$

Using the property (1.11), we relax the moment assumption made by Escanciano (2009) by establishing the strong consistency of the QMLE ($\hat{\boldsymbol{\theta}}_n$) of the semi-strong GARCH models, without moment condition on $\boldsymbol{\epsilon}_t$. This result is based on the classical assumptions considered in Francq and Zakoian (2004), as well as on the following additional condition,

$$\log^- \boldsymbol{\eta}_0 \text{ is integrable.} \tag{1.15}$$

This assumption is verified by most classical distributions. It prevents the marginal distribution of innovations from admitting too much mass around 0.

The proof of consistency follows the steps $i) - iv)$ of that of Theorem 2.1 of Francq and Zakoian (2004), with important modifications on the arguments. The property (1.11) is used in step $i)$ instead of the existence of a small order moment to show the asymptotic negligibility of the initial values in the definition of the quasi-likelihood. New techniques, as well as the condition (1.15), are used in steps $iii)$ and $iv)$ to circumvent the fact that the standard limit criterion based on the expectation of the criterion cannot be defined without the existence of a small order moment.

The asymptotic normality of the QMLE of semi-strong GARCHs remains difficult to establish due to the absence of moment. We discuss at the end of the chapter why this property might not hold.

Chapter 4: Inference on GARCH-MIDAS models without any small-order moment.

Recent developments in the financial econometric literature suggest that the dynamics of volatility may be better described by multi-component models. Engle and Lee (1999) introduced a long-run and short-run volatility model with additive components. In recent years, a class of models called GARCH-Mixed DAta Sampling (MIDAS), proposed by Engle et al. (2013), has gained popularity for its ability to capture the dependencies between daily stock returns and low-frequency (e.g. monthly, quarterly) explanatory variables. Instead of modelling conditional volatility as the sum of two components, these models specify volatility as the product of long- and short-term components. A GARCH component with unit variance (short term) fluctuates around a long term component with regular time variation that is a function of the explanatory variables, as in the model (1.16) and figure (2.2) below. For similar approaches, see also the work of Conrad and Engle (2021), Conrad and Schienle (2020) and their references.

The probabilistic structure and statistical properties of GARCH-MIDAS models have not received much attention. Preliminary results can be found in Wang and Ghysels

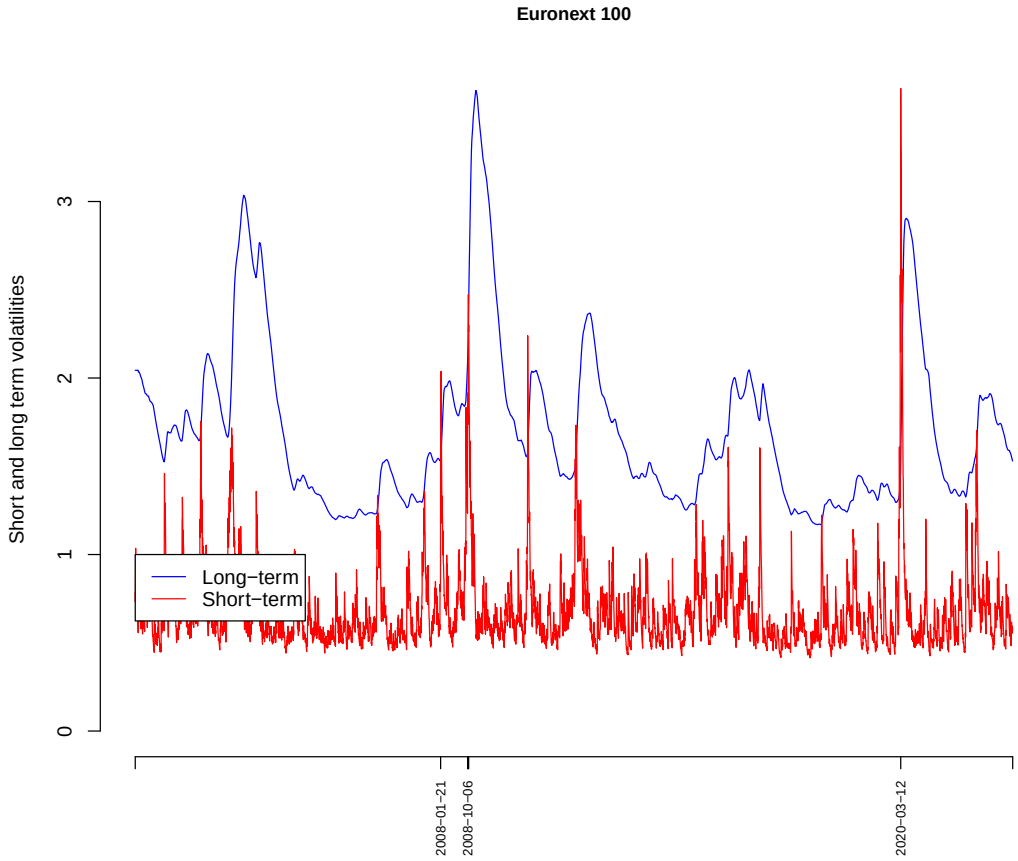

Figure 1.2 – Decomposition of short- and long-term volatilities for the Euronext 100 stock index from 31.12.1999 to 16.02.2023 using the GARCH-MIDAS model with $N = 22$ and $Q = 250$.

(2015). The aim of this chapter is to address some important econometric and statistical issues related to such models. In the chapter we consider the following two-factor model:

$$\begin{cases} r_t = \tau_t \epsilon_t, & \tau_t^2 = 1 + a_0 \sum_{i=1}^{Q} \varphi_i(\vartheta_0) RV_{t-i}, \\ \epsilon_t = \sigma_t \eta_t, & \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \end{cases} \tag{1.16}$$

where (η_t) is iid, $\mathbb{E}\eta_t^2 = 1$, $a_0 \geq 0$, $\omega_0 > 0$, $\alpha_0 \geq 0$, $\beta_0 \geq 0$, $RV_t = \sum_{i=0}^{N-1} r_{t-i}^2$ is a volatility performed on a sliding window, Q and N are positive integers and ϑ_0 is a real. The $\varphi_i(\cdot)$ are positive weights. Engle et al. (2013) suggests for example the Beta weights:

$$\varphi_i(\vartheta_0) = \frac{\{1 - i/(Q+1)\}^{\vartheta_0 - 1}}{\sum_{j=1}^{Q} \{1 - j/(Q+1)\}^{\vartheta_0 - 1}}, \quad \vartheta_0 \in (0, \infty) \tag{1.17}$$

The figure 2.2 above shows the decomposition of the volatility of the Euronext 100 returns from 1999 to 2023 using the GARCH-MIDAS Beta weighted model with $N = 22$ and $Q = 250$ and the parameter $(\omega_0, \alpha_0, \beta_0, a_0, \vartheta_0)$ estimated at

$(0.0285, 0.128, 0.815, 0.056, 2.932)$.

The parameterisation used in (1.16) is somewhat more general than that of Engle et al. (2013). Indeed, we do not require that the short-term volatility σ_t has a finite variance. The equation (1.16) can be rewritten in the following autoregressive form:

$$\boldsymbol{r}_t = \boldsymbol{A}_t \boldsymbol{r}_{t-1} + \boldsymbol{b}_t, \tag{1.18}$$

where $\boldsymbol{r}_t = (r_t^2, \ldots, r_{t-q+1}^2)'$, $\boldsymbol{b}_t = (\epsilon_t^2, \boldsymbol{0}'_{q-1})'$ and $\boldsymbol{A}_t = \boldsymbol{A}(\epsilon_t)$ is a companion type matrix. We can see that (1.18) is an SRE generated by the GARCH $(1,1)$ (non-iid) (ϵ_t).

Assuming that $a_0 \sup_{1 \leq i \leq q} \varphi_i(\vartheta_0) > 0$, we guarantee that the long-run volatility is not degenerate. Under this assumption and other regularity conditions, the first result of the chapter establishes that the model (1.16) admits a strictly stationary solution if and only if $\gamma(\boldsymbol{A}) < 0$. More surprisingly, we show that this stationary solution verifies:

$$\mathbb{E}|r_t|^s = \infty \quad \text{for all } s > 0. \tag{1.19}$$

These results show that the trajectories of returns that satisfy the model are not explosive but can be very erratic, unlike one-factor GARCH-type processes for which there are at least a small moment. This property leads to formidable difficulties for statistical inference. Wang and Ghysels (2015) studied the asymptotic distribution of the QMLE under the assumption that $\mathbb{E}|r_t|^s$ is finite for some $s > 0$. As noted above, this is a key assumption for showing the consistency and asymptotic normality of the QMLE of GARCH models. To our knowledge, the asymptotic properties of the QMLE have never been established without an assumption implying the existence of a small-order moment. However, in our framework, the property (1.19) shows that this moment condition does not hold. One of the main motivations of the chapter was to establish the asymptotic properties of GARCH-MIDAS models with the difficulty induced by (1.19). Under some relatively weak regularity and identifiability assumptions, we show the strong consistency of the QMLE $(\hat{\boldsymbol{\theta}}_n)$ and its asymptotic normality:

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)\boldsymbol{J}^{-1}),$$

where $\kappa_\eta := E\eta_t^4$ and $\boldsymbol{J} := E\left(\frac{1}{V_t^2(\boldsymbol{\theta}_0)} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0) \nabla'_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\right)$ is a positive-definite matrix. The proof of the consistency partly uses the control property (1.11), established in the chapter 3, and some arguments used in the proof of consistency of the QMLE of the semi-strong GARCH. The proof of asymptotic normality consists in establishing a sequence of properties, such as the existence and invertibility of the asymptotic covariance matrix $\boldsymbol{J}$, taking into account the absence of moments.

Chapter 4 also presents test procedures to detect the existence of a long-run volatility component, i.e. the null hypothesis $H_0 : a_0 = 0$. Because of the non-identifiability of the ϑ parameter under H_0, usual tests such as the Wald test may have non-standard asymptotic distributions. Therefore, we consider two approaches. In the first, we set ϑ to a value ϑ^*. In this approach, statistical tests with and without bootstrap are proposed. In the second approach, we estimated all parameters by QMLE, including ϑ, and estimated the critical value of the resulting Wald test by a bootstrap procedure based on the residuals.

The theoretical validity of the first method is established in Proposition 4.2 and Theorem 4.3.

In our simulations, we observed that the second method outperforms the first one, in which the power of the tests depends on the value of ϑ^*. We also found that the asymptotic distributions of the statistical tests of the first method without bootstrap are not well approximated in finite samples. Conrad and Schienle (2020) proposed a test of the model based on the score, not requiring a bootstrap. However, a moment assumption is made on the returns process and furthermore, the parameterisation used in their paper does not imply an identifiability issue.

The chapter ends with numerical applications: our simulations show that the lack of moments does not significantly affect the estimation of the parameters in a finite sample. In addition, the applications conducted on stock market indices NASDAQ, CAC40, DAX, and HSI, provide evidence of the existence of long-run volatility in these processes.

Chapter 5: On the growth rate of superadditive processes and the stability of functional GARCH models.

This chapter is divided into two parts. The main motivation of the first part is to provide an extension of the property of Kesten (1975) on random walks with stationary increments and to derive the contraction property of products of random operators. Kesten's result states that if $(\boldsymbol{z}_t)$ is a stationary process, then

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s. \text{ on } \{\boldsymbol{S}_n := \sum_{k=1}^{n} \boldsymbol{z}_k \to \infty \text{ lorsque } n \to \infty\}. \tag{1.20}$$

This result shows that the random walk cannot diverge to infinity at a rate smaller than n. The property (1.5) of Bougerol and Picard (1992b), established under additional ergodicity and integrability conditions, is a partial extension of Kesten's property to products of random matrices. The result of Kesten (1975) is often expressed using notions from ergodic theory. To share this expression, let us briefly recall some definitions and basic notions of ergodic theory we will use, as well as a brief review of the literature to contextualize.

A dynamical system is a quadruplet $(\Omega, \mathcal{B}, \mu, T)$, where $(\Omega, \mathcal{B}, \mu)$ is a probability space and $T : \Omega \to \Omega$ is a measurable application which preserves μ, that is, for any $A \in \mathcal{B}$, $\mu(T^{-1}(A)) = \mu(A)$. A set I in $\mathcal{B}$ is said to be invariant if $\mu\left(I \Delta T^{-1}(I)\right) = 0$ We say that T is ergodic if, for any invariant set I, $\mu(I) = 0$ or $\mu(I) = 1$..

A process $(\boldsymbol{S}_n)_{n\geq 1}$ is said to be superadditive if for all n, s, $\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s}$ $a.s.$ These processes were introduced by Hammersley and Welsh (1965) and one of the most important contributions to the asymptotic properties of these processes is the sub-additive ergodic theorem of Kingman (1973) which establishes the almost sure convergence of $(n^{-1}\boldsymbol{S}_n)_n$ under the integrability of $\boldsymbol{S}^-$. This result generalises the Birkhoff's well-known ergodic theorem on additive processes, i.e. when for all n, s, $\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n = \boldsymbol{S}_{n+s}$ $a.s.$ The result of Kesten (1975) says that if $(\boldsymbol{S}_n)_n$ is additive then

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s. \text{ on } \{\boldsymbol{S}_n \to \infty, n \to \infty\}. \tag{1.21}$$

This property holds even if $\boldsymbol{S}_1^-$ is not integrable. This result has found many applications in ergodic theory and was a precursor to the study of recurrence of random walks with stationary increments.

One of the main results of the chapter 5, established in Theorem 5.1, generalises the property (1.21) to superadditive processes. In this theorem, we show that if $(\boldsymbol{S}_n)_n$ is superadditive, then

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s. \text{ on } \liminf\{\boldsymbol{S}_n > 0\}. \tag{1.22}$$

This result is obtained under weaker conditions than those of Kesten. Indeed, it suffices to note that $\liminf\{\boldsymbol{S}_n > 0\}$ is the set of $\omega \in \Omega$ such that $(\boldsymbol{S}_n)$ is positive for n large enough, and thus includes the set where $(\boldsymbol{S}_n)$ diverges to infinity. Our proof relies on new techniques based on concepts from ergodic theory and differs from those used by Kesten and Bougerol-Picard. Several stronger versions of the property (1.22) are proposed through the theorems 5.2 and 5.3 as well as the corollaries 5.1 and 5.2. However, they require additional conditions. We also discuss the necessity and optimality of these assumptions throughout the chapter. These results state that the condition that the sequence $(\boldsymbol{S}_n)$ is strictly positive from a certain period can be weakened in some cases. In Theorem 5.2 we show that if E is an invariant subset of $\{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0\} \cap \liminf\{\boldsymbol{S}_n \geq 0\}$, then

$$\liminf_{n\to\infty} n^{-1}\boldsymbol{S}_n > 0 \text{ a.s. on } E. \tag{1.23}$$

Noting that Ω is invariant, a direct consequence of this result is that if, almost surely, the sequence $(\boldsymbol{S}_n)$ takes at least one strictly positive value and is non-negative from a certain time, then $\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s.$

Let $A \in \mathcal{B}$ et E' be the set of elements of A such that the sequence $(\boldsymbol{S}_n : \ T^n \in A)$ is positive from a certain period. Theorem 5.3 states that if $\boldsymbol{S}_1^-$ is integrable, then

$$\lim n^{-1}\boldsymbol{S}_n > 0 \quad a.s. \text{ on } E'.$$

This property shows that, in the integrability case, conditionally on A, the positivity condition involves only the values of $(\boldsymbol{S}_n)$ with indices in $(n : \ T^n \in A)$. We deduce from these results that if T is ergodic and $(\boldsymbol{\Lambda}_n)_{n\geq 1}$ is a positive submultiplicative process, i.e., $(-\ln \boldsymbol{\Lambda}_n)$ is a superadditive process, then for any $A \in \mathcal{B}$ such that $\mu(A) > 0$:

1. if $\mu(\liminf_n\{\boldsymbol{\Lambda}_n < 1\}) > 0$, then $\gamma := \limsup_n n^{-1}\log \boldsymbol{\Lambda}_n$ is almost surely constant in $\overline{\mathbb{R}}$ et $\gamma < 0 \quad a.s.$,
2. if $\mathbb{E}\log^+ \boldsymbol{\Lambda}_1$ is finite and, almost surely, the sequence $(\boldsymbol{\Lambda}_n : \ T^n \in A)$ is strictly less than 1 from a certain period, then

$$\gamma = \lim_n n^{-1}\log \boldsymbol{\Lambda}_n = \lim_n n^{-1}\mathbb{E}\log \boldsymbol{\Lambda}_n = \inf_n n^{-1}\mathbb{E}\log \boldsymbol{\Lambda}_n < 0 \quad a.s.$$

The statement of this result can be found in the corollary 5.3. Noting that $(\boldsymbol{\Lambda}_0^{(n)})$ is a special case of sub-multiplicative process, these properties generalise the property (1.5) to products of random operators and, more generally, to sub-multiplicative processes.

The second part of the chapter focuses on the study of the stationarity of SREs with positive coefficients in the space F of continuous functions with compact support. The model is as follows

$$\boldsymbol{h}_t = \delta(\boldsymbol{\eta}_{t-1}) + \gamma(\boldsymbol{\eta}_{t-1})\boldsymbol{h}_{t-1} = \boldsymbol{\Psi}_t\left(\boldsymbol{h}_{t-1}\right), \tag{1.24}$$

where $(\boldsymbol{\eta}_t)$ is a strictly stationary ergodic process in F. By denoting F_+ the set of positive functions of F, $\delta(\cdot)$ is a measurable function with values in F_+ and $\gamma(\cdot)$ take values in the space of continuous linear operators which maps the elements of F_+ to F_+. This class of models encompasses most conditional volatility models, including the continuous fGARCH.

The main result of the second part is Theorem 5.4. It establishes, under relatively weak assumptions, necessary and sufficient conditions for the existence of positive strictly stationary solutions to Equation (1.24) by relating the sign of the Lyapunov coefficient $(\boldsymbol{\Psi})$ of $(\boldsymbol{\Psi}_t)$ to this existence. The proof of this result uses the contraction properties established earlier and new arguments to overcome other technical difficulties, which we will discuss in the chapter.

In the corollary 5.5 we derive necessary and sufficient conditions for the existence of positive strictly stationary solutions for fGARCH models (1.7).

These results are, to our knowledge, the first to establish the necessity of (1.6) for the existence of stationary solutions for models of type (1.24) when the dimension of F is not finite. More discussion of this topic is given in the chapter.

Chapter 2

Introduction (Français)

Abstract

Cette thèse s'intéresse à l'étude des propriétés théoriques des équations récurrentes stochastiques et de leurs applications en économétrie financière. Nous nous intéressons à l'existence de solutions stationnaires, ainsi qu'aux propriétés statistiques et probabilistes de ces solutions. En particulier, nous allons étudier la stationnarité et l'inférence statistique des modèles Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) (p, q) semi-forts, dans lesquels le processus d'innovation n'est pas nécessairement indépendant. Nous nous intéresserons également aux modèles GARCH à deux facteurs de volatilité, un facteur de volatilité à long terme et un à court terme. Nous illustrerons nos résultats avec des simulations et des applications sur des données financières réelles. Enfin, nous nous intéresserons à la propriété de contraction des produits d'opérateurs aléatoires, avec une application à l'étude de la stationnarité des modèles de volatilité conditionnelle dans l'espace des fonctions continues (fGARCH).

Contents

1 Contexte 14
2 Synthèse des principaux résultats 18

1 Contexte

Au cours des dernières années, les processus autorégressifs de la forme

$$\boldsymbol{X}_t = \Psi\left(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1}\right) = \boldsymbol{\Psi}_t\left(\boldsymbol{X}_{t-1}\right), \quad t \in \mathbb{Z}, \tag{2.1}$$

où $\Psi(\theta, \cdot)$ est un opérateur d'un espace polonais $(F,\ d)$ dans lui même et $(\boldsymbol{\theta}_t)$ est un processus strictement stationnaire et ergodique, ont attiré beaucoup d'attention. Cette équation, communément appelé Iterated Function Systems (IFS) regroupe une grande variété de modèles. Parmis eux, on peut citer les Stochastic Recurrence Equation (SRE) introduits par Kesten (1973), dans lesquels $F = \mathbb{R}^d$, $\boldsymbol{\theta}_t = (\boldsymbol{A}_t, \boldsymbol{B}_t)$ où $\boldsymbol{A}_t$ est une matrice $d \times d$ et $\boldsymbol{B}_t$ est un vecteur $d \times 1$, et $\Psi(\theta, \cdot)$ est l'opérateur affine tel que

$$\boldsymbol{X}_t = \boldsymbol{A}_t \boldsymbol{X}_{t-1} + \boldsymbol{B}_t, \quad t \in \mathbb{Z}. \tag{2.2}$$

Les modèles SREs incluent notamment les processus de type AutoregRessive–Moving-Average (ARMA) et de type GARCH, voir Eq. (2.13). Une classe de SREs bivariés est également souvent utilisée pour dessiner des images de fractales en deux dimensions, telles que la fougère de Barnsley, voir figure 2.1, le flocon de Koch ou la courbe du dragon. La procédure de construction est la suivante: considérons une séquence finie de transformations affines de $\mathbb{R}^2$, c'est-à-dire $(\mathbf{a}_i, \mathbf{b}_i), i = 1, \ldots, k$, pour certains $k \geq 1$, où $(\mathbf{a}_i)$ sont des matrices 2×2 et $(\mathbf{b}_i)$ sont des vecteurs dans $\mathbb{R}^2$. Supposons que les $\mathbf{a}_i$ sont des matrices contractantes, c'est-à-dire que leurs rayons spectraux respectifs ρ_i sont strictement inférieurs à 1. En considérant des poids positifs $p_1, \ldots, p_k$, avec $p_1 + \cdots + p_k = 1$, on peut définir la chaîne de Markov $(\boldsymbol{X}_t)$ se déplaçant dans $\mathbb{R}^2$: à partir de x, la chaîne procède en choisissant i au hasard avec une probabilité de p_i et en se déplaçant vers $\boldsymbol{X}_{t+1} = a_i \boldsymbol{X}_t + b_i$. L'image fractale est obtenue en représentant le nuage de points $(\boldsymbol{X}_t)$ dans le plan à partir de d'un n arbitrairement grand afin de réduire l'effet l'initialisation pour s'approcher de la distribution stationnaire. Pour plus de détails sur les SREs fractales, voir Barnsley (2014). La classe des modèles AutoRégressifs (AR) avec un bruit suivant un modèle ARCH introduite par Weiss (1984) constitue un exemple de processus vérifiant (2.1) et qui ne sont pas des SREs. Un cas particulier de ces processus est le Modèle AR(1) avec bruits ARCH(1) défini par :

$$\boldsymbol{X}_t = \alpha \boldsymbol{X}_{t-1} + \sqrt{\beta + \lambda \boldsymbol{X}_{t-1}^2}\boldsymbol{\eta}_t, \quad t \in \mathbb{Z}, \tag{2.3}$$

où $(\boldsymbol{\eta}_t)$ sont des variables aléatoires indépendantes et identiquement distribuées (iid), $\alpha \in \mathbb{R}$, $\beta > 0$ et $\lambda > 0$. Le modèle (2.3) a été étudié dans Borkovec and Klüppelberg (2001) et Borkovec (2000). D'autres exemples d'applications des IFS peuvent être trouvés dans Buraczewski et al. (2016, Examples 1.0.1 à 1.0.5) et dans Diaconis and Freedman (1999).

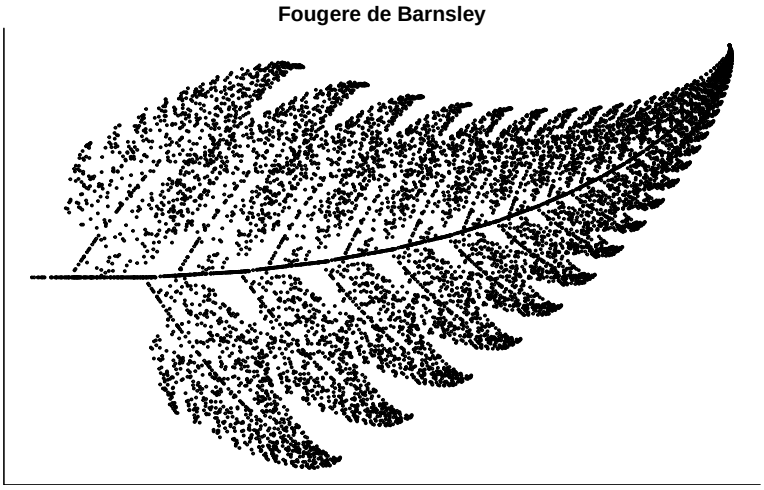

Figure 2.1 – Fougere de Barnsley obtenue avec 10 000 points avec les paramètres (voir Barnsley (2014, table 3.8.3 and figure 3.8.3)):

$$
\begin{aligned}
(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4) &= \left(\begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.16 \end{pmatrix}, \begin{pmatrix} 0.85 & 0.04 \\ -0.04 & 0.85 \end{pmatrix}, \begin{pmatrix} 0.20 & -0.26 \\ 0.23 & 0.22 \end{pmatrix}, \begin{pmatrix} -0.15 & 0.28 \\ 0.26 & 0.24 \end{pmatrix} \right) \\
(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4) &= \left(\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 1.6 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 1.6 \end{pmatrix}, \begin{pmatrix} 0.0 \\ 0.44 \end{pmatrix} \right), \text{ et} \\
(p_1, p_2, p_3, p_4) &= (0.01, 0.85, 0.07, 0.07)
\end{aligned}
$$

1.1 Stationnarité

En raison de leur utilisation généralisée, les propriétés théorique des modèles (2.1) a également suscité beaucoup d'intérêt au cours des 50 dernières années. Dans le cas univarié des SREs (2.2) avec des innovations $\boldsymbol{\theta}_t = (\boldsymbol{A}_t, \boldsymbol{B}_t) \in \mathbb{R}^2$ iid, l'existence de solutions stationnaires a été étudiée, entre autres, par Kesten (1973), Vervaat (1979), Goldie (1991) et Goldie and Maller (2000). Ces derniers fournissent une caractérisation complète du problème en proposant un ensemble de conditions nécessaires et suffisantes. Dans le cas iid multivarié, les conditions les plus faibles sur l'existence de solution stationnaire sont dues à Bougerol and Picard (1992b). Sous une condition d'irréductibilité du modèle, ils montrent que si $\mathbb{E} \log^+ \boldsymbol{A}_0$ et $\mathbb{E} \log^+ \boldsymbol{B}_0$ sont finies, alors l'equation (2.2) admet une solution strictement stationnaire et non anticipative $(\boldsymbol{X}_t)$, c'est-à-dire que $\boldsymbol{X}_t$ est $\sigma(\boldsymbol{\theta}_k,\ k \leq t)$-measurable, si et seulement si

$$
\boldsymbol{\gamma} := \lim_{n \to +\infty} \frac{1}{n} \log(\boldsymbol{\Lambda}_0^{(n)}) = \inf_{n \geq 1} \frac{1}{n} \mathbb{E} \log(\|\boldsymbol{A}_n \boldsymbol{A}_{n-1} \cdots \boldsymbol{A}_1\|) < 0, \tag{2.4}
$$

où $\boldsymbol{\Lambda}_t^{(0)} = 1$ et $\boldsymbol{\Lambda}_t^{(k)} = \|\boldsymbol{A}_t \boldsymbol{A}_{t-1} \cdots \boldsymbol{A}_{t-k+1}\|$ pour tout entier $k > 0$. La preuve de ce résultat repose en partie sur une autre propriété établie dans le même article. Communément nommée **propriété de contraction de produit de matrices aléatoires**, cette propriété affirme que si $(\boldsymbol{A}_t)$ est strictement stationnaire et ergodique, et que $\mathbb{E} \log^+ \boldsymbol{A}_0$

est finie, alors

$$\boldsymbol{\gamma}(\boldsymbol{A}) < 0 \ \ p.s. \quad \text{si} \quad \boldsymbol{\Lambda}_0^{(n)} \to 0 \ \ p.s. \text{ lorsque } \ n \to \infty. \tag{2.5}$$

Cette propriété, est couramment utilisée pour établir la nécessité de (2.4) pour l'existence de solutions stationnaires de modèles de type GARCH.

Pour le modèle général (2.1), une condition suffisante d'existence de solutions stationnaires a été établie par Elton (1990, Théorème 3) (voir aussi Bougerol (1993, Théorème 3.1)). Comme dans le modèle affine en dimension finie, ce résultat lie le signe du coefficient de Lyapunov $\boldsymbol{\gamma}(\boldsymbol{\Psi})$, défini ci-dessous, du processus $(\boldsymbol{\Psi}_t)$ à l'existence de solutions stationnaires. La contraction de $(\boldsymbol{\Psi}_t)$ est obtenue quand

$$\boldsymbol{\gamma}(\boldsymbol{\Psi}) := \limsup_{n\to+\infty} \frac{1}{n} \log \boldsymbol{\Lambda}_0^{(n)} < 0, \tag{2.6}$$

où $\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t) = \sup_{x_1,x_2\in F, x_1\neq x_2} \frac{d(\boldsymbol{\Psi}_t(x_1),\boldsymbol{\Psi}_t(x_2))}{d(x_1,x_2)}$ avec $\boldsymbol{\Lambda}_t^{(0)} = 1$ et $\boldsymbol{\Lambda}_t^{(k)} = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t \circ \cdots \circ \boldsymbol{\Psi}_{t-k+1})$ pour tout entier $k > 0$. Pour le modèle (2.3), on peut noter que la condition d'existence d'une solution stationnaire proposée par Borkovec and Klüppelberg (2001, Eq. 3.2) implique (2.6). Cependant, à notre connaissance, il n'existe pas de résultat établissant la nécessité de l'équation (2.6) dans le contexte général des IFS (non affines) ou pour des opérateurs affines $\boldsymbol{\Psi}_t$ en dimension infinie. Les modèles ARCH/GARCH fonctionnels dans lesquels le processus de volatilité est une fonction constituent une nouvelle classe de modèles dans lesquels $\boldsymbol{\Psi}_t$ est un opérateur affine dans un espace L^p ou dans celui des fonctions continues. Un processus GARCH fonctionnel d'ordres $(1,1)$ (fGARCH$(1,1)$) dans $\mathcal{C}[0,1]$ est une séquence $(\boldsymbol{r}_t : t \in \mathbb{Z})$ d'éléments aléatoires où $\boldsymbol{r}_t$ est une fonction, $(\boldsymbol{r}_t(u) : u \in [0,1])$, continue sur $[0,1]$ satisfaisant:

$$\boldsymbol{r}_t = \boldsymbol{\sigma}_t\boldsymbol{\eta}_t, \ \boldsymbol{\sigma}_t^2 \ = \delta + \alpha \boldsymbol{r}_{t-1}^2 + \beta\boldsymbol{\sigma}_{t-1}^2 = \delta + \int \boldsymbol{\gamma}_t(\cdot,s)\boldsymbol{\sigma}_{t-1}^2(s)ds = \delta + \boldsymbol{\gamma}_t\boldsymbol{\sigma}_{t-1}, \tag{2.7}$$

où $(\boldsymbol{\eta}_t : t \in \mathbb{Z})$ est une séquence iid de fonctions aléatoires dans $\mathcal{C}[0,1]$, δ est une fonction positive et les opérateurs intégraux α et β, c'est-à-dire $(\alpha x)(u) = \int \alpha(u,s)x(s)ds$ et $(\beta x)(u) = \int \beta(u,s)x(s)ds$, sont positifs, c'est-à-dire qu'ils font correspondre des fonctions non négatives à des fonctions non négatives. $\boldsymbol{\gamma}_t(u,s) = \alpha(u,s)\boldsymbol{\eta}_{t-1}^2(s) + \beta(u,s)$ est un élément de $\mathcal{C}[0,1]^2$. On peut en effet remarquer que $(\boldsymbol{\sigma}_t^2)$ définie dans (2.7) vérifie (2.1) avec $F = \mathcal{C}[0,1]$, $\boldsymbol{\theta}_t = \boldsymbol{\eta}_t$ et $\boldsymbol{\Psi}_t = \delta + \boldsymbol{\gamma}_t(\cdot)$. Les fGARCH ont été introduit par Hörmann et al. (2013) et Aue et al. (2017) pour modéliser la dynamique de la courbe de volatilité conditionnelle en s'inspirant la structure des modèles GARCH classiques. Des conditions suffisantes d'existence de solutions stationnaires pour différentes classes de ces modèles ont été établies dans Aue et al. (2017), Hörmann et al. (2013), Kühnert (2020) et dans Cerovecki et al. (2019). Ce dernier article considère le fGARCH(p,q) dans l'espace $L^2([0,1])$. Dans ce contexte, les auteurs remarquent qu'un des principaux défis pour établir la nécessité de (2.6) est d'étendre la propriété de contraction des matrices aléatoires (2.5) aux processus d'opérateurs $(\boldsymbol{\Psi}_t)$. À notre connaissance, cette extension n'a jusqu'à présent pas été établie.

1.2 Queue de distribution

Une propriété importante de la distribution marginale de la solution stationnaire $(\boldsymbol{X}_t)$ de l'équation (2.2) nous vient de Kesten (1973). Sous des hypothèses relativement faibles, il établit dans le cas iid du modèle univarié que la queue de cette distribution est une loi à queue de puissance, dans le sens suivant :

$$\mathbb{P}(|\boldsymbol{X}| > x) \sim cx^{-\alpha}, \quad x \to \infty, \tag{2.8}$$

où c et α sont des constantes strictement positives. Ce résultat implique en particulier que $\boldsymbol{X}$ a des moments d'ordre $s < \alpha$. Les travaux de Goldie (1991) démontrent que la propriété (2.8) peut être généralisée aux cas iid du modèle (2.1), sous certaines hypothèses de régularité sur $\boldsymbol{\Psi}_t$ qui ne sont pas nécessairement limitées au cas affine (2.2). Les travaux de Borkovec and Klüppelberg (2001) sur le modèle (2.3) fournissent également des résultats similaires. Plus généralement, on peut montrer dans le cas iid, sous (2.6) et quelques hypothèses de régularité sur $\boldsymbol{\Psi}_0$, que la solution stationnaire de (2.1) admet un moment de petit ordre. Cette propriété de moment est souvent utilisée dans l'inférence statistique des modèles (2.1) tels que ceux de type GARCH à innovation iid, voir Berkes et al. (2003) ou Francq and Zakoian (2004) pour le GARCH(p, q) classique. Cependant, lorsque $(\boldsymbol{\theta}_t)$ est stationnaire et ergodique, mais non indépendante, nous démontrerons sur une classe de modèles GARCH à deux facteurs de volatilité que la solution strictement stationnaire peut n'admettre aucun moment.

1.3 Motivation

Comme nous venons de le remarquer, les propriétés théoriques du modèle (2.1) restent peu connues lorsque $(\boldsymbol{\theta}_t)$ est non-iid ou lorsque le modèle affine (2.2) est considéré en dimension infinie. Plusieurs classes de modèles de volatilité conditionnelle appartenant à (2.1) s'insèrent dans ces deux catégories. On peut citer les fGARCH et les modèles GARCH semi-forts introduits par Lee and Hansen (1994) et dans lesquels le processus des innovations suit un processus à différence martingale non-iid. Les modèles GARCH à deux facteurs de volatilité (GARCH-MIDAS) de Engle et al. (2013) constituent une autre classe de modèles IFS qui ne sont pas engendrés par une séquence iid. Un autre exemple est donné par les modèles GARCH-X qui sont des IFS pilotés par une séquence d'innovations et de covariables généralement non iid. Un exemple simple de GARCH-X est donnée par le processus de la forme

$$\boldsymbol{\epsilon}_t = \boldsymbol{\sigma}_t \boldsymbol{\eta}_t \quad \boldsymbol{\sigma}_t^2 = \omega + \alpha \boldsymbol{\epsilon}_{t-1}^2 + \beta \boldsymbol{\sigma}_{t-1}^2 + \pi' \boldsymbol{x}_{t-1} = \omega + \pi' \boldsymbol{x}_{t-1} + (\alpha \boldsymbol{\eta}_{t-1}^2 + \beta) \boldsymbol{\sigma}_{t-1}^2, \tag{2.9}$$

où $\boldsymbol{x}_t = (\boldsymbol{x}_{1,t}, \ldots, \boldsymbol{x}_{r,t})$ est un vecteur de r covariables exogènes telles que le volume quotidien des transactions, des données macro-économiques ou d'autres données de marchées. Dans ces modèles le processus de volatilité dépends d'un processus de covariables en plus des carrés des rendement passés. Cette information supplémentaire permet d'avoir de meilleurs prévisions du carré des rendements. Ces différents exemples motivent l'étude des équations SRE affines en dimension infinie ou des IFS conduites par des innovations non iid.

2 Synthèse des principaux résultats

Chapitre 3: Contrôle exponentiel de la trajectoire des IFS et application aux modèles GARCH(p,q) semi-forts.

Nous savons, d'après Elton (1990), que s'il existe un élément c dans F telle que:

$$\mathbb{E}\ln^+ d\left(\mathbf{\Lambda}_0(c), c\right) \text{ et } \mathbb{E}\ln^+ \mathbf{\Lambda}_0 \text{ sont finis,} \tag{2.10}$$

alors la condition (2.6) est suffisante pour garantir l'existence d'une solution stationnaire $(\boldsymbol{X}_t)$ à l'équation (2.1). Si $(\boldsymbol{\theta}_t)$ est iid et que $\mathbb{E}(\mathbf{\Lambda}_r)^u$ et $\mathbb{E}d\left(\mathbf{\Psi}_0(c), c\right)^u$ sont finis pour un certain $r \geq 1$ et $u > 0$, on peut montrer que la distribution marginale de cette solution admet un moment de petit ordre, de la manière suivante:

$$\text{il existe } s > 0, \quad \text{tel que } \mathbb{E}d(\boldsymbol{X}, c)^s < \infty. \tag{2.11}$$

Dans le cas où l'hypothèse d'indépendance n'est pas vérifiée, nous montrons par des exemples que la propriété (2.11) n'est pas toujours vraie. Dans ce contexte, le résultat principal de ce chapitre propose une propriété sur la trajectoire de la solution stationnaire, plus faible que (2.11). L'énoncé complet de ce résultat se trouve dans le théorème 3.2. Il affirme que pour tout entier relatif t,

$$1. \quad \limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c) = 0 \quad \text{et} \quad 2. \quad \limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t-n}, c) = 0 \quad p.s. \tag{2.12}$$

Cette propriété peut être interprétée comme un contrôle exponentiel de la trajectoire. Nous avons donné des justifications sur le fait que le résultat en (2.11) implique (2.12) mais que la réciproque n'est pas toujours vraie. La section 3 est entièrement consacrée à la preuve de ce résultat. Une des distinctions entre les points 1. et 2. vient du fait que leurs preuves utilisent des arguments différents. La preuve de ce résultat utilise des méthodes nouvelles et repose en partie sur le théorème sous-additif ergodique de Kingman (1973).

La section 4 du chapitre est dédiée à l'inférence des modèles GARCH(p, q) semi-forts suivants:

$$\begin{aligned} &\boldsymbol{\epsilon}_t = \sqrt{\boldsymbol{h}_t}\boldsymbol{\eta}_t, \\ &\boldsymbol{h}_t = \omega_0 + \textstyle\sum_{i=1}^{q} \alpha_{0i}\boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j}\boldsymbol{h}_{t-j}, \quad \forall t \in \mathbb{Z} \end{aligned} \tag{2.13}$$

où $E\boldsymbol{\eta}_t^2 = 1$, $\omega_0 > 0, \alpha_{0i} \geqslant 0$ $(i = 1, \ldots, q)$ et $\beta_{0j} \geqslant 0$ $(j = 1, \ldots, p)$. Le modèle (2.13) est un cas particulier de (2.1) où $\boldsymbol{\theta}_t = \boldsymbol{\eta}_t$. Lorsque $(\boldsymbol{\eta}_t)$ est iid, nous savons par Berkes et al. (2003, Lemma 2.3) (voir aussi (2.11)) que la solution stationnaire $(\boldsymbol{\epsilon}_t)$ du GARCH fort admet un moment de petit ordre. La preuve classique de la consistance et de la normalité asymptotique du Quasi-Maximum de Vraisemblance (QMV) de ces modèles repose sur cette propriété (voir Berkes et al. (2003) et Francq and Zakoian (2004)). En se basant sur les travaux de Francq and Zakoian (2004), Escanciano (2009) relâche l'hypothèse d'indépendance, mais en faisant l'hypothèse de l'existence de moment, et établit la consistance du QMV des modèles GARCH semi-fort, c'est à dire quand $(\boldsymbol{\eta}_t)$

vérifie:

$$\begin{array}{l}(\boldsymbol{\eta}_t) \text{ est stationnaire et ergodique,} \\ \boldsymbol{\eta}_t^2 \text{ est non dégénérée et } \mathbb{E}\left[\boldsymbol{\eta}_t^2 \mid \mathcal{F}_{t-1}\right] = 1 \text{ p.s.,}\end{array} \tag{2.14}$$

où $\mathcal{F}_{t-1}$ est la filtration engendrée par $(\boldsymbol{\epsilon}_{t-1}, \boldsymbol{\epsilon}_{t-2}, \ldots)$. Cette hypothèse est bien évidemment plus faible que la condition d'iid. Une construction explicite d'un tel processus qui n'est pas réduit au cas iid est proposée dans l'Annexe 5.5 du chapitre. Comme nous l'avons dit plus haut, nous savons que l'existence de ce moment n'est pas garantie sans l'hypothèse d'indépendance et donc en particulier sous (2.14).

En considérant que $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ constituent une réalisation (de longueur n) du processus GARCH semi-fort (2.13), pour la valeur $\boldsymbol{\theta}_0 = (\omega, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)^{\mathrm{T}}$ du paramètre appartant à un espace de paramètres $\boldsymbol{\Theta} \subset]0, +\infty[\times [0, \infty[^{p+q}$. Conditionnellement aux valeurs initiales $\boldsymbol{\epsilon}_0, \ldots, \boldsymbol{\epsilon}_{1-q}$, $\tilde{\boldsymbol{\sigma}}_0^2, \ldots, \tilde{\boldsymbol{\sigma}}_{1-p}^2$, la quasi-vraisemblance gaussienne conditionnelle est donnée par

$$L_n(\boldsymbol{\theta}) = L_n\left(\boldsymbol{\theta}; \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n\right) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\boldsymbol{\sigma}}_t^2}} \exp\left(-\frac{\boldsymbol{\epsilon}_t^2}{2\tilde{\boldsymbol{\sigma}}_t^2}\right),$$

où les $\tilde{\boldsymbol{\sigma}}_t^2$ sont récursivement définies, pour $t \geqslant 1$, par $\tilde{\boldsymbol{\sigma}}_t^2 = \tilde{\boldsymbol{\sigma}}_t^2(\boldsymbol{\theta}) = \omega + \sum_{i=1}^{q} \alpha_i \boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{\boldsymbol{\sigma}}_{t-j}^2$. Le QMV du paramètre $\boldsymbol{\theta}_0$ est défini comme toute solution mesurable $\hat{\boldsymbol{\theta}}_n$ de:

$$\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L_n(\boldsymbol{\theta}) \tag{2.15}$$

En utilisant la propriété (2.12), nous relâchons l'hypothèse de moment faite par Escanciano (2009) en établissant la consistance forte du QMV $(\hat{\boldsymbol{\theta}}_n)$ des modèles GARCH semi-forts, sans condition de moment sur $\boldsymbol{\epsilon}_t$. Ce résultat est basé sur les hypothèses classiques considérées dans Francq and Zakoian (2004), ainsi que sur la condition supplémentaire suivante,

$$\log^- \boldsymbol{\eta}_0 \text{ est intégrable.} \tag{2.16}$$

Cette hypothèse est vérifiée par la plupart des distributions classiques. Elle empêche la distribution marginale des innovations d'avoir trop de masse autour de 0.

La preuve de la consistance suit les étapes $i) - iv)$ de celle du théorème 2.1 de Francq and Zakoian (2004), avec des modifications importantes sur les arguments utilisés. La propriété (2.12) est utilisée dans l'étape $i)$ à la place de l'existence d'un moment de petit ordre pour montrer la négligeabilité asymptotique des valeurs initiales dans la définition de la quasi-vraisemblance. De nouvelles techniques, ainsi que la condition (2.16), sont utilisées dans les étapes $iii)$ et $iv)$ pour contourner le fait que le critère standard limite basé sur l'espérance du critère ne peut pas être défini sans l'existence d'un moment de petit ordre.

La normalité asymptotique du QMV des GARCH semi-forts reste difficile à établir en raison de l'absence de moment. Nous discutons à la fin du chapitre des raisons pour lesquelles cette propriété pourrait ne pas être vérifiée.

Chapitre 4: Inférence des modèles GARCH-MIDAS sans moment de petit

ordre.

Les récents développements dans la littérature en économétrie financière suggèrent que la dynamique de la volatilité pourrait être mieux décrite par des modèles à plusieurs composantes. Engle and Lee (1999) ont introduit un modèle de volatilité à composantes additives, à long et à court terme. Ces dernières années, une classe de modèles appelée GARCH-Mixed Data Sampling (MIDAS), proposée par Engle et al. (2013), a gagné en popularité pour sa capacité à capter les dépendances entre les rendements boursiers quotidiens et les variables explicatives à basse fréquence (par exemple, mensuelles, trimestrielles). Au lieu de modéliser la volatilité conditionnelle comme la somme de deux composantes, ces modèles spécifient la volatilité comme le produit de composantes à long et à court terme. Une composante GARCH à variance unitaire (à court terme) fluctue autour d'une composante à long terme à variation temporelle régulière qui est fonction de variables explicatives, comme par exemple dans le modèle (2.17) et la figure 2.2 ci-dessous. Pour des approches similaires, voir également les travaux de Conrad and Engle (2021), Conrad and Schienle (2020) et leurs références.

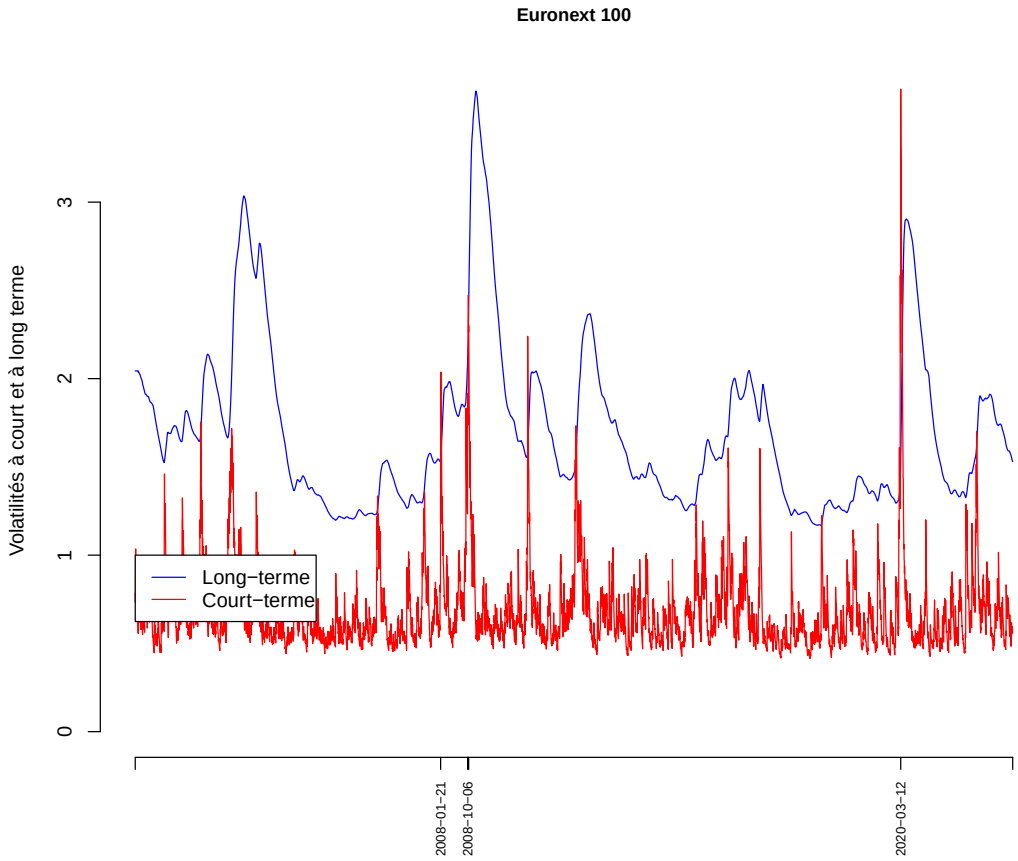

Figure 2.2 – Décomposition des volatilités à court et long terme pour l'indice boursier Euronext 100 de 1999-12-31 à 2023-02-16 avec le modèle GARCH-MIDAS avec $N = 22$ et $Q = 250$.

La structure probabiliste, ainsi que les propriétés statistiques des modèles GARCH-MIDAS, n'ont pas reçu beaucoup d'attention jusqu'à présent. Des résultats préliminaires peuvent cependant être trouvés dans Wang and Ghysels (2015). L'objectif du chapitre 4 est d'aborder certaines questions économétriques et statistiques importantes liées à de tels modèles. Dans le chapitre, nous considérons le modèle à deux facteurs suivant :

$$\begin{cases} r_t = \tau_t \epsilon_t, & \tau_t^2 = 1 + a_0 \sum_{i=1}^{Q} \varphi_i(\boldsymbol{\vartheta}_0) RV_{t-i}, \\ \epsilon_t = \sigma_t \eta_t, & \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \end{cases} \tag{2.17}$$

où (η_t) est iid avec $\mathbb{E}\eta_t^2 = 1$, $a_0 \geq 0$, $\omega_0 > 0$, $\alpha_0 \geq 0$ et $\beta_0 \geq 0$, $RV_t = \sum_{i=0}^{N-1} r_{t-i}^2$ est une volatilité réalisée sur une fenêtre glissante, Q et N sont des entiers positifs, $\boldsymbol{\vartheta}_0$ est un réel et les $\varphi_i(\cdot)$ sont des pondérations positives. Engle et al. (2013) suggère par exemples les pondérations Bêta:

$$\varphi_i(\vartheta_0) = \frac{\{1 - i/(Q+1)\}^{\vartheta_0 - 1}}{\sum_{j=1}^{Q} \{1 - j/(Q+1)\}^{\vartheta_0 - 1}}, \quad \vartheta_0 \in (0, \infty) \tag{2.18}$$

La figure 2.2 ci-dessous montre la décomposition de la volatilité du Euronext 100 de 1999 à 2023 avec le modèle GARCH-MIDAS à pondération Béta avec $N = 22$ et $Q = 250$ et du paramètre $(\omega_0, \alpha_0, \beta_0, a_0, \boldsymbol{\vartheta}_0)$ estimés à $(0.0285, 0.128, 0.815, 0.056, 2.932)$.

La paramétrisation utilisée est légèrement plus générale que celle de Engle et al. (2013). En effet, nous n'imposons pas nécessairement que la volatilité à court terme σ_t soit à variance finie. Nous nous sommes intéressés en premier lieu à la stationnarité du modèle. l'équation (2.17) peut être réécrite sous la forme autorégressive suivante :

$$\boldsymbol{r}_t = \boldsymbol{A}_t \boldsymbol{r}_{t-1} + \boldsymbol{b}_t, \tag{2.19}$$

où $\boldsymbol{r}_t = (r_t^2, \ldots, r_{t-q+1}^2)'$, $\boldsymbol{b}_t = (\epsilon_t^2, \boldsymbol{0}_{q-1}')'$ et $\boldsymbol{A}_t = \boldsymbol{A}(\epsilon_t)$ est une matrice de type compagnon. On peut remarquer que (2.19) est une SRE engendrée par le GARCH $(1,1)$ (non-iid) (ϵ_t).

En supposant que $a_0 \sup_{1 \leq i \leq q} \varphi_i(\boldsymbol{\vartheta}_0) > 0$, on garantit que la volatilité à long terme n'est pas dégénérée. Sous cette hypothèse, et avec d'autres conditions de régularité, le premier résultat du chapitre établit que le modèle (2.17) admet une solution strictement stationnaire si et seulement si $\gamma(\boldsymbol{A}) < 0$. De manière plus surprenante, nous montrons que cette solution stationnaire vérifie:

$$\mathbb{E}|r_t|^s = \infty \quad \text{pour tout } s > 0. \tag{2.20}$$

Ces résultats montrent que les trajectoires des rendements satisfaisants le modèle ne sont pas explosives mais peuvent être très erratiques, contrairement aux processus de type GARCH pour lesquels au moins de petits moments existent. Cette propriété entraîne de formidables difficultés pour l'inférence statistique. Wang and Ghysels (2015) ont étudié la distribution asymptotique du QMV du modèle sous l'hypothèse que $\mathbb{E}|r_t|^s$ est finie pour un certain $s > 0$. Comme nous l'avons remarqué plus en haut, c'est une hypothèse clé pour montrer la consistance et la normalité asymptotique du QMV des modèles GARCH. À notre connaissance, les propriétés asymptotiques du QMV n'ont jamais été établies

sans une hypothèse qui implique l'existence de moment. Dans notre cadre, la propriété (2.20) montre cependant que cette condition de moment n'est pas vérifiée. Une des principales motivations du chapitre 4 a été d'établir les propriétés asymptotiques des modèles GARCH-MIDAS avec la difficulté induite par (2.20). Sous quelques hypothèses, relativement faibles, de régularité et d'identifiabilité, nous montrons la consistance forte du QMV $(\hat{\boldsymbol{\theta}}_n)$ ainsi sa normalité asymptotique:

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)\boldsymbol{J}^{-1}),$$

où $\kappa_\eta := E\eta_t^4$ et $\boldsymbol{J} := E\left(\frac{1}{V_t^2(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\right)$ est une matrice définie-positive. La preuve de la consistance utilise en partie la propriété de contrôle (2.12), établie dans le chapitre 3, et quelques arguments utilisés dans la preuve de la consistance du QMV des GARCH semi-forts. La preuve de la normalité asymptotique consiste à établir une suite de propriétés, comme l'existence et l'inversibilité de la matrice de covariance asymptotique $\boldsymbol{J}$, en tenant compte de l'absence de moments.

Le chapitre 4 présente également des procédures de test pour détecter l'existence d'une composante de volatilité à long terme, c'est-à-dire l'hypothèse nulle $H_0 : a_0 = 0$. À cause de la non-identifiabilité du paramètre $\boldsymbol{\vartheta}$ sous H_0, les tests habituels tels que le test de Wald peuvent avoir des distributions asymptotiques non standard. Ainsi, nous avons envisagé deux approches. Dans la première, nous avons fixé $\boldsymbol{\vartheta}$ à une valeur $\boldsymbol{\vartheta}^*$. Dans cette approche, des statistiques de test avec et sans bootstrap sont proposées. Dans la seconde, nous avons estimé tous les paramètres par QMV, y compris $\boldsymbol{\vartheta}$, et avons estimé la valeur critique du test de Wald résultant par une procédure de bootstrap basée sur les résidus. La validité théorique de la première méthode est établie dans la proposition 4.2 et le théorème 4.3.

Durant nos simulations, nous avons observé que la seconde méthode sur-performe la première, dans laquelle la puissance des tests dépend de la valeur de $\boldsymbol{\vartheta}^*$. Nous avons aussi remarqué que les distributions asymptotiques des statistiques de test de la première méthode sans bootstrap ne n'est pas bien approximées à échantillon fini. Conrad and Schienle (2020) ont proposé un test du modèle basé sur le score, ne nécessitant pas de bootstrap. Cependant, une hypothèse de moment est faite sur le processus de rendement et de plus, la paramétrisation utilisée dans leur article n'implique pas de problème d'identifiabilité.

Le chapitre 4 se termine par des applications numériques: nos simulations montrent que l'absence de moment, n'a pas trop d'impact sur l'estimation des paramètres à échantillon fini. Les applications réalisées sur les indices boursiers NASDAQ, CAC40, DAX et HSI confirment la présence de volatilité à long terme sur ces processus.

Chapitre 5: Sur le taux de croissance des processus super-additifs et la stabilité des fGARCHs dans $\mathcal{C}^0$.

Ce chapitre est divisé en deux grandes parties. La motivation principale de la première est de proposer une extension de la propriété de Kesten (1975) sur les marches aléatoires à termes stationnaires et d'en déduire la propriété de contraction des produits d'opérateurs

aléatoires. Le résultat de Kesten affirme que si $(\boldsymbol{z}_t)$ est un processus stationnaire, alors

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad p.s. \text{ dans } \{\boldsymbol{S}_n := \sum_{k=1}^{n} \boldsymbol{z}_k \to \infty \text{ lorsque } n \to \infty\}. \tag{2.21}$$

Ce résultat démontre que la marche aléatoire ne peut diverger vers l'infini à un taux plus petit que n. La propriété (2.5) de Bougerol and Picard (1992b), établie sous des conditions supplémentaires d'ergodicité et d'intégrabilité, est une extension partielle de cette propriété aux produits de matrices aléatoires. Le résultat de Kesten (1975) est souvent exprimé en utilisant des notions de la théorie ergodique. Pour partager cette expression, rappelons brièvement quelques définitions et notions de base de la théorie ergodique que nous allons utiliser, ainsi qu'une petite revue de la littérature pour contextualiser.

Un système dynamique mesuré est un quadruplet $(\Omega, \mathcal{B}, \mu, T)$ où $(\Omega, \mathcal{B}, \mu)$ est un espace probabilisé et $T : \Omega \to \Omega$ est une application mesurable qui préserve μ, c'est-à-dire que pour tout $A \in \mathcal{B}$, $\mu(T^{-1}(A)) = \mu(A)$. Un ensemble $I \in \mathcal{B}$ est dit invariant si $\mu\left(I \Delta T^{-1}(I)\right) = 0$. Nous dirons que T est ergodique si pour tout ensemble invariant I, $\mu(I) = 0$ ou $\mu(I) = 1$.

Un processus $(\boldsymbol{S}_n)_{n\geq 1}$ est dit super-additif si pour tout n, s, $\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s}$ $p.s.$ Ces processus ont été introduits par Hammersley and Welsh (1965) et l'une des contributions les plus importantes sur les propriétés asymptotiques de ces processus est le théorème ergodique sous-additif de Kingman (1973) qui établit la convergence presque sûre de $(n^{-1}\boldsymbol{S}_n)_n$ sous l'hypothèse d'intégrabilité de $\boldsymbol{S}^-$. Ce résultat généralise le théorème ergodique très connu de Birkhoff sur les processus additifs, c.à.d quand pour tout n, s, $\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n = \boldsymbol{S}_{n+s}$ $p.s.$ Le résultat de Kesten (1975) dit que si $(\boldsymbol{S}_n)_n$ est additif alors

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad p.s. \text{ dans } \{\boldsymbol{S}_n \to \infty, n \to \infty\}. \tag{2.22}$$

Cette propriété reste vraie même lorsque $\boldsymbol{S}_1^-$ n'est pas intégrable. Ce résultat a trouvé de nombreuses applications en théorie ergodique et a été principalement précurseur dans l'étude de la récurrence des marches aléatoires stationnaires.

L'un des principaux résultats du chapitre 5, établi dans le théorème 5.1, généralise la propriété (2.22) au processus super-additif. Nous montrons dans ce théorème que si $(\boldsymbol{S}_n)_n$ est super-additif alors

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad p.s. \text{ dans } \liminf\{\boldsymbol{S}_n > 0\}. \tag{2.23}$$

Ce résultat est établi sous des conditions plus faibles que celles de Kesten, qui suppose que $(\boldsymbol{S}_n)$ diverge vers l'infini. En effet, il suffit de remarquer que $\liminf\{\boldsymbol{S}_n > 0\}$ est l'ensemble des $\omega \in \Omega$ tel que la $(\boldsymbol{S}_n)$ est positive pour n assez grand. Notre preuve repose sur des techniques nouvelles, basées sur des concepts de la théorie ergodique, et diffère de celles utilisées par Kesten et Bougerol-Picard. Différentes versions plus fortes de la propriété (2.23) sont proposées à travers les théorèmes 5.2 et 5.3 ainsi que les corollaires 5.1 et 5.2. Elles requièrent toutefois des conditions supplémentaires. Nous discutons également de la nécessité et de l'optimalité de ces hypothèses tout au long du chapitre. Ces résultats stipulent que l'hypothèse que la suite $(\boldsymbol{S}_n)$ soit strictement positive a partir d'un certain

rang peut être affaiblie dans certains cas. Dans le théorème 5.2, nous montrons que si E est un sous-ensemble invariant de $\{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0\} \cap \liminf\{\boldsymbol{S}_n \geq 0\}$, alors

$$\liminf_{n\to\infty} n^{-1}\boldsymbol{S}_n > 0 \text{ p.s. dans } E. \tag{2.24}$$

En remarquant que Ω est invariant, une conséquence directe de ce résultat est que si, presque sûrement, la séquence $(\boldsymbol{S}_n)$ prend au moins une valeur strictement positive et qu'elle est positive (ou nulle) à partir d'un certain rang alors $\liminf n^{-1}\boldsymbol{S}_n > 0 \;\; p.s.$

Soit $A \in \mathcal{B}$ et E' l'ensemble des éléments de A tels que la séquence $(\boldsymbol{S}_n : \; T^n \in A)$ est positive à partir d'un certain rang. le théorème 5.3 établit que si $\boldsymbol{S}_1^-$ est intégrable, alors

$$\lim n^{-1}\boldsymbol{S}_n > 0 \;\; p.s. \text{ dans } E'.$$

Cette propriété montre, sous l'hypothèse d'intégrabilité, que dans A, la condition de positivité ne fait intervenir que les valeurs de $(\boldsymbol{S}_n)$ avec des indices dans $(n : \; T^n \in A)$. Nous déduisons de ces résultats que si T est ergodique et que $(\boldsymbol{\Lambda}_n)_{n\geq 1}$ est un processus positif sous-multiplicatif, c'est-à-dire que $(-\ln \boldsymbol{\Lambda}_n)$ est un processus super-additif, alors pour tout $A \in \mathcal{B}$ tel que $\mu(A) > 0$:

1. si $\mu(\liminf_n\{\boldsymbol{\Lambda}_n < 1\}) > 0$, alors $\gamma := \limsup_n n^{-1}\log \boldsymbol{\Lambda}_n$ est presque sûrement constante dans $\overline{\mathbb{R}}$ et $\gamma < 0 \;\; p.s.$,
2. si $\mathbb{E}\log^+ \boldsymbol{\Lambda}_1$ est finie et, presque sûrement, la séquence $(\boldsymbol{\Lambda}_n : \; T^n \in A)$ est strictement inférieure à 1 à partir d'un certain rang, alors

$$\gamma = \lim_n n^{-1}\log \boldsymbol{\Lambda}_n = \lim_n n^{-1}\mathbb{E}\log \boldsymbol{\Lambda}_n = \inf_n n^{-1}\mathbb{E}\log \boldsymbol{\Lambda}_n < 0 \;\; p.s.$$

L'énoncé de ce résultat se trouve dans le corollaire 5.3. En remarquant que $(\boldsymbol{\Lambda}_0^{(n)})$ est un cas particulier de processus sous-multiplicatif, ces propriétés généralisent la propriété (2.5) aux produits d'opérateurs aléatoires et, plus généralement, aux processus sous-multiplicatifs ergodiques.

La deuxième partie du chapitre 5 se concentre sur l'étude de la stationnarité des SREs à coefficients positifs dans l'espace F des fonctions continues à support compact. Le modèle s'écrit :

$$\boldsymbol{h}_t = \delta(\boldsymbol{\eta}_{t-1}) + \gamma(\boldsymbol{\eta}_{t-1})\boldsymbol{h}_{t-1} = \boldsymbol{\Psi}_t\left(\boldsymbol{h}_{t-1}\right), \tag{2.25}$$

où $(\boldsymbol{\eta}_t)$ un processus strictement stationnaire ergodique dans F. En notant F_+ l'ensemble des fonctions positives de F, $\delta(\cdot)$ et $\gamma(\cdot)$ sont des fonctions mesurables respectivement à valeurs dans F_+ et dans l'espace des opérateurs linéaires continus qui envoient les éléments de F_+ dans F_+. Cette classe de modèles regroupe la plupart des modèles de volatilité conditionnelle, y compris les fGARCH continus.

Le résultat principal de la deuxième partie est le théorème 5.4. Il établit, sous des hypothèses relativement faibles, des conditions nécessaires et suffisantes d'existence de solutions positives strictement stationnaires à l'équation (2.25) en reliant le signe du coefficient de Lyapunov $\boldsymbol{\gamma}(\boldsymbol{\Psi})$ de $(\boldsymbol{\Psi}_t)$ à cette existence. La preuve de ce résultat utilise les propriétés de contraction établies précédemment et des arguments nouveaux pour

surmonter d'autres difficultés techniques, que nous détaillerons dans le chapitre.

Nous déduisons dans le corollaire 5.5 des conditions nécessaires et suffisantes d'existence de solutions positives strictement stationnaires pour les modèles fGARCH dans $\mathcal{C}([0,1])$ introduits par Aue et al. (2017) et Hörmann et al. (2013).

Ces résultats sont, à notre connaissance, les premiers de ce type à établir la nécessité de (2.6) dans l'existence de solutions stationnaires pour les modèles de type (2.25) lorsque la dimension de F n'est pas finie. Plus de discussions à ce sujet sont apportées dans ce chapitre.

Chapter 3

Exponential control of the trajectories of Iterated Function Systems and application to semi-strong GARCH(p,q) models

A version of this chapter is forthcoming in the *Journal of Applied Probability.*

Abstract

The chapter establishes new results on the strictly stationary solution to an Iterated Function System (IFS). When the driving sequence is stationary and ergodic, though not independent, the strictly stationary solution may admit no moment but we show an exponential control of the trajectories. We exploit these results to prove, under mild conditions, the consistency of the quasi-maximum likelihood estimator of GARCH(p, q) models with non independent innovations.

Contents

1 Introduction 28
2 Stochastic IFS without moments 29
3 Proof of the main result 31
4 Inference for semi-strong GARCH 36
5 Appendix: Complementary Proofs 41

1 Introduction

Since Kesten (1973), the theoretical properties of the Stochastic Recurrence Equation (SRE) $\boldsymbol{X}_t = \boldsymbol{A}_t\boldsymbol{X}_{t-1} + \boldsymbol{B}_t$ has received much attention. This equation gathers a large class of classical econometric processes such as the GARCH and ARMA models, and their numerous variants. Brandt (1986) proposed a sufficient condition of existence and uniqueness of a strictly stationary solution in the case where $(\boldsymbol{A}_t, \boldsymbol{B}_t)_t$ is stationary and ergodic. Under an irreducibility condition, Bougerol and Picard (1992b) established that this condition is also necessary when the sequence $(\boldsymbol{A}_t, \boldsymbol{B}_t)$ is independent and identically distributed (iid). The probabilistic properties of the stationary solution of SRE model in the iid case are well known. In the scalar case, Kesten (1973) showed that $\mathbb{P}(\pm\boldsymbol{X}_1 > x) \sim c_\pm x^{-a}$ *as* $x \to \infty$ for some positive constants $c_\pm$. A thorough study of SRE models, in particular their tail behavior, is presented in Buraczewski et al. (2016). The SRE model is the affine mapping particular case of the so-called Stochastic Iterated Function Systems (IFS) $\boldsymbol{X}_t = \Psi(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1})$. Most of the theoretical properties established for SRE models (stationary, tail properties) can be extended to IFS equations.

One important application of SREs in time series analysis is the study of the stationarity properties of GARCH processes. Assuming iid innovations, Bougerol and Picard (1992a) deduced from Brandt (1986) a necessary and sufficient condition for the existence of a unique stationary solution of a general GARCH(p, q) model. In recent years, the iid assumption on the innovations has often been replaced by a less restrictive conditional moment assumption (the model is then called « semi-strong » GARCH). See Escanciano (2009) for the classical GARCH(p, q) model, Francq and Thieu (2019) and Han and Kristensen (2014) for GARCH-X models. The GARCH-MIDAS models of Engle et al. (2013) constitute other class of IFS models which are not driven by an iid sequence. Another example is given by GARCH-X models which are IFS driven by a–generally non iid– sequence of innovations and covariates. This motivates studying IFS equations driven by non iid innovations.

However, strict stationarity generally does not suffice for establishing the asymptotic properties of estimators, such as the Quasi-Maximum Likelihood Estimator (QMLE). To our knowledge, all existing works on the QML inference of IFS models *assume* the existence of a small-order moment of the observed process. Surprisingly, however, the strictly stationary solutions of IFS equations with non-iid innovations may not admit any finite moment.

The aim of this chapter is to establish that the stationary trajectories of the IFS equations enjoy an exponential control property. We also show that this property is sufficient to establish the consistency of the QMLE of semi-strong GARCH models.

The rest of the chapter is organized as follows. In Section 2 we present our main result and Section 3 is devoted to its proof. Section 4 investigates the estimation of the semi-strong GARCH(p, q) model. Complementary proofs are displayed in the appendices.

2 Stochastic IFS without moments

Let $(E, \mathcal{E})$ be a measurable space and (F, d) a complete and separable metric space (Polish space). Let $(\boldsymbol{\theta}_t)_{t\in\mathbb{Z}}$ be a stationary and ergodic process valued in E, and let $\Psi : E \times F \to F$ a function such that $x \mapsto \Psi(\theta, x)$ is Lipschitz continuous for all $\theta \in E$. Let

$$\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t) = \sup_{x_1,x_2\in F, x_1\neq x_2} \frac{d\left(\boldsymbol{\Psi}_t\left(x_1\right), \boldsymbol{\Psi}_t\left(x_2\right)\right)}{d\left(x_1, x_2\right)}$$

where $\boldsymbol{\Psi}_t = \Psi(\boldsymbol{\theta}_t, \cdot)$. Let $\boldsymbol{\Lambda}_t^{(0)} = 1$ and $\boldsymbol{\Lambda}_t^{(r)} = \boldsymbol{\Lambda}(\boldsymbol{\Psi}_t \circ \cdots \circ \boldsymbol{\Psi}_{t-r+1})$ for all $r > 0$.

Consider the IFS

$$\boldsymbol{X}_t = \Psi\left(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1}\right) = \boldsymbol{\Psi}_t\left(\boldsymbol{X}_{t-1}\right), \quad \text{for all } t \in \mathbb{Z}. \tag{3.1}$$

A solution $(\boldsymbol{X}_t)$ of (3.1) is said to be causal if, for every t, $\boldsymbol{X}_t$ is $\sigma\left(\boldsymbol{\theta}_k,\ k \leq t\right)$-measurable.

Under a slightly different form, the following result has been established by Elton (1990, Theorem 3) and Bougerol (1993, Theorem 3.1), see also Straumann and Mikosch (2006, Theorem 2.8) and the review of Diaconis and Freedman (1999).

Theorem 3.1. *Assume the following conditions hold: (i) there exists a constant $c \in F$ such that $\mathbb{E}\ln^+ d\left(\boldsymbol{\Psi}_0(c), c\right) < \infty$, (ii) $\mathbb{E}\ln^+ \boldsymbol{\Lambda}_0 < \infty$ and (iii) $\lim_{r\to\infty} \frac{1}{r}\ln \boldsymbol{\Lambda}_0^{(r)} < 0$ a.s. Then there exists a unique stationary (causal and ergodic) solution $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ to Equation (3.1).*

Moreover we have:

$$\text{for all } \ t \in \mathbb{Z}, \quad d(\boldsymbol{X}_t, c) \leq \sum_{n=0}^{\infty} \boldsymbol{\Lambda}_t^{(n)} d\left(\boldsymbol{\Psi}_{t-n}(c), c\right) < \infty \quad a.s. \tag{3.2}$$

Note that $(\ln \boldsymbol{\Lambda}_0^{(r)})_{r>1}$ is a sub-additive sequence. Therefore, by the sub-additive ergodic theorem of Kingman (1973), the limit in assumption (iii) exists.

For the reader's convenience and because we have not been able to find Equation (3.2) exactly under this form, we provide a proof for Theorem 3.1 in the appendix.

In L_p space with iid driven sequences, a consequence of theorem 3.1 is that the stationary solution is L_p-approximable, see Hörmann and Kokoszka (2010, Definition 2.1). This property is often used to study the theoretical properties of non-linear time series, such as the IFSs in (3.1), see Hörmann and Kokoszka (2010), Berkes et al. (2009), Hörmann (2008). However, in the general setup this property is difficult to apply because this L_p-approximable notion requires the existence of higher-order moments.

Remark 3.1. If $(\boldsymbol{\theta}_t)$ is iid, it is possible to prove in particular cases, including the affine mapping, that $d(\boldsymbol{X}_1, c)$ has a power-law tail, see Buraczewski et al. (2016, Theorem 5.3.6). More generally, it can be shown that, under the conditions of Theorem 3.1, there exists $s > 0$ such that $\mathbb{E}d(\boldsymbol{X}_1, c)^s < \infty$. This small moment property is often used in the statistical inference of IFS models, for example, to prove the consistency of GARCH

models and its derivatives (see Berkes et al. (2003) for GARCH model and Francq et al. (2018) for EGARCH and Log-GARCH model). If $(\boldsymbol{\theta}_t)$ is not iid, the examples below show that the stationary solution may not admit any small-order moment.

Example 3.1. *Let $\delta \in (0,1)$ and let $(\boldsymbol{z}_t)_{t\in\mathbb{Z}}$ be an iid non negative real process with $\mathbb{E}\boldsymbol{z}_t = \frac{1-\delta}{2}$ and $\mathbb{E}\boldsymbol{z}_t^2 = \infty$. The process $(\boldsymbol{\theta}_t)$, defined by $\boldsymbol{\theta}_t = \sum_{k=0}^{\infty} \delta^k \boldsymbol{z}_{t-k}$ for all $t \in \mathbb{Z}$ satisfies $\mathbb{E}\boldsymbol{\theta}_t = \frac{1}{2}$ and is such that for all $t \in \mathbb{Z}$, $\boldsymbol{x}_t = 1 + \sum_{k=1}^{\infty} \prod_{j=1}^{k} \boldsymbol{\theta}_{t-j+1}$ exists a.s. Moreover $(\boldsymbol{x}_t)$ is the unique stationary solution of*

$$\boldsymbol{x}_t = \boldsymbol{\theta}_t \boldsymbol{x}_{t-1} + 1, \qquad t \in \mathbb{Z}. \tag{3.3}$$

Note that $\boldsymbol{x}_t \geq \prod_{j=1}^{k} \boldsymbol{\theta}_{t-j+1} \geq \delta^{\frac{k(k-1)}{2}} (\boldsymbol{z}_{t-k+1})^k$ for all $k \in \mathbb{N}^*$. For all $s > 0$, we thus have $\mathbb{E}\boldsymbol{x}_0^s \geq \mathbb{E}\delta^{\frac{sk(k-1)}{2}} (\boldsymbol{z}_0)^{sk} = \infty$ for k such that $sk > 2$.

The previous example is simple but probably a little artificial. We now give an example of commonly used econometric models, for which it was recently proven that the strictly stationary solution does not admit any finite moment.

Example 3.2. *Consider the following GARCH-MIDAS model of Engle et al. (2013)*

$$\begin{cases} \boldsymbol{r}_t = \sqrt{\boldsymbol{\tau}_t}\boldsymbol{\sigma}_t\boldsymbol{\eta}_t \\ \boldsymbol{\tau}_t = a + b\boldsymbol{r}_{t-1}^2 \\ \boldsymbol{\sigma}_t^2 = 1 - \alpha - \beta + \alpha\frac{\boldsymbol{r}_{t-1}^2}{\boldsymbol{\tau}_t} + \beta\boldsymbol{\sigma}_{t-1}^2 \end{cases} \tag{3.4}$$

where $(\boldsymbol{\eta}_t)_t$ is a zero mean and unit variance iid sequence, $\alpha > 0$, $\beta \geq 0$, $\alpha+\beta < 1$, $a > 0$ and $b > 0$. Noting that $\boldsymbol{\epsilon}_t := \boldsymbol{\sigma}_t\boldsymbol{\eta}_t$ is a GARCH process, we see that $(\boldsymbol{\tau}_t)$ follows a SRE

$$\boldsymbol{\tau}_t = a + b\boldsymbol{r}_{t-1}^2 = a + (b\boldsymbol{\epsilon}_{t-1}^2)\boldsymbol{\tau}_{t-1}$$

driven by a non iid sequence $\boldsymbol{\epsilon}_t$. It can be shown that, when $b \leq 1$, the process $(\boldsymbol{r}_t)$ is strictly stationary but, when $\boldsymbol{\eta}_0$ has a unbounded support, then

$$\text{for any } s > 0, \quad E|\boldsymbol{r}_t|^s = \infty.$$

See Proposition 4.1 for the proof of the previous result.

We now state our main result, which provides a way to circumvent the non existence of small order moments for models such as those of Examples 3.1 and 3.2. Section 4 will be devoted to the statistical study of a class of econometric models where the existence of moments is not guaranteed.

Theorem 3.2. *Under the conditions of Theorem 3.1, if (i) $\mathbb{P}(d(\boldsymbol{X}_1, c) = 0) < 1$ then: for all $t \in \mathbb{Z}$*

1. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c) = 0$ *and* 2. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t-n}, c) = 0$ *a.s.*

The condition *(i)* of the theorem means that $d(\boldsymbol{X}_t, c)$ is not almost surely zero. In the case where $d(\boldsymbol{X}_t, c) = 0$ *a.s.*, it is easy to see that the limit superior become $-\infty$. Theorem

3.2 can be interpreted as an exponential control of the trajectory of the stationary solution. Note that the property $\mathbb{E}\ln^+ d(\boldsymbol{X}_1, c) < \infty$ (a weaker condition than the existence of a small-order moment) implies the results of Theorem 3.2 (see Appendix 5.2). However, the converse is false, see Example (a) of Tanny (1974). Note that in the general case of a stationary and ergodic sequence $(\boldsymbol{Z}_t)$ such that $\mathbb{P}(\boldsymbol{Z}_1 > -\infty) > 0$, we show in Appendix 5.3 that:

$$(i)\ \limsup_{n\to\infty} \frac{1}{n}\boldsymbol{Z}_n = 0\ a.s.\ \text{or}\ (ii)\ \limsup_{n\to\infty} \frac{1}{n}\boldsymbol{Z}_n = \infty\ a.s. \tag{3.5}$$

Theorem 3.2 shows that the sequence $(d(\boldsymbol{X}_t, c))_t$ belongs to case (i) of Eq. (3.5).

As a consequence of the previous theorem, we obtain the following result. Its proof is provided in Appendix 5.4.

Corollary 3.1. *Under the conditions of Theorem 3.2, almost surely*

$$\lim_{|n|\to\infty} \frac{1}{|n|} \ln^+ d(\boldsymbol{X}_{t+n}, c) \text{ exists and is equal to } 0,$$

and if $\mathbb{E}\ln^- d(\boldsymbol{X}_1, c) < \infty$ *then*

$$\lim_{|n|\to\infty} \frac{1}{|n|} \ln d(\boldsymbol{X}_{t+n}, c) \text{ exists and is equal to } 0. \tag{3.6}$$

3 Proof of the main result

To show Theorem 3.2, we first define a SRE which bounds the distance between $\boldsymbol{X}_t$ and c.

Note that, by Kingman (1973),

$$\lim_{r\to\infty} \frac{1}{r} \ln \boldsymbol{\Lambda}_0^{(r)} = \inf_{r\in\mathbb{N}^*} \frac{1}{r}\mathbb{E}\ln \boldsymbol{\Lambda}_0^{(r)} = \lim_{r\to\infty} \frac{1}{r}\mathbb{E}\ln \boldsymbol{\Lambda}_0^{(r)}\ a.s., \tag{3.7}$$

so by *iii)* of Theorem 3.1 there exists a positive integer r_0 such that $\mathbb{E}\ln \boldsymbol{\Lambda}_0^{(r_0)} < 0$. It can be shown that $\mathbb{E}\left[\ln\left((\boldsymbol{\Lambda}_0^{(r_0)} + u)\right)\right] \overset{u\to 0}{\to} \mathbb{E}\ln \boldsymbol{\Lambda}_0^{(r_0)}$, see Straumann and Mikosch (2006, proof of Theorem 2.10). Therefore $\exists\ u_0 > 0,\ \ln(u_0) \le \gamma_0 := \mathbb{E}\left[\ln\left((\boldsymbol{\Lambda}_0^{(r_0)} + u_0)\right)\right] < 0$. We thus have, for all $v \in [\gamma_0, 0)$,

$$\mathbb{E}\left[\ln\left(\delta(v)(\boldsymbol{\Lambda}_0^{(r_0)} + u_0)\right)\right] = v \tag{3.8}$$

with $\delta(v) = \exp(v - \gamma_0) \ge 1$.

Now, for any integer $p \in [0, r_0 - 1]$, define $(\boldsymbol{a}_{p,t}(v), \boldsymbol{b}_{p,t})_{t\in\mathbb{Z}}$ by

$$\boldsymbol{a}_{p,t}(v) = \delta(v)(\boldsymbol{\Lambda}_{r_0 t+p}^{(r_0)} + u_0),\ \text{and}\ \boldsymbol{b}_{p,t} = 1 + \sum_{k=0}^{r_0-1} \boldsymbol{\Lambda}_{r_0 t+p}^{(k)} d\left(\boldsymbol{\Psi}_{r_0 t+p-k}(c), c\right).$$

By Assumptions (i) and (ii) of Theorem 3.1 and by the elementary inequality $\ln\left(\sum_{i=1}^n a_i\right) \leq \ln n + \sum_{i=1}^n \ln^+ a_i$ for non-negative $\{a_i\}_{i=1}^n$, we have $E\ln^+ \boldsymbol{a}_{p,t}(v) < \infty$ and $E\ln^+ \boldsymbol{b}_{p,t}(v) < \infty$. Therefore, in view of (3.8), there exists a unique stationary solution $(\boldsymbol{z}_{p,t}(v))_t$ to the equation

$$\boldsymbol{z}_{p,t}(v) = \boldsymbol{a}_{p,t}(v)\boldsymbol{z}_{p,t-1}(v) + \boldsymbol{b}_{p,t}. \tag{3.9}$$

Note that by Brandt (1986)

$$\boldsymbol{z}_{p,t}(v) = \sum_{q=0}^{\infty}\left(\prod_{i=0}^{q-1}\boldsymbol{a}_{p,t-i}(v)\right)\boldsymbol{b}_{p,t-q}. \tag{3.10}$$

By iterating Equation (3.9) we have

$$\boldsymbol{z}_{p,t}(v) = \sum_{q=0}^{n}\left(\prod_{i=0}^{q-1}\boldsymbol{a}_{p,t-i}(v)\right)\boldsymbol{b}_{p,t-q} + \left(\prod_{i=0}^{n}\boldsymbol{a}_{p,t-i}(v)\right)\boldsymbol{z}_{p,t-(n+1)}(v), \quad \forall n \geq 1. \tag{3.11}$$

By (3.10) and (3.11), $\left(\prod_{i=0}^n \boldsymbol{a}_{p,t-i}(v)\right)\boldsymbol{z}_{p,t-(n+1)}(v)$ is the remainder of a convergent series, hence it almost surely converges to 0. i.e.

$$\left(\prod_{k=0}^{n-1}\boldsymbol{a}_{p,t-k}(v)\right)\boldsymbol{z}_{p,t-n}(v) \overset{n\to\infty}{\to} 0 \quad a.s. \tag{3.12}$$

We now give a technical lemma linking the processes $(\boldsymbol{X}_t)$ and $(\boldsymbol{z}_{p,t}(v))_t$.

Lemma 3.1. *For all $v \in [\gamma_0, 0)$, $0 \leq p \leq r_0 - 1$, and $t \in \mathbb{Z}$, we have*

$$d(\boldsymbol{X}_{r_0t+p}, c) \leq \boldsymbol{z}_{p,t}(v) \quad a.s. \tag{3.13}$$

Proof of Lemma 3.1. For any integer n, let q and m denote the quotient and remainder of the Euclidean division of n by r_0: $n = qr_0 + m$. By sub-multiplicativity we have

$$\boldsymbol{\Lambda}_t^{(n)} \leq \left(\prod_{i=0}^{q-1}\boldsymbol{\Lambda}_{t-ir_0}^{(r_0)}\right)\boldsymbol{\Lambda}_{t-qr_0}^{(m)}, \quad \text{with} \quad \prod_{i=0}^{-1}\boldsymbol{\Lambda}_{t-ir_0}^{(r_0)} = 1.$$

For all $q \in \mathbb{N}$, we then obtain

$$\sum_{n=qr_0}^{(q+1)r_0-1}\boldsymbol{\Lambda}_t^{(n)} d\left(\boldsymbol{\Psi}_{t-n}(c), c\right) \leq \left(\prod_{i=0}^{q-1}\boldsymbol{\Lambda}_{t-ir_0}^{(r_0)}\right)\sum_{m=0}^{r_0-1}\boldsymbol{\Lambda}_{t-qr_0}^{(m)} d\left(\boldsymbol{\Psi}_{t-qr_0-m}(c), c\right).$$

It follows that

$$\sum_{n=0}^{\infty} \mathbf{\Lambda}_t^{(n)} d\left(\mathbf{\Psi}_{t-n}(c), c\right) = \sum_{q=0}^{\infty} \sum_{n=qr_0}^{(q+1)r_0-1} \mathbf{\Lambda}_t^{(n)} d\left(\mathbf{\Psi}_{t-n}(c), c\right)$$
$$\leq \sum_{q=0}^{\infty} \left(\prod_{i=0}^{q-1} \mathbf{\Lambda}_{t-ir_0}^{(r_0)} \right) \sum_{m=0}^{r_0-1} \mathbf{\Lambda}_{t-qr_0}^{(m)} d\left(\mathbf{\Psi}_{t-qr_0-m}(c), c\right).$$

Since $\delta(v) \geq 1$ and $u_0 > 0$, we obtain

$$\left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q} \geq \left(\prod_{i=0}^{q-1} \mathbf{\Lambda}_{(r_0t+p)-ir_0}^{(r_0)} \right) \sum_{m=0}^{r_0-1} \mathbf{\Lambda}_{(r_0t+p)-qr_0}^{(m)} d\left(\mathbf{\Psi}_{(r_0t+p)-qr_0-m}(c), c\right).$$

In view of the last two inequalities, together with (3.10) and (3.2), we have

$$\boldsymbol{z}_{p,t}(v) \geq \sum_{n=0}^{\infty} \mathbf{\Lambda}_{r_0t+p}^{(n)} d\left(\mathbf{\Psi}_{r_0t+p-n}(c), c\right) \geq d(\boldsymbol{X}_{r_0t+p}, c),$$

which proves (3.13). □

Let **Aff** denote the set of affine maps from $\mathbb{R}$ into $\mathbb{R}$. An element $\boldsymbol{f}_{a,b}$ of **Aff** can be written as

$$\boldsymbol{f}_{a,b}(x) = ax + b, \quad x \in \mathbb{R}, \quad \text{where } (a,b) \in \mathbb{R}^2$$

Lemma 3.2. *Let us define a function Φ from **Aff** to $\mathbb{R}_+$ by $\Phi(\boldsymbol{f}_{a,b}) = |a| + |b|$.*

1. *For any x, $|x| \geq 1$, $|\boldsymbol{f}_{a,b}(x)| \leq \Phi(\boldsymbol{f}_{a,b})|x|$.*
2. *If $|d| \geq 1$ then $\Phi(\boldsymbol{f}_{a,b} \circ \boldsymbol{f}_{c,d}) \leq \Phi(\boldsymbol{f}_{a,b})\Phi(\boldsymbol{f}_{c,d})$.*

Since Lemma 3.2 is elementary, its proof is skipped. Note that Φ is the 1-norm in the vector space of affine maps.

Lemma 3.3. *For all $p \in \{0, \ldots, r_0 - 1\}$ and $t \in \mathbb{Z}$, letting $Q_p(t) = r_0t + p$, we have*

1. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t+n)}, c) \leq 0$,
2. $\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{Q_p(t-n)}, c) \leq 0$ *a.s.*

In the previous lemma we distinguished cases 1. and 2. because their proofs are different.

Proof of Lemma 3.3. We start by proving Point 1. Let $\boldsymbol{f}_t$ be the random affine map defined by

$$\boldsymbol{f}_t(x) = \boldsymbol{a}_{p,t}(v)x + \boldsymbol{b}_{p,t}$$

for all $x \in \mathbb{R}$. Define also the maps

$$\boldsymbol{\gamma}_{t,n} = \boldsymbol{f}_t \circ \boldsymbol{f}_{t-1} \cdots \circ \boldsymbol{f}_{t-n+1} \quad \text{and} \quad \boldsymbol{\zeta}_{t,n} = \boldsymbol{f}_{t+n} \circ \boldsymbol{f}_{t+n-1} \cdots \circ \boldsymbol{f}_{t+1}$$

for all $(t,n) \in \mathbb{Z} \times \mathbb{N}^*$. Note that,

$$\boldsymbol{\zeta}_{t,n} = \boldsymbol{\gamma}_{t+n,n},\ \boldsymbol{z}_{p,t}(v) = \boldsymbol{\gamma}_{t,n}(\boldsymbol{z}_{p,t-n}(v)) \text{ and } \boldsymbol{z}_{p,t+n}(v) = \boldsymbol{\zeta}_{t,n}(\boldsymbol{z}_{p,t}(v)) \quad a.s. \tag{3.14}$$

Since $\boldsymbol{b}_{p,t} \geq 1$, by *2.)* of Lemma 3.2

$$(\boldsymbol{u}_{t,n})_n := (\ln \Phi(\boldsymbol{\gamma}_{t,n}))_n \quad \text{and} \quad (\boldsymbol{w}_{t,n})_n := (\ln \Phi(\boldsymbol{\zeta}_{t,n}))_n \tag{3.15}$$

are sub-additive sequences. By already used arguments, we have $\mathbb{E}|\ln \Phi(\boldsymbol{\gamma}_{t,1})| = \mathbb{E}|\ln \Phi(\boldsymbol{\zeta}_{t,1})| = \mathbb{E}|\ln \Phi(\boldsymbol{f}_t)| < \infty$. In view of (3.14) and *1.* of Lemma 3.2,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \leq \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{w}_{t,n} + \limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t}(v) \quad a.s.$$

Because $\boldsymbol{z}_{p,t}(v)$ does not depend on n, we have $\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t}(v) = 0$ $a.s.$ Therefore

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \leq \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{w}_{t,n} \quad a.s. \tag{3.16}$$

Since for any $n \in \mathbb{N}^*$, $\boldsymbol{u}_{t,n}$ and $\boldsymbol{w}_{t,n}$ have the same law, by (3.15) and Kingman sub-additive ergodic theorem,

$$\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{w}_{t,n} = \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}\boldsymbol{u}_{t,n} = \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}_{t,n} \quad a.s. \tag{3.17}$$

On the other hand, in view of (3.11), we have by the positivity of the coefficients,

$$\Phi(\boldsymbol{\gamma}_{t,n+1}) = \sum_{q=0}^{n} \left(\prod_{i=0}^{q-1} \boldsymbol{a}_{p,t-i}(v) \right) \boldsymbol{b}_{p,t-q} + \left(\prod_{i=0}^{n} \boldsymbol{a}_{p,t-i}(v) \right) \overset{n\to\infty}{\to} \boldsymbol{z}_{p,t}(v) \quad a.s.$$

Therefore

$$\lim_{n\to\infty} \boldsymbol{u}_{t,n} = \ln \boldsymbol{z}_{p,t}(v) \quad a.s.,$$

which entails

$$\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{u}_{t,n} = 0 \quad a.s. \tag{3.18}$$

By (3.16), (3.17) and (3.18) we get

$$\limsup_{n\to\infty} \frac{1}{n} \ln \boldsymbol{z}_{p,t+n}(v) \leq 0 \quad a.s.,$$

which implies, by Equation (3.13), Point 1. of the lemma

For the second point, by (3.13), (3.12), (3.8) and the ergodic theorem, we have

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{Q_p(t-n)},c) &\leq \limsup_{n\to\infty}\frac{1}{n}\ln \boldsymbol{z}_{p,t-n}(v)\\ &\leq \limsup_{n\to\infty}\frac{1}{n}\ln\left(\prod_{i=0}^{n-1}\boldsymbol{a}_{p,t-i}(v)\right)\boldsymbol{z}_{p,t-n}(v)\\ &\quad -\liminf_{n\to\infty}\frac{1}{n}\ln\left(\prod_{i=0}^{n-1}\boldsymbol{a}_{p,t-i}(v)\right)\\ &\leq -v\ a.s.\end{aligned}$$

For all $v\in[\gamma_0,0)$. Letting $v\to 0^-$ we get the result. □

We are now ready to prove Theorem 3.2.

Proof of Theorem 3.2. For all $t\in\mathbb{Z}$, let $t'\in\mathbb{Z}$ and p', $0\leq p'\leq r_0-1$ such that $t=r_0t'+p'$. Note that

$$\{t+k,k\in\mathbb{N}\}\subset\bigcup_{0\leq p\leq r_0-1}\{r_0(t'+k)+p,k\in\mathbb{N}\}.$$

The previous relation and the first point of Lemma 3.3 imply that

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{t+n},c) &\leq \max_{0\leq p\leq r_0-1}\left(\limsup_{n\to\infty}\frac{1}{Q_p(t'+n)}\ln d(\boldsymbol{X}_{Q_p(t'+n)},c)\right)\\ &\leq C\max_{0\leq p\leq r_0-1}\left(\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{Q_p(t'+n)},c)\right)\leq 0\ a.s.,\end{aligned}$$

for $C=\max_{0\leq p\leq r_0-1}\left(\sup_{n\geq 0}\frac{n}{Q_p(t'+n)}\right)$.

We now show that this inequality is reduces to an equality. Let us argue by contradiction. Suppose that $\mathbb{P}\left(\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{t+n},c)<0\right)>0$. Since $\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{t+n},c)$ is almost surely constant by ergodicity, it follows that $\limsup_{n\to\infty}\frac{1}{n}\ln d(\boldsymbol{X}_{t+n},c)<K<0\ a.s.$ for some constant K. Thus, $d(\boldsymbol{X}_{t+n},c)\leq\exp(Kn)$ from some (random) n_0 $a.s.$ This implies that $(d(\boldsymbol{X}_{t+n},c))_n$ converges to 0 $a.s.$ This contradicts the fact that $(d(\boldsymbol{X}_{t+n},c))_n$ is strictly stationary and not almost surely equal to zero. This concludes the proof of the first point of the theorem.

Without using a proof by contradiction, we could conclude by showing the inequality in the opposite sense. We argue as follows. By the condition *(i)* of Theorem 3.2, we have $\mathbb{P}\left(\ln d(\boldsymbol{X}_1,c)\geq K\right)>0$ for some real K. It follows from Birkhoff's ergodic theorem that

$$n^{-1}\sum_{k=1}^{n}\mathbb{1}_{\{\ln d(\boldsymbol{X}_{t+n},c)\geq K\}}\longrightarrow\mathbb{P}\left(\ln d(\boldsymbol{X}_1,c)\geq K\right)>0\quad a.s\quad as\quad n\to\infty.$$

Thus, almost surely, the set $\{n : \ln d(\boldsymbol{X}_{t+n}, c) \geq K\}$ is not finite. This implies that

$$\limsup_{n\to\infty} \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c) \geq \liminf_{n\to\infty} \frac{1}{n} K = 0.$$

The second point of the theorem follows with similar arguments. □

4 Inference for semi-strong GARCH(p,q)

Consider the GARCH (p, q) model

$$\begin{aligned} \boldsymbol{\epsilon}_t &= \sqrt{\boldsymbol{h}_t}\boldsymbol{\eta}_t, \\ \boldsymbol{h}_t &= \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j}\boldsymbol{h}_{t-j}, \quad \forall t \in \mathbb{Z} \end{aligned} \tag{3.19}$$

where $\omega_0 > 0, \alpha_{0i} \geqslant 0$ $(i = 1, \ldots, q)$ and $\beta_{0j} \geqslant 0$ $(j = 1, \ldots, p)$. When $(\boldsymbol{\eta}_t)$ is iid, Model (3.19) is a standard strong GARCH, for which the statistical inference has been thoroughly studied. In particular Berkes et al. (2003) and Francq and Zakoian (2004) studied the (gaussian) QMLE under the stationarity of $(\boldsymbol{\epsilon}_t)$, and Jensen and Rahbek (2004) explored the asymptotic behavior of the QMLE in the explosive case. There are alternative methods to inference for GARCH-type models, such as the Laplacian QMLE, which differ from the standard QMLE. A paper by Bardet et al. (2017) shows that the asymptotic theory of the Laplacian-QMLE works well in a wide range of econometric models, including GARCH models. In the stationary framework, Escanciano (2009) proved the consistency and asymptotic normality of the QMLE without iidness for $(\boldsymbol{\eta}_t)$, but had to assume that $E|\boldsymbol{\epsilon}_t|^s < \infty$ for some small $s > 0$. The aim of this section is to relax this extra moment assumption.

4.1 Property of the strictly stationary solution

Let

$$\boldsymbol{A}_t = \begin{pmatrix} \alpha_{01}\boldsymbol{\eta}_t^2 & \cdots & \alpha_{0q}\boldsymbol{\eta}_t^2 & \beta_{01}\boldsymbol{\eta}_t^2 & \cdots & \beta_{0p}\boldsymbol{\eta}_t^2 \\ & I_{q-1} & & & 0_{(q-1)\times p} & \\ \alpha_{01} & \cdots & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0p} \\ & 0_{(p-1)\times q} & & & I_{p-1} & \end{pmatrix} \quad \text{and} \quad \boldsymbol{b}_t = \begin{pmatrix} \omega_0\boldsymbol{\eta}_t^2 \\ 0_{q-1} \\ \omega_0 \\ 0_{p-1} \end{pmatrix}$$

with standard notations.

Model (3.19) is a special case of (3.1) where we use the notations $\boldsymbol{X}_t = \left(\boldsymbol{\epsilon}_t^2, \ldots, \boldsymbol{\epsilon}_{t-q+1}^2, \boldsymbol{h}_t^2, \ldots, \boldsymbol{h}_{t-p+1}^2\right)'$, $\theta_t = (\boldsymbol{A}_t, b_t)$, $\Psi(\theta, x) = Ax + b$, and $d(x, y) = \|x - y\|$ for any norm $\|\cdot\|$ on $\mathbb{R}^{p+q}$. Remark that $\boldsymbol{\Lambda}_t^{(r)} = \|\boldsymbol{A}_t\boldsymbol{A}_{t-1}\ldots\boldsymbol{A}_{t-r+1}\|$.

In the sequel, we do not assume that $(\boldsymbol{\eta}_t)$ is iid, we only assume that it is stationary and ergodic. If $\mathbb{E}\ln^+ \boldsymbol{\eta}_1^2 < \infty$, Theorem 3.1 applies with $c = 0_{p+q}$. Therefore, in view of (3.7), there exists a unique non-anticipative strictly stationary solution $(\boldsymbol{\epsilon}_t)$ to Model

(3.19) if

$$\begin{aligned}\gamma(\mathbf{A}) &:= \inf_{r\in\mathbb{N}^*}\frac{1}{r}\mathbb{E}\left(\ln\|\boldsymbol{A}_0\boldsymbol{A}_{-1}\ldots\boldsymbol{A}_{-r+1}\|\right)\\ &= \lim_{r\to\infty}\frac{1}{r}\ln\|\boldsymbol{A}_0\boldsymbol{A}_{-1}\ldots\boldsymbol{A}_{-r+1}\| < 0 \text{ a.s.}\end{aligned}$$

By Theorem 3.2, it follows that the strictly stationary solution of (3.19) satisfies

$$\limsup_{n\to\infty}\frac{1}{n}\ln\boldsymbol{\epsilon}_{t+n}^2 = 0, \quad \limsup_{n\to\infty}\frac{1}{n}\ln\boldsymbol{\epsilon}_{t-n}^2 = 0 \quad \text{a.s.}, \tag{3.20}$$

for all $t\in\mathbb{Z}$.

In the GARCH(1,1) case, it is easy to check that $\gamma(\mathbf{A}) = \mathbb{E}\ln(\alpha_{01}\boldsymbol{\eta}_t^2+\beta_{01})$. For general GARCH$(p,q)$ of the form (3.19), it seems impossible to compute $\gamma(\mathbf{A})$ explicitly. The issue has been discussed in several papers, see Bougerol and Picard (1992a, page 117) and Buraczewski et al. (2016, pages 148 and 149). Both papers recommend estimation by computer simulations.

4.2 QML estimator

Let $\{\boldsymbol{\epsilon}_t\}_{t=1}^n$ be a sample of size n of the unique non-anticipative strictly stationary solution of Model (3.19). The vector of parameters

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_{p+q+1})^{\mathrm{T}} = (\omega,\alpha_1,\ldots,\alpha_q,\beta_1,\ldots,\beta_p)^{\mathrm{T}}$$

belongs to a parameter space $\boldsymbol{\Theta}\subset]0,+\infty[\times[0,\infty[^{p+q}$. The true value of the parameter is unknown and is denoted by $\boldsymbol{\theta}_0 = (\omega_0,\alpha_{01},\ldots,\alpha_{0q},\beta_{01},\ldots,\beta_{0p})^{\mathrm{T}}$. Conditionally on initial values $\boldsymbol{\epsilon}_0,\ldots,\boldsymbol{\epsilon}_{1-q}$, $\tilde{\boldsymbol{\sigma}}_0^2,\ldots,\tilde{\boldsymbol{\sigma}}_{1-p}^2$, the Gaussian quasi-likelihood is defined by

$$L_n(\boldsymbol{\theta}) = L_n(\boldsymbol{\theta};\boldsymbol{\epsilon}_1,\ldots,\boldsymbol{\epsilon}_n) = \prod_{t=1}^n\frac{1}{\sqrt{2\pi\tilde{\boldsymbol{\sigma}}_t^2}}\exp\left(-\frac{\boldsymbol{\epsilon}_t^2}{2\tilde{\boldsymbol{\sigma}}_t^2}\right),$$

where the $\tilde{\boldsymbol{\sigma}}_t^2$ are defined recursively, for $t\geqslant 1$, by

$$\tilde{\boldsymbol{\sigma}}_t^2 = \tilde{\boldsymbol{\sigma}}_t^2(\boldsymbol{\theta}) = \omega + \sum_{i=1}^q\alpha_i\boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^p\beta_j\tilde{\boldsymbol{\sigma}}_{t-j}^2.$$

For instance, the initial values can be chosen as

$$\boldsymbol{\epsilon}_0^2 = \ldots = \boldsymbol{\epsilon}_{1-q}^2 = \tilde{\boldsymbol{\sigma}}_0^2 = \ldots = \tilde{\boldsymbol{\sigma}}_{1-p}^2 = c \tag{3.21}$$

with $c=\omega$ or $\boldsymbol{\epsilon}_1^2$. The standard estimator of the GARCH parameter $\boldsymbol{\theta}_0$ is the QMLE defined as any measurable solution $\hat{\boldsymbol{\theta}}_n$ of

$$\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L_n(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\tilde{\mathbf{l}}_n(\boldsymbol{\theta}) \tag{3.22}$$

where $\tilde{\mathbf{l}}_n(\boldsymbol{\theta}) = n^{-1}\sum_{t=1}^n \tilde{\ell}_t$ and $\tilde{\ell}_t = \tilde{\ell}_t(\boldsymbol{\theta}) = \frac{\boldsymbol{\epsilon}_t^2}{\tilde{\boldsymbol{\sigma}}_t^2} + \ln \tilde{\boldsymbol{\sigma}}_t^2$.

Let $\mathcal{A}_{\boldsymbol{\theta}}(z) = \sum_{i=1}^q \alpha_i z^i$ and $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1 - \sum_{j=1}^p \beta_j z^j$. It is not restrictive to assume that $q \geq 1$. By convention $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1$ if $p = 0$. Let $\mathcal{F}_{t-1}$ be the σ-field generated by $(\boldsymbol{\epsilon}_{t-1}, \boldsymbol{\epsilon}_{t-2}, \ldots)$. To show the strong consistency, the following assumptions will be made.

A3.1 $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$ and $\boldsymbol{\Theta}$ is compact.

A3.2 $\gamma(\mathbf{A}_0) < 0$ and $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\sum_{j=1}^p \beta_j < 1$.

A3.3 $(\boldsymbol{\eta}_t)$ is stationary and ergodic, $\boldsymbol{\eta}_t^2$ has a non-degenerate distribution with *i)* $\mathbb{E}[\boldsymbol{\eta}_t^2 \mid \mathcal{F}_{t-1}] = 1$ a.s. and *ii)* $E \ln \boldsymbol{\eta}_t^2 > -\infty$.

A3.4 If $p > 0$, $\mathcal{A}_{\boldsymbol{\theta}_0}(z)$ and $\mathcal{B}_{\boldsymbol{\theta}_0}(z)$ have no common root, $\mathcal{A}_{\boldsymbol{\theta}_0}(1) \neq 0$, and $\alpha_{0q} + \beta_{0p} \neq 0$

Remark 3.2. Assumptions **A3.1**, **A3.2** and **A3.4** are standard (see Francq and Zakoian (2004) for comments on these assumptions). Condition **A3.3** *i)* is obviously less restrictive than the iid assumption with finite second-order moments. In Appendix 5.5, we provide an explicit example of semi-strong GARCH based on a non-iid martingale difference innovation satisfying **A3.3** *i)*. This assumption was first used by Lee and Hansen (1994) for the inference of GARCH models. Escanciano (2009) established the consistency of the QMLE under this assumption, with a small-order moment condition on the observed process instead of our assumption **A3.3** *ii)*. Note that the latter assumption precludes densities with too much mass around zero, but is satisfied by most commonly used distributions. It is also weaker than the regularity condition on the $\boldsymbol{\eta}_t$ law ($\lim_{t\to 0} t^{-\mu}\mathbb{P}\{\boldsymbol{\eta}_0^2 \leqslant t\} = 0$, for some $\mu > 0$) used by Berkes et al. (2003) [1].

Assumption **A3.2** implies that the roots of $\mathcal{B}_{\boldsymbol{\theta}}(z)$ are outside the unit disc. Therefore, by the second inequality of (3.20), we can define $(\boldsymbol{\sigma}_t^2) = \{\boldsymbol{\sigma}_t^2(\boldsymbol{\theta})\}$ as the (unique) strictly stationary, ergodic and non-anticipative solution of

$$\boldsymbol{\sigma}_t^2 = \omega + \sum_{i=1}^q \alpha_i \boldsymbol{\epsilon}_{t-i}^2 + \sum_{j=1}^p \beta_j \boldsymbol{\sigma}_{t-j}^2, \quad \forall t, \tag{3.23}$$

see Appendix 5.6.

Note that $\boldsymbol{\sigma}_t^2(\boldsymbol{\theta}_0) = \boldsymbol{h}_t$. Let

$$\mathbf{l}_n(\boldsymbol{\theta}) = \mathbf{l}_n(\boldsymbol{\theta}; \boldsymbol{\epsilon}_n, \boldsymbol{\epsilon}_{n-1} \ldots,) = n^{-1}\sum_{t=1}^n \ell_t, \quad \ell_t = \ell_t(\boldsymbol{\theta}) = \frac{\boldsymbol{\epsilon}_t^2}{\boldsymbol{\sigma}_t^2} + \ln \boldsymbol{\sigma}_t^2.$$

We are now able to establish the strong consistency of the QMLE.

Theorem 3.3. *Let $(\hat{\boldsymbol{\theta}}_n)$ be a sequence of QMLE satisfying (3.22), with any initial condition (3.21). Then, under* **A3.1-A3.4**, $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ *a.s.* as $n \to \infty$.

1. Knowing that $\mathbb{E}\left(\ln^+(\boldsymbol{\eta}_1^2)\right) < \infty$ by **A3.3** *i)*, to establish **A3.3** *ii)* it is therefore sufficient to prove that $\mathbb{E}\left(\ln^-(\boldsymbol{\eta}_1^2)\right) < \infty$. Using $\mathbb{E}\left(\ln^-(\boldsymbol{\eta}_1^2)\right) = \int_0^\infty \mathbb{P}(\ln^+(\frac{1}{\boldsymbol{\eta}_1^2}) \geq s)ds = \int_0^\infty \mathbb{P}(\ln(\frac{1}{\boldsymbol{\eta}_1^2}) \geq s)ds = \int_0^\infty \mathbb{P}(\frac{1}{\boldsymbol{\eta}_1^2} \geq \exp(s))ds = \int_0^\infty \mathbb{P}(\boldsymbol{\eta}_1^2 \leq \exp(-s))ds$, we have under the condition of Berkes et al. (2003) that $\mathbb{P}(\boldsymbol{\eta}_1^2 \leq \exp(-s)) = o(\exp(-\mu s))$ when $s \to \infty$, which gives the result.

Remark 3.3. *Escanciano (2009) established the asymptotic normality of the QMLE under the assumption that a small-order moment exists. This moment condition is mainly used to justify the existence of the asymptotic covariance of the QMLE. To the best of our knowledge, the asymptotic normality has never been shown without an hypothesis that implies the existence of a small-order moment. In some cases, the asymptotic covariance matrix may not exist without a finite moment of sufficiently large order (see Francq and Zakoian (2007, Section 3.1)). The study of the asymptotic distribution of the semi-strong GARCH without any moment condition is left for future work.*

Proof of Theorem 3.3. The proof relies on the following intermediate results.

$$
\begin{aligned}
&i)\ \lim_{n\to\infty}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\mathbf{l}_n(\boldsymbol{\theta})-\tilde{\mathbf{l}}_n(\boldsymbol{\theta})|=0,\quad a.s.\\
&ii)\ \text{if}\ \ \sigma_t^2(\boldsymbol{\theta})=\sigma_t^2(\boldsymbol{\theta}_0)\quad a.s.,\quad \text{then}\ \ \boldsymbol{\theta}=\boldsymbol{\theta}_0,\\
&iii)\ \text{if}\ \ \boldsymbol{\theta}\neq\boldsymbol{\theta}_0,\ \ \text{then}\ \ \mathbb{E}\{\ell_1(\boldsymbol{\theta})-\ell_1(\boldsymbol{\theta}_0)\}>0,\\
&iv)\ \text{any}\ \boldsymbol{\theta}\neq\boldsymbol{\theta}_0\ \text{has a neighborhood}\ V(\boldsymbol{\theta})\ \text{such that}\\
&\qquad\qquad\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*)-\tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)\right)>0\quad a.s.
\end{aligned}
$$

To prove *i)*, note that Francq and Zakoian (2004, Equation (4.7)) show that almost surely,

$$
\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\mathbf{l}_n(\boldsymbol{\theta})-\tilde{\mathbf{l}}_n(\boldsymbol{\theta})\right|\leqslant\left\{\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\frac{1}{\omega^2}\right\}Cn^{-1}\sum_{t=1}^{n}\rho^t\boldsymbol{\epsilon}_t^2+\left\{\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\frac{1}{\omega}\right\}Cn^{-1}\sum_{t=1}^{n}\rho^t,
$$

for some constants $C>0$ and $0<\rho<1$ (independent of n). The point *i)* thus follows by Cesàro lemma, since the first inequality of (3.20) implies that $\rho^t\boldsymbol{\epsilon}_t^2\to 0$ a.s. as $t\to\infty$:

$$
\limsup_{n\to\infty}\frac{1}{k}\ln\rho^k\boldsymbol{\epsilon}_{t+k}^2\leq\ln\rho+\limsup_{n\to\infty}\frac{1}{k}\ln\boldsymbol{\epsilon}_{t+k}^2=\ln\rho<0.
$$

The proof of *ii)* uses the same arguments as those of step *ii)* in Theorem 2.1 of Francq and Zakoian (2004).

Now let us turn to the proof of *iii)*. For strong GARCH models it is known that $\mathbb{E}\ell_1(\boldsymbol{\theta}_0)$ is finite. This may not be the case in our framework. This is why we give an alternative proof of *iii*). We will first establish the existence of $\mathbb{E}\{\ell_1(\boldsymbol{\theta})-\ell_1(\boldsymbol{\theta}_0)\}$. Let $W_t(\boldsymbol{\theta})=\sigma_t^2(\boldsymbol{\theta}_0)/\sigma_t^2(\boldsymbol{\theta})$ and, for $K>0$, $A_K=[K^{-1},K]$, write

$$
\ell_t(\boldsymbol{\theta})-\ell_t(\boldsymbol{\theta}_0)=g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}+g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}
$$

where, for $x>0, y\geq 0$, $g(x,y)=-\log x+y(x-1)$. Introducing the negative part $x^-=\max(-x,0)$ of any real number x, we thus have

$$
\ell_t(\boldsymbol{\theta})-\ell_t(\boldsymbol{\theta}_0)\geq g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}-\left\{g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\right\}^-\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}.\tag{3.24}
$$

Noting that $W_t(\boldsymbol{\theta})$ is $\mathcal{F}_{t-1}$- measurable and by **A3.3** *i)*, $\mathbb{E}[g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mid\mathcal{F}_{t-1}]=$

$g(W_t(\boldsymbol{\theta}),1)$, the expectation of the first term in the r.h.s. of (3.24) is well-defined and satisfies

$$\mathbb{E}[g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] = \mathbb{E}[g(W_t(\boldsymbol{\theta}),1)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] \geq 0$$

since $g(x,1)\geq 0$ for any $x\geq 0$, with equality only if $x=1$. By *ii)* we have that $W_t(\boldsymbol{\theta})=1$ a.s. if and only if $\boldsymbol{\theta}=\boldsymbol{\theta}_0$. We thus have, by Beppo-Levi's theorem,

$$\begin{aligned}\lim_{K\to\infty}\mathbb{E}[g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}] &= \mathbb{E}[g(W_t(\boldsymbol{\theta}),1)\lim_{K\to\infty}\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K}]\\ &= \mathbb{E}[g(W_t(\boldsymbol{\theta}),1)] > 0 \quad \text{for} \quad \boldsymbol{\theta}\neq\boldsymbol{\theta}_0.\end{aligned}$$

To deal with the expectation of the second term in the r.h.s. of (3.24) we use the fact that for $y>0$, $g(x,y)\geq g(1/y,y)$. It follows that

$$-\mathbb{E}\left[\left\{g(W_t(\boldsymbol{\theta}),\boldsymbol{\eta}_t^2)\right\}^-\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}\right] \geq -\mathbb{E}\left[\left\{g(1/\boldsymbol{\eta}_t^2,\boldsymbol{\eta}_t^2)\right\}^-\mathbb{1}_{W_t(\boldsymbol{\theta})\in A_K^c}\right]\to 0 \quad \text{as } K\to\infty,$$

because, by **A3.3** *ii)*, $\mathbb{E}\left[\{g(1/\boldsymbol{\eta}_t^2,\boldsymbol{\eta}_t^2)\}^-\right] < \infty$ and thus the convergence holds by Lebesgue's dominated convergence theorem. This completes the proof of Step *iii)*.

Now we prove *iv)*. As for Step *iii)*, the possible non existence of $\mathbb{E}\ell_1(\boldsymbol{\theta})$, requires a modification of the standard proof. For any $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ we have

$$\tilde{\mathbf{l}}_n(\boldsymbol{\theta})-\tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0) \geq \mathbf{l}_n(\boldsymbol{\theta})-\mathbf{l}_n(\boldsymbol{\theta}_0)-|\tilde{\mathbf{l}}_n(\boldsymbol{\theta})-\mathbf{l}_n(\boldsymbol{\theta})|-|\tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)-\mathbf{l}_n(\boldsymbol{\theta}_0)|.$$

Hence, using *i)*

$$\begin{aligned}&\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*)-\tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)\right)\\ \geq&\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\mathbf{l}_n(\boldsymbol{\theta}^*)-\mathbf{l}_n(\boldsymbol{\theta}_0)\right)-2\limsup_{n\to\infty}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\tilde{\mathbf{l}}_n(\boldsymbol{\theta})-\mathbf{l}_n(\boldsymbol{\theta})|\\ =&\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\mathbf{l}_n(\boldsymbol{\theta}^*)-\mathbf{l}_n(\boldsymbol{\theta}_0)\right).\end{aligned} \tag{3.25}$$

For any $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ and any positive integer k, let $V_k(\boldsymbol{\theta})$ the open ball of center $\boldsymbol{\theta}$ and radius $1/k$. We have

$$\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\mathbf{l}_n(\boldsymbol{\theta}^*)-\mathbf{l}_n(\boldsymbol{\theta}_0)\right)\geq\liminf_{n\to\infty}\frac{1}{n}\sum_{t=1}^n\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\ell_t(\boldsymbol{\theta}^*)-\ell_t(\boldsymbol{\theta}_0). \tag{3.26}$$

By arguments already given, under **A3.3** *ii)*,

$$\mathbb{E}\left(\inf_{\boldsymbol{\theta}^*\in V_k(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\ell_t(\boldsymbol{\theta}^*)-\ell_t(\boldsymbol{\theta}_0)\right)^- \leq \mathbb{E}\left(g(1/\boldsymbol{\eta}_t^2,\boldsymbol{\eta}_t^2)\right)^- < \infty.$$

Therefore $\mathbb{E}\left(\inf_{\boldsymbol{\theta}^* \in V_k(\boldsymbol{\theta}) \cap \boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)\right)$ exists in $\mathbb{R} \cup \{+\infty\}$, and the ergodic theorem applies (see Francq and Zakoian (2019, Exercises 7.3 and 7.4)). From (3.26) we obtain

$$\liminf_{n\to\infty} \left(\inf_{\boldsymbol{\theta}^* \in V_k(\boldsymbol{\theta}) \cap \boldsymbol{\Theta}} \mathbf{l}_n(\boldsymbol{\theta}^*) - \mathbf{l}_n(\boldsymbol{\theta}_0) \right) \geq \mathbb{E} \left(\inf_{\boldsymbol{\theta}^* \in V_k(\boldsymbol{\theta}) \cap \boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0) \right).$$

The latter term into parentheses converges to $\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0)$ as $k \to \infty$, and, by standard arguments using the positive and negative parts of $\inf_{\boldsymbol{\theta}^* \in V_k(\boldsymbol{\theta}) \cap \boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0)$, we have that

$$\lim_{k\to\infty} \mathbb{E} \left(\inf_{\boldsymbol{\theta}^* \in V_k(\boldsymbol{\theta}) \cap \boldsymbol{\Theta}} \ell_t(\boldsymbol{\theta}^*) - \ell_t(\boldsymbol{\theta}_0) \right) = \mathbb{E} \left\{ \ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0) \right\},$$

which by *i)* is strictly positive. In view of (3.25), the proof of *iv)* is complete.

Now we complete the proof of the theorem. The set $\boldsymbol{\Theta}$ is covered by the union of an arbitrary neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ and, for any $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, by neighborhoods $V(\boldsymbol{\theta})$ satisfying *iv)*. Obviously, $\inf_{\boldsymbol{\theta}^* \in V(\boldsymbol{\theta}_0) \cap \boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) \leq \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)$, $a.s.$ Moreover, by compactness of $\boldsymbol{\Theta}$, there exists a finite subcover of the form $V(\boldsymbol{\theta}_0), V(\boldsymbol{\theta}_1), \ldots, V(\boldsymbol{\theta}_M)$. By *iv)*, for $i = 1, \ldots, M$, there exists n_i such that for $n \geq n_i$,

$$\inf_{\boldsymbol{\theta}^* \in V(\boldsymbol{\theta}_i) \cap \boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) > \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0), \quad a.s.$$

Thus for $n \geq \max_{i=1,\ldots,M}(n_i)$,

$$\inf_{\boldsymbol{\theta}^* \in \bigcup_{i=1,\ldots,M} V(\boldsymbol{\theta}_i) \cap \boldsymbol{\Theta}} \tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) > \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0), \quad a.s.$$

from which we deduce that $\widehat{\boldsymbol{\theta}}_n$ belongs to $V(\boldsymbol{\theta}_0)$ for sufficiently large n.

□

5 Appendix: Complementary Proofs

5.1 Proof of Theorem 3.1

Proof. For all $t \in \mathbb{Z}$ and $n \in \mathbb{N}$, let

$$\boldsymbol{X}_{t,n} = \Psi\left(\boldsymbol{\theta}_t, \boldsymbol{X}_{t-1,n-1}\right) \tag{3.27}$$

with $\boldsymbol{X}_{t,0} = c$. Note that

$$\boldsymbol{X}_{t,n} = \psi_n\left(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t-1}, \ldots, \boldsymbol{\theta}_{t-n+1}\right)$$

for some measurable function $\psi_n : (E^n, \mathcal{B}_{E^n}) \to (F, \mathcal{B}_F)$, with the usual notation. For all n, the sequence $(\boldsymbol{X}_{t,n})_{t\in\mathbb{Z}}$ is thus stationary and ergodic. If for all t, the limit $\boldsymbol{X}_t = \lim_{n\to\infty} \boldsymbol{X}_{t,n}$ exists a.s., then by taking the limit of both sides of Equation (3.27), it can be seen that the process $(\boldsymbol{X}_t)$ is solution of Equation (3.1). When it exists, the limit is

a measurable function of the form $\boldsymbol{X}_t = \psi(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t-1}, \ldots)$ [2] and is therefore stationary and ergodic. The existence of $\lim_{n\to\infty} \boldsymbol{X}_{t,n}$ has been proven by Elton (1990). He showed that, a.s., the sequence $(\boldsymbol{X}_{t,n})_{n\in\mathbb{N}}$ is a Cauchy sequence in the complete space F.

By iterating equation (3.27) we have

$$\boldsymbol{X}_{t,n} = \boldsymbol{\Psi}_t \circ \cdots \circ \boldsymbol{\Psi}_{t-n+1}(c).$$

If follows that

$$d(\boldsymbol{X}_{t,n}, \boldsymbol{X}_{t,n-1}) \leq \boldsymbol{\Lambda}_t^{(n-1)} d(\boldsymbol{\Psi}_{t-n+1}(c), c).$$

For $n < m$, we thus have

$$\begin{aligned} d(\boldsymbol{X}_{t,m}, \boldsymbol{X}_{t,n}) &\leq \sum_{k=0}^{m-n-1} d(\boldsymbol{X}_{t,m-k}, \boldsymbol{X}_{t,m-k-1}) \\ &\leq \sum_{k=0}^{m-n-1} \boldsymbol{\Lambda}_t^{(m-k-1)} d(\boldsymbol{\Psi}_{t-m+k+1}(c), c) \\ &\leq \sum_{j=n}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d(\boldsymbol{\Psi}_{t-j}(c), c). \end{aligned} \tag{3.28}$$

Note that

$$\limsup_{j\to\infty} \ln\left(\boldsymbol{\Lambda}_t^{(j)} d(\boldsymbol{\Psi}_{t-j}(c), c)\right)^{1/j} = \limsup_{j\to\infty} \frac{1}{j}\left(\ln \boldsymbol{\Lambda}_t^{(j)} + \ln d(\boldsymbol{\Psi}_{t-j}(c), c)\right) < 0$$

under (i) and (ii), by using Kingman's sub-additive ergodic theorem (see Kingman (1973)) and Francq and Zakoian (2019, Exercises 4.12). We conclude, from the Cauchy criterion for the convergence of series with positive terms, that

$$\sum_{j=1}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d(\boldsymbol{\Psi}_{t-j}(c), c)$$

is a.s. finite, under (i) and (ii). It follows that $(\boldsymbol{X}_{t,n})_{n\in\mathbb{N}}$ is a.s. a Cauchy sequence in F. The existence of a stationary and ergodic solution to Equation (3.1) follows.

Assume that there exists another stationary process $(\boldsymbol{X}_t^*)$ such that $\boldsymbol{X}_t^* = \boldsymbol{\Psi}_t(\boldsymbol{X}_{t-1}^*)$. For all $N \geq 0$, we have

$$d(\boldsymbol{X}_t, \boldsymbol{X}_t^*) \leq \boldsymbol{\Lambda}_t^{(N+1)} d(\boldsymbol{X}_{t-N}, \boldsymbol{X}_{t-N}^*). \tag{3.29}$$

Since $\boldsymbol{\Lambda}_t^{(N+1)} \to 0$ a.s. as $N \to \infty$, and $d(\boldsymbol{X}_{t-N}, \boldsymbol{X}_{t-N}^*) = O_P(1)$ by stationarity, the right-hand side of Equation (3.29) tends to zero in probability. Since the left-hand side does not depend on N, we have $\mathbb{P}(d(\boldsymbol{X}_t, \boldsymbol{X}_t^*) > \epsilon) = 0$ for all $\epsilon > 0$, and thus

2. For the measurability of $\boldsymbol{X}_t$, one can consider $\boldsymbol{X}_{t,n}$ as functions of $(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t-1}, \cdots)$ and argue that in metric space, a limit of measurable functions is measurable.

$\mathbb{P}\left(\boldsymbol{X}_t = \boldsymbol{X}_t^*\right) = 1$, which establishes the uniqueness. In view of Equation (3.28), we have

$$d\left(\boldsymbol{X}_t, c\right) \leq \sum_{j=0}^{\infty} \boldsymbol{\Lambda}_t^{(j)} d\left(\boldsymbol{\Psi}_{t-j}\left(c\right), c\right)$$

and Equation (3.2) follows. □

5.2 Proof of the comment of Theorem 3.2

For all $\epsilon > 0$, since $\mathbb{P}(\ln d(\boldsymbol{X}_1, c) > \epsilon) = \mathbb{P}(\ln^+ d(\boldsymbol{X}_1, c) > \epsilon)$, then

$$\begin{aligned}
\sum_{n=1}^{\infty} P\left(n^{-1} \ln d(\boldsymbol{X}_{t+n}, c) > \epsilon\right) &= \sum_{n=1}^{\infty} P\left(n^{-1} \ln^+ d(\boldsymbol{X}_1, c) > \epsilon\right) \\
&\leq \int_0^{\infty} P\left(t^{-1} \ln^+ d(\boldsymbol{X}_1, c) > \epsilon\right) dt \\
&= \int_0^{\infty} P\left(\epsilon^{-1} \ln^+ d(\boldsymbol{X}_1, c) > t\right) dt \\
&= \epsilon^{-1} E \ln^+ d(\boldsymbol{X}_1, c) < \infty.
\end{aligned}$$

It follows by the Borel–Cantelli lemma that $\limsup n^{-1} \ln d(\boldsymbol{X}_{t+n}, c) \leq 0$ *a.s.* The second result is obtained by the same arguments.

5.3 Proof of Equation (3.5)

The condition on $\boldsymbol{Z}_1$ implies that $\mathbb{P}(\boldsymbol{Z}_1 > K) > 0$ for some real K. It follows from Birkhoff's ergodic theorem that $n^{-1} \sum_{k=1}^{n} \mathbb{1}_{\{\boldsymbol{Z}_n \geq K\}} \longrightarrow \mathbb{P}\left(\boldsymbol{Z}_1 \geq K\right) > 0$ *a.s. as* $n \to \infty$. Thus, almost surely, the set $\{n : \boldsymbol{Z}_n \geq K\}$ is not finite. This implies that $\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{Z}_n = \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{Z}_n \mathbb{1}_{\{\boldsymbol{Z}_n \geq K\}}$ *a.s.* Since $\frac{1}{n} \boldsymbol{Z}_n \mathbb{1}_{\{K \leq \boldsymbol{Z}_n \leq 0\}} \longrightarrow 0$ *a.s as* $n \to \infty$, it follows that $\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{Z}_n = \limsup_{n\to\infty} \frac{1}{n} \boldsymbol{Z}_n^+$ *a.s.* Therefore, Eq. (3.5) follows from Tanny (1974, Theorem 1).

5.4 Proof of Corollary 3.1

We have for all $n \geq 1$,

$$\sup_{k \geq n} \max(0, \ln d(\boldsymbol{X}_{t+k}, c)) = \max(0, \sup_{k \geq n} \ln d(\boldsymbol{X}_{t+k}, c)).$$

It follows that

$$\limsup_n \frac{1}{n} \ln^+ d(\boldsymbol{X}_{t+n}, c) = \max(0, \limsup_n \frac{1}{n} \ln d(\boldsymbol{X}_{t+n}, c)) = 0 \ a.s.$$

Since, in addition, $\ln^+ d(\boldsymbol{X}_{t+n}, c)$ is non-negative, $\lim_{n\to\infty} \frac{1}{n} \ln^+ d(\boldsymbol{X}_{t+n}, c)$ exists and is equal to 0 a.s. We get $\lim_{n\to\infty} \frac{1}{n} \ln^+ d(\boldsymbol{X}_{t-n}, c)$ by the same arguments. which gives the first part of the corollary.

For (3.6), we have $\ln d(\boldsymbol{X}_{t+n}, c) = \ln^+ d(\boldsymbol{X}_{t+n}, c) - \ln^- d(\boldsymbol{X}_{t+n}, c)$. Since $\frac{1}{|n|} \ln^+ d(\boldsymbol{X}_{t-n}, c)$ converges a.s to 0 and $\frac{1}{|n|} \ln^- d(\boldsymbol{X}_{t-n}, c)$ also converges a.s to 0 as $|n| \to \infty$ (see for instance Francq and Zakoian (2019, Exercice 2.13)) then $\frac{1}{|n|} \ln d(\boldsymbol{X}_{t+n}, c)$ converges a.s to 0 as $|n| \to \infty$.

5.5 Construction of a semi-strong GARCH

We first define a non iid martingale difference process. Consider a sequence $(\boldsymbol{x}_t)_{t\in\mathbb{Z}}$ of iid random variables with standard normal distribution. Since for all $z \in \mathbb{R}_+$, $\boldsymbol{x}_t\sqrt{2z} - z \sim \mathcal{N}(-z, 2z)$, using the moment-generating function of the Gaussian distribution, we have

$$\mathbb{E}\left[\exp(\boldsymbol{x}_t\sqrt{2z} - z)\right] = 1. \tag{3.30}$$

If $(\boldsymbol{z}_t)$ is a positive process, independent of $(\boldsymbol{x}_t)$, we also have $\mathbb{E}\boldsymbol{\eta}_t^2 = 1$, where $\boldsymbol{\eta}_t^2 = \exp\left(\boldsymbol{x}_t\sqrt{2\boldsymbol{z}_t} - \boldsymbol{z}_t\right)$. This is the case if, for instance, $\boldsymbol{z}_t$ follows a causal AR(1) model of the form $\boldsymbol{z}_t = \phi\boldsymbol{z}_{t-1} + \boldsymbol{u}_t$ with $\phi \in (0, 1)$ and $\boldsymbol{u}_t$ iid with positive variance. It is easy to see that $\mathrm{Cov}(\boldsymbol{z}_1, \boldsymbol{z}_0) \neq 0$, and thus

$$\begin{aligned}\mathrm{Cov}\left\{\ln(\boldsymbol{\eta}_1^2), \ln(\boldsymbol{\eta}_0^2)\right\} &= 2\mathbb{E}\{\boldsymbol{x}_1\sqrt{\boldsymbol{z}_1}\boldsymbol{x}_0\sqrt{\boldsymbol{z}_0}\} - \mathbb{E}\{\boldsymbol{x}_1\sqrt{2\boldsymbol{z}_1}\boldsymbol{z}_0\} \\ &\quad - \mathbb{E}\{\boldsymbol{z}_1\boldsymbol{x}_0\sqrt{2\boldsymbol{z}_0}\} + \mathbb{E}\{\boldsymbol{z}_1\boldsymbol{z}_0\} - \mathbb{E}\boldsymbol{z}_1\mathbb{E}\boldsymbol{z}_0 \\ &= \mathrm{Cov}\left\{\boldsymbol{z}_1, \boldsymbol{z}_0\right\} \neq 0.\end{aligned}$$

It follows that $(\boldsymbol{\eta}_t^2)$ is not iid. We now define $(\boldsymbol{\eta}_t)$. Let $(\boldsymbol{r}_t)$ be an iid sequence of Rademacher variables (uniform distribution on $\{-1, 1\}$), independent of the two sequences $(\boldsymbol{x}_t)$ and $(\boldsymbol{u}_t)$. We thus define $(\boldsymbol{\eta}_t)$ by $\boldsymbol{\eta}_t = \boldsymbol{r}_t\sqrt{\boldsymbol{\eta}_t^2}$.

Let $(\mathcal{F}_t)$ be the canonical filtration of $(\boldsymbol{\eta}_t)$, i.e. $\mathcal{F}_t = \sigma\left(\boldsymbol{\eta}_k, k \leq t\right)$. Define a second filtration $\mathcal{H}_t = \sigma\left(\boldsymbol{r}_k, \boldsymbol{x}_{k+1}, \boldsymbol{u}_{k+1}, k \leq t\right)$. Since $\mathcal{F}_t \subset \mathcal{H}_t$ and $\boldsymbol{r}_t$ is independent of $\mathcal{H}_{t-1}$, we have

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{\eta}_t \mid \mathcal{F}_{t-1}\right] &= \mathbb{E}\{\mathbb{E}\left[\boldsymbol{\eta}_t \mid \mathcal{H}_{t-1}\right] \mid \mathcal{F}_{t-1}\} \\ &= \mathbb{E}\{\exp([\boldsymbol{x}_t\sqrt{2\boldsymbol{z}_t} - \boldsymbol{z}_t]/2)\mathbb{E}\left[\boldsymbol{r}_t \mid \mathcal{H}_{t-1}\right] \mid \mathcal{F}_{t-1}\} \\ &= 0.\end{aligned}$$

Define a new filtration $\mathcal{I}_t = \sigma\left(\boldsymbol{r}_k, \boldsymbol{x}_k, \boldsymbol{u}_{k+1}, k \leq t\right)$. Since $\mathcal{F}_t \subset \mathcal{I}_t$, $\boldsymbol{z}_t$ is $\mathcal{I}_{t-1}$-measurable, and $\boldsymbol{x}_t$ is independent of $\mathcal{I}_{t-1}$, by (3.30) we have

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{\eta}_t^2 \mid \mathcal{F}_{t-1}\right] &= \mathbb{E}\{\mathbb{E}\left[\boldsymbol{\eta}_t^2 \mid \mathcal{I}_{t-1}\right] \mid \mathcal{F}_{t-1}\} \\ &= \mathbb{E}\{\mathbb{E}\left[\exp(\boldsymbol{x}_t\sqrt{2\boldsymbol{z}_t} - \boldsymbol{z}_t) \mid \mathcal{I}_{t-1}\right] \mid \mathcal{F}_{t-1}\} = 1.\end{aligned}$$

We thus have shown the existence of a non degenerate unit martingale difference

sequence, that is a stationary and ergodic sequence $(\boldsymbol{\eta}_t)$ satisfying the conditions

$$\mathbb{E}\left[\boldsymbol{\eta}_t^2\right] < \infty, \quad \mathbb{E}\left[\boldsymbol{\eta}_t \mid \mathcal{F}_{t-1}\right] = 0, \quad \mathbb{E}\left[\boldsymbol{\eta}_t^2 \mid \mathcal{F}_{t-1}\right] = 1, \quad \left(\boldsymbol{\eta}_t^2\right) \text{ are not iid.}$$

It is then easy to define a semi-strong GARCH with innovations $(\boldsymbol{\eta}_t)$.

5.6 Proof of the existence of a unique strictly stationary solution to (3.23)

Rewrite (3.23) in vector form as

$$\underline{\boldsymbol{\sigma}}_t^2 = \underline{\boldsymbol{c}}_t + B\underline{\boldsymbol{\sigma}}_{t-1}^2,$$

where

$$\underline{\boldsymbol{\sigma}}_t^2 = \begin{pmatrix} \boldsymbol{\sigma}_t^2 \\ \boldsymbol{\sigma}_{t-1}^2 \\ \vdots \\ \boldsymbol{\sigma}_{t-p+1}^2 \end{pmatrix}, \quad \underline{\boldsymbol{c}}_t = \begin{pmatrix} \omega + \sum_{i=1}^q \alpha_i \boldsymbol{\epsilon}_{t-i}^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \\ 1 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

we have by the second inequality of (3.20) that $\limsup_{n\to\infty} \frac{1}{n} \ln \|\underline{\boldsymbol{c}}_n\| = 0$. By Assumption **A3.2**, we deduce that

$$\limsup_{n\to\infty} \frac{1}{n} \ln \|B^n \underline{\boldsymbol{c}}_{n-1}^2\| \leq \limsup_{n\to\infty} \frac{1}{n} \ln \|B^n\| + \limsup_{n\to\infty} \frac{1}{n} \ln \|\underline{\boldsymbol{c}}_n\| < 0.$$

From this, we deduce by the Cauchy rule that the series $\hat{\underline{\boldsymbol{\sigma}}}_t^2 := \sum_{n=0}^{\infty} B^n \underline{\boldsymbol{c}}_{t-n}^2$ converges almost surely. We note that $(\hat{\underline{\boldsymbol{\sigma}}}_t^2)$ is a strictly stationary, ergodic and non-anticipative solution of (3.23).

To show the uniqueness, assume that there exists another stationary process $(\underline{\boldsymbol{\sigma}}_t^2 *)$ of (3.23). For all $n \geq 0$, we have $\|\underline{\boldsymbol{\sigma}}_t^2 * - \hat{\underline{\boldsymbol{\sigma}}}_t^2\| = \|B^n \underline{\boldsymbol{\sigma}}_{t-n}^2 * - B^n \hat{\underline{\boldsymbol{\sigma}}}_{t-n}^2\| \leq \|B^n\| \|\underline{\boldsymbol{\sigma}}_{t-n}^2 *\| + \|B^n\| \|\hat{\underline{\boldsymbol{\sigma}}}_{t-n}^2\|$. Since $\|B^n\| \to 0$ a.s. as $n \to \infty$ and $\|\underline{\boldsymbol{\sigma}}_{t-n}^2 *\|$ and $\|\hat{\underline{\boldsymbol{\sigma}}}_{t-n}^2\|$ converges in law by stationary, Slutsky's theorem entails that $\|\underline{\boldsymbol{\sigma}}_t^2 * - \hat{\underline{\boldsymbol{\sigma}}}_t^2\|$ converges in law to 0 as $n \to \infty$. Since $\|\underline{\boldsymbol{\sigma}}_t^2 * - \hat{\underline{\boldsymbol{\sigma}}}_t^2\|$ does not depend on n, we conclude that $\|\underline{\boldsymbol{\sigma}}_t^2 * - \hat{\underline{\boldsymbol{\sigma}}}_t^2\| = 0$ *a.s.*

Chapter 4

Inference on GARCH-MIDAS models without any small-order moment

A version of this chapter, co-authored with Christian Francq and Jean-Michel Zakoïan, is forthcoming in *Econometric Theory*.

Abstract

In GARCH-mixed-data sampling (GARCH-MIDAS) models, the volatility is decomposed into the product of two factors, which are often interpreted as « short run » (high frequency) and « long run » (low frequency) components. While two-component volatility models are widely used in applied works, some of their theoretical properties remain unexplored. We show that the strictly stationary solutions of such models do not admit any small-order finite moment, contrary to classical GARCH. It is shown that the strong consistency and the asymptotic normality of the Quasi-Maximum Likelihood estimator hold despite the absence of moments. Tests for the presence of a long-run volatility relying on the asymptotic theory and a bootstrap procedure are proposed. Our results are illustrated via Monte Carlo experiments and real financial data.

Contents

1 Introduction 48
2 Stationarity and existence of moments 49
3 QMLE of GARCH-MIDAS 52
4 Testing the existence of a long-run volatility 55
5 Numerical results 59
6 Conclusion 70
7 Appendix: proofs 70

1 Introduction

Despite their ability to capture a number of empirical characteristics of financial returns, the restrictive features of « one-factor » classical GARCH models are well known. The parameter β in a GARCH(1,1) has to be close to 1 to ensure high volatility persistence, but this may induce undesirable restrictions on the marginal distribution of the returns. Moreover, parameters governing the short-run effect of shocks (α in the usual GARCH(1,1) parametrization, as in the equation of σ_t^2 in Model (4.1) below) also impact the long-run response through the coefficients ($\alpha\beta^i$) of the asymptotic expansion of the volatility as a function of the past squared returns. This lack of flexibility, in particular the necessity to disentangle short and long run impacts of shocks, has motivated the introduction of alternative volatility specifications in the econometric and finance literatures. Additive component GARCH models were introduced by Ding and Granger (1996), and Engle and Lee (1999) but, in recent years, multiplicative component GARCH processes have attracted more attention. In such models, called GARCH-mixed-data sampling (GARCH-MIDAS), the volatility is decomposed into the product of two factors which may receive different interpretations, generally in terms of « short run » (high frequency) and « long run » (low frequency) components. To cite just a few recent references, the reader is referred to Engle et al. (2013), Wang and Ghysels (2015), Amado and Teräsvirta (2017), Conrad et al. (2018), Conrad and Engle (2021).

While GARCH-MIDAS volatility models are widely used in applied works, some of their theoretical properties remain unexplored. An exception is the paper by Wang and Ghysels (2015) who consider stationarity and ergodicity, as well as asymptotic theory for the Quasi-Maximum Likelihood (QML) estimator, under assumptions we will further discuss. In this chapter, we consider three issues: first, the existence of small-order moments for the strictly stationary solution of the two-component volatility model, second, the consistency and asymptotic normality of the QML estimator, and third, testing the existence of a long-run volatility. The first two issues are closely related because all existing proofs of the consistency and asymptotic normality of QML estimators in standard GARCH models rely on the existence of small-order moments. The third issue was also considered by Conrad and Schienle (2020) who proposed a score-based test in a general multiplicative component model.

One characteristic of most commonly used GARCH-type models is that strict stationarity entails the existence of a small-order moment. Hence, even if stationary solutions (r_t) of standard GARCH models are generally characterized by heavy-tails (a desirable property for the modelling of financial returns), there exists a sufficiently small power s (depending on both the volatility parameters and the innovations distribution) such that $\mathbb{E}|r_t|^s < \infty$. In a sense, this means that such one-factor volatility models are too constrained, as the conditions ensuring stability of the dynamics produce unexpected restrictions on the marginal distributions. By contrast, the models we consider in this paper have the surprising property of admitting strictly stationary solutions that do not have any power moment (unless a very restrictive condition is imposed on the errors distribution). This heavyness of the tails of the marginal distribution entails formidable statistical difficulties for proving the consistency and asymptotic normality of the QML estimator.

Indeed, the existence of a small moment for the observed process is crucial to derive the asymptotic properties of the QMLE in most GARCH-type models (see for instance Francq and Zakoian (2019, Section 7.4)). In particular, contrary to the standard GARCH case, the proof of the consistency cannot rely on the existence of a limiting QML criterion. To circumvent the absence of moments, we use a property of exponential control of the trajectories which will be detailed below.

The rest of the chapter is organized as follows. In the next section we study the existence of strictly stationary solutions to the GARCH-MIDAS volatility model and their moment properties. Section 3 considers the estimation by QML of the model parameters. In Section 4 we propose tests for the existence of a long-run volatility. Two approaches are considered to handle the problem of unidentified parameters under the null and bootstrap procedures are proposed. Numerical and empirical results are presented in Section 5. Section 6 concludes. Proofs are given in the Appendix.

2 Model and an unexpected property of the stationary solution

We study in this chapter a class of GARCH-MIDAS processes (r_t) defined by

$$\left\{ \begin{array}{ll} r_t = \tau_t \epsilon_t, & \tau_t^2 = 1 + a_0 \sum_{i=1}^{Q} \varphi_i(\boldsymbol{\vartheta}_0) RV_{t-i}, \\ \epsilon_t = \sigma_t \eta_t, & \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \end{array} \right. , \tag{4.1}$$

where σ_t is the positive square root of σ_t^2, (η_t) is an iid sequence with $\mathbb{E}\eta_t^2 = 1$, $a_0 \geq 0$, $\omega_0 > 0$, $\alpha_0 \geq 0$ and $\beta_0 \geq 0$, $RV_t = \sum_{i=0}^{N-1} r_{t-i}^2$ is a rolling window realized volatility, Q and N are positive integers, and $\varphi_i(\boldsymbol{\vartheta}_0)$ are positive weights, depending on some d-variate parameter $\boldsymbol{\vartheta}_0$. For the specification of the functions φ_i used to smooth the realized volatilities Engle et al. (2013), under a slightly different parametrization [1], suggest exponential weights

$$\varphi_i(\vartheta_0) = \frac{\vartheta_0^i}{\sum_{j=1}^{Q} \vartheta_0^i}, \quad \vartheta_0 \in (0, \infty) \tag{4.2}$$

or Beta weights

$$\varphi_i(\vartheta_0) = \frac{\{1 - i/(Q+1)\}^{\vartheta_0 - 1}}{\sum_{j=1}^{Q} \{1 - j/(Q+1)\}^{\vartheta_0 - 1}}, \quad \vartheta_0 \in (0, \infty). \text{ [2]} \tag{4.3}$$

1. Engle et al. (2013) considered a unit-variance GARCH(1,1) equation, $\sigma_t^2 = 1 - \alpha_0 - \beta_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2$ for the short-run volatility and introduced an intercept m in the equation of τ_t^2. This choice is guided by the necessity to identify short- and long-run volatilities. The alternative identifiability condition we adopt here is a unit intercept, $m = 1$, in the long-term volatility dynamics. This constraint is not restrictive, whereas imposing a unit-variance for the short-run volatility requires $\alpha_0 + \beta_0 < 1$, which is not necessary for strict stationarity. Note that Engle et al. (2013) also allow for an intercept in the equation of r_t.

2. In these examples, the weight parameter ϑ_0 is scalar, and therefore is not shown in bold.

The standard GARCH(1,1) is obtained for $a_0 = 0$. For $a_0 > 0$, the volatility component τ_t^2 is often referred to as the *long-run volatility* (for large q), while the *short-run volatility* σ_t^2 is a function of the normalized (long-run detrended) squared returns r_{t-i}^2/τ_{t-i}^2.

Model (4.1) can be written under the following form, which will be used throughout,

$$\begin{cases} r_t = \tau_t \epsilon_t, & \tau_t^2 = 1 + a_0 \sum_{i=1}^q \phi_i(\boldsymbol{\vartheta}_0) r_{t-i}^2, \\ \epsilon_t = \sigma_t \eta_t, & \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \end{cases} \tag{4.4}$$

where the $\phi_i(\boldsymbol{\vartheta}_0)$'s are nonnegative, with at least one stricly positive coefficient. Model (4.4) is the model we focus on, and is more general than the GARCH-MIDAS for which we have $q = N + Q - 1$. Without loss of generality assume that $\sum_{i=1}^q \phi_i(\boldsymbol{\vartheta}_0) = 1$.

Next, we turn to the existence of strictly stationary solutions to Model (4.4).

Let $\delta_t = \alpha_0 \eta_t^2 + \beta_0$. Under the assumption

A4.1 $\gamma := \mathbb{E} \log \delta_1 < 0$,

the GARCH(1,1) equation in (4.4) admits the strictly stationary, non anticipative and ergodic solution

$$\epsilon_t = \sigma_t \eta_t, \quad \sigma_t^2 = \omega_0 \left(1 + \sum_{i=1}^{\infty} \prod_{j=1}^{i} \delta_{t-j} \right). \tag{4.5}$$

Note that **A4.1** is less restrictive than the condition $\alpha_0 + \beta_0 < 1$ used in Wang and Ghysels (2015).

It is known that, for $r > 0$,

$$\mathbb{E}(\sigma_t^{2r}) < \infty \text{ if and only if } \mathbb{E}\delta_1^r < 1. \tag{4.6}$$

Note that σ_t, and thus ϵ_t, cannot admit moments of any order when δ_t is not almost surely bounded by 1, *i.e.* when

A4.2 $P(\delta_1 > 1) \neq 0$.

Indeed, for $\iota > 0$ such that $P(\delta_1 > 1 + \iota) > 0$, we have

$$\mathbb{E}\delta_1^r \geq (1 + \iota)^r P(\delta_1 > 1 + \iota) \to \infty$$

as $r \to \infty$. Note that **A4.2** is satisfied when η_t^2 is not bounded and $\alpha_0 \neq 0$.[3] It follows from (4.6) that $\mathbb{E}(\sigma_t^{2r}) = \infty$ for r large enough. This is a well-known property dating back to Kesten (1973), see also Mikosch and Starica (2000).

Write (4.4) in matrix form as

$$\boldsymbol{r}_t = \boldsymbol{A}_t \boldsymbol{r}_{t-1} + \boldsymbol{b}_t, \tag{4.7}$$

3. This assumption is therefore very mild. Moreover, it can be verified in practice by estimating $P(\delta_1 > 1)$.

where $\boldsymbol{r}_t = (r_t^2, \ldots, r_{t-q+1}^2)'$, $\boldsymbol{b}_t = (\epsilon_t^2, \mathbf{0}'_{q-1})'$ and $\boldsymbol{A}_t = \boldsymbol{A}(\epsilon_t)$ is a companion-like matrix:

$$\boldsymbol{A}_t = \begin{pmatrix} a_0\phi_1(\boldsymbol{\vartheta}_0)\epsilon_t^2 & \ldots & a_0\phi_{q-1}(\boldsymbol{\vartheta}_0)\epsilon_t^2 & a_0\phi_q(\boldsymbol{\vartheta}_0)\epsilon_t^2 \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{pmatrix}.$$

Noting that, under **A4.1**, the sequence $(\boldsymbol{A}_t, \boldsymbol{b}_t)$ is strictly stationary and ergodic, Equation (4.7) admits, by Brandt (1986, Theorem 1), the strictly stationary solution

$$\boldsymbol{r}_t = \boldsymbol{b}_t + \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i} \boldsymbol{A}_{t+1-j} \right) \boldsymbol{b}_{t-i} \tag{4.8}$$

under the assumption

A4.3 $\gamma_{\boldsymbol{A}} < 0$, where $\gamma_{\boldsymbol{A}} = \lim_{k\to\infty} \frac{1}{k} \mathbb{E} \log \|\boldsymbol{A}_k \boldsymbol{A}_{k-1} \ldots \boldsymbol{A}_1\| < 0$.

Note that the top-Lyapounov exponent $\gamma_{\boldsymbol{A}}$ involved in **A4.3** is well defined in $[-\infty, \infty)$ because $\mathbb{E} \log^+ \|\boldsymbol{A}_t\| < \infty$, in view of (4.9) below. Wang and Ghysels (2015) obtained explicit conditions entailing **A4.3** for particular sub-models. The next assumption guarantees that the long and short-run volatilities τ_t and σ_t are not degenerate.

A4.4 $a_0 > 0$ and $\alpha_0 > 0$.

According to Lemma 2.3 in Berkes et al. (2003), the strictly stationary solution ϵ_t of the standard GARCH(1,1) equation satisfies

$$\mathbb{E}|\epsilon_t|^s < \infty \quad \text{for some } s > 0. \tag{4.9}$$

The following proposition shows that, surprisingly, this feature does not extend to the solution (r_t) of the GARCH-MIDAS model (4.4).

We start by proving the following lemma, of independent interest as it concerns the GARCH(1,1) process (ϵ_t).

Lemma 4.1. *Assume **A4.1**-**A4.2**. For all integer $k \geq 2$, all real numbers $p_j > 0$ and integers i_j, $j = 1, \ldots k$, there exists $K \in (0, \infty]$ such that*

$$\mathbb{E}|\epsilon_{t-i_1}|^{p_1} |\epsilon_{t-i_1-i_2}|^{p_2} \ldots |\epsilon_{t-i_1-\cdots-i_k}|^{p_k} \geq K \mathbb{E}|\epsilon_1|^{p_1+\cdots+p_k}.$$

The right-hand side, and thus the left-hand side, of the inequality is infinite when $p_1 + \cdots + p_k$ is large enough.

Proposition 4.1. *There exists a strictly stationary and ergodic solution (r_t) to (4.4) if and only if both **A4.1** and **A4.3** are satisfied. If in addition **A4.2** and **A4.4** hold, this solution does not admit any moments, in the sense that*

$$\mathbb{E}|r_t|^s = \infty \quad \textit{for all } s > 0. \tag{4.10}$$

Note that Wang and Ghysels (2015) showed that $\mathbb{E}|r_t|^2 = \infty$, under slightly more restrictive assumptions on the distribution of η_t (see their Proposition 3.9).

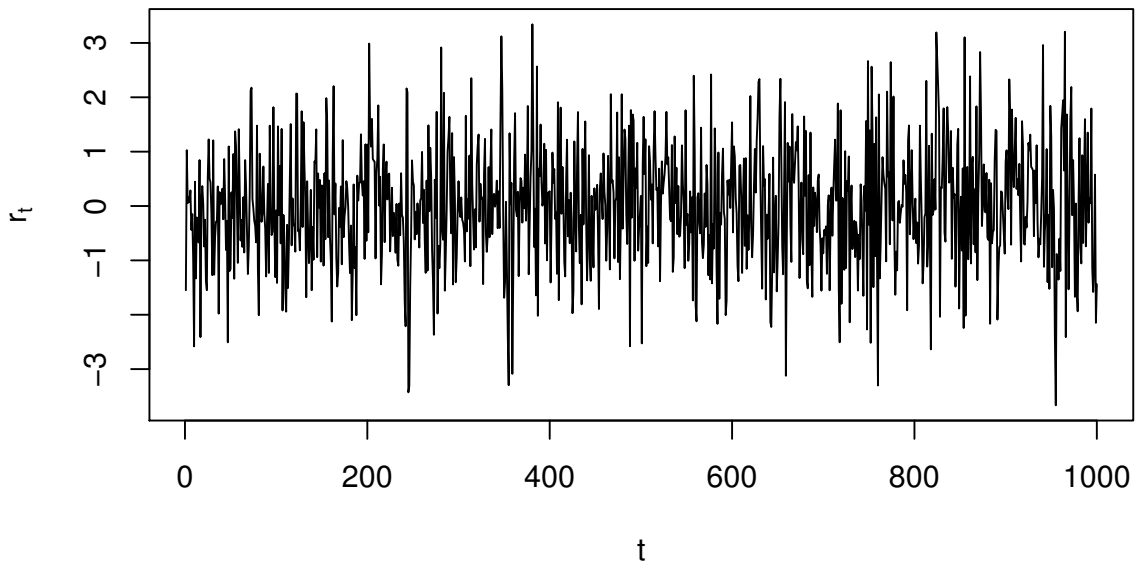

Figure 4.1 – Simulation of $r_t = \sqrt{1 + 0.1 r_{t-1}^2}\epsilon_t$ with $\epsilon_t = \sqrt{1 + 0.05\epsilon_{t-1}^2}\eta_t$, $\eta_t \sim \mathcal{N}(0,1)$.

Remark 4.1. Without Assumption **A4.2**, the GARCH-MIDAS process may admit moments at any order. Indeed, suppose that $\delta_1 \in [0,1]$ with probability 1. It follows that, for any $s > 0$, $\mathbb{E}|\delta_t|^s < 1$ using (4.5). Since $|\eta_t|$ is bounded when $\delta_t < 1$, both σ_t^2 and ϵ_t^2 admit finite moments at any order. If in addition ϵ_t^2 is bounded with probability 1 (which holds when $|\delta_1| < \overline{\delta} < 1$ with probability 1), let $\overline{\boldsymbol{A}}$ the upper bound of the matrices $\boldsymbol{A}_t$ componentwise. If the spectral radius of $\overline{\boldsymbol{A}}$ is less than one, then Assumption **A4.3** is satisfied and, by (4.8), r_t^2 admits moments at any order.

Example 4.1 (Trajectory of a process without any finite moment)**.** *Figure 4.1 displays a simulated trajectory of the simplest version of Model (4.4), which we know, from Proposition 4.1 that it is a strictly stationary process without any finite moment. Other simulations have been carried out, but the absence of any finite moment is, to say the least, difficult to detect on the trajectories.*

3 QMLE without moment assumption on the observed process

In this section, we study the estimation of the true parameter value $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0, a_0, \boldsymbol{\vartheta}_0')'$ in Model (4.4), assuming the functions ϕ_i are known and such that $\sum_{i=1}^q \phi_i(\cdot) = 1$. We start by introducing a consequence of the strict stationarity which will replace the existence of a small moment in the proof of the consistency and asymptotic

normality (CAN) of the QMLE.

3.1 Exponential control of the trajectories

Wang and Ghysels (2015) studied the asymptotic distribution of the QMLE of the GARCH-MIDAS under the assumption that

$$\mathbb{E}|r_t|^s < \infty \text{ for some } s > 0. \tag{4.11}$$

This is a key assumption to show the CAN of the QMLE of GARCH (see Berkes et al. (2003) and Francq and Zakoian (2004)). To the authors' knowledge, the consistency of the QMLE has never been shown without an assumption that implies (4.11). Proposition 4.1 however entails that (4.11) cannot be assumed in our framework.

To circumvent the failure of the small-order moment assumption, we will use the following Lemma, which is a consequence of Theorem 3.2.

Lemma 4.2. *Under **A4.1** and **A4.3**, the strictly stationary solution of (4.4) satisfies*

$$\limsup_{k\to\infty} \frac{1}{k} \log r_{t+k}^2 = 0, \quad \limsup_{k\to\infty} \frac{1}{k} \log r_{t-k}^2 = 0 \quad a.s. \tag{4.12}$$

for all $t \in \mathbb{Z}$.

This property can be interpreted as an exponential control of the trajectories. It is easy to see that (4.11) implies (4.12),[4] but the converse is false.[5]

Assume that the observations $r_1, \dots, r_n$ constitute a realization (of length n) of the two-factor GARCH process defined by (4.4), for the value $\boldsymbol{\theta}_0$ of the parameter. Let $\boldsymbol{\Theta}$ a compact subset of $(0,\infty) \times [0,\infty)^2 \times [0,1) \times \mathbb{R}^d$ and assume $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$. For initial values $r_0, \dots, r_{-q}, \tilde{\sigma}_0^2$, and for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the conditional Gaussian quasi-likelihood is given by

$$\tilde{L}_n(\boldsymbol{\theta}) = \tilde{L}_n(\boldsymbol{\theta}; r_1, \dots, r_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\tau}_t^2\tilde{\sigma}_t^2}} \exp\left(-\frac{r_t^2}{2\tilde{\tau}_t^2\tilde{\sigma}_t^2}\right),$$

where the $\tilde{\tau}_t^2$ and $\tilde{\sigma}_t^2$ are recursively defined, for $t \geq 1$, by

$$\tilde{\tau}_t^2 = \tilde{\tau}_t^2(\boldsymbol{\theta}) = 1 + a \sum_{i=1}^{q} \phi_i(\boldsymbol{\vartheta}) r_{t-i}^2,$$

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\boldsymbol{\theta}) = \omega + \alpha\tilde{\epsilon}_{t-1}^2 + \beta\tilde{\sigma}_{t-1}^2, \quad \tilde{\epsilon}_t^2 = \frac{r_t^2}{\tilde{\tau}_t^2}.$$

4. See for instance Exercise 4.12 in Francq and Zakoian (2019).

5. Let a sequence (X_t) of identically distributed random variables such that $\mathbb{E}|X_t| < \infty$ but $EX_t^2 = \infty$. Then $r_t = e^{|X_t|/2}$ satisfies (4.12) because $k^{-1}\log r_{t+k}^2 = k^{-1}|X_{t+k}| \to 0$ a.s. (see for instance Exercise 2.13 in Francq and Zakoian (2022)). On the other hand (4.11) is not satisfied because $\mathbb{E}|r_t|^s = Ee^{s|X_t|/2} \geq \frac{1}{2}\mathbb{E}(s|X_t|/2)^2 = \infty$, for any $s > 0$.

A QMLE of $\boldsymbol{\theta}_0$ is defined as any measurable solution of

$$\widehat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \tilde{L}_n(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{n}\sum_{t=1}^{n} \tilde{\ell}_t(\boldsymbol{\theta}) := \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \tilde{\mathrm{l}}_n(\boldsymbol{\theta}),$$

where $\tilde{\ell}_t(\boldsymbol{\theta}) = \frac{r_t^2}{\tilde{\tau}_t^2\tilde{\sigma}_t^2} + \log\tilde{\tau}_t^2 + \log\tilde{\sigma}_t^2$.

3.2 Asymptotic properties of the QMLE

To establish the strong consistency of the QMLE, we need the following additional assumptions.

A4.5 The support of the law of η_t^2 contains three distinct points.

A4.6 $(\phi_i(\boldsymbol{\vartheta}))_{i=1...,q} = (\phi_i(\boldsymbol{\vartheta}_0))_{i=1...,q} \;\Rightarrow\; \boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$.

A4.7 $\mathbb{E}\log\eta_t^2 > -\infty$.

For the volatility of a standard GARCH to be non-degenerate, we know that the support of the law of η_t must contain three distinct points. To show the identifiability of the GARCH-MIDAS, we need the slightly stronger assumption **A4.5**. This is due to the fact that the volatility of this model can be written as a polynomial of order 2 in η_{t-1}^2 (instead of order 1 in the GARCH case), with coefficients belonging to the sigma-field generated by $\{\eta_u, u \leq t-2\}$. Assumption **A4.6** is another identifiability condition which is satisfied, in particular, for the exponential weights $\phi_i(\vartheta) = \vartheta^i/\sum_{j=1}^{q}\vartheta^i$ (except when $q=1$). The assumption is also satisfied for the Beta weighting schemes (4.3) (with obvious change of notation). Assumption **A4.7**, precluding densities with too much mass around zero, is satisfied by most commonly used distributions. It is not required for the consistency of the standard GARCH (see Berkes et al. (2003), Francq and Zakoian (2004)) but it is introduced here to circumvent the absence of any moments (Proposition 4.1), which constitutes the major difficulty of the proof of the next consistency result.

Theorem 4.1. *Under Assumptions **A4.1**, **A4.3**-**A4.7**, we have*

$$\widehat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0, \quad a.s.\ as\ n\to\infty.$$

We now turn to the asymptotic normality. We introduce the following additional assumptions.

A4.8 $\boldsymbol{\theta}_0 \in \overset{\circ}{\boldsymbol{\Theta}}$, where $\overset{\circ}{\boldsymbol{\Theta}}$ denotes the interior of $\boldsymbol{\Theta}$.

A4.9 $\kappa_\eta := \mathbb{E}\eta_t^4 < \infty$.

Denote by $\nabla_{\boldsymbol{\theta}}$ (resp. $\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}$) the partial derivative operator (resp. the second-order derivative operator) with respect to $\boldsymbol{\theta}$ (resp. $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$). Similarly, we denote by ∇_{θ_i} the partial derivative with respect to any component θ_i of $\boldsymbol{\theta}$.

A4.10 The functions $\phi_i(\cdot)$, for $i=1,\ldots,q$, admit continuous second-order derivatives and the matrix $[\nabla_{\boldsymbol{\vartheta}}\phi_1(\boldsymbol{\vartheta}_0),\ldots,\nabla_{\boldsymbol{\vartheta}}\phi_q(\boldsymbol{\vartheta}_0)]$ has full-row rank.

A4.11 For $i=1,\ldots,q$, either $\phi_i(\cdot)=0$ or $\phi_i(\boldsymbol{\vartheta}_0)\neq 0$.

Assumption **A4.8** and **A4.9** are also made in the standard GARCH case. Note that **A4.10** and **A4.11** are satisfied in the cases of exponential and Beta weights. The next result establishes the asymptotic normality of the QMLE. Let $V_t(\boldsymbol{\theta}) = \sigma_t^2(\boldsymbol{\theta})\tau_t^2(\boldsymbol{\vartheta})$.

Theorem 4.2. *Under the Assumptions of Theorem 4.1 and **A4.8-A4.11**,*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\kappa_\eta - 1)\boldsymbol{J}^{-1}),$$

where

$$\boldsymbol{J} := \mathbb{E}\left(\frac{1}{V_t^2(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\right) \tag{4.13}$$

is a positive definite matrix.

Despite the absence of moments for the return process (which complicates the proof) the form of the asymptotic variance is thus the same as in the standard GARCH model (with obviously a multiplicative component volatility in the definition of $\boldsymbol{J}$).

4 Testing the existence of a long-run volatility

To test the existence of a long term volatility component, *i.e.* the null hypothesis $H_0 : a_0 = 0$, usual tests such as the Wald test may have non standard asymptotic distributions due to the presence of the unidentified parameter $\boldsymbol{\vartheta}$ under the null. Indeed, it is known that in similar situations (see *e.g.* Figure 1 in Francq et al. (2010)) the Wald, score and Likelihood-Ratio (LR) test statistics do not follow the standard distributions [6] under the null. To solve the problem, we consider two approaches. First, we fix the unidentified parameter to some value $\boldsymbol{\vartheta}^*$. This gives rise to test procedures which have standard, χ^2 or chi-bar-square, asymptotic distributions under the null, but whose power properties depend on the arbitrary choice of $\boldsymbol{\vartheta}^*$. We thus consider a second approach consisting in estimating by QMLE all the parameters, including the unidentified parameter $\boldsymbol{\vartheta}$, and estimating the critical value of the resulting Wald test by a residual-based bootstrap procedure. Note that the identifiability problem is not present in the framework of Conrad and Schienle (2020), in which a score-based test, not requiring the bootstrap, is developed.

4.1 Fixing ϑ

The first approach relies on the auxiliary model

$$\begin{cases} r_t = \tau_t \epsilon_t, & \tau_t^2 = 1 + a_0 \sum_{i=1}^q \phi_i(\boldsymbol{\vartheta}^*) r_{t-i}^2, \\ \epsilon_t = \sigma_t \eta_t, & \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \end{cases} \tag{4.14}$$

where $\boldsymbol{\vartheta}^*$ is given, and the unknown parameter is $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0, a_0)'$. Let $\widehat{\boldsymbol{\theta}}_n = \widehat{\boldsymbol{\theta}}_n(\boldsymbol{\vartheta}^*) = (\widehat{\omega}_n, \widehat{\alpha}_n, \widehat{\beta}_n, \widehat{a}_n)'$ be the QMLE of $\boldsymbol{\theta}_0$. Denote also by $\widehat{\boldsymbol{\theta}}_G = (\widehat{\omega}^c, \widehat{\alpha}^c, \widehat{\beta}^c)'$ the QMLE of a

6. χ^2 for the score, chi-bar-square for the Wald and LR statistics due to the positivity constraints on the estimator of a_0.

standard GARCH(1,1) model. In other words, $\widehat{\boldsymbol{\theta}}_n^c = (\widehat{\boldsymbol{\theta}}_G', 0)'$ is the QMLE of $\boldsymbol{\theta}_0$ under H_0. Let $\boldsymbol{e}_i$ be the i-th column of the 4×4 identity matrix. Let also $\widehat{\eta}_t = r_t/\widetilde{V}_t^{1/2}(\widehat{\boldsymbol{\theta}}_n)$, where $\widetilde{V}_t(\boldsymbol{\theta}) = \widetilde{\sigma}_t^2(\boldsymbol{\theta})\widetilde{\tau}_t^2(\boldsymbol{\theta})$, and

$$\widehat{\eta}_t^c = r_t/\widetilde{V}_t^{1/2}(\widehat{\boldsymbol{\theta}}_n^c) = r_t/\widetilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_G),$$

$\widehat{\kappa}_n = n^{-1}\sum_{t=1}^n |\widehat{\eta}_t|^4$ and $\widehat{\kappa}_n^c = n^{-1}\sum_{t=1}^n |\widehat{\eta}_t^c|^4$. The Wald, score and likelihood ratio test statistics are defined respectively by

$$\mathrm{W}_n = \frac{n}{\widehat{\kappa}_n - 1}\frac{\widehat{a}_n^2}{\boldsymbol{e}_4'\widehat{\boldsymbol{J}}_n^{-1}\boldsymbol{e}_4}, \qquad \widehat{\boldsymbol{J}}_n = \frac{1}{n}\sum_{t=1}^n \frac{1}{\widetilde{V}_t^2}\nabla_{\boldsymbol{\theta}}\widetilde{V}_t\nabla_{\boldsymbol{\theta}}'\widetilde{V}_t(\widehat{\boldsymbol{\theta}}_n),$$

$$\mathrm{R}_n = \frac{n}{\widehat{\kappa}_n^c - 1}\nabla_{\boldsymbol{\theta}}'\tilde{\mathrm{l}}_n(\widehat{\boldsymbol{\theta}}_n^c)\left(\widehat{\boldsymbol{J}}_n^c\right)^{-1}\nabla_{\boldsymbol{\theta}}\tilde{\mathrm{l}}_n(\widehat{\boldsymbol{\theta}}_n^c), \quad \widehat{\boldsymbol{J}}_n^c = \frac{1}{n}\sum_{t=1}^n \frac{1}{\widetilde{V}_t^2}\nabla_{\boldsymbol{\theta}}\widetilde{V}_t\nabla_{\boldsymbol{\theta}}'\widetilde{V}_t(\widehat{\boldsymbol{\theta}}_n^c),$$

and

$$\mathrm{L}_n = 2\frac{n}{\widehat{\kappa}_n - 1}\left\{\tilde{\mathrm{l}}_n(\widehat{\boldsymbol{\theta}}_n^c) - \tilde{\mathrm{l}}_n(\widehat{\boldsymbol{\theta}}_n)\right\}.$$

Denote by χ_1^2 the chi-square distribution with one degree of freedom, and the chi-bar-square distribution $\frac{1}{2}\delta_0 + \frac{1}{2}\chi_1^2$ that is the equally weighted mixture of the Dirac measure at 0 and the χ_1^2 distribution. The following proposition gives the asymptotic distributions of the previous test statistics under the null.

Proposition 4.2. *Assume* ***A4.1****,* ***A4.2****,* ***A4.3****,* ***A4.5****,* ***A4.7****,* ***A4.9*** *and that* $(\omega_0, \alpha_0, \beta_0)' \in \overset{\circ}{\boldsymbol{\Theta}}_G$*, where* $\overset{\circ}{\boldsymbol{\Theta}}_G$ *denotes the interior of the GARCH(1,1) parameter space* $\boldsymbol{\Theta}_G$*, a compact subset of* $(0,\infty)^2 \times [0,1)$*. Under* H_0 *we have,* $\mathrm{W}_n \xrightarrow{\mathcal{L}} \frac{1}{2}\delta_0 + \frac{1}{2}\chi_1^2$*,* $\mathrm{R}_n \xrightarrow{\mathcal{L}} \chi_1^2$ *and* $\mathrm{L}_n \xrightarrow{\mathcal{L}} \frac{1}{2}\delta_0 + \frac{1}{2}\chi_1^2$ *as* $n \to \infty$*.*

We will see in the numerical section that the finite sample distributions of the test statistics are not always well approximated by their asymptotic laws. To solve the problem we will approximate the test statistic distributions by means of a residual-based bootstrap procedure. Recent papers dealing with similar bootstrap inference procedures are Leucht et al. (2015), Beutner et al. (2018), Cavaliere et al. (2022).

Because the Wald test was found to be more powerful than the other tests in our Monte Carlo experiments, we present the resampling scheme and study its asymptotic behavior for the Wald-type statistic only. The algorithm is the following.

1. On the observations $r_1, \ldots, r_n$, compute the QMLE $\widehat{\boldsymbol{\theta}}_G = (\widehat{\omega}, \widehat{\alpha}, \widehat{\beta})'$ of a GARCH(1,1) model and compute the standardized residuals (discarding the first n_0 values the alleviate the effect of the initial values) $\widehat{\eta}_t^0 = (\widehat{\eta}_t^c - m_n)/s_n$, for $t = n_0 + 1, \ldots, n$, where $\widehat{\eta}_t^c$, m_n and s_n are respectively the non-standardized GARCH residuals, their empirical mean and standard deviation. Denote by F_n the empirical distribution of these standardized residuals. Also compute the QMLE of the auxiliary GARCH-MIDAS model (4.14). Let $\widehat{a}_n$ be the estimator of the parameter a.
2. Simulate a trajectory of length n of a GARCH(1,1) model with parameter $\widehat{\boldsymbol{\theta}}_G$ and

iid noise (η_t^*) with distribution F_n, compute the QMLE $\widehat{\boldsymbol{\theta}}_n^* = (\widehat{\omega}_n^*, \widehat{\alpha}_n^*, \widehat{\beta}_n^*, \widehat{a}_n^*)'$ of the GARCH-MIDAS model (4.14).

3. Repeat B times Step 2, and denote by $\widehat{a}_n^{*1}, \ldots, \widehat{a}_n^{*B}$ the bootstrap estimates of a. Approximate the p-value of the test $H_0 : a_0 = 0$ against $H_1 : a_0 > 0$ by $p_B^* = (1 + \#\{\widehat{a}_n^{*j} \geq \widehat{a}_n; j = 1, \ldots, B\})/(B+1)$.

To reduce the computational burden of bootstrap procedures, Kreiss et al. (2011) and Shimizu (2013) proposed to simulate the distribution of the (Q)MLE by using a Newton-Raphson type iteration. This trick can not be used directly here because $\boldsymbol{\theta}_0$ belongs to the boundary of the parameter space under H_0, which implies that the Bahadur-type approximation

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \boldsymbol{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\eta_t^2 - 1\right) \frac{1}{V_t} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0) + o_P(1),$$

used for the Newton-Raphson iteration, is not valid when $a_0 = 0$. By the arguments of Francq and Zakoïan (2009), it can however be seen that in this case

$$\sqrt{n}\widehat{a}_n = \max\left\{ \boldsymbol{e}_4' \boldsymbol{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\eta_t^2 - 1\right) \frac{1}{V_t} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0), 0 \right\} + o(1) \quad \text{a.s.}$$

This suggests replacing $\widehat{a}_n^*$ in Step 2 by

$$\widehat{a}_n^* = \max\left\{ \boldsymbol{e}_4' \left(\widehat{\boldsymbol{J}}_n^c\right)^{-1} \frac{1}{n} \sum_{t=1}^{n} \left(\eta_t^{*\,2} - 1\right) \frac{1}{\widetilde{V}_t} \nabla_{\boldsymbol{\theta}} \widetilde{V}_t(\widehat{\boldsymbol{\theta}}_n^c), 0 \right\}. \tag{4.15}$$

Since White (1982) it is known that the (Q)MLE of a misspecified model generally converges to some pseudo-true value. The resampling algorithm is valid in the following sense.

Theorem 4.3. *Let the assumptions of Proposition 4.2 hold. Assume also that the distribution of η_t admits a bounded density with respect to the Lebesgue measure. Let $\widehat{a}_n^*$ defined by (4.15). Under H_0, for almost all realization (r_t), as $n \to \infty$ we have, given (r_t),*

$$\sqrt{n}\widehat{a}_n^* \xrightarrow{\mathcal{L}} N \mathbb{1}_{N \geq 0}, \quad N \sim \mathcal{N}\left(0, \sigma^2 := (\kappa - 1)\boldsymbol{e}_4 \boldsymbol{J}^{-1} \boldsymbol{e}_4\right), \tag{4.16}$$

and thus

$$\mathrm{W}_n^* := \frac{n}{\widehat{\kappa}_n - 1} \frac{(\widehat{a}_n^*)^2}{\boldsymbol{e}_4' \widehat{\boldsymbol{J}}_n^{-1} \boldsymbol{e}_4} \xrightarrow{\mathcal{L}} \frac{1}{2}\delta_0 + \frac{1}{2}\chi_1^2.$$

Under $H_1 : a_0 > 0$, for almost all realization (r_t), if $\widehat{\boldsymbol{\theta}}_G$ converges to some pseudo-true value $\boldsymbol{\theta}_G \in \boldsymbol{\Theta}_G$ such that

$$\boldsymbol{J} := \mathbb{E} \frac{1}{V_t^2} \nabla_{\boldsymbol{\theta}} V_t \nabla_{\boldsymbol{\theta}} V_t' \begin{pmatrix} \boldsymbol{\theta}_G \\ 0 \end{pmatrix}$$

exists and is invertible and if $\widehat{a}_n \to a_0$ *then* $p^* \to 0$ *as* $n \to \infty$, *where* $p^* = \lim_{B\to\infty} p_B^*$ *a.s.*

The previous result thus shows that the distribution of $\widehat{a}_n^*$ (resp. W_n^*) given (r_t) well mimics the (unconditional) distribution of $\widehat{a}_n$ (resp. W_n) under H_0 when n is large. It is also expected that in finite samples the bootstrap distribution of $\sqrt{n}\widehat{a}_n^*$ better approaches the distribution of $\sqrt{n}\widehat{a}_n$ than its asymptotic distribution. The consistency of the bootstrap is also ensured as soon as $\liminf_{n\to\infty} \widehat{a}_n > 0$ and $\sqrt{n}\widehat{a}_n^* = O_P(1)$, which holds under the conditions of the theorem, but also under more general conditions.

4.2 Bootstrapping the full Wald test

The asymptotic properties of the test statistics defined in the previous section do not depend on the fixed value of the parameter $\boldsymbol{\vartheta}^*$ in (4.14). However, the illustrations presented in the numerical section show that the finite sample behavior of the tests depends on this parameter. In addition, there is no obvious choice of the parameter that one could recommend to the practitioner. When $\boldsymbol{\vartheta}$ is estimated by QMLE, together with the other parameters, the test statistics have non standard asymptotic distributions under the null, and the bootstrap techniques become particularly appealing. The resampling scheme is then modified as follows.

1. On the observations $r_1, \ldots, r_n$, compute the GARCH(1,1) QMLE $\widehat{\boldsymbol{\theta}}_G = (\widehat{\omega}, \widehat{\alpha}, \widehat{\beta})'$ and the standardized residuals $\widehat{\eta}_t^0 \sim F_n$, exactly as in the previous algorithm. Compute the QMLE of the GARCH-MIDAS model (4.4). Let $\widehat{a}_n$ be the estimator of the parameter a.
2. Simulate a trajectory of length n of a GARCH(1,1) model with parameter $\widehat{\boldsymbol{\theta}}_G$ and iid noise (η_t^*) with distribution F_n, compute the QMLE $\widehat{\boldsymbol{\theta}}_n^* = \left(\widehat{\omega}_n^*, \widehat{\alpha}_n^*, \widehat{\beta}_n^*, \widehat{a}_n^*, \widehat{\boldsymbol{\vartheta}}_n^{*\prime}\right)'$ of the GARCH-MIDAS model (4.4).
3. Repeat B times Step 2, and compute the bootstrap estimated p-value p_B^* exactly as in the previous algorithm.

Under H_0, the distribution of $\widehat{a}_n$ in the QML estimation of the full GARCH-MIDAS model (4.4) is an unknown function $G_n(\boldsymbol{\theta}_0, F)$ of the GARCH parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)'$ and the distribution F of the noise η_t. The previous residual bootstrap algorithm estimates $G_n(\boldsymbol{\theta}_0, F)$ by $G_n(\widehat{\boldsymbol{\theta}}_G, F_n)$. A formal justification, similar to that given in Theorem 4.3, would certainly rely on the strong consistency of $\widehat{\boldsymbol{\theta}}_G$ and on the consistency of F_n, in the sense of (4.29), and would require establishing a kind of continuity of $G_n(\cdot)$ and/or the asymptotic form of G_n as $n \to \infty$. To obtain the latter, techniques used to obtain the asymptotic distribution of sup-type test statistics, as in Hansen (1996), could be considered but the problem seems difficult because the parameter is on the boundary under the null hypothesis (see Andrews (2001)). To our knowledge there is no available result dealing with sup-tests when the parameter is on the boundary of the parameter set.

Note that the choice of B has little effect on the size and power of the test. Consider the test which rejects the null when $p_B^* \leq 5\%$. If $B = 19$ or $B = 99$, the size is exactly 5%.

Note also that the bootstrap is a randomized procedure, in the sense that the statistical decision depends not only on the observations $r_1, \ldots, r_n$, but also on the random bootstrap trials (for a formal definition, see *e.g.* Page 98 in Van der Vaart (2000)). Taking a large value of B (we took $B = 999$ for the numerical illustrations of Section 5.2) has the advantage of reducing the test randomness. To assess the performance of the bootstrap test on Monte-Carlo simulation experiments, the randomness of the procedure is not an issue. We thus follow the so-called « warp-speed » methodology of Giacomini et al. (2013) by computing $\widehat{a}_n$ on a large number K of Monte Carlo replications of a GARCH-MIDAS model (4.1). For each of the K Monte Carlo simulations, we generated $B = 1$ bootstrap simulation and computed the corresponding bootstrap statistic $\widehat{a}_n^*$. Let ξ_α^* be the α-quantile of the K values of $\widehat{a}_n^*$. The size (resp. power) of the bootstrap test of nominal level α is then approximated by the proportion of $\widehat{a}_n > \xi_{1-\alpha}^*$ over the K replications when $a_0 = 0$ (resp. $a_0 > 0$) in the simulated GARCH-MIDAS model.

5 Numerical results

We first present the results of Monte Carlo experiments. Our objectives are twofold: *i)* evaluating the effect of the absence of moments on the accuracy of the QMLE, and *ii)* assessing the performance of the QML in detecting and estimating the two volatility components. Then, we will present an application on real financial data.

5.1 Monte Carlo experiments

The aim of our first Monte Carlo experiment is to study the effect of the absence or presence of marginal moments on the empirical accuracy of the QMLE. We simulated the simplest version of model (4.4) with $q = 1$, $\phi_i(\boldsymbol{\vartheta}) \equiv 1$ and parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0, a_0)$ given in the column « True » of Table 4.1. For the first data generating process (DGP A) the noise η_t is $\mathcal{N}(0,1)$-distributed, so that **A4.2** is satisfied, and the DGP is stationary but does not admit any moment. For the second data generating process (DGP B) the noise η_t follows an equally weighted mixture of $\mathcal{N}(m,1)$ and $\mathcal{N}(-m,1)$ distributions truncated on the interval $[-b, b]$, where m is chosen such that $\mathbb{E}\eta_t^2 = 1$ and $b = \sqrt{(1-\iota-\beta)/\alpha}$ with $0 < \iota < 1 - \beta$. Since $a_t < 1 - \iota$ a.s. we have $\epsilon_t^2 \leq b\omega/\iota$. If $\iota > ab\omega$ then $a\epsilon_t^2 < 1$, which entails that r_t is bounded. For DGP B, we took $\iota = 0.05$, so that $b = \sqrt{3}$, $0 < \iota < 1 - \beta = 0.2$ and $\iota > ab\omega = 0.02\sqrt{3}$. This DGP thus admits moments of any order.

The number of replications of each simulation is $R = 1000$, with sample sizes $n = 2000$ and $n = 4000$. The two DGPs have been estimated by QMLE. Table 4.1 displays the results of these Monte Carlo experiments. The columns « Min », « Q1 », « Q2 », « Q3 », « Max », « Bias » and « RMSE » provide respectively the minimum, the first quartile, the median, the third quartile, the maximum, the bias and the root mean square error (RMSE) of the R estimated values of the parameter. The column « MASE » refers to the estimated standard error based on the asymptotic theory. The i-th Mean Asymptotic Standard Error (MASE) is defined as the empirical mean over the R replications of the estimated

n		True	Min	Q1	Q2	Q3	Max	Bias	RMSE	MASE
DGP A satisfying **A4.2** (no moments)										
2000	ω	0.2	0.023	0.146	0.221	0.343	1.391	0.076	0.206	0.760
	α	0.05	0.000	0.037	0.054	0.082	0.240	0.015	0.045	0.043
	β	0.8	0.000	0.676	0.781	0.849	0.978	-0.064	0.174	0.642
	a	0.1	0.000	0.061	0.089	0.115	0.236	-0.012	0.044	0.046
4000	ω	0.2	0.008	0.153	0.210	0.283	0.901	0.037	0.139	0.112
	α	0.05	0.000	0.038	0.052	0.068	0.253	0.007	0.031	0.024
	β	0.8	0.212	0.730	0.790	0.841	0.991	-0.031	0.120	0.098
	a	0.1	0.000	0.076	0.096	0.115	0.193	-0.005	0.032	0.029
DGP B that does not satisfy **A4.2** (moments at any order)										
2000	ω	0.2	0.008	0.149	0.227	0.340	1.030	0.073	0.197	0.310
	α	0.05	0.000	0.038	0.055	0.081	0.205	0.014	0.041	0.040
	β	0.8	0.161	0.680	0.774	0.846	0.992	-0.061	0.167	0.256
	a	0.1	0.000	0.067	0.091	0.112	0.187	-0.012	0.039	0.039
4000	ω	0.2	0.020	0.154	0.208	0.280	1.010	0.034	0.130	0.107
	α	0.05	0.005	0.041	0.051	0.066	0.222	0.006	0.027	0.023
	β	0.8	0.160	0.731	0.791	0.838	0.975	-0.028	0.111	0.093
	a	0.1	0.000	0.081	0.097	0.113	0.181	-0.005	0.026	0.025

Table 4.1 – Distribution of the QMLE over 1000 replications

standard errors $\sqrt{\widehat{\boldsymbol{\Sigma}}(i,i)/n}$, where $\widehat{\boldsymbol{\Sigma}}$ is the empirical estimator of the asymptotic variance $\boldsymbol{\Sigma} = (\kappa_\eta - 1)\boldsymbol{J}^{-1}$ of the QMLE. As expected, bias and RMSE decrease when the sample size increases. The values of RMSE and MASE get closer as the sample size increases, which means that the empirical distribution of the estimator becomes closer to its asymptotic distribution. Unsurprisingly, the QMLE turns out to be more accurate when all moments exist (DGP B) than when there is no moment (DGP A), but the difference in accuracy is quite small.

In a second set of Monte Carlo experiments, we assess the ability of our estimation approach to estimate and detect the presence of long-term volatility. We chose to estimate the GARCH-MIDAS specification of τ_t in (4.1), with Beta weights given by (4.3). We thus simulated 1000 trajectories of size $n = 4000$ of Model (4.1) with $N = 22$, $Q = 250$ and $(\omega_0, \alpha_0, \beta_0, \vartheta_0, a_0) = (0.028, 0.115, 0.831, 2.067, 0.056)$.[7] For the distribution of η_t we took a standardized Student distribution with $\nu = 5.41$ degrees of freedom[8]. The estimation

7. These parameters are those estimated on the NASDAQ index considered in Section 5.2, with RVs computed over one month and one MIDAS lag year, on a set of historical data of size $n = 12654$ (for our simulations, we consider the smallest sample size $n = 4000$).

8. the kurtosis thus corresponds to the empirical kurtosis of the residuals of the model fitted to the NASDAQ series.

results are presented in the top panel of Table 4.2. Interestingly, the parameter a_0 is estimated with a small bias, and its estimated standard deviation is on average very close to the observed RMSE. We have redone the estimation exercise on simulations of a standard GARCH (corresponding to Model (4.4) with $a_0 = 0$). The bottom panel of Table 4.2 shows that at least one half of the estimated values of a are exactly equal to zero. Unsurprisingly, the estimations of ϑ, whose true value is undefined when $a_0 = 0$, are erratic. Figure 4.2 displays a typical example of estimates of the short and long term volatilities of the two DGPs of Table 4.2. The distinction between the dynamics of the two DGPs is clear from the figure, and can be confirmed by a formal test of the null hypothesis $H_0 : a_0 = 0$. Figure 4.3 shows that the estimation of the volatilities is fortunately not too sensitive to the choice of the integers N and Q in (4.1). Finally, we estimated a (misspecified) standard GARCH(1,1) on simulations of a GARCH-MIDAS (with same parameters as in the first part of Table 4.2). Table 4.3 presents the estimation results. The columns « Mean » and « SD » stand for the mean and standard deviation of the estimates over the 1000 replications. It can be noted that the estimated value of $\alpha + \beta$ is always very close to 1, a stylized fact that is often observed on real series. Over a small sub-period of a randomly chosen simulation, Figure 4.4 graphically compares the volatility estimates obtained by the correctly specified GARCH-MIDAS model with those obtained by the misspecified standard GARCH(1,1). Even if the volatility estimation of the standard GARCH is, as expected, dominated by the GARCH-MIDAS estimation, the difference is not huge. Table 4.4 confirms that the estimates obtained from the GARCH-MIDAS model are indeed better, but only slightly better, than those obtained from the GARCH model, as measured by the QLIK loss defined by

$$QLIK = \frac{1}{n} \sum_{t=r_0+1}^{n} \frac{V_t^2}{\widehat{V}_t^2} + \log \widehat{V}_t^2,$$

where V_t denotes the true volatility and $\widehat{V}_t$ denotes the estimated volatility (for the GARCH or the GARCH-MIDAS). We took $r_0 = 100$ to avoid the effect of the initial values required to compute the volatility estimates. The reader is referred to Patton (2011) for arguments in favor of the QLIK loss to compare volatility forecasts/estimates. We did not use the MSE loss because we know from Proposition 4.1 that σ_t^2 does not admit any moment.

Table 4.5 gives the empirical relative frequency of rejection of the score, Wald and LR tests of Section 4.1 for the null of no long-run volatility. The DGP is that used in Table 4.2, except that $a_0 = 0$ (under the null) or $a_0 \in \{0.01, 0.05\}$ (under the alternative). The number of replications is 1000. Different values of $\vartheta \geq 1$ are used. With $\vartheta = 1$ all the RVs involved in (4.1) have the same weight; the larger ϑ, the higher the weights of the most recent RVs. It can be seen from this table that the 3 tests are conservative, but the size is better controlled with the score test. It can also be seen that the ranking of the 3 tests, in terms of power, vary a lot with a_0, the nominal level, and the parameter ϑ. Other numeric experiments, not presented here, show that the Wald test seems slightly more powerful than the two other tests when the sample size n is larger. The poor control of the error of the first kind, as well as the sensitivity to the choice of the fixed parameter

	True	Min	Q1	Q2	Q3	Max	Bias	RMSE	MASE
ω	0.028	0.009	0.025	0.033	0.042	0.139	0.007	0.016	0.013
α	0.115	0.057	0.103	0.116	0.128	0.203	0.001	0.019	0.020
β	0.831	0.572	0.804	0.828	0.846	0.922	-0.008	0.037	0.033
ϑ	2.067	0.000	1.447	2.122	3.173	68.541	0.743	3.329	5.925
a	0.056	0.000	0.030	0.045	0.064	0.256	-0.005	0.033	0.032
ω	0.028	0.010	0.024	0.028	0.034	0.084	0.002	0.009	0.009
α	0.115	0.062	0.105	0.116	0.129	0.204	0.002	0.020	0.021
β	0.831	0.534	0.802	0.823	0.841	0.908	-0.012	0.038	0.033
ϑ	UD	0.000	2.067	2.067	2.067	4650.425	10.653	207.455	199.696
a	0	0.000	0.000	0.000	0.013	0.116	0.010	0.020	0.022

Table 4.2 – Distribution of the QMLE of a GARCH-MIDAS, when the DGP is a GARCH-MIDAS (first part) and when it is a standard GARCH (second part of the table). In the latter case, the parameter ϑ is undefined (UD).

	Min	Q1	Q2	Q3	Max	Mean	SD
ω	0.004	0.025	0.033	0.043	0.131	0.036	0.015
α	0.027	0.092	0.105	0.118	0.170	0.105	0.019
β	0.747	0.857	0.873	0.890	0.969	0.872	0.026
$\alpha + \beta$	0.876	0.968	0.979	0.989	1.015	0.977	0.017

Table 4.3 – Distribution of the QMLE of a GARCH(1,1) when the DGP is the GARCH-MIDAS of Table 4.2 (top panel).

ϑ, motivated us to consider the bootstrapped Wald test of Section 4.2. Table 4.6 shows that this bootstrap test much better controls the error of the first kind, without degrading the power. Note that these empirical sizes and powers are obtained from the warp-speed methodology of Giacomini et al. (2013), as explained in Section 4.2, with $K = 1000$.

5.2 Application to stock indices

We estimated the GARCH-MIDAS model (4.1) with exponential weights on the daily returns of the CAC 40, DAX, NASDAQ and Hang Seng indices, from 1990-03-01 to 2021-04-08. Table 4.7 displays the estimated coefficients when $N = 65$ (corresponding to RVs over a quarter) and $Q = 1000$ (corresponding to 4 MIDAS lag years). These values were advocated by Engle et al. (2013). We checked that the short and long term volatilities are not much modified with other choices of these parameters (in particular with biannual rolling window RV, *i.e.* $N = 125$, and 2 MIDAS lag years, or with $N = 22$ and $Q = 250$, *i.e.* RVs over one month and one MIDAS lag year). The last column of Table 4.7 displays the estimated p-values of the bootstrap Wald test of Section 4.2 (with $B = 999$). The most noticeable output of that Table is that these p-values are small and the estimated value of a is always clearly significant, except perhaps for the HSI series, showing the

Model	Min	Q1	Q2	Q3	Max	Mean	SD
MIDAS	0.589	0.956	1.101	1.265	3.285	1.133	0.263
GARCH	0.591	0.959	1.108	1.272	3.301	1.139	0.265

Table 4.4 – Distribution of the QLIK losses over 1000 replications when the GARCH-MIDAS volatility is estimated by the GARCH-MIDAS model or by the GARCH model.

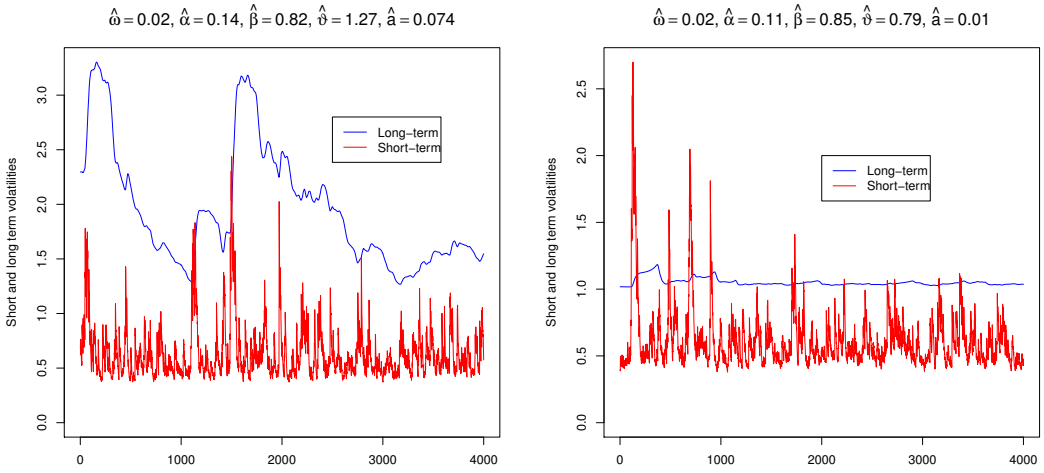

Figure 4.2 – Examples of estimated short and long term volatilities when the GDP is a GARCH-MIDAS (left figure) or a standard GARCH (right figure)

existence of time-varying long term volatilities. Figure 4.5 confirms that the GARCH-MIDAS parameter estimate $\widehat{a}_n$ is well on the right of its estimated distribution under the null $H_0 : a = 0$. The latter distribution, which is a mixture of a Dirac mass at zero and a continuous distribution on $(0, \infty)$, has been estimated by a Kernel density estimator—using the reflection method for boundary correction. Figure 4.6 displays the estimated short and long-term volatilities. The most striking feature of this figure is that long-term volatility varies strongly, but as expected slowly, over time. The volatilities of the CAC and DAX indices are surprisingly similar, with in particular a strong increase in long-term volatility after the 2008 crisis and the recent Covid crisis. The Nasdaq behaves similarly in the most recent period, but reacted much more to the 2001 recession. The HSI behaves quite differently, with an increase in long-term volatility after the Asian Crisis of 1997 and after the Global Financial Crisis of 2008, but with little response to the Covid pandemic.

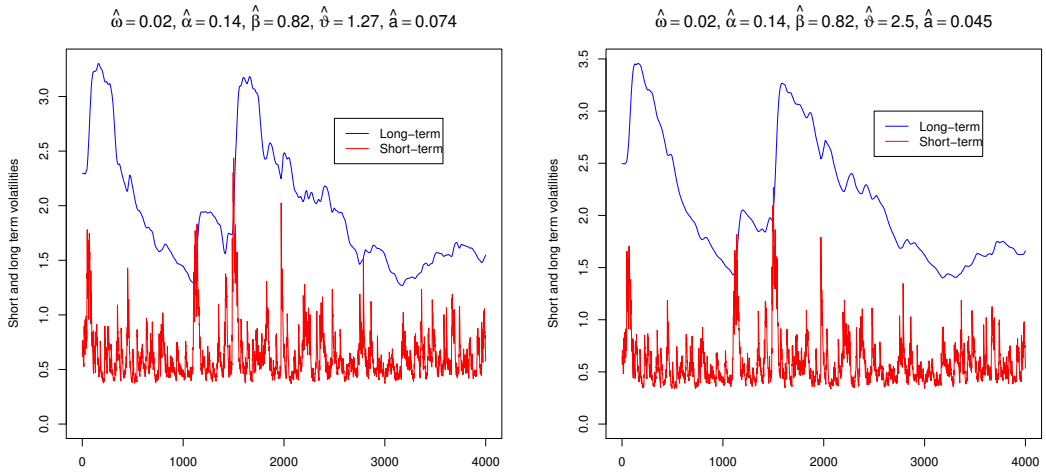

Figure 4.3 – Estimated short and long term volatilities of the GARCH-MIDAS model with $N = 22$ and $Q = 250$ (left figure) and with $N = 44$ and $Q = 500$ (right figure)

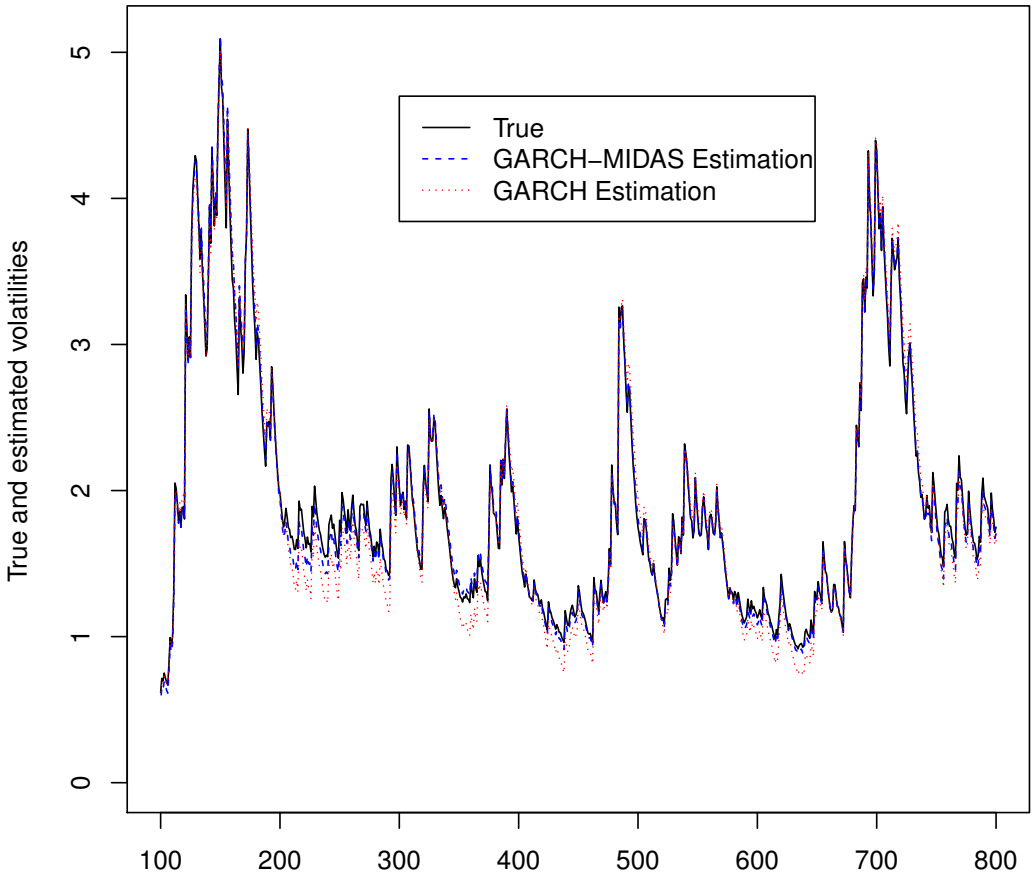

Figure 4.4 – True and estimated volatility estimated by a GARCH-MIDAS and by a standard GARCH

a_0	ϑ	Test	0.1%	1%	2%	3%	4%	5%	6%	7%	10%	20%
0	1	R_n	0	1	2	3.3	3.9	4.9	6.2	6.9	10.1	18.9
		W_n	0	0	0	0	0	0	0.2	0.4	1.6	9.7
		L_n	0	0.2	0.6	1.4	1.8	1.9	2	2.5	3.2	6.6
	2	R_n	0	0.8	1.5	3	4.1	4.9	5.4	6.6	9.2	19.2
		W_n	0	0	0	0	0.1	0.1	0.3	0.6	2.1	8.8
		L_n	0	0.2	0.6	1	1.8	2.2	2.6	2.7	3.4	6.1
	3	R_n	0	0.4	1.5	2.9	3.8	4.8	5.8	6.6	9.3	18.6
		W_n	0	0	0	0.1	0.1	0.4	0.7	1.2	3.1	10.1
		L_n	0	0.3	0.6	1.2	1.6	2.1	2.4	2.6	3.3	5.9
	9	R_n	0.3	0.8	1.5	2.1	3.2	3.6	4.1	5.1	7.5	15.9
		W_n	0	0	0.5	1.4	1.7	2.3	3.2	3.6	5.8	12.6
		L_n	0	0.6	1.1	1.8	2.2	2.4	2.7	3.2	3.9	6.7
0.01	1	R_n	0.6	2.5	3.8	4.8	5.3	6.8	7.7	8.6	11.5	20.8
		W_n	0	0	0.1	0.3	0.4	0.6	1.3	1.9	4.9	20.6
		L_n	0.3	1.9	2.6	3.5	4.3	5	5.6	6.5	8.9	14.6
	2	R_n	0.2	2.2	3.5	4.7	5.6	6.4	7.6	8.5	11.3	20.8
		W_n	0	0.1	0.2	0.3	0.6	1.1	2	3	7.3	22.7
		L_n	0.5	2.1	3.4	4.1	4.6	5.6	6.3	6.9	9.4	16.4
	3	R_n	0.4	2.2	3.1	3.9	4.7	5.4	6.2	7.2	10.1	20.3
		W_n	0	0.1	0.2	0.4	1.1	1.9	3	4.5	8.8	24.3
		L_n	0.3	2.2	3	4.1	4.9	5.4	6	6.6	9.5	15.7
	9	R_n	0.4	0.9	1.2	1.4	2.1	3.1	3.5	4	6	15
		W_n	0	0.4	1.2	2.8	4.1	5.5	7.1	8.3	12.2	25.2
		L_n	0.3	2	2.9	3.7	4.7	5.4	5.8	6.3	7.9	12.6
0.05	1	R_n	11.1	29.4	37.4	43.2	48.1	52.8	55.4	57.1	62.4	73.3
		W_n	0.1	1.3	4.6	9.5	17.2	25.1	34	42.1	61.2	84.4
		L_n	17.3	38.2	46.9	52.6	57.6	60.4	63.2	65.2	70.9	81
	2	R_n	9.2	24.2	32.6	38.6	41	43.8	47	50.3	55.2	65.6
		W_n	0	1.4	7.2	17.9	29.3	38.2	48.1	55.3	71.3	90.2
		L_n	24.4	46.8	54.9	60.6	65	67.7	70.2	71.8	76.7	85
	3	R_n	4.1	14.6	20.1	23.4	26.7	28.7	31	33.3	38.6	51.3
		W_n	0	2	11.8	25.9	38.1	47.7	55.2	61.9	74.4	90.6
		L_n	22.4	44.6	53.5	58.5	61.5	64.3	67.5	69.6	74.5	83.1
	9	R_n	0.6	2.6	3.4	4.9	6.1	6.9	8.2	9.5	12.8	23.2
		W_n	0	13.6	33.4	44.1	52.4	58.4	62	66.2	73.2	84.9
		L_n	8.8	21.6	27.7	32.7	36.5	39.9	43.3	45.1	51.5	62.3

Table 4.5 – Empirical relative frequency of rejection of the null that there exists no long-run volatility (*i.e.* $a_0 = 0$) using the score, Wald and LR tests with a fixed value of ϑ, for nominal levels varying from 0.1% to 20%.

a_0	0.1%	1%	2%	3%	4%	5%	6%	7%	10%	20%
0	0.2	1.2	3.2	3.6	4.3	4.9	6.5	7.7	10.9	23.4
0.01	0.0	1.1	4.2	7.8	10.2	11.5	13.8	15.6	20.5	36.5
0.05	0.5	5.3	20.7	32.4	45.9	56.2	61.2	66.2	75.2	93.1

Table 4.6 – Empirical relative frequency of rejection of the null that there exists no long-run volatility (*i.e.* $a_0 = 0$) using the bootstrapped version of the Wald test, for nominal levels varying from 0.1% to 20%.

	ω	α	β	ϑ	a	p-value
CAC	0.031 0.007	0.110 0.011	0.846 0.017	16.308 6.656	0.013 0.005	0.003
DAX	0.027 0.008	0.095 0.012	0.867 0.018	11.724 5.729	0.012 0.005	0.010
NASDAQ	0.026 0.005	0.113 0.011	0.840 0.015	10.813 3.227	0.017 0.005	0.001
HSI	0.034 0.009	0.080 0.011	0.884 0.016	11.316 5.889	0.008 0.003	0.031

Table 4.7 – GARCH-MIDAS fitted on stock returns. The estimated standard deviations are displayed in small font, under the estimated values of the coefficients. The last column gives the bootstrap estimated p-value of the Wald test of $H_0 : a = 0$.

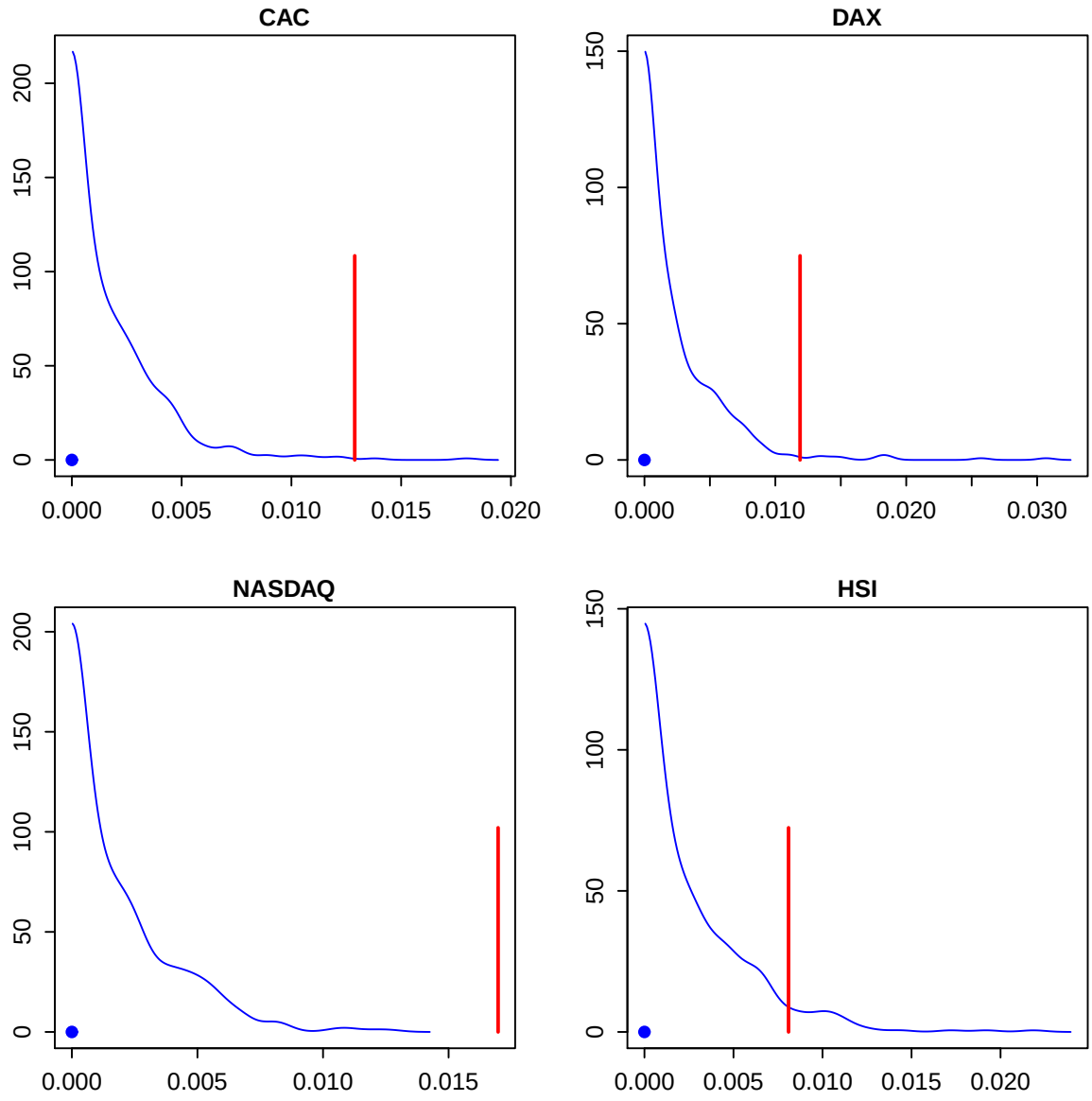

Figure 4.5 – Bootstrap estimate of the distribution of $\widehat{a}_n$ when $a = 0$ (in blue) and observed value of $\widehat{a}_n$ (red vertical line).

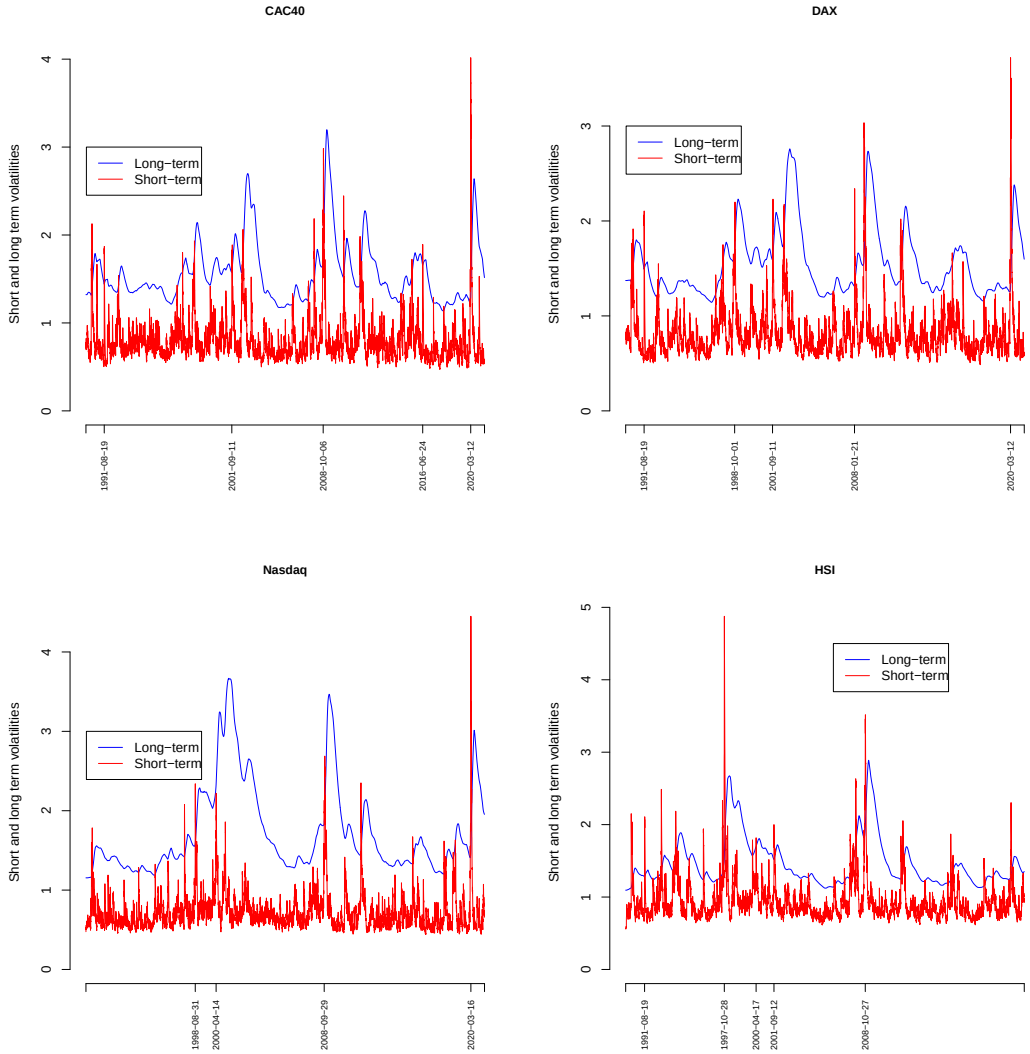

Figure 4.6 – GARCH-MIDAS short and long term volatilities for four stock indices from 1990-03-01 to 2021-04-08.

6 Conclusion

In this chapter, we studied a class of models enabling long and short run volatilities. We showed that strictly stationary solutions are so heavy tailed that not even a small power moment exists. The main theoretical novelty with respect to the literature on GARCH estimation comes from showing that strong consistency and asymptotic normality of the QMLE hold despite the absence of moments. We also proposed tests of the existence of a long-run volatility component. Our numerical applications illustrated the ability of the QML to distinguish and accurately estimate the two components in finite sample, but also confirmed that a misspecified GARCH model can deliver reliable estimates of volatility. Other specifications of the long-run variance could be considered in further work, in particular those including exogenous variables (such as macroeconomic factors) in the dynamics of τ_t, as in Conrad and Loch (2015), or Conrad and Schienle (2020) among many others.

GARCH-MIDAS is a complex model and several difficult questions remain open. In particular, does stationarity entail the existence of log-moments? It seems difficult to conjecture the result. On the one hand, it can be shown that (4.12) is equivalent to the existence of a finite log-moment when (r_t^2) is iid and bounded away from zero [9]. On the other hand, Tanny (1974) provided an example of stationary and ergodic sequence (r_t) where (4.12) is true and the log-moment is infinite. At least from a theoretical point of view, it would be interesting to know if this is also the case for GARCH-MIDAS processes. Other interesting questions concern the practical implications of the absence of moments. Starting from the general principle that it is better for a model to share the same characteristics as the data to which it applies, the question is whether financial returns (or other real time series) are devoid of a finite moment. It is too difficult a problem to solve here. It seems from our experiments that the existence of moments might not be detectable from the trajectories (see the graph in Example 4.1). Another interesting question raised by a referee is the behavior of sample autocorrelations in the absence of theoretical autocorrelations. From Davis and Resnick (1986), the empirical ACF of an AR(1) with heavy-tailed iid innovations is known to converge to the AR coefficient. More recently, Skrobotov et al. (2021) derived the asymptotic distribution of empirical autocorrelations of powers of absolute returns under heavy-tailed assumptions. Do these results hold true when innovations follow the GARCH-MIDAS model? The numerical experiments we have done lead us to believe that the convergence holds but proving the result is beyond the scope of this paper.

7 Appendix: proofs

7.1 Proof of Lemma 4.1

Let $\mathcal{F}_t$ be the sigma-field generated by $\{\eta_u, u \leq t\}$. Let $\mu_p = \mathbb{E}|\eta_1|^p$ for any $p > 0$. Note that $\mu_{p_i} \in (0, \infty]$ because $\mu_2 = 1$ implies that $|\eta_1|$ can not be equal to zero with probability one. Without loss of generality, assume $i_2 \geq 1$. We can also assume $\mu_{p_1} < \infty$, otherwise

9. using (i)-(iii) in Tanny (1974)

the result is trivial. Since $\sigma_t \geq \alpha_0^{1/2}|\epsilon_{t-1}|$, for all positive random variable $X_{t-2} \in \mathcal{F}_{t-2}$ we have

$$\mathbb{E}|\epsilon_{t-1}|^{p_1} X_{t-2} = \mu_{p_1} \mathbb{E}\sigma_{t-1}^{p_1} X_{t-2} \geq \mu_{p_1} \alpha_0^{\frac{p_1}{2}} \mathbb{E}|\epsilon_{t-2}|^{p_1} X_{t-2}.$$

By succesive applications of this inequality, it follows that

$$\begin{aligned} &\mathbb{E}|\epsilon_{t-i_1}|^{p_1}|\epsilon_{t-i_1-i_2}|^{p_2}\dots|\epsilon_{t-i_1-\dots-i_k}|^{p_k} \\ &\geq \left(\mu_{p_1}\alpha_0^{\frac{p_1}{2}}\right)^{i_2} \mathbb{E}|\epsilon_{t-i_1-i_2}|^{p_1+p_2}|\epsilon_{t-i_1-i_2-i_3}|^{p_3}\dots|\epsilon_{t-i_1-\dots-i_k}|^{p_k}. \end{aligned}$$

Iterating the argument, we obtain the result with

$$K = \left(\mu_{p_1}\alpha_0^{\frac{p_1}{2}}\right)^{i_2}\left(\mu_{p_1+p_2}\alpha_0^{\frac{p_1+p_2}{2}}\right)^{i_3}\cdots\left(\mu_{p_1+\dots+p_{k-1}}\alpha_0^{\frac{p_1+\dots+p_{k-1}}{2}}\right)^{i_k}.$$

Under **A4.2**, $\mathbb{E}|\epsilon_t|^{2r} = \infty$ for r large enough and the conclusion follows.

7.2 Proof of Lemma 4.2

First note that $\boldsymbol{X}_t = \boldsymbol{r}_t$ satisfies the Stochastic Iterated Function Systems (3.1), with $\theta_t = (\boldsymbol{A}_t, \boldsymbol{b}_t)$. Note that the assumptions of Theorem 3.1 are satisfied: i) is satisfied because GARCH possess small order moments, ii) is satisfied with Λ_t the operator norm of $\boldsymbol{A}_t$, and iii) is satisfied under **A4.3**. The lemma is thus a consequence of Theorem 3.2, since $\log r_{t+k}^2 \leq \frac{1}{2}\log d(\boldsymbol{r}_{t+k}, c)$ with $c = 0$ and d the euclidean norm.

7.3 Proof of Proposition 4.1

Let us show the first assertion. We begin with the sufficient condition. Under **A4.1** and **A4.3**, the strictly stationary solution is obtained from (4.5) and (4.8) by taking r_t equal to the square root of the first component of $\boldsymbol{r}_t$ multiplied by the sign of η_t. For the necessary condition, consider that (4.4) admits a strictly stationary and ergodic solution (r_t). Note that Klüppelberg et al. (2004) consider the GARCH$(1,1)$ with the condition $\mathbb{P}(\eta_1 = 0) = 0$ but, in their Theorem 2.1, they show without this assumption that, if **A4.1** does not hold then $\sigma_t \to \infty$ in probability as $t \to \infty$. Since for all t, $r_t^2 \geq \sigma_t^2\eta_t^2$, it follows that for all $M > 0$, $\mathbb{P}(r_t^2 > M) \geq \mathbb{P}(\sigma_t^2 > M, \eta_t^2 \geq 1) \to \mathbb{P}(\eta_1^2 \geq 1)$ as $t \to \infty$. We know that $\mathbb{P}(\eta_1^2 \geq 1) > 0$ because $\mathbb{E}\eta_1^2 = 1$. Thus, for all $M > 0$, there exists t_0 such that $t \geq t_0$, $\mathbb{P}(r_t^2 > M) > \mathbb{P}(\eta_1^2 \geq 1)/2 > 0$. This contradicts the stationarity of (r_t) that implies that for all t, $\mathbb{P}(r_t^2 > M) = \mathbb{P}(r_1^2 > M) \to 0$ as $M \to \infty$ The condition **A4.3** follows from (4.8) and the point 1 of Theorem 5.4. Indeed, by (4.9), the assumption **A5.2** holds and the condition (5.35) is satisfied because $\boldsymbol{b}_t + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\boldsymbol{A}_{t+1-j}\right)\boldsymbol{b}_{t-i} \geq (\epsilon_t^2, \dots, \epsilon_{t-q+1}^2)' \geq \omega_0(\eta_t^2, \dots, \eta_{t-q+1}^2)'$ and $\mathbb{P}(\eta_t^2 > 0, \dots, \eta_{t-q+1}^2 > 0) = \mathbb{P}(\eta_1^2 > 0)^q > 0$.

Now we show the second assertion, let i_0 such that $\phi_0 = \phi_{i_0}(\boldsymbol{\vartheta}_0) > 0$. We have

$$\tau_t^2 \geq 1 + a_0\phi_0\epsilon_{t-i_0}^2\tau_{t-i_0}^2 = 1 + a_0\phi_0\epsilon_{t-i_0}^2 + a_0^2\phi_0^2\epsilon_{t-i_0}^2\epsilon_{t-2i_0}^2 + \cdots$$

We thus have $|r_t|^s \geq (a_0\phi_0)^{ks/2}|\epsilon_t|^s|\epsilon_{t-i_0}|^s\cdots|\epsilon_{t-ki_0}|^s$ for any $s > 0$ and any $k \geq 1$. By

A4.2, for any $s > 0$ there exists $k \geq 1$ such that $\mathbb{E}|\epsilon_t|^{ks} = \infty$. The conclusion follows from Lemma 4.1.

7.4 Proof of Theorem 4.1

Let

$$\mathbf{l}_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n} \ell_t(\boldsymbol{\theta}), \qquad \ell_t(\boldsymbol{\theta}) = \frac{r_t^2}{\tau_t^2\sigma_t^2} + \log\sigma_t^2 + \log\tau_t^2,$$

where $\tau_t^2 = \tau_t^2(\boldsymbol{\theta}) = 1 + a\sum_{i=1}^{q}\phi_i(\boldsymbol{\vartheta})r_{t-i}^2$ and $\sigma_t^2 = \sigma_t^2(\boldsymbol{\theta}) = \omega + \alpha\epsilon_{t-1}^2(\boldsymbol{\theta}) + \beta\sigma_{t-1}^2$, with $\epsilon_t^2(\boldsymbol{\theta}) = r_t^2/\tau_t^2$. Note that σ_t^2 is well defined because

$$\sum_{k=0}^{\infty}\beta^k\epsilon_{t-k-1}^2(\boldsymbol{\theta}) \leq \sum_{k=0}^{\infty}\beta^k r_{t-k-1}^2 < \infty, \quad a.s.$$

Since $\limsup_{k\to\infty}\frac{1}{k}\log\left(\beta^k r_{t-k-1}^2\right) = \log\beta < 0$ by the second inequality in (4.12), the convergence of the latter infinite sum follows from the Cauchy rule.

However, contrary to the standard GARCH case, the limiting criterion $\mathbb{E}\ell_t(\boldsymbol{\theta})$ might not be defined, even at $\boldsymbol{\theta}_0$, because if **A4.2** holds the observed process has no moment.

The proof therefore relies on the following intermediate results which, contrary to the standard GARCH case (see for instance Francq and Zakoian (2019) Section 7.4), do not involve a limiting criterion :

i) $\lim_{n\to\infty}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\mathbf{l}_n(\boldsymbol{\theta}) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta})| = 0, \quad a.s.$

ii) if $\sigma_t^2(\boldsymbol{\theta})\tau_t^2(\boldsymbol{\theta}) = \sigma_t^2(\boldsymbol{\theta}_0)\tau_t^2(\boldsymbol{\theta}_0) \quad a.s.$, then $\boldsymbol{\theta} = \boldsymbol{\theta}_0$,

iii) if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, then $\mathbb{E}\{\ell_t(\boldsymbol{\theta}) - \ell_t(\boldsymbol{\theta}_0)\} > 0$,

iv) any $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ has a neighborhood $V(\boldsymbol{\theta})$ such that

$$\liminf_{n\to\infty}\left(\inf_{\boldsymbol{\theta}^*\in V(\boldsymbol{\theta})\cap\boldsymbol{\Theta}}\tilde{\mathbf{l}}_n(\boldsymbol{\theta}^*) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta}_0)\right) > 0 \quad a.s.$$

We first show *i)*. We have

$$\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\mathbf{l}_n(\boldsymbol{\theta}) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta})| \\
&\leq \frac{1}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\{\left|\log\left(\frac{\sigma_t^2}{\tilde{\sigma}_t^2}\right)\right| + r_t^2\frac{|\sigma_t^2 - \tilde{\sigma}_t^2|}{\tilde{\tau}_t^2\sigma_t^2\tilde{\sigma}_t^2} + \left|\log\left(\frac{\tau_t^2}{\tilde{\tau}_t^2}\right)\right| + r_t^2\frac{|\tau_t^2 - \tilde{\tau}_t^2|}{\tilde{\tau}_t^2\tau_t^2\sigma_t^2}\right\}.
\end{aligned}$$

Noting that $\tau_t^2 = \tilde{\tau}_t^2$ for $t > q$, the last two terms asymptotically vanish and we have, for $t > q$,

$$|\sigma_t^2 - \tilde{\sigma}_t^2| \leq \beta|\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2| \leq \beta^{t-q}|\sigma_q^2 - \tilde{\sigma}_q^2|. \tag{4.17}$$

Using the inequality $|\log(x/y)| \leq |x-y|/(x \vee y)$ for $x, y > 0$, we deduce

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |\mathbf{l}_n(\boldsymbol{\theta}) - \tilde{\mathbf{l}}_n(\boldsymbol{\theta})| \leq \frac{K}{n}\sum_{t=1}^{n} \rho^t(1+r_t^2),$$

where $\rho = \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \beta < 1$ and, in view of (4.17), K is $\mathcal{F}_q$-measurable random variable. By the first inequality in (4.12), we have

$$\limsup_{t\to\infty} \frac{1}{t}\log \rho^t r_t^2 = \log\rho + \limsup_{t\to\infty}\frac{1}{t}\log r_t^2 = \log\rho < 0, \quad \text{a.s.}$$

from which it follows that $\rho^t r_t^2 \to 0$, and then $\rho^t(1+r_t^2) \to 0$, a.s. as $t \to \infty$. By Cesàro's lemma $\frac{1}{n}\sum_{t=1}^{n}\rho^t(1+r_t^2) \to 0$, a.s. Because K is fixed (independent of n), the conclusion follows.

Next we turn to *ii)*. Letting $V_t(\boldsymbol{\theta}) = \sigma_t^2(\boldsymbol{\theta})\tau_t^2(\boldsymbol{\theta})$, we have

$$\begin{aligned} V_t(\boldsymbol{\theta}) =& \left\{\omega + \alpha\frac{V_{t-1}(\boldsymbol{\theta}_0)}{\tau_{t-1}^2(\boldsymbol{\theta})}\eta_{t-1}^2 + \beta\sigma_{t-1}^2(\boldsymbol{\theta})\right\} \\ &\times\left\{1 + a\phi_1(\boldsymbol{\vartheta})V_{t-1}(\boldsymbol{\theta}_0)\eta_{t-1}^2 + a\sum_{i=2}^{q}\phi_i(\boldsymbol{\vartheta})r_{t-i}^2\right\} \\ :=& b_{t-1}(\boldsymbol{\theta})\eta_{t-1}^4 + c_{t-1}(\boldsymbol{\theta})\eta_{t-1}^2 + d_{t-1}(\boldsymbol{\theta}), \end{aligned}$$

where $b_{t-1}(\boldsymbol{\theta}), c_{t-1}(\boldsymbol{\theta}), d_{t-1}(\boldsymbol{\theta}) \in \mathcal{F}_{t-2}$. By Assumption **A4.5**, $V_t(\boldsymbol{\theta}) = V_t(\boldsymbol{\theta}_0)$ entails $b_{t-1}(\boldsymbol{\theta}) = b_{t-1}(\boldsymbol{\theta}_0), c_{t-1}(\boldsymbol{\theta}) = c_{t-1}(\boldsymbol{\theta}_0)$ and $d_{t-1}(\boldsymbol{\theta}) = d_{t-1}(\boldsymbol{\theta}_0)$. First consider the case $\phi_1(\boldsymbol{\vartheta}_0) \neq 0$. The equality $b_{t-1}(\boldsymbol{\theta}) = b_{t-1}(\boldsymbol{\theta}_0)$ then implies

$$\frac{\tau_{t-1}^2(\boldsymbol{\theta})}{\tau_{t-1}^2(\boldsymbol{\theta}_0)} = \frac{a\alpha\phi_1(\boldsymbol{\vartheta})}{a_0\alpha_0\phi_1(\boldsymbol{\vartheta}_0)} := c. \tag{4.18}$$

Now $\tau_{t-1}^2(\boldsymbol{\theta}) = c\tau_{t-1}^2(\boldsymbol{\theta}_0)$ writes

$$\sum_{i=1}^{q}\{a\phi_i(\boldsymbol{\vartheta}) - ca_0\phi_i(\boldsymbol{\vartheta}_0)\}V_{t-i}(\boldsymbol{\theta}_0)\eta_{t-i}^2 = c - 1$$

which, because $V_{t-i}(\boldsymbol{\theta}_0) > 0$ and by already given arguments, entails $a\phi_i(\boldsymbol{\vartheta}) = a_0\phi_i(\boldsymbol{\vartheta}_0)$, for $i = 1, \ldots, q$ and $c = 1$. Because the $\phi_i(\cdot)$'s sum up to 1, we deduce $a = a_0$ and then, by Assumptions **A4.4** and **A4.6**, $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$. By (4.18) we also have $\alpha = \alpha_0$. In view of $c_{t-1}(\boldsymbol{\theta}) = c_{t-1}(\boldsymbol{\theta}_0)$ we obtain $\omega = \omega_0$. In view of $d_{t-1}(\boldsymbol{\theta}) = d_{t-1}(\boldsymbol{\theta}_0)$ we get $\beta\sigma_{t-1}^2(\boldsymbol{\theta}) = \beta_0\sigma_{t-1}^2(\boldsymbol{\theta}_0)$ from which we deduce $\beta = \beta_0$ by already given arguments. Now consider the case $\phi_1(\boldsymbol{\vartheta}_0) = 0$. The equality $b_{t-1}(\boldsymbol{\theta}) = b_{t-1}(\boldsymbol{\theta}_0)$ then implies $\phi_1(\boldsymbol{\vartheta}) = 0$, and $c_{t-1}(\boldsymbol{\theta}) = c_{t-1}(\boldsymbol{\theta}_0)$ in turn implies $\tau_{t-1}^2(\boldsymbol{\theta}_0) = c\tau_{t-1}^2(\boldsymbol{\theta})$ with $c = \alpha_0/\alpha$, which allows us to conclude by the previous arguments. Step *ii)* is thus established.

The proof of *iii)* and *iv)* uses the same arguments as those of steps *iii)* and *iv)* in Theorem 3.3. We also complete the proof with the same arguments used in the conclusion

of the latter.

7.5 Proof of Theorem 4.2

The proof relies on the following steps. There exists a neighborhood $\mathcal{V}(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ such that

$$\text{a) } \mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\|\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})\nabla'_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})\|<\infty, \quad \mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\ell_t(\boldsymbol{\theta})\right\|<\infty,$$

$$\text{b) } \boldsymbol{J} \text{ is invertible and } \sqrt{n}\nabla_{\boldsymbol{\theta}}\mathbf{l}_n(\boldsymbol{\theta}_0)\xrightarrow{\mathcal{L}}\mathcal{N}(0,(\kappa_\eta-1)\boldsymbol{J}),$$

$$\text{c) } \sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|n^{-1/2}\sum_{t=1}^{n}\left\{\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})-\nabla_{\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\right\}\right\|\to 0 \text{ in probability as } n\to\infty,$$

$$\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|n^{-1}\sum_{t=1}^{n}\left\{\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\ell_t(\boldsymbol{\theta})-\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\right\}\right\|\to 0 \text{ in probability as } n\to\infty,$$

$$\text{d) } n^{-1}\sum_{t=1}^{n}\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta}^*)\to\boldsymbol{J}(i,j) \text{ a.s.} \quad \text{for any } \boldsymbol{\theta}^* \text{ between } \widehat{\boldsymbol{\theta}}_n \text{ and } \boldsymbol{\theta}_0.$$

We have

$$\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})=\left(1-\frac{V_t(\boldsymbol{\theta}_0)\eta_t^2}{V_t}\right)\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}V_t,$$

$$\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\ell_t(\boldsymbol{\theta})=\left(1-\frac{V_t(\boldsymbol{\theta}_0)\eta_t^2}{V_t}\right)\frac{1}{V_t}\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}V_t(\boldsymbol{\theta})+\left(2\frac{V_t(\boldsymbol{\theta}_0)\eta_t^2}{V_t}-1\right)\frac{1}{V_t^2}\nabla_{\boldsymbol{\theta}}V_t\nabla'_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta}).$$

To establish a), by the Hölder inequality it thus suffices to show

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left|\frac{V_t(\boldsymbol{\theta}_0)}{V_t}\right|^{2p_1}<\infty, \quad \mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|\frac{1}{V_t^2}\nabla_{\boldsymbol{\theta}}V_t\nabla'_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta})\right\|^{q_1}<\infty, \tag{4.19}$$

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left|\frac{V_t(\boldsymbol{\theta}_0)}{V_t}\right|^{p_2}<\infty, \quad \mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|\frac{1}{V_t}\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}V_t(\boldsymbol{\theta})\right\|^{q_2}<\infty, \tag{4.20}$$

for some conjugate numbers $p_i,q_i>1$ such that $p_i^{-1}+q_i^{-1}=1$, with $i=1,2$. We have $\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta})=\frac{1}{\tau_t^2(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}\tau_t^2(\boldsymbol{\theta})+\frac{1}{\sigma_t^2(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta})$ and, omitting the dependence with respect to $\boldsymbol{\theta}$, for $a,\alpha>0$ and $\beta\in(0,1)$ (which holds in a neighborhood of $\boldsymbol{\theta}_0$),

$$|\tau_t^{-2}\nabla_a\tau_t^2|\le 1/a, \quad |\sigma_t^{-2}\nabla_\alpha\sigma_t^2|\le\frac{1}{\alpha}, \quad |\sigma_t^{-2}\nabla_\omega\sigma_t^2|\le 1/\{\omega(1-\beta)\},$$

$$|\sigma_t^{-2}\nabla_a\sigma_t^2|\le\sigma_t^{-2}\alpha\sum_{k\ge 0}\beta^k\epsilon^2_{t-k-1}|\tau^{-2}_{t-k-1}\nabla_a\tau^2_{t-k-1}|\le\frac{1}{a}.$$

Let I the set of the indices $i\in\{1,\dots,q\}$ such that $\phi_i(\boldsymbol{\vartheta}_0)>0$. Using **A4.11** and the continuity of $\phi_i(\cdot)$, I is also the set of the indices $i\in\{1,\dots,q\}$ such that $\phi_i(\boldsymbol{\vartheta})>0$ for

$\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}_0)$, when $\mathcal{V}(\boldsymbol{\theta}_0)$ is small enough. We thus obtain for $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}_0)$

$$\|\tau_t^{-2}\nabla_{\boldsymbol{\vartheta}}\tau_t^2\| \leq \sum_{i\in I}\|\nabla_{\boldsymbol{\vartheta}}\log\phi_i(\boldsymbol{\vartheta})\|,$$
$$\|\sigma_t^{-2}\nabla_{\boldsymbol{\vartheta}}\sigma_t^2\| \leq \sigma_t^{-2}\alpha\sum_{k\geq 0}\beta^k\epsilon_{t-k-1}^2\|\tau_{t-k-1}^{-2}\nabla_{\boldsymbol{\vartheta}}\tau_{t-k-1}^2\| \leq \sum_{i\in I}\|\nabla_{\boldsymbol{\vartheta}}\log\phi_i(\boldsymbol{\vartheta})\|.$$

Moreover, for all $s_0 \in (0,1)$, using $x/(1+x) \leq x^{s_0}$ when $x \geq 0$,

$$\begin{aligned}|\sigma_t^{-2}\nabla_\beta\sigma_t^2| &= \sigma_t^{-2}\sum_{k\geq 0}(k+1)\beta^k(\omega+\alpha\epsilon_{t-k-2}^2)\\ &\leq \frac{1}{(1-\beta)^2}+\frac{1}{\beta}\sum_{k\geq 0}(k+1)\frac{\alpha\beta^{k+1}\epsilon_{t-k-2}^2}{\omega+\alpha\beta^{k+1}\epsilon_{t-k-2}^2}\\ &\leq \frac{1}{(1-\beta)^2}+\frac{1}{\beta}\sum_{k\geq 0}(k+1)\left(\frac{\alpha\beta^{k+1}\epsilon_{t-k-2}^2}{\omega}\right)^{s_0}.\end{aligned}$$

The inequality

$$\frac{\tau_t^2(\boldsymbol{\theta}_0)}{\tau_t^2(\boldsymbol{\theta})} \leq 1+\frac{a_0}{a}\sum_{i\in I}\frac{\phi_i(\boldsymbol{\vartheta}_0)}{\phi_i(\boldsymbol{\vartheta})} \quad \forall \boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0), \tag{4.21}$$

A4.11 and (4.9) entail $\mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}|\epsilon_t(\boldsymbol{\theta})|^s < \infty$. It follows that there exist $K \in (0,\infty)$ and $\rho \in (0,1)$ such that, for all $q_1 > 1$ and s_0 small enough,

$$\left\|\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left|\sigma_t^{-2}\nabla_\beta\sigma_t^2\right|\right\|_{2q_1} \leq K+K\sum_{k\geq 0}k\rho^k\left\|\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}|\epsilon_{t-k-2}(\boldsymbol{\theta})|^{2s_0}\right\|_{2q_1} < \infty.$$

The existence of the second expectation in (4.19) follows.

Let $\iota > 0$ and $\mathcal{V}(\boldsymbol{\theta}_0)$ such that $\beta_0/\beta < 1+\iota$. For all $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}_0)$, using (4.21) and already given arguments, there exist a generic $K \in (0,\infty)$ such that, for $s_0 \in (0,1)$,

$$\frac{\sigma_t^2(\boldsymbol{\theta}_0)}{\sigma_t^2(\boldsymbol{\theta})} \leq K+K\sum_{i=0}^{\infty}\frac{\beta_0^i\frac{r_{t-i-1}^2}{\tau_{t-i-1}(\boldsymbol{\theta}_0)}}{\omega+\alpha\beta^i\frac{r_{t-i-1}^2}{\tau_{t-i-1}(\boldsymbol{\theta})}} \leq K+K\sum_{i=0}^{\infty}(1+\iota)^i\beta^{is_0}\epsilon_{t-i-1}^{2s_0}(\boldsymbol{\theta}).$$

By choosing ι such that $\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}(1+\iota)\beta^{s_0} < 1$ and s_0 sufficiently small, the expectation of the supremum over $\mathcal{V}(\boldsymbol{\theta}_0)$ of the last sum is finite. The existence of the first expectations in (4.19) and (4.20) follows, for all values of p_1 and p_2.

Turning to second-order derivatives, we have

$$\frac{1}{V_t}\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}^2V_t = \frac{1}{\sigma_t^2}\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}^2\sigma_t^2+\frac{1}{\tau_t^2}\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}^2\tau_t^2+\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}\tau_t^2\nabla_{\boldsymbol{\theta}'}\sigma_t^2+\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla_{\boldsymbol{\theta}'}\tau_t^2. \tag{4.22}$$

The matrix $\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\tau_t^2$ has the form

$$\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\tau_t^2 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sum_{i=1}^q \nabla_{\boldsymbol{\vartheta}}\phi_i(\boldsymbol{\vartheta}) r_{t-i}^2 \\ \mathbf{0} & \sum_{i=1}^q \nabla_{\boldsymbol{\vartheta}'}\phi_i(\boldsymbol{\vartheta}) r_{t-i}^2 & a\sum_{i=1}^q \nabla^2_{\boldsymbol{\vartheta}\boldsymbol{\vartheta}'}\phi_i(\boldsymbol{\vartheta}) r_{t-i}^2 \end{pmatrix}.$$

Hence by **A4.11** and already used arguments $\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \left\|\tau_t^{-2}\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\tau_t^2\right\|$ is bounded by a constant when $\mathcal{V}(\boldsymbol{\theta}_0)$ is sufficiently small. We similarly show that $\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \left\|\sigma_t^{-2}\nabla^2_{\boldsymbol{\theta}\boldsymbol{\theta}'}\sigma_t^2\right\|$ admits moments of any order, which, using the triangle and Cauchy-Schwarz inequalities in (4.22), allows to show the existence of the second expectation in (4.20) and to complete the proof of a).

Now we turn to b). Suppose there exists a vector $\boldsymbol{x} = (x_1, x_2, x_3, x_4, \boldsymbol{x}_5')' \in \mathbb{R}^{d+4}$ such that $\boldsymbol{x}'\boldsymbol{J}\boldsymbol{x} = 0$. Then, in view of $\nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0) = \sigma_t^2(\boldsymbol{\theta}_0)\nabla_{\boldsymbol{\theta}}\tau_t^2(\boldsymbol{\theta}_0) + \tau_t^2(\boldsymbol{\theta}_0)\nabla_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta}_0)$, we have

$$\begin{aligned} 0 =& \boldsymbol{x}'\nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0) \\ =& \sigma_t^2 \boldsymbol{x}' \left\{ (\nabla_{\boldsymbol{\theta}} a_0) \sum_{i=1}^q \phi_i(\boldsymbol{\vartheta}_0) r_{t-i}^2 + a_0 \sum_{i=1}^q \nabla_{\boldsymbol{\theta}}\phi_i(\boldsymbol{\vartheta}_0) r_{t-i}^2 \right\} \\ &+ \tau_t^2 \boldsymbol{x}' \left\{ \nabla_{\boldsymbol{\theta}}\omega_0 + \epsilon_{t-1}^2 \nabla_{\boldsymbol{\theta}}\alpha_0 - \alpha_0 \epsilon_{t-1}^2 \nabla_{\boldsymbol{\theta}} \log \tau_{t-1}^2 + \sigma_{t-1}^2 \nabla_{\boldsymbol{\theta}}\beta_0 + \beta_0 \nabla_{\boldsymbol{\theta}}\sigma_{t-1}^2 \right\} \\ :=& e_{t-1}\eta_{t-1}^4 + f_{t-1}\eta_{t-1}^2 + g_{t-1}, \quad a.s. \end{aligned} \tag{4.23}$$

where $e_{t-1}, f_{t-1}, g_{t-1} \in \mathcal{F}_{t-2}$. By Assumption **A4.5**, we must have $e_{t-1} = f_{t-1} = g_{t-1} = 0$, a.s. Therefore,

$$\begin{aligned} 0 = e_{t-1} =& \alpha_0 V_{t-1}\sigma_{t-1}^2 \boldsymbol{x}' \left\{ \phi_1(\boldsymbol{\vartheta}_0)\nabla_{\boldsymbol{\theta}} a_0 + a_0 \nabla_{\boldsymbol{\theta}}\phi_1(\boldsymbol{\vartheta}_0) \right\} \\ &+ a_0 \phi_1(\boldsymbol{\vartheta}_0) V_{t-1}\sigma_{t-1}^2 \boldsymbol{x}' \left\{ \nabla_{\boldsymbol{\theta}}\alpha_0 - \alpha_0 \nabla_{\boldsymbol{\theta}} \log \tau_{t-1}^2 \right\}, \end{aligned}$$

from which we deduce

$$\begin{aligned} & a_0\phi_1(\boldsymbol{\vartheta}_0)\alpha_0 \boldsymbol{x}'\nabla_{\boldsymbol{\theta}} \log \tau_{t-1}^2 \\ =& \alpha_0 \boldsymbol{x}' \left\{ \phi_1(\boldsymbol{\vartheta}_0)\nabla_{\boldsymbol{\theta}} a_0 + a_0 \nabla_{\boldsymbol{\theta}}\phi_1(\boldsymbol{\vartheta}_0) \right\} + a_0 \phi_1(\boldsymbol{\vartheta}_0) \boldsymbol{x}'\nabla_{\boldsymbol{\theta}}\alpha_0 := c. \end{aligned}$$

We thus have

$$a_0\phi_1(\boldsymbol{\vartheta}_0)\alpha_0 \boldsymbol{x}'\nabla_{\boldsymbol{\theta}}\tau_{t-1}^2 = c\tau_{t-1}^2,$$

that is,

$$a_0 \sum_{i=1}^q \left[\phi_1(\boldsymbol{\vartheta}_0)\alpha_0 \boldsymbol{x}' \left\{ a_0 \nabla_{\boldsymbol{\theta}}\phi_i(\boldsymbol{\vartheta}_0) + \phi_i(\boldsymbol{\vartheta}_0)\nabla_{\boldsymbol{\theta}} a_0 \right\} - c\phi_i(\boldsymbol{\vartheta}_0) \right] r_{t-i}^2 = c.$$

By **A4.5**, it can be shown that any equality of the form $\sum_{i=1}^\infty b_i r_{t-i}^2 = b_0$, where the b_i's are real constants, entails $b_i = 0$ for all $i \geq 0$. We thus have $c = 0$ and, since $a_0\alpha_0 > 0$,

$$\phi_1(\boldsymbol{\vartheta}_0) \left\{ x_4 \phi_i(\boldsymbol{\vartheta}_0) + a_0 \boldsymbol{x}_5' \nabla_{\boldsymbol{\vartheta}}\phi_i(\boldsymbol{\vartheta}_0) \right\} = 0, \quad i = 1, \ldots, q.$$

First suppose $\phi_1(\boldsymbol{\vartheta}_0) \neq 0$. Then, since $\sum_{i=1}^q \phi_i(\boldsymbol{\vartheta}_0) = 1$ and $\sum_{i=1}^q \nabla_{\boldsymbol{\vartheta}}\phi_i(\boldsymbol{\vartheta}_0) = \mathbf{0}$, we get $x_4 = 0$. Thus $\boldsymbol{x}_5' \left[\nabla_{\boldsymbol{\vartheta}}\phi_1(\boldsymbol{\vartheta}_0), \ldots, \nabla_{\boldsymbol{\vartheta}}\phi_q(\boldsymbol{\vartheta}_0)\right] = 0$, which by **A4.10** entails $\boldsymbol{x}_5 = \mathbf{0}$. The definition of c thus implies $x_2 = 0$. Turning back to (4.23), we obtain

$$0 = x_1 + x_3\sigma_{t-1}^2 + \beta(x_1\nabla_\omega\sigma_{t-1}^2 + x_3\nabla_\beta\sigma_{t-1}^2) = x_3(1+\beta)\sigma_{t-1}^2 + y_{t-2},$$

where $y_{t-2} \in \mathcal{F}_{t-3}$. Using again **A4.5**, we deduce $x_3 = 0$ and finally $x_1 = 0$. We have shown that $\boldsymbol{x} = \mathbf{0}$ and the proof of the first part of *b)* is now complete. We have

$$\sqrt{n}\nabla_{\boldsymbol{\theta}}\mathbf{l}_n(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}}\sum_{t=1}^n (1-\eta_t^2)\nabla_{\boldsymbol{\theta}}\log V_t(\boldsymbol{\theta}_0).$$

The convergence in distribution follows from the central limit theorem for square integrable stationary and ergodic martingale differences Billingsley (1961)).

Now we turn to *c)*. Note that

$$\begin{aligned}&\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\\ =&\frac{r_t^2}{V_t\tilde{V}_t}(V_t - \tilde{V}_t)\nabla_{\boldsymbol{\theta}}\log V_t + \left(1 - \frac{r_t^2}{\tilde{V}_t}\right)(\nabla_{\boldsymbol{\theta}}\log V_t - \nabla_{\boldsymbol{\theta}}\log\tilde{V}_t).\end{aligned} \tag{4.24}$$

We have, for t large enough, $\nabla_{\boldsymbol{\theta}}\tau_t^2 = \nabla_{\boldsymbol{\theta}}\tilde{\tau}_t^2$. Moreover, $\tilde{\sigma}_t^2 = \omega + \alpha\tilde{\epsilon}_{t-1}^2 + \beta\tilde{\sigma}_{t-1}^2$, where $\tilde{\epsilon}_t = r_t/\tilde{\tau}_t$, thus

$$\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2 = \nabla_{\boldsymbol{\theta}}\omega + \tilde{\epsilon}_{t-1}^2\nabla_{\boldsymbol{\theta}}\alpha + \alpha\nabla_{\boldsymbol{\theta}}\tilde{\epsilon}_{t-1}^2 + \tilde{\sigma}_{t-1}^2\nabla_{\boldsymbol{\theta}}\beta + \beta\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_{t-1}^2.$$

Therefore, for t large enough,

$$\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2 = (\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2)\nabla_{\boldsymbol{\theta}}\beta + \beta\{\nabla_{\boldsymbol{\theta}}\sigma_{t-1}^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_{t-1}^2\}.$$

By (4.17), this entails, for t large enough,

$$\left\|\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\right\| \leq Kt\beta^t,$$

and, given that $\tilde{\sigma}_t^2$ and σ_t^2 are uniformly bounded below, it is straightforward to deduce

$$\left\|\nabla_{\boldsymbol{\theta}}\log\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\log\tilde{\sigma}_t^2\right\| \leq K\beta^t\left\{t + \left\|\nabla_{\boldsymbol{\theta}}\log\sigma_t^2\right\|\right\}.$$

By $\nabla_{\boldsymbol{\theta}}\log V_t = \nabla_{\boldsymbol{\theta}}\log\sigma_t^2 + \nabla_{\boldsymbol{\theta}}\log\tau_t^2$, we also have

$$\left\|\nabla_{\boldsymbol{\theta}}\log V_t - \nabla_{\boldsymbol{\theta}}\log\tilde{V}_t\right\| \leq K\beta^t\left\{t + \left\|\nabla_{\boldsymbol{\theta}}\log\sigma_t^2\right\|\right\}, \tag{4.25}$$

for large enough t. Noting that $V_t - \tilde{V}_t = (\sigma_t^2 - \tilde{\sigma}_t^2)\tau_t^2$ for large t, we deduce from (4.24) that

$$\left\|\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\right\| \leq K\left\{1 + \epsilon_t^2(\boldsymbol{\theta})\right\}\left\{t + \left\|\nabla_{\boldsymbol{\theta}}\log V_t\right\|\right\}\beta^t.$$

From the proof of a), we have

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}|\epsilon_t(\boldsymbol{\theta})|^{4s_0}<\infty \text{ and } \mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\|\nabla_{\boldsymbol{\theta}}\log V_t\|^{2s_0}<\infty$$

for sufficiently small $s_0\in(0,1)$. By the triangle and Hölder inequalities, for $K\in(0,\infty)$ and $\rho\in(0,1)$ we then have

$$\mathbb{E}\left(\sum_{t=1}^{\infty}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})-\nabla_{\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\right\|\right)^s\leq K\sum_{t=1}^{\infty}(t^s+K)\rho^{ts}<\infty,$$

which entails that $\sum_{t=1}^{\infty}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)}\left\|\nabla_{\boldsymbol{\theta}}\ell_t(\boldsymbol{\theta})-\nabla_{\boldsymbol{\theta}}\tilde{\ell}_t(\boldsymbol{\theta})\right\|$ is finite almost surely. The convergence in the first part of c) follows. The second convergence can be established along the same lines.

Turning to d) we note that, by a) and the ergodic theorem

$$n^{-1}\sum_{t=1}^{n}\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta}_0)\to\boldsymbol{J}(i,j)\quad\text{a.s. as}\quad n\to\infty.$$

For all $\varepsilon>0$, by the same argument, the continuity of the second derivatives and the dominated convergence theorem, there exists a sufficiently small neighborhood $\mathcal{V}(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ such that

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta})}\left|\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta})-\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta}_0)\right|\\ =&\mathbb{E}\sup_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta})}\left|\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta})-\nabla^2_{\theta_i\theta_j}\ell_t(\boldsymbol{\theta}_0)\right|\leq\varepsilon.\end{aligned}$$

The point d) is thus a consequence of the consistency of $\widehat{\boldsymbol{\theta}}_n$.

The proof of the theorem then follows from a Taylor expansion of the criterion around $\boldsymbol{\theta}_0$ and classical arguments.

7.6 Proof of Proposition 4.2

Conditional on (r_t), the bootstrap statistics a_n^* and W_n^* remain random because they depend on $\epsilon_t^*\sim F_n$. The proof is standard and uses the same arguments as those of Theorem 2 and Proposition 2 in Francq and Zakoïan (2009).

7.7 Proof of Theorem 4.3

Throughout the proof, we assume a fixed trajectory $(r_t)_{t\in\mathbb{Z}}$, belonging to a subset of events of probability one such that (4.25) holds uniformly in $\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)$ and $\widehat{\boldsymbol{J}}_n^c\to\boldsymbol{J}$ as $n\to\infty$. This sequence exists by the arguments used to show c) and d) in thepProof of

Theorem 4.2. We thus have

$$\left(\widehat{\boldsymbol{J}}_n^c\right)^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\eta_t^{*2}-1\right)\frac{1}{\widetilde{V}_t}\nabla_{\boldsymbol{\theta}}\widetilde{V}_t(\widehat{\boldsymbol{\theta}}_n^c)=\boldsymbol{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{x}_{t,n}+o(1)$$

with $\boldsymbol{x}_{t,n}=(\eta_t^{*2}-1)\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta}_0)$. Conditional on (r_t), the previous quantity remains random because it depends on $\epsilon_t^*\sim F_n$. To establish (4.16), by the Wold-Cramer device, it thus suffices to show that for any $\boldsymbol{\lambda}\neq\mathbf{0}\in\mathbb{R}^4$

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\xrightarrow{\mathcal{L}}\mathcal{N}\left(0,(\kappa_\eta-1)\boldsymbol{\lambda}'\boldsymbol{J}\boldsymbol{\lambda}\right). \tag{4.26}$$

Note that, still conditioning by $(r_t)_{t\in\mathbb{Z}}$, for each n the random variables $\boldsymbol{\lambda}'\boldsymbol{x}_{1,n},\boldsymbol{\lambda}'\boldsymbol{x}_{2,n},\ldots$ are independent and centered, with finite second-order moments. By the Lindeberg's CLT for triangular arrays of square integrable martingale increments, it remains to show that

$$\frac{1}{n}\sum_{t=1}^{n}\mathrm{Var}\left(\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\right)\to(\kappa_\eta-1)\boldsymbol{\lambda}'\boldsymbol{J}\boldsymbol{\lambda}>0\quad\text{as } n\to\infty, \tag{4.27}$$

and for all $\varepsilon>0$

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(\left\{\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\right\}^2\mathbb{1}_{\{|\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}|\geq\sqrt{n}\varepsilon\}}\right)\to0\quad\text{as } n\to\infty. \tag{4.28}$$

In Lemma A.1 in Francq and Zakoian (2022), it has been shown that, for standard GARCH, the distribution F_n of the standardized residuals tends to the (unconditional) distribution F of η_t. More precisely, for any almost everywhere continuous function h such that $|h(x|\leq ax^4+b$ where $a,b>0$, for almost all realization $(r_t)_{t\in\mathbb{Z}}$ we have

$$\int h(x)F_n(dx)\to\int h(x)F(dx)\text{ as } n\to\infty. \tag{4.29}$$

It can be assumed that $(r_t)_{t\in\mathbb{Z}}$ is such that (4.29) holds. Since $\eta_t^*\sim F_n$, given $(r_t)_{t\in\mathbb{Z}}$, under H_0 we have

$$\mathbb{E}\eta_t^*=0,\quad\mathbb{E}\eta_t^{*2}=1\quad\text{and}\quad\mathbb{E}\eta_t^{*i}=\frac{1}{n-n_0}\sum_{k=n_0+1}^{n}\widehat{\eta}_k^{0i}\to\mathbb{E}\eta_1^i\quad\text{for } i\leq4$$

as $n\to\infty$. For t fixed we then have

$$\begin{aligned}\mathrm{Var}(\boldsymbol{\lambda}'\boldsymbol{x}_{t,n})&=\left\{\boldsymbol{\lambda}'\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta}_0)\right\}^2\left(\frac{1}{n-n_0}\sum_{k=n_0+1}^{n}\left(\widehat{\eta}_k^0\right)^4-1\right)\\&\to\left\{\boldsymbol{\lambda}'\frac{1}{V_t}\nabla_{\boldsymbol{\theta}}V_t(\boldsymbol{\theta}_0)\right\}^2(\kappa_\eta-1)\quad\text{as } n\to\infty,\end{aligned}$$

from which (4.27) follows.

Given (r_t), when $\boldsymbol{\lambda}' \frac{1}{V_t} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0) \neq 0$ we have

$$\begin{aligned} &\mathbb{E}\left\{\boldsymbol{\lambda}' \boldsymbol{x}_{t,n}\right\}^2 \mathbb{1}_{\{|\boldsymbol{\lambda}' \boldsymbol{x}_{t,n}| \geq \sqrt{n}\varepsilon\}} \\ &= \left\{\boldsymbol{\lambda}' \frac{1}{V_t} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\right\}^2 \mathbb{E}\left\{ \left|\eta_t^{*2} - 1\right|^2 \mathbb{1}_{\left(|\eta_t^{*2}-1| \geq \frac{\sqrt{n}\varepsilon}{\left|\boldsymbol{\lambda}' \frac{1}{V_t} \nabla_{\boldsymbol{\theta}} V_t(\boldsymbol{\theta}_0)\right|}\right)} \right\}. \end{aligned} \tag{4.30}$$

For any $A > 0$ there exists n_A such that if $n > n_A$ then the expectation in the right-hand side of (4.30) is bounded by

$$\mathbb{E}\left|\eta_t^{*2} - 1\right|^2 \mathbb{1}_{\left\{|\eta_t^{*2}-1| \geq A\right\}}.$$

By (4.29), this term tends as $n \to \infty$ to

$$\int_{|x^2-1| \geq A} \left|x^2 - 1\right|^2 F(dx)$$

which is arbitrarily small when A is sufficiently large. We then obtain (4.28) by the Cesàro Mean Theorem. The convergence (4.16) follows. The second convergence is obtained by noting that $\frac{1}{\sigma^2} N^2 \mathbb{1}_{N \geq 0} \sim \frac{1}{2}\delta_0 + \frac{1}{2}\chi_1^2$.

Under H_1 and the conditions given in the theorem, a careful examination of the proof of Lemma A.1 in Francq and Zakoian (2022) shows that (4.29) holds if F denotes the marginal distribution of $r_t/\sigma_t(\boldsymbol{\theta}_G)$. It follows that

$$\left(\widehat{\boldsymbol{J}}_n^c\right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\eta_t^{*2} - 1\right) \frac{1}{\widetilde{V}_t} \nabla_{\boldsymbol{\theta}} \widetilde{V}_t(\widehat{\boldsymbol{\theta}}_n^c) = O_P(1),$$

and thus $\sqrt{n}\widehat{a}_n^* = O_P(1)$. Since $\sqrt{n}\widehat{a}_n \to \infty$ as $n \to \infty$, we have $P(\sqrt{n}\widehat{a}_n^* \geq \sqrt{n}\widehat{a}_n) \to 0$ as $n \to \infty$.

Chapter 5

On the growth rate of superadditive processes and the stability of functional GARCH models

Abstract

We extend the result of Kesten (1975) on the growth rate of random walks with stationary increments to superadditive processes. We show that superadditive processes which remain positive after a certain time diverge at least linearly to infinity. Our proof relies on new techniques based on concepts from ergodic theory. Different versions of this result are also given, generalizing Lemma 3.4 of Bougerol and Picard (1992) on the contraction property of products of random matrices. We use our results to provide necessary and sufficient conditions for the stability of a class of Stochastic Recurrent Equations (SRE) with positive coefficients in the space of continuous functions with compact support, including continuous functional GARCH models.

Contents

1 Introduction 82
2 The growth rate of superadditive processes 83
3 Stationarity of fGARCH models in $\mathcal{C}^0$ 94
4 Perspective 101
5 Appendix: Complementary Proofs 101

1 Introduction

Let $(\Omega, \mathcal{B}, \mu, T)$ a measure-preserving dynamical system, i.e $(\Omega, \mathcal{B}, \mu)$ is a probability space and for all $A \in \mathcal{B}$, $\mu(T^{-1}(A)) = \mu(A)$. A $\mathcal{B}$-measurable sequence $\{\boldsymbol{S}_n\}_{n\geq 1}$ with value in $(-\infty, \infty]$ is said to be superadditive if

$$\text{for all } \; n, s \in \mathbb{N}^* \;\; \boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s} \;\; a.s. \tag{5.1}$$

A subadditive sequence is defined as the opposite of a superadditive process. Since their introduction by Hammersley and Welsh (1965), one of the most significant contributions to the study of subadditive stochastic processes is the Kingman's subadditive ergodic theorem (see Kingman (1973)). Kingman showed that if $\{\boldsymbol{S}_n\}_{n\geq 1}$ is a superadditive process and $\boldsymbol{S}_1^-$ is integrable then $n^{-1}\boldsymbol{S}_n$ converges $a.s.$ to a function $\boldsymbol{S} : \Omega \to \overline{\mathbb{R}}$. Moreover, $\boldsymbol{S}^-$ is integrable and

$$\int \boldsymbol{S} d\mu = \lim_{n\to\infty} \frac{1}{n} \int \boldsymbol{S}_n d\mu = \sup_n \frac{1}{n} \int \boldsymbol{S}_n d\mu \in (-\infty, +\infty].$$

This result is a generalization of the well-known ergodic theorem of Birkhoff for additive processes, such that for all $n, s \in \mathbb{N}^* \;\; \boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n = \boldsymbol{S}_{n+s} \;\; a.s.$ For these additive processes, even if the integrability condition does not hold, Kesten (1975) (see also Atkinson (1976)) showed that

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \;\; a.s \text{ on the set } \; \{\boldsymbol{S}_n \to \infty, n \to \infty\}. \tag{5.2}$$

This well-known result has found numerous applications in ergodic theory and was a precursor in the study of the recurrence of stationary random walks, see Atkinson (1976), Berbee (1981) and Schmidt (2006)). A similar result under an integrability condition has been obtained by Bougerol and Picard (1992b, Lemma 3.4) for the product of random matrices, which characterizes the case where the so-called top-Lyapunov coefficient is negative. As in Bougerol and Picard (1992b), this contraction property is often used to establish necessary and sufficient conditions for the existence of stationary solutions for Stochastic Recurrence Equations (SRE) in $\mathbb{R}^n$.

In this paper, we extend Kesten's result to superadditive processes by showing that a superadditive process that stays positive for a certain period grows at least linearly to infinity. As a corollary, we deduce the lemma 3.4 of Bougerol and Picard (1992b). Our results provide a characterization of the top-Lyapunov's exponent sign for a class of discrete-time dynamical systems. The top-Lyapunov exponent is used to quantify the stability or instability of a system, and is often associated with stability when it is negative. For instance, we use our result to provide, under mild conditions, a necessary and sufficient condition for the existence of stationary solutions of functional GARCH models in the space of continuous functions introduced by Aue et al. (2017) and Hörmann et al. (2013).

The rest of the paper is organized as follows. Section 2 is reserved for the main results. The study of the existence of stationary solution of functional GARCH models is the object of Section 3. Section 4 discusses perspectives for future work.

2 The growth rate of superadditive processes

Let us start with some remainders and conventions. A set $I \in \mathcal{B}$ is said to be invariant if $\mu\left(I \Delta T^{-1}(I)\right)=0$. The invariant σ-algebra $\mathcal{I}_\mu$ is the collection of all such invariant sets I. It is easy to verify that for all $A \in \mathcal{B}, \quad \mu(A)=1$ implies that $A \in \mathcal{I}_\mu$.

We set $\boldsymbol{S}_0 = 0$ throughout the paper. The convention that $\inf \varnothing = \infty$ and $\sup \varnothing = -\infty$ is used, a sum over an empty set will be equal to zero, and $T^0 = id_\Omega$. Pointwise convergence will be denoted by $\xrightarrow{\text{pw}}$. For all measurable functions Y from Ω to a measurable space $(F, \mathcal{F})$, and all $A \in \mathcal{F}$ and $(B, C) \in \mathcal{B}^2$, we say that:

$$\begin{aligned} &Y \in A \ \ a.s \ \ on \quad B \quad &&\text{if} \ \ \mu(\{Y \in A\} \cap B)=\mu(B), \\ &C \subset B \ \ a.s &&\text{if} \quad \mathbb{1}_C \leq \mathbb{1}_B \ \ a.s. \end{aligned}$$

Remark that results obtained for superadditive processes can be easily adapted for subadditive processes. Let us state our main result.

Theorem 5.1. *Let $\{\boldsymbol{S}_n\}_{n\geq 1}$ be a superadditive sequence and let $\boldsymbol{\tau}_0 = \sup_{n\in\mathbb{N}}\{n : \quad \boldsymbol{S}_n \leq 0\}$. We have*

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \ \ a.s \quad on \ \left\{\boldsymbol{\tau}_0 < \infty\right\}.$$

Noting that $\{\boldsymbol{\tau}_0 < \infty\} = \liminf\{\boldsymbol{S}_n > 0\}$, Theorem 5.1 includes Kesten's result for additive sequences. Unlike Kesten's assumption that $\boldsymbol{S}_n$ goes to infinity, we only require that the process is positive for sufficiently large values of n.

We need two technical lemmas before proving the theorem.

Lemma 5.1. *Let $\{\boldsymbol{S}_n\}_{n\geq 1}$ be a real valued superadditive sequence and let $\boldsymbol{\tau}_0 = \sup_{n\in\mathbb{N}}\{n : \quad \boldsymbol{S}_n \leq 0\}$. We have*

$$\liminf \boldsymbol{S}_n > 0 \ \ a.s \quad on \ \left\{\boldsymbol{\tau}_0 < \infty\right\}. \tag{5.3}$$

Proof. Let $A = \{\liminf \boldsymbol{S}_n = 0, \ \ \boldsymbol{\tau}_0 < \infty\}$, it is clear that (5.3) is equivalent to $\mu(A) = 0$. We argue by contradiction: suppose that $\mu(A) > 0$, let

$$V = A \cap \{\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s} \quad \text{for all} \ \ n, s \in \mathbb{N}^*\}.$$

By countability of $\mathbb{N}^2$, we have

$$\mu(\{\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s} \quad \text{for all} \ \ n, s \in \mathbb{N}^*\}) = 1,$$

so we have $\mu(V) = \mu(A) > 0$. By Birkhoff's ergodic theorem we have

$$g_n := n^{-1}\sum_{k=1}^{n} \mathbb{1}_V \circ T^k \longrightarrow g := \mathbb{E}^\mu(\mathbb{1}_V | \mathcal{I}_\mu) \ \ a.s \ \ as \quad n \to \infty.$$

Since $\mathbb{E}^{\mu}(g) = \mu(V)$, then $\mu(\{g > 0\}) = \mu(\{\liminf g_n > 0\}) > 0$ and so $\{\liminf g_n > 0\} \neq \varnothing$. Letting $\omega \in \{\liminf g_n > 0\}$, this implies that $\{k : \ T^k(\omega) \in V\}$ is not finite, and so there is a strictly increasing sequence of integers $\{n_k(\omega)\}_{n\geq 1}$ such that $T^{n_k}(\omega) \in V$. Since $\omega' := T^{n_1}(\omega) \in V$ and $V \subset \{\boldsymbol{\tau}_0 < \infty\}$, then $\boldsymbol{\tau}_0(\omega') < \infty$. Let $p(\omega)$ such that $s := n_p - n_1 \geq \boldsymbol{\tau}_0(\omega') + 1$. Since $s > \boldsymbol{\tau}_0(\omega')$ it follows that

$$\boldsymbol{S}_s(\omega') > 0. \tag{5.4}$$

The fact that $T^{n_p}(\omega) \in V$ implies that

$$\liminf \boldsymbol{S}_n(T^{n_p}(\omega)) = 0. \tag{5.5}$$

By the fact that $\omega' \in V$, for all $n \geq s$,

$$\boldsymbol{S}_s(\omega') + \boldsymbol{S}_{n-s}\left(T^{n_p}(\omega)\right) = \boldsymbol{S}_s(\omega') + \boldsymbol{S}_{n-s} \circ T^s(\omega') \leq \boldsymbol{S}_n(\omega'), \tag{5.6}$$

It follows by (5.4)-(5.6), that $\liminf \boldsymbol{S}_n(\omega') \geq \boldsymbol{S}_s(\omega') > 0$, which contradicts the fact that $\omega' \in V$ and thus $\mu(A) = 0$. □

The second lemma gives a property on series with terms in $\{0, 1\}$. Let $p > 0$ and let $\boldsymbol{u} = \{\boldsymbol{u}_n\}_{n\geq 0}$ a sequence of elements of $\{0, 1\}^{\mathbb{N}}$. For all $n > p$, define $\{\boldsymbol{v}_k^n(\boldsymbol{u})\}_{k=0}^{\infty}$ a sequence of elements of $\{n, n-1, \cdots, 0, \ -\infty\}^{\mathbb{N}}$ by

$$\begin{array}{l} \boldsymbol{v}_0^n(\boldsymbol{u}) = n \quad \text{and} \quad \text{for all} \quad k > 0 \quad \boldsymbol{v}_k^n(\boldsymbol{u}) = \sup\{r \in \mathbb{N} : r \leq \boldsymbol{v}_{k-1}^n(\boldsymbol{u}) - p, \ \boldsymbol{u}_r = 1\}. \\ \text{Define also} \quad q^n(\boldsymbol{u}) = \sup\{k : \ \boldsymbol{v}_k^n(\boldsymbol{u}) > -\infty\} \text{ and } s(\boldsymbol{u}) = \inf\{k : \ \boldsymbol{s}_k \geq p\}, \\ \text{where } \boldsymbol{s}_k := \displaystyle\sum_{i=0}^{k} \boldsymbol{u}_i. \end{array} \tag{5.7}$$

It is clear that the sequence $\{\boldsymbol{v}_k^n(\boldsymbol{u})\}_{k=0}^{\infty}$ is decreasing and becomes $-\infty$ eventually. Hence, the largest index k for which $\boldsymbol{v}_k^n(\boldsymbol{u})$ is finite, denoted as $q^n(\boldsymbol{u})$, is well-defined. Note also that

$$\text{for all } k < q^n(\boldsymbol{u}), \quad \boldsymbol{v}_k^n(\boldsymbol{u}) - \boldsymbol{v}_{k+1}^n(\boldsymbol{u}) \geq p. \tag{5.8}$$

If $\sum_{i=1}^{\infty} \boldsymbol{u}_i \geq p$ then $s(\boldsymbol{u})$ is finite and, since $\{\boldsymbol{u}_n\}_{n\geq 0}$ takes values in $\{0, 1\}$, there exists an integer n such that $\sum_{k=0}^{n} \boldsymbol{u}_k = p$, which implies that $\sum_{k=0}^{s(\boldsymbol{u})} \boldsymbol{u}_k = p$.

Lemma 5.2. *If* $\liminf n^{-1}\boldsymbol{s}_n > 0$ *then*

1. *for all* $n > s(\boldsymbol{u})$, $\boldsymbol{v}_{q^n(\boldsymbol{u})}^n(\boldsymbol{u}) \leq s(\boldsymbol{u})$,
2. $\liminf n^{-1} q^n(\boldsymbol{u}) > 0$.

Proof. For the first part, we have $s(\boldsymbol{u}) + 1 \geq p$ and then

$$\begin{aligned} \sum_{k=0}^{s(\boldsymbol{u})+1-p} \boldsymbol{u}_k = \sum_{k=0}^{s(\boldsymbol{u})} \boldsymbol{u}_k - \sum_{k=s(\boldsymbol{u})+1-p+1}^{s(\boldsymbol{u})} \boldsymbol{u}_k &= p - \sum_{k=s(\boldsymbol{u})+2-p}^{s(\boldsymbol{u})} \boldsymbol{u}_k \\ &\geq p - (p-1) = 1 > 0. \end{aligned}$$

It follows that there exists $r_0 \leq s(\boldsymbol{u}) + 1 - p$ such that $\boldsymbol{u}_{r_0} = 1$. Therefore, if $\boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u}) > s(\boldsymbol{u})$, i.e $s(\boldsymbol{u}) + 1 \leq \boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})$, then $r_0 \leq \boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u}) - p$. It follows that $r_0 \in \{r \in \mathbb{N} : r \leq \boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u}) - p,\ \boldsymbol{u}_r = 1\}$ and thus $\boldsymbol{v}^n_{q^n(\boldsymbol{u})+1} \neq -\infty$. This contradicts the definition of $q^n(\boldsymbol{u})$. We thus have shown 1. To show 2, noting that $\boldsymbol{v}^n_0(\boldsymbol{u}), \boldsymbol{v}^n_1(\boldsymbol{u}), \cdots, \boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})$ is a strictly decreasing sequence of integers with $\boldsymbol{v}^n_0(\boldsymbol{u}) = n$, one has

$$s_n = \sum_{k=\boldsymbol{v}^n_1(\boldsymbol{u})+1}^{\boldsymbol{v}^n_0(\boldsymbol{u})} \boldsymbol{u}_k + \sum_{k=\boldsymbol{v}^n_2(\boldsymbol{u})+1}^{\boldsymbol{v}^n_1(\boldsymbol{u})} \boldsymbol{u}_k + \cdots + \sum_{k=0}^{\boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})} \boldsymbol{u}_k. \tag{5.9}$$

Since, by definition, $\boldsymbol{v}^n_{l+1}(\boldsymbol{u})$ is the largest index l below $\boldsymbol{v}^n_l(\boldsymbol{u}) - p$ such that $\boldsymbol{u}_l = 1$ then $\boldsymbol{u}_k = 0$ for all $\boldsymbol{v}^n_{l+1}(\boldsymbol{u}) < k \leq \boldsymbol{v}^n_l(\boldsymbol{u}) - p$. We thus have

$$\sum_{k=\boldsymbol{v}^n_{l+1}(\boldsymbol{u})+1}^{\boldsymbol{v}^n_l(\boldsymbol{u})} \boldsymbol{u}_k = \sum_{k=\boldsymbol{v}^n_l(\boldsymbol{u})-p+1}^{\boldsymbol{v}^n_l(\boldsymbol{u})} \boldsymbol{u}_k \leq p, \text{ for all } l < q^n(\boldsymbol{u}). \tag{5.10}$$

By the first part of the lemma we have

$$\sum_{k=0}^{\boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})} \boldsymbol{u}_k \leq \sum_{k=0}^{s(\boldsymbol{u})} \boldsymbol{u}_k = p, . \tag{5.11}$$

It follows by (5.9), (5.10) and (5.11) that $\boldsymbol{s}_n \leq p(q^n(\boldsymbol{u}) + 1)$. Therefore,

$$\liminf n^{-1} q^n(\boldsymbol{u}) \geq \liminf n^{-1}(p^{-1}\boldsymbol{s}_n - 1) = p^{-1} \liminf n^{-1}\boldsymbol{s}_n > 0,$$

which concludes the proof. □

We are now ready to give the proof of the theorem.

Proof of Theorem 5.1. The proof does not follow that of Kesten and only uses the ergodic theorem as an external result. Since the real valued sequence $\tilde{\boldsymbol{S}} := \{\min(\boldsymbol{S}_n, n)\}_{n\geq 1}$ is superadditive and $\liminf\{\boldsymbol{S}_n > 0\} = \liminf\{\tilde{\boldsymbol{S}}_n > 0\}$, and $\boldsymbol{S} \geq \tilde{\boldsymbol{S}}$ then one can assume without loss of generality that $\boldsymbol{S}$ is a real valued process. By (5.3), it suffices to prove that

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s \quad \text{on} \quad \{\liminf \boldsymbol{S}_n > 0\}. \tag{5.12}$$

Let $B = \{\liminf n^{-1}\boldsymbol{S}_n = 0,\ \liminf \boldsymbol{S}_n > 0\}$. Since on $\{\liminf \boldsymbol{S}_n > 0\}$, one has $\liminf n^{-1}\boldsymbol{S}_n = 0$ or $\liminf n^{-1}\boldsymbol{S}_n > 0$ $a.s.$ then to show (5.12), it is equivalent to prove that $\mu(B) = 0$. We argue by contradiction: assume that $\mu(B) > 0$. Let $f = \liminf n^{-1}\boldsymbol{S}_n$. Note that for all $\omega \in \Omega$,

$$f(\omega) = \liminf \frac{1}{n}\boldsymbol{S}_{n+1}(\omega) \geq \liminf_{n\to\infty} \frac{\boldsymbol{S}_n \circ T(\omega) + \boldsymbol{S}_1(\omega)}{n} = f\left(T(\omega)\right), \quad a.s$$

hence for all $a \in \overline{\mathbb{R}}$, $\{\omega : \ f \circ T(\omega) > a\} \subset \{\omega : \ f(\omega) > a\}$ $a.s.$ i.e.

$$T^{-1}\left(\{f > a\}\right) \subset \{f > a\} \quad a.s.$$

Because $\mu\left(T^{-1}(\{f > a\})\right) = \mu(\{f > a\})$, we have $\mu\left(\{f > a\} \Delta T^{-1}(\{f > a\})\right) = 0$, and therefore

$$\text{for all } a \in \overline{\mathbb{R}}, \ \{f > a\} \in \mathcal{I}_\mu. \tag{5.13}$$

Let $N = \{f \leq 0\}$. Since $B \subset N$, then $\mu(N) > 0$. Let ν the probability measure in $(\Omega, \ \mathcal{B})$ given by the conditional probability given N. By Lemma 5.5, $(\Omega, \mathcal{B}, \nu, T)$ is a measure-preserving dynamical system and since ν is absolutely continuous with respect to μ then $\{\boldsymbol{S}_n\}_{n\geq 1}$ is a superadditive sequence on $(\Omega, \mathcal{B}, \nu, T)$. Noting that $\{f > 0\} \cap N = \{f > 0\} \cap \{f \leq 0\} = \varnothing$, we have

$$\nu(f > 0) = \mu(N)^{-1}\mu(\{f > 0\} \cap N) = 0. \tag{5.14}$$

Let us now show that under the condition $\mu(B) > 0$, one also has $\nu(f > 0) > 0$, which contradicts (5.14). Since $\{\liminf \boldsymbol{S}_n > 0\} \cap N = B$, one has

$$\nu(\liminf \boldsymbol{S}_n > 0) = \mu(N)^{-1}\mu(B) > 0$$

and thus

$$\text{there exists } \eta > 0 \text{ such that } \nu(\liminf \boldsymbol{S}_n > \eta) > 0. \tag{5.15}$$

Since

$$\nu(\liminf \boldsymbol{S}_n > \eta) = \nu(\cup_n\{\inf_{k\geq n} \boldsymbol{S}_k > \eta\}) = \lim_{n\to\infty} \nu(\inf_{k\geq n} \boldsymbol{S}_k > \eta),$$

it follows by (5.15) that

$$\text{there exists } p > 0 \text{ such that } \nu(\inf_{k\geq p} \boldsymbol{S}_k > \eta) > 0. \tag{5.16}$$

For this p, let

$$W = \{\inf_{k\geq p} \boldsymbol{S}_k > \eta\}.$$

By Birkhoff's ergodic theorem we have

$$h_n := n^{-1}\sum_{k=1}^{n} \mathbb{1}_W \circ T^k \longrightarrow h := \mathbb{E}^\nu(\mathbb{1}_W|\mathcal{I}_\nu) \ \ \nu - a.s \ \ as \quad n \to \infty.$$

Since $\mathbb{E}^\nu(h) = \nu(W) > 0$, then $\nu(\{\liminf h_n > 0\}) = \nu(\{h > 0\}) > 0$. Let

$$U = \{\liminf h_n > 0\} \cap \{\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s} \quad \text{for all} \ \ n, s \in \mathbb{N}^*\}.$$

By arguments already given, we have

$$\nu(U) = \nu(\{\liminf h_n > 0\}) > 0. \tag{5.17}$$

Let $\boldsymbol{u} = \{\boldsymbol{u}_n\}_{n\geq 0} := \{\mathbb{1}_W \circ T^n\}_{n\geq 0}$, on U, i.e. for all $\omega \in U$, $\boldsymbol{u}^\omega = \{\mathbb{1}_W \circ T^n(\omega)\}_{n\geq 0}$. Define $\{\boldsymbol{v}_k^n(\boldsymbol{u})\}_{k=0}^\infty$, $\{q^n(\boldsymbol{u})\}_{n\geq 0}$, $s(\boldsymbol{u})$ and $\{\boldsymbol{S}_n\}_{n\geq 1}$ as in (5.7) with p defined in (5.16). Note that $s(\boldsymbol{u}) < \infty$ and for all $n > s(\boldsymbol{u})$, $q^n(\boldsymbol{u}) \geq 1$. Remark also that $n^{-1}\boldsymbol{s}_n = h_n$ for all n and thus on U

$$\liminf n^{-1}\boldsymbol{s}_n = \liminf h_n > 0 \tag{5.18}$$

Since $\boldsymbol{v}_0^n(\boldsymbol{u}) = n$, then on U

$$\text{for all } n \geq s(\boldsymbol{u}), \quad \boldsymbol{S}_n = \boldsymbol{S}_{\boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})} + \sum_{k=0}^{q^n(\boldsymbol{u})-1} \left(\boldsymbol{S}_{\boldsymbol{v}_k^n(\boldsymbol{u})} - \boldsymbol{S}_{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})}\right). \tag{5.19}$$

By the first point of Lemma 5.2 one has $\boldsymbol{S}_{\boldsymbol{v}^n_{q^n(\boldsymbol{u})}(\boldsymbol{u})} \geq \inf_{i\leq s(\boldsymbol{u})}(\boldsymbol{S}_i)$ and by the definition of U, for all $n \geq s(\boldsymbol{u})$ and $k < q^n(\boldsymbol{u})$,

$$\boldsymbol{S}_{\boldsymbol{v}_k^n(\boldsymbol{u})} - \boldsymbol{S}_{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \geq \boldsymbol{S}_{\boldsymbol{v}_k^n(\boldsymbol{u})-\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \circ T^{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \quad \text{on} \quad U.$$

It follows by (5.19) that on U,

$$\text{for all } n \geq s(\boldsymbol{u}), \quad \boldsymbol{S}_n \geq \inf_{i\leq s(\boldsymbol{u})}(\boldsymbol{S}_i) + \sum_{k=0}^{q^n(\boldsymbol{u})-1} \boldsymbol{S}_{\boldsymbol{v}_k^n(\boldsymbol{u})-\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \circ T^{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})}. \tag{5.20}$$

Since on U for all $k < q^n(\boldsymbol{u})$, $\boldsymbol{u}_{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} = 1$, i.e. $T^{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \in W$ and by Eq. (5.8) $\boldsymbol{v}_k^n(\boldsymbol{u}) - \boldsymbol{v}_{k+1}^n(\boldsymbol{u}) \geq p$, then the definition of W implies that

$$\text{for all } k < q^n(\boldsymbol{u}), \ \boldsymbol{S}_{\boldsymbol{v}_k^n(\boldsymbol{u})-\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} \circ T^{\boldsymbol{v}_{k+1}^n(\boldsymbol{u})} > \eta \quad \text{on } U.$$

Thus, by (5.20), one has on U

$$\text{for all } n \geq s(\boldsymbol{u}), \quad \boldsymbol{S}_n \geq \inf_{i\leq s(\boldsymbol{u})}(\boldsymbol{S}_i) + \eta q^n(\boldsymbol{u}).$$

It follows by (5.18) and the second point of Lemma 5.2, that

$$f = \liminf n^{-1}\boldsymbol{S}_n \geq \eta \liminf n^{-1} q^n(\boldsymbol{u}) > 0 \ \text{ on } U.$$

Thus $\nu(\{f > 0\}) \geq \nu(U) > 0$, where the last inequality is due to (5.17). This contradicts (5.14) and concludes the proof. □

Remark 5.1. *Following the result of Theorem 5.1, we can wonder if*

$$\limsup n^{-1}\boldsymbol{S}_n < 0 \ \ a.s \quad \textit{on} \ \liminf\{\boldsymbol{S}_n < 0\}.$$

However, this statement is incorrect. A simple counter-example is the superadditive process that is identically equal to -1. *A counter-example of non-a.s. constant process can be constructed.* [1].

1. Let $\{\boldsymbol{X}_n\}_{n=1}^\infty$ a positive strictly stationary and ergodic process with a positive finite moment and let

Theorem 5.1 could be stated in a weaker form if the set where $(\boldsymbol{S}_n)_{n\geq 1}$ is not always non-positive and becomes non-negative for sufficiently large values of n is invariant. This variant result is the following:

Theorem 5.2. *Let $\{\boldsymbol{S}_n\}_{n\geq 1}$ be a superadditive sequence, let $\boldsymbol{\tau} = \sup_{n\in\mathbb{N}}\{n : \ \boldsymbol{S}_n < 0\}$ and let E be an invariant subset of $\{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0, \ \boldsymbol{\tau} < \infty\}$. One has*

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \quad a.s \quad on \ E. \tag{5.21}$$

Proof of Theorem 5.2. Note that (5.21) is equivalent to $\mu(\{\liminf n^{-1}\boldsymbol{S}_n > 0\} \cap E) = \mu(E)$. Letting $\upsilon = \inf_{n\in\mathbb{N}}\{n : \ \boldsymbol{S}_n > 0\}$, one has $\{\upsilon < \infty\} = \{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0\}$ and then $E \subset \{\upsilon < \infty, \boldsymbol{\tau} < \infty\}$. Let $P = \{\boldsymbol{S}_n + \boldsymbol{S}_s \circ T^n \leq \boldsymbol{S}_{n+s} \ \text{ for all } \ n, s \in \mathbb{N}^*\}$ and let

$$E' = E \cap P \text{ and } C(E') = \bigcap_{n\geq 0} T^{-n}(E').$$

Since $\mu(P) = 1$ then P is invariant and $\mu(E') = \mu(E)$. Hence E' is also invariant and it follows by Lemma 5.6, that

$$\mu(C(E')) = \mu(E). \tag{5.22}$$

Since $C(E') \subset E \subset \{\upsilon < \infty, \boldsymbol{\tau} < \infty\}$, it follows by Lemma 5.6 that for all $\omega \in C(E')$ one has,

$$\boldsymbol{S}_\upsilon(\omega) > 0, \ \boldsymbol{S}_n(\omega) \geq 0 \ \text{ for all } n > \boldsymbol{\tau} \ \text{ and } \ T^n(\omega) \in C(E') \ \text{ for all } n \geq 0$$

and then on $C(E')$,

$$\text{for all } n > \boldsymbol{\tau} \circ T^\upsilon + \upsilon, \quad \boldsymbol{S}_n(\omega) \geq \boldsymbol{S}_{n-\upsilon} \circ T^\upsilon + \boldsymbol{S}_\upsilon \geq \boldsymbol{S}_\upsilon > 0. \tag{5.23}$$

The second last inequality comes from the fact that $\boldsymbol{S}_{n-\upsilon} \circ T^\upsilon \geq 0$ because $n - \upsilon > \boldsymbol{\tau} \circ T^\upsilon$ and $T^\upsilon \in C(E')$ on $C(E') \subset \{\boldsymbol{\tau} < \infty\}$. It follows by (5.23) that $C(E') \subset \liminf\{\boldsymbol{S}_n > 0\} \cap E$ and thus, by Theorem 5.1, one has

$$\mu(C(E')) \leq \mu(\liminf\{\boldsymbol{S}_n > 0\} \cap E) = \mu(\{\liminf n^{-1}\boldsymbol{S}_n > 0\} \cap E). \tag{5.24}$$

Hence, in views of (5.22) and (5.24) we have $\mu(\{\liminf n^{-1}\boldsymbol{S}_n > 0\} \cap E) \geq \mu(E)$. Since $\{\liminf n^{-1}\boldsymbol{S}_n > 0\} \cap E \subset E$, it follows that

$$\mu(\{\liminf n^{-1}\boldsymbol{S}_n > 0\} \cap E) = \mu(E).$$

This concludes the proof. □

Remark 5.2. *It is clear that the condition $\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0$ in Theorem 5.2 is necessary and cannot be replaced by a weaker condition. Moreover, the invariance assumption can-*

$\alpha \in (0, 1)$. Using the inequality $(a+b)^\alpha \leq a^\alpha + b^\alpha$ for all $a, b \geq 0$, we have $\{\boldsymbol{S}_n := -(\sum_{k=1}^n \boldsymbol{X}_k)^\alpha\}_{n=0}^\infty$ is superadditive. However, since the ergodic theorem implies that $n^{-1}\sum_{k=1}^n \boldsymbol{X}_k \to \mathbb{E}\boldsymbol{X}_1 \in (0, \infty)$ *a.s.* as $n \to \infty$ then $\boldsymbol{S}_n \to -\infty$ *a.s.* as $n \to \infty$ and $\limsup n^{-1}\boldsymbol{S}_n = \limsup n^{-(1-\alpha)}(n^{-1}\sum_{k=1}^n \boldsymbol{X}_k)^\alpha = 0$ *a.s.*

not be weakened without adding a supplementary condition. To illustrate this, we consider a process from Kesten (1975). Let $\Omega = \mathbb{R}^{\mathbb{N}}$, T *the left shift operator and* $\mathbb{P}[\{(-1)^n\}_n] = \mathbb{P}[\{(-1)^{n+1}\}_n] = 1/2$. *The sequence* $\boldsymbol{S}_n(\{x_i\}_n) := \sum_{i=0}^{n-1} x_i$ *is an additive process. However,* $(\boldsymbol{S}_n)_n$ *is almost surely bounded on* $\{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0\} \cap \liminf\{\boldsymbol{S}_n \geq 0\} = \{[(-1)^n]_n\}$. *It can be observed that* $\{[(-1)^n]_n\}$ *is not invariant.*

Remark 5.3. *An interesting consequence of Theorem 5.2 is that*

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \ a.s \quad \text{if} \quad \sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0 \ \ \mu\text{-a.s and } \mu(\liminf\{\boldsymbol{S}_n \geq 0\}) = 1, \tag{5.25}$$

(i.e if $\mu(\{\sup_{n\in\mathbb{N}} \boldsymbol{S}_n > 0\} \cap \liminf\{\boldsymbol{S}_n \geq 0\}) = 1$*). Indeed, since for all* $A \in \mathcal{B}$, $\mu(A) = 1$ *implies that* $A \in \mathcal{I}_\mu$, *then (5.25) follows from Theorem 5.2.*

We say that T is ergodic if for all $I \in \mathcal{I}_\mu$, $\mu(I) \in \{0, 1\}$.

Corollary 5.1. *Let* $\{\boldsymbol{S}_n\}_{n\geq 1}$ *a superadditive sequence. If* T *is ergodic then* $\liminf n^{-1}\boldsymbol{S}_n$ *is almost surely constant in* $\overline{\mathbb{R}}$ *and*

$$\liminf n^{-1}\boldsymbol{S}_n > 0 \ \ a.s \quad \text{if and only if} \quad \mu(\liminf\{\boldsymbol{S}_n > 0\}) > 0. \tag{5.26}$$

Proof. For the first point, using (5.13) and the ergodicity of T, we can deduce that for all $a \in \overline{\mathbb{R}}$, the function $F(a) := \mu(\liminf n^{-1}\boldsymbol{S}_n \leq a)$ takes its values on $\{0, 1\}$. Since $F(\infty) = 1$, and $a \mapsto F(a)$ is right continuous and non-decreasing on $\overline{\mathbb{R}}$, we can conclude that there exists an $a_0 \in \overline{\mathbb{R}}$ such that for all $a < a_0$, $F(a) = 0$ and $F(a_0) = 1$. Therefore, we can conclude that $\mu(\liminf n^{-1}\boldsymbol{S}_n = a_0) = 1$, which implies that $\liminf n^{-1}\boldsymbol{S}_n = a_0$ almost surely.

The necessary condition in (5.26) is trivial. To show the sufficient condition, using Theorem 5.1, we deduce that $\mu(\liminf n^{-1}\boldsymbol{S}_n > 0) = \mu(\liminf \boldsymbol{S}_n > 0) > 0$. Therefore, in view of Equation (5.13) and the ergodicity of T, it follows that $\mu(\liminf n^{-1}\boldsymbol{S}_n > 0) = 1$. □

In the next theorem, we state the last main result of this section. Let $A \in \mathcal{B}$ with $\mu(A) > 0$. Let $\tau = \inf\{n \geq 1 : \ T^n \in A\}$. Define L by $L = T^\tau$ if τ is finite, and $L = id_\Omega$ otherwise. Let ν be the probability measure given by the conditional probability given A. By the Poincaré recurrence theorem we know that the set of points ω of A for which $T^n(\omega) \notin A$ for all $n \geq 1$ has zero measure. Therefore, τ is almost surely finite under ν and then $L = T^\tau$ ν-a.s. We can also define the ν-a.s. finite sequence of integers $(\tau_n)_{n\geq 1}$: $\tau_n = \tau \circ L^{n-1}$. For all $n \geq 0$, let $\upsilon_n = \sum_{k=1}^{n} \tau_k$. It is easy to see that υ_n is the index k where $T^k \in A$ for the n-th time. Therefore (υ_n) is ν-a.s. strictly increasing and grows to infinity. We have the following:

Theorem 5.3. *Let* $\{\boldsymbol{S}_n\}_{n\geq 1}$ *a superadditive sequence with* $\boldsymbol{S}_1^-$ *integrable, one has*

$$\lim n^{-1}\boldsymbol{S}_n > 0 \ \ \nu\text{-a.s} \ \ \text{on} \ \liminf_n \{\boldsymbol{S}_{\upsilon_n} > 0\}.$$

Noting that $(\upsilon_n) = (n : T^n \in A)$, Theorem 5.3 states that if $\boldsymbol{S}_1^-$ is integrable, then $\lim n^{-1}\boldsymbol{S}_n > 0$ μ-$a.s$ on the intersection of set A and the set where the sequence $(\boldsymbol{S}_n : T^n \in A)$ is positive from a certain period. This means that under the set A, the positivity condition only involves the values of $(\boldsymbol{S}_n)$ with indices in $(n : T^n \in A)$. Note also that, under the integrability of $\boldsymbol{S}_1^-$, this result is more general than Theorem 5.1, which is obtained by taking $A = \Omega$.

In Remark 5.4 below, we show that the integrability condition in Theorem 5.3 is not superfluous.

We immediately deduce the following result that extends a variant of Kesten's result, established by Eskin and Mirzakhani (2018, Lemma C.8), for additive sequences to superadditive processes.

Corollary 5.2. *Suppose that T is ergodic and let $\{\boldsymbol{S}_n\}_{n\geq 1}$ a superadditive sequence with $\boldsymbol{S}_1^-$ integrable. Let $A \in \mathcal{B}$ with $\mu(A) > 0$. If, almost surely, the sequence $(\boldsymbol{S}_n : T^n \in A)$ is positive from a certain period, then*

$$\lim n^{-1}\boldsymbol{S}_n = \lim n^{-1}\mathbb{E}\boldsymbol{S}_n = \sup_n n^{-1}\mathbb{E}\boldsymbol{S}_n > 0 \quad a.s. \tag{5.27}$$

Proof. The first two equalities in (5.27) follows from Kingman's subadditive ergodic theorem. To prove that $\lim n^{-1}\boldsymbol{S}_n > 0$ $a.s.$, observe that if $\mu(\{(\boldsymbol{S}_n : T^n \in A) \text{ is positive from a certain period}\}) = 1$, then by Theorem 5.3, $\mu(\lim n^{-1}\boldsymbol{S}_n > 0,\ A) = \mu(A) > 0$. This means that $\mu(\lim n^{-1}\boldsymbol{S}_n > 0) > 0$, which implies the result by already given arguments. □

The proof of Theorem 5.3 is based on Theorem 5.1 and the following additional result.

Lemma 5.3. *We claim that: i) $(\Omega, \mathcal{B}, \nu, L)$ is a measure-preserving dynamical system and, ii)*

$$\mathbb{E}^\nu \tau = \mu(\cup_{k\geq 1}\{T^k \in A\})/\mu(A) < \infty. \tag{5.28}$$

Moreover, iii) $(\boldsymbol{S}_{\upsilon_n})_n$ is a superadditive sequence on $(\Omega, \mathcal{B}, \nu, L)$.

Proof. We prove *i). Recall that τ is almost surely finite under ν and $L = T^\tau$ ν-a.s. Thus, we must show that for all $B \in \mathcal{B}$, $\nu(T^\tau \in B)) = \nu(B)$. We have*

$$\begin{aligned}
\nu(T^\tau \in B)) &= \sum_{k=1}^{\infty} \nu(\tau = k, T^k \in B) \\
&= \mu(A)^{-1} \sum_{k=1}^{\infty} \mu(A, T^1 \notin A, \cdots, T^{k-1} \notin A, T^k \in A, T^k \in B) \\
&= \mu(A)^{-1} \sum_{k=1}^{\infty} \mu(\boldsymbol{X}_0 = 1, \boldsymbol{X}_1 = 0, \cdots, \boldsymbol{X}_{k-1} = 0, \boldsymbol{X}_k = 1, \boldsymbol{Y}_k = 1),
\end{aligned}$$

where $(\boldsymbol{X}_n, \boldsymbol{Y}_n) = (\mathbb{1}_A \circ T^n, \mathbb{1}_B \circ T^n)$ for all $n \geq 0$. It is clear that $(\boldsymbol{X}_n, \boldsymbol{Y}_n)_{n\geq 0}$ is a stationary sequence on $(\Omega, \mathcal{B}, \mu)$. Therefore, it is well-known that we can extend that

sequence into the past to obtain a full stationary process $(\boldsymbol{X}_n, \boldsymbol{Y}_n)_{n\in\mathbb{Z}}$, *see for instance Elton (1990, Lemma 1). Hence*

$$\begin{aligned}\nu(T^\tau \in B)) &= \mu(A)^{-1}\sum_{k=1}^{\infty}\mu(\boldsymbol{X}_{-k}=1,\boldsymbol{X}_{-k+1}=0,\cdots,\boldsymbol{X}_{-1}=0,\boldsymbol{X}_0=1,\boldsymbol{Y}_0=1)\\ &= \mu(A)^{-1}\sum_{k=1}^{\infty}\mu(\boldsymbol{X}_0=1,\boldsymbol{Y}_0=1,\cup_{k\geq 1}\{\boldsymbol{X}_{-k}=1\})\\ &= \mu(A)^{-1}\mu(\boldsymbol{X}_0=1,\boldsymbol{Y}_0=1\})\\ &= \nu(B).\end{aligned}$$

The second equality is derived from the fact that the sets $(\{\boldsymbol{X}_{-k}=1,\boldsymbol{X}_{-k+1}=0,\cdots,\boldsymbol{X}_{-1}=0\})_{k\geq 1}$ *are disjoint and their union constitutes* $\cup_{k\geq 1}\{\boldsymbol{X}_{-k}=1\}$. *The third equality follows from the Poincaré recurrence theorem, which implies that* $\{\boldsymbol{X}_0=1\}\subset\cup_{k\geq 1}\{\boldsymbol{X}_{-k}=1\}$. *The conclusion follows.*

The proof of ii) (Eq. (5.28)) uses similarly arguments. We have

$$\begin{aligned}\mathbb{E}^\nu\tau=\sum_{k=1}^{\infty}\nu(\tau\geq k) &= \mu(A)^{-1}\sum_{k=1}^{\infty}\mu(\boldsymbol{X}_{-k}=1,\boldsymbol{X}_{-k+1}=0,\cdots,\boldsymbol{X}_{-1}=0)\\ &= \mu(A)^{-1}\mu(\cup_{k\geq 1}\{\boldsymbol{X}_{-k}=1\})\\ &= \mu(\cup_{k\geq 1}\{T^k\in A\})/\mu(A),\end{aligned}$$

because $\mu(\cup_{k\geq 1}\{\boldsymbol{X}_{-k}=1\})=\lim_n\mu(\cup_{k=1}^n\{\boldsymbol{X}_{-k}=1\})=\lim_n\mu(\cup_{k=1}^n\{\boldsymbol{X}_k=1\})=\mu(\cup_{k\geq 1}\{\boldsymbol{X}_k=1\})=\mu(\cup_{k\geq 1}\{T^k\in A\})$.

To show iii), first note that,

$$\text{for all } n,s\in\mathbb{N}^*,\ L^n=T^{\upsilon_n}\ \text{ and }\ \upsilon_n+\upsilon_s\circ T^{\upsilon_n}=\upsilon_{n+s}\ \ \nu\text{-a.s.}\tag{5.29}$$

Thus, by superadditivity and the fact that ν *is absolutely continuous with respect to* μ, *for all* $n,s\in\mathbb{N}^*$,

$$\begin{aligned}\boldsymbol{S}_{\upsilon_n}+\boldsymbol{S}_{\upsilon_s\circ L^n}\circ L^n &= \boldsymbol{S}_{\upsilon_n}+\boldsymbol{S}_{\upsilon_s\circ T^{\upsilon_n}}\circ T^{\upsilon_n}\\ &\leq \boldsymbol{S}_{\upsilon_n+\upsilon_s\circ T^{\upsilon_n}}\\ &= \boldsymbol{S}_{\upsilon_{n+s}}\ \ \nu\text{-a.s.}\end{aligned}$$

This concludes the proof. □

Proof of Theorem 5.3. By superadditivity, for all n, one has ν-almost surely $\boldsymbol{S}_n-\boldsymbol{S}_{\upsilon_k}\geq$

$\boldsymbol{S}_{n-\upsilon_k} \circ T^{\upsilon_k}$ where $\upsilon_k \leq n < \upsilon_{k+1}$. Thus

$$\begin{aligned}
\boldsymbol{S}_n &= (\boldsymbol{S}_n - \boldsymbol{S}_{\upsilon_k}) + \boldsymbol{S}_{\upsilon_k} \\
&\geq \boldsymbol{S}_{n-\upsilon_k} \circ T^{\upsilon_k} + \boldsymbol{S}_{\upsilon_k} \\
&\geq \min_{0 \leq i < \upsilon_{k+1} - \upsilon_k} \boldsymbol{S}_i \circ T^{\upsilon_k} + \boldsymbol{S}_{\upsilon_k} \\
&\geq \min_{0 \leq i < \tau \circ L^k} \boldsymbol{S}_i \circ L^k + \boldsymbol{S}_{\upsilon_k}, \quad \text{by Eq. (5.29).} \\
&\geq \min_{0 \leq i < \tau \circ L^k} -\boldsymbol{S}_i^- \circ L^k + \boldsymbol{S}_{\upsilon_k} \ \nu\text{-}a.s.
\end{aligned}$$

Since, by superadditivity $\boldsymbol{S}_i \geq \sum_{j=0}^{i-1} \boldsymbol{S}_1 \circ T^j$ ν-a.s. for all $0 \leq i < \tau \circ L^k$, we have $-\boldsymbol{S}_i^- \circ L^k \geq (\sum_{j=0}^{i-1} -\boldsymbol{S}_1^- \circ T^j) \circ L^k \geq -(\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^k$ ν-a.s. Therefore,

$$\begin{aligned}
n^{-1}\boldsymbol{S}_n &\geq -n^{-1}(\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^k + n^{-1}\boldsymbol{S}_{\upsilon_k} \\
&\geq -k^{-1}(\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^k + \upsilon_{k+1}^{-1}\boldsymbol{S}_{\upsilon_k} \ \nu\text{-}a.s.
\end{aligned}$$

The last inequality is due to the fact that υ_k is ν-a.s. strictly increasing, which implies that $k \leq \upsilon_k \leq n$. Under the integrablity of $\boldsymbol{S}_1^-$, the Kingman ergodic theorem implies that $\lim n^{-1}\boldsymbol{S}_n$ exists μ-a.s. Thus the limit also exists ν-a.s. Since k grows to infinity with n, it follows that

$$\liminf_n n^{-1}\boldsymbol{S}_n \geq -\limsup_k k^{-1}(\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^k + \liminf_k \upsilon_{k+1}^{-1}\boldsymbol{S}_{\upsilon_k} \ \nu\text{-}a.s.$$

To conclude, it suffices to show that on $\liminf_k \{\boldsymbol{S}_{\upsilon_k} > 0\}$, one has ν-$a.s$:

$$i) \ \liminf_k \upsilon_{k+1}^{-1}\boldsymbol{S}_{\upsilon_k} > 0, \text{ and } ii) \ \limsup_k k^{-1}(\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^k = 0.$$

Let us show i). Since $\upsilon_n = \sum_{k=0}^{n-1} \tau \circ L^k$ ν-a.s., it follows by Lemma 5.3 and the Birkhoff ergodique theorem that

$$\lim_k k^{-1}\upsilon_{k+1} \text{ exists and is finite} \quad \nu\text{-}a.s. \tag{5.30}$$

We also have, by Lemma 5.3 and Theorem 5.1, that

$$\liminf_k k^{-1}\boldsymbol{S}_{\upsilon_k} > 0 \ \ \nu\text{-}a.s \quad \text{on} \ \liminf_k \{\boldsymbol{S}_{\upsilon_k} > 0\}.$$

Therefore, the result follows:

$$\liminf_k \upsilon_{k+1}^{-1} \boldsymbol{S}_{\upsilon_k} = \liminf_k (k^{-1}\upsilon_{k+1})^{-1}(k^{-1}\boldsymbol{S}_{\upsilon_k}) > 0 \quad \nu\text{-}a.s \quad \text{on} \quad \liminf_k \{\boldsymbol{S}_{\upsilon_k} > 0\}.$$

Now let us turn to the proof of *ii*). Since $\boldsymbol{S}_1^-$ is integrable, we have by the Birkhoff ergodique theorem and the « absolutely continuous » argument that

$$\lim_n n^{-1} \sum_{i=1}^{n} \boldsymbol{S}_1^- \circ T^i \text{ exists and is finite } \nu\text{-}a.s. \tag{5.31}$$

Since $\upsilon_k \to \infty$ ν-a.s., it follows that $\lim_k \upsilon_k^{-1} \sum_{i=0}^{\upsilon_k - 1} \boldsymbol{S}_1^- \circ T^i$ exists and is finite ν-a.s. Letting $f = \sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j$, it is no difficult to see that

$$\sum_{i=0}^{\upsilon_k - 1} \boldsymbol{S}_1^- \circ T^i = \sum_{i=0}^{k-1} (\sum_{j=0}^{\tau-1} \boldsymbol{S}_1^- \circ T^j) \circ L^i = \sum_{i=0}^{k-1} f \circ L^i \ \nu\text{-}a.s.$$

It follows by (5.30) and (5.31) that

$$\lim_k k^{-1} \sum_{i=0}^{k-1} f \circ L^i = \lim_k (k^{-1}\upsilon_k)(\upsilon_k^{-1} \sum_{i=0}^{\upsilon_k - 1} \boldsymbol{S}_1^- \circ T^i) \text{ exists and is finite } \nu\text{-}a.s.$$

This implies that $k^{-1} f \circ L^k$ converges to 0 ν-a.s., which concludes the proof. □

The following result, which we state without proof, follows directly from Corollary 5.1 for Point 1. and Corollary 5.2 for Point 2. through the application of the function $-\log$.

Corollary 5.3. *Suppose that T is ergodic and let $\{\boldsymbol{\gamma}_n\}_{n\geq 1} \in \mathbb{R}^{\mathbb{N}}$ be a positive sub-multiplicative process (i.e for all $n, s \in \mathbb{N}^*$ $0 \leq \boldsymbol{\gamma}_{n+s} \leq \boldsymbol{\gamma}_n \times \boldsymbol{\gamma}_s \circ T^n$ a.s.). Let $A \in \mathcal{B}$ with $\mu(A) > 0$.*

1. *If $\mu(\liminf_n \{\boldsymbol{\gamma}_n < 1\}) > 0$, then $\boldsymbol{\gamma}$ is almost surely constant in $\overline{\mathbb{R}}$ and*

$$\gamma := \limsup_n n^{-1} \log \boldsymbol{\gamma}_n < 0 \quad a.s.$$

2. *If $\mathbb{E} \log^+ \boldsymbol{\gamma}_1$ is finite and, almost surely, the sequence $(\boldsymbol{\gamma}_n : T^n \in A)$ is strictly less than 1 from a certain period, then*

$$\gamma = \lim_n n^{-1} \log \boldsymbol{\gamma}_n = \lim_n n^{-1} \mathbb{E} \log \boldsymbol{\gamma}_n = \inf_n n^{-1} \mathbb{E} \log \boldsymbol{\gamma}_n < 0 \quad a.s.$$

Recall that γ is almost surely constant. The previous corollary is more general than Lemma 3.4 of Bougerol and Picard (1992b), see the next corollary. Point 1. of Corollary 5.3 does not require any integrability condition, it applies to all sub-multiplicative ergodic sequences and only needs its values to be negative for n large enough. This result also enables the characterization of the case where the top-lyapunov exponent of a class of cocycles on a measure preserving transformation is negative.

Remark 5.4. *If* $\mathbb{E}\log^+ \boldsymbol{\gamma}_1$ *is not finite, then the conclusion of Point 2. of Corollary 5.3 is no longer valid. Indeed, for all* n*, let* $\boldsymbol{\alpha}_n = e^{-1}\frac{\boldsymbol{u}_n}{\boldsymbol{u}_{n-1}}$ *where* $(\boldsymbol{u}_n)_{n\in\mathbf{z}}$ *is a positive iid sequence such that* $\mathbb{E}\ln^+ \boldsymbol{u}_0 = \infty$*. Consider the measure-preserving dynamical system given by the quadruplet:* $R^{\mathbb{Z}}$ *and its Borel* σ*-algebra, the push-forward probability measure* $\mathbb{P}_{\boldsymbol{u}}$ *of* $(\boldsymbol{u}_n)$ *and the shift operator* T*. It is easy to see that the process* $(\boldsymbol{\gamma}_n)_{n\geq 1}$*, where* $\boldsymbol{\gamma}_n = \prod_{k=1}^n \boldsymbol{\alpha}_n$*, is a sub-multiplicative process. We show, in Appendix 5.1, that there exists a measurable set* A *with* $\mathbb{P}_{\boldsymbol{u}}(A) > 0$ *such that the sequence* $(\boldsymbol{\gamma}_n : T^n \in A)$ *converges almost surely to* 0 *and on the other hand that* $\limsup_n n^{-1}\log \boldsymbol{\gamma}_n \geq 0$*. This is because* $\mathbb{E}\ln^+ \boldsymbol{\gamma}_1 = \infty$*.*

Corollary 5.4. *(Bougerol and Picard, 1992b, Lemma 3.4) Let* $\{M_n\}_{n\geq 1}$ *be an ergodic strictly stationary sequence in the space of the* $d \times d$ *real matrices. We suppose that* $\mathbb{E}\left(\log^+ \|M_0\|\right)$ *is finite and that, almost surely,* $\lim_{n\to+\infty} \|M_n M_{n-1}\cdots M_1\| = 0$*. Then*

$$\gamma := \inf_{n\in\mathbb{N}} \frac{1}{n}\mathbb{E}\left(\log \|M_n M_{n-1}\cdots M_1\|\right) < 0.$$

Proof. Let $\boldsymbol{\gamma}_n = \log \|M_n M_{n-1}\cdots M_1\|$ for all $n \in \mathbb{N}$. Note that $\mathbb{P}(\liminf_n\{\boldsymbol{\gamma}_n < 1\}) = 1$ if $\lim_{n\to\infty} \boldsymbol{\gamma}_n = 0$ *a.s.* Since $\{\boldsymbol{\gamma}_n\}_{n\geq 1}$ is a sub-multiplicative sequence, the result follows from Corollary 5.3 and the Kingman's subadditive ergodic theorem (see also Furstenberg and Kesten (1960)). □

3 Stationarity of fGARCH models in $\mathcal{C}^0$

In this section, we study the existence of a stationary solution of the functional GARCH models in the space of continuous functions (see Aue et al. (2017)).

A sequence $(\boldsymbol{r}_t : t \in \mathbb{Z})$ of random elements where each random object $\boldsymbol{r}_t$ is a curve $(\boldsymbol{r}_t(u) : u \in [0,1])$ in $\mathcal{C}[0,1]$, the space of continuous functions on $[0,1]$, is called a functional GARCH process of orders $(1,1)$, abbreviated by fGARCH$(1,1)$, if it satisfies the equations

$$\begin{aligned} \boldsymbol{r}_t &= \boldsymbol{\sigma}_t \boldsymbol{\eta}_t, \\ \boldsymbol{\sigma}_t^2 &= \delta + \alpha \boldsymbol{r}_{t-1}^2 + \beta \boldsymbol{\sigma}_{t-1}^2 = \delta + \int \boldsymbol{\gamma}_t(\cdot, s)\boldsymbol{\sigma}_{t-1}^2(s)ds = \delta + \boldsymbol{\gamma}_t \boldsymbol{\sigma}_{t-1}, \end{aligned} \tag{5.32}$$

where $(\boldsymbol{\eta}_t : t \in \mathbb{Z})$ is a sequence of independent and identically distributed (iid) random functions in $\mathcal{C}[0,1]$, δ is a positive function and the integral operators α and β, i.e. $(\alpha x)(u) = \int \alpha(u,s)x(s)ds$ and $(\beta x)(u) = \int \beta(u,s)x(s)ds$ are positive, i.e. they map nonnegative functions to nonnegative function. $\boldsymbol{\gamma}_t(u,s) = \alpha(u,s)\boldsymbol{\eta}_{t-1}^2(s) + \beta(u,s)$ is an element of $\mathcal{C}[0,1]^2$.

By extending our considerations to include strictly stationary and ergodic but non-id innovations, and by replacing the interval $[0,1]$ with an arbitrary compact set K, we can generalize the autoregressive model with non-negative random functional coefficients with Eq. (5.33) below to include a wide range of conditional volatility models. Furthermore, by allowing the coefficients δ, α, and β to be stochastic processes rather than constants,

and by dropping the assumption that $\gamma = \alpha + \beta$ must be an integral operator, we can obtain even more flexibility in our modeling approach. The model is as follows:

$$\boldsymbol{h}_t = \delta(\boldsymbol{\eta}_{t-1}) + \gamma(\boldsymbol{\eta}_{t-1})\boldsymbol{h}_{t-1}, \tag{5.33}$$

where the positive stochastic curve $\boldsymbol{\delta}_t = \delta(\boldsymbol{\eta}_t)$ and linear operator $\boldsymbol{\gamma}_t = \gamma(\boldsymbol{\eta}_t)$ are measurable functions of $\boldsymbol{\eta}_t$.

We can see that Model (5.33) include the functional GARCH considered in (5.32). If $K = \{1\}$, we obtain the univariate class of GARCH(1, 1) model of He and Teräsvirta (1999) and if K is finite we get the class of multivariate-constant conditional correlation and univariate asymmetric power GARCH(p,q), see the AR(1) representation of Maïnassara et al. (2022).

Across the different normed vector spaces, we will unambiguously use the classical notation of the norm, $\|\cdot\|$. We recall that $F := \mathcal{C}(K)$ equipped with the uniform norm $\|x\| = \sup\{|x(u)|, \quad u \in K\}$ is a Banach space. The space of the linear endomorphisms in F is equipped with the usual operator norm $\|\alpha\| = \sup\{\|\alpha(x)\|, \quad \|x\| \leq 1\}$. Denoting $\boldsymbol{e} : K \ni u \longmapsto 1$, remark that for all positive operator α, $\|\alpha\| = \|\alpha\boldsymbol{e}\|$. For all $x \in F$, let $\inf x = \inf\{|x(u)|, \quad u \in K\}$.

The stationarity of Model (5.32) has been studied in (Aue et al., 2017, Theorem 2.2) and in (Hörmann et al., 2013, Theorem 2.3) when the model is reduced to a pure functional ARCH. In both papers, they give a sufficient condition for the existence of a stationary solution. The weakest condition is obtained by Aue et al. (2017). They show that if

$$-\infty \leq \mathbb{E}\log\|\boldsymbol{\gamma}_0\| < 0, \tag{5.34}$$

then Model (5.32) have a unique, strictly stationary and nonanticipative solution in $\mathcal{C}[0, 1]$.

Contrary to the multivariate setup, to our knowledge, necessary and sufficient conditions for the existence of a stationary solution of Model (5.32) have never been established. As noted by Cerovecki et al. (2019, Remark 1), one of the main challenges in establishing these conditions is to extend the contraction property of random matrices to linear operators. Since we have established this result in Corollary 5.3, we are ready to provide necessary and sufficient conditions for the existence of a stationary solution for the general functional GARCH models considered in (5.33). To establish theses conditions, the following assumptions will be made.

A5.1 $(\boldsymbol{\eta}_t)$ is iid.

A5.2 $(\boldsymbol{\eta}_t)$ is strictly stationary and ergodic and $\mathbb{E}(\log^+\|\boldsymbol{\gamma}_0\|)$ is finite.

For all $t \geq 0$, let

$$\boldsymbol{\gamma}_t^{(0)} = id_F \text{ and } \boldsymbol{\gamma}_t^{(n)} = \boldsymbol{\gamma}_t \circ \cdots \circ \boldsymbol{\gamma}_{t-n+1} \text{ for all } n \geq 1.$$

Consider the following assumption.

$$\mathbb{P}(\inf\{\sum_{k=0}^{+\infty} \boldsymbol{\gamma}_0^{(k)}\boldsymbol{\delta}_{-k}, \quad u \in K\} = 0) < 1. \tag{5.35}$$

Note that if $\mathbb{P}(\inf\{\boldsymbol{\delta}_0(u),\ u \in K\} = 0) < 1$ then we have (5.35). Since we deal with volatility curves, it is not restrictive to assume that (5.35) holds. This condition is satisfied by most commonly used volatility models and ensures that the solutions are positive on the entire curve in a non-negligible set. Indeed, by iterating (5.33), we can see that any non-negative solution $(\boldsymbol{h}_t)$ of (5.33) satisfies: $\boldsymbol{h}_t \geq \sum_{k=0}^{\infty} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{t-k}$. Thus, (5.35) implies that $\mathbb{P}(\inf\{\boldsymbol{h}_0(u),\ \ u \in K\} = 0) < 1$.

For all $t \in \mathrm{Z}$, define $\boldsymbol{w}_t := \sum_{k=0}^{+\infty} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{t-k} \in [0,\infty]^K$. Note that the sequence of continuous, positive, and non-decreasing functions $(\sum_{k=0}^{n} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{t-k})_n$ converges pointwise to $\boldsymbol{w}_t$ a.s., even though the limit may not be finite at some points. It is also important to note that the convergence may not be uniform. Therefore, $\boldsymbol{w}_t$ is not necessarily continuous.

Now we state the main result of this section.

Theorem 5.4. *Let* $\boldsymbol{\gamma} = \limsup_n \frac{1}{n} \log(\|\boldsymbol{\gamma}_0^{(n)}\|)$.

1. *Suppose that* (5.35) *hold. If* ***A5.1*** *or* ***A5.2*** *hold and Equation* (5.33) *has a positive stationary solution in* F *then*
$$\boldsymbol{\gamma} < 0 \quad a.s.$$
2. *Conversely, if* $\mathbb{E}(\log^+ \|\boldsymbol{\delta}_0\|) < \infty$ *and* $\boldsymbol{\gamma} < 0$ *then* $(\sum_{k=0}^{n} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{t-k})_n$ *converges in* F *to* $\boldsymbol{w}_t$ *and* $(\boldsymbol{w}_t)$ *is the unique (continuous, positives and non-anticipative) stationary solution of* (5.33).

Remark 5.5.

1. *In views of Corollary* 5.3*, under* ***A5.1****,* $\boldsymbol{\gamma}$ *is almost surely constant with value in* $[-\infty,\infty[$. *Under* ***A5.2****, the subadditive ergodic theorem implies that*
$$\begin{aligned} \boldsymbol{\gamma} &= \lim_{n\to+\infty} \frac{1}{n} \log(\|\boldsymbol{\gamma}_0^{(n)}\|) \\ &= \lim_{n\to+\infty} \frac{1}{n} \mathbb{E} \log(\|\boldsymbol{\gamma}_0^{(n)}\|) = \lim_{n\to+\infty} \frac{1}{n} \mathbb{E} \log(\|\boldsymbol{\gamma}_n \circ \cdots \circ \boldsymbol{\gamma}_1\|) \\ &= \inf_{n\geq 1} \frac{1}{n} \mathbb{E} \log(\|\boldsymbol{\gamma}_n \circ \cdots \circ \boldsymbol{\gamma}_1\|). \end{aligned} \tag{5.36}$$
2. *If* $(\boldsymbol{\eta}_t)$ *is strictly stationary and ergodic and* $\mathbb{E}(\log^+ \|\boldsymbol{\gamma}_0\|)$ *is not finite, the following example, in the scalar case, shows that it is possible to have a stationary solution and at the same time* $\boldsymbol{\gamma} \geq 0$. *Let us take* $\boldsymbol{\eta}_t = \boldsymbol{u}_t$, $\boldsymbol{\delta}_t = 1/\boldsymbol{u}_t$, $\boldsymbol{\gamma}_t = e^{-1}\boldsymbol{u}_{t-1}/\boldsymbol{u}_t$, *where* $(\boldsymbol{u}_t)$ *is defined in Remark* 5.4*. We have* $\sum_{k=0}^{+\infty} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{-k} = (1/\boldsymbol{u}_t) \sum_{k=0}^{+\infty} e^{-k} < \infty$ *a.s. It is easy to see that this process is a strictly stationary (and ergodic) solution. However, using the arguments used in Appendix* 5.1*, we can see that* $\boldsymbol{\gamma} = \limsup_n \frac{1}{n} \log \boldsymbol{\gamma}_0^{(n)} \geq 0$ *and* $\mathbb{E} \log^+ \boldsymbol{\gamma}_0 = \infty$.

For all $\delta \in \mathcal{C}[0,1]$, we can remark that if
$$\text{for all } u \in [0,1],\ \delta(u) > 0,\ \textit{i.e.}\ \inf_{u\in[0,1]} \delta(u) > 0, \tag{5.37}$$
then (5.35) holds. Therefore, we have the following immediate corollary.

Corollary 5.5. *If Eq. (5.37) hold and* ***A5.1*** *or* ***A5.2*** *is verified, Model (5.32) admits a (unique and non-anticipative) positive and strictly stationary (and ergodic) solution in* $\mathcal{C}([0,1])$ *if and only if*

$$\limsup_n \frac{1}{n} \log(\|\boldsymbol{\gamma}_0 \circ \cdots \circ \boldsymbol{\gamma}_{-n+1}\|) < 0 \quad a.s. \tag{5.38}$$

Moreover, under ***A5.2****, Eq. (5.38) is equivalent to*

$$\mathbb{E}\left[\log \|\boldsymbol{\gamma}_n \circ \cdots \circ \boldsymbol{\gamma}_1\|\right] < 0 \text{ for some } n. \tag{5.39}$$

Proof. Under **A5.2**, the equivalence between equations (5.38) and (5.39) comes from Eq. (5.36). Thus, Corollary 5.5 is a direct consequence of Theorem 5.4. □

Since (5.39) is necessary and sufficient under **A5.2**, which does not require the iid assumption, it is clear that this condition is weaker than the sufficient condition, Eq. (5.34), given by Aue et al. (2017).

In order to prove Theorem 5.4, we will use the following general result. It is used, under **A5.1**, to address the other challenge mentioned in (Cerovecki et al., 2019, Remark 1), which consists in showing (5.52) from (5.51). Point 2. of Corollary 5.3 is used to handle this step under **A5.2**.

Lemma 5.4. *Let* $(\boldsymbol{x}_n)_{n\geq 0}$ *and* $(\boldsymbol{y}_n)_{n\geq 0}$ *be real value processes. If* (i) $(\boldsymbol{x}_n)_{n\geq 0}$ *is identically distributed,* (ii) $\boldsymbol{x}_{n+1}$ *and* $\sigma((\boldsymbol{x}_s, \boldsymbol{y}_{s+1}), s \leq n)$ *are independent and* (iii) $\mathbb{P}(\boldsymbol{x}_0 = 0) < 1$ *then*

$$\boldsymbol{y}_n \to 0 \text{ a.s when } n \to \infty \text{ on } G := \{\boldsymbol{x}_n \boldsymbol{y}_n \to 0 \text{ a.s when } n \to \infty\}$$

By replacing the condition $\mathbb{P}(\boldsymbol{x}_0 = 0) < 1$ in the previous lemma by the slightly stronger assumption that $\boldsymbol{x}_0$ is not almost surely constant, we can establish the following more general result:

$$\boldsymbol{y}_n \to 0 \text{ a.s when } n \to \infty \text{ on } \{(\boldsymbol{x}_n \boldsymbol{y}_n) \text{ converges}\},$$

see Proposition 5.1 in Appendix 5.2.

Proof of Lemma 5.4. It suffices to prove that

$$\text{for all } \varepsilon > 0,\ \mathbb{P}\left(\liminf_n \{|\boldsymbol{y}_n| < \varepsilon\} \cap G\right) = \mathbb{P}(\{\sum_{n\geq 1} \mathbb{1}_{\{|\boldsymbol{y}_n| \geq \varepsilon\}} < \infty\} \cap G) = \mathbb{P}(G).$$

Define $G_\varepsilon = \{\sum_{n\geq 1} \mathbb{1}_{\{|\boldsymbol{x}_n \boldsymbol{y}_n| \geq \varepsilon\}} < \infty\} = \liminf_n \{|\boldsymbol{x}_n \boldsymbol{y}_n| < \varepsilon\}$. Note that (iii) implies that

$$\text{there exists } \delta > 0 \text{ such that } \mathbb{P}(|\boldsymbol{x}_0| \geq \delta) > 0. \tag{5.40}$$

Fix ε and for this δ, let's show first that for all $0 < \varepsilon' \leq \delta\varepsilon$,

$$\mathbb{P}(\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon\}}<\infty\}\cap G_{\varepsilon'})=\mathbb{P}(G_{\varepsilon'}). \tag{5.41}$$

Since $\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}<\infty\}\subset\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon\}}<\infty\}$ for all $\varepsilon'\leq\delta\varepsilon$, to prove (5.41), it suffices to show that for all $\varepsilon'\leq\delta\varepsilon$,

$$\mathbb{P}(\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}<\infty\}\cap G_{\varepsilon'})=\mathbb{P}(G_{\varepsilon'}). \tag{5.42}$$

To prove this, we will use a conditional version of the Borel-Cantelli lemma. Let $\varepsilon'>0$. Since

$$|\boldsymbol{x}_n\boldsymbol{y}_n|\geq\delta\mathbb{1}_{\{|\boldsymbol{x}_n|\geq\delta\}}|\boldsymbol{y}_n| \quad\text{and}\quad \mathbb{1}_{\{\delta\mathbb{1}_{\{|\boldsymbol{x}_n|\geq\delta\}}|\boldsymbol{y}_n|\geq\varepsilon'\}}=\mathbb{1}_{\{|\boldsymbol{x}_n|\geq\delta\}}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}},$$

we have

$$\mathbb{1}_{\{|\boldsymbol{x}_n\boldsymbol{y}_n|\geq\varepsilon'\}}\geq\mathbb{1}_{\{|\boldsymbol{x}_n|\geq\delta\}}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}.$$

Define, for all $n\geq 0$, $\boldsymbol{z}_n=\mathbb{1}_{\{|\boldsymbol{x}_n|\geq\delta\}}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}$, it follows that

$$\sum_{n\geq 1}\boldsymbol{z}_n<\infty \quad a.s \quad \text{on} \quad G_{\varepsilon'}. \tag{5.43}$$

Let $\mathcal{F}_n=\sigma((\boldsymbol{x}_s,\boldsymbol{y}_{s+1}),s\leq n)$, for all $n\geq 0$. Since $\boldsymbol{y}_n$ is $\mathcal{F}_{n-1}$-measurable and in view of (ii), $\boldsymbol{x}_n$ and $\mathcal{F}_{n-1}$ are independent, then, also by (i), for all $n\geq 1$,

$$\boldsymbol{m}_n:=\mathbb{E}(\boldsymbol{z}_n|\mathcal{F}_{n-1})=\mathbb{P}(|\boldsymbol{x}_0|\geq\delta)\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}},$$

hence, by (5.40), $\{\sum_{n\geq 1}\boldsymbol{m}_n<\infty\}=\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}<\infty\}$ $a.s$. This result, the fact that $(\mathcal{F}_n)_{n\geq 0}$ is a sequence of nondecreasing σ-algebras, $\boldsymbol{z}_n$ is $\mathcal{F}_n$-measurable, and Theorem 1 of Chen (1978) (see also Freedman (1973, Eq. 5 and 6)) imply that

$$\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon'/\delta\}}<\infty \quad a.s \quad \text{on} \quad \{\sum_{n\geq 1}\boldsymbol{z}_n<\infty\}. \tag{5.44}$$

Equation (5.42) is a direct consequence of (5.43) and (5.44). This show (5.41). Since $(G_{1/m})_{m\in\mathbb{N}}$ is a nonincreasing sequence of sets and

$$G=\cap_{m\geq 1}\cup_{N\geq 1}\cap_{n\geq N}\{|\boldsymbol{x}_n\boldsymbol{y}_n|<1/m\}=\cap_{m\geq 1}G_{1/m},$$

it follows by the monotone convergence theorem and (5.41) that

$$\begin{aligned}\mathbb{P}(G)&=\lim_{m\to\infty}\mathbb{P}(G_{1/m})\\&=\lim_{m\to\infty}\mathbb{P}(\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon\}}<\infty\}\cap G_{1/m})\\&=\mathbb{P}(\{\sum_{n\geq 1}\mathbb{1}_{\{|\boldsymbol{y}_n|\geq\varepsilon\}}<\infty\}\cap G),\end{aligned}$$

which completes the proof. □

Proof of Theorem 5.4.

We prove *1*. Let us first consider that **A5.1** holds. For all $t \in \mathbb{Z}$ and $n \in \mathbb{N}$, let $\boldsymbol{w}_{t,n} = \sum_{k=0}^{n} \boldsymbol{\gamma}_t^{(k)} \boldsymbol{\delta}_{t-k}$. Suppose that (5.33) has a positive stationary solution $(\boldsymbol{h}_t)_{t\in\mathbb{Z}}$ in F. By iterating (5.33), it follows that for all $n \geq 0$ and $t \in \mathbb{Z}$, $\boldsymbol{w}_{t,n} \leq \boldsymbol{h}_t$. This implies that *a.s.* $(\boldsymbol{w}_{t,n})_{n\geq 0}$ is a sequence of nondecreasing functions bounded by $\boldsymbol{h}_t$. Therefore, *a.s.* $(\boldsymbol{w}_{t,n})_n$ converges pointwise (with finite limit) to $\boldsymbol{w}_t$. Noting that for all n and t, $\boldsymbol{w}_0 = \boldsymbol{w}_{t,n-1} + \boldsymbol{\gamma}_0^{(n)} \boldsymbol{w}_{-n}$, one has a.s.

$$\boldsymbol{\gamma}_0^{(n)} \boldsymbol{w}_{-n} \xrightarrow{\text{pw}} 0 \quad when \quad n \to \infty$$

Since

$$(\inf \boldsymbol{w}_{-n}) \boldsymbol{\gamma}_0^{(n)} \boldsymbol{e} \leq \boldsymbol{\gamma}_0^{(n)} \boldsymbol{w}_{-n}, \tag{5.45}$$

we have

$$(\inf \boldsymbol{w}_{-n}) \boldsymbol{\gamma}_0^{(n)} \boldsymbol{e} \xrightarrow{\text{pw}} 0 \quad when \quad n \to \infty \quad a.s. \tag{5.46}$$

Consider that (5.35) holds. The proof relies on the following intermediate results.

a) there exists $n_0 \geq 1$ such that $\mathbb{P}(\inf \boldsymbol{w}_{0,n_0} = 0) < 1$.

b) $\limsup_k (n_0 k)^{-1} \log(\|\boldsymbol{\gamma}_0^{(n_0 k)}\|) < 0 \quad a.s,$

c) $\limsup_k (n_0 k + p)^{-1} \log(\|\boldsymbol{\gamma}_0^{(n_0 k + p)}\|) < 0 \quad a.s$ for all $p = 0, 1, \cdots n_0 - 1$

Let us proceed by contradiction to prove *a)*. Suppose that

$$\text{for all } n \geq 1, \quad \inf \boldsymbol{w}_{0,n} = 0 \ a.s. \tag{5.47}$$

For all $n \geq 1$, let $\boldsymbol{J}_n^{\omega} = \{u \in K : \ \boldsymbol{w}_{0,n}(u) = 0\}$. The sequence $(\boldsymbol{w}_{0,n})_{n\geq 0}$ is continuous, positive, and non-decreasing, therefore, a.s., $(\boldsymbol{J}_n)_{n\geq 0}$ is a sequence of non-empty, non-increasing, random, compact sets. By Cantor's intersection theorem and Eq. (5.47), $\boldsymbol{J} := \bigcap_{n\geq 0} \boldsymbol{J}_n \neq \varnothing$ *a.s.* The sequence $(\boldsymbol{w}_{0,n})_{n\geq 1}$ converges pointwise to $\boldsymbol{w}_0$, and almost surely $\boldsymbol{w}_{0,n} = 0$ on $\boldsymbol{J}$ for all $n \geq 1$. This implies that $\boldsymbol{w}_0 = 0$ on $\boldsymbol{J}$ *a.s.*, which contradicts equation (5.35). This completes the proof of part *a)*.

We now prove *b)*. By iteration, note that for all $n \geq 0$,

$$\boldsymbol{h}_0 = \boldsymbol{w}_{t,n-1} + \boldsymbol{\gamma}_0^{(n)} \boldsymbol{h}_{-n}. \tag{5.48}$$

It follows that $(\boldsymbol{\gamma}_0^{(n)} \boldsymbol{h}_{-n})$ is a sequence of nonincreasing functions, pointwise bounded by $\boldsymbol{h}_t$ and then, almost surely, it converges pointwise. Since, by continuity, $\boldsymbol{h}_0$ is almost surely bounded then $\lim_{K\to\infty} \mathbb{P}(\sup \boldsymbol{h}_0 < K) = 1$. It follows by (5.35) that there exists $K > 0$ and $\epsilon >$ such that $\mathbb{P}(\sup \boldsymbol{h}_0 < K, \ \inf \boldsymbol{w}_0 > \epsilon) > 0$. Noting that

$$\boldsymbol{\gamma}_0^{(n)} \boldsymbol{h}_{-n} \mathbb{1}_{\{\sup \boldsymbol{h}_{-n} \leq K, \ \inf \boldsymbol{w}_{-n} > \epsilon\}} < (K/\epsilon) \inf \boldsymbol{w}_{-n} \boldsymbol{\gamma}_0^{(n)} \boldsymbol{e},$$

it follows by (5.46) that

$$\boldsymbol{\gamma}_0^{(n)}\boldsymbol{h}_{-n}\mathbb{1}_{\{\sup \boldsymbol{h}_{-n}<K,\ \inf \boldsymbol{w}_{-n}>\epsilon\}} \xrightarrow{\text{pw}} 0 \quad as \quad n\to\infty. \tag{5.49}$$

Since the ergodic theorem implies that almost surely, $\mathbb{1}_{\{\sup \boldsymbol{h}_{-n}<K,\ \inf \boldsymbol{w}_{-n}>\epsilon\}} = 1$ for an infinite number of n, it follows that the sub-sequence $(\boldsymbol{\gamma}_0^{(n)}\boldsymbol{h}_{-n} : \ \mathbb{1}_{\{\sup \boldsymbol{h}_{-n}<K,\ \inf \boldsymbol{w}_{-n}>\epsilon\}} = 1)$ converges, almost surely, pointwise to the limit of $(\boldsymbol{\gamma}_0^{(n)}\boldsymbol{h}_{-n})$. It follows by (5.49) that this limit is 0, i.e.

$$\boldsymbol{\gamma}_0^{(n)}\boldsymbol{h}_{-n} \xrightarrow{\text{pw}} 0 \quad when \quad n\to\infty \quad a.s. \tag{5.50}$$

From this and Eq. (5.48) one has $F \ni \boldsymbol{h}_0 = \boldsymbol{w}_0$ *a.s.* It follows by Dini's Theorem that $(\boldsymbol{w}_{0,n})_n$ converges uniformly to $\boldsymbol{w}_0$ *a.s.* Hence,

$$\|\boldsymbol{w}_0 - \boldsymbol{w}_{0,n}\| = \|\boldsymbol{\gamma}_0^{(n)}\boldsymbol{w}_{-n}\| \longrightarrow 0 \quad when \quad n\to\infty \quad a.s. \tag{5.51}$$

For all $k \geq 0$, let $\boldsymbol{x}_k = \inf \boldsymbol{w}_{-n_0k,n_0}$ and $\boldsymbol{y}_k = \|\boldsymbol{\gamma}_0^{(n_0k)}\|$. Note that $(\boldsymbol{x}_k,\ \boldsymbol{y}_k(u))$ verifies the conditions of Lemma 5.4 because of *a)* and the fact that $(\boldsymbol{\eta}_t)$ is iid. Since $inf\boldsymbol{w}_{-n_0k,n_0}\|\boldsymbol{\gamma}_0^{(n_0k)}\boldsymbol{e}\| = inf\boldsymbol{w}_{-n_0k,n_0}\|\boldsymbol{\gamma}_0^{(n_0k)}\| \leq \|\boldsymbol{\gamma}_0^{(n_0k)}\boldsymbol{w}_{-n_0k}\|$, it follows by (5.51), and Lemma 5.4 that

$$\|\boldsymbol{\gamma}_0^{(n_0k)}\| \longrightarrow 0 \quad when \quad k\to\infty \quad a.s. \tag{5.52}$$

Since $(\|\boldsymbol{\gamma}_0^{(n_0k)}\|)_k$ is a sub-multiplicative sequence, by (5.52) and Point 1. of Corollary 5.3, we have

$$\limsup_k (n_0k)^{-1}\log(\|\boldsymbol{\gamma}_0^{(n_0k)}\|) = \limsup_k (\frac{k}{n_0k})[k^{-1}\log(\|\boldsymbol{\gamma}_0^{(n_0k)}\|)] < 0 \quad a.s.$$

This concludes the proof of *b)*. To prove *c)*, remark by stationarity and *b)* that for all $p = 0, 1, \cdots n_0 - 1$, $\limsup_k (n_0k)^{-1}\log(\|\boldsymbol{\gamma}_{-p}^{(n_0k)}\|) < 0$. Therefore

$$\begin{aligned}
\psi :&= \limsup_k (n_0k+p)^{-1}\log(\|\boldsymbol{\gamma}_0^{(n_0k+p)}\|) \\
&\leq \limsup_k (n_0k+p)^{-1}\log(\|\boldsymbol{\gamma}_0^{(p)}\|) + \limsup_k (n_0k+p)^{-1}\log(\|\boldsymbol{\gamma}_{-p}^{(n_0k)}\|) \\
&= \limsup_k (\frac{n_0k}{n_0k+p})[(n_0k)^{-1}\log(\|\boldsymbol{\gamma}_{-p}^{(n_0k)}\|)] < 0 \quad a.s.
\end{aligned} \tag{5.53}$$

Noting that $\mathrm{N} = \cup_{0\leq p\leq n_0-1}\{n_0k+p : \ k\in\mathbb{N}\}$, it follows that

$$\limsup_n \frac{1}{n}\log\|\boldsymbol{\gamma}_0^{(n)}\| \leq \max_{0\leq p\leq n_0-1}\left(\limsup_k (n_0k+p)^{-1}\log(\|\boldsymbol{\gamma}_0^{(n_0k+p)}\|)\right) < 0.$$

which gives the first point under **A5.1**.

We now prove the claim under **A5.2**. First observe that the iid assumption is only used, in Lemma 5.4, to derive (5.52) from (5.51). Therefore, all the results showed before

(5.51) hold under **A5.2**. Hence, (5.51) implies that

$$\mathbb{1}_{\{\inf \boldsymbol{w}_{-n}>\epsilon\}}\|\boldsymbol{\gamma}_0^{(n)}\| \longrightarrow 0 \quad when \quad k \to \infty \quad a.s.$$

Therefore, the sequence $(\|\boldsymbol{\gamma}_0^{(n)}\| : \inf \boldsymbol{w}_{-n} > \epsilon\})$ converges almost surely to 0. The result follows from arguments used in Remark 5.4 (to define the dynamic system), Appendix 5.1 (to verifies the condition of the corollary) and Point 2. of Corollary 5.3.

We now prove 2. We have [2]

$$\begin{aligned}\limsup_n \frac{1}{n}\log(\|\boldsymbol{\gamma}_0^{(n)}\boldsymbol{\delta}_{-n}\|) &\leq \limsup_n \frac{1}{n}[\|\boldsymbol{\gamma}_0^{(n)}\| + \log^+(\|\boldsymbol{\delta}_{-n}\|)] \\ &\leq \boldsymbol{\gamma} + \limsup_n \frac{1}{n}\log^+(\|\boldsymbol{\delta}_{-n}\|) \\ &= \boldsymbol{\gamma} < 0.\end{aligned}$$

Therefore, by Cauchy's rule, $(\boldsymbol{w}_{0,n})_{n\geq 1}$ converges absolutely almost surely. Thus, $\boldsymbol{w}_0 \in F$ *a.s.* It is easy to verify that the continuous, positive, stationary process $(\boldsymbol{w}_t)_n$ is non-anticipative and satisfies (5.33). The proof of the uniqueness is standard, see Appendix 5.1. This completes the proof. □

4 Perspective

The main result of this chapter extends the result of Kesten (1975) on the growth rate of sums of stationary sequences to superadditive processes. Our result is established under weaker conditions than those of Kesten (1975). Using a result from Tanny (1974), Kesten show in the same paper that if $\{\boldsymbol{S}_n\}_{n=1}^{\infty}$ is an additive sequence, then on $\{\boldsymbol{S}_n \to \infty\}$, $0 < \liminf n^{-1}\boldsymbol{S}_n < \limsup n^{-1}\boldsymbol{S}_n = \infty$ *a.s.* or $\lim n^{-1}\boldsymbol{S}_n$ exists and $\lim n^{-1}\boldsymbol{S}_n > 0$. An interesting question that could be considered for further work is to see if this result also generalizes to superadditive processes. That is, in which cases can the limit superior in Theorem 5.1 be replaced by a limit.

5 Appendix: Complementary Proofs

5.1 Complement to Remark 5.4

Let $\boldsymbol{\delta}_n = \frac{1}{\boldsymbol{u}_{n-1}}$. Since $\boldsymbol{u}_0$ is not almost surely constant then there exists $s > 0$ such that $\mathbb{P}_{\boldsymbol{u}}(\boldsymbol{\delta}_0 > s) > 0$. Let $A = \{\boldsymbol{\delta}_0 > s\}$. It is easy to see that $\boldsymbol{\gamma}_n\boldsymbol{\delta}_n = \frac{e^{-n}}{\boldsymbol{u}_0} \to 0$ *a.s.* as $n \to 0$. Thus $\boldsymbol{\gamma}_n\mathbb{1}_A \circ T^n = \boldsymbol{\gamma}_n\mathbb{1}_{\boldsymbol{\delta}_n>s} \leq \boldsymbol{\gamma}_n\boldsymbol{\delta}_n \to 0$ *a.s.* It follows that $(\boldsymbol{\gamma}_n : T^n \in A)$ converges almost surely to 0.

2. For all non negative stationary process $(X_n)_{n\geq 1}$ such that $\mathbb{E}X_1 < \infty$, one has $\limsup_{n\to\infty} n^{-1}X_n = 0$. Indeed, for all $\epsilon > 0$, noting that the function $f(t) = P\left(t^{-1}X_1 > \epsilon\right)$ is decreasing, we have $\sum_{n=1}^{\infty} P\left(n^{-1}X_n > \epsilon\right) \leq \int_0^{\infty} P\left(\epsilon^{-1}X_1 > t\right) dt = \epsilon^{-1}EX_1 < \infty$. The convergence follows from the Borel-Cantelli lemma.

On the other hand, since $\boldsymbol{\gamma}_n = \frac{\boldsymbol{u}_n}{\boldsymbol{u}_0} e^{-n}$, then $\limsup_n n^{-1} \log \boldsymbol{\gamma}_n < 0$ *a.s.* implies that $(\boldsymbol{u}_n e^{-n})$ converges to 0 *a.s.* However, one has

$$\sum_{n=0}^{\infty} \mathbb{P}(\boldsymbol{u}_n e^{-n} > 1) = \sum_{n=0}^{\infty} \mathbb{P}(\ln^+ \boldsymbol{u}_0 > n) \geq \int_0^{\infty} \mathbb{P}(\ln^+ \boldsymbol{u}_0 > t) dt$$
$$= \mathbb{E}(\ln^+ \boldsymbol{u}_0) = \infty.$$

It follows by the second Borel-Cantelli lemma that $\mathbb{P}(\limsup\{\boldsymbol{u}_n e^{-n} > 1\}) = 1$ and then $(\boldsymbol{z}_n e^{-n})$ does not converge to 0.

Now we compute $\mathbb{E} \ln^+ \boldsymbol{\gamma}_1$. Let a real $K > 0$ such that $\mathbb{P}(\ln^+ \boldsymbol{u}_0 \leq K) > 0$. Since $\ln^+ \boldsymbol{\gamma}_1 \geq \ln^+ \boldsymbol{u}_1 - \ln^+ \boldsymbol{u}_0$, it follows that $\ln^+ \boldsymbol{\gamma}_1 \geq \ln^+ \boldsymbol{u}_1 \mathbb{1}_{\ln^+ \boldsymbol{u}_0 \leq K} - \ln^+ \boldsymbol{u}_0 \mathbb{1}_{\ln^+ \boldsymbol{u}_0 \leq K}$. Hence $\mathbb{E} \ln^+ \boldsymbol{\gamma}_1 \geq \mathbb{E} \ln^+ \boldsymbol{u}_1 \mathbb{P}(\ln^+ \boldsymbol{u}_0 \leq K) - \mathbb{E} \ln^+ \boldsymbol{u}_0 \mathbb{1}_{\ln^+ \boldsymbol{u}_0 \leq K} = \infty$, because the second term is finite.

5.2 On the convergence of the product of two independent random elements

The following result, which generalises Lemma 5.4 is of independent interest.

Proposition 5.1. *Let $(\boldsymbol{x}_n)_{n\geq 0}$ and $(\boldsymbol{y}_n)_{n\geq 0}$ be real value processes. If (i) $(\boldsymbol{x}_n)_{n\geq 0}$ is identically distributed, (ii) $\boldsymbol{x}_{n+1}$ and $\mathcal{F}_n := \sigma((\boldsymbol{x}_s, \boldsymbol{y}_{s+1}), s \leq n)$, are independent and (iii) $\boldsymbol{x}_0$ is not almost surely constant, then*

1. $\boldsymbol{x}_n \boldsymbol{y}_n \to 0$ *a.s when* $n \to \infty$ *on* $\{\boldsymbol{x}_n \boldsymbol{y}_n$ *converges*$\}$
2. $\boldsymbol{y}_n \to 0$ *a.s when* $n \to \infty$ *on* $\{\boldsymbol{x}_n \boldsymbol{y}_n$ *converges*$\}$
3. *If $(\boldsymbol{x}_n \boldsymbol{y}_n)$ converges in probability then the limit is* 0.

Proof. For the first point, it suffices to prove that

$$\mathbb{P}\left(\{|\limsup \boldsymbol{x}_n \boldsymbol{y}_n| > 0,\ \boldsymbol{x}_n \boldsymbol{y}_n \text{ converges}\}\right) = 0. \tag{5.54}$$

For all $\epsilon > 0$ and $t \in \mathbb{R}$, let $B(t, \epsilon) := (t - \epsilon,\ t + \epsilon)$. Let $\boldsymbol{z} = \limsup \boldsymbol{x}_n \boldsymbol{y}_n$ and $G = \{\boldsymbol{x}_n \boldsymbol{y}_n$ converges$\}$. Note that on G, $\boldsymbol{z}$ is finite and $(\boldsymbol{x}_n \boldsymbol{y}_n)$ converges to $\boldsymbol{z}$. We argue by contradiction: suppose that $\mathbb{P}\left(\{|\boldsymbol{z}| > 0\} \cap G\right) > 0$. Since this condition implies that $\mathbb{P}(G) > 0$, let $\mathbb{P}^G$ be the conditional probability given G. Noting that $\mathbb{P}^G(|\boldsymbol{z}| > 0) > 0$ we have that the support of $\boldsymbol{z}$ under $\mathbb{P}^G$ contains a non-zero element z_0. Thus, for all $\epsilon > 0$ we have $\mathbb{P}^G\left(\boldsymbol{z} \in B(z_0, \epsilon)\right) > 0$ i.e.

$$\mathbb{P}\left(\{\boldsymbol{z} \in B(z_0, \epsilon)\} \cap G\right) > 0. \tag{5.55}$$

The condition *(iii)* implies that the support of $\boldsymbol{x}_0$ under $\mathbb{P}$ contains at least two different elements x_1 and x_2. Since $x_1 \neq x_2$, we can assume without loss of generality that $x_1 \neq 0$.

Let $y_0 = z_0/x_1$, and $\epsilon_0,\ \epsilon_1,\ \epsilon_2,\ \epsilon_3 > 0$ such that

$$\begin{array}{l} 1)\ \{xy:\ (x,y) \in B(x_2,\epsilon_2) \times B(y_0,\epsilon_3)\} \cap B(z_0,\epsilon_0) = \varnothing, \\ 2)\ \text{if } (z,x) \in B(z_0,\epsilon_0) \times B(x_1,\epsilon_1) \text{ then } z/x \in B(y_0,\epsilon_3). \end{array} \tag{5.56}$$

The point *1)* in (5.56) comes from the fact that $(x,y) \mapsto xy$ is continuous at (x_2, y_0) and $x_2 y_0 = z_0(x_2/x_1) \neq z_0$. The second point is because $(z,x) \mapsto z/x$, defined on $B(z_0,\epsilon) \times B(x_1,\epsilon)$ for ϵ small enough, is continuous at (z_0, x_1). Indeed, for *1)*, take $\epsilon_0 > 0$ and $\delta > 0$ such that

$$B(x_2 y_0, \delta)\} \cap B(z_0,\epsilon_0) = \varnothing.$$

Choose $\epsilon_2,\ \epsilon_3 > 0$ such that

$$(x,y) \in B(x_2,\epsilon_2) \times B(y_0,\epsilon_3),$$

we have $xy \in B(x_2 y_0, \delta)$ and thus

$$\{xy:\ (x,y) \in B(x_2,\epsilon_2) \times B(y_0,\epsilon_3)\} \cap B(z_0,\epsilon_0) = \varnothing.$$

Noting that this statement remains true for smaller ϵ_0, ϵ_2 and ϵ_3, let's fix ϵ_2 and ϵ_3, and choose ϵ_0 smaller than its previous value and take also ϵ_1 such that

$$(z,x) \in B(z_0,\epsilon_0) \times B(x_1,\epsilon_1).$$

We thus have $z/x \in B(y_0,\epsilon_3)$.

For $i \in \{1,\ 2\}$, we have by the strong law of large numbers that

$$n^{-1} \sum_{k=0}^{n} \mathbb{1}_{\{\boldsymbol{x}_k \in B(x_1,\epsilon_1)\}} \to \mathbb{P}(\boldsymbol{x}_0 \in B(x_1,\epsilon_1))\ a.s \text{ when } n \to \infty.$$

Since $\mathbb{P}(\boldsymbol{x}_0 \in B(x_1,\epsilon_1)) > 0$, then $\mathbb{P}(S) = 1$ where $S = \{\sum_{k=0}^{n} \mathbb{1}_{\{\boldsymbol{x}_k \in B(x_1,\epsilon_1)\}} \to \infty\}$. We have by this result and (5.55) that $\mathbb{P}(E) > 0$ where

$$E = S \cap \{\boldsymbol{z} \in B(z_0,\epsilon_0)\} \cap G.$$

Since on E $(\boldsymbol{x}_n \boldsymbol{y}_n)$ converges to $\boldsymbol{z}$, which is in the open set $B(z_0,\epsilon_0)$, then there exists an integer N (random integer) such that if $n \geq N$, then $\boldsymbol{x}_n \boldsymbol{y}_n \in B(z_0,\epsilon_0)$. It follows by *1)* (5.56) that

$$\sum_{k \geq 1} \mathbb{1}_{\{\boldsymbol{x}_k \in B(x_2,\epsilon_2),\ \boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} < \infty \quad a.s \quad on \quad E. \tag{5.57}$$

Since on S, and thus on E, $\boldsymbol{x}_n \in B(x_1,\epsilon_1)$ for infinitely many n, it follows also that

$$\{n:\ \boldsymbol{x}_n \boldsymbol{y}_n \in B(z_0,\epsilon_0),\ \boldsymbol{x}_n \in B(x_1,\epsilon_1)\}$$

is infinite. Therefore, we have by *2)* (5.56) that

$$\sum_{k\geq 1} \mathbb{1}_{\{\boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} = \infty \quad on \quad E. \tag{5.58}$$

To arrive at a contradiction, let's also show that

$$\sum_{k\geq 1} \mathbb{1}_{\{\boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} < \infty \quad a.s \quad on \quad E.$$

In views of (5.57), it is equivalent to show that

$$\sum_{k\geq 1} \mathbb{1}_{\{\boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} < \infty \quad a.s \quad on \quad \{\sum_{k\geq 1} \boldsymbol{z}_k < \infty\} \tag{5.59}$$

where

$$\boldsymbol{z}_k = \mathbb{1}_{\{\boldsymbol{x}_k \in B(x_2,\epsilon_2),\ \boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} = \mathbb{1}_{\{\boldsymbol{x}_k \in B(x_2,\epsilon_2)} \mathbb{1}_{\boldsymbol{y}_k \in B(y_0,\epsilon_3)\}}.$$

To get this result, remark that $\boldsymbol{z}_n$ is $\mathcal{F}_n$-measurable and since $\boldsymbol{y}_n$ is $\mathcal{F}_{n-1}$-measurable and $\boldsymbol{x}_n$ and $\mathcal{F}_{n-1}$ are independent, then for all $n \geq 1$

$$\boldsymbol{m}_n := \mathbb{E}(\boldsymbol{z}_n | \mathcal{F}_{n-1}) = \mathbb{P}(\boldsymbol{x}_0 \in B(x_2, \epsilon_2)) \mathbb{1}_{\{\boldsymbol{y}_n \in B(y_0,\epsilon_3)\}}.$$

It follows from the converse part of Theorem 1 of Chen (1978) that

$$\sum_{k\geq 1} \boldsymbol{m}_k < \infty \quad a.s \quad on \quad \{\sum_{k\geq 1} \boldsymbol{z}_k < \infty\}.$$

Noting that

$$\{\sum_{k\geq 1} \boldsymbol{m}_n < \infty\} = \{\sum_{k\geq 1} \mathbb{1}_{\{\boldsymbol{y}_k \in B(y_0,\epsilon_3)\}} < \infty\} \quad a.s,$$

by the fact that $\mathbb{P}(\boldsymbol{x}_0 \in B(x_2, \epsilon_2)) > 0$, (5.59) follows from the previous result. This contradicts (5.58) since $\mathbb{P}(E) > 0$, and thus we have (5.54).

The second point follows from the first point and Lemma 5.4.

For the last point, note that the convergence in probability implies convergence on a sub-sequence $(\boldsymbol{x}_{\phi(n)}\boldsymbol{y}_{\phi(n)})$ almost surely. Since $(\boldsymbol{x}_{\phi(n)}, \boldsymbol{y}_{\phi(n)})$ checks the conditions of Proposition 5.1, the result follows from the first point. This concludes the proof. □

5.3 Ergodic Lemmas

This results may not be new. Since we have not been able to find it in the literature, we provide a proof.

Lemma 5.5. *Let $I \in \mathcal{I}_\mu$ with $\mu(I) > 0$. Let ν be the probability measure in $(\Omega,\ \mathcal{B})$ given by the conditional probability given I (i.e for all $A \in \mathcal{B}$, $\quad \nu(A) = \mu(I)^{-1}\mu(A \cap I)$). Then $(\Omega, \mathcal{B}, \nu, T)$ is a measure-preserving dynamical system, i.e.*

$$\textit{for all } A \in \mathcal{B},\ \nu(T^{-1}(A)) = \nu(A),$$

Proof. For all $I \in \mathcal{I}_\mu$ and $A \in \mathcal{B}$, because

$$\begin{aligned}
&\mu(I \Delta T^{-1}(I)) = 0, \\
&I \cup I\Delta T^{-1}(I) = T^{-1}(I) \cup I\Delta T^{-1}(I) \text{ and} \\
&T^{-1}(A) \cap T^{-1}(I) = T^{-1}(A \cap I),
\end{aligned}$$

one has

$$\begin{aligned}
\mu(T^{-1}(A) \cap I) &= \mu(T^{-1}(A) \cap (I \cup I\Delta T^{-1}(I)) \\
&= \mu\left(T^{-1}(A) \cap (T^{-1}(I) \cup I\Delta T^{-1}(I))\right) \\
&= \mu(T^{-1}(A) \cap T^{-1}(I)) = \mu(T^{-1}(A \cap I)) \\
&= \mu(A \cap I).
\end{aligned} \tag{5.60}$$

The result follows by dividing by $\mu(I)$. □

Lemma 5.6. *For all invariant set I, Let $C(I) := \bigcap_{n=0}^{\infty} T^{-n}(I)$, where $T^0 = Id_\Omega$. One has*

1. $C(I) \in \mathcal{I}_\mu$
2. $\mu\left(C(I)\right) = \mu(I)$
3. *for all $\omega \in C(I)$ and $n \geq 0$, $T^n(\omega) \in C(I)$.*

Proof. Let show the first point. If $I \in \mathcal{I}$ then

$$\begin{aligned}
\mu\left(T^{-2}(I)\Delta T^{-1}(I)\right) &= \mu(T^{-2}(I) \cup T^{-1}(I)) - \mu(T^{-2}(I) \cap T^{-1}(I)) \\
&= \mu(T^{-1}(I) \cup I) - \mu(T^{-1}(I) \cap I) \\
&= \mu\left(T^{-1}(I)\Delta I\right) = 0,
\end{aligned}$$

and then $T^{-1}(I) \in \mathcal{I}_\mu$. Hence, by recurrence, we show that $C(I)$ is the intersection of elements of the σ-algebra $\mathcal{I}_\mu$, and thus $C(I) \in \mathcal{I}_\mu$.

For the second statement, using (5.60) for $A = I$, one has $\mu(T^{-1}(I) \cap I) = \mu(I)$. Therefore, by doing the same operation on $\bigcap_{k=0}^{n} T^{-k}(I)$ for $n = 1, 2, \cdots$, one has by recurrence and the monotone convergence theorem that $\mu\left(C(I)\right) = \mu(I)$.

For the last one, note that $\omega \in C(I)$ is equivalent to, for all $n \geq 0$, $T^n(\omega) \in I$. It follows that, $\omega \in C(I)$ implies that for all $p \geq 0$ and for all $n \geq 0$, $T^{n+p}(\omega) = T^n(T^p(\omega)) \in I$, i.e. for all $p \geq 0$, $T^p(\omega) \in C(I)$.

□

Conclusion and perspectives

This thesis focuses on the study of the theoretical properties of stochastic IFSs (Iterated Function Systems) and their applications in financial econometrics. Despite their numerous applications in economics and finance, the theoretical properties of IFSs when considered in infinite dimensions or when driven by a non-independent process are not well understood.

We know that IFSs driven by non-iid (independent and identically distributed) sequences can have stationary solutions with heavy tails marginal distribution, such that this solution admits no moment, even of small-order. In Chapter 3, we studied these stationary solutions when the innovation sequences are stationary and ergodic. We established a kind of stability property on the trajectory of this solution, which is weaker than the existence of moments. By exploiting this result, we relaxed the moment existence condition that assumed by Escanciano (2009) and established, under weak assumptions, the strong consistency of the Quasi-Maximum Likelihood Estimator (QMLE) for semi-strong GARCH(p, q) models with non-independent innovations. In Chapter 4, we focused on the stationarity and inference of GARCH-MIDAS models, which have the particularity of not admitting small-order moments. The control property established in Chapter 3 is used to overcome the difficulties caused by this characteristic for the statistical inference of these models. We establish the consistency and asymptotic normality of the QMLE and propose statistical tests to detect the presence of long-term volatility. Our results were illustrated using both simulated and real financial data. In Chapter 5, we first established various extensions of the contraction property of random matrices to products of stochastic operators. Our approach is based on concepts from ergodic theory and dynamical systems. These results were then applied to the study of the stability of affine IFSs in the space of continuous positive functions with compact support. We provided necessary and sufficient conditions for the existence of stationary solutions of these models as well as for functional GARCH models in $\mathcal{C}([0, 1])$.

Our results contribute to a better understanding of the existence of stationary solutions and their probabilistic properties, as well as to the inference of a variety of IFS models in financial econometrics. However, they are far from addressing all the unexplored issues related to these models. We have identified some of these issues for future potential works:

Tails of IFSs driven by non-iid sequence. We know from Goldie (1991) that stationary IFS driven by an idd sequence have a power-tailed marginal distribution under general regularity conditions. However, we have shown that this property does not hold when independence is no longer assumed. Therefore, in addition to the stability property we have proposed, it may be interesting to conduct a thorough study of the tail distribution of these processes. This could lead to potential applications in risk management in finance,

such as studying the extremes of GARCH-MIDAS models.

Asymptotic normality of the QMLE of semi-strong GARCH. In Chapter 3, We discussed why the asymptotic normality of the QMLE for semi-strong GARCH models, which has been left hanging, may fail due to the possible non-existence of moments for these models. As asymptotic distributions are crucial in statistical inference, it would be interesting to consider alternatives to the QMLE that allows to obtain the asymptotic properies of estimators. Horváth and Liese (2004) consider L^p estimators for GARCH models and show the consistency and asymptotic normality of these estimators without any moment assumptions. Similarly, Zhu and Ling (2011) study the self-weighted quasi-maximum exponential likelihood estimator of ARMA-GARCH without moment assumptions on the observed process. These classes of estimators could be considered as potential alternatives to QMLE in the inference of semi-strong GARCH.

Multivariate extension and financial applications of GARH-MIDAS. Univariate volatility models are commonly used to model a single financial instruments. However, in practical applications such as asset management, financial time series are often considered in a multivariate setting. In this context, a multivariate extension of the GARCH-MIDAS model can be useful. This can be achieved by decomposing the conditional covariance matrix into short-term and long-term covariance matrices. Furthermore, since multi-component volatility models have been shown to provide better long-term volatility forecasts than classical conditional volatility models, we believe that multivariate multi-component extensions will enable us to obtain better long-term forecasts of risk measurements, such as portfolio Value at Risk. The usual one-component multiplicative form of GARCH models allows to write standard conditional risk measures as the product of the volatility and a characteristic of the errors distribution (e.g. a quantile in the case of VaR). How the additional volatility component affects the definition of such risk measures, and how they can be estimated, would also be an important feature to explore.

Stationarity of fGARCH in L^p spaces. The necessary and sufficient conditions given in Chapter 5 for the stability of a class of functional GARCH$(1,1)$ models in the space of continuous functions can be easily extended, in the same space, to higher-order GARCH(p,q) models using the same argument. However, since norms are not equivalent in the infinite-dimensional setting, it would be interesting to investigate whether these conditions remain true when considering a different space, such as L^p spaces.

Bibliography

C. Amado and T. Teräsvirta. Specification and testing of multiplicative time-varying garch models with applications. *Econometric Reviews*, 36:421–446, 2017.

D. W. Andrews. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica*, 69:683–734, 2001.

G. Atkinson. Recurrence of co-cycles and random walks. *Journal of the London Mathematical Society*, 2:486–488, 1976.

A. Aue, L. Horváth, and D. F. Pellatt. Functional generalized autoregressive conditional heteroskedasticity. *Journal of Time Series Analysis*, 38:3–21, 2017.

J.-M. Bardet, Y. Boularouk, and K. Djaballah. Asymptotic behavior of the Laplacian quasi-maximum likelihood estimator of affine causal processes. *Electronic Journal of Statistics*, 11:452 – 479, 2017.

M. F. Barnsley. *Fractals everywhere*. Academic press, 2014.

H. Berbee. Recurrence and transience for random walks with stationary increments. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 56:531–536, 1981.

I. Berkes, L. Horváth, and P. Kokoszka. GARCH processes: structure and estimation. *Bernoulli*, 9:201 – 227, 2003.

I. Berkes, S. Hörmann, and J. Schauer. Asymptotic results for the empirical process of stationary sequences. *Stochastic processes and their applications*, 119(4):1298–1324, 2009.

E. Beutner, A. Heinemann, and S. Smeekes. A residual bootstrap for conditional value-at-risk. *arXiv preprint arXiv:1808.09125*, 2018. URL https://EconPapers.repec.org/RePEc:arx:papers:1808.09125.

P. Billingsley. The lindeberg-levy theorem for martingales. *Proceedings of the American Mathematical Society*, 12:788–792, 1961.

M. Borkovec. Extremal behavior of the autoregressive process with arch (1) errors. *Stochastic Processes and their applications*, 85:189–207, 2000.

M. Borkovec and C. Klüppelberg. The Tail of the Stationary Distribution of an Autoregressive Process with ARCH(1) Errors. *The Annals of Applied Probability*, 11(4):1220 – 1241, 2001.

P. Bougerol. Kalman filtering with random coefficients and contractions. *Siam Journal on Control and Optimization*, 31:942–959, 1993.

P. Bougerol and N. Picard. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52:115–127, 1992a.

P. Bougerol and N. Picard. Strict stationarity of generalized autoregressive processes. *The Annals of Probability*, 20:1714 – 1730, 1992b.

A. Brandt. The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability*, 18:211–220, 1986.

D. Buraczewski, E. Damek, and T. Mikosch. *Stochastic models with power-law tails.* Springer, 2016.

G. Cavaliere, H. B. Nielsen, R. S. Pedersen, and A. Rahbek. Bootstrap inference on the boundary of the parameter space, with application to conditional volatility models. *Journal of Econometrics*, 227:241–263, 2022.

C. Cerovecki, C. Francq, S. Hörmann, and J.-M. Zakoïan. Functional GARCH models: The quasi-likelihood approach and its applications. *Journal of econometrics*, 209:353–375, 2019.

L. H. Y. Chen. A short note on the conditional Borel-Cantelli lemma. *The Annals of Probability*, 6(4):699 – 700, 1978.

C. Conrad and R. F. Engle. Modelling volatility cycles: the $(\mathrm{MF})^2$ GARCH model. *NYU Stern School of Business Forthcoming*, 2021.

C. Conrad and K. Loch. Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30:1090–1114, 2015.

C. Conrad and M. Schienle. Testing for an omitted multiplicative long-term component in garch models. *Journal of Business & Economic Statistics*, 38:229–242, 2020.

C. Conrad, A. Custovic, and E. Ghysels. Long-and short-term cryptocurrency volatility components: A garch-midas analysis. *Journal of Risk and Financial Management*, 11: 23, 2018.

R. Davis and S. Resnick. Limit theory for the sample covariance and correlation functions of moving averages. *The Annals of Statistics*, 14:533–558, 1986.

P. Diaconis and D. Freedman. Iterated random functions. *SIAM review*, 41:45–76, 1999.

Z. Ding and C. W. Granger. Modeling volatility persistence of speculative returns: a new approach. *Journal of econometrics*, 73:185–215, 1996.

J. H. Elton. A multiplicative ergodic theorem for lipschitz maps. *Stochastic Processes and their Applications*, 34:39–47, 1990.

R. F. Engle and G. Lee. A long-run and short-run component model of stock return volatility. *Cointegration, causality, and forecasting: A Festschrift in honour of Clive WJ Granger*, pages 475–497, 1999.

R. F. Engle, E. Ghysels, and B. Sohn. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95:776–797, 2013.

J. C. Escanciano. Quasi-maximum likelihood estimation of semi-strong GARCH models. *Econometric Theory*, 25:561–570, 2009.

A. Eskin and M. Mirzakhani. Invariant and stationary measures for the action on moduli space. *Publications mathématiques de l'IHÉS*, 127:95–324, 2018.

C. Francq and L. Q. Thieu. QML Inference for volatility models with covariates. *Econometric Theory*, 35:37–72, 2019.

C. Francq and J.-M. Zakoian. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10:605 – 637, 2004.

C. Francq and J.-M. Zakoian. Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, 117:1265–1284, 2007.

C. Francq and J.-M. Zakoïan. Testing the nullity of GARCH coefficients: correction of the standard tests and relative efficiency comparisons. *Journal of the American Statistical Association*, 104:313–324, 2009.

C. Francq and J.-M. Zakoian. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons, Second Edition, 2019.

C. Francq and J.-M. Zakoian. Testing the existence of moments for GARCH processes. *Journal of Econometrics*, 227:47–64, 2022.

C. Francq, L. Horvath, and J.-M. Zakoïan. Sup-tests for linearity in a general nonlinear AR(1) model. *Econometric Theory*, 26:965–993, 2010.

C. Francq, O. Wintenberger, and J.-M. Zakoian. Goodness-of-fit tests for Log-GARCH and EGARCH models. *Test*, 27:27–51, 2018.

D. Freedman. Another note on the Borel–Cantelli lemma and the strong law, with the Poisson approximation as a by-product. *The Annals of Probability*, (6):910 – 925, 1973.

H. Furstenberg and H. Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31:457–469, 1960.

R. Giacomini, D. N. Politis, and H. White. A warp-speed method for conducting monte carlo experiments involving bootstrap estimators. *Econometric theory*, 29:567–589, 2013.

C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, 1:126 – 166, 1991.

C. M. Goldie and R. A. Maller. Stability of perpetuities. *The Annals of Probability*, 28: 1195–1218, 2000.

J. M. Hammersley and D. J. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Bernoulli 1713, Bayes 1763, Laplace 1813*. 1965.

H. Han and D. Kristensen. Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *Journal of Business & Economic Statistics*, 32:416–429, 2014.

B. E. Hansen. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, 64:413–430, 1996.

C. He and T. Teräsvirta. Properties of moments of a family of GARCH processes. *Journal of Econometrics*, 92:173–192, 1999.

S. Hörmann. Augmented GARCH sequences: Dependence structure and asymptotics. *Bernoulli*, 14(2):543 – 561, 2008.

S. Hörmann and P. Kokoszka. Weakly dependent functional data. *The Annals of Statistics*, 38(3):1845 – 1884, 2010.

S. Hörmann, L. Horváth, and R. Reeder. A functional version of the ARCH model. *Econometric Theory*, 29:267–288, 2013.

L. Horváth and F. Liese. Lp-estimators in ARCH models. *Journal of Statistical Planning and Inference*, 119:277–309, 2004.

S. T. Jensen and A. Rahbek. Asymptotic inference for nonstationary GARCH. *Econometric Theory*, 20:1203–1226, 2004.

H. Kesten. Random difference equations and Renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

H. Kesten. Sums of stationary sequences cannot grow slower than linearly. *Proceedings of the American Mathematical Society*, 49:205–211, 1975.

J. F. C. Kingman. Subadditive Ergodic Theory. *The Annals of Probability*, 1:883 – 899, 1973.

C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *Journal of Applied Probability*, 41:601–622, 2004.

J.-P. Kreiss, E. Paparoditis, and D. N. Politis. On the range of validity of the autoregressive sieve bootstrap. *The Annals of Statistics*, 39:2103 – 2130, 2011.

S. Kühnert. Functional ARCH and GARCH models: A Yule-Walker approach. *Electronic Journal of Statistics*, 14:4321 – 4360, 2020.

S.-W. Lee and B. E. Hansen. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric theory*, 10:29–52, 1994.

A. Leucht, J.-P. Kreiss, and M. H. Neumann. A model specification test for GARCH(1,1) processes. *Scandinavian Journal of Statistics*, 42:1167–1193, 2015.

Y. B. Maïnassara, O. Kadmiri, and B. Saussereau. Estimation of multivariate asymmetric power garch models. *Journal of Multivariate Analysis*, 192:–, 2022.

T. Mikosch and C. Starica. Is it really long memory we see in financial returns. *Extremes and integrated risk management*, 12:149–168, 2000.

A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160:246–256, 2011.

K. Schmidt. Recurrence of cocycles and stationary random walks. *Lecture Notes-Monograph Series*, pages 78–84, 2006.

K. Shimizu. The bootstrap does not alwayswork for heteroscedasticmodels. *Statistics & Risk Modeling*, 30:189–204, 2013.

A. Skrobotov, R. Pedersen, and R. Ibragimov. New approaches to robust inference on market (non-) efficiency, volatility clustering and nonlinear dependence. *Volatility Clustering and Nonlinear Dependence*, 2021.

D. Straumann and T. Mikosch. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, 34:2449 – 2495, 2006.

D. Tanny. A zero-one law for stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 30:139–148, 1974.

A. W. Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

W. Vervaat. On a stochastic difference equation and a representation of non–negative infinitely divisible random variables. *Advances in Applied Probability*, 11:750–783, 1979.

F. Wang and E. Ghysels. Econometric analysis of volatility component models. *Econometric Theory*, 31:362–393, 2015.

A. A. Weiss. ARMA models with ARCH errors. *Journal of time series analysis*, 5:129–143, 1984.

H. White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25, 1982.

K. Zhu and S. Ling. Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models. *The Annals of Statistics*, 39:2131 – 2163, 2011.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Equations récurrentes stochastiques : structure, inférence statistique et applications financières

Mots clés : Inférence sans moments, Systèmes stochastiques de fonctions itérées, Exposant de Lyapunov, GARCH à composantes multiplicatives, Bootstrap, Processus sous-additif

Résumé : Nous nous intéressons à l'étude des propriétés théoriques des équations récurrentes stochastiques (SRE) et de leurs applications en finance. Ces modèles sont couramment utilisés en économétrie, y compris en économétrie de la finance, pour styliser la dynamique d'une variété de processus tels que la volatilité des rendements financiers. Cependant, la structure de probabilité ainsi que les propriétés statistiques de ces modèles sont encore mal connues, particulièrement lorsque le modèle est considéré en dimension infinie ou lorsqu'il est généré par un processus non indépendant. Ces deux caractéristiques entraînent de formidables difficultés à l'étude théorique de ces modèles. Dans ces contextes, nous nous intéressons à l'existence de solutions stationnaires, ainsi qu'aux propriétés statistiques et probabilistes de ces solutions. Nous établissons de nouvelles propriétés sur la trajectoire de la solution stationnaire des SREs que nous exploitons dans l'étude des propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance (QMLE) des modèles de volatilité conditionnelle de type GARCH. En particulier, nous avons étudié la stationnarité et l'inférence statistique des modèles GARCH(p,q) semi-forts dans lesquels le processus d'innovation n'est pas nécessairement indépendant. Nous établissons la consistance du QMLE des GARCH (p,q) semi-forts sans hypothèses d'existence de moment, couramment supposée pour ces modèles, sur la distribution stationnaire. De même, nous nous sommes intéressés aux modèles GARCH à deux facteurs (GARCH-MIDAS), un facteur de volatilité à long terme et un autre à court terme. Ces récents modèles introduits par Engle et al. (2013) ont la particularité d'avoir des solutions stationnaires avec des distributions à queue épaisse. Ces modèles sont maintenant fréquemment utilisés en économétrie, cependant, leurs propriétés statistiques n'ont pas reçu beaucoup d'attention jusqu'à présent. Nous montrons la consistance et la normalité asymptotique du QMLE des modèles GARCH-MIDAS et nous proposons différentes procédures de test pour évaluer la présence de volatilité à long terme dans ces modèles. Nous illustrons nos résultats avec des simulations et des applications sur des données financières réelles.
Enfin, nous étendons le résultat de Kesten (1975) sur le taux de croissance des séquences additives aux processus superadditifs. Nous déduisons de ce résultat des généralisations de la propriété de contraction des matrices aléatoires aux produits d'opérateurs stochastiques. Nous utilisons ces résultats pour établir des conditions nécessaires et suffisantes d'existence de solutions stationnaires du modèle affine à coefficients positifs des SREs dans l'espace des fonctions continues. Cette classe de modèles regroupe la plupart des modèles de volatilité conditionnelle, y compris les GARCH fonctionnels.

Institut Polytechnique de Paris
91120 Palaiseau, France

Title : Stochastic recurrent equations: structure, statistical inference, and financial applications

Keywords : Inference Without Moments, Iterated Function Systems, Lyapunov exponent, Multiplicative Component GARCH, Residual Bootstrap, Subadditive sequence

Abstract : We are interested in the theoretical properties of Stochastic Recurrent Equations (SRE) and their applications in finance. These models are widely used in econometrics, including financial econometrics, to explain the dynamics of various processes such as the volatility of financial returns. However, the probability structure and statistical properties of these models are still not well understood, especially when the model is considered in infinite dimensions or driven by non-independent processes. These two features lead to significant difficulties in the theoretical study of these models. In this context, we aim to explore the existence of stationary solutions and the statistical and probabilistic properties of these solutions.

We establish new properties on the trajectory of the stationary solution of SREs, which we use to study the asymptotic properties of the quasi-maximum likelihood estimator (QMLE) of GARCH-type (generalized autoregressive conditional heteroskedasticity) conditional volatility models. In particular, we study the stationarity and statistical inference of semi-strong GARCH(p,q) models where the innovation process is not necessarily independent. We establish the consistency of the QMLE of semi-strong GARCHs without assuming the commonly used condition that the stationary distribution admits a small-order moment. In addition, we are interested in the two-factor volatility GARCH models (GARCH-MIDAS), a long-run, and a short-run volatility. These models were recently introduced by Engle et al. (2013) and have the particularity to admit stationary solutions with heavy-tailed distributions. These models are now widely used but their statistical properties have not received much attention. We show the consistency and asymptotic normality of the QMLE of the GARCH-MIDAS models and provide various test procedures to evaluate the presence of long-run volatility in these models. We also illustrate our results with simulations and applications to real financial data.

Finally, we extend a result of Kesten (1975) on the growth rate of additive sequences to superadditive processes. From this result, we derive generalizations of the contraction property of random matrices to products of stochastic operators. We use these results to establish necessary and sufficient conditions for the existence of stationary solutions of the affine case with positive coefficients of SREs in the space of continuous functions. This class of models includes most conditional volatility models, including functional GARCHs.

Institut Polytechnique de Paris
91120 Palaiseau, France