Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 24-2023

Limit Theory for Martingale Transforms with Heavy-Tailed Noise

Stelios Arvanitis and Alexandros Louka

December 2023

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

The Working Papers in this series circulate mainly for early presentation and discussion, as well as for the information of the Academic Community and all interested in our current research activity.

The authors assume full responsibility for the accuracy of their paper as well as for the opinions expressed therein.

Limit Theory for Martingale Transforms with Heavy-Tailed Noise

Stelios Arvanitis[†] and Alexandros Louka[**]

[†]Department of Economics, AUEB, Patision str. 76, Athens 10434, Greece

[**]Foundation for Economic and Industrial Research, Tsami Karatasou str. 11, Athens 117 42, Greece

Abstract

A limit theorem for partial sums of martingale transforms with multiplicative noise and ergodic transform processes is established, resulting to regularly varying rates and stable limits. Its' establishment is facilitated by the derivation of an extension to empirical distributions of Breiman's Theorem along with the Principle of Conditioning. The theorem is applied for the derivation of the limit theory OLSE in regressions with heavy tailed noise, as well as of the limit theory of the Gaussian QMLE in GARCH-type models. Depending on the index of stability, regularly varying rates and asymptotic stable distributions or inconsistency, are obtained.

MSC2020: 60, 62, 91.

Keywords: Martingale Limit Theorem; Principle of Conditioning; Domain of Attraction; α-Stable Distribution; Regular Variation; Breiman's Theorem; Heavy-tailed noise regression; OLSE; GARCH-type model; Gaussian QMLE; Inconsistency.

1 Introduction

Weak convergence theorems for partial sums of stationary weakly dependent processes with limiting stable laws exist in the literature; for two quite general formulations, see Davis and Hsing (1995) or the monograph of Kulik and Soulier (2020), for a point processes approach, and Bartkiewicz et al. (2011) for an approach based on characteristic functions. When the underlying process is a martingale transform, those results depend on conditions that may be either tedious to verify and/or more restrictive than necessary. In this paper weaker sufficient conditions are discussed for weak convergence to stable laws of partial sums of martingale transforms structured by multiplicative iid noise and ergodic transform processes. The Principle of Conditioning (see Jakubowski (1986)) is utilized, which allows for deriving limit theorems

*Corresponding author, email address: stelios@aueb.gr

for sums of dependent random variables from existing limit theory for independent processes combined with ergodic laws of large numbers.

Our motivation stems from the GARCH models literature (see for example the references in Ch. 3-7 of Straumann (2004)) that focuses on the determination of the asymptotic properties of the computationally convenient Gaussian Quasi Maximum Likelihood Estimator (QMLE). Under mild conditions, the relevant score process takes the form of a multiplicative martingale transform where the squared innovations act as noise. When the iid innovations possess fourth moments, the estimator is known to be $\sqrt{n}$-consistent and asymptotically normal (see Ch. 5 of Straumann (2004)). However, the empirically relevant possibility of the non-existence of the fourth moment (see e.g. Rachev and Mittnik (1988) and Mittnik et al. (1998)) raised the issue of its implications on the limit theory of the QMLE. Hall and Yao (2003) and Mikosch and Straumann (2006) derived martingale limit theorems that facilitated the derivation of the asymptotic properties of the estimator in this context.

We extend the results of Hall and Yao (2003), and Mikosch and Straumann (2006)) to a limit theorem for martingale transforms with regularly varying rates and stable limits under weaker conditions regarding temporal dependence and existence of moments. Our derivations exploit the multiplicative structure of the transform and the local representations of the characteristic function of the noise (see Ibragimov and Linnik (1971) and Aaronson and Denker (1998)). In the non-Gaussian domains of attraction, the derivations are facilitated by our asymptotic extension of the Denisov and Zwart (2007) version of Breiman's Theorem (see Breiman (1965)) to empirical distributions. Our approach does not require the verification of extremal index conditions (see Davis and Hsing (1995)) in order to establish asymptotic non-degeneracy. It also avoids the use of restrictions regarding mixing and/or the existence of higher order moments for the transform process, at the cost of restrictions on the regular variation properties of the martingale difference process.

The limiting distributions are stable with parameters that depend on the analogous parameters of the noise as well as on appropriate moments of the transform process. The results are readily extendable to multivariate transform processes via Cramer's Theorem. The case where the index of stability equals 2 yields Gaussian limits. The results thus incorporate asymptotic normality, with rates of convergence potentially slower than $\sqrt{n}$. This occurs when the truncated second moment of the noise slowly diverges to infinity, yet the second moments for the transform process exist. This is a restriction weaker than the condition implied in Hall and Yao (2003), while the conditional Lindberg condition (see Jeganathan (1982)) is also avoided; the latter would fail in this framework. Finally, our results constitute an essentially unified limit theory since they also incorporate the classical case involving stationary and ergodic square integrable transforms.

The remaining paper is organized as follows: in the following section, the needed notation and assumption framework are established, and the main result is derived and discussed. Section 3 discusses the assumptions, and derives auxiliary results. It contains the aforementioned extension of the Denisov and Zwart (2007) Breiman's Theorem to empirical distributions. In Section 4 we apply the theorem in order to derive the limiting properties of the OLSE in regressions with heavy-tailed errors. In

Section 5 we apply it to the Gaussian QMLE for GARCH-type models and discuss several examples. The final section contains the proofs.

2 Main result

We work in the context of a complete probability space $(\Omega, \mathcal{G}, \mathbb{P})$. The abbreviation $\mathbb{P}$ a.s. signifies almost sure events with respect to $\mathbb{P}$. We denote convergence in distribution of sequences of random elements with $\rightsquigarrow$ and equality in distribution with $\overset{d}{=}$. All limits are considered as $n \to +\infty$ unless otherwise specified. The stochastic processes under consideration are defined on $\mathbb{Z}$ or $\mathbb{N}$. The results are presented for the first case, yet they also hold for the second with appropriate modifications that involve initial values. $\mathcal{F}$ will denote some filtration on $\mathcal{G}$, i.e. an increasing double sequence $(\mathcal{G}_i)_{i\in\mathbb{Z}}$ of $\mathcal{G}$ sub-σ-algebras. Given a non empty set A, $\ell^\infty(A)$ denotes the set of bounded real functions on A equipped with the uniform metric. $\|\cdot\|$ denotes the Euclidean norm, and $\mathbf{i}$ denotes $\sqrt{-1}$.

We are interested in the asymptotic behavior of the partial sums of the process $((\xi_i - \gamma) V_i)_{i\in\mathbb{Z}}$ where $(\xi_i)_{i\in\mathbb{Z}}$ is an iid process, γ is a location real parameter, and $(V_i)_{i\in\mathbb{Z}}$ is a stationary ergodic process. We employ measurability properties of the constituent processes w.r.t. $\mathcal{F}$, that enable characterization of the pointwise product $((\xi_i - \gamma) V_i)_{i\in\mathbb{Z}}$ as a martingale transform of $(\xi_i)_{i\in\mathbb{Z}}$ by $(V_i)_{i\in\mathbb{Z}}$. The term is inaccurate whenever $\mathbb{E}[|\xi_0 V_0|] = +\infty$, yet we universally adopt it in the spirit of Mikosch and Straumann (2006).

The assumption framework that is introduced below deals with probabilistic and dynamical properties of the constituent processes. Its' formulation additionally uses the following notation:

$S_\alpha(\beta, c, \gamma)$ denotes the (univariate) stable distribution with parameters α, β, c, γ denoting stability, skewness, scale and location respectively (see Ibragimov and Linnik (1971)). When $\alpha = 2$, then the restriction $\beta = 0$ also holds, and $S_2(0, c, \gamma) = N(\gamma, c)$. For any $\alpha \in (0, 2]$ the distribution of ξ_0 belongs to the Domain of Attraction (hereafter DoA) of $S_\alpha(\beta, c, \gamma)$ if and only if its log-characteristic function has the following representation as $t \to 0$:

$$\begin{cases} \gamma \mathbf{i}t - c|t|^\alpha L(|t|^{-1}) \left(1 - \mathbf{i}\beta \operatorname{sgn}(t) \tan\left(\frac{1}{2}\pi\alpha\right)\right) + o\left(|t| L\left(|t|^{-1}\right)\right) & , \alpha \neq 1 \\ \left(\gamma + H(|t|^{-1})\right) \mathbf{i}t - c|t| L(|t|^{-1}) \left(1 - 2C\mathbf{i}\frac{\beta}{\pi}\operatorname{sgn}(t)\right) + o\left(|t| L\left(|t|^{-1}\right)\right) & , \alpha = 1 \end{cases}, \tag{1}$$

where $L(x)$ is a slowly varying function in the sense of Karamata (see for example Bingham et al. (1989)), $H(\lambda) = \int_0^\lambda \frac{x}{1+x^2} L(x)(2\beta c\pi^{-1} + k(x))dx$ and $k(x) \to 0$ as $x \to \infty$, and $-C$ is the Euler-Mascheroni constant. This is due to Theorem 2.6.5 in Ibragimov and Linnik (1971), with the case $\alpha = 1$ clarified in Aaronson and Denker (1998) (see Theorems 1 and 2 there).

Also, $c(x)$ is a function converging to a strictly positive constant, while U and $U^\star$ are independent long-tailed random variables with hazard rates converging to 0. A random variable U is long tailed iff $\mathbb{P}[U > x] \sim \mathbb{P}[U > x + y]$ as x tends to ∞ for any y, which then implies that $\mathbb{P}[U > \log x]$ is slowly varying. Long tailed random variables appear in a convenient characterization of slowly varying functions at infinity in Lemma 2.1 of Denisov and Zwart (2007).

Assumption 1. *For some $\alpha \in (0,2]$, $\beta \in [-1,1]$, $c > 0$ and $\gamma \in \mathbb{R}$, $(\xi_i)_{i\in\mathbb{Z}}$ is iid and the distribution of ξ_0 lies in the domain of attraction (DoA) of $S_\alpha(\beta, c, \gamma)$.*

Assumption 2. *For some filtration $\mathcal{F} \equiv (\mathcal{G}_i)_{i\in\mathbb{Z}}$, $(\xi_i V_i)_{i\in\mathbb{Z}}$ is $\mathcal{F}$-adapted, ξ_i is independent of $\mathcal{G}_{i-1}$ and V_i is $\mathcal{G}_{i-1}$-measurable for all $i \in \mathbb{Z}$.*

Assumption 3. *$(V_i)_{i\in\mathbb{Z}}$ is stationary and ergodic with $\mathbb{E}[|V_0|^\alpha] < \infty$ and $\mathbb{P}[|V_0| > x] = o(\mathbb{P}[|\xi_0| > x])$ as $x \to +\infty$.*

Assumption 4. *$\alpha < 2$ and either one of the following conditions hold:*

1. $\limsup_{x\to+\infty} \sup_{1\leq y\leq x} L(y)/L(x) < \infty$, *or*
2. *for some $\delta > 0$, $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right] < \infty$.*

Assumption 5. *$\alpha < 2$, $(V_i)_{i\in\mathbb{Z}}$ is strongly mixing and either one of the following conditions hold:*

1. *L is of the form $L(x) = c(x)\mathbb{P}(U > \log x)$, or of the form $L(x) = c(x)\mathbb{P}[U > \log x]/\mathbb{P}[U^\star > \log x]$. Moreover,*
$$\lim_{x\to+\infty} \int_0^x \frac{\mathbb{P}[U > x-y]}{\mathbb{P}[U > x]} \mathbb{P}[U > y]\, dy = 2\int_0^\infty \mathbb{P}[U > y]\, dy < +\infty,$$
and, $\frac{x^\alpha \mathbb{P}[|V_0|>x]}{\mathbb{P}[U>\log x]} \to 0$ as $x \to +\infty$, or,
2. $\limsup_{x\to+\infty} \sup_{\sqrt{x}\leq y\leq x} \frac{L(y)}{L(x)} < +\infty$ *and,*
$$\frac{\mathbb{P}[|V_0| > x]}{\mathbb{P}[|\xi_0| > x]} \int_0^x t^\alpha d\mathbb{P}[|\xi_0| \leq t] \to 0,$$
as $x \to +\infty$.

A discussion on the plausibility of the assumption framework is provided in the following section. We define r_n by the asymptotic relation $\frac{L(n^{\frac{1}{\alpha}} r_n^{\frac{1}{\alpha}})}{r_n} \to 1$; this constitutes a well defined slowly varying sequence-see the proof of Theorems 2.6.1-2, result (2.2.18) of Ibragimov and Linnik (1971), and Remark 2 of Aaronson and Denker (1998).

Our main result about the asymptotic behavior of $\sum_{i=1}^n (\xi_i - \gamma) V_i$ follows:

Theorem 1. *Suppose that Assumptions 1-3 hold. Furthermore, if $\alpha < 2$ suppose that either Assumption 4 or Assumption 5 also hold. Then, if $\alpha \neq 1$,*
$$\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n (\xi_i - \gamma) V_i \rightsquigarrow S_\alpha\left(\beta \frac{\mathbb{E}[|V_0|^\alpha \operatorname{sgn}(V_0)]}{\mathbb{E}[|V_0|^\alpha]}, c\mathbb{E}[|V_0|^\alpha], 0\right). \tag{2}$$
If $\alpha = 1$ and
$$\mathbb{E}[|V_0| |\log(|V_0|)|] < \infty, \text{ when } \beta \neq 0$$

then,

$$\frac{1}{nr_n}\sum_{i=1}^{n}\left(\xi_i-\gamma-H\left(nr_n\right)\right)V_i-2\beta c\pi^{-1}\left(C\mathbb{E}\left(V_0\right)-\mathbb{E}\left[V_0|\log(|V_0|)|\right]\right)$$
$$\rightsquigarrow S_1\left(\beta\frac{\mathbb{E}\left[V_0\right]}{\mathbb{E}\left[|V_0|\right]},c\mathbb{E}\left[|V_0|\right],0\right). \tag{3}$$

If $\alpha<1$ and either for any $M>0$,

$$\mathbb{P}\left(\max_{1\le i\le n}|V_i|>Mq_n^{\frac{1}{\alpha}}n^{\frac{1}{\alpha}}\right)\to 0, \tag{4}$$

where $q_n=O(r_n^{1/(1-\alpha)})$, or for some $\delta>0$,

$$\mathbb{E}\left[|V_0|^{\alpha+\delta}\right]<\infty, \tag{5}$$

then,

$$\frac{1}{n^{1/\alpha}r_n^{1/\alpha}}\sum_{i=1}^{n}\xi_iV_i\rightsquigarrow S_\alpha\left(\beta\frac{\mathbb{E}\left[|V_0|^\alpha\operatorname{sgn}\left(V_0\right)\right]}{\mathbb{E}\left[|V_0|^\alpha\right]},c\mathbb{E}\left[|V_0|^\alpha\right],0\right). \tag{6}$$

The convergence rate $n^{1/\alpha}r_n^{1/\alpha}$ of the sum reflects tail variation patterns of the distribution of ξ_0 and is not affected by properties of the $(V_i)_{i\in\mathbb{Z}}$ process. The limiting distribution is stable with stability parameter strictly determined by Assumption 1. The distribution of V_0 affects only the scale and symmetry parameters of the limit when $\alpha\neq1$. It also affects the form of the translating constants in the case where $\alpha=1$; the result is obtained without imposing conditions that nullify the translating sequence-compare with Theorem 1, Condition 5 of Bartkiewicz et al. (2011). When the attractor is symmetric, V_0 affects only the limiting scale, in which case $(V_i)_{i\in\mathbb{Z}}$ can be characterized as a stochastic scaling process for the transform. In this case and for $\alpha=1$, the term $2\beta c\pi^{-1}\left(C\mathbb{E}\left(V_0\right)-\mathbb{E}\left[V_0|\log(|V_0|)|\right]\right)$ disappears from the centering sequence, thus (3) coincides with (2) and the limit becomes a Cauchy distribution. When $\alpha<1$, and due to the zero location, when $c_2=0$ (resp. $c_1=0$), the limiting distribution is supported on $[0,+\infty)$ (resp. $(-\infty,0]$). Furthermore, when $\alpha<1$ and either (4)-that strengthens (8), or the stricter (5) hold, then the translating sequence $\frac{\gamma}{n^{1/\alpha}r_n^{1/\alpha}}\sum_{i=1}^{n}V_i$ becomes asymptotically negligible. Finally, when $V_i=1$, $\mathbb{P}$ a.s. for all i, then the classical iid results are recovered for all α (see Remark 2 of Aaronson and Denker (1998)).

When $\alpha=2$ and r_n converges (necessarily to $\mathbb{E}\left[\xi_0^2\right]$), a version of the classical CLT for stationary and ergodic martingale difference sequences is obtained with distributional limit the $N\left(0,\mathbb{E}\left[V_0^2\right]\right)$ distribution. When r_n diverges a CLT is still obtained, albeit with slower rates of the form $\sqrt{\frac{n}{r_n}}$ and the same limiting distribution. The rate reflects the divergence properties of the truncated second moment of ξ_0. Due to the monotonicity of the mapping $x\to\mathbb{E}\left[\xi_0^2 1_{|\xi_0|\le x}\right]$, this case is derivable even if $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right]=+\infty$ for every $\delta>0$, and/or without any mixing conditions for $(V_i)_{i\in\mathbb{Z}}$. Furthermore, then, the conditional Lindberg condition requiring that $\sum_{i=1}^{n}\mathbb{E}\left[\frac{1}{nr_n}\left(\xi_i-\gamma\right)^2V_i^2 1_{|(\xi_i-\gamma)V_i|>M\sqrt{nr_n}}/\mathcal{G}_{i-1}\right]\to0$ in probability for any $M > 0$, fails. Jeganathan (1982) considers it as in some sense necessary condition for the martingale convergence, hence such cases constitute counter-examples. For a simple example suppose that $\xi_0 \sim t_2$. Then the result assumes the form $\frac{1}{\sqrt{n \log n}} \sum_{i=1}^{n} \xi_i V_i \rightsquigarrow N\left(0, \mathbb{E}\left[V_0^2\right]\right)$ by a simple calculation. This essentially generalizes the results of Abadir and Magnus (2004) to dependent processes.

Theorem 1 can be readily extended when V_i is an $\mathbb{R}^d$-valued random vector via the use of the Cramer's Theorem. Suppose that for any $\lambda \in \mathbb{R}^d$ different than zero, when $\alpha \neq 1$, we have that $\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^{n} (\xi_i - \gamma) \lambda^T V_i$ converges in distribution to $S_\alpha \left(\beta \frac{\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha \operatorname{sgn}\left(\lambda^T V_0\right)\right]}{\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha\right]}, c\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha\right], 0 \right)$, while when $\alpha = 1$ we obtain the same result by re-centering with $2\beta c \pi^{-1}(C\mathbb{E}\left[\lambda^T V_0\right] - \mathbb{E}\left[\lambda^T V_0 | \log(|\lambda^T V_0|)|\right])$. Then, when $\alpha \geq 1$, Example 3.3.4 of Samorodnitsky and Taqqu (1994) implies that the multivariate limits are identified as multivariate α-stable distributions with spectral measures Γ determined by

$$\beta \frac{\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha \operatorname{sgn}\left(\lambda^T V_0\right)\right]}{\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha\right]} = \frac{\int_{\mathbb{S}^{d-1}} \left|s^T \lambda\right|^\alpha \operatorname{sgn}(s^T \lambda) \Gamma(ds)}{\int_{\mathbb{S}^{d-1}} \left|s^T \lambda\right|^\alpha \Gamma(ds)}, \; c\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha\right](\lambda) = \int_{\mathbb{S}^{d-1}} \left|s^T \lambda\right|^\alpha \Gamma(ds) \;, \tag{7}$$

where $\mathbb{S}^{d-1}$ denotes the $d-1$ dimensional sphere. Theorem 2.3 of Gupta et al. (1994) and the zero location in (2) implies that the same is true for $\alpha < 1$. However, our methodology cannot accommodate the case where ξ_0 is an $\mathbb{R}^d$-valued random vector since this would require a non trivial extension of results similar to Proposition 1 (see the following section) to spectral measures. Such a consideration is delegated to future research.

The multivariate analogue of Theorem 1 extends the results of Mikosch and Straumann (2006), and partly the ones in the GARCH-type, model specific framework of Hall and Yao (2003). This is due to that: i) it avoids simultaneously requiring the existence of $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right]$ for some $\delta > 0$ and mixing conditions for the $(V_i)_{i \in \mathbb{Z}}$ process. A fortiori, in the case corresponding to Assumption 4.1, neither mixing nor the existence of $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right]$ is required. ii) It avoids conditions requiring strict positivity for the extremal index of the process $((\xi_i - \gamma) V_i)_{i \in \mathbb{Z}}$. This index reflects information on the clustering of the process above large thresholds-see for example Davis and Hsing (1995), and its evaluation need not be trivial in applications (see the monograph of Kulik and Soulier (2020) for issues of statistical inference about the index). Its' positivity ensures that the limiting distribution is not degenerate at zero. In the present framework this follows simply when $\mathbb{E}\left[|V_0|^\alpha\right]$ is strictly positive (or more generally, in the multivariate case when $\mathbb{E}\left[\left|\lambda^T V_0\right|^\alpha\right]$ is strictly positive for some allowable λ). iii) It is more informative on the characterization of the limiting distributions. iv) It generally allows for $\alpha \leq 1$. Theorem 1 extends the results in Surgailis (2008) since it allows for $\mathbb{E}\left[|V_0|\right]^{\alpha+\delta} = +\infty$ for any $\delta > 0$, $\alpha \leq 1$ and does not require normal DoAs-consider for example Assumption 4.1 for L monotonically diverging. Due to that, it analogously extends the results of Jakubowski (2012).

Theorem 1 can also be useful for processes that admit decompositions involving multiplicative martingale transforms. Suppose for example that $(X_i, \mathcal{G}_i)_{i \in \mathbb{Z}}$ is a stationary L_1-mixingale of size -1 (see Ch. 16 of Davidson (1994)). By the proof of Theorem 5.4 of Hall and Heyde (2014), it admits the decomposition $X_i = Y_i + Z_i - Z_{i-1}$, where $(Y_i, \mathcal{G}_i)_{i\in\mathbb{Z}}$ is a martingale difference sequence and $(Z_i)_{i\in\mathbb{Z}}$ is stationary with $\mathbb{E}[|Z_0|] < +\infty$. Suppose that for all $i \in \mathbb{Z}$, Y_i can be factored as $(\xi_i - \gamma) V_i$, for which Assumptions 4 or 5 are valid for some $\alpha > 1$. Then, $\max_{i\le n} \mathbb{P}\left[\frac{|Z_i|}{n^{1/\alpha} r_n^{1/\alpha}} > M\right] \to 0$ for all $M > 0$, and thereby due to Theorem 1, $\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n X_i \rightsquigarrow S_\alpha\left(\beta\frac{\mathbb{E}[|V_0|^\alpha \operatorname{sgn}(V_0)]}{\mathbb{E}[|V_0|^\alpha]}, c\mathbb{E}\left[|V_0|^\alpha\right], 0\right)$.

In the case where V_i is $\mathbb{P}$ a.s. positive for all i and the transform process is independent of the noise, $(\xi_i V_i)_{i\in\mathbb{Z}}$ can be characterized as a stochastic volatility process (see Andersen et al. (2009)). Then Theorem 1 extends the results of Proposition 2 of Bartkiewicz et al. (2011) without assuming that $(\ln V_i)_{i\in\mathbb{Z}}$ is a Gaussian ARMA process, or that, when $\alpha = 1$, ξ_0 is symmetric, while it allows for Gaussian limits even if the martingale difference process has infinite variance.

Theorem 1 is inapplicable in cases where $\prod_{i=1}^n \mathbb{E}\left[\exp\left(itn^{-\frac{1}{\alpha}} r_n^{-\frac{1}{\alpha}}(\xi_i - \gamma V_i)\right)/\mathcal{G}_i\right]$ can only converge weakly to some non-degenerate limit. Such a result would generalize the results in Wang (2014), and this could in turn be useful in several applications including non linear co-integration (see for example Wang (2014)), or the limit theory of the OLSE in frameworks of moderate deviations from a unit root (see for example Phillips and Magdalinos (2007)). We delegate this investigation to further research.

Finally, suppose that (Θ, d) is a totally bounded metric space, $(V_i(\theta))_{i\in\mathbb{Z}}$ is stationary ergodic and $\mathbb{E}\left[|V_0|^{\alpha+\delta}(\theta)\right] < +\infty$ for all $\theta \in \Theta$ and some $\delta > 0$, $|V_i(\theta) - V_i(\theta^\star)| \le v_i d(\theta, \theta^\star)$, for all $i \in \mathbb{Z}$ and θ, $\theta^\star \in \Theta$, where $(v_i)_{i\in\mathbb{Z}}$ is stationary ergodic, and $\mathbb{E}\left[v_0^{\alpha+\delta}\right]$ exists. Suppose also that $\gamma = 0$ if $\alpha \ge 1$, and $\beta = 0$ if $\alpha = 1$. Then, for any $\varepsilon, \eta > 0$ there exists $\delta > 0$ small enough, such that

$$\mathbb{P}\left[\sup_{\theta,\theta^\star\in\Theta, d(\theta,\theta^\star)<\delta} \frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n \xi_i \left|V_i(\theta) - V_i(\theta^\star)\right| > \varepsilon\right]$$
$$\le \mathbb{P}\left[\sup_{\theta,\theta^\star\in\Theta, d(\theta,\theta^\star)<\delta} \frac{d(\theta,\theta^\star)}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n \xi_i v_i > \varepsilon\right] \le \mathbb{P}\left[O_p(1) > \frac{\varepsilon}{\delta}\right] \le \eta,$$

where the $O_p(1)$ term in the previous display is obtained by the application of Theorem 1 (in the case where $\alpha < 1$ then (6) holds) to $\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n \xi_i v_i$. (2) then implies stochastic equicontinuity and therefore along with the application of Theorem 1 to the fidis of $\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n \xi_i V_i(\theta)$, the weak convergence of the latter to a stochastic process with α-stable marginals in $\ell^\infty[\Theta]$. In the special case where $\alpha < 1$ and $\min\{\xi_0, V_0\} \ge 0$ $\mathbb{P}$ a.s., the limiting process is equal in distribution to $Z\mathbb{E}[V_0^\alpha(\theta)]$, where $Z \sim S_\alpha(1, c, 0)$ for some $c > 0$. In that case, finally suppose that $\arg\max_{\theta\in\Theta} \mathbb{E}[V_0^\alpha(\theta)] = \{\theta_0\} \subset \Theta$. Then the CMT implies that $\arg\max_\Theta \frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^n \xi_i V_i(\theta) \rightsquigarrow \{\theta_0\}$ since Z has positive support. This could be useful for establishing consistency of M-estimators in very heavy tailed frameworks even when the appropriately scaled objective functions have stochastic limits, as long as appropriate conditions for parameter identification hold.

3 Discussion and auxiliary results

A discussion on the assumptions that appeared in the previous section is presented here. Some auxiliary results to the proof of Theorem 1 are also derived.

The motivation behind Assumption 1 is the issue of the fourth moment existence for the innovations of GARCH-type models in the framework of empirical finance. In this respect $(\xi_i)_{i\in\mathbb{Z}}$ represents the squared innovations-potentially translated by $-\gamma = -1$-as those appear in parts of the Gaussian quasi log-likelihood function and its derivatives. It encompasses the usual case where $\alpha = 2$ and $\mathbb{E}[\xi_0^2] < +\infty$. It also allows for cases where $\mathbb{E}[\xi_0^2] = +\infty$ and $\mathbb{E}\left[\xi_0^2 1_{|\xi_0|\leq x}\right]$ is either regularly varying at infinity with index $1 - \frac{\alpha}{2}$ or it is slowly varying. Then, the usual ergodic square integrable martingale difference CLT is inapplicable to $\sum_{i=1}^n (\xi_i - \gamma) V_i$.

In some applications the filtration $\mathcal{F}$ appearing in Assumption 2, represents the history of the $(\xi_i)_{i\in\mathbb{Z}}$ process, i.e. it is defined by $\mathcal{G}_i = \sigma(\xi_{i-j}, j \geq 0)$. Then, independence of ξ_i from $\mathcal{G}_{i-1}$ follows readily from Assumption 1. The $\mathcal{G}_{i-1}$-measurability of V_i could follow when the random element is defined as some measurable transformation of causal solutions to stochastic recurrence equations (SRE) measurable w.r.t. $\mathcal{G}_{i-1}$. In any case Assumption 2 along with the Principle of Conditioning (see Jakubowski (1986)) implies that the limiting distribution of $\sigma_n^{-1}\left(\sum_{i=1}^n (\xi_i - \gamma) V_i - \tau_n\right)$ for the appropriate choice of (τ_n, σ_n), if any, will be determined by the limiting behavior of $\exp\left(\frac{\tau_n}{\sigma_n}\right)\prod_{i=1}^n \mathbb{E}\left[\exp\left(\mathbf{i}t\sigma_n^{-1}(\xi_i - \gamma) V_i\right)/\mathcal{G}_i\right]$, $t \neq 0$.

In Assumption 3, stationarity and ergodicity for the $(V_i)_{i\in\mathbb{Z}}$ process, whenever this is defined as some measurable transformation of solutions of stationary and ergodic SREs, can be established via conditions on the Liapunov exponents of the relevant dynamical systems; see for example Bougerol and Picard (1992). The existence of the α moment for the stationary distribution can in a similar set up be derivable via results like the ones in Goldie (1991). The final part of the assumption employs a comparison between the tails of the stationary distributions of ξ_0 and V_0. In view of Assumption 1, it certainly holds whenever $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right] < +\infty$, but this is not necessary. Regular variation of the tails of V_0 with index α would suffice as long as the slowly varying part is asymptotically dominated by L. This allows for cases where $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right] = +\infty$ for any $\delta > 0$ even when $\mathbb{E}[|\xi_0|^\alpha] < +\infty$. The particular assumption controls the behavior of the maximum order statistic for the $(|V_i|)_{i\in\mathbb{Z}}$ process; the first auxiliary result below shows that the latter cannot diverge at a rate faster than $n^{\frac{1}{\alpha}} r_n^{\frac{1}{\alpha}}$ with probability converging to one:

Lemma 1. *Suppose that Assumptions 1-3 hold. Then for any $M > 0$,*

$$\mathbb{P}\left[n^{-\frac{1}{\alpha}} r_n^{-\frac{1}{\alpha}} \max_{1\leq i\leq n} |V_i| > M\right] \to 0. \tag{8}$$

Assumption 1 also implies that for any $M > 0$,

$$\mathbb{P}\left[n^{-\frac{1}{\alpha}} r_n^{-\frac{1}{\alpha}} \max_{1\leq i\leq n} |\xi_i| > M\right] \to \Phi_\alpha(M),$$

where Φ_α denotes the Frechet distribution; see Embrechts et al. (2013). This and Lemma 1 partially imply that the index of the tail variation properties for the limiting distribution of the scaled partial sum of the transform is essentially determined by the distribution of ξ_0.

Lemma 1 also implies that $\mathbb{E}\left[\exp\left(\mathbf{i}tn^{-\frac{1}{\alpha}}r_n^{-\frac{1}{\alpha}}(\xi_i-\gamma)V_i\right)/\mathcal{G}_i\right]$ can be approximated with high probability by the local representation

$$-\frac{c|t|^\alpha}{nr_n}|V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|tV_i|^{-1}\right)+1\{\alpha=1\}H(n^{1/\alpha}r_n^{1/\alpha}|tV_i|^{-1})\mathbf{i}t\frac{V_i}{n^{1/\alpha}r_n^{1/\alpha}}$$
$$\frac{|t|^\alpha}{nr_n}\mathbf{i}\beta c\operatorname{sgn}(t)\tan\left(\tfrac{1}{2}\pi\alpha\right)\operatorname{sgn}(V_i)|V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|tV_i|^{-1}\right).$$

This is due to Assumption 1 and the fact that $tn^{-\frac{1}{\alpha}}r_n^{-\frac{1}{\alpha}}\max_{1\leq i\leq n}|V_i|$ lies in decreasing neighborhoods of validity of the representation in 1 with probability converging to one, for any t by Lemma 1. A simple calculation (see the proof of Theorem 1) then shows that the derivation of the main result would be greatly facilitated by the determination of the asymptotic behavior of terms similar to $\frac{1}{nr_n}\sum_{i=1}^n|V_i|^\alpha L\left(\frac{n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}}{(|t||V_i|)}\right)$. Given Assumption 3, and Birkhoff's LLN, if some Tauberian-type result would allow for the asymptotic factoring out from the sum of the slowly varying component, then the required limits would be easily derivable from ergodicity.

To this end, when $\alpha<2$, due to Assumptions 1-3, $\frac{1}{nr_n}\sum_{i=1}^n|V_i|^\alpha L\left(\frac{n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}}{(|t||V_i|)}\right)$ is $\mathbb{P}$ a.s. asymptotically equivalent to $\frac{\mathbb{P}_n^\star\left[v_n|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}{\mathbb{P}\left[|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}$, where $\mathbb{P}_n^\star$ denotes the stochastic product measure between the empirical distribution of $(|V_i|)_{1\leq i\leq n}$ and $\mathbb{P}$, and v_n is a random variable that follows this empirical distribution and is independent of ξ_0. This ratio is partially an empirical analogue of the random variables' product that appears Breiman's Theorem (see Breiman (1965)). This in our setting states that if, in addition to Assumption 3, $\mathbb{E}\left[|V_0|^{\alpha+\delta}\right]<+\infty$ for some $\delta>0$, then as $x\to+\infty$, $\frac{\mathbb{P}_n^\star\left[v_n|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}{\mathbb{P}\left[|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}\to\mathbb{E}\left[|V_0|^\alpha\right]$. Denisov and Zwart (2007) extend Breiman's Theorem by essentially assuming the moment existence part of Assumption (3), while imposing further conditions on the properties of L and/or the comparison between the tails of the implicated distributions. The question that arises is whether the extension of Denisov and Zwart (2007), can be modified so as to hold for the empirical ratio $\frac{\mathbb{P}_n^\star\left[v_n|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}{\mathbb{P}\left[|\xi_0|>\frac{1}{|t|}n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}\right]}$. Sufficient conditions for this modification are essentially provided by the final pair of assumptions.

The first part of Assumption 4, covers normal DoAs or eventually monotonically diverging slow variation for ξ_0 (for sufficient conditions on the existence of monotone versions of L see Buldygin et al. (2013)). The second part requires the existence of $\alpha+\delta$ moment for $|V_0|$. It corresponds to the empirical distribution extension of the original result of Breiman (1965).

In Assumption 5.1, the integrability requirement for $\mathbb{P}[U>x]$ is equivalent to that the tail distribution function is a sub-exponential density, or equivalently that the random variable U belongs to the class S^* (see Klüppelberg (1988)). In the case where $L(x) = c(x)\mathbb{P}[U > \log x]$, this is equivalent to $\int_0^x \frac{L(e^{x-y})}{L(e^x)} L(e^y)\, dy = 2\int_0^\infty L(e^y)\, dy < +\infty$ (see Denisov and Zwart (2007)).

Assumption 5.2 essentially handles cases where $\int_0^\infty \mathbb{P}[U > y]\, dy$ diverges. A sufficient condition for the required asymptotic boundedness, is that $L \circ \exp$ is of bounded variation; the functionality of the assumption is retained if $\sqrt{x}$ is replaced with x^β for any $\beta \in (0,1)$ (see again Denisov and Zwart (2007)). The remaining asymptotic negligibility requirements in Assumption 5 are exactly the ones in Propositions 2.2-3 of Denisov and Zwart (2007). They constitute refined comparisons between the tails of the stationary distributions of ξ_0 and V_0. They, along with strong mixing for $(V_i)_{i\in\mathbb{Z}}$, allow for the asymptotic control of a scaled approximation error of the empirical distribution of $(|V_i|)_{1\leq i\leq n}$ to the stationary distribution of $|V_0|$, uniformly over appropriate classes of events. Strong mixing can be substituted by a more general mixingale restriction (see Theorem 1 of Hill (2010) and the derivation of (16) in the proof of the proposition) without affecting the validity of the results.

Assumptions 1-5 enable our second auxiliary result when $\alpha < 2$. It is to the best of our knowledge new, and it provides the useful extension of the Theorem of Breiman (see Breiman (1965)) and of the corresponding extension of Denisov and Zwart (2007), to the aforementioned ratio that involves the empirical $\mathbb{P}_n^\star$:

Proposition 1. *Suppose that Assumptions 1-3 hold for $\alpha < 2$. If Assumption 4 holds, then, for any real sequence $m_n \to +\infty$, as $n \to +\infty$,*

$$\frac{\mathbb{P}_n^\star [v_n |\xi_0| > m_n]}{\mathbb{P}[|\xi_0| > m_n]} \to \mathbb{E}[|V_0|^\alpha], \mathbb{P}\ a.s. \tag{9}$$

where $\mathbb{P}_n^\star$ denotes the product measure between the empirical distribution of $(|V_i|)_{1\leq i\leq n}$ and $\mathbb{P}$, and v_n is a random variable that follows this empirical distribution and is independent of ξ_0.

If Assumption 5 holds, then for any real sequence $m_n \to +\infty$, as $n \to +\infty$,

$$\frac{\mathbb{P}_n^\star [v_n |\xi_0| > m_n]}{\mathbb{P}[|\xi_0| > m_n]} \to \mathbb{E}[|V_0|^\alpha],\ \textit{in probability.} \tag{10}$$

The result (9) implies the strong array LLN; $\frac{1}{nr_n}\sum_{i=1}^n |V_i|^\alpha L\left(\frac{n^{\frac{1}{\alpha}} r_n^{\frac{1}{\alpha}}}{(|t||V_i|)}\right) \to \mathbb{E}[|V_0|^\alpha]$ $\mathbb{P}$ a.s. without any further dependence restrictions for $(V_i)_{i\in\mathbb{Z}}$. Analogously, the result (10) obtains a weak array LLN; $\frac{1}{nr_n}\sum_{i=1}^n |V_i|^\alpha L\left(\frac{n^{\frac{1}{\alpha}} r_n^{\frac{1}{\alpha}}}{(|t||V_i|)}\right) \to \mathbb{E}[|V_0|^\alpha]$ in probability.

The asymptotic representation of $\frac{1}{nr_n}\sum_{i=1}^n |V_i|^\alpha L\left(\frac{n^{\frac{1}{\alpha}} r_n^{\frac{1}{\alpha}}}{(|t||V_i|)}\right)$ by the aforementioned probability ratio breaks down when $\alpha = 2$. In this case a simple calculation shows that the sum equals $\sum_{i=1}^n \mathbb{E}\left[(\xi_i - \gamma)^2 V_i^2 1_{|(\xi_i-\gamma)V_i| > \frac{1}{|t|}} / \mathcal{G}_{i-1}\right]$. Our final auxiliary result handles the asymptotic properties of this representation:

Proposition 2. *Suppose that Assumptions 1-3 hold for $\alpha = 2$. Then for any $M > 0$,*

$$\sum_{i=1}^n \mathbb{E}\left[\frac{1}{nr_n}(\xi_i - \gamma)^2 V_i^2 1_{|(\xi_i-\gamma)V_i| \leq M\sqrt{nr_n}} / \mathcal{G}_{i-1}\right] \to \mathbb{E}[|V_0|^\alpha], \mathbb{P}\ a.s. \tag{11}$$

The proof closely parallels the proof of the first part of Proposition 1 due to the fact that when $\alpha = 2$ then L is necessarily monotonic. When r_n converges, then (11) implies the conditional Lindberg condition typically encountered in several martingale limit theorems (see for example Hall and Heyde (2014); Jeganathan (1982)).

4 Application: regressions with heavy-tailed errors

The application of Theorem 1 to the derivation of the limiting behavior of the OLSE in a regression with heavy-tailed errors is investigated. Specifically, the regression model at hand is of the form:

$$y_i = \mathbf{x}_i\beta + \varepsilon_i,\ i \in \mathbb{Z}, \tag{12}$$

where $(y_i, \mathbf{x}_i)_{i\in\mathbb{Z}}$ is an $1 \times (d+1)$ dimensional observable stochastic process, $(\varepsilon_i)_{i\in\mathbb{Z}}$ is an one-dimensional latent heavy tailed error process, and the parameter $\beta \in \mathbb{R}^d$ is unknown. Given a sample $(y, \mathbf{x}) := (y_i, \mathbf{x}_i)_{i=1,\dots,n}$, and if the stochastic matrix $\mathbf{x}$ has rank equal to p, the OLSE for β is $\beta_n = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T y$, which, given (12), obtains also the latent form $\beta + (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\varepsilon$, where $\varepsilon := (\varepsilon_i)_{i=1,\dots,n}$. We are interested in the limiting behavior of β_n, under the following conditions:

- B.1 The error process $(\varepsilon_i)_{i\in\mathbb{Z}}$ has the form $\varepsilon_i := \xi_i u_i$, $\forall i \in \mathbb{Z}$, where $(\xi_i)_{i\in\mathbb{Z}}$ is iid, and $((u_i, \mathbf{x}_i)_{i\in\mathbb{Z}}$ is adapted to some filtration $\mathcal{F} := (\mathcal{G}_i)_{i\in\mathbb{Z}}$. ξ_i is independent of $\mathcal{G}_i$ for all $i \in \mathbb{Z}$.
- B.2 For some $\alpha \in (0, 2)$, $c > 0$, $\xi_0 \sim S_\alpha(0, c, 0)$.
- B.3 $\mathbb{P}[u_0 > 0] = 1$, and (i) there exists some $\eta > 2$, for which as $x \to +\infty$, $\mathbb{P}[u_0] = O(x^{-\alpha}\ln^{-\eta}(x))$, or (ii) for some $0 < \delta^\star < 2 - \alpha$, and some slowly varying function at infinity $\ell^\star$, $\mathbb{P}[u_0] = O(x^{-\alpha-\delta^\star}\ell^\star(x))$ as $x \to +\infty$.
- B.4 $(u_i)_{i\in\mathbb{Z}}$ is stationary and strongly mixing.
- B.5 For all $i \in \mathbb{Z}$, $\mathbf{x}_i$ admits the causal representation $\mathbf{x}_i = \mathbf{c} + \sum_{k=0}^{\infty} \Psi_k \epsilon_{i-k}$, where $\mathbf{c} \in \mathbb{R}^d$, and $(\epsilon_i)_{i\in\mathbb{Z}}$ is a stationary and strong mixing vector white noise process with $\mathbb{E}[\epsilon_0\epsilon_0^T] = \mathrm{Id}_p$. Furthermore, the matrix $\sum_{k=0}^{\infty} \Psi_k\Psi_k^T$ is well defined and has rank equal to d. u_i is independent of ϵ_j, for all $i, j \in \mathbb{Z}$.
- B.6 Either (i) $\max_{i,j=1,\dots,d} |\Psi_k(i,j)| + \min_{i,j=1,\dots,d} |\Psi_k(i,j)| = O(k^{-1})$, or (ii), there exists a $\theta \in (0, 1)$, and a $C > 0$ such that for large enough k, $\max_{i,j=1,\dots,d} |\Psi_k(i,j)| \le C\theta^k$, and the distribution of ϵ_0 has a differentiable density.

B.1 posits the multiplicative martingale transform structure for the error process. u_i can be perceived as a scale, that-given the independence of the noise-produces temporal dependence. The information structure considered via the employed filtration enables the use of the principle of conditioning. $\mathcal{G}_{i-1}$ may be formed by the noise up to $i - 1$, as well as by the regressors and the stochastic constituents of the scale up to i.

B.2 states that the noise follows a symmetric, zero location stable distribution. It can be easily extended to asymmetric distributions inside the respective domain of attraction. B.3 and B.4 take care of the marginal and temporal characteristics of the scale process. The scale is strictly positive with a marginal distribution with regularly varying tail of index α in case (i), or $\alpha+\delta$ in case (ii). The power logarithmic part of the tail in case (i) implies that $\mathbb{E}[u_0^{\alpha}] < +\infty$, and $\mathbb{E}[u_0^{\alpha+\delta}] = +\infty$ for all $\delta > 0$. In case (ii) $\mathbb{E}[u_0^{\alpha+\delta}] < +\infty$ for all $0 < \delta < \delta^{\star}$. In both cases $\mathbb{P}[u_0 > x] = o(\mathbb{P}[\xi_0 > x])$, as $x \to +\infty$. B.4 implies ergodicity of the scale process. It allows for example for linear processes of the form $u_i = c_u + \sum_{k=0}^{\infty} \beta_u^k \zeta_{i-k}$ with $c_u > 0$, $0 \leq \beta_u < 1$, $(\zeta_i)_{i\in\mathbb{Z}}$ is an iid sequence of positive random variables, and such that the distribution of ζ_0 has a differentiable density-see Theorem 2.1 of Chanda (1974).

B.5 states that the regressors' process is linear over a stationary and strong mixing unit variance vector white noise. Under either B.6.(i) or B.6.(ii), Proposition 2.1.1 of Straumann (2004) implies that $(\mathbf{x}_i)_{i\in\mathbb{Z}}$ is stationary and ergodic. The rank condition ensures asymptotic identification and it is ensured if Ψ_0 has full rank. Then, B.5 along with B.6, the ergodic LLN and the CMT imply that $n(\mathbf{x}_i^T \mathbf{x}_i)^{-1} \rightsquigarrow (\sum_{k=0}^{\infty} \Psi_k \Psi_k^T)^{-1}$. The independence between the scale and the white noise directly imply then the independence between the regressors' process and the scale. If u_i is structured as the aforementioned linear process, this part of B.5 would follow from independence between the associated white noise processes. B.4 and B.5 then imply that for any $\lambda \in \mathbb{R}^d$, $(\lambda \mathbf{x}_i u_i)_{i\in\mathbb{Z}}$ is stationary and ergodic. B.3 and B.5 then imply via Breiman's Theorem, that $\mathbb{P}[|\lambda \mathbf{x}_0| u_0 > x] = O(x^{-\alpha} \ln^{-\eta}(x))$, which then implies that $\mathbb{E}[(|\lambda \mathbf{x}_0| u_0)^{\alpha}] < +\infty$, and $\mathbb{E}[(|\lambda \mathbf{x}_0| u_0)^{\alpha+\delta}] = +\infty$ for all $\delta > 0$, and that $\mathbb{P}[|\lambda \mathbf{x}_0| u_0 > x] = o(\mathbb{P}[\xi_0 > x])$, as $x \to +\infty$. B.6.(i) says that the regressors' process has long range dependence as the decay of the temporal covariances is harmonic. It can be readily extended so as to include non trivial slowly varying parts attached to the harmonic coefficients. B.6.(i) is compatible with the regressors' process satisfying some stationary and ergodic version of the VARFIMA model (see for example Kechagias and Pipiras (2015)). This implies that for any $\lambda \in \mathbb{R}^d$, $(\lambda \mathbf{x}_i u_i)_{i\in\mathbb{Z}}$ is not strong mixing. B.6.(ii) is a condition of weak dependence for the regressors' process-see again Theorem 2.1 of Chanda (1974). It along with B.4 and B.5 implies the strong mixing property for the multiplicative scalar process $(\lambda \mathbf{x}_i u_i)_{i\in\mathbb{Z}}$. B.6 is in both cases compatible with the existence of the matrix series in B.5.

The discussion above shows that regarding the transform $V_i \xi_i$, where $V_i := \lambda \mathbf{x}_i u_i$, Assumptions 1-3 follow from B.1-B.5. Additionally, B.3.(ii) and B.6.(i) verify Assumption 4.(1), while B.3.(i) and B.6.(ii) verify Assumption 5.(2), since $\eta > 1$, and $\sup_{\sqrt{x} \leq y \leq x} (\frac{\ln(x)}{\ln(y)})^{\eta} = 2^{\eta}$. Furthermore, when $\alpha < 1$, B.2 implies (4) via Lemma 1, and thereby the totality of Theorem 1 is applicable to the partial sums of $(V_i \xi_i)_{i\in F}$, which along with the aforementioned limiting properties of $n(\mathbf{x}_i^T \mathbf{x}_i)^{-1}$ and the latent representation of the estimator establish the following result:

Theorem 2. *Suppose that B.1, B.2, B.4 and B.5, and either B.3.(ii) and B.6.(i), or B.3.(i) and B.6.(ii), hold. Then, as $n \to +\infty$, if $\alpha \neq 1$,*

$$n^{\frac{\alpha-1}{\alpha}} (\beta_n - \beta) \rightsquigarrow (\sum_{k=0}^{\infty} \Psi_k \Psi_k^T)^{-1} z_{\beta},$$

where z_β follows a p-dimensional stable distribution characterized via the projections $\lambda^T z_\beta \sim S_\alpha(0, c(\lambda), 0)$, with $c(\lambda) := c\mathbb{E}\left[(|\lambda^T \mathbf{x}_0|u_0)^\alpha\right]$, for any non-zero $\lambda \in \mathbb{R}^d$. If $\alpha = 1$, then,

$$\beta_n - \beta - H(n)(\mathbf{x}^T\mathbf{x})^{-1}\sum_{i=1}^{n}\mathbf{x}_i^T u_i \rightsquigarrow (\sum_{k=0}^{\infty}\Psi_k\Psi_k^T)^{-1} z_\beta,$$

where z_β follows now a p-dimensional stable distribution characterized via the projections $\lambda^T z_\beta \sim S_1(0, c(\lambda), 0)$, with $c(\lambda) := c\mathbb{E}\left[(|\lambda^T \mathbf{x}_0|u_0)\right]$, for any non-zero $\lambda \in \mathbb{R}^d$.

When $\alpha > 1$, the result shows weak consistency, regularly varying rates and asymptotic symmetric and zero location stable distributions for the OLSE. In the very heavy-tailed case where $\alpha < 1$, the result implies inconsistency and asymptotic non-tightness. When $\alpha = 1$, and due to the implications of the ergodic theorem and the CMT on the limiting behavior of $(\mathbf{x}^T\mathbf{x})^{-1}\sum_{i=1}^{n}\mathbf{x}_i^T u_i$, the results also imply inconsistency. Whenever either $n(\mathbf{x}^T\mathbf{x})^{-1}$ and/or $\sum_{i=1}^{n}\mathbf{x}_i^T u_i$ have regularly varying rates of convergence, then the result can be re-expressed as $\beta_n - \beta - (\sum_{k=0}^{\infty}\Psi_k\Psi_k^T)^{-1}H(n)\mathbb{E}[\mathbf{x}_0^T u_0] \rightsquigarrow (\sum_{k=0}^{\infty}\Psi_k\Psi_k^T)^{-1}z_\beta$, given that H is slowly varying at infinity. For example, whenever B.3.(i) and B.6.(ii) hold, then $\sqrt{n}$ rates exist for $n(\mathbf{x}^T\mathbf{x})^{-1} - (\sum_{k=0}^{\infty}\Psi_k\Psi_k^T)^{-1}$ as long as $\mathbb{E}[\|\mathbf{x}_0\|^{2+\eta^\star}] < +\infty$ holds for some $\eta^\star > 0$.

The results above are derivable without utilizing any point process theory associated with the random elements involved, and without any calculus for the derivation of properties of the relevant extremal index.

5 Application: The Gaussian QMLE in GARCH-type models

The consideration of heavy-tailed distributions for the squared innovation process in GARCH-type models became of interest in financial applications (see e.g. Rachev and Mittnik (1988) and Mittnik et al. (1998)). Focusing on the issue of its implications on the limit theory of the Gaussian QMLE for the GARCH(p, q) model, Mikosch and Straumann (2006) employ their martingale limit theorem which among others depends on a mixing condition for the volatility process and strict positivity for the extremal index of the associated martingale transform appearing in the score vector. They indicate that such conditions seem indispensable and thereby have to be in each case confirmed, in order for their results to be extended to other GARCH-type models. Our martingale limit theory avoids mixing in cases that adhere to Assumption 4 and satisfy (9) of Proposition 1. It also avoids the explicit consideration of the relevant extremal index. Case 2 of Assumption 4 is relevant in many such models, hence one obvious application of our main result concerns the subsequent extension of the Mikosch and Straumann (2006) limit theory in such examples. Hence in this section we derive the limit theory of the Gaussian QMLE for the parameters of the volatility process, in the context of a general stationary and ergodic conditionally heteroskedastic formulation via among others the use of Theorem 1.

5.1 Framework and results

The following general model is considered:

$$\begin{cases} y_i = \sigma_i z_i \\ \sigma_i^2 = g_\theta \left(y_{i-1}, y_{i-2}, \dots, y_{i-p}, \sigma_{i-1}^2, \dots, \sigma_{i-q}^2\right) \end{cases}, \ i \in \mathbb{Z} \tag{13}$$

where for each θ lying in some compact parameter space $\Theta \subset \mathbb{R}^d$, g_θ is a non-negative measurable function on $\mathbb{R}^p \times [0, \infty)^q$, and the sequence $(z_t)_{t\in\mathbb{Z}}$ consists of i.i.d. random variables with $\mathbb{E}[z_0] = 0$, while σ_{i+1}^2 is measurable w.r.t. $\mathcal{F}_i = \sigma(z_i, z_{i-1}, \dots)$. Examples of models that adhere to the formulation in (13) are (see among others Paragraph 3.3 in Straumann (2004)):

- GARCH(p, q), Bollerslev (1986), where $\sigma_i^2 = \omega + \sum_{j=1}^p a_j y_{i-j}^2 + \sum_{k=1}^q \beta_k \sigma_{i-k}^2$,
- GQARCH(p, q), Sentana (1995), where $\sigma_i^2 = \omega + \sum_{j=1}^p (a_j y_{i-j} - \gamma_j)^2 + \sum_{k=1}^q \beta_k \sigma_{i-k}^2$,
- AGARCH(p, q), Ding et al. (1993), where $\sigma_i^2 = \omega + \sum_{j=1}^p (a_j |y_{i-j}| - \gamma y_{i-j})^2 + \sum_{k=1}^q \beta_k \sigma_{i-k}^2$,
- EGARCH(p, q), Nelson (1991), where $\ln \sigma_i^2 = \omega + \sum_{j=1}^p \left(a_j \frac{|y_{i-j}|}{\sigma_{i-j}} + \gamma_j \frac{y_{i-j}}{\sigma_{i-j}}\right) + \sum_{k=1}^q \beta_k \ln \sigma_{i-k}^2$.

Using a stochastic recurrence equation (SRE) approach, Straumann and Mikosch (2006) give sufficient conditions for the existence of a stationary and ergodic solution to (13) by imposing conditions on the Lipschitz coefficient of the random map that generates σ_i^2. Suppose that $((y_i, \sigma_i^2))_{i\in\mathbb{Z}}$ is such a solution for some $\theta = \theta_0$. In order to infer the unknown parameter θ_0, since in practice only $(y_i)_{i=1,\dots,n}$ is observable, the volatility process (σ_i^2) is reconstructed under the parameter hypothesis θ using the following volatility filter for some random vector of initial values $(\varsigma_0, \varsigma_1, \dots, \varsigma_{q-1})' \in \mathbb{R}^q$,

$$\hat{h}_i(\theta) = \begin{cases} \varsigma_i, & q-1 \le i \le 0, \\ g_\theta \left(y_{i-1}, y_{i-2}, \dots, y_{i-p}; \hat{h}_{i-1}(\theta), \dots, \hat{h}_{i-q}(\theta)\right), & i > 0. \end{cases}$$

which is non-stationary in general. Thus, in order to use established results, it is useful to be able to uniformly approximate the latter using a stationary and ergodic filter. This is ensured when the model (13) is uniformly invertible, which in turn implies that there exists a stationary ergodic sequence $(h_i(\theta))_{i\in\mathbb{Z}}$ for all $\theta \in \Theta$ such that $h_i(\theta_0) = \sigma_i^2$, $\mathbb{P}$ a.s. and $\sup_{\theta\in\Theta} \left| h_i(\theta) - \hat{h}_i(\theta) \right| \to 0$, $\mathbb{P}$ a.s. at an exponential rate, as $i \to \infty$. Sufficient conditions for uniform invertibility can be found in Proposition 3.12 of Straumann and Mikosch (2006). Next, given $(y_i)_{i=1,\dots,n}$, the Gaussian quasi likelihood function is proportional to

$$\hat{c}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \hat{\ell}_i(\theta),$$

where $\hat{\ell}_i(\theta) = \log \hat{h}_i(\theta) + \frac{y_i^2}{\hat{h}_i(\theta)}$, and then the Gaussian QMLE θ_n of θ_0 is defined by

$$\hat{c}_n(\theta_n) = \inf_{\theta \in \Theta} \hat{c}_n(\theta).$$

Standard lower semi-continuity and measurability conditions imply existence of the estimator (see for example Definition 3.5 and Proposition 3.6 in Chapter 5 of Molchanov (2011)), while the definition-and the subsequent results-can be easily extended so that approximate/numerical optimization is allowed. Then strong consistency of the QMLE is implied by the following conditions that are usually met in the literature (see e.g. Straumann and Mikosch (2006)):

C.1 For $\theta = \theta_0$, (13) admits a unique stationary ergodic solution $((y_i, \sigma_i^2))_{i \in \mathbb{Z}}$ with $\mathbb{E}\left[\log^+ \sigma_0^2\right] < +\infty$.

C.2 The model (13) is uniformly invertible in Θ.

C.3 $\inf_{\theta \in \Theta} \inf_{\mathbf{x} \in \mathbb{R}^p \times [0,\infty)^q} g_\theta(\mathbf{x}) > 0$

C.4 For any $\theta \in \Theta$, $h_0(\theta) = \sigma_0^2 \Leftrightarrow \theta = \theta_0$, $\mathbb{P}$ a.s.

C.5 $\mathbb{E}[z_0^2] = 1$.

Conditions C.1-C.3 enable us to work with a stationary and ergodic version of $\hat{c}_n$, say c_n, where $\left(\hat{h}_i(\theta)\right)_{i=1,\dots,n}$ is replaced by the stationary $(h_i(\theta))_{i=1,\dots,n}$ since we have that for some $\eta > 0$, $\sup_{\Theta \cap \bar{B}(\theta,\eta)} |c_n - \hat{c}_n| \to 0$ $\mathbb{P}$ a.s. as in Straumann and Mikosch (2006). They also allow for the $\mathbb{P}$ a.s. epi-convergence of c_n to a non stochastic lower semi-continuous limit that is a proper extended real-valued function on Θ. Conditions C.4-C.5 ensure asymptotic identification, i.e. that the limit function is uniquely minimized at θ_0. C.5 fails when z_0^2 adheres to Assumption 1 and $\alpha = 1$, with $r_n \nrightarrow 0$, or $\alpha < 1$. From C.1-5 strong consistency follows by results like Theorem 7.33 of Rockafellar (1997):

Theorem 3. *(See Theorem 5.3.1 of Straumann (2004)) Suppose that Conditions C.1-C.5 hold. Then, the QMLE is strongly consistent.*

The condition of uniform invertibility may be quite strong in some cases. It can be relaxed to continuous invertibility, a notion introduced by Wintenberger (2013); Wintenberger and Cai (2011) who use it in the context of the EGARCH$(1,1)$ model. Since our main focus lies in the establishment of the rate of convergence and the limiting distribution, as those are essentially established by the asymptotic behavior of the ergodic approximation to the score, we do not insist on the fine details of such framework extensions for brevity. Our results in Theorem 4 below can be readily extended in such cases. Given consistency, we provide with sufficient conditions for the derivation of the rate and the asymptotic distribution of the QMLE under heavy-tailedness for the marginals of the squared innovation process. Except for the possibility of the non-existence of $\mathbb{E}[z_0^4]$, those are identical to the ones in Straumann and Mikosch (2006).

D.1 Conditions C.1-C.5 hold and the true parameter vector θ_0 lies in the interior of Θ.

D.2 Assumption 1 holds with $\xi_i := z_i^2 - 1,\ i \in \mathbb{Z}$.

D.3 h_i has continuous partial derivatives of order 2 on some open neighborhood of θ_0 inside Θ. Those satisfy conditions D.2-D.3 of Proposition 5.5.2 in Straumann (2004).

D.4 $\mathbb{E}\left[\sup_{\theta\in K}\|\ell_0''\|\right] < +\infty$ where K denotes the aforementioned neighborhood.

D.5 The components of the vector $\frac{\partial g_\theta}{\partial \theta}\left(y_0, \sigma_0^2\right)\big|_{\theta=\theta_0}$ are linearly independent random variables.

Conditions D.1 and D.2 imply together that $\alpha > 1$ while they allow for cases where $\alpha = 1$ and $\int^{+\infty} \frac{L(x)}{x} dx$ converges. Condition D.2 is essentially a generalization of the usual framework in that it enables the possibility that $\mathbb{E}\left[z_0^4\right] = +\infty$, while it encompasses the standard case. Condition D.3 ensures the existence of unique stationary and ergodic approximations of the first and second order partial derivatives of $\hat{h}_t$, say h_t' and h_t'', and that the error of approximation decays exponentially fast $\mathbb{P}$ a.s. This in turn ensures that we may replace the derivatives of the likelihood function by their ergodic counterparts, in our derivations. As noted above such conditions are establishable by conditions that ensure either uniform invertibility or the weaker continuous invertibility for the solutions of the relevant recursions satisfied by the filter derivatives. Condition D.4 implies $\mathbb{E}\left[\|h_0'(\theta_0)\|^2 \sigma_0^{-4}\right] < \infty$, which is standard for the classical limit theory, and is stronger when $\alpha \in [1, 2)$ to the $\mathbb{E}\left[\left\|h_0'(\theta_0)\sigma_0^{-2}\right\|^{\alpha+\delta}\right] < +\infty$ which can be sufficient for Theorem 1 to be applicable. However D.4 could be relaxed considering the results of Wintenberger (2013) for the EGARCH(1,1) and this generates interesting refinements as we will discuss later. D.2-D.4 imply that the score vector evaluated at θ_0, as approximated by a stationary ergodic process adheres to Theorem 1 via case 1 of Proposition 4. Condition D.5 ensures that $\mathbf{J}_{\theta_0} := \mathbb{E}\left[\ell_0''(\theta_0)\right] = \mathbb{E}\left[h_0'(\theta_0)\left[h_0'(\theta_0)\right]^T \sigma_0^{-4}\right]$ is positive definite (see Lemma 5.6.3. in Straumann (2004)) and thereby in conjunction with the previous that the Hessian is asymptotically invertible when K is small enough. Then the main result of this section is easily derivable by standard Taylor type approximations of the first order conditions that the estimator satisfies with probability converging to 1. As in Theorem 1 the formulation of the result depends on the value of α since the case of $\alpha = 1$ involves the relevant translating sequence.

Theorem 4. *Suppose that Conditions D.1-D.5 hold. Then, for $\alpha \in (1, 2]$*

$$n^{\frac{\alpha-1}{\alpha}} r_n^{-\frac{1}{\alpha}} \left(\theta_n - \theta_0\right) \rightsquigarrow \mathbf{J}_{\theta_0}^{-1} z_{\theta_0}, \tag{14}$$

where z_{θ_0} follows a multivariate α-stable distribution characterized by the projections $\lambda^T z_{\theta_0} \sim S_\alpha\left(\beta(\lambda), c(\lambda), 0\right)$, with $\beta(\lambda) := \frac{\mathbb{E}\left[\left|\lambda^T h_0'(\theta_0)\right|^\alpha \operatorname{sgn}(\lambda^T h_0'(\theta_0))\right]}{\mathbb{E}\left[\left|\lambda^T h_0'(\theta_0)\right|^\alpha\right]}$ and with $c(\lambda) := c\mathbb{E}\left[\frac{1}{\sigma_0^{2\alpha}}\left|\lambda^T h_0'(\theta_0)\right|^\alpha\right]$, for any non-zero $\lambda \in \mathbb{R}^d$.

For $\alpha = 1$ and for some $\bar{\theta}_n$ lying between θ_n and θ_0,

$$\frac{1}{r_n}(\theta_n - \theta_0) - \frac{\gamma + H(nr_n)}{r_n}\left[c_n''(\bar{\theta}_n)\right]^{-1} \frac{1}{n}\sum_{i=1}^n \frac{h_i'(\theta_0)}{\sigma_i^2} \rightsquigarrow \mathbf{J}_{\theta_0}^{-1} z_{\theta_0}', \tag{15}$$

where z'_{θ_0} *follows a multivariate* 1*-stable distribution characterized by* $\lambda^T z'_{\theta_0} \sim S_\alpha(\beta(\lambda), c(\lambda), \gamma(\lambda))$*, with* $\beta(\lambda)$ *and* $c(\lambda)$ *as above, and also with* $\gamma(\lambda) := 2\beta c\pi^{-1}\left\{C\mathbb{E}\left[\sigma_0^{-2}\lambda^T h_0'(\theta_0)\right] + \mathbb{E}\left[\sigma_0^{-2}\lambda^T h_0'(\theta_0)\log\left|\sigma_0^{-2}\lambda^T h_0'(\theta_0)\right|\right]\right\}$*, for any* $\lambda \in \mathbb{R}^d$*.*

The result above avoids any invocation of any mixing condition, in contrast to the approach of Mikosch and Straumann (2006) (see their page 6 for a mixing condition implied by strong mixing) as well as the need to explicitly verify anti-clustering type conditions like the ones in Bartkiewicz et al. (2011) or Davis and Hsing (1995). Moreover, whenever suchlike mixing conditions are known, they typically involve restrictions on the support of the distribution of z_0 and the existence of densities. Such restrictions are avoided here. The result also avoids explicit considerations on the positivity of the extremal index of the score vector process. Hence, it implies direct applicability for a wide class of models for which the validity of the combination of such properties is to the best of our knowledge, currently unknown. The rate is identified as $n^{\frac{\alpha-1}{\alpha}} r_n^{-\frac{1}{\alpha}}$ and the distribution of the limiting random vector z_{θ_0} is multivariate α stable. Given the limiting behavior of the Hessian, the limiting distribution is multivariate α stable as in the results of the relevant literature. As noted earlier, the result clearly provides more information on the rate and the limiting distribution compared to the analogous literature.

When $\alpha = 2$ we have that $z_{\theta_0} \sim N(0, \mathbf{J}_{\theta_0})$ and thereby that $\sqrt{\frac{n}{r_n}}(\theta_n - \theta_0) \rightsquigarrow N\left(0, \mathbf{J}_{\theta_0}^{-1}\right)$, a result not obtainable by the approach of Mikosch and Straumann (2006). When furthermore $\mathbb{E}[z_0^4] < \infty$ we obtain the classical result since then $r_n \rightsquigarrow \mathbb{E}[z_0^4] - 1$. However, we still obtain asymptotic normality with a slower than the usual rate, when the truncated fourth moment is slowly varying and diverging at infinity. For example when $\sqrt{2}z_0 \sim t_4$ then simple calculations show that $\sqrt{\frac{n}{\log n}}(\theta_n - \theta_0) \rightsquigarrow N\left(0, \frac{3}{2}\mathbf{J}_{\theta_0}^{-1}\right)$.

When $\alpha = 1$ and Condition C.5 holds the result involves the non tight stochastic translating sequence $\frac{\gamma+H(nr_n)}{r_n}\left[c_n''(\bar{\theta}_n)\right]^{-1}\frac{1}{n}\sum_{i=1}^n \frac{h_i'(\theta_0)}{\sigma_i^2}$. Given Conditions D.3-4, if $\frac{1}{r_n}\left\|\left[c_n''(\bar{\theta}_n)\right]^{-1}\frac{1}{n}\sum_{i=1}^n\left(\frac{h_i'(\theta_0)}{\sigma_i^2} - \mathbb{E}\left[\frac{h_0'(\theta_0)}{\sigma_0^2}\right]\right)\right\| = o_p(1)$, the sequence can be replaced by $\frac{\gamma+H(nr_n)}{r_n}\mathbf{J}_{\theta_0}^{-1}\mathbb{E}\left[\frac{h_0'(\theta_0)}{\sigma_0^2}\right]$. This would hold for example whenever a mixing-type CLT holds for the partial sum process $\sum_{i=1}^n\left(\frac{h_i'(\theta_0)}{\sigma_i^2} - \mathbb{E}\left[\frac{h_0'(\theta_0)}{\sigma_0^2}\right]\right)$. If the translating sequence can be chosen as $\frac{\gamma+H(nr_n)}{r_n}\mathbf{J}_{\theta_0}^{-1}\mathbb{E}\left[\frac{h_0'(\theta_0)}{\sigma_0^2}\right]$, then the proof of this part of the theorem implies the inconsistency of the estimator in case where $r_n \to +\infty$: assuming consistency we would arrive at the contradictory result that the translating sequence is tight.

When $\alpha > 1$ the results can be easily extended to the case where θ_0 lies on the boundary of Θ, via the generalization of Condition D.3 to one-sided derivatives in the spirit of Andrews (1999), the consideration of the Painleve-Kuratowski limit of the set $n^{\frac{\alpha-1}{\alpha}} r_n^{-\frac{1}{\alpha}}(\Theta - \theta_0)$ and a generalization of Theorem 7.12 of van der Vaart (2000). The rates would remain the same but the limiting distribution would be characterized via an appropriate projection of $\mathbf{J}_{\theta_0}^{-1} z_{\theta_0}$ on the limiting parameter space. Such a formulation would not work for the case where $\alpha = 1$ due to the presence of the translating sequences. This consideration is delegated to future research.

The information present in the results on the properties of the limiting distributions could be useful for the construction and the determination of properties of inferential procedures about θ_0 based on the QMLE. For example, in the GARCH(p,q) case, it is not difficult to see that when $\alpha = 2$ yet $\mathbb{E}[z_0^4] = +\infty$, the usual Wald-type inferential procedures remain asymptotically robust (see for example Section 3 of Hall and Yao (2003)). Furthermore, the classical Wald-type statistic remains self-normalized in the more general case where $\alpha > 1$, but then, the limiting null distribution is not the usual and asymptotic inference based on the classical limiting rejection regions may be asymptotically conservative; see among others Loretan and Phillips (1994) for an analogous discussion in related frameworks. One way to avoid such shortcomings is via the use of the m out of n parametric bootstrap in the spirit of Hall and Yao (2003). It may thus be possible to use the information present in the results above in order to further normalize the statistic so as to obtain a limiting distribution where several nuisance parameters are eliminated. Furthermore, it is possible that the bootstrap resampling can be performed via the use of the information present in condition D.2 in the spirit of Cornea-Madeira and Davidson (2015) so as to avoid subsampling rates. Such considerations also involving further extensions to other models are delegated to further research.

5.2 Examples

The previous imply that the existing results for the Gaussian QMLE in the GARCH(p,q) case can be extended for a variety of conditionally heteroskedastic models without necessarily imposing conditions that are either difficult to verify or unnecessarily restrict the parameters of interest. Simultaneously we can identify in more detail both the rate of convergence and the properties of the limiting distributions.

As illustrative examples consider the GQARCH$(1,1)$, the AGARCH(p,q) and the EGARCH$(1,1)$ models. The first one has been treated in Arvanitis and Louka (2015) assuming normal domains of attraction for $\alpha \in (1,2]$ and those results can be clearly extended to general domains of attraction as well as including the case where $\alpha = 1$. Analogous is the case for the AGARCH(p,q) model as the required conditions for the asymptotic normality of the QMLE which have been shown in Straumann and Mikosch (2006) are also sufficient for obtaining stable limits when the assumption on the innovation process is relaxed to condition D.2. Hence our main theorem enables the straightforward establishment of the examined limit theory in those classes of models by avoiding non trivial derivations of mixing and extremal index properties as in Mikosch and Straumann (2006). The EGARCH$(1,1)$ model can be also easily examined, via the results in Wintenberger (2013), and by imposing the condition $\mathbb{E}\left[(\beta_0 - \frac{1}{2}(a_0|z_0| + \gamma_0 z_0))^2\right] < 1$ which ensures that $\mathbb{E}\left[\left\|h_0'(\theta_0)\sigma_0^{-2}\right\|^2\right] < +\infty$ via Lemma 1 in the Appendix of the aforementioned paper. Thereby, for the aforementioned models the results of Theorem 4 follow easily.

Theorem 1 could also be used in order to obtain the results on the limit theory of the QMLE in non-stationary versions of suchlike models. Consider for example the results in Arvanitis and Louka (2017) for the non-stationary GARCH(1,1) case, or in Arvanitis (2019) for the non-stationary version of the Asymmetric GARCH(1,1) model of Francq and Zakoïan (2013). There, results analogous to Theorem 4 hold, with differences in the expressions for the $\beta(\lambda)$ and $c(\lambda)$ parameters as well as the limiting form of the Hessian, that reflect properties of the stationary approximations to the score process.

6 Proofs

6.1 Proofs of auxiliary results

Proof of Lemma 1. We have that $\mathbb{P}\left[\max_{1\leq i\leq n}|V_i| > Mr_n^{\frac{1}{\alpha}}n^{\frac{1}{\alpha}}\right]$ equals

$$\mathbb{P}[\cup_{i=1}^{n}\left\{|V_i| > Mr_n^{\frac{1}{\alpha}}n^{\frac{1}{\alpha}}\right\}] \leq \sum_{i=1}^{n}\mathbb{P}\left[|V_i| > Mr_n^{\frac{1}{\alpha}}n^{\frac{1}{\alpha}}\right] = n\mathbb{P}\left[|V_0| > Mr_n^{\frac{1}{\alpha}}n^{\frac{1}{\alpha}}\right],$$

due to the stationarity of $(V_i)_{i\in\mathbb{Z}}$. Then the result follows by Assumptions 1 and 3. □

Proof of Proposition 1. As in Denisov and Zwart (2007) (see the proofs of Propositions 2.1, 2.2 and 2.3) we can work by replacing $|V_0|$, $|\xi_0|$ and m_n with $|V_0|^{\alpha}$, $|\xi_0|^{\alpha}$ and $m_n^{\frac{1}{\alpha}}$ respectively, and thereby we can assume that $\alpha = 1$, and without loss of generality we can assume that $\mathbb{P}[|V_0| = 0] = 0$. Notice first that

$$\liminf_{n\to\infty}\frac{\mathbb{P}_n^{\star}[v_n|\xi_0| > m_n]}{\mathbb{P}[|\xi_0| > m_n]} = \liminf_{n\to\infty}\int_0^{\infty}\frac{\mathbb{P}[t|\xi_0| > m_n]}{\mathbb{P}[|\xi_0| > m_n]}d\mathbb{P}_n[v_n \leq t] \geq \int_0^{\infty} td\mathbb{P}[|V_0| \leq t],$$

where the last inequality in the previous display follows from Fatou's lemma, the Glivenko-Cantelli Theorem (see Theorem 1 of Adams and Nobel (2010)), Birkhoff's LLN, and the regular variation of index -1 of the tail of $|\xi_0|$. For an upper bound consider

$$\mathbb{P}_n^{\star}[v_n|\xi_0| > m_n] = \sum_{j=1}^{3}\mathbb{P}_n^{\star}[v_n|\xi_0| > m_n, v_n \in A_{i,n}],$$

with $A_{1,n} = (0,\varepsilon)$, $A_{2,n} = [\varepsilon, g_n m_n)$, and $A_{3,n} = [g_n m_n, \infty)$ for some $\varepsilon > 0$, and, $g_n \downarrow 0$ with $g_n m_n \to +\infty$, such that $\mathbb{P}[|V_0| > g_n m_n] = o(\mathbb{P}[|\xi_0| > m_n])$. Denote the respective terms as $\mathcal{I}_{1,n}, \mathcal{I}_{2,n}, \mathcal{I}_{3,n}$. Due to the UCT for regularly varying functions with negative index (see Theorem 1.5.2 of Bingham et al. (1989)) and Birkhoff's LLN we obtain

$$\limsup_{n\to+\infty}\frac{\mathcal{I}_{1,n}}{\mathbb{P}[|\xi_0| > m_n]} = \mathbb{E}[|V_0|\,1(0 < |V_0| < \varepsilon)],\ \text{a.s.}$$

Furthermore by the construction of g_n and the Glivenko-Cantelli Theorem (see Theorem 1 of Adams and Nobel (2010))

$$\limsup_{n\to+\infty}\frac{\mathcal{I}_{3,n}}{\mathbb{P}[|\xi_0| > m_n]} \leq \limsup_{n\to+\infty}\frac{\mathbb{P}_n[v_n > g_n m_n]}{\mathbb{P}[|\xi_0| > m_n]} = \limsup_{n\to+\infty}\frac{\mathbb{P}_n[|V_0| > g_n m_n]}{\mathbb{P}[|\xi_0| > m_n]} = 0,\ \text{a.s.},$$

since $\mathbb{P}[|V_0| > g_n m_n] = o(\mathbb{P}[|\xi_0| > m_n])$. Using the above we have that

$$\limsup_{n\to\infty} \frac{\mathbb{P}_n^\star[v_n|\xi_0| > m_n]}{\mathbb{P}[|\xi_0| > m_n]} \leq \mathbb{E}[|V_0| 1(0 < |V_0| < \varepsilon)] + \limsup_{n\to+\infty} \frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0| > m_n]}.$$

Thus, it suffices for (9) that the last term in the rhs of the previous display converges a.s. to zero as $n \to +\infty$ and then $\varepsilon \to +\infty$. Notice that for this term we have that

$$\frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0| > m_n]} = \int_{\varepsilon}^{g_n m_n} \frac{L(\frac{m_n}{t})}{L(m_n)} t d\mathbb{P}_n[v_n \leq t].$$

Suppose that Assumption 4.1 holds. Then for $C_n := \sup_{y\in[1,m_n]} \frac{L(y)}{L(m_n)} = \sup_{t\in[1,m_n]} \frac{L(\frac{m_n}{t})}{L(m_n)}$, and due to Birkhoff's LLN we obtain that

$$\limsup_{n\to+\infty} \frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0| > m_n]} \leq \limsup_{n\to+\infty} C_n \mathbb{E}_n[u_n 1(u_n > \varepsilon)] \leq \mathbb{E}[|V_0| 1(|V_0| > \varepsilon)] \limsup_{n\to+\infty} C_n, \mathbb{P}\text{a.s.}$$

Since $\lim\sup_{n\to+\infty} C_n < \infty$ and $\mathbb{E}[|V_0|] < \infty$ the result follows.

Suppose now that Assumption 4.2 holds. Given δ, we can choose ε large enough so that Potter's Theorem (see Theorem 1.5.6 (i)-(ii) of Bingham et al. (1989)) is applicable and via Birkhoff's LLN we obtain that

$$\limsup_{n\to+\infty} \frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0| > m_n]} \leq \int_{\varepsilon}^{\infty} t^{1+\frac{\delta}{\alpha}} d\mathbb{P}[|V_0| \leq t] = \mathbb{E}\left[|V_0|^{1+\frac{\delta}{\alpha}} 1(|V_0| > \varepsilon)\right],$$

and the result follows since $\mathbb{E}\left[|V_0|^{1+\frac{\delta}{\alpha}}\right] < \infty$.

For establishing (10) notice first that due to the previous, it suffices that $\frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0|>m_n]}$ converges in probability to zero. Secondly, notice that due to Markov's, the generalized von Bahr-Esseen (see Theorem 1 of Hill (2010)) and the C_r inequalities, for any $\delta > 0$, and for $\eta > 0$ such that $(1+\eta)\epsilon < 1$, we have that for some $C > 0$ and $r = 1+\eta$

$$\mathbb{P}\left[\frac{|\mathbb{P}_n[v_n > t^-] - \mathbb{P}[|V_0| > t]|}{(\mathbb{P}[|V_0| > t])^\epsilon} > \delta\right] \leq \frac{C}{\delta^{1+\eta} n^\eta} \left((\mathbb{P}[|V_0| > t])^{1-(1+\eta)\epsilon} + (\mathbb{P}[|V_0| > t])^{1-\epsilon}\right), \tag{16}$$

and the rhs of the previous display converges to zero uniformly in t. Hence $\frac{\mathbb{P}_n[v_n>t^-]-\mathbb{P}[|V_0|>t]}{(\mathbb{P}[|V_0|>t])^\epsilon}$ is bounded with probability converging to one uniformly in t.

Suppose now that Assumption 5.1 holds. Define $c_n^\star := \sup_{t\in[\varepsilon,g_n m_n]} c\left(\frac{m_n}{t}\right)/c(m_n)$ and notice that $(c_n^\star)$ is bounded since by construction $c(x)$ converges to a positive real number as $x \to +\infty$. Then

$$\frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0| > m_n]} \leq c_n^\star \int_{\varepsilon}^{g_n m_n} \frac{\mathbb{P}[U > \ln m_n - \ln t]}{\mathbb{P}[U > \ln m_n]} t d\mathbb{P}_n[v_n \leq t].$$

Hence it suffices to show that integral in the rhs of the previous display a.s. converges to zero as $n \to +\infty$ and then $\varepsilon \to +\infty$. Let $s(x) := \frac{x\mathbb{P}[|V_0|>x]}{\mathbb{P}[U>\ln x]}$ which converges to zero as $x \to +\infty$ due to the final part of Assumption 5.1. Using the integration by parts formula for the Lebesgue-Stieljes integral-see Theorem 21.67 in Hewitt and Stromberg (2013)-we have that

$$\frac{1}{\mathbb{P}\left[U>\ln m_n\right]}\int_{\varepsilon}^{g_n m_n}\mathbb{P}\left[U>\ln m_n-\ln t\right]td\mathbb{P}_n\left[v_n\leq t\right]$$

$$=-\mathbb{P}_n\left[v_n>g_n m_n\right]\frac{\mathbb{P}\left[U>-\ln g_n\right]}{\mathbb{P}\left[U>\ln m_n\right]}g_n m_n+\mathbb{P}_n\left[v_n>\varepsilon^-\right]\frac{\mathbb{P}\left[U>\ln m_n-\ln\varepsilon\right]}{\mathbb{P}\left[U>\ln m_n\right]}\varepsilon$$

$$+\int_{\varepsilon}^{g_n m_n}\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}\mathbb{P}_n\left[v_n>t^-\right]\mathrm{d}t+\int_{\varepsilon}^{g_n m_n}\mathbb{P}_n\left[v_n>t^-\right]td_t\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}.$$

The first term is the rhs of the previous display is less than or equal to zero and thereby can be ignored for the construction of an upper bound. For the second term since $U\in S^*$ (i.e. U has a sub-exponential tail distribution function; see Klüppelberg (1988)) and due to the Glivenko-Cantelli theorem we have that

$$\limsup_{n\to+\infty}\mathbb{P}_n\left[v_n>\varepsilon^-\right]\frac{\mathbb{P}\left[U>\ln m_n-\ln\varepsilon\right]}{\mathbb{P}\left[U>\ln m_n\right]}\varepsilon\leq\varepsilon\mathbb{P}\left[|V_0|>\varepsilon\right],\text{ a.s.}$$

For the third term we have that for large enough ε,

$$\int_{\varepsilon}^{g_n m_n}\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}\mathbb{P}_n\left[v_n>t^-\right]\mathrm{d}t$$

$$\leq\sup_{t\geq\varepsilon}s\left(t\right)\int_{\varepsilon}^{g_n m_n}\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]\mathbb{P}\left[U>\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}\frac{\mathbb{P}_n\left[v_n>t^-\right]}{\mathbb{P}\left[|V_0|>t\right]}d\ln t,\qquad(17)$$

and thereby the rhs of (17) is bounded from above with probability converging to one by

$$\left(1+o_p\left(1\right)\right)\sup_{t\geq\varepsilon}s\left(t\right)\int_{\varepsilon}^{g_n m_n}\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]\mathbb{P}\left[U>\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}d\ln t\leq 2\mathbb{E}\left[U\right]\sup_{t\geq\varepsilon}s\left(t\right),$$

since $U\in S^*$. For the fourth term we analogously have that for large enough ε, and with probability converging to one

$$\int_{\varepsilon}^{g_n m_n}\mathbb{P}_n\left[v_n>t^-\right]td_t\frac{\mathbb{P}\left[U>\ln m_n-\ln t\right]}{\mathbb{P}\left[U>\ln m_n\right]}$$

$$\leq\sup_{t\geq\varepsilon}s\left(t\right)\int_{\varepsilon}^{g_n m_n}\frac{\mathbb{P}_n\left[v_n>t^-\right]}{\mathbb{P}\left[|V_0|>t\right]}d_t\mathbb{P}\left[U>\ln m_n-\ln t\right]$$

$$=\left(1+O_p\left(1\right)\right)\sup_{t\geq\varepsilon}s\left(t\right)\int_{\varepsilon}^{g_n m_n}d_t\mathbb{P}\left[U>\ln m_n-\ln t\right]\leq 2\sup_{t\geq\varepsilon}s\left(t\right).$$

The previous imply that with probability converging to one, for large enough ε,

$$\frac{1}{\mathbb{P}\left[U>\ln m_n\right]}\int_{\varepsilon}^{g_n m_n}\mathbb{P}\left[U>\ln m_n-\ln t\right]td\mathbb{P}_n\left[v_n\leq t\right]$$

$$\leq \varepsilon \mathbb{P}\left[|V_0| \geq \varepsilon\right] + 2\left(\mathbb{E}\left[U\right]+1\right)\left(1+O_p\left(1\right)\right)\sup_{t\geq\varepsilon} s\left(t\right),$$

and the latter converges to zero as $\varepsilon \to +\infty$.

Suppose finally that Assumption 5.2 holds. As in the proof of Proposition 2.3 of Denisov and Zwart (2007) we can assume that eventually $g_n > \frac{1}{\sqrt{m_n}}$. Using this we have that $\frac{\mathcal{I}_{2,n}}{\mathbb{P}[|\xi_0|>m_n]} = \int_{\varepsilon}^{\sqrt{m_n}} \frac{L(\frac{m_n}{t})}{L(m_n)} t d\mathbb{P}_n\left[v_n \leq t\right] + \int_{\sqrt{m_n}}^{g_n m_n} \frac{L(\frac{m_n}{t})}{L(m_n)} t d\mathbb{P}_n\left[v_n \leq t\right]$. Then, due to Birkhoff's LLN

$$\limsup_{n\to+\infty} \int_{\varepsilon}^{\sqrt{m_n}} \frac{L(\frac{m_n}{t})}{L(m_n)} t d\mathbb{P}_n\left[v_n \leq t\right] \leq \mathbb{E}\left[|V_0|\, 1\left(|V_0| > \varepsilon\right)\right] \limsup_{n\to+\infty} \sup_{\sqrt{m_n}\leq t\leq m_n} \frac{L(\frac{m_n}{t})}{L(m_n)}, \text{a.s.},$$

and as $\varepsilon \to +\infty$, the term in the rhs of the previous display converges to zero. Furthermore, for the second integral above, using the integration by parts formula for the Lebesgue-Stieljes integral we obtain

$$\int_{\sqrt{m_n}}^{g_n m_n} \frac{L(\frac{m_n}{t})}{L(m_n)} t d\mathbb{P}_n\left[v_n \leq t\right] = -\frac{L(\frac{1}{g_n})}{L(m_n)} \mathbb{P}_n\left[v_n > g_n m_n\right] g_n m_n$$

$$+\frac{L(\sqrt{m_n})}{L(m_n)} \mathbb{P}_n\left[v_n > \sqrt{m_n}^{-}\right]\sqrt{m_n} + \frac{1}{L(m_n)} \int_{\sqrt{m_n}}^{g_n m_n} \mathbb{P}_n\left[v_n > t^{-}\right] d_t\left(tL\left(\frac{m_n}{t}\right)\right).$$

As previously the first term in the rhs of the previously display is less than or equal to zero and therefore can be ignored. For the second term we have that it is less than or equal to

$$\frac{1}{n}\sum_{i=1}^{n} |V_0|\, 1\left(|V_0| > \sqrt{m_n}\right) \sup_{\sqrt{m_n}\leq t\leq m_n} \frac{L(\frac{m_n}{t})}{L(m_n)},$$

and the latter converges a.s. to zero as $n \to +\infty$ since $\mathbb{E}\left[|V_0|\right] < \infty$. For the final term let $Q\left(x\right) := \int_0^x t^{\alpha} d\mathbb{P}\left[|\xi_0| \leq x\right]$. Due to the previous we have that as $n \to +\infty$ with probability tending to 1

$$\frac{1}{L(m_n)} \int_{\sqrt{m_n}}^{g_n m_n} \mathbb{P}_n\left[v_n > t^{-}\right] d_t\left(tL\left(\frac{m_n}{t}\right)\right) \leq$$

$$o(1)\left(1+O_p\left(1\right)\right) \int_{\sqrt{m_n}}^{g_n m_n} m_n \frac{\mathbb{P}\left[|\xi_0| > t\right]}{L\left(m_n\right) Q\left(t\right)} d\mathbb{P}\left[|\xi_0| \leq \frac{m_n}{t}\right],$$

and the integral in the rhs of the previous display is bounded from above exactly as in the proof of Proposition 2.3 of Denisov and Zwart (2007). □

Proof of Proposition 2. Notice first that since as $x \to +\infty$, $L\left(x\right) \sim \mathbb{E}\left[\left(\xi_i - \gamma\right)^2 1_{|\xi_i-\gamma|>x}\right]$, L has an equivalent monotone version. Assumption 3 and Birkhoff's LLN implies that
$\sum_{i=1}^{n} \mathbb{E}\left[\frac{1}{nr_n}\left(\xi_i-\gamma\right)^2 V_i^2 1_{|(\xi_i-\gamma)V_i|>M\sqrt{nr_n}}/\mathcal{G}_{i-1}\right]$ is asymptotically equivalent to
$\frac{L(M\sqrt{nr_n})}{nr_n}\sum_{i=1}^{n} V_i^2 \frac{L\left(\frac{M\sqrt{nr_n}}{|V_i|}\right)}{L(M\sqrt{nr_n})}$. The same argument implies that we can replace L by

its monotone equivalent version. Denote the latter with L for brevity. Let $\varepsilon > 0$ and consider

$$\frac{L\left(M\sqrt{nr_n}\right)}{nr_n}\sum_{i=1}^{n}V_i^2 1_{|V_i|>\varepsilon}\frac{L\left(\frac{M\sqrt{nr_n}}{|V_i|}\right)}{L\left(M\sqrt{nr_n}\right)} \leq \frac{L\left(\frac{M\sqrt{nr_n}}{\varepsilon}\right)}{L\left(M\sqrt{nr_n}\right)}\frac{L\left(M\sqrt{nr_n}\right)}{nr_n}\sum_{i=1}^{n}V_i^2 1_{|V_i|>\varepsilon}.$$

Due to Assumption 3, Birkhoff's LLN, the slow variation of L and the defining property of r_n, the rhs of the previous display converges $\mathbb{P}$ a.s. to $\mathbb{E}\left[V_0^2 1_{|V_0|>\varepsilon}\right]$. Furthermore, let $g(x) = x^{-2}L(x)$ and notice that

$$\frac{1}{n}\sum_{i=1}^{n}\left|\frac{g(M\sqrt{nr_n}|V_i|^{-1})}{g(M\sqrt{nr_n})} - |V_i|^{-2}\right| 1_{|V_i|\leq\varepsilon} \leq \sup_{\lambda\in[\varepsilon^{-1},\infty)}\left|\frac{g(\lambda M\sqrt{nr_n})}{g(M\sqrt{nr_n})} - \lambda^{-2}\right|\frac{1}{n}\sum_{i=1}^{n}1_{|V_i|\leq\varepsilon}.$$

Due to Assumption 3 and Birkhoff's LLN $\frac{1}{n}\sum_{i=1}^{n}1_{|V_i|>\varepsilon} = O(1)$ $\mathbb{P}$ a.s. Due to the UCT for regularly varying functions with negative index (see Theorem 1.5.2 of Bingham et al. (1989)) $\sup_{\lambda\in[\varepsilon^{-1},\infty)}\left|\frac{g(\lambda n^{1/\alpha}r_n^{1/\alpha}|t|^{-1})}{g(n^{1/\alpha}r_n^{1/\alpha}|t|^{-1})} - \lambda^{-\alpha}\right| = o(1)$. Hence

$$\frac{L\left(M\sqrt{nr_n}\right)}{nr_n}\sum_{i=1}^{n}V_i^2 1_{|V_i|\leq\varepsilon}\frac{L\left(\frac{M\sqrt{nr_n}}{|V_i|}\right)}{L\left(M\sqrt{nr_n}\right)} = \frac{L\left(M\sqrt{nr_n}\right)}{r_n}\left(\frac{1}{n}\sum_{i=1}^{n}V_i^2 1_{|V_i|\leq\varepsilon} + o_{\mathbb{P}\text{ a.s.}}(1)\right).$$

Again, due to Assumption 3, Birkhoff's LLN, the slow variation of L and the defining property of r_n, the rhs of the previous display converges $\mathbb{P}$ a.s. to $\mathbb{E}\left[V_0^2 1_{|V_0|>\varepsilon}\right]$. The result then follows from Assumption 3 by letting $\varepsilon \to +\infty$, since $\mathbb{E}\left[V_0^2\right] < +\infty$. □

6.2 Proofs of main results

Proof of Theorem 1. By the Main Lemma for Sequences in Jakubowski (2012) (see, equivalently, Theorem 1.1 along with Paragraph 3 of Jakubowski (1986)) the result would follow if we would prove that for all $t \in \mathbb{R}$,

$$\prod_{i=1}^{n}\mathbb{E}\left[\exp\left(\mathbf{i}t\frac{1}{n^{\frac{1}{\alpha}}r_n^{\frac{1}{\alpha}}}\rho_{i,\alpha}\right)/\mathcal{G}_{i-1}\right] \tag{18}$$

pointwise converges $\mathbb{P}$ a.s. to the characteristic function of $S_\alpha\left(\beta\left(\frac{\mathbb{E}[|V_0|^\alpha \operatorname{sgn}(V_0)]}{\mathbb{E}(|V_0|^\alpha)}\right), c\mathbb{E}\left(|V_0|^\alpha\right), 0\right)$, where

$$\rho_{i,\alpha} = \left\{\begin{array}{cc}(\xi_i-\gamma)V_i, & \alpha\neq 1\\ (\xi_i-\gamma-H(nr_n))V_i - r_n 2\beta c\pi^{-1}\left(C\mathbb{E}\left(V_0\right) - \mathbb{E}\left[V_0|\log(|V_0|)|\right]\right), & \alpha = 1.\end{array}\right\}$$

For any $t \neq 0$, by defining the event

$$C_{n,K} \equiv \left\{\omega\in\Omega : |V_i| \leq K_t\left(nr_n\right)^{\frac{1}{\alpha}} \quad, \forall i = 1,\dots,n\right\}$$

where $K_t < \frac{M_t}{|t|}$, for some sequence $0 < M_t \to 0$ which exists due to Lemma 1, we have that $\mathbb{P}(C_{n,K}^c) \to 0$ again by Lemma 1. When $\alpha \neq 1$, due to Assumption 1 if $\omega \in C_{n,K}$ then the logarithm of (18) equals

$$\begin{array}{c}-\frac{c|t|^\alpha}{nr_n}\sum_{i=1}^{n}|V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}\left|tV_i\right|^{-1}\right) + \\ \frac{|t|^\alpha}{nr_n}\mathbf{i}\beta c\operatorname{sgn}(t)\tan\left(\frac{1}{2}\pi\alpha\right)\sum_{i=1}^{n}\operatorname{sgn}(V_i)\left|V_i\right|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}\left|tV_i\right|^{-1}\right).\end{array} \tag{19}$$

When $\alpha < 2$, due to Lemma 1, Assumption 1, the asymptotic representation of the tail of the distribution of $|\xi_0|$ (see Appendix 1 in Ibragimov and Linnik (1971)), and the definition of r_n, we have that

$$\frac{\mathbb{P}_n^\star[v_n|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]} = \frac{\mathbb{E}_n^\star\left[\mathbb{P}\left[|\xi_0|>\frac{n^{1/\alpha}r_n^{1/\alpha}}{u_n}\right]\right]}{\mathbb{P}\left[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}\right]} = \frac{nr_n}{(1+o(1))}\mathbb{E}_n^\star\left[\frac{(1+o(1))u_n^\alpha}{nr_n}\frac{L\left(n^{1/\alpha}r_n^{1/\alpha}|V_i|^{-1}\right)}{L\left(n^{1/\alpha}r_n^{1/\alpha}\right)}\right]$$
$$= \frac{(1+o(1))}{(1+o(1))^2 r_n}\frac{1}{n}\sum_{i=1}^n |V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|V_i|^{-1}\right),$$

and the latter is asymptotically equivalent to $\frac{1}{nr_n}\sum_{i=1}^n |V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|V_i|^{-1}\right)$, which, due to Lemma 1, the fact that t is fixed and non-zero, and that L is slowly varying to infinity, is asymptotically equivalent to $\frac{1}{nr_n}\sum_{i=1}^n |V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|tV_i|^{-1}\right)$. In an analogous manner it is easy to show that $\frac{\mathbb{P}_n^\star[v_n 1\{v_n\geq 0\}|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]} - \frac{\mathbb{P}_n^\star[|v_n|1\{v_n<0\}|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}$ is asymptotically equivalent to $\frac{1}{nr_n}\sum_{i=1}^n \mathrm{sgn}(V_i)|V_i|^\alpha L\left(n^{1/\alpha}r_n^{1/\alpha}|tV_i|^{-1}\right)$. Hence, (19) is asymptotically equivalent to

$$-c|t|^\alpha\frac{\mathbb{P}_n^\star[v_n|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}+$$
$$|t|^\alpha \mathbf{i}\beta c\,\mathrm{sgn}(t)\tan\left(\tfrac{1}{2}\pi\alpha\right)\left[\frac{\mathbb{P}_n^\star[v_n 1\{v_n\geq 0\}|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]} - \frac{\mathbb{P}_n^\star[|v_n|1\{v_n<0\}|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}{\mathbb{P}[|\xi_0|>n^{1/\alpha}r_n^{1/\alpha}]}\right].$$

When $\alpha = 2$, (19) is asymptotically equivalent to

$$-c|t|^2\sum_{i=1}^n \mathbb{E}\left[\frac{1}{nr_n}(\xi_i-\gamma)^2 V_i^2 1_{|(\xi_i-\gamma)V_i|>M\sqrt{nr_n}}/\mathcal{G}_{i-1}\right].$$

The result then follows from Propositions 1 and 2 respectively.

When $\alpha = 1$, by Assumption 1, if $\omega \in C_{n,K}$ then the logarithm of (18) equals

$$-c|t|\frac{1}{nr_n}\sum |V_i| L\left(nr_n|tV_i|^{-1}\right) + \mathbf{i}2\beta c\pi^{-1}Ct\frac{1}{nr_n}\sum V_i L\left(nr_n|tV_i|^{-1}\right)$$
$$+\mathbf{i}t\frac{1}{nr_n}\sum V_i\left[H\left(nr_n|tV_i|^{-1}\right) - H(nr_n)\right],$$

where the first two terms of the above expression can be treated analogously to the above in order to obtain their $\mathbb{P}$ a.s. limit as

$$-c|t|\mathbb{E}[|V_0|] + \mathbf{i}2\beta c\pi^{-1}Ct\mathbb{E}[V_0] = -c\mathbb{E}[|V_0|]|t|\left[1 - \mathbf{i}2\beta\pi^{-1}C\mathrm{sgn}(t)\frac{\mathbb{E}[V_0]}{\mathbb{E}[|V_0|]}\right].$$

For the third term, first notice that

$$H(k\lambda) - H(\lambda) = \int_1^k \frac{\lambda^2 x}{1+\lambda^2 x^2}\left(c_1 - c_2 + k(\lambda x)\right)L(\lambda x)dx.$$

Then we have that

$$\frac{1}{nr_n}\sum V_i\left[H\left(nr_n|tV_i|^{-1}\right) - H(nr_n)\right]$$
$$= \frac{L(nr_n)}{r_n}\left(2\beta c\pi^{-1} + o(1)\right)\frac{1}{n}\sum V_i\int_1^{|tV_i|^{-1}}\frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{L(xnr_n)}{L(nr_n)}dx,$$

since for any constant A, $\sup_{x\in[(\max|V_i|)^{-1},A]} k(nr_n|t|^{-1}x) = k(nr_n|t|^{-1}x_n^*)$ for some x_n^*. Lemma 1 implies that $(\max|V_t|)^{-1}n^{1/\alpha}r_n^{1/\alpha} \to \infty$ in $\mathbb{P}$ probability, hence we obtain that $k(nr_n|t|^{-1}x_n^*) = o(1)$. Furthermore using a similar to the above truncation argument we have that

$$\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{L(xnr_n)}{L(nr_n)}dx$$
$$=\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}dx$$
$$+\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\left(\frac{L(xnr_n)}{L(nr_n)}-1\right)dx.$$

Then notice that for some $A_2=[a_1,a_2]$ with $0<a_1\le a_2$ and possibly dependent on the choice of ε,

$$\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\left(\frac{L(xnr_n)}{L(nr_n)}-1\right)dx$$
$$\le\int_{x\in A_2}\frac{x}{\frac{1}{n^2r_n^2}+x^2}\left|\frac{L(xnr_n)}{L(nr_n)}-1\right|dx\frac{1}{n}\sum_{i=1}^n |V_i|\,1\{|V_i|\le\varepsilon\},$$

and the dominant part of the previous display converges to zero $\mathbb{P}$ a.s. via the use of the Dominated Convergence Theorem and Assumption 2. Regarding the first term, first notice that

$$\int_1^{|tV_i|^{-1}}\frac{x}{\frac{1}{n^2r_n^2}+x^2}dx1\{|V_i|\le\varepsilon\}=\frac{1}{2}\left[\log\left(1+n^2r_n^2x^2\right)\right]_1^{|tV_i|^{-1}}1\{|V_i|\le\varepsilon\}$$
$$=\frac{1}{2}\log\left(\frac{1+n^2r_n^2|tV_i|^{-2}}{1+n^2r_n^2}\right)1\{|V_i|\le\varepsilon\}=\log|tV_i|^{-1}1\{|V_i|\le\varepsilon\}+o(1),$$

where the $o(1)$ term is independent of V_i using the fact that

$$\sup_{x\in[(t\varepsilon)^{-1},\infty)}\left|\log\left(\frac{1+\lambda^2x}{1+\lambda^2}\right)-\log x\right|\to 0 \text{ as } \lambda\to+\infty.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{h(xnr_n)}{h(nr_n)}dx$$
$$=\log\frac{1}{|t|}\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|\le\varepsilon\}-\frac{1}{n}\sum_{i=1}^n V_i\log|V_i|1\{|V_i|\le\varepsilon\}+o(1).$$

Next, we treat the analogous term obtained by truncating $V_i1\{|V_i|>\varepsilon\}$, i.e.

$$\frac{1}{n}\sum_{i=1}^n V_i 1\{|V_i|>\}\int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{L(xnr_n)}{L(nr_n)}dx$$

by noticing that $|tV_i|^{-1} < 1$, since ε can be chosen large enough. Suppose first that Assumption 4.1 holds. Notice that,

$$\left|\int_{|tV_i|^{-1}}^{1} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\left(\frac{L(xnr_n)}{L(nr_n)}-1\right)dx\right| \leq \int_{|tV_i|^{-1}}^{1} \frac{x^{1-\delta}}{\frac{1}{n^2r_n^2}+x^2}x^{\delta}\left|\frac{L(xnr_n)}{L(nr_n)}-1\right|dx$$
$$= o\left(1\right)\int_{|tV_i|^{-1}}^{1} \frac{x^{1-\delta}}{\frac{1}{n^2r_n^2}+x^2}dx \leq o\left(1\right)\int_{|tV_i|^{-1}}^{1} x^{-(1+\delta)}dx = o\left(1\right)\left(1+|V_i|^{\delta}\right),$$

where the second inequality follows from the UCT for regularly varying functions with positive index (see Theorem 1.5.2 of Bingham et al. (1989)). Hence, and due to that $\frac{x}{\frac{1}{n^2r_n^2}+x^2} \leq \frac{1}{x}$ for all x, we obtain that

$$\left|\frac{1}{n}\sum_{i=1}^{n} V_i 1\{|V_i|>\varepsilon\}\int_{1}^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{h(xnr_n)}{h(nr_n)}dx\right|$$
$$\leq o\left(1\right)\frac{1}{n}\sum_{i=1}^{n}|V_i|\left(1+|V_i|^{\delta}\right)1\{|V_i|>\varepsilon\}+\frac{|\ln|t||}{n}\sum_{i=1}^{n}|V_i|\ln|tV|\,1\{|V_i|>\varepsilon\},$$

and the result follows by letting $n \to +\infty$ and then $\varepsilon \to +\infty$.

Suppose now that Assumption 4.2 holds. Notice that,

$$\int_{|tV_i|^{-1}}^{1} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{L(xnr_n)}{L(nr_n)}dx = \frac{1}{L(nr_n)}\int_{|tV_i|^{-1}}^{1} \frac{x^2}{\frac{1}{n^2r_n^2}+x^2}nr_n\frac{1}{xnr_n}L(xnr_n)dx$$
$$= \frac{nr_n}{L(nr_n)}\int_{|tV_i|^{-1}}^{1} \frac{x^2}{\frac{1}{n^2r_n^2}+x^2}\mathbb{P}\left[|\xi_0|>xnr_n\right]dx$$
$$= \frac{1}{L(nr_n)}\int_{nr_n|tV_i|^{-1}}^{nr_n} \frac{u^2}{1+u^2}\mathbb{P}\left[|\xi_0|>u\right]du = \int_{nr_n|tV_i|^{-1}}^{nr_n} \frac{u}{1+u^2}\frac{L(u)}{L(nr_n)}du. \tag{20}$$

Then, since with probability converging to one $nr_n|tV_i|^{-1} \geq 1$, $\frac{u}{1+u^2} \leq \frac{1}{u}$ for all u, the rhs integral in (20) is for ε large enough with the same probability less than or equal to

$$\limsup_{n\to+\infty}\sup{}_{1\leq y\leq nr_n}\frac{L(y)}{L(nr_n)}\int_{nr_n|tV_i|^{-1}}^{nr_n}\frac{1}{u}du \leq C\ln|tV_i|\,.$$

Hence, as before we obtain that with probability converging to one

$$\left|\frac{1}{n}\sum_{i=1}^{n} V_i 1\{|V_i|>\varepsilon\}\int_{1}^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2r_n^2}+x^2}\frac{h(xnr_n)}{h(nr_n)}dx\right| \leq C\frac{\ln|t|}{n}\sum_{i=1}^{n}|V_i|\,|\ln|V_i||\,1\{|V_i|>\varepsilon\},$$

and the result follows again by letting $n \to +\infty$ and then $\varepsilon \to +\infty$.

For the cases where either Assumption 5.1 or Assumption 5.2 hold, we have that

$$H\left(nr_n|tV_i|^{-1}\right)-H\left(nr_n\right) = -\left(2\beta c\pi^{-1}+o(1)\right)\int_{1}^{|tV_i|} \frac{x}{\frac{|tV_i|^2}{n^2r_n^2}+x^2}\frac{L(xnr_n|tV_i|^{-1})}{L(nr_n)}dx,$$

and that due to that $\frac{x^2}{\frac{|tV_i|^2}{n^2 r_n^2}+x^2} \leq 1$ uniformly in n, x and i,

$$\int_1^{|tV_i|} \frac{x}{\frac{|tV_i|^2}{n^2 r_n^2}+x^2} \frac{L(xnr_n\,|tV_i|^{-1})}{L(nr_n)} dx \leq |V_i|^{-1} \int_1^{|tV_i|} \frac{\mathbb{P}\left[|V_i|\,|\xi_0| > \frac{x}{|t|} nr_n\right]}{\mathbb{P}\left[|\xi_0| > \frac{x}{|t|} nr_n\right]} dx.$$

Notice that the results in the proofs of Propositions 2.2-3 of Denisov and Zwart (2007), concerning upper bounds for $\frac{\mathbb{P}[|V_i||\xi_0|>\delta nr_n]}{\mathbb{P}[|\xi_0|>\delta nr_n]}$ hold uniformly in δ as long as this is bounded away from zero. Hence the rhs of the previous display is less than or equal to
$\mathbb{E}\left[|V_0|\right]\left(1+|V_i|^{-1}\right)$, and thereby,

$$\left|\frac{1}{n}\sum_{i=1}^{n} V_i 1\{|V_i|>\varepsilon\} \int_1^{|tV_i|^{-1}} \frac{x}{\frac{1}{n^2 r_n^2}+x^2} \frac{h(xnr_n)}{h(nr_n)} dx\right|$$
$$\leq \mathbb{E}\left[|V_0|\right] \frac{1}{n}\sum_{i=1}^{n} |V_i| \left(1+|V_i|^{-1}\right) 1\{|V_i|>\varepsilon\}.$$

The result follows by letting first $n \to +\infty$, and then $\varepsilon \to +\infty$. Combining the above results we obtain (3).

Finally, when $\alpha < 1$ and under (4), observe that

$$\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^{n} |V_i| \leq \frac{1}{n^{1/\alpha-1} r_n^{1/\alpha}} \max_{1\leq i\leq n} |V_i|^{1-\alpha} \frac{1}{n}\sum_{i=1}^{n} |V_i|^{\alpha} \leq O_{\mathbb{P}\text{ a.s.}}(1) \frac{q_n^{\frac{1-\alpha}{\alpha}}}{r_n^{1/\alpha}} M,$$

with $\mathbb{P}$ probability approaching 1 as $n \to \infty$. The result follows as we can choose M arbitrarily small. Under (5), note that for δ small enough (so that $\alpha+\delta<1$), $\left(\frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^{n} |V_i|\right)^{\alpha+\delta} \leq \frac{1}{n^{\frac{\delta}{\alpha}} r_n^{1+\frac{\delta}{\alpha}}} \frac{1}{n} \sum_{i=1}^{n} |V_i|^{\alpha+\delta}$ and the result in (6) follows since $n^k r_n \to \infty$ for any $k>0$ due to that r_n is slowly varying. □

Proof of Theorem 4. Given Theorem 3, and since θ_0 is an interior point, we have that the arguments that lead to equations (5.62-63) in Straumann (2004) are also valid here. Then for $\alpha \in (1,2]$ we have that with probability converging to one

$$c_n''(\bar{\theta}_n)(\theta_n-\theta_0) = \frac{1}{n}\sum_{i=1}^{n}(z_i^2-1)\frac{h_i'(\theta_0)}{\sigma_i^2} \Rightarrow$$
$$n^{\frac{\alpha-1}{\alpha}} r_n^{-\frac{1}{\alpha}}(\theta_n-\theta_0) = \left[c_n''(\bar{\theta}_n)\right]^{-1} \frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \sum_{i=1}^{n}(z_i^2-1)\frac{h_i'(\theta_0)}{\sigma_i^2}.$$

Condition D.4 implies that $c_n''(\bar{\theta}_n) \rightsquigarrow \mathbf{J}_{\theta_0}$ due to the uniform Birkhoff's LLN. Also, as mentioned earlier, condition D.4 implies that $\mathbb{E}\left[\|h_0'(\theta_0)\|^2 \sigma_0^{-4}\right] < \infty$, thus Condition 2 of Proposition 4 holds. Thus, the result follows by an application of Theorem 1.

Analogously for the case where $\alpha = 1$ the result follows by noting that

$$\frac{1}{nr_n}\sum_{i=1}^{n}(z_i^2-1)\frac{h_i'(\theta_0)}{\sigma_i^2} = \frac{1}{nr_n}\sum_{i=1}^{n}(z_i^2-1-\gamma-H(nr_n))\frac{h_i'(\theta_0)}{\sigma_i^2}$$
$$+ \frac{\gamma+H(nr_n)}{r_n}\frac{1}{n}\sum_{i=1}^{n}\frac{h_i'(\theta_0)}{\sigma_i^2}.$$

Then, condition D.4 implies that

$$\mathbb{E}\left[\left|\sigma_0^{-2}\lambda^T h_0'(\theta_0)\log\left|\sigma_0^{-2}\lambda^T h_0'(\theta_0)\right|\right|\right] < +\infty \quad \forall\lambda\in\mathbb{R}^d,$$

hence the $\alpha = 1$ case of Theorem 1 is directly applicable. □

References

Aaronson, J. and M. Denker (1998): "Characteristic functions of random variables attracted to 1-stable laws," *The Annals of Probability*, 26, 399–415.

Abadir, K. and J. Magnus (2004): "03.6.1 The Central Limit Theorem for Student's Distribution–Solution," *Econometric Theory*, 20, 1261–1263.

Adams, T. M. and A. B. Nobel (2010): "Uniform convergence of Vapnik-Chervonenkis classes under ergodic sampling," *The Annals of Probability*, 1345–1367.

Andersen, T. G., R. A. Davis, J.-P. Kreiss, and T. V. Mikosch (2009): *Handbook of financial time series*, Springer Science & Business Media.

Andrews, D. W. K. (1999): "Estimation When a Parameter is on a Boundary," *Econometrica*, 67, 1341–1383.

Arvanitis, S. (2019): "Stable limit theory for the Gaussian QMLE in a non-stationary asymmetric GARCH model," *Statistics & Probability Letters*, 145, 166–172.

Arvanitis, S. and A. Louka (2015): "Limit Theory for the QMLE of the GQARCH (1, 1) model," *Communications in Statistics-Theory and Methods*, 44, 3549–3575.

——— (2017): "Stable limits for the Gaussian QMLE in the non-stationary GARCH (1, 1) model," *Economics Letters*, 161, 135–137.

Bartkiewicz, K., A. Jakubowski, T. Mikosch, and O. Wintenberger (2011): "Stable limits for sums of dependent infinite variance random variables," *Probability theory and related fields*, 150, 337–372.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989): *Regular variation*, vol. 27, Cambridge university press.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bougerol, P. and N. Picard (1992): "Strict stationarity of generalized autoregressive processes," *The Annals of Probability*, 1714–1730.

Breiman, L. (1965): "On Some Limit Theorems Similar to the Arc-Sin Law," *Theory Probab. Appl.*, 10, 323–331.

BULDYGIN, V., O. KLESOV, AND J. STEINEBACH (2013): "Equivalent monotone versions of PRV functions," *Journal of Mathematical Analysis and Applications*, 401, 526–533.

CHANDA, K. C. (1974): "Strong mixing properties of linear stochastic processes," *Journal of Applied Probability*, 11, 401–408.

CORNEA-MADEIRA, A. AND R. DAVIDSON (2015): "A parametric bootstrap for heavy-tailed distributions," *Econometric Theory*, 31, 449–470.

DAVIDSON, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricicans (Advanced Texts in Econometrics)*, Oxford University Press.

DAVIS, R. A. AND T. HSING (1995): "Point process and partial sum convergence for weakly dependent random variables with infinite variance," *The Annals of Probability*, 879–917.

DENISOV, D. AND B. ZWART (2007): "On a theorem of Breiman and a class of random difference equations," *Journal of Applied Probability*, 44, 1031–1046.

DING, Z., C. W. GRANGER, AND R. F. ENGLE (1993): "A long memory property of stock market returns and a new model," *Journal of empirical finance*, 1, 83–106.

EMBRECHTS, P., C. KLÜPPELBERG, AND T. MIKOSCH (2013): *Modelling extremal events: for insurance and finance*, vol. 33, Springer Science & Business Media.

FRANCQ, C. AND J.-M. ZAKOÏAN (2013): "Inference in nonstationary asymmetric GARCH models," *The Annals of Statistics*, 41, 1970–1998.

GOLDIE, C. M. (1991): "Implicit renewal theory and tails of solutions of random equations," *The Annals of Applied Probability*, 126–166.

GUPTA, A. K., T. T. NGUYEN, AND W.-B. ZENG (1994): "Conditions for Stability of Laws with All Projections Stable," *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, 56, 438–443.

HALL, P. AND C. C. HEYDE (2014): *Martingale limit theory and its application*, Academic press.

HALL, P. AND Q. YAO (2003): "Inference in ARCH and GARCH Models with Heavy-Tailed Errors," *Econometrica*, 71, 285–317.

HEWITT, E. AND K. STROMBERG (2013): *Real and abstract analysis: a modern treatment of the theory of functions of a real variable*, Springer-Verlag.

HILL, J. B. (2010): "A New Moment Bound and Weak Laws for Mixingale Arrays without Memory or Heterogeneity Restrictions, with Applications to Tail Trimmed Arrays," *University of North Carolina'Chapel Hill.*

IBRAGIMOV, I. A. AND Y. V. LINNIK (1971): *Independent and stationary sequences of random variables*, Wolters-Noordhoff.

Jakubowski, A. (1986): "Principle of Conditioning in Limit Theorems for Sums of Random Variables," *The Annals of Probability*, 14, 902–915.

——— (2012): "Principle of Conditioning revisited," *Demonstratio Mathematica*, 45, 325–36.

Jeganathan, P. (1982): "A solution of the martingale central limit problem, Part I," *Sankhyā: The Indian Journal of Statistics, Series A*, 299–318.

Kechagias, S. and V. Pipiras (2015): "Definitions and representations of multivariate long-range dependent time series," *Journal of Time Series Analysis*, 36, 1–25.

Klüppelberg, C. (1988): "Subexponential distributions and integrated tails," *Journal of Applied Probability*, 25, 132–141.

Kulik, R. and P. Soulier (2020): *Heavy-tailed time series*, Springer.

Loretan, M. and P. C. Phillips (1994): "Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets," *Journal of Empirical Finance*, 1, 211 – 248.

Mikosch, T. and D. Straumann (2006): "Stable Limits of Martingale Transforms with Application to the Estimation of Garch Parameters," *The Annals of Statistics*, 34, 493–522.

Mittnik, S., S. T. Rachev, and M. S. Paolella (1998): "Stable Paretian modeling in finance: Some empirical and theoretical aspects," *A Practical Guide to Heavy Tails*, 79–110.

Molchanov, I. (2011): *Theory of Random Sets (Probability and Its Applications)*, Springer.

Nelson, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Phillips, P. C. and T. Magdalinos (2007): "Limit theory for moderate deviations from a unit root," *Journal of Econometrics*, 136, 115–130.

Rachev, S. and S. Mittnik (1988): *Stable Paretian Models in Finance, Series in Financial Economics and Quantitative Analysis*, John Wiley & Sons Ltd., England.

Rockafellar, R. T. (1997): *Variational Analysis*, Springer.

Samorodnitsky, G. and M. S. Taqqu (1994): *Stable non-Gaussian processes*, Chapman and Hall, New York.

Sentana, E. (1995): "Quadratic ARCH Models," *The Review of Economic Studies*, 62, 639–661.

STRAUMANN, D. (2004): *Estimation in Conditionally Heteroscedastic Time Series Models*, Springer.

STRAUMANN, D. AND T. MIKOSCH (2006): "Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach," *The Annals of Statistics*, 34, 2449–2495.

SURGAILIS, D. (2008): "A Quadratic ARCH(∞) Model with Long Memory and Lévy Stable Behavior of Squares," *Advances in Applied Probability*, 40, 1198–1222.

VAN DER VAART, A. W. (2000): *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*, Cambridge University Press.

WANG, Q. (2014): "Martingale Limit Theorem Revisited and Nonlinear Cointegrating Regression," *Econometric Theory*, 30, 509–535.

WINTENBERGER, O. (2013): "Continuous Invertibility and Stable QML Estimation of the EGARCH(1,1) Model," *Scandinavian Journal of Statistics*, 40, 846–867.

WINTENBERGER, O. AND S. CAI (2011): "Parametric inference and forecasting in continuously invertible volatility models," *arXiv preprint arXiv:1106.4983*.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ — ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS — ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ / SCHOOL OF ECONOMIC SCIENCES — ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ / DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2021

01-21 **Historical Cycles of the Economy of Modern Greece From 1821 to the Present, George Alogoskoufis**
02-21 **Greece Before and After the Euro: Macroeconomics, Politics and the Quest for Reforms, George Alogoskoufis**
03-21 **Commodity money and the price level, George C. Bitros. Published in: *Quarterly Journal of Austrian Economics*, 2022**
04-21 **Destabilizing asymmetries in central banking: With some enlightenment from money in classical Athens, George C. Bitros. Published in: *Journal of Economic Asymmetries*, 2021**
05-21 **Exploring the Long-Term Impact of Maximum Markup Deregulation, Athanasios Dimas and Christos Genakos**
06-21 **A regularization approach for estimation and variable selection in high dimensional regression models, Y. Dendramis, L. Giraitis, G. Kapetanios**
07-21 **Tax Competition in the Presence of Environmental Spillovers, Fabio Antoniou, Panos Hatzipanayotou, Michael S. Michael, Nikos Tsakiris**
08-21 **Firm Dynamics by Age and Size Classes and the Choice of Size Measure, Stelios Giannoulakis and Plutarchos Sakellaris**
09-21 **Measuring the Systemic Importance of Banks, Georgios Moratis, Plutarchos Sakellaris**
10-21 **Firms' Financing Dynamics Around Lumpy Capacity Adjustments, Christoph Görtz, Plutarchos Sakellaris, John D. Tsoukalas**
11-21 **On the provision of excludable public goods General taxes or user prices? George Economides and Apostolis Philippopoulos**
12-21 **Asymmetries of Financial Openness in an Optimal Growth Model, George Alogoskoufis**
13-21 **Evaluating the impact of labour market reforms in Greece during 2010-2018 , Georgios Gatopoulos, Alexandros Louka, Ioannis Polycarpou, Nikolaos Vettas**
14-21 **From the Athenian silver to the bitcoin standard: Private money in a state-enforced free banking model, George C. Bitros**
15-21 **Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities. Stelios Arvanitis and Thierry Post**
16-21 **Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka**
17-21 **Competition and Pass-Through: Evidence from Isolated Markets, Christos Genakos and Mario Pagliero**
18-21 **Exploring Okun's Law Asymmetry: An Endogenous Threshold LSTR Approach, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis**
19-21 **Limit Theory for Martingale Transforms with Heavy-Tailed Multiplicative Noise, Stelios Arvanitis and Alexandros Louka**
20-21 **Optimal taxation with positional considerations, Ourania Karakosta and Eleftherios Zacharias**

21-21 **The ECB's policy, the Recovery Fund and the importance of trust: The case of Greece, Vasiliki Dimakopoulou, George Economides and Apostolis Philippopoulos**

2022

01-22 **Is Ireland the most intangible intensive economy in Europe? A growth accounting perspective, Ilias Kostarakos, KieranMcQuinn and Petros Varthalitis**
02-22 **Common bank supervision and profitability convergence in the EU, Ioanna Avgeri, Yiannis Dendramis and Helen Louri**
03-22 **Missing Values in Panel Data Unit Root Tests, Yiannis Karavias, Elias Tzavalis and Haotian Zhang**
04-22 **Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities, Stelios Arvanitis and Thierry Post**
05-22 **Concentration Inequalities for Kernel Density Estimators under Uniform Mixing, Stelios Arvanitis**
06-22 **Public Sector Corruption and the Valuation of Systemically Important Banks, Georgios Bertsatos, Spyros Pagratis, Plutarchos Sakellaris**
07-22 **Finance or Demand: What drives the Responses of Young and Small Firms to Financial Crises? Stelios Giannoulakis and Plutarchos Sakellaris**
08-22 **Production function estimation controlling for endogenous productivity disruptions, Plutarchos Sakellaris and Dimitris Zaverdas**
09-22 **A panel bounds testing procedure, Georgios Bertsatos, Plutarchos Sakellaris, Mike G. Tsionas**
10-22 **Social policy gone bad educationally: Unintended peer effects from transferred students, Christos Genakos and Eleni Kyrkopoulou**
11-22 **Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka**
12-22 **Time to question the wisdom of active monetary policies, George C. Bitros**
13-22 **Investors' Behavior in Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou and Georgios Tsomidis**
14-22 **On the asking price for selling Chelsea FC, Georgios Bertsatos and Gerassimos Sapountzoglou**
15-22 **Hysteresis, Financial Frictions and Monetary Policy, Konstantinos Giakas**
16-22 **Delay in Childbearing and the Evolution of Fertility Rates, Evangelos Dioikitopoulos and Dimitrios Varvarigos**
17-22 **Human capital threshold effects in economic development: A panel data approach with endogenous threshold, Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis**
18-22 **Distributional aspects of rent seeking activities in a Real Business Cycle model, Tryfonas Christou, Apostolis Philippopoulos and Vanghelis Vassilatos**

2023

01-23 **Real interest rate and monetary policy in the post Bretton Woods United States, George C. Bitros and Mara Vidali**
02-23 **Debt targets and fiscal consolidation in a two-country HANK model: the case of Euro Area, Xiaoshan Chen, Spyridon Lazarakis and Petros Varthalitis**
03-23 **Central bank digital currencies: Foundational issues and prospects looking forward, George C. Bitros and Anastasios G. Malliaris**
04-23 **The State and the Economy of Modern Greece. Key Drivers from 1821 to the Present, George Alogoskoufis**
05-23 **Sparse spanning portfolios and under-diversification with second-order stochastic dominance, Stelios Arvanitis, Olivier Scaillet, Nikolas Topaloglou**

06-23 **What makes for survival? Key characteristics of Greek incubated early-stage startup(per)s during the Crisis: a multivariate and machine learning approach, Ioannis Besis, Ioanna Sapfo Pepelasis and Spiros Paraskevas**

07-23 **The Twin Deficits, Monetary Instability and Debt Crises in the History of Modern Greece, George Alogoskoufis**

08-23 **Dealing with endogenous regressors using copulas; on the problem of near multicollinearity, Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis**

09-23 **A machine learning approach to construct quarterly data on intangible investment for Eurozone, Angelos Alexopoulos and Petros Varthalitis**

10-23 **Asymmetries in Post-War Monetary Arrangements in Europe: From Bretton Woods to the Euro Area, George Alogoskoufis, Konstantinos Gravas and Laurent Jacque**

11-23 **Unanticipated Inflation, Unemployment Persistence and the New Keynesian Phillips Curve, George Alogoskoufis and Stelios Giannoulakis**

12-23 **Threshold Endogeneity in Threshold VARs: An Application to Monetary State Dependence, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis**

13-23 **A DSGE Model for the European Unemployment Persistence, Konstantinos Giakas**

14-23 **Binary public decisions with a status quo: undominated mechanisms without coercion, Efthymios Athanasiou and Giacomo Valletta**

15-23 **Does Agents' learning explain deviations in the Euro Area between the Core and the Periphery? George Economides, Konstantinos Mavrigiannakis and Vanghelis Vassilatos**

16-23 **Mild Explocivity, Persistent Homology and Cryptocurrencies' Bubbles: An Empirical Exercise, Stelios Arvanitis and Michalis Detsis**

17-23 **A network and machine learning approach to detect Value Added Tax fraud, Angelos Alexopoulos, Petros Dellaportas, Stanley Gyoshev, Christos Kotsogiannis, Sofia C. Olhede, Trifon Pavkov**

18-23 **Time Varying Three Pass Regression Filter, Yiannis Dendramis, George Kapetanios, Massimiliano Marcellino**

19-23 **From debt arithmetic to fiscal sustainability and fiscal rules: Taking stock, George Economides, Natasha Miouli and Apostolis Philippopoulos**

20-23 **Stochastic Arbitrage Opportunities: Set Estimation and Statistical Testing, Stelios Arvanitis and Thierry Post**

21-23 **Behavioral Personae, Stochastic Dominance, and the Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou, and Georgios Tsomidis**

22-23 **Block Empirical Likelihood Inference for Stochastic Bounding: Large Deviations Asymptotics Under *m*-Dependence, Stelios Arvanitis and Nikolas Topaloglou**

23-23 **A Consolidation of the Neoclassical Macroeconomic Competitive General Equilibrium Theory via Keynesianism (Part 1 and Part 2), Angelos Angelopoulos**

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/