

1 ASYMPTOTIC INFERENCE FOR AR 2 MODELS WITH HEAVY-TAILED 3 G-GARCH NOISES

4 RONGMAO ZHANG
5 *Zhejiang University*

6 SHIQING LING
7 *Hong Kong University of Science and Technology*

8 It is well known that the least squares estimator (LSE) of an AR(p) model with i.i.d.
9 (independent and identically distributed) noises is $n^{1/\alpha}L(n)$ -consistent when the
10 tail index α of the noise is within $(0, 2)$ and is $n^{1/2}$ -consistent when $\alpha \geq 2$, where
11 $L(n)$ is a slowly varying function. When the noises are not i.i.d., however, the case
12 is far from clear. This paper studies the LSE of AR(p) models with heavy-tailed
13 G-GARCH(1,1) noises. When the tail index α of G-GARCH is within $(0, 2)$, it is
14 shown that the LSE is not a consistent estimator of the parameters, but converges to a
15 ratio of stable vectors. When $\alpha \in [2, 4]$, it is shown that the LSE is $n^{1-2/\alpha}$ -consistent
16 if $\alpha \in (2, 4)$, log n -consistent if $\alpha = 2$, and $n^{1/2}/\log n$ -consistent if $\alpha = 4$, and its
17 limiting distribution is a functional of stable processes. Our results are significantly
18 different from those with i.i.d. noises and should warn practitioners in economics
19 and finance of the implications, including inconsistency, of heavy-tailed errors in
20 the presence of conditional heterogeneity.

Query 3

21 1. INTRODUCTION

22 Since the seminal work by Engle (1982) and Bollerslev (1986), the G/ARCH-
23 type models have been extensively applied in economics and finance. This paper
24 considers the following strictly stationary autoregressive [AR(p)] process:

$$25 Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t, \quad (1.1)$$

26 where $\{\varepsilon_t : t = 1, 2, \dots\}$ is generated by the general GARCH(1,1) process
[G-GARCH(1,1)]:

$$\varepsilon_t = \eta_t h_t \text{ and } h_t^\delta = g(\eta_{t-1}) + c(\eta_{t-1})h_{t-1}^\delta, \quad (1.2)$$

We thank Ms. Alice Cheng for her editing comments and three referees, the co-editor Giuseppe Cavaliere, and the editor Peter C.B. Phillips for their very helpful and professional comments. Zhang's research was supported by NSFC grants 11371318 and 11171074, the Fundamental Research Funds for the Central Universities, and Scientific Research Fund of Zhejiang Provincial Education Department (Y201009944). Ling's research was supported by the Hong Kong Research Grants Council (Grants HKUST641912, 603413 and FSGRF12SC12).

Query 2

2 RONGMAO ZHANG AND SHIQING LING

1 where $\delta > 0$, $Pr\{h_t^\delta > 0\} = 1$, $c(0) < 1$, $c(\cdot)$ and $g(\cdot)$ are nonnegative functions,
2 and $\{\eta_t\}$ is a sequence of i.i.d. (independent and identically distributed) symmetric
3 white noises. The general model (1.2) was defined by He and Terasvirta (1999).
4 It includes many models as special cases, for example, the GARCH(1,1) model
5 of Bollerslev (1986), the absolute value GARCH(1,1) model of Taylor (1986) and
6 Schwert (1989), the nonlinear GARCH(1,1) model of Engle (1990), the volatil-
7 ity switching GARCH(1,1) model of Fornari and Mele (1997), the threshold
8 GARCH(1,1) model of Zakoian (1994), and the generalized quadratic ARCH(1,1)
9 model of Sentana (1995).

10 Let $\phi = (\phi_1, \dots, \phi_p)'$ be the unknown parameter vector and its true value be
11 ϕ_0 . When ε_t is i.i.d. (i.e., h_t is a constant) with $E\varepsilon_t^2 = \infty$, the estimated ϕ_0 has
12 been well studied in the literature. Hannan and Kanter (1977) proved the least
13 squares estimator (LSE) of ϕ_0 is $n^{1/\nu}$ -consistent, where $\nu > \alpha$ and $\alpha \in (0, 2)$ is
14 the tail index of ε_t (see also Knight, 1987) and n is the sample size. The same rate
15 of convergence was obtained by An and Chen (1982) for the least absolute devia-
16 tion (LAD) estimator of ϕ_0 . The limiting distribution of the LSE was not available
17 until Davis and Resnick (1986). Based on a point process technique and assum-
18 ing that ε_t has a regular varying tail index α , they showed that under condition
19 $\lim_{t \rightarrow \infty} P(|\varepsilon_1 \varepsilon_2| > t) / P(|\varepsilon_1| > t) = 2E|\varepsilon_1|^\alpha < \infty$, the LSE converges weakly to
20 a ratio of two stable random variables at the rate $n^{1/\alpha} L(n)$, where $L(n)$ is a slowly
21 varying function. The asymptotic theory of the LAD and M-estimators of ϕ_0 was
22 fully established by Davis, Knight, and Liu (1992). We refer to Mikosch, Gadrich,
23 Klüppelberg, and Adler (1995) and Kokoszka and Taqqu (1996) for infinite vari-
24 ance ARMA and long-memory ARFIMA models. Up to date, it is well known
25 that all the classical estimators have a faster rate of convergence when $\alpha \in (0, 2)$
26 than those when $\alpha \geq 2$ if the noises are i.i.d. However, when the noises are not
27 i.i.d., the case is far from clear. A few exceptions can be found in Mikosch and
28 Stărică (2000), Lange (2011), and the references therein.

29 In this paper, we show that due to the dependence, the cross-product terms re-
30 lated to $\varepsilon_t \varepsilon_{t-j}$, $j > 0$ do not vanish asymptotically, and the limiting distribution
31 of the sample autocovariance $\{Y_t Y_{t-k}\}$ and $\{Y_{t-k} \varepsilon_t\}$ depends on an infinite num-
32 ber of point processes, and thus it differs substantially from that in Davis and
33 Resnick (1986) for i.i.d. noises. When the tail index α of G-GARCH(1,1) noise is
34 within $(0, 2)$, $\sum_{t=1}^n Y_t Y_{t-k}$ and $\sum_{t=1}^n Y_{t-k} \varepsilon_t$ have the same rate of convergence.
35 As a result, the LSE is not a consistent estimator of the parameters, but tends to
36 be a function of stable vectors. However, when $\alpha > 2$, the rate of convergenc
37 of $\sum_{t=1}^n Y_t Y_{t-k}$ is controlled by its centralized constant whose rate is faster than
38 that of $\sum_{t=1}^n Y_{t-k} \varepsilon_t$ (see Lemma 3.2). This leads to the consistency of the LSE.
39 In particular, when $\alpha \in [2, 4]$, it is shown that the LSE is $n^{1-2/\alpha}$ -consistent if
40 $\alpha \in (2, 4)$, $\log n$ -consistent if $\alpha = 2$, and its limiting distribution is a stable ran-
41 dom vector. When $\alpha = 4$, the LSE is $n^{1/2}/\log n$ -consistent and asymptotically
42 normal. Our results are significantly different from those with i.i.d noises and
43 should warn practitioners in economics and finance of the implications, including
44 inconsistency of heavy-tailed errors in the presence of conditional heterogeneity.

1 This paper is organized as follows. The main results are provided in Section 2
 2 and the technical proofs are given in Section 3. Simulation results and additional
 3 proofs are reported in the online supplementary material.

4 2. MAIN RESULTS

5 Given observations Y_1, \dots, Y_n , the LSE of ϕ_0 for model (1.1) is defined by

$$\hat{\phi}_n = \left(\sum_{i=p+1}^n \mathbf{Y}_{i-1} \mathbf{Y}_{i-1}^T \right)^{-1} \left(\sum_{i=p+1}^n \mathbf{Y}_{i-1} Y_i \right), \quad (2.1)$$

6 where $\mathbf{Y}_i = (Y_i, Y_{i-1}, \dots, Y_{i-p+1})$. Throughout the paper, we make the following
 7 assumptions:

- 8 **H1.** $E\log(c(\eta_t)) < 0$.
- 9 **H2.** There exists a $k_0 > 0$ such that $E(c(\eta_t))^{k_0} \geq 1$, $E[(c(\eta_t))^{k_0} \log^+(c(\eta_t))] < \infty$,
 10 and $E[g(\eta_t) + |\eta_t|^\delta]^{k_0} < \infty$, where $\log^+(x) = \max\{0, \log(x)\}$.
- 11 **H3.** The density $f(x)$ of η_1 is positive in the neighborhood of zero.

12 Condition H1 is a necessary and sufficient condition for the existence of a
 13 stationary solution of h_t^2 (see Nelson, 1990). If condition H2 holds, then con-
 14 dition H1 is equivalent to $E(c(\eta_1))^\mu < 1$ for some $\mu > 0$ (see Remark 2.9 of
 15 Basrak, Davis, and Mikosch, 2002). H3 also implies that h_t is not a constant
 16 and hence excludes the i.i.d. case. Suppose that there exists a $t_0 > 0$ such that
 17 $E|\eta_t|^{t_0} = \infty$ and $E|\eta_t|^t < \infty$ for all $t < t_0$, then conditions H1 and H2 are sat-
 18 isfied. Condition H3 is a mixing condition for model (1.2) and can be relaxed
 19 to some certain (see Francq and Zakoian, 2006). Furthermore, the symmetry as-
 20 sumption is used to simplify the proof for the case when $\alpha = 2$. Note that, when
 21 $\alpha = 2$, $\varepsilon_t \varepsilon_{t-j}$, $j \geq 1$ (see Lemma 3.1), is a regularly varying variable with tail
 22 index $\alpha/2 = 1$. If η_t is not symmetric, then a centralizing constant is required to
 23 derive the limiting distribution $\sum_{t=p}^n \varepsilon_t \varepsilon_{t-j}$. We first give a lemma for the tail
 24 index of $\{\varepsilon_t\}$. Its proof is similar to those of Lemmas A.1 and A.3 in Chan and
 25 Zhang (2010).

Query 4

26 LEMMA 2.1. *Under conditions H1, H2, and H3, there exists a unique*
 27 $\alpha \in (0, \delta k_0]$ *such that $E(c(\eta_t))^{\alpha/\delta} = 1$ and*

$$P(|\varepsilon_1| > x) \sim c_0^{(\alpha)} E|\eta_1|^\alpha x^{-\alpha},$$

28 where

$$c_0^{(\alpha)} = \frac{E \left([g(\eta_1) + c(\eta_1)\sigma_1^\delta]^{\alpha/\delta} - [c(\eta_1)\sigma_1^\delta]^{\alpha/\delta} \right)}{\alpha E \left(c(\eta_1)^{\alpha/\delta} \log^+(c(\eta_1)) \right)}.$$

4 RONGMAO ZHANG AND SHIQING LING

We further make the following assumption:

H4. $\phi(z) = 1 - \sum_{i=1}^p \phi_i z^i \neq 0$ for $|z| \leq 1$. Under condition H4, model (1.1) is stationary and has the following expansion:

$$Y_t = \sum_{l=0}^{\infty} \varphi_l \varepsilon_{t-l}, \quad (2.2)$$

and $\Sigma \equiv E[Y_1 Y_1^T]$ exists and is positive definite when $\alpha > 2$.

Let $\xrightarrow{\mathcal{L}}$ denote convergence in distribution. We now state our main results as follows:

THEOREM 2.1. Let α be given as in Lemma 2.1. Then under conditions H1–H4, it follows that

(a) when $\alpha \in (0, 2)$,

$$\hat{\phi}_n - \phi_0 \xrightarrow{\mathcal{L}} \Sigma_{\alpha/2}^{-1} Z_{\alpha/2},$$

where $Z_{\alpha/2}$ is a p -dimensional stable vector with index $\alpha/2$ and $\Sigma_{\alpha/2}$ is a $p \times p$ matrix whose elements are composed of stable variables with index $\alpha/2$;

(b) when $\alpha = 2$,

$$\log n (\hat{\phi}_n - \phi_0) \xrightarrow{\mathcal{L}} \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|} \right)^{-1}_{p \times p} Z_{\alpha/2};$$

(c) when $2 < \alpha < 4$,

$$n^{1-2/\alpha} (\hat{\phi}_n - \phi_0) \xrightarrow{\mathcal{L}} \Sigma^{-1} Z_{\alpha/2};$$

(d) when $\alpha = 4$,

$$(n/\log n)^{1/2} (\hat{\phi}_n - \phi_0) \xrightarrow{\mathcal{L}} \Sigma^{-1} N(0, A),$$

where $A = (c_0^{(4)} E \eta_1^2)(a_{ij})_{p \times p}$ is positive definite with $a_{ij} = \lim_{M \rightarrow \infty} E [u_{t,i,M} u_{t,j,M}]$ and

$$u_{t,i,M} = \sum_{l=i}^M \varphi_{l-i} \eta_{t-l} \prod_{i=1}^l c^{\frac{1}{\delta}}(\eta_{t-i}) \prod_{k=l+1}^M c^{\frac{2}{\delta}}(\eta_{t-k}).$$

We should mention that when

$$h_t^2 = \omega + a \varepsilon_{t-1}^2 + \beta h_{t-1}^2, \quad (2.3)$$

where $\omega > 0$, $a > 0$, and $\beta > 0$, Chan and Zhang (2010) showed that for all $\alpha > 0$, the rate of convergence of the LSE in the unit-root case is of order n

and (c) was obtained by Lange (2011). Our theorem completely characterizes the feature of the LSE under a more general setup when $E\varepsilon_t^4 = \infty$. Many empirical examples have shown the evidence that the fourth moment does not exist in economics and finance (see Mikosch and Stărică, 2000). Theorem 2.1 indicates that the statistical inference based on the LSE may be misleading and it is necessary to consider other approaches in this case. In fact, the tail trimming QMLE in Hill and Renault (2010) and modified QMLE in Lange, Rahbek, and Jensen (2011) are \sqrt{n} -consistent and asymptotically normal for models (1.1) and (1.2). Furthermore, if one can specify the form of G-GARCH model as (2.3), then the self-weighted QMLE in Ling (2007) and Zhu and Ling (2011) is also \sqrt{n} -consistent and asymptotically normal. When $\alpha < 2$, we conjecture that the M-estimators of Knight (1991), the LAD asymptotics of Phillips (1991), or the dummy-based estimators of Cavalier and Georgiev (2013) with some kind of weights can achieve a fast rate of convergence.

The nonstandard results are mainly because of condition H2 which generates the heavy tails of volatility h_t . Let us consider the special T-CHARM model proposed by Chan, Li, Ling, and Tong (2012):

$$h_t = \sigma_1 I\{\varepsilon_{t-1} > r\} + \sigma_2 I\{\varepsilon_{t-1} \leq r\}. \quad (2.4)$$

We can rewrite (2.4) as follows:

$$h_t = (\sigma_1 + \sigma_2 a_{t-1} - \sigma_1 b_{t-1}) + (b_{t-1} - a_{t-1})h_{t-1},$$

where $a_t = I\{\eta_t \sigma_1 \leq r\}$ and $b_t = I\{\eta_t \sigma_2 \leq r\}$. We can see that $c(\eta_{t-1}) = b_{t-1} - a_{t-1}$ does not satisfy condition H2. Under (2.4), it is not difficult to show that

$$\sqrt{n}(\hat{\phi}_n - \phi_0) \xrightarrow{\mathcal{L}} N(0, \Sigma^{-1}\Omega\Sigma^{-1}),$$

where $\Omega = E[\mathbf{Y}_1 \mathbf{Y}_1^T h_t]$. This means that the classical statistical inference for the AR model is always valid under (2.4) specification. Otherwise, one needs to pay special attention to the tail index α of ε_t . The tail index α of ε_t is unknown in practice, but it is identical to that of Y_t . We can estimate it by using Hill's estimator.

3. TECHNICAL PROOFS

In this section, we prove Theorem 2.1. Note that

$$\hat{\phi}_n - \phi = \left(\sum_{t=p+1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\sum_{t=p+1}^n \mathbf{Y}_{t-1} \varepsilon_t \right].$$

The limiting distribution of $\hat{\phi}_n - \phi$ will follow by the asymptotic behaviors of $\sum_{t=p+1}^n Y_{t-k} \varepsilon_t$ and $\sum_{t=p+1}^n Y_{t-k} Y_{t-j}$ for $1 \leq k, j \leq p$. Intuitively, by equation (2.2), we know that $\sum_{t=p+1}^n Y_{t-k} \varepsilon_t$ can be approximated by

6 RONGMAO ZHANG AND SHIQING LING

¹ $\sum_{l=0}^H \varphi_l \sum_{t=p+1}^n \varepsilon_{t-k-l} \varepsilon_t$ and $\sum_{t=p+1}^n Y_{t-k} Y_{t-j}$ can be approximated by
² $\sum_{l=0}^H \sum_{m=0}^H \varphi_l \varphi_m \sum_{t=p+1}^n \varepsilon_{t-k-l} \varepsilon_{t-j-m}$ as $H \rightarrow \infty$. So, by the continuous map-
³ ping theorem, it is enough to show:

⁴ (a) The union convergence of $(\sum_{t=p+1}^n \varepsilon_{t-k-l} \varepsilon_{t-j-m}, l = 1, 2, \dots, H)$. This
⁵ will be proved in Lemma 3.1 by using a point process convergence technique.

⁶ (b) The limiting distribution of $a_n^{-2} \lim_{H \rightarrow \infty} \sum_{l=0}^H \varphi_l \sum_{t=p+1}^n \varepsilon_{t-k-l} \varepsilon_t$ and
⁷ $a_n^{-2} \lim_{H \rightarrow \infty} \sum_{l=0}^H \sum_{m=0}^H \varphi_l \varphi_m \sum_{t=p+1}^n \varepsilon_{t-k-l} \varepsilon_{t-j-m}$. This will be proved in
⁸ Lemmas 3.2 and 3.3.

⁹ Denote $a_n^{(\alpha)} = (c_0^{(\alpha)} E|\eta_1|^\alpha n)^{1/\alpha}$. For simplicity, we write $c_0 = c_0^{(\alpha)}$, $a_n = a_n^{(\alpha)}$
¹⁰ and assume $E\eta_1^2 = 1$ when $\alpha \geq 2$.

¹¹ For any given integers l and H , we define two $(H+1)$ -dimensional random
¹² vectors:

$$\mathbf{X}_{t,l,H} = (\varepsilon_{t-l}, \dots, \varepsilon_{t-l-H}) \text{ and}$$

$$\mathbf{Z}_{t,l,H} = (\varepsilon_{t-l}^2 - c_n, \varepsilon_{t-l} \varepsilon_{t-l-1}, \dots, \varepsilon_{t-l} \varepsilon_{t-l-H}), \quad (3.1)$$

¹³ where $c_n = 0$ as $0 < \alpha < 2$, $c_n = E\varepsilon_1^2 I_{(|\varepsilon_1| \leq \sqrt{c_0 n})} = c_0 \log n$ as $\alpha = 2$, and $c_n = E\varepsilon_1^2$
¹⁴ as $\alpha > 2$. Under the assumptions of Theorem 2.1, $\{\mathbf{X}_{t,l,H}, t \geq 1\}$ is a varying reg-
¹⁵ular random vector sequence with index α . By Theorem 2.8 of Davis and Mikosch
¹⁶ (1998) (see also Theorem 3.1 of Mikosch and Stărică, 2000), there exists a Pois-
¹⁷son process $\sum_{i=1}^{\infty} \delta_{P_i}$ defined on \mathbb{R}_+ with intensity measure $v(dy) = \Upsilon \alpha y^{-\alpha-1} dy$
¹⁸ and a sequence of i.i.d. point processes $\{\sum_{j=1}^{\infty} \delta_{Q_{ij,H}}\}$ with distribution Q_H ,
¹⁹ which depends on α , such that

$$\sum_{t=1}^n \delta_{\mathbf{X}_{t,l,H}/a_n} \xrightarrow{\mathcal{L}} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \delta_{P_i} Q_{ij,H},$$

²⁰ where $\mathbf{Q}_{ij,H} = (Q_{ij}^0, Q_{ij}^1, \dots, Q_{ij}^H)$, $\{\sum_{j=1}^{\infty} \delta_{Q_{ij,H}}\}$ is independent of the process
²¹ $\{P_i\}$, and Υ and Q_H are similarly defined as in Davis and Mikosch (1998). Thus,
²² by Proposition 3.3 of Davis and Mikosch (1998) and the continuous mapping
²³ theorem, we have

$$\begin{aligned} \frac{1}{a_n^2} \sum_{t=1}^n \mathbf{Z}_{t,j,H} &\xrightarrow{\mathcal{L}} \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 (Q_{ij}^0)^2, \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^0 Q_{ij}^1, \dots, \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^0 Q_{ij}^H \right) \\ &=: (S^0(\alpha), S^1(\alpha), \dots, S^H(\alpha)) = S_H(\alpha), \end{aligned} \quad (3.2)$$

²⁴ if $0 < \alpha < 2$ and

$$\begin{aligned} \frac{1}{a_n^2} \sum_{t=1}^n \mathbf{Z}_{t,j,H} &\xrightarrow{\mathcal{L}} \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^0 Q_{ij}^h I(|P_i^2 Q_{ij}^0 Q_{ij}^h| > 0) \right. \\ &\quad \left. - \int_{(\mathbf{x}: |x_0 x_h| > 0)} x_0 x_h d\mu(\mathbf{x}) \right)_{h=0, \dots, H} \\ &=: (S^0(\alpha), S^1(\alpha), \dots, S^H(\alpha)) = S_H(\alpha), \end{aligned} \quad (3.3)$$

1 if $2 \leq \alpha < 4$, where $\mu(\cdot) = \lim_{n \rightarrow \infty} n P(\mathbf{Z}_{t,j,H}/n^{\frac{2}{\alpha}} \in \cdot)$ is a measure on \mathbb{R}^{H+1}
 2 and $\mathbf{x} = (x_0, x_1, \dots, x_H)$. $\mathbf{S}_H(\alpha)$ is a stable random vector with index $\alpha/2$.

3 LEMMA 3.1. *Under the conditions of Theorem 2.1, for any positive integer K and H ,*

$$\begin{aligned} \frac{1}{a_n^2} \left(\sum_{t=1}^n \mathbf{Z}_{t,0,H}, \sum_{t=1}^n \mathbf{Z}_{t,1,H}, \dots, \sum_{t=1}^n \mathbf{Z}_{t,K,H} \right) \mathcal{L} \\ \longrightarrow (\mathbf{S}_H(\alpha), \mathbf{S}_H(\alpha), \dots, \mathbf{S}_H(\alpha))_{1 \times (K+1)}, \end{aligned}$$

5 as $n \rightarrow \infty$, where $0 < \alpha < 4$ and $\mathbf{S}_H(\alpha)$ is defined in (3.2) and (3.3).

6 **Proof.** We first have the expansion:

$$h_t^\delta = \sum_{i=1}^h \prod_{j=1}^{i-1} c(\eta_{t-j}) g(\eta_{t-i}) + \prod_{i=1}^h c(\eta_{t-i}) h_{t-h}^\delta, \quad (3.4)$$

7 where $\prod_{j=1}^0 c(\eta_{t-j}) = 1$. Furthermore, we have

$$|\varepsilon_t \varepsilon_{t-h}|^\delta = |\eta_t|^\delta \sum_{i=1}^h \prod_{j=1}^{i-1} c(\eta_{t-j}) g(\eta_{t-i}) |\eta_{t-h}|^\delta h_{t-h}^\delta + |\eta_t|^\delta \prod_{i=1}^h c(\eta_{t-i}) |\eta_{t-h}|^\delta h_{t-h}^{2\delta}.$$

8 By Lemma 2.1 and Proposition 3 of Breiman (1965), we have that

$$\begin{aligned} \lim_{y \rightarrow \infty} P \left\{ |\eta_t|^\delta \sum_{i=1}^h \prod_{j=1}^{i-1} c(\eta_{t-j}) g(\eta_{t-i}) |\eta_{t-h}|^\delta h_{t-h}^\delta > y \right\} \\ = c(\alpha) E \left\{ |\eta_t|^\delta \sum_{i=1}^h \prod_{j=1}^{i-1} c(\eta_{t-j}) g(\eta_{t-i}) |\eta_{t-h}|^\delta \right\}^{\alpha/\delta} y^{-\alpha/\delta} =: c_0(\alpha) y^{-\alpha/\delta} \end{aligned}$$

9 and

$$\begin{aligned} \lim_{y \rightarrow \infty} P \left\{ |\eta_t|^\delta \prod_{i=1}^h c(\eta_{t-i}) |\eta_{t-h}|^\delta h_{t-h}^{2\delta} > y \right\} \\ = c(\alpha) E[n_0^2 c(\eta_0)]^{\frac{\alpha}{2\delta}} [E c(\eta_0)^{\frac{\alpha}{2\delta}}]^{h-1} [E |\eta_0|^{\frac{\alpha}{\delta}}]^h y^{-\frac{\alpha}{2\delta}}, \\ := c_1(\alpha) \rho^h y^{-\frac{\alpha}{2\delta}}, \text{ for some } 0 < \rho < 1. \end{aligned}$$

10 Combining the previous two equations, we have, for any given h and a large enough x ,

$$P(|\varepsilon_t \varepsilon_{t-h}| > x) = P(|\varepsilon_t \varepsilon_{t-h}|^\delta > x^\delta) \simeq c_1(\alpha) \rho^h x^{-\alpha/2}. \quad (3.5)$$

12 For any $0 \leq i < j \leq K$, (3.5) yields that

$$\frac{1}{a_n^2} \sum_{t=1}^n (\mathbf{Z}_{t,i,H} - \mathbf{Z}_{t,j,H}) \xrightarrow{p} 0.$$

8 RONGMAO ZHANG AND SHIQING LING

Thus, by (3.2) and (3.3), we can show that the conclusion holds. ■

¹ Let $S^i(\alpha)$ be given as in (3.2) for $\alpha < 2$ and (3.3) for $2 \leq \alpha < 4$ and define

$$Z_{\alpha/2}^{(k)}(Y) = \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+k} \right) S^0(\alpha) + \sum_{h=1}^{\infty} \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+h+k} + \sum_{l=0 \vee (k-h)}^{\infty} \varphi_l \varphi_{l+h-k} \right) S^h(\alpha)$$

³ and $Z_{\alpha/2}^{(k)}(\varepsilon) = \sum_{h=k+1}^{\infty} \varphi_{h-k-1} S^h(\alpha)$. We have the following lemma.

⁴ **LEMMA 3.2.** *Under the conditions of Theorem 2.1, we have for $0 < \alpha < 4$,*

$$\begin{aligned} & \frac{1}{a_n^2} \left\{ \sum_{t=1}^n \left[Y_t Y_{t-k} - \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+k} \right) c_n \right], \sum_{t=1}^n Y_{t-1-k} \varepsilon_t, 0 \leq k \leq p \right\} \\ & \xrightarrow{\mathcal{L}} \left\{ Z_{\alpha/2}^{(k)}(Y), Z_{\alpha/2}^{(k)}(\varepsilon), 0 \leq k \leq p \right\}, \end{aligned}$$

⁵ where c_n is defined as in (3.1).

⁶ **Proof.** By Cramér–Wold's lemma, it is sufficient to show that, for any real
⁷ number $f_k, g_k, 0 \leq k \leq p$,

$$\begin{aligned} & \frac{1}{a_n^2} \left\{ \sum_{k=0}^p \sum_{t=1}^n f_k \left[Y_t Y_{t-k} - \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+k} \right) c_n \right] + \sum_{k=0}^p \sum_{t=1}^n g_k Y_{t-1-k} \varepsilon_t \right\} \\ & \xrightarrow{\mathcal{L}} \sum_{k=0}^p f_k Z_{\alpha/2}^{(k)}(Y) + \sum_{k=0}^p g_k Z_{\alpha/2}^{(k)}(\varepsilon). \end{aligned} \tag{3.6}$$

⁸ Let $d_j = \sum_{k=0}^{j \wedge p} f_k \varphi_{j-k}$ and $e_j = \sum_{k=0}^{j \wedge p} g_k \varphi_{j-k}$. By (2.2), we have

$$\begin{aligned} & \sum_{k=0}^p f_k \left[Y_t Y_{t-k} - \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+k} \right) c_n \right] + \sum_{k=0}^p g_k Y_{t-1-k} \varepsilon_t \\ & = \sum_{j=-\infty}^t \sum_{h=-\infty, h \neq j}^t d_{t-h} \varphi_{t-j} \varepsilon_h \varepsilon_j + \sum_{h=-\infty, h \neq j}^{t-1} e_{t-1-h} \varepsilon_h \varepsilon_t + \sum_{j=0}^{\infty} d_j \varphi_j (e_{t-j}^2 - c_n) \\ & = \left\{ \sum_{j=0}^K d_j \varphi_j (e_{t-j}^2 - c_n) + \sum_{h=1}^H \sum_{j=0}^K (d_{j+h} \varphi_j + d_j \varphi_{j+h}) \varepsilon_{t-j-h} \varepsilon_{t-j} + \sum_{h=1}^H e_{h-1} \varepsilon_{t-h} \varepsilon_t \right\} \\ & \quad + \left\{ \left(\sum_{h=H+1}^{\infty} \sum_{j=0}^{\infty} + \sum_{h=1}^H \sum_{j=K+1}^{\infty} \right) (d_{j+h} \varphi_j + d_j \varphi_{j+h}) \varepsilon_{t-j-h} \varepsilon_{t-j} \right. \\ & \quad \left. + \sum_{j=K+1}^{\infty} d_j \varphi_j (e_{t-j}^2 - c_n) + \sum_{h=H+1}^{\infty} e_{h-1} \varepsilon_{t-h} \varepsilon_t \right\} \\ & =: I_t^1(H, K) + I_t^2(H, K). \end{aligned}$$

1 Thus, by Lemma 3.1, along the lines of the proof for Theorem 3.1 of Zhang, Sin,
 2 and Ling (2013), we can show that

$$\begin{aligned} \frac{1}{a_n^2} \sum_{t=1}^n I_t^1(H, K) &\xrightarrow{\mathcal{L}} \sum_{l=0}^K d_l \varphi_l S^0(\alpha) + \sum_{h=1}^H \left[\sum_{l=0}^K (d_{h+l} \varphi_l + d_l \varphi_{h+l}) + e_{h-1} \right] S^h(\alpha) \\ &\xrightarrow{p} \sum_{k=0}^p f_k Z_{\alpha/2}^{(k)}(Y) + \sum_{k=0}^p g_k Z_{\alpha/2}^{(k)}(\varepsilon) \end{aligned}$$

3 by letting $H \rightarrow \infty$ and $K \rightarrow \infty$ and for any $\delta > 0$

$$\lim_{H, K \rightarrow \infty} \lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{a_n^2} \sum_{t=1}^n I_t^2(H, K) \right| > \delta \right\} = 0.$$

Combining the previous equations, we can see that (3.6) holds. ■

4

5 LEMMA 3.3. Let A be the matrix given in Theorem 2.1. If $\alpha = 4$, then

$$\frac{1}{\sqrt{n \log n}} \sum_{t=1}^n \{Y_{t-k} \varepsilon_t, 1 \leq k \leq p\} \xrightarrow{\mathcal{L}} N(0, A). \quad (3.7)$$

6 **Proof.** Since the proof is extremely technical, we put it in the supplementary material. ■

7

8 **The proof of Theorem 2.1** (i) When $0 < \alpha < 2$,

$$\widehat{\phi}_n - \phi = \left(\frac{1}{a_n^2} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\frac{1}{a_n^2} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \varepsilon_t \right].$$

9 By Lemma 3.2 and continuous mapping theorem, (a) holds. (ii) When $\alpha = 2$,

$$(\log n)(\widehat{\phi}_n - \phi) = \left(\frac{1}{c_0 n \log n} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\frac{1}{c_0 n} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \varepsilon_t \right].$$

10 By Lemma 3.2, (b) holds. (iii) When $2 < \alpha < 4$,

$$\frac{n}{a_n^2} (\widehat{\phi}_n - \phi) = \left(\frac{1}{n} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\frac{1}{a_n^2} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \varepsilon_t \right].$$

1 By Lemma 3.2, (c) holds. (iv) When $\alpha = 4$,

$$\frac{n}{\sqrt{n \log n}} (\hat{\phi}_n - \phi) = \left(\frac{1}{n} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\frac{1}{\sqrt{n \log n}} \sum_{t=p+1}^n \mathbf{Y}_{t-1} \varepsilon_t \right].$$

2 By the ergodic theorem and Lemma 3.3, (d) holds. This completes the proof. ■

3

4 REFERENCES

- 5 An, H.Z. & Z.G. Chen (1982) On convergence of LAD estimates in autoregression with infinite variance. *Journal of Multivariate Analysis* 12, 335–345.
- 6 Basrak, B., R.A. Davis, & T. Mikosch (2002) Regular variation of GARCH processes. *Stochastic Processes and Their Applications* 99, 95–115.
- 7 Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- 8 Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory of Probability and Its Applications* 10, 323–331.
- 9 Query 5 Cavaliere, G. & I. Georgiev (2013) Exploiting infinite variance through dummy variables in non-stationary autoregressions. *Econometric Theory*, to appear.
- 10 Chan, K.S., D. Li, S. Ling, & H. Tong (2012) Threshold modeling of martingale differences. Submitted.
- 11 Chan, N.H. & R.M. Zhang (2010) Inference for unit-root models with infinite variance GARCH errors. *Statistica Sinica* 20, 1363–1393.
- 12 Davis, R.A., K. Knight, & J. Liu (1992) M-estimation for autoregressions with infinite variance. *Stochastic Processes and Their Applications* 40, 145–180.
- 13 Davis, R.A. & T. Mikosch (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics* 26, 2049–2080.
- 14 Davis, R.A. & S. Resnick (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* 14, 533–558.
- 15 Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica* 50, 987–1008.
- 16 Engle, R.F. (1990) Discussion: Stock market volatility and the crash of 1987. *Review of Financial Studies* 3, 103–106.
- 17 Fornari, F. & A. Mele (1997) Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics* 12, 1779–1801.
- 18 Francq, C. & J.M. Zakoïan (2006) Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22, 815–834.
- 19 Hannan, E.J. & M. Kanter (1977) Autoregressive processes with infinite variance. *Journal of Applied Probability* 14, 411–415.
- 20 He, C. & T. Teräsvirta (1999) Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.
- 21 Hill, J.B. & E. Renault (2010) Generalized Method of Moments with Tail Trimming. Working paper, University of North Carolina.
- 22 Knight, K. (1987) Rate of convergence of centred estimates of autoregressive parameters for infinite variance autoregressions. *Journal of Time Series Analysis* 8, 51–60.
- 23 Knight, K. (1991) Limit theory for M-estimates in an integrated infinite variance. *Econometric Theory* 7, 200–212.
- 24 Kokoszka, P.S. & M.S. Taqqu (1996) Parameter estimation for infinite variance fractional ARIMA. *Annals of Statistics* 24, 1880–1913.

- 1 Lange, T. (2011) Tail behavior and OLS estimation in AR-GARCH models. *Statistica Sinica* 21,
2 1191–1200.
- 3 Lange, T., A. Rahbek, & S.T. Jensen (2011) Estimation and asymptotic inference in the AR-ARCH
4 model. *Econometric Reviews* 30, 129–153.
- 5 Ling, S. (2007) Self-weighted and local quasi-maximum likelihood estimators for ARMA-
6 GARCH/IGARCH models. *Journal of Econometrics* 140, 849–873.
- 7 Mikosch, T., T. Gadrich, C. Klüppelberg, & R.J. Adler (1995) Parameter estimation for ARMA models
8 with infinite variance innovations. *Annals of Statistics* 23, 305–326.
- 9 Mikosch, T. & C. Stărică (2000) Limit theory for the sample autocorrelations and extremes of a
10 GARCH (1,1) process. *Annals of Statistics* 28, 1427–1451.
- 11 Nelson, D.B. (1990) Stationary and persistence in GARCH (1,1) models. *Econometric Theory* 6,
12 318–334.
- 13 Phillips, P.C.B. (1991) A shortcut to LAD asymptotics. *Econometric Theory* 7, 450–463.
- 14 Schwert, G.W. (1989) Why does stock market volatility change over time? *Journal of Finance* 45,
15 1129–1155.
- 16 Sentana, E. (1995) Quadratic ARCH models. *Review of Economic Studies* 62, 639–661.
- 17 Taylor, S. (1986) *Modelling Financial Time Series*. Wiley.
- 18 Zakoian, J.M. (1994) Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*
19 18, 931–955.
- 20 Zhang, R.M., C.Y. Sin, & S. Ling (2013) Limit theories of linear processes of possibly heavy-tailed
21 GARCH(1,1) errors. *Stochastic Processes and Their Applications*, Revised.
- 22 Zhu, K. & S. Ling (2011) Global self-weighted and local quasi-maximum exponential likelihood
23 estimators for ARMA–GARCH/IGARCH models. *Annals of Statistics* 39, 2131–2163.

Author Query Form

Journal: Econometric Theory

Article: S0266466614000632

Dear Author, During the preparation of your manuscript for publication, the questions listed below have arisen. The numbers pertain to the numbers in the margin of the proof. Please attend to these matters and return the form with this proof.

Many thanks for your assistance.

Query no.	Query	Response
Q1	Please check that the suggested running head is correct.	
Q2	Please provide corresponding author details for this article.	
Q3	Please check that the inserted definition for “i.i.d.” is correct.	
Q4	The meaning of the sentence “Condition H3 is...” is not clear. Please clarify.	
Q5	Please update the references “Cavaliere & Georgiev (2013),” “Chan et al. (2012),” and “Zhang et al. (2013).”	