
Volume 138 Issue 11 1 November 2008 ISSN 0378-3758
ELSEVIER
journal of statistical planning and inference
Special Issue in Honor of Junjiro Ogawa (1915–2000): Design of Experiments, Multivariate Analysis and Statistical Inference
Edited by: Takeshi Hayakawa
Sanpei Kageyama
Madan L. Puri
For Contents see page iii
Available online at
ScienceDirect
www.sciencedirect.com

Journal of Statistical Planning and Inference 138 (2008) 3578 – 3589

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Distributional analysis of empirical volatility in GARCH processes

Lajos Horváth[a,*,1], Piotr Kokoszka[b,2], Ričardas Zitikis[c,3]

[a]Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, Utah 84112-0090, USA
[b]Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, Utah 84322-3900, USA
[c]Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7

ARTICLE INFO

Article history:
Received 31 March 2005
Accepted 27 February 2007
Available online 16 March 2008

MSC:
60G35
60F05

Keywords:
GARCH model
Limit theory
Volatility

ABSTRACT

Conditionally heteroskedastic time series given by $y_k = \sigma_k \varepsilon_k$ are frequently used in econometrics. The conditional variance σ_k^2 is defined by a parametric function of past observations and volatilities. Since several conditionally heteroskedastic time series models have been suggested in the literature, we want to test if a given model fits well the data. The method we propose in this paper is based on comparing the distributions of the observed and implied volatilities. Our results can be used to assess the validity of the GARCH(p,q) model.

1. Introduction

Following the seminal work of Engle (1982), the most commonly used model for stationary conditionally heteroskedastic econometric time series can be written as

$$y_k = \sigma_k \varepsilon_k, \tag{1.1}$$

where $\sigma_k \geqslant 0$ is a function of $\{y_j, -\infty < j \leqslant k-1\}$, and the random variables $\{\varepsilon_k, -\infty < k < \infty\}$ are independent and identically distributed. The function defining σ_k is unknown, but its form is typically specified by means of recursive equations. In this paper, we focus on the GARCH(p,q) model (cf. Bollerslev, 1986):

$$\sigma_k^2 = \omega + \sum_{1 \leqslant i \leqslant p} \alpha_i y_{k-i}^2 + \sum_{1 \leqslant j \leqslant q} \beta_j \sigma_{k-j}^2, \tag{1.2}$$

where $\omega > 0$, $\alpha_i \geqslant 0$ and $\beta_j \geqslant 0$ are parameters. There are, of course, many other recursive equations commonly used in the econometric literature (cf., e.g., Carrasco and Chen, 2002), but the GARCH model remains the most important archetypal example. Throughout the paper we assume that $\alpha_i > 0$ and $\beta_j > 0$.

Necessary and sufficient conditions under which the GARCH(p,q) equations have a unique, strictly stationary, and non-anticipative solution were found by Nelson (1990) for $p = 1$ and $q = 1$, and by Bougerol and Picard (1992a,b) for arbitrary

* Corresponding author.
E-mail addresses: horvath@math.utah.edu (L. Horváth), Piotr.Kokoszka@usu.edu (P. Kokoszka), zitikis@stats.uwo.ca (R. Zitikis).

1 Supported by the NSF Grant DMS 060670 and the Grant RGC-HKUST 6428/06H.
2 Supported by the NSF Grants DMS-0413653 and INT-0223262, and the NATO Grant PST.EAP.CLG 980599.
3 Supported by a Discovery Research Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

0378-3758/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2007.02.014

$p \geqslant 1$ and $q \geqslant 1$. (Non-anticipative solution means that $\{\sigma_k^2, k \leqslant i\}$ and $\{\varepsilon_k, k \geqslant i\}$ are independent for all $-\infty < i < \infty$.) We assume throughout the paper that these conditions are satisfied, and we also note that these conditions imply the existence of $\rho < 1$ such that $\beta_1 + \cdots + \beta_q < \rho$, which we use explicitly later.

If we assume (as always done in this context) that $\mathbf{E}(\varepsilon_k^2) = 1$, then the processes $\{y_k\}$ and $\{\sigma_k\}$ have the same variances, provided that their second moments are finite, which we assume throughout. More generally, assuming the GARCH(p, q) model, the volatility of the observed time series can be associated with the behavior of either $\{y_k^2\}$ or $\{\sigma_k^2\}$. In fact, in order to obtain an idea about the volatility of an econometric zero-mean time series, it is customary to examine the plots of $y_1^2, \ldots, y_n^2$, where n is the length of the observed series, and of $\sigma_k^2, \ldots, \sigma_n^2$. The latter series is not observable and is therefore replaced by its estimated counterpart. Zivot and Wang (2003, Chapter 7) give a practical introduction to these ideas.

The objective of this paper is to explore the consequences of Eq. (1.1) by comparing functions of the squared observations y_k^2 to their analogs based on the right-hand side (r.h.s.) of Eq. (1.1). The first of the functions is the sample second moment $n^{-1}\sum_{1 \leqslant k \leqslant n} y_k^2$, and the second function is the empirical distribution function $n^{-1}\sum_{1 \leqslant k \leqslant n} \mathbf{1}\{y_k^2 \leqslant x\}$, where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Since the process $\{y_k^2\}$ carries information about the volatility, the two functions are measures of volatility.

The first issue that we need to address is how to define analogs of the sample second moment and the empirical distribution function using the r.h.s. of Eq. (1.1) and employing only the observations $y_1^2, \ldots, y_n^2$ and the parameters of the GARCH(p, q) process. Such analogs may be used to develop procedures for verifying the validity of assumption (1.1), especially sequential monitoring schemes in which the constancy of the specified model and its parameters is postulated under the null hypothesis. Procedures of this type can be based on differences between the corresponding functionals of $\{y_k^2\}$ and their counterparts corresponding to the r.h.s. of Eq. (1.1). In Section 2, we establish the existence of limiting distributions for these differences. Theoretical justifications of the testing procedures alluded to above—whether based on the asymptotic, Monte Carlo, or various resampling distributions—require the existence of the limits derived in Section 2. This paper focuses on the relevant underlying theory and, due to space limitations, is not particularly concerned with the development of statistical procedures; nevertheless, a small numerical illustration is presented in Section 3.

The paper is organized as follows. Section 2 contains main results of the paper which are briefly illustrated by numerical results in Section 3. Sections 4–7 contain proofs. The proofs are fairly complex, but we think they are instructive. Specifically, we demonstrate how Davydov's (1996) sufficient conditions for weak convergence of general stochastic processes can be utilized in the context of econometric time series. The required moment-type inequalities are proved using classical martingale inequalities as well as the 'near-independent' structure of GARCH models developed by Berkes and Horváth (2001). We believe that this combination of the aforementioned techniques opens up a convenient route for establishing asymptotic results in the theory of GARCH processes.

2. Measures of volatility

Denote the parameter-vector of the GARCH(p, q) process by $\boldsymbol{\theta}$. Under the assumptions stated in Section 1, we can write (cf., e.g., Berkes et al., 2003) the equation $\sigma_k^2 = w_k(\boldsymbol{\theta})$, where

$$w_k(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{1 \leqslant i < \infty} c_i(\boldsymbol{\theta}) y_{k-i}^2. \tag{2.1}$$

The coefficient-functions $c_i(\boldsymbol{\theta}) \geqslant 0$ are deterministic and known; they are defined as solutions of certain recursive equations (cf. Berkes et al., 2003). Let $\mathscr{F}_k$ be the σ-algebra generated by $\{\varepsilon_i, -\infty < i \leqslant k\}$. Since $\mathbf{E}(y_k^2 | \mathscr{F}_{k-1}) = \sigma_k^2$, the expectations of y_k^2 and σ_k^2 coincide. Since the process $\{y_k^2\}$ is stationary, the ergodic theorem implies that, almost surely (a.s.),

$$\frac{1}{n} \sum_{1 \leqslant k \leqslant n} y_k^2 \to \mathbf{E}(y_0^2)[=\mathbf{E}(\sigma_0^2)]. \tag{2.2}$$

Similarly, since $\{w_k(\boldsymbol{\theta})\}$ is stationary, we have that

$$\frac{1}{n} \sum_{1 \leqslant k \leqslant n} w_k(\boldsymbol{\theta}) \left[= \frac{1}{n} \sum_{1 \leqslant k \leqslant n} \sigma_k^2 \right] \to \mathbf{E}(w_k(\boldsymbol{\theta})) \quad \text{a.s.} \tag{2.3}$$

Note, however, that the average on the left-hand side (l.h.s.) of Eq. (2.3) depends on $\{y_i, -\infty < i \leqslant n-1\}$, whereas we observe only $\{y_1, \ldots, y_n\}$. This suggests averaging $s_k^2 = \widetilde{w}_k(\boldsymbol{\theta})$ instead of $w_k(\boldsymbol{\theta})$, where

$$\widetilde{w}_k(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{1 \leqslant i \leqslant k-1} c_i(\boldsymbol{\theta}) y_{k-i}^2.$$

We thus study the difference between $n^{-1}\sum_{1\leqslant k\leqslant n} y_k^2$ and $n^{-1}\sum_{1\leqslant k\leqslant n} s_k^2$. Specifically, in Theorem 2.1 we derive the asymptotic distribution of the process

$$S_n(t)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant nt}(y_k^2-s_k^2),\quad 0\leqslant t\leqslant 1.$$

Theorem 2.1. *Assume that* ε_0 *is non-degenerate and the moments* $\mathbf{E}(\varepsilon_0^4)$ *and* $\mathbf{E}(\sigma_0^4)$ *are finite. Then the process* S_n *converges weakly to* $\tau\mathscr{W}$, *where* $\tau^2=\mathbf{E}(\sigma_0^4)(\mathbf{E}(\varepsilon_0^4)-1)$ *and* $\mathscr{W}$ *is a standard Wiener process.*

The asymptotic variance τ^2 is a known function of the parameter-vector $\boldsymbol{\theta}$. Relevant formulas for GARCH(p,q) and other models from the GARCH family have been derived in He and Teräsvirta (1999, 2004).

To obtain analogous results for the aforementioned empirical distribution function, we observe that the function is an estimator of $\mathbf{P}(y_0^2\leqslant x)$ since the process $\{y_k,-\infty<k<\infty\}$ is strictly stationary and ergodic. Note that $\mathbf{P}(y_0^2\leqslant x)$ is equal to $\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$. The GARCH$(p,q)$ model therefore implies the equality of the expected values of $\mathbf{P}\{y_0^2\leqslant x|\ \varepsilon_i, i<k\}$ and $F_{\varepsilon_0^2}(x/\sigma_0^2)$ for all $x\geqslant 0$.

Constructing an empirical analog for the expectation $\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$ for which a satisfactory asymptotic theory can be developed is a challenge. For this, we first note that $\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$ can be approximated by $n^{-1}\sum_{1\leqslant k\leqslant n}F_{\varepsilon_0^2}(x/w_k(\boldsymbol{\theta}))$. Furthermore, it is natural to expect that the latter arithmetic mean is close to $n^{-1}\sum_{1\leqslant k\leqslant n}F_{\varepsilon_0^2}(x/s_k^2)$. We want to approximate $F_{\varepsilon_0^2}(z)$ using the empirical distribution function $n^{-1}\sum_{1\leqslant k\leqslant n}\mathbf{1}\{\varepsilon_k^2\leqslant z\}$, but ε_k's are not observable. Hence, we rewrite the average as $n^{-1}\sum_{1\leqslant k\leqslant n}\mathbf{1}\{y_k^2/\sigma_k^2\leqslant z\}$ and then replace σ_k^2 by s_k^2. In this way we arrive at the empirical approximation $F_n(z)=n^{-1}\sum_{1\leqslant l\leqslant n}\mathbf{1}\{y_l^2/s_l^2\leqslant z\}$ of $F_{\varepsilon_0^2}(z)$. Given the discussion above, we expect that $n^{-1}\sum_{1\leqslant k\leqslant n}F_n(x/s_k^2)$ is a valid approximation of $\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$. A more tractable asymptotic theory can, however, be derived if instead of $F_n(z)$ we use

$$F_{k,n}(z)=\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\mathbf{1}\left\{\frac{y_l^2}{s_l^2}\leqslant z\right\}$$

and, consequently, consider $n^{-1}\sum_{1\leqslant k\leqslant n}F_{k,n}(x/s_k^2)$ as an approximation to the expectation $\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$. We therefore focus our attention on the process

$$Z_n(x)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\left(\mathbf{1}\{y_k^2\leqslant x\}-F_{k,n}\left(\frac{x}{s_k^2}\right)\right)\tag{2.4}$$

and establish its limiting distribution so that various functionals of interest would be possible to consider and thus statistical inferential results derived. To state assumptions under which we can achieve these goals, we need additional notation. Namely, let $\boldsymbol{\tau}_n=(\beta_1+\alpha_1\varepsilon_n^2,\beta_2,\ldots,\beta_{q-1})\in\mathbf{R}^{q-1}$, $\boldsymbol{\xi}_n=(\varepsilon_n^2,0,\ldots,0)\in\mathbf{R}^{q-1}$, and $\boldsymbol{\alpha}=(\alpha_2,\ldots,\alpha_{p-1})\in\mathbf{R}^{p-2}$. With these vectors, let A_n be a $(p+q-1)\times(p+q-1)$ matrix written in the block form as

$$A_n=\begin{bmatrix}\boldsymbol{\tau}_n & \beta_q & \boldsymbol{\alpha} & \alpha_p\\ I_{q-1} & 0 & 0 & 0\\ \boldsymbol{\xi}_n & 0 & 0 & 0\\ 0 & 0 & I_{p-2} & 0\end{bmatrix},$$

where I_{q-1} and I_{p-2} are the identity matrices of size $q-1$ and $p-2$, respectively. For any matrix M, we use $\|M\|$ to denote the spectral norm $\sup\{\|M\mathbf{x}\|/\|\mathbf{x}\|:\mathbf{x}\neq\mathbf{0}\}$.

Theorem 2.2. *Assume that the density* $f_{|\varepsilon_0|}$ *is bounded, the moments* $\mathbf{E}(\sigma_0^4)$ *and* $\mathbf{E}((\log^+\|A_0\|)^\mu)$ *are finite, for some* $\mu>4$. *Then* Z_n *converges weakly to a mean-zero Gaussian process* Z *whose covariance kernel* $K(x,y)$ *is defined by*

$$K(x,y)=\mathbf{E}\left(F_{\varepsilon_0^2}\left(\frac{x\wedge y}{\sigma_0^2}\right)-F_{\varepsilon_0^2}\left(\frac{x}{\sigma_0^2}\right)F_{\varepsilon_0^2}\left(\frac{y}{\sigma_0^2}\right)\right).$$

Theorem 2.2 shows that even though the difference $F_{y_0^2}(x)-\mathbf{E}(F_{\varepsilon_0^2}(x/\sigma_0^2))$ is zero when the GARCH(p,q) model holds, its empirical analog converges to a non-degenerate distribution at the rate $n^{-1/2}$.

3. Numerical examples

We present here results of a simple numerical experiment related to Theorem 2.1 in which the limiting distribution is normal. The empirical rate of convergence of the normalized random variable $S_n(1)/\tau$ to the standard normal distribution depends on the

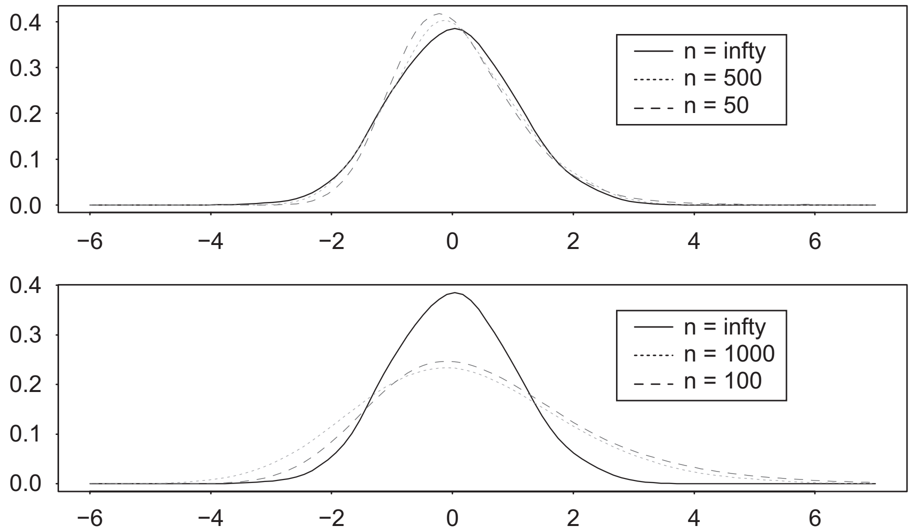

Fig. 1. The densities of $S_n(1)/\tau$ with the parameters $\omega = 0.1$, $\alpha = 0.2$, $\beta = 0$ (top), and $\omega = 0.1$, $\alpha = 0.1$, $\beta = 0.8$ (bottom).

Table 1
The power of the Monte Carlo test based on the statistic $S_n(1)/\tau$ with the parameters: $\omega = 0.1$, $\alpha = 0.1$, $\beta = 0.8$ (null), and $\omega = 0.1$, $\alpha = 0.1$, $\beta = 0.7$ (alternative)

n	500		250		100		50		25	
Size (%)	10	5	10	5	10	5	10	5	10	5
Power	1.000	1.000	1.000	0.995	0.884	0.811	0.6635	0.5240	0.4225	0.2780

parameters of the GARCH process. The smaller the parameters α_i and β_i, the faster the convergence. This is intuitively clear because larger values of α_i and β_i imply not only more dependence, but also heavier tails of the marginal distribution of the process $\{y_k^2\}$ (cf. Mikosch and Stărică, 2000; Basrak et al., 2002). To illustrate this point, Fig. 1 shows the simulated densities of $S_n(1)/\tau$, together with the limiting distribution, for the GARCH(1, 1) model with two sets of parameters. Other simulations also show that large values of β lead to slow rates of convergence.

The simulated distribution of $S_n(1)/\tau$ can, in principle, be used to test if the parameter $\boldsymbol{\theta}$ is equal to a specified value. A test of this type can, for example, be incorporated into a monitoring scheme designed to quickly detect a change in parameters. As noted in Section 1, we are not concerned here with the practical development of such procedures, but merely present a simple illustration. Under the null hypothesis, the observations follow a GARCH(1, 1) model with the parameters $\omega = 0.1$, $\alpha = 0.1$, and $\beta = 0.8$; under the alternative the parameter β changes to $\beta = 0.7$. The test rejects at the nominal significance level s if $S_n(1)/\tau$ is smaller than $Q(s/2)$ or greater than $Q(1 - s/2)$, where $Q(x)$ is the xth quantile of the simulated distribution of $S_n(1)/\tau$ under the null hypothesis. When computing the statistic $S_n(1)/\tau$ from the data, τ is computed assuming the null hypothesis. Such a test has an empirical size equal to the nominal size up to a chance error due to the finite number of replications (we have used two thousands). Table 1 shows that the empirical power is very good. Generally speaking, the theory developed here can be used to provide asymptotic justification for monitoring procedures in which a departure from a model with a specified parameter value must be detected early.

4. Proof of Theorem 2.1

We start with the elementary equation

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant nt} (y_k^2 - w_k(\boldsymbol{\theta})) + \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant nt} (w_k(\boldsymbol{\theta}) - \tilde{w}_k(\boldsymbol{\theta})) \tag{4.1}$$

and show that its right-most term converges to 0. Note that the absolute difference $|w_k(\boldsymbol{\theta}) - \tilde{w}_k(\boldsymbol{\theta})|$, which is $\sum_{k \leqslant i < \infty} c_i(\boldsymbol{\theta}) y_{k-i}^2$, does not exceed $c \sum_{k \leqslant i < \infty} \rho^{i/q} y_{k-i}^2$, which we rewrite as $c\rho^{k/q} \sum_{0 \leqslant j < \infty} \rho^{j/q} y_{-j}^2$. By Lemma 2.2 of Berkes et al. (2003), $\sum_{0 \leqslant j < \infty} \rho^{j/q} y_{-j}^2 < \infty$ a.s. Hence, uniformly over all t, the right-most term of Eq. (4.1) converges to 0 when $n \to \infty$, and so the limit of the process S_n is determined by the first term, which is the normalized partial sum of the martingale differences $\sigma_k^2(\varepsilon_k^2 - 1)$ with respect to the σ-algebras $\mathscr{F}_k$. Using a martingale convergence theorem (cf., e.g., Hall and Heyde, 1980), we have that the first term on the r.h.s. of Eq. (4.1) has the same limit process as the one specified in Theorem 2.1. This concludes the proof. □

Note 4.1. The above proof shows that $\sum_{1\leqslant k<\infty}|\sigma_k^2-s_k^2|<\infty$ a.s. The ergodic theorem yields that $n^{-1}\sum_{1\leqslant k\leqslant n}\sigma_k^2\to\mathbf{E}(\sigma_0^2)$ a.s. These statements imply that $n^{-1}\sum_{1\leqslant k\leqslant n}s_k^2\to\mathbf{E}(\sigma_0^2)$ a.s. Also, with the help of the bound

$$\sum_{k\leqslant l\leqslant n}\frac{\sigma_l^2}{s_l^2}\leqslant(n-k+1)+\frac{1}{c_0(\boldsymbol{\theta})}\sum_{k\leqslant l\leqslant n}|\sigma_l^2-s_l^2|,$$

the above statements imply that $\max_{1\leqslant k\leqslant n}(n-k+1)^{-1}\sum_{k\leqslant l\leqslant n}(\sigma_l^2/s_l^2)$ is bounded in probability when $n\to\infty$. We shall frequently use these 'convergence' statements in the proofs below.

5. Proof of Theorem 2.2

The proof is technically demanding but interesting. We start with the note that instead of the process $\{Z_n(x)\ :\ x\geqslant 0\}$, we equivalently investigate weak convergence of $\{Z_n(x^2):x\geqslant 0\}$, which we write as

$$Z_n(x^2)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\left(\mathbf{1}\left\{|\varepsilon_k|\leqslant\frac{x}{\sigma_k}\right\}-\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\mathbf{1}\left\{|\varepsilon_l|\leqslant\frac{xs_l}{s_k\sigma_l}\right\}\right).$$

Next, we write $Z_n(x^2)=Z_n^*(x)+Z_n^{**}(x)-Z_n^{***}(x)$, where

$$Z_n^*(x)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\left(\mathbf{1}\left\{|\varepsilon_k|\leqslant\frac{x}{\sigma_k}\right\}-F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)\right),$$

$$Z_n^{**}(x)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\left(F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)-\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)\right),$$

$$Z_n^{***}(x)=\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\left(\mathbf{1}\left\{|\varepsilon_l|\leqslant\frac{xs_l}{s_k\sigma_l}\right\}-F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)\right).$$

We show next that $\sup_x|Z_n^{**}(x)|\to_{\mathbf{P}}0$, which reduces weak convergence of Z_n to that of $Z_n^*-Z_n^{***}$. To prove the aforementioned convergence to zero, we split the supremum into two halves: $x\leqslant x_n$ and $x>x_n$, where $x_n=n^{1/3}$. When $x\leqslant x_n$, we use the assumption that $f_{|\varepsilon_0|}$ is bounded, apply the bounds $\sigma_k,s_k\geqslant c_0(\boldsymbol{\theta})>0$ for all k, and have that

$$|Z_n^{**}(x)|\leqslant\frac{cx}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\left|\frac{1}{\sigma_k}-\frac{s_l}{s_k\sigma_l}\right|\leqslant\frac{cx_n\log n}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}|\sigma_k-s_k|. \tag{5.1}$$

Since $\sum_{1\leqslant k<\infty}|\sigma_k^2-s_k^2|<\infty$ a.s. (cf. Note 4.1), we have $\sup_{0\leqslant x\leqslant x_n}|Z_n^{**}(x)|\to_{\mathbf{P}}0$. Hence, we are left with the case $x>x_n$. We estimate $|F_{|\varepsilon_0|}(x_n/\sigma_k)-F_{|\varepsilon_0|}(x_ns_l/(s_k\sigma_l))|$ by the sum of two survival functions. Then we estimate the survival functions using the Markov inequality and obtain that

$$\sup_{x\geqslant x_n}|Z_n^{**}(x)|\leqslant\frac{c}{x_n^2\sqrt{n}}\left(\sum_{1\leqslant k\leqslant n}\sigma_k^2+\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}s_k^2\sum_{k\leqslant l\leqslant n}\frac{\sigma_l^2}{s_l^2}\right). \tag{5.2}$$

By Note 4.1, $n^{-1}\sum_{1\leqslant k\leqslant n}\sigma_k^2\to\mathbf{E}(\sigma_0^2)$ a.s. and $n^{-1}\sum_{1\leqslant k\leqslant n}s_k^2\to\mathbf{E}(\sigma_0^2)$ a.s. We also have that $\max_{1\leqslant k\leqslant n}(n-k+1)^{-1}\sum_{k\leqslant l\leqslant n}(\sigma_l^2/s_l^2)=\mathcal{O}_{\mathbf{P}}(1)$. Hence, the r.h.s. of bound (5.2) is $\mathcal{O}_{\mathbf{P}}(\sqrt{n}/x_n^2)$, which converges to zero and establishes the desired statement.

In view of the above, we now need to establish weak convergence of $Z_n^*-Z_n^{***}$. Naturally, we first check that all finite dimensional distributions (f.d.d.'s) of $Z_n^*-Z_n^{***}$ converge to the corresponding f.d.d.'s of the process $\{Z(x^2),0\leqslant x<\infty\}$. To this end, we write $Z_n^*-Z_n^{***}=\sum_{1\leqslant l\leqslant n}\xi_{n,l}$, where

$$\xi_{n,l}(x)=\frac{1}{\sqrt{n}}\left(\mathbf{1}\left\{|\varepsilon_l|\leqslant\frac{x}{\sigma_l}\right\}-F_{|\varepsilon_0|}\left(\frac{x}{\sigma_l}\right)\right)-\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left(\mathbf{1}\left\{|\varepsilon_l|\leqslant\frac{xs_l}{s_k\sigma_l}\right\}-F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)\right).$$

For any $m\geqslant 1$, $\lambda_i\in\mathbf{R}$ and $x_i\geqslant 0$, we have $\sum_{1\leqslant i\leqslant m}\lambda_i(Z_n^*-Z_n^{***})(x_i)=\sum_{1\leqslant l\leqslant n}\eta_{n,l}$ with $\eta_{n,l}=\sum_{1\leqslant i\leqslant m}\lambda_i\xi_{n,l}(x_i)$. Note that $\mathbf{E}(\xi_{n,l}(x)|\mathscr{F}_{l-1})=0$ for any l. Hence, for every fixed n, the process $(\sum_{1\leqslant l\leqslant m}\eta_{n,l},\mathscr{F}_m,1\leqslant m\leqslant n)$ is a martingale. We can therefore employ the CLT for martingales (cf., e.g., Hall and Heyde, 1980) to prove convergence of $\sum_{1\leqslant l\leqslant m}\eta_{n,l}$ in distribution. Note that $|\eta_{n,l}|\leqslant c(|\lambda_1|+\cdots+|\lambda_m|)(\log n)/\sqrt{n}$. Hence, Corollary 3.1 of Hall and Heyde (1980) implies that $\sum_{1\leqslant l\leqslant n}\eta_{n,l}$ converges in distribution to $\sum_{1\leqslant i\leqslant m}\lambda_i Z(x_i^2)$,provided that

$$\sum_{1\leqslant l\leqslant n}\mathbf{E}(\eta_{n,l}^2|\mathscr{F}_{l-1})\to\mathbf{E}\left(\left(\sum_{1\leqslant i\leqslant m}\lambda_i Z(x_i^2)\right)^2\right)\quad\text{a.s.}\tag{5.3}$$

To prove the latter statement, we first write the expectation $\mathbf{E}(\eta_{n,l}^2|\mathscr{F}_{l-1})$ as the double sum $\sum_{1\leqslant i,r\leqslant m}\lambda_i\lambda_r\mathbf{E}(\xi_{n,l}(x_i)\xi_{n,l}(x_r)|\mathscr{F}_{l-1})$. Next we analyze the expectation $\mathbf{E}(\xi_{n,l}(x_i)\xi_{n,l}(x_r)|\mathscr{F}_{l-1})$. Trivial algebra shows that this expectation can be written as $n^{-1}(a_{n,l,1}(x_i,x_r)+\cdots+a_{n,l,4}(x_i,x_r))$, where the four summands are, respectively,

$$F_{|\varepsilon_0|}\left(\frac{x_i\wedge x_r}{\sigma_l}\right)-F_{|\varepsilon_0|}\left(\frac{x_i}{\sigma_l}\right)F_{|\varepsilon_0|}\left(\frac{x_r}{\sigma_l}\right),\tag{5.4}$$

$$-\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left(F_{|\varepsilon_0|}\left(\frac{x_is_l}{s_k\sigma_l}\wedge\frac{x_r}{\sigma_l}\right)-F_{|\varepsilon_0|}\left(\frac{x_is_l}{s_k\sigma_l}\right)F_{|\varepsilon_0|}\left(\frac{x_r}{\sigma_l}\right)\right),\tag{5.5}$$

$$-\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left(F_{|\varepsilon_0|}\left(\frac{x_rs_l}{s_k\sigma_l}\wedge\frac{x_i}{\sigma_l}\right)-F_{|\varepsilon_0|}\left(\frac{x_i}{\sigma_l}\right)F_{|\varepsilon_0|}\left(\frac{x_rs_l}{s_k\sigma_l}\right)\right),\tag{5.6}$$

$$\sum_{1\leqslant k\leqslant l}\sum_{1\leqslant i\leqslant l}\frac{1}{n-k+1}\frac{1}{n-i+1}\left(F_{|\varepsilon_0|}\left(\frac{x_is_l}{s_k\sigma_l}\wedge\frac{x_rs_l}{s_i\sigma_l}\right)-F_{|\varepsilon_0|}\left(\frac{x_is_l}{s_k\sigma_l}\right)F_{|\varepsilon_0|}\left(\frac{x_rs_l}{s_i\sigma_l}\right)\right).\tag{5.7}$$

By the ergodic theorem, $n^{-1}\sum_{1\leqslant l\leqslant n}a_{n,l,1}(x_i,x_r)\to K(x_i^2,x_r^2)$ a.s., and so we have

$$\sum_{1\leqslant i,r\leqslant m}\lambda_i\lambda_r\left(\frac{1}{n}\sum_{1\leqslant l\leqslant n}a_{n,l,1}(x_i,x_r)\right)\to\sum_{1\leqslant i,r\leqslant m}\lambda_i\lambda_rK(x_i^2,x_r^2)\quad\text{a.s.}\tag{5.8}$$

Note that the r.h.s. of statement (5.8) equals the r.h.s. of statement (5.3). Hence, the latter statement follows if $n^{-1}(a_{n,l,2}(x_i,x_r)+a_{n,l,3}(x_i,x_r)+a_{n,l,4}(x_i,x_r))\to 0$ a.s. Using the boundedness of $f_{|\varepsilon_0|}$ and also Note 4.1, we replace—with an error terms of up to $\mathcal{O}_{\text{a.s.}}(n^{-1}\log n)$—all s's by σ's in quantities (5.5)–(5.7)) and then see that the resulting quantities add up to zero. This finishes the proof of statement (5.3). Hence, all f.d.d.'s of $Z_n^*-Z_n^{***}$ converge to the corresponding f.d.d.'s of the process $\{Z(x^2),0\leqslant x<\infty\}$. Theorem 2.2, however, requires us to establish *weak convergence* of $Z_n^*-Z_n^{***}$. To this end, we first transform the process $Z_n^*-Z_n^{***}$ into

$$V_n(t)=(Z_n^*-Z_n^{***})(F_{|y_0|}^{-1}(t)),\quad 0\leqslant t\leqslant 1.\tag{5.9}$$

We already know that all f.d.d.'s of V_n converge to the corresponding f.d.d.'s of the process V, which is defined by

$$V(t)=Z([F_{|y_0|}^{-1}(t)]^2),\quad 0\leqslant t\leqslant 1.$$

This verifies assumption (i) of the following theorem, which has recently been extended to multi-parameter stochastic processes by Davydov and Zitikis (2008). □

Theorem 5.1 (*Davydov, 1996*). *Let $V_n=\{V_n(t),0\leqslant t\leqslant 1\}$ be a stochastic process with paths in $D[0,1]$ a.s. Let the following three assumptions hold*:

(i) *There exists a process $V=\{V(t),\ 0\leqslant t\leqslant 1\}$ such that all f.d.d.'s of V_n converge to the corresponding f.d.d.'s of V.*
(ii) *There exist constants $\kappa>\gamma>1$, $c>0$, and a sequence of positive numbers $a_n\downarrow 0$ such that, for all $n\geqslant 1$ and whenever $|t-s|\geqslant a_n$, we have that*

$$\mathbf{E}(|V_n(t)-V_n(s)|^{\kappa})\leqslant c|t-s|^{\gamma}.$$

(iii) *The process V_n can be written as the difference of two non-decreasing processes $G_n^*-G_n^{**}$ such that at least one of them, denote it by G_n, satisfies the property*

$$\max_i|G_n(t_{i+1})-G_n(t_i)|\to_{\mathbf{P}}0,$$

where, for every fixed n, the maximum is taken over all $i = 0, 1, \ldots, i_n$ with $i_n = [1/a_n]$, and $t_i = ia_n$ for all $i = 0, 1, \ldots, i_n$ and $t_{i_n+1} = 1$.

Then V_n converges weakly to V in $D[0, 1]$, and the limit process V is continuous a.s.

To complete the proof of Theorem 2.2, we need to verify assumptions (ii) and (iii). Since the verification is involved, we subdivide the task into two sections.

6. Verification of assumption (ii) of Theorem 5.1

We verify assumption (ii) with $\kappa = 4$ and $a_n = n^{-(1-\lambda)}$, where $0 < \lambda < 1$. With $y = F^{-1}_{|y_0|}(t)$ and $x = F^{-1}_{|y_0|}(s)$, we write $V_n(t) - V_n(s)$ as $\sum_{i=1}^n \xi_{n,l}((x, y])$, where $\xi_{n,l}((x, y]) = \xi_{n,l}(y) - \xi_{n,l}(x)$. (In general, for any function h, we use the notation $h((x, y]) = h(y) - h(x)$.) Hence, we need to prove that, when $|t - s| \geqslant a_n$,

$$\mathbf{E}\left(\left|\sum_{i=1}^{n} \xi_{n,l}((x, y])\right|^4\right) \leqslant c|t - s|^{\gamma}. \tag{6.1}$$

For fixed $x < y$ the process $(\sum_{1 \leqslant l \leqslant m} \xi_{n,l}((x, y]), \mathscr{F}_m, 1 \leqslant m \leqslant n)$ is a martingale. The Rosenthal inequality (cf., e.g., Hall and Heyde, 1980, pp. 23–24) implies that up to a universal constant c the moment in (6.1) does not exceed the sum of

$$\sum_{1 \leqslant l \leqslant n} \mathbf{E}(\xi^4_{n,l}((x, y])) \quad \text{and} \quad \mathbf{E}\left(\left[\sum_{1 \leqslant l \leqslant n} \mathbf{E}(\xi^2_{n,l}((x, y])|\mathscr{F}_{l-1})\right]^2\right). \tag{6.2}$$

We need to show that the two quantities in (6.2) do not exceed $c|t - s|^{\gamma}$ for some $\gamma > 1$. To this end, we first express $\xi_{n,l}((x, y])$ as the sum of $n^{-1/2}A_{n,l}((x, y])$ and $n^{-1/2}B_{n,l}((x, y])$, where

$$A_{n,l}(x) = \mathbf{1}\left\{|\varepsilon_l| \leqslant \frac{x}{\sigma_l}\right\} - F_{|\varepsilon_0|}\left(\frac{x}{\sigma_l}\right)$$

and $B_{n,l}(x) = \sum_{1 \leqslant k \leqslant l} B_{n,l,k}(x)$ with the notation

$$B_{n,l,k}(x) = \frac{1}{n-k+1}\left(\mathbf{1}\left\{\varepsilon_l \leqslant \frac{xs_l}{s_k\sigma_l}\right\} - F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)\right).$$

We next prove the upper bound $c|t - s|^{\gamma}$ for the first quantity in (6.2). Up to a constant c, the quantity does not exceed the sum of

$$\frac{1}{n^2}\sum_{1 \leqslant l \leqslant n} \mathbf{E}(A^4_{n,l}((x, y])) \quad \text{and} \quad \frac{1}{n^2}\sum_{1 \leqslant l \leqslant n} \mathbf{E}(B^4_{n,l}((x, y])). \tag{6.3}$$

Hence, we need to establish the upper bound $c|t - s|^{\gamma}$ for both quantities in (6.3). For this, we first note that since $|A_{n,l}(x)| \leqslant 1$, the expectation $\mathbf{E}(A^4_{n,l}((x, y]))$ does not exceed $\mathbf{E}(|A_{n,l}((x, y])|)$. The latter expectation does not exceed $2\mathbf{E}(|F_{|\varepsilon_0|}((x, y]/\sigma_l)|)$, which is of course $2|F_{|y_0|}(y) - F_{|y_0|}(x)|$ and thus $2|t - s|$. Consequently, the first quantity in (6.3) does not exceed $cn^{-1}|t - s|$. Since $a_n = n^{-(1-\lambda)}$ and $|t - s| \geqslant a_n$, we have that $cn^{-1}|t - s|$ does not exceed $c|t - s|^{\gamma}$ for some $\gamma > 1$, which is the desired result. We next bound the second quantity in (6.3), and start with the bound

$$\mathbf{E}(B^4_{n,l}((x, y])) \leqslant \left(\sum_{1 \leqslant k \leqslant l} [\mathbf{E}(B^4_{n,l,k}((x, y]))]^{1/4}\right)^4. \tag{6.4}$$

To estimate the expectation on the r.h.s. of bound (6.4), we first use the fact that $B_{n,l,k}((x, y])$ does not exceed 1 and then obtain

$$\begin{aligned}\mathbf{E}(B^4_{n,l,k}((x, y])) \leqslant{}& \frac{c}{(n-k+1)^4}\sup_{0 \leqslant x < \infty}\left|F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right) - F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)\right| \\ &+ \frac{c}{(n-k+1)^4}\mathbf{E}\left(\left|F_{|\varepsilon_0|}\left(\frac{(x, y]}{\sigma_k}\right)\right|\right).\end{aligned} \tag{6.5}$$

We have already noted above that the second expectation on the r.h.s. of bound (6.5) does not exceed $c|t-s|$. To estimate the supremum on the r.h.s. of bound (6.5), we split it into two halves: $x \leqslant x_n$ and $x > x_n$, with $x_n = 1/a_n$. Using the assumption that the density $f_{|\varepsilon_0|}$ is bounded, we estimate the supremum with respect to $x \leqslant x_n$ by the quantity $cx_n|\sigma_k - s_k| + cx_n|\sigma_l - s_l|$. To estimate the supremum with respect to $x > x_n$, we split it into the sum of the survival function of $|\varepsilon_0|$ evaluated at the points $x_n s_l/(s_k\sigma_l)$ and x_n/σ_k, and then apply the Markov inequality. In summary, we have from bound (6.5) that $\mathbf{E}(B^4_{n,l,k}((x,y]))$ does not exceed

$$\frac{c}{(n-k+1)^4}\left(x_n\mathbf{E}(|\sigma_k - s_k|) + x_n\mathbf{E}(\sigma_l - s_l|) + \frac{1}{x_n}\mathbf{E}(s_k\sigma_l + \sigma_k) + |t-s|\right).$$

Note 4.1 shows that $\mathbf{E}(|\sigma_k - s_k|)$ does not exceed $c\rho^{k/q}$. Furthermore, the expectation $\mathbf{E}(s_k\sigma_l + \sigma_k)$ does not exceed a constant c, which is independent of k and l. Hence,

$$\sum_{1\leqslant k\leqslant l}[\mathbf{E}(B^4_{n,l,k}((x,y]))]^{1/4} \leqslant \frac{cx_n^{1/4}}{(n-l+1)} + c\log n\left(x_n\rho^{l/q} + \frac{1}{x_n} + |t-s|\right)^{1/4}. \tag{6.6}$$

We raise the r.h.s. of bound (6.6) to the power 4 (cf. bound (6.4)), sum up the result with respect to $1 \leqslant l \leqslant n$, divide it by n^2, and arrive at the bound

$$\frac{1}{n^2}\sum_{1\leqslant l\leqslant n}\mathbf{E}(B^4_{n,l}((x,y])) \leqslant c(\log n)^4\left(\frac{x_n}{n^2} + \frac{1}{x_n n} + \frac{1}{n}|t-s|\right). \tag{6.7}$$

Since $x_n = 1/a_n$, $a_n = n^{-(1-\lambda)}$ and $|t-s| \geqslant a_n$, the r.h.s. of bound (6.7) does not exceed $c|t-s|^\gamma$ with $\gamma > 1$. This finishes the estimation of the first quantity in (6.2).

In order to obtain the upper bound $c|t-s|^\gamma$ for the second quantity of (6.2), we estimate it by the sum of

$$\frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\mathbf{E}(A^2_{n,l}((x,y])|\mathscr{F}_{l-1})\right]^2\right) \quad \text{and} \quad \frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\mathbf{E}(B^2_{n,l}((x,y])|\mathscr{F}_{l-1})\right]^2\right). \tag{6.8}$$

The first quantity in (6.8) does not exceed $n^{-2}\mathbf{E}([\sum_{1\leqslant l\leqslant n}Y_l]^2)$, where the random variables $Y_l = F_{|\varepsilon_0|}((x,y]/\sigma_l)$ are identically distributed. If the random variables Y_l were independent, then $\mathbf{E}([\sum_{1\leqslant l\leqslant n}Y_l]^2)$ would not exceed $n^2(\mathbf{E}(Y_0))^2 + n\mathbf{E}(Y_0^2)$; note also $\mathbf{E}(Y_0^2) \leqslant \mathbf{E}(Y_0)$. The expectation $\mathbf{E}(Y_0)$ is equal to $F_{|y_0|}(y) - F_{|y_0|}(x)$, which is $t-s$. (Recall that we consider the case $x < y$.) Hence, under the independence assumption, the first quantity in (6.8) does not exceed $|t-s|^2 + n^{-1}|t-s|$, which is a desired bound. Note, however, that the variables Y_l are not independent; fortunately, they are nearly independent as proved by Berkes and Horváth (2001, Lemma 2.1) and recorded for our convenience as Lemma 6.1.

Lemma 6.1. *If $\mathbf{E}((\log^+\|A\|)^\mu) < \infty$, then there are constants c_1 and c_2 such that $\sigma_k^2 = \tau_{k,R}^2 + \eta_{k,R}\sigma_{k-R}^2$, where $\tau_{k,R}^2 \in \mathscr{F}(\varepsilon_{k-1},\ldots,\varepsilon_{k-R})$ is not smaller than a constant $c(\boldsymbol{\theta}) > 0$, and the bound $\mathbf{P}(|\eta_{k,R}| \geqslant e^{-c_1R}) \leqslant c_2R^{-\mu/2}$ holds.*

The variables $\tau_{k,R}$ and $\tau_{l,R}$ are independent when $|k-l| > R$, and they are close to σ_k and σ_l, respectively. This is the main idea of our proof below showing that, for some $\gamma > 1$,

$$\frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}Y_l\right]^2\right) \leqslant c|t-s|^\gamma \tag{6.9}$$

whenever $|t-s| \geqslant a_n$. We start the proof of bound (6.9) as follows:

$$\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}Y_l\right]^2\right) \leqslant \mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}Y_{l,R}\right]^2\right) + cn\sum_{1\leqslant l\leqslant n}\Lambda_l, \tag{6.10}$$

where $Y_{l,R} = F_{|\varepsilon_0|}((x,y]/\tau_{l,R})$ and

$$\Lambda_l = \mathbf{E}\left(\sup_{0\leqslant x<\infty}\left|F_{|\varepsilon_0|}\left(\frac{x}{\sigma_l}\right) - F_{|\varepsilon_0|}\left(\frac{x}{\tau_{l,R}}\right)\right|\right).$$

We write the second moment on the r.h.s. of (6.10) as a double sum, split it into two parts according to whether $|k-l| \leqslant R$ or $|k-l|>R$, and then use the independence of $\tau_{k,R}$ and $\tau_{l,R}$ when $|k-l|>R$. This gives the bound

$$\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n} Y_{l,R}\right]^2\right) \leqslant cnR + \left(\sum_{1\leqslant k\leqslant n} \mathbf{E}(Y_{k,R})\right)^2. \tag{6.11}$$

The sum $\sum \mathbf{E}(Y_{k,R})$ does not exceed $2\sum \Lambda_l + \sum \mathbf{E}(Y_k)$, and $\mathbf{E}(Y_k) = |t-s|$. Hence,

$$\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n} Y_l\right]^2\right) \leqslant cnR + \left(2\sum_{1\leqslant l\leqslant n} \Lambda_l + n|t-s|\right)^2 + cn\sum_{1\leqslant l\leqslant n} \Lambda_l. \tag{6.12}$$

We next estimate Λ_l by splitting the supremum into two halves: $x \leqslant x_n$ and $x > x_n$. To estimate the first part, we use the assumption that $f_{|\varepsilon_0|}$ is bounded and, with the help of Lemma 6.1, obtain that

$$\begin{aligned}\mathbf{E}\left(\sup_{x\leqslant x_n}\left|F_{|\varepsilon_0|}\left(\frac{x}{\sigma_l}\right) - F_{|\varepsilon_0|}\left(\frac{x}{\tau_{l,R}}\right)\right|\right) &\leqslant cx_n\mathbf{E}\left(\left|\frac{1}{\sigma_l} - \frac{1}{\tau_{l,R}}\right| \mathbf{1}\{|\eta_{k,R}| \leqslant \mathrm{e}^{-cR}\}\right) + cR^{-\mu/2}\\ &\leqslant cx_n\mathrm{e}^{-cR} + cR^{-\mu/2}.\end{aligned}$$

To estimate the second part, we bound the supremum by the sum of two survival functions, split one of them into two parts as suggested by Lemma 6.1, apply the Markov inequality, and obtain that

$$\begin{aligned}\mathbf{E}\left(\sup_{x> x_n}\left|F_{|\varepsilon_0|}\left(\frac{x}{\sigma_l}\right) - F_{|\varepsilon_0|}\left(\frac{x}{\tau_{l,R}}\right)\right|\right) &\leqslant \frac{1}{x_n^2}\mathbf{E}(\sigma_l^2) + \frac{1}{x_n^2}\mathbf{E}(\tau_{l,R}^2\mathbf{1}\{|\eta_{k,R}| \leqslant \mathrm{e}^{-cR}\}) + cR^{-\mu/2}\\ &\leqslant \frac{c}{x_n^2} + cR^{-\mu/2}.\end{aligned}$$

The above two bounds imply that $\Lambda_l \leqslant c(x_n^{-2} + x_n\mathrm{e}^{-cR} + R^{-\mu/2})$. Choosing the parameters appropriately, we have that $\Lambda_l \leqslant c|t-s|^\gamma$ when $|t-s| \geqslant a_n$. Indeed, with $x_n = 1/a_n$ and $a_n = 1/n^{1-\lambda}$ as before, and with $R = n^\rho$ for some $0<\rho<1$ to be specified later, we have that x_n^{-2} and $x_n\mathrm{e}^{-cR}$ do not exceed $c|t-s|^2$, whereas $R^{-\mu/2}$ does not exceed $|t-s|^\gamma$ when $\mu > 2(1-\lambda)/\rho$. The latter condition follows from $\mu > 2(1-\lambda)/\lambda$, provided that we can choose ρ as close to λ as desired, which we can do. Hence, we have the bound $n^{-2}\mathbf{E}([\sum_{1\leqslant l\leqslant n} Y_l]^2) \leqslant cRn^{-1} + c|t-s|^\gamma$. Finally, Rn^{-1} does not exceed $|t-s|^\gamma$ with $\gamma = (1-\rho)/(1-\lambda)$, which is greater than 1 since $\rho<\lambda$. This proves that the first quantity in (6.8) does not exceed $|t-s|^\gamma$ when $|t-s| \geqslant a_n$.

We next establish an analogous bound for the second quantity in (6.8) under the same choices of parameters as above, except that we now restrict the values of λ to $\frac{1}{3} < \lambda < 1$. Using the equation $B_{n,l}(x) = \sum_{1\leqslant k\leqslant l} B_{n,l,k}(x)$, writing $B_{n,l}^2((x,y])$ as a double sum, and applying the Cauchy–Schwarz inequality, we have that

$$\begin{aligned}\mathbf{E}(B_{n,l}^2((x,y])|\mathscr{F}_{l-1})) &\leqslant c\left(\sum_{1\leqslant k\leqslant l} \frac{1}{n-k+1}\left(\left|F_{|\varepsilon_0|}\left(\frac{(x,y]s_l}{s_k\sigma_l}\right)\right|\right)^{1/2}\right)^2\\ &\leqslant c\log n \sum_{1\leqslant k\leqslant l} \frac{1}{n-k+1}\left|F_{|\varepsilon_0|}\left(\frac{(x,y]s_l}{s_k\sigma_l}\right)\right|.\end{aligned} \tag{6.13}$$

To estimate the r.h.s. of bound (6.13), we replace $s_l/(s_k\sigma_l)$ by $1/\sigma_k$ and $1/\sigma_k$ by $1/\tau_{k,R}$; the latter replacement gives the needed independence structure. That is, we have that $\mathbf{E}(B_{n,l}^2((x,y])|\mathscr{F}_{l-1}))$ does not exceed the sum of the three quantities

$$\Theta_{n,l}^*((x,y]) = c\log n\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left|F_{|\varepsilon_0|}\left(\frac{(x,y]s_l}{s_k\sigma_l}\right) - F_{|\varepsilon_0|}\left(\frac{(x,y]}{\sigma_k}\right)\right|,$$

$$\Theta_{n,l}^{**}((x,y]) = c\log n\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left|F_{|\varepsilon_0|}\left(\frac{(x,y]}{\sigma_k}\right) - F_{|\varepsilon_0|}\left(\frac{(x,y]}{\tau_{k,R}}\right)\right|,$$

$$\Theta_{n,l}^{***}((x,y]) = c\log n\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\left|F_{|\varepsilon_0|}\left(\frac{(x,y]}{\tau_{k,R}}\right)\right|.$$

We next show that the r.h.s. of the following bound does not exceed $c|t-s|^{\gamma}$ whenever $|t-s| \geqslant a_n$ (cf. notes below bound (6.5) for hints):

$$\begin{aligned}\frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\Theta^{*}_{n,l}((x,y])\right]^2\right) &\leqslant c\frac{x_n^2(\log n)^2}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}|\sigma_k-s_k|\right]^2\right)\\ &+c\frac{x_n^2(\log n)^4}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}|\sigma_l-s_l|\right]^2\right)\\ &+c\frac{(\log n)^2}{x_n^2n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\sigma_l\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}s_k\right]^2\right)\\ &+c\frac{(\log n)^2}{x_n^2n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\sum_{1\leqslant k\leqslant l}\frac{1}{n-k+1}\sigma_k\right]^2\right). \end{aligned}\tag{6.14}$$

The first term on the r.h.s. of bound (6.14) does not exceed the second one. Note 4.1 implies that the expectation in the second term does not exceed a constant since $\mathbf{E}(y_0^4)<\infty$. The third expectation on the r.h.s. of bound (6.14) does not exceed cn^2 since $\mathbf{E}(\sigma_0^4)<\infty$. The fourth expectation does not exceed $\mathbf{E}([\sum_{1\leqslant k\leqslant n}\sigma_k]^2)$, which does not exceed cn^2. Hence, the r.h.s. of bound (6.14) does not exceed $cx_n^2(\log n)^4n^{-2}+c(\log n)^2x_n^{-2}$. With $x_n=a_n^{-1}$ and $a_n=1/n^{1-\lambda}$ as before, we only need to show that $c(\log n)^4/(a_n^2n^2)$ and $ca_n^2(\log n)^2$ do not exceed $|t-s|^{\gamma}$ when $|t-s|\geqslant a_n$. The quantity $ca_n^2(\log n)^2$ obviously satisfies the bound. As to $c(\log n)^4/(a_n^2n^2)$, we first note that this quantity equals $c(\log n)^4/n^{2\lambda}$, which does not exceed $|t-s|^{\gamma}$ when $|t-s|\geqslant a_n$, provided that $2\lambda/(1-\lambda)>1$. The latter holds when $\lambda>\frac{1}{3}$, which we assume.

We next consider the quantity $\Theta^{**}_{n,l}((x,y])$ and start with the bound

$$\frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\Theta^{**}_{n,l}((x,y])\right]^2\right)\leqslant\frac{c(\log n)^2}{n}\sum_{1\leqslant k\leqslant n}\mathbf{E}\left(\left|F_{|\varepsilon_0|}\left(\frac{(x,y]}{\sigma_k}\right)-F_{|\varepsilon_0|}\left(\frac{(x,y]}{\tau_{k,R}}\right)\right|^2\right).\tag{6.15}$$

Using the earlier obtained estimate of Λ_l, we have that the r.h.s. of bound (6.15) does not exceed $c(\log n)^2(x_n^{-2}+x_n\mathrm{e}^{-cR}+R^{-\mu/2})$, which in turn does not exceed $|t-s|^{\gamma}$ when $|t-s|\geqslant a_n$. Finally, we consider $\Theta^{***}_{n,l}((x,y])$ and write the equation

$$\frac{1}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant l\leqslant n}\Theta^{**}_{n,l}((x,y])\right]^2\right)=\frac{(\log n)^2}{n^2}\mathbf{E}\left(\left[\sum_{1\leqslant k\leqslant n}\left|F_{|\varepsilon_0|}\left(\frac{(x,y]}{\tau_{k,R}}\right)\right|\right]^2\right).\tag{6.16}$$

Arguments below bound (6.11) imply that the r.h.s. of Eq. (6.16) does not exceed $c|t-s|^{\gamma}$ when $|t-s|\geqslant a_n$. This completes the verification of condition (ii).

7. Verification of assumption (iii) of Theorem 5.1

With $x=F^{-1}_{|y_0|}(t)$ we write $V_n(t)=Z_n^{*}(x)-Z_n^{***}(x)$ as follows:

$$\begin{aligned}V_n(t)=&\left[\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\mathbf{1}\left\{|\varepsilon_k|\leqslant\frac{x}{\sigma_k}\right\}+\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)\right]\\ &-\left[\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\mathbf{1}\left\{|\varepsilon_l|\leqslant\frac{xs_l}{s_k\sigma_l}\right\}+\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)\right].\end{aligned}\tag{7.1}$$

This is the difference of two non-decreasing processes, and since we can choose any of them to verify assumption (iii), we choose the one in the first square brackets. Denote this processes by G_n and rewrite it in a slightly different form

$$\begin{aligned}G_n(x)=&\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}A_{n,k}(x)+\frac{2}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)\\ &+\frac{1}{\sqrt{n}}\sum_{1\leqslant k\leqslant n}\frac{1}{n-k+1}\sum_{k\leqslant l\leqslant n}\left(F_{|\varepsilon_0|}\left(\frac{xs_l}{s_k\sigma_l}\right)-F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right)\right).\end{aligned}\tag{7.2}$$

We need to show that $\max_i |G_n((x_i, x_{i+1}])| \to_{\mathbf{P}} 0$, where $x_i = F^{-1}_{|y_0|}(t_i)$, $i = 0, 1, \ldots, i_n$, and $x_{i_n+1} = F^{-1}_{y_0^2}(1)$. The third term on the r.h.s. of Eq. (7.2) converges to 0 uniformly in x since the difference between the two distribution functions there does not exceed $cx_n|\sigma_k - s_k| + cx_n|\sigma_l - s_l| + cx_n^{-2}(s_k^2\sigma_l^2 + \sigma_k^2)$ with an appropriate choice of x_n; see the estimation of the supremum on the r.h.s. of bound (6.5) for detail. Hence, we need to prove only that $\max_i |H_n((x_i, x_{i+1}])| \to_{\mathbf{P}} 0$, where $H_n(x)$ is the sum of the first two terms on the r.h.s. of bound (7.2), that is,

$$H_n(x) = \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} A_{n,k}(x) + \frac{2}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} \left(F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right) - \mathbf{E}\left(F_{|\varepsilon_0|}\left(\frac{x}{\sigma_0}\right)\right)\right) + 2\sqrt{n}\mathbf{E}\left(F_{|\varepsilon_0|}\left(\frac{x}{\sigma_0}\right)\right).$$

The right-most expectation equals $F_{|y_0|}(x)$. The quantity $|F_{|y_0|}((x_i, x_{i+1}])|$ equals $|t_{i+1} - t_i|$, which does not exceed a_n. Hence, $\sqrt{n} \max_i |F_{|y_0|}((x_i, x_{i+1}])|$ does not exceed $\sqrt{n}\, a_n$, which converges to 0 since $\lambda < \frac{1}{2}$. Consequently, the verification of assumption (iii) reduces to proving the statements:

$$\max_i \left| \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} A_{n,k}((x_i, x_{i+1}]) \right| \to_{\mathbf{P}} 0, \tag{7.3}$$

$$\max_i \left| \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} \left(F_{|\varepsilon_0|}\left(\frac{(x_i, x_{i+1}]}{\sigma_k}\right) - \mathbf{E}\left(F_{|\varepsilon_0|}\left(\frac{(x_i, x_{i+1}]}{\sigma_0}\right)\right)\right)\right| \to_{\mathbf{P}} 0. \tag{7.4}$$

To prove statement (7.3), we write the bound

$$\mathbf{P}\left\{ \max_i \left| \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} A_{n,k}((x_i, x_{i+1}]) \right| > \varepsilon \right\} \leqslant \frac{1}{\varepsilon^4 n^2} \sum_i \mathbf{E}\left(\left| \sum_{1 \leqslant k \leqslant n} A_{n,k}((x_i, x_{i+1}]) \right|^4 \right). \tag{7.5}$$

For fixed $x < y$, the process $(\sum_{1 \leqslant k \leqslant m} A_{n,k}((x, y]), \mathscr{F}_m, 1 \leqslant m \leqslant n)$ is a martingale. Thus, the Rosenthal inequality (cf., e.g., Hall and Heyde, 1980, pp. 23–24) implies

$$\frac{1}{n^2}\mathbf{E}\left(\left| \sum_{1 \leqslant k \leqslant n} A_{n,k}((x_i, x_{i+1}]) \right|^4 \right) \leqslant \frac{c}{n^2} \sum_{1 \leqslant k \leqslant n} \mathbf{E}(A^4_{n,k}((x_i, x_{i+1}])) + \frac{c}{n^2}\mathbf{E}\left(\left[\sum_{1 \leqslant k \leqslant n} \mathbf{E}(A^2_{n,k}((x_i, x_{i+1}]) | \mathscr{F}_{l-1}) \right]^2 \right). \tag{7.6}$$

We have encountered quantities on the r.h.s. of bound (7.6) in the proofs above. Specifically, we already know that the first term on the r.h.s. of (7.6) does not exceed $(c/n)|t_{i+1} - t_i|$. Summing it up with respect to i, we have a quantity that converges to 0 when $n \to \infty$. The second term on the r.h.s. of (7.6) does not exceed $c(x_n^{-2} + x_n \mathrm{e}^{-cR} + R^{-\mu/2} + R/n + |t_{i+1} - t_i|^2)$ with, say, $x_n = 1/a_n$. Summing it up with respect to i and using the facts that $|t_{i+1} - t_i| \leqslant a_n$ and that there are at most $\mathcal{O}(1/a_n)$ summands in the sum $\sum_i$, we obtain a quantity of the order $n^{-1} + a_n + a_n^{-2}\mathrm{e}^{-cR} + a_n^{-1}R^{-\mu/2} + Ra_n^{-1}n^{-1}$. It converges to 0 since $a_n = 1/n^{1-\lambda}$, $R = n^\rho$, ρ is a sufficiently close to λ, and $\mu > 2(1-\lambda)/\lambda$. This completes the proof of statement (7.3).

To prove statement (7.4), we first note that

$$\frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} \mathbf{E}\left(\sup_{0 \leqslant x < \infty} \left| F_{|\varepsilon_0|}\left(\frac{x}{\sigma_k}\right) - F_{|\varepsilon_0|}\left(\frac{x}{\tau_{k,R}}\right)\right|\right) \to_{\mathbf{P}} 0.$$

Indeed, from results above, it follows that the r.h.s. of the above statement does not exceed $c\sqrt{n}(x_n^{-2} + x_n \mathrm{e}^{-cR} + R^{-\mu/2})$, which converges to 0 with our choices of parameters. With the help of the just proved statement, we reduce statement (7.4) to an analogous one with $\tau_{k,R}$ instead of σ_k and σ_0. Denote

$$D_{k,R}(x_i, x_{i+1}] = F_{|\varepsilon_0|}\left(\frac{(x_i, x_{i+1}]}{\tau_{k,R}}\right) - \mathbf{E}\left(F_{|\varepsilon_0|}\left(\frac{(x_i, x_{i+1}]}{\tau_{k,R}}\right)\right).$$

Our task becomes to show that the r.h.s. of the following bound converges to 0:

$$\mathbf{P}\left\{ \max_i \left| \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant n} D_{k,R}(x_i, x_{i+1}] \right| > \varepsilon \right\} \leqslant \frac{1}{\varepsilon^4 n^2} \sum_i \mathbf{E}\left(\left| \sum_{1 \leqslant k \leqslant n} D_{k,R}(x_i, x_{i+1}] \right|^4 \right). \tag{7.7}$$

To estimate the right-most expectation, we rewrite it as the sum of (1) the sum over all indices $1 \leqslant k_1, k_2, k_3, k_4 \leqslant n$ such that the distance between any of the k_i's does not exceed $3R$ and (2) the sum over all $1 \leqslant k_1, k_2, k_3, k_4 \leqslant n$ such that the above '$3R$-distance rule' is violated but the following property holds: for each of the four k_i's there is a k_j such that the distance between k_i and k_j does not exceed R. Hence,

$$\mathbf{E}\left(\left|\sum_{1 \leqslant k \leqslant n} D_{k,R}(x_i, x_{i+1}]\right|^4\right) \leqslant cnR^3\mathbf{E}(|D_{0,R}(x_i, x_{i+1}]|) + cn^2R^2(\mathbf{E}(|D_{0,R}(x_i, x_{i+1}]|))^2. \tag{7.8}$$

Estimates obtained in proofs above show that the r.h.s. of bound (7.8) does not exceed (with our usual choice of parameters; see below bound (7.9))

$$cnR^3\left(\frac{1}{x_n^2} + \frac{x_n}{\mathrm{e}^{cR}} + \frac{1}{R^{\mu/2}}\right) + cn^2R^2\left(\frac{1}{x_n^2} + \frac{x_n}{\mathrm{e}^{cR}} + \frac{1}{R^{\mu/2}}\right)^2 + cnR^3|t_{i+1} - t_i| + cn^2R^2|t_{i+1} - t_i|^2.$$

Since the sum $\sum_i$ has $\mathcal{O}(1/a_n)$ summands, the above bound implies that

$$\sum_i \frac{1}{n^2}\mathbf{E}\left(\left|\sum_{1 \leqslant k \leqslant n} D_{k,R}(x_i, x_{i+1}]\right|^4\right) \leqslant \frac{cR^3}{na_n}\left(\frac{1}{x_n^2} + \frac{x_n}{\mathrm{e}^{cR}} + \frac{1}{R^{\mu/2}}\right) + \frac{cR^2}{a_n}\left(\frac{1}{x_n^2} + \frac{x_n}{\mathrm{e}^{cR}} + \frac{1}{R^{\mu/2}}\right)^2 + \frac{cR^3}{n} + ca_nR^2. \tag{7.9}$$

We choose $a_n = 1/n^{1-\lambda}$, $x_n = 1/a_n$, $R = n^\rho$, and $\lambda > \frac{1}{3}$, and then check that there exists $\rho < \frac{1}{3}$ such that, under the assumption $\mu > 4$, the r.h.s. of bound (7.9) converges to 0. This completes the verification of condition (iii). The proof of Theorem 5.1 is finished.

Acknowledgements

Sincere thanks are due to Youri Davydov for discussions concerning his paper, Davydov (1996), as well as to Javier Hidalgo for sharing his work with us prior to its publication. We are also grateful to two anonymous referees and the editor, Takeshi Hayakawa, for their constructive criticism and suggestions that helped to reshape the manuscript.

References

Basrak, B., Davis, R., Mikosch, T., 2002. Regular variation of GARCH processes. Stochastic Process. Appl. 99, 95–115.
Berkes, I., Horváth, L., 2001. Strong approximation of the empirical process of GARCH sequences. Ann. Appl. Probab. 11, 789–809.
Berkes, I., Horváth, L., Kokoszka, P.S., 2003. GARCH processes: structure and estimation. Bernoulli 9, 201–227.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31, 307–327.
Bougerol, P., Picard, N., 1992a. Strict stationarity of generalized autoregressive processes. Ann. Probab. 20, 1714–1730.
Bougerol, P., Picard, N., 1992b. Stationarity of GARCH processes and of some nonnegative time series. J. Econometrics 52, 115–127.
Carrasco, M., Chen, X., 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory 18, 17–39.
Davydov, Y., 1996. Weak convergence of discontinuous processes to continuous ones. In: Probability Theory and Mathematical Statistics. Gordon and Breach, Amsterdam, pp. 15–18.
Davydov, Y., Zitikis, R., 2008. On weak convergence of random fields. Ann. Inst. Statist. Math., in press, doi:10.1007/s10463-006-0090-4.
Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.
Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Applications. Academic Press, New York.
He, C., Teräsvirta, T., 1999. Properties of moments of a family of GARCH processes. J. Econometrics 92, 173–192.
He, C., Teräsvirta, T., 2004. An extended constant conditional correlation GARCH model and its fourth-moment structure. Econometric Theory 20, 904–926.
Mikosch, T., Stărică, C., 2000. Limit theory for the sample autocorrelation and extremes of a GARCH(1, 1) process. Ann. Statist. 28, 1427–1451.
Nelson, D.B., 1990. Stationarity and persistence in the GARCH(1, 1) model. Econometric Theory 6, 318–334.
Zivot, E., Wang, J., 2003. Modeling Financial Time Series with S-Plus. Springer, Berlin.