Manuscript submitted to *The Econometrics Journal*, pp. 1–26.

Supplementary Material to "Testing for parameter change epochs in GARCH time series"

Stefan Richter[†], Weining Wang[‡] and Wei Biao Wu[*]

[†] *Institute of Applied Mathematics, Im Neuenheimer Feld 205, Heidelberg, Germany.*
E-mail: stefan.richter@iwr.uni-heidelberg.de
[‡] *University of York*[‡].
E-mail: weining.wang@york.ac.uk
[*] *University of Chicago.*
E-mail: wbwu@galton.uchicago.edu

Summary
Section A presents the simulations results of the proposed tests, and Section B provides the proofs for the theorems and technical lemmata.

A. SIMULATIONS

In this section, we conduct a simulation study for evaluating the performance of our methodology. The algorithm is summarized in Section 2. We consider the GARCH(1,1)-model with time-varying parameters $\theta(i) = (\alpha_0(i), \alpha_1(i), \beta_1(i))'$,

$$\theta(i) = \begin{cases} \theta^*, & i \le \lfloor n\tau_1^* \rfloor, \\ \theta^* + H\Delta^*, & \lfloor n\tau_1^* \rfloor + 1 \le i \le \lfloor n\tau_2^* \rfloor, \\ \theta^*, & i > \lfloor n\tau_2^* \rfloor, \end{cases}$$

with $H_0 : \Delta^* = 0$ and $H_1 : \Delta^* > 0$, where

i) $H = (0, 1, 1)'$ and $\alpha_0^* = 0.2$, $\alpha_1^* = 0.5$, $\beta_1^* = 0.25$,
ii) $H = (0, 1, 1)'$ and $\alpha_0^* = 0.3$, $\alpha_1^* = 0.39$, $\beta_1^* = 0.6$,
iii) $H = (0, 1, 1)'$ and $\alpha_0^* = 0.5$, $\alpha_1^* = 0.07$, $\beta_1^* = 0.91$,

and

$$X_i^2 = \zeta_i^2 \sigma_i^2, \qquad \sigma_i^2 = \alpha_0(i) + \alpha_1(i) X_{i-1}^2 + \beta_1(i)\sigma_{i-1}^2.$$

We now check the behaviour of the test under the null hypothesis H_0 of no change. We use the test proposed in Algorithm 1 in Section 2 with $\kappa = \kappa' = 0.1$ and a grid approximation of $L = 10, 30, 50$. In addition we also show the comparison with the method in Chen and Hong (2016). We adopt the test statistics as in Equation (4.5) in CH. We use the uniform kernel with the support $[-1, 1]$. The bandwidth here is selected using a grid search and a cross validation method.

For $N = 1000$ replications and $n \in \{500, 1000, 2000\}$, $\delta \in \{0.90, 0.95\}$, we obtain the quantiles $\hat{q}_{W,0.90} \approx 3.031$, $\hat{q}_{W,0.95} \approx 3.285$ and the results given in Table A.1 (cf. Step 4 in the algorithm in Section 2). We can see from the table that as the sample size increases, the coverage probabilities approach the nominal level for both $\delta = 0.90, 0.95$, which illustrates that the asymptotic results hold true already for relatively small sample sizes. We can see that with increasing L, the sizes would drop a bit. The CH test in the most cases has slightly better sizes, but our method is still quite competitive.

Table A.1. The averaged acceptance rate under null hypothesis H_0.

$L = 10$	i)		ii)		iii)	
n / δ	0.90	0.95	0.90	0.95	0.90	0.95
500	0.944	0.970	0.932	0.976	0.924	0.968
1000	0.932	0.963	0.918	0.968	0.917	0.961
2000	0.911	0.955	0.910	0.965	0.912	0.957
$L = 30$	i)		ii)		iii)	
n / δ	0.90	0.95	0.90	0.95	0.90	0.95
500	0.867	0.901	0.862	0.904	0.839	0.893
1000	0.892	0.905	0.880	0.911	0.890	0.906
2000	0.899	0.923	0.887	0.915	0.894	0.918
$L = 50$	i)		ii)		iii)	
n / δ	0.90	0.95	0.90	0.95	0.90	0.95
500	0.836	0.842	0.833	0.861	0.825	0.883
1000	0.851	0.872	0.846	0.907	0.844	0.902
2000	0.886	0.916	0.881	0.919	0.890	0.924
CH	i)		ii)		iii)	
n / δ	0.90	0.95	0.90	0.95	0.90	0.95
500	0.831	0.911	0.824	0.902	0.845	0.892
1000	0.846	0.917	0.841	0.909	0.862	0.918
2000	0.883	0.930	0.886	0.922	0.896	0.941

To evaluate the test performance under the alternative, we consider $\delta = 0.95, 0.90$ and the cases

$$\Delta^* \in \{0.05, 0.1, 0.2\}$$

with a shock period of $\tau_2^* - \tau_1^* \in \{0.1, 0.2\}$, where we have chosen $\tau_1^* = 0.5$. The choice of τ_1^* does not exert a significant influence on the performance of the test; therefore we do not present the simulation results for different τ_1^* here. The test results of different scenarios can be found in Table A.2 and A.3. It can be seen that our test shows good power under the alternative hypothesis, which is robust against different choices of break sizes and time length of breaks. We also find that as the sample size increases, the power drastically increases. In Table A.3, we show the power of the CH test. The CH test sometimes shows a better power than our test. As our test serves a different purpose of identifying mildly explosive segments, we conclude that our method is complementary to the CH test.

Table A.2. The rejection rate of the test (test power) under the alternative H_1 with different Δ^* and $\tau_2^* - \tau_1^*$. $\delta = 95\%$

$L = 50$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.398	0.685	0.900	0.899	0.937	0.952
	1000	0.681	0.696	0.912	0.926	1.000	1.000
	2000	0.994	0.908	0.917	1.000	1.000	1.000
ii)	500	0.397	0.543	0.837	0.908	0.946	0.964
	1000	0.846	0.865	0.875	0.969	0.981	1.000
	2000	0.838	0.864	0.922	1.000	1.000	1.000
iii)	500	0.436	0.486	0.858	0.915	0.953	0.974
	1000	0.811	0.817	0.863	0.957	0.990	1.000
	2000	0.829	0.833	0.939	1.000	1.000	1.000
$L = 30$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.321	0.436	0.828	0.886	0.917	0.939
	1000	0.583	0.599	0.898	0.904	0.953	0.969
	2000	0.812	0.826	0.912	1.000	1.000	1.000
ii)	500	0.357	0.475	0.830	0.902	0.930	0.949
	1000	0.798	0.804	0.870	0.963	0.988	0.997
	2000	0.805	0.819	0.904	1.000	1.000	1.000
iii)	500	0.431	0.577	0.845	0.901	0.949	0.958
	1000	0.783	0.826	0.860	0.956	0.981	0.998
	2000	0.792	0.831	0.924	1.000	1.000	1.000
$L = 10$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.148	0.154	0.414	0.843	0.869	0.881
	1000	0.176	0.195	0.491	0.877	0.938	0.896
	2000	0.404	0.473	0.716	0.943	0.976	1.000
ii)	500	0.271	0.424	0.752	0.812	0.827	0.941
	1000	0.414	0.668	0.799	0.830	0.853	0.962
	2000	0.589	0.715	0.836	1.000	1.000	1.000
iii)	500	0.193	0.301	0.803	0.828	0.869	0.914
	1000	0.288	0.716	0.834	0.851	0.916	0.937
	2000	0.405	0.817	0.906	1.000	1.000	1.000

Table A.3. The rejection rate of the test (test power) under the alternative H_1 with different Δ^* and $\tau_2^* - \tau_1^*$. $\delta = 90\%$

$L = 50$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.441	0.725	0.932	0.941	0.986	0.990
	1000	0.722	0.771	0.956	0.955	1.000	1.000
	2000	0.989	0.924	0.971	1.000	1.000	1.000
ii)	500	0.513	0.782	0.884	0.951	0.964	0.976
	1000	0.890	0.903	0.913	0.994	1.000	1.000
	2000	0.946	0.987	1.000	1.000	1.000	1.000
iii)	500	0.465	0.754	0.882	0.946	0.967	0.980
	1000	0.832	0.858	0.927	1.000	1.000	1.000
	2000	0.957	0.976	0.998	1.000	1.000	1.000
$L = 30$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.388	0.506	0.834	0.914	0.937	0.950
	1000	0.657	0.693	0.901	0.925	0.996	0.997
	2000	0.883	0.894	0.923	1.000	1.000	1.000
ii)	500	0.395	0.701	0.875	0.932	0.957	0.962
	1000	0.808	0.837	0.900	0.997	0.998	0.999
	2000	0.939	0.972	0.995	1.000	1.000	1.000
iii)	500	0.435	0.486	0.858	0.934	0.951	0.977
	1000	0.807	0.817	0.863	0.997	0.998	0.999
	2000	0.941	0.831	0.945	1.000	1.000	1.000
$L = 10$		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.187	0.205	0.486	0.863	0.911	0.893
	1000	0.260	0.303	0.530	0.895	0.946	0.917
	2000	0.456	0.547	0.814	0.957	1.000	1.000
ii)	500	0.336	0.491	0.762	0.839	0.861	0.956
	1000	0.532	0.755	0.858	0.858	0.889	0.993
	2000	0.651	0.824	0.849	1.000	1.000	1.000
iii)	500	0.275	0.608	0.851	0.873	0.923	0.936
	1000	0.420	0.802	0.869	0.892	0.978	0.945
	2000	0.571	0.930	0.956	1.000	1.000	1.000

Table A.4. The rejection rate of the test (test power) under the alternative H_1 with different Δ^* and $\tau_2^* - \tau_1^*$. ($L = 30$), $\delta = 95\%$, 90%, CH.

		$\tau_2^* - \tau_1^* = 0.1$			$\tau_2^* - \tau_1^* = 0.2$		
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.346	0.418	0.810	0.875	0.886	0.939
	1000	0.563	0.541	0.893	0.911	0.913	0.995
	2000	0.805	0.834	0.918	1.000	1.000	1.000
ii)	500	0.334	0.416	0.852	0.881	0.915	0.968
	1000	0.783	0.875	0.880	0.939	0.968	0.987
	2000	0.805	0.882	0.901	1.000	1.000	1.000
iii)	500	0.419	0.500	0.763	0.806	0.919	0.932
	1000	0.625	0.776	0.839	0.847	0.936	0.981
	2000	0.794	0.824	0.884	1.000	1.000	1.000
	n/Δ^*	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$	$\Delta^* = 0.05$	$\Delta^* = 0.1$	$\Delta^* = 0.2$
i)	500	0.428	0.434	0.912	0.894	0.922	0.984
	1000	0.609	0.697	0.898	0.943	0.938	1.000
	2000	0.864	0.916	0.929	1.000	1.000	1.000
ii)	500	0.414	0.539	0.894	0.907	0.926	0.991
	1000	0.817	0.893	0.915	0.954	0.980	0.996
	2000	0.826	0.897	0.963	1.000	1.000	1.000
iii)	500	0.452	0.596	0.849	0.856	0.922	0.954
	1000	0.768	0.847	0.875	0.953	0.961	0.998
	2000	0.872	0.850	0.914	1.000	1.000	1.000

B. TECHNICAL PROOFS AND LEMMATA

For some sequence $(y_j)_{j\in\mathbb{N}}$ of real numbers and some sequence $(\chi_j)_{j\in\mathbb{N}}$ of nonnegative real numbers, we define the weighted seminorm

$$|y|_{\chi,q} := \Big(\sum_{j=1}^{\infty} \chi_j |y_j|^q\Big)^{1/q}.$$

For some random variable Z, we define $\|Z\|_q := (\mathbb{E}|Z|^q)^{1/q}$. If $\|\cdot\|_q$ is applied to a matrix, this is meant by a component-wise operation. For the i.i.d. random variables ζ_i, $i \in \mathbb{Z}$ used in the model definition (1.1), let $\mathcal{F}_i := (\zeta_i, \zeta_{i-1}, ...)$. With some abuse of notation, we refer to $\mathcal{F}_i$ also as the σ algebra generated by the entries of $\mathcal{F}_i$. For $X_i = h(\mathcal{F}_i)$, $i \in \mathbb{Z}$ with some measurable function h, we define the functional dependence measure (cf. Wu and Shao (2004)) as

$$\delta_q(i) := \|X_i - X_i^*\|_q,$$

where $X_i^* = h(\mathcal{F}_i^*)$ and $\mathcal{F}_i^* := (\zeta_i, ..., \zeta_1, \zeta_0^*, \zeta_{-1}, \zeta_{-2}, ...)$ with ζ_0^* being an independent copy of ζ_0. It is worth noting that in the following context we add the superscript X as $\delta_q^X(i)$ for the dependence measure of the process X.

B.1. Existence of GARCH models

PROPOSITION B.1. (EXISTENCE OF THE GARCH MODEL) *If Assumption 3.1 and H_0 holds, then the following statements hold true.*

(i) *(1.1) has a unique stationary solution $X_i^2 = H(\mathcal{F}_i)$, $i \in \mathbb{Z}$.*
(ii) *There exists $q > 0$ with $\|X_0^2\|_q \le D$ and $\delta_q^{X^2}(k) = O(c^k)$ (recall that $\delta_q^{X^2}(k)$ is the functional dependence measure of X^2) for some $0 < c < 1$.*
(iii) *$\lambda_{max}(B(\theta)) < 1$, where*

$$B(\theta) = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \dots & \beta_s \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

PROOF. (PROOF OF PROPOSITION B.1:) Following the proof of Lemma 2.3 in Berkes et al. (2003), there exists $m \in \mathbb{N}$ such that $\mathbb{E}\log|A_m(\theta^*)A_{m-1}(\theta^*)\dots A_1(\theta^*)|_2 < 0$. The function $[0,\delta) \to [0,\infty)$, $s(t,\theta^*) = \mathbb{E}|A_m(\theta^*)A_{m-1}(\theta^*)\dots A_1(\theta^*)|_2^t$ fulfills $s'(0,\theta^*) = \mathbb{E}\log|A_m(\theta^*)A_{m-1}(\theta^*)\dots A_1(\theta^*)|_2 < 0$, thus $t \mapsto s(t,\theta^*)$ decreases in a neighborhood of 0. Since $s(0) = 1$, this implies that there exists $0 < q < \delta$ such that

$$\mathbb{E}|A_m(\theta^*)A_{m-1}(\theta^*)\dots A_1(\theta^*)|_2^q = s(q,\theta^*) < 1. \qquad \text{(B.1)}$$

Define

$$P_i := (X_i^2, \dots, X_{i-r+1}^2, \sigma_i^2, \dots, \sigma_{i-s+1}^2)',$$
$$a_i(\theta^*) := (\alpha_0^*\zeta_i^2, 0, \dots, 0, \alpha_0^*, 0, \dots, 0)'.$$

Following Section 3.1 in Wu and Min (2005), the model (1.1) admits the representation

$$P_i = A_i(\theta^*)P_{i-1} + a_i(\theta^*). \qquad \text{(B.2)}$$

Therefore, $P_i = G_{\zeta_i,\theta^*}(P_{i-1})$ with $G_{\zeta_i,\theta^*}(y) = A_i(\theta^*) \cdot y + a_i(\theta^*)$. Let $W_n(y,\theta^*) := G_{\zeta_n,\theta^*} \circ G_{\zeta_{n-1},\theta^*} \circ ... \circ G_{\zeta_1,\theta^*}(y)$. Then we have

$$W_n(y,\theta^*) - W_n(y',\theta^*) = A_n(\theta^*)A_{n-1}(\theta^*) \cdot ... \cdot A_1(\theta^*) \cdot (y - y').$$

Using the submultiplicativity of $|\cdot|_2$, we therefore have with (B.1) and some suitable constant $C > 0$:

$$\begin{aligned}\| |W_n(y,\theta^*) - W_n(y',\theta^*)|_2 \|_q &\le \| |A_n(\theta^*)A_{n-1}(\theta^*) \cdot ... \cdot A_1(\theta^*)|_2 \|_q |y - y'|_2 \\ &\le C(s(q,\theta^*)^{q/m})^n |y - y'|_2.\end{aligned}$$

By Theorem 2 in Wu and Shao (2004), we obtain existence and a.s. uniqueness of $X_i^2 = H(t,\mathcal{F}_i)$, $\|X_0^2\|_q < \infty$ and $\delta_q^{X^2}(k) = O(c^k)$ with some $0 < c < 1$, i.e. (i) and (ii). (iii) is due to Proposition 1 in Francq and Zakoïan (2004).

B.2. Proofs for asymptotic theory

Observe that

$$L^c_{n,\tau_1,\tau_2}(\theta) = L^c_{n,\tau_2}(\theta) - L^c_{n,\tau_1}(\theta), \qquad L^c_{n,r}(\theta) := \frac{1}{n}\sum_{i=1}^{\lfloor nr \rfloor} \ell(X_i^2, Y_i^c, \theta).$$

Define $L(\theta) := \mathbb{E}\ell(X_i^2, Y_i, \theta)$. This is well-defined due to $\mathbb{E}\max\{-\ell(X_i^2, Y_i, \theta), 0\} < \infty$ (cf. Francq and Zakoïan (2004), Proof of Theorem 2.1).

To prove Theorem 3.1, we introduce some notation. For some sequence of real-valued random variables W_n, we write $W_n \xrightarrow{p} \infty$ if for each $M \in \mathbb{N}$, $\mathbb{P}(W_n < M) \to 0$ $(n \to \infty)$.

LEMMA B.1. *Let Assumption 3.1 hold. Let $\kappa > 0$, and $R := \{(\tau_1,\tau_2) \in [0,1]^2 : \tau_1 < \tau_2, |\tau_1 - \tau_2| \ge \kappa\}$. For fixed $k \in \mathbb{N}$, let $V_k(\theta) := \{\theta' \in \Theta : |\theta' - \theta|_1 < 1/k\}$. Define $W_i^{(k)}(\theta) := \inf_{\theta' \in V_k(\theta)} \ell(X_i^2, Y_i, \theta)$. Then:*

(i) $\mathbb{E}W_1^{(k)}(\theta) \in \mathbb{R} \cup \{\infty\}$ *and*

$$\liminf_{n\to\infty} \inf_{(\tau_1,\tau_2)\in R} \frac{1}{n(\tau_2 - \tau_1)} \sum_{i=\lfloor n\tau_1 \rfloor + 1}^{\lfloor n\tau_2 \rfloor} W_i^{(k)}(\theta) \ge \mathbb{E}W_1^{(k)}(\theta) \quad a.s.$$

(ii) $L(\theta^*)$ *is finite and*

$$\sup_{(\tau_1,\tau_2)\in R} \frac{1}{n(\tau_2 - \tau_1)} \sum_{i=\lfloor n\tau_1 \rfloor + 1}^{\lfloor n\tau_2 \rfloor} \ell(X_i^2, Y_i, \theta^*) \to L(\theta^*) \quad a.s.$$

PROOF. (i) Fix some $M \in \mathbb{N}$. It holds that $\mathbb{E}W_1^{(k)} \in \mathbb{R} \cup \{\infty\}$ and $\mathbb{E}[W_1^{(k)} \wedge M] \in \mathbb{R}$ since $\mathbb{E}\max\{-\ell(X_1^2, Y_1, \theta), 0\} < \infty$. Define $S_m := \sum_{i=1}^m W_i^{(k)}(\theta) \wedge M$. By the ergodic theorem, we have

$$\lim_{m\to\infty} \frac{1}{m} S_m = \mathbb{E}S_1 \quad a.s.$$

It holds that $\lfloor n\tau_2 \rfloor \to \infty$ uniformly in $\tau_2 \in [\kappa, 1]$, and thus

$$\sup_{\tau_2 \in [\kappa,1]} \Big| \frac{S_{\lfloor n\tau_2 \rfloor}}{\lfloor n\tau_2 \rfloor} - \mathbb{E}S_1 \Big| \to 0 \quad a.s. \tag{B.3}$$

Furthermore we have that $\frac{S_m}{m}$ is a.s. bounded. We conclude that

$$\sup_{\tau_1\in[0,1]}\Big|\tau_1\cdot\Big(\frac{S_{\lfloor n\tau_1\rfloor}}{\lfloor n\tau_1\rfloor}-\mathbb{E}S_1\Big)\Big|\to 0\quad a.s. \tag{B.4}$$

[Proof: Fix some $\omega\in\Omega$ and let $C>0$ be such that $|\frac{S_m(\omega)}{m}|\le C$ for all $m\in\mathbb{N}$. For $\varepsilon>0$,

$$\sup_{\tau_1\in[0,1]}\Big|\tau_1\cdot\Big(\frac{S_{\lfloor n\tau_1\rfloor}}{\lfloor n\tau_1\rfloor}-\mathbb{E}S_1\Big)\Big|\le\sup_{\tau_1\in[\frac{\varepsilon}{C},1]}\Big|\tau_1\cdot\Big(\frac{S_{\lfloor n\tau_1\rfloor}}{\lfloor n\tau_1\rfloor}-\mathbb{E}S_1\Big)\Big|+\sup_{\tau_1\in[0,\frac{\varepsilon}{C})}\Big|\tau_1\cdot\Big(\frac{S_{\lfloor n\tau_1\rfloor}}{\lfloor n\tau_1\rfloor}-\mathbb{E}S_1\Big)\Big|$$

The second term is bounded by $\frac{\varepsilon}{C}\cdot C=\varepsilon$, while for the first term we can choose n large enough such that it is $\le\varepsilon$.]

With the decomposition

$$\frac{1}{n(\tau_2-\tau_1)}\sum_{i=\lfloor n\tau_1\rfloor+1}^{\lfloor n\tau_2\rfloor}W_i^{(k)}(\theta)\wedge M=\frac{1}{\tau_2-\tau_1}\Big[\frac{\lfloor n\tau_2\rfloor}{n}\cdot\frac{S_{\lfloor n\tau_2\rfloor}}{\lfloor n\tau_2\rfloor}-\frac{\lfloor n\tau_1\rfloor}{n\tau_1}\cdot\tau_1\cdot\frac{S_{\lfloor n\tau_1\rfloor}}{\lfloor n\tau_1\rfloor}\Big]$$

and (B.3), (B.4), $\sup_{\tau_2\in[\kappa,1]}|\frac{\lfloor n\tau_2\rfloor}{n}-\tau_2|\le n^{-1}$ we obtain

$$\inf_{(\tau_1,\tau_2)\in R}\frac{1}{n(\tau_2-\tau_1)}\sum_{i=\lfloor n\tau_1\rfloor+1}^{\lfloor n\tau_2\rfloor}W_i^{(k)}(\theta)\wedge M\to\mathbb{E}[W_1^{(k)}(\theta)\wedge M]\quad a.s.$$

Since $W_i^{(k)}(\theta)\ge W_i^{(k)}\wedge M$ and applying $M\to\infty$ on the r.h.s., we obtain

$$\liminf_{n\to\infty}\inf_{(\tau_1,\tau_2)\in R}\frac{1}{n(\tau_2-\tau_1)}\sum_{i=\lfloor n\tau_1\rfloor+1}^{\lfloor n\tau_2\rfloor}W_i^{(k)}(\theta)\ge\mathbb{E}W_1^{(k)}(\theta)\quad a.s.$$

(ii) The argument follows the same lines as (i). We obtain convergence since no truncation with M is needed.

PROOF. (PROOF OF THEOREM 3.1) We make use of some results obtained in the proof of Theorem 2.1. in Francq and Zakoïan (2004). It was shown therein that $L(\theta):=\mathbb{E}\ell(X_i^2,Y_i,\theta)$ fulfills

$$\mathbb{E}|\ell(X_i^2,Y_i,\theta^*)|<\infty,\qquad\forall\theta\ne\theta^*:\quad L(\theta)>L(\theta^*). \tag{B.5}$$

Let $k\in\mathbb{N}$. We use the notation from Lemma B.1. Let $\theta\ne\theta^*$. By Beppo-Levi's theorem, we have

$$\mathbb{E}W_1^{(k)}(\theta)\uparrow L(\theta)>L(\theta^*).$$

Thus, for each $\theta\ne\theta^*$ we can find $k(\theta)\in\mathbb{N}$ such that $\mathbb{E}W_1^{(k(\theta))}(\theta)>L(\theta^*)$.

Let $\varepsilon>0$ and $\Theta_\varepsilon:=\{\theta\in\Theta:|\theta-\theta^*|\ge\varepsilon\}$. Then Θ_ε is compact, and there exist finitely many $\theta_1,...,\theta_l$ with $\Theta_\varepsilon\subset\bigcup_{i=1}^l V_{k(\theta_i)}(\theta_i)$. Let

$$\delta:=\min\{\inf_{i=1,\dots,l}\mathbb{E}W_1^{(k(\theta_i))}(\theta_i)-L(\theta^*),1\}>0.$$

Suppose that $\sup_{(\tau_1,\tau_2)\in R}|\hat\theta_{n,\tau_1,\tau_2}-\theta^*|\ge\varepsilon$. By the minimal property of $\hat\theta_{n,\tau_1,\tau_2}$ and dividing by $\tau_2 - \tau_1$, we conclude that

$$
\begin{aligned}
0 &\geq \inf_{(\tau_1,\tau_2)\in R} \frac{1}{\tau_2-\tau_1}\big\{L^c_{n,\tau_1,\tau_2}(\hat{\theta}_{n,\tau_1,\tau_2}) - L^c_{n,\tau_1,\tau_2}(\theta^*)\big\} \\
&= \inf_{i=1,\ldots,l} \inf_{(\tau_1,\tau_2)\in R} \inf_{\theta'\in V_{k(\theta_i)}(\theta_i)} \frac{1}{\tau_2-\tau_1}\{L^c_{n,\tau_1,\tau_2}(\theta') - L^c_{n,\tau_1,\tau_2}(\theta^*)\} \\
&\geq \inf_{i=1,\ldots,l} \inf_{(\tau_1,\tau_2)\in R} \inf_{\theta'\in V_{k(\theta_i)}(\theta_i)} \frac{L^c_{n,\tau_1,\tau_2}(\theta')}{\tau_2-\tau_1} - \sup_{(\tau_1,\tau_2)\in R} \frac{L^c_{n,\tau_1,\tau_2}(\theta^*)}{\tau_2-\tau_1}
\end{aligned}
\tag{B.6}
$$

We furthermore have:

$$
\begin{aligned}
&\inf_{(\tau_1,\tau_2)\in R} \inf_{\theta'\in V_{k(\theta_i)}(\theta_i)} \frac{L^c_{n,\tau_1,\tau_2}(\theta')}{\tau_2-\tau_1} \\
&\geq \inf_{(\tau_1,\tau_2)\in R} \inf_{\theta'\in V_{k(\theta_i)}(\theta_i)} \frac{L_{n,\tau_1,\tau_2}(\theta')}{\tau_2-\tau_1} - \kappa\cdot \sup_{(\tau_1,\tau_2)\in R} \sup_{\theta'\in\Theta} |L^c_{n,\tau_1,\tau_2}(\theta') - L_{n,\tau_1,\tau_2}(\theta')| \\
&\geq \inf_{(\tau_1,\tau_2)\in R} \frac{1}{n(\tau_2-\tau_1)} \sum_{i=\lfloor n\tau_1\rfloor}^{\lfloor n\tau_2\rfloor} W_i^{(k(\theta_i))}(\theta_i) \\
&\qquad -\kappa\cdot \sup_{(\tau_1,\tau_2)\in R} \sup_{\theta'\in\Theta} |L^c_{n,\tau_1,\tau_2}(\theta') - L_{n,\tau_1,\tau_2}(\theta')|.
\end{aligned}
\tag{B.7}
$$

By Lemma B.2(i),

$$
\mathbb{P}\Big(\sup_{(\tau_1,\tau_2)\in R} \sup_{\theta'\in\Theta} |L^c_{n,\tau_1,\tau_2}(\theta') - L_{n,\tau_1,\tau_2}(\theta')| > \frac{\delta}{8}\Big) = o(1). \tag{B.8}
$$

By Lemma B.1(i),

$$
\mathbb{P}\Big(\inf_{i=1,\ldots,l} \inf_{(\tau_1,\tau_2)\in R} \frac{1}{n(\tau_2-\tau_1)} \sum_{i=\lfloor n\tau_1\rfloor}^{\lfloor n\tau_2\rfloor} W_i^{(k(\theta_i))}(\theta_i) \leq L(\theta^*) + \frac{\delta}{2}\Big) = o(1). \tag{B.9}
$$

By Lemma B.1(ii),

$$
\mathbb{P}\Big(\sup_{(\tau_1,\tau_2)\in R} \frac{L^c_{n,\tau_1,\tau_2}(\theta^*)}{\tau_2-\tau_1} \geq L(\theta^*) + \frac{\delta}{8}\Big) = o(1). \tag{B.10}
$$

Inserting (B.8), (B.9) into (B.7) and afterwards using (B.10), we have

$$
\begin{aligned}
&\mathbb{P}\Big(\sup_{(\tau_1,\tau_2)\in R} |\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*| \geq \varepsilon\Big) \\
&\leq \mathbb{P}\Big(0 \geq \inf_{i=1,\ldots,l} \inf_{(\tau_1,\tau_2)\in R} \inf_{\theta'\in V_{k(\theta_i)}(\theta_i)} \frac{L^c_{n,\tau_1,\tau_2}(\theta')}{\tau_2-\tau_1} - \sup_{(\tau_1,\tau_2)\in R} \frac{L^c_{n,\tau_1,\tau_2}(\theta^*)}{\tau_2-\tau_1}\Big) \\
&\leq \mathbb{P}\Big(0 \geq (L(\theta^*) + \frac{\delta}{2}) - \frac{\delta}{8} - (L(\theta^*) + \frac{\delta}{8}) = \frac{\delta}{4}\Big) + o(1) = o(1),
\end{aligned}
$$

showing the assertion.

Proof. (Proof of Proposition 3.1) By Lemma B.3 and Lemma B.2, we have

$$
\sup_{\theta\in\Theta} \big|L^c_{n,\tau_1^*,\tau_2^*}(\theta) - \tilde{L}_{n,\tau_1^*,\tau_2^*}(\theta)\big| \to 0 \quad a.s., \tag{B.11}
$$

where

$$\tilde{L}_{n,\tau_1^*,\tau_2^*}(\theta) = \frac{1}{n}\sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor} \ell(\tilde{X}_i^2, \tilde{Y}_i, \theta)$$

and $\tilde{X}_i$ denotes a GARCH(r,s) process with constant parameters $\theta(i) \equiv \bar{\theta}^* := \theta^* + H\Delta^*$. The proof is now a simpler version of the proof of Lemma B.1 and Theorem 3.1. For the sake of completeness, we provide here the main steps. For fixed $k \in \mathbb{N}$, let $V_k(\theta) := \{\theta' \in \Theta : |\theta' - \theta|_1 < \frac{1}{k}\}$ and $W_i^{(k)} := \inf_{\theta' \in V_k(\theta)} \ell(\tilde{X}_i^2, \tilde{Y}_i, \theta)$. As in Lemma B.1, we obtain that

i) $\mathbb{E}W_1^{(k)}(\theta) \in \mathbb{R} \cup \{\infty\}$ and

$$\liminf_{n\to\infty} \frac{1}{n(\tau_2^* - \tau_1^*)} \sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor} W_i^{(k)}(\theta) \geq \mathbb{E}W_1^{(k)}(\theta) \quad a.s., \tag{B.12}$$

ii) $\tilde{L}(\bar{\theta}^*)$ is finite and

$$\frac{1}{n(\tau_2^* - \tau_1^*)} \sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor} \ell(\tilde{X}_i^2, \tilde{Y}_i, \bar{\theta}^*) \to \tilde{L}(\bar{\theta}^*), \tag{B.13}$$

where $\tilde{L}(\theta) := \mathbb{E}\ell(\tilde{X}_i^2, \tilde{Y}_i, \theta)$ and $\forall \theta \neq \bar{\theta}^*$, $\tilde{L}(\theta) > \tilde{L}(\bar{\theta}^*)$.

Let $\varepsilon > 0$. For $\theta \neq \theta^*$, we now construct $k(\theta) \in \mathbb{N}$ such that $\mathbb{E}W_1^{(k(\theta))} > \tilde{L}(\bar{\theta}^*)$ and define $\bigcup_{i=1}^{l} V_{k(\theta_i)}(\theta_i)$ as a finite covering of $\Theta_\varepsilon := \{\theta \in \Theta : |\theta - \bar{\theta}^*| \geq \varepsilon\}$. Define

$$\delta := \min\{\inf_{i=1,\dots,l} \mathbb{E}W_1^{(k(\theta_i))}(\theta_i) - \tilde{L}(\bar{\theta}^*), 1\} > 0.$$

Suppose that $|\hat{\theta}_{n,\tau_1^*,\tau_2^*} - \bar{\theta}^*| \geq \varepsilon$. By the minimal property of $\hat{\theta}_{n,\tau_1^*,\tau_2^*}$, we have

$$\begin{aligned} 0 &\geq L^c_{n,\tau_1^*,\tau_2^*}(\hat{\theta}_{n,\tau_1^*,\tau_2^*}) - L^c_{n,\tau_1^*,\tau_2^*}(\bar{\theta}^*) \\ &\geq \inf_{i=1,\dots,l} \inf_{\theta' \in V_{k(\theta_i)}(\theta_i)} L^c_{n,\tau_1^*,\tau_2^*}(\theta') - L^c_{n,\tau_1^*,\tau_2^*}(\bar{\theta}^*). \end{aligned}$$

For n large enough, (B.11) and (B.12), (B.13) imply

$$\begin{aligned} 0 &\geq \inf_{i=1,\dots,l} \inf_{\theta' \in V_{k(\theta_i)}(\theta_i)} \tilde{L}_{n,\tau_1^*,\tau_2^*}(\theta') - \tilde{L}_{n,\tau_1^*,\tau_2^*}(\bar{\theta}^*) - \frac{\delta}{4} \\ &\geq \inf_{i=1,\dots,l} \mathbb{E}W_1^{(k(\theta_i))}(\theta_i) - \tilde{L}(\bar{\theta}^*) - \frac{\delta}{2} \geq \frac{\delta}{2}, \end{aligned}$$

a contradiction. Thus for n large enough, $|\hat{\theta}_{n,\tau_1^*,\tau_2^*} - \bar{\theta}^*| \leq \varepsilon$.

Proof. (Proof of Theorem 3.2) Let $\kappa > 0$ und define $R := \{(\tau_1, \tau_2) \in [0,1]^2 : \tau_1 < \tau_2, |\tau_1 - \tau_2| \geq \kappa\}$. By Theorem 3.1, we have

$$\sup_{(\tau_1,\tau_2)\in R} |\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*|_1 \to 0 \quad a.s. \tag{B.14}$$

Therefore, $\hat{\theta}_{n,\tau_1,\tau_2} \in int(\Theta)$ uniformly in $(\tau_1, \tau_2) \in R$ for n large enough. Thus there exists $\bar{\Theta} \subset int(\Theta)$ with $\hat{\theta}_{n,\tau_1,\tau_2} \in \bar{\Theta}$ (for n large enough, $(\tau_1, \tau_2) \in R$.

By a Taylor expansion, we have

$$\begin{aligned}\hat{\theta}_{n,\tau_1,\tau_2} - \theta^* &= -[\nabla_\theta^2 L_{n,\tau_1,\tau_2}^c(\bar{\theta})]^{-1}\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*)\\ &= -[(\tau_2-\tau_1)V(\theta^*) + T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2})]^{-1}\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*), \qquad \text{(B.15)}\end{aligned}$$

where $T_{n,\tau_1,\tau_2}(\bar{\theta}_n) = \nabla_\theta^2 L_{n,\tau_1,\tau_2}^c(\bar{\theta}_n) - (\tau_2-\tau_1)V(\theta^*)$ and $\bar{\theta}_{n,\tau_1,\tau_2} \in \Theta$ with $|\bar{\theta}_{n,\tau_1,\tau_2} - \theta^*|_1 \leq |\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*|_1$. By Lemma B.2 and Lemma B.5 and $|\frac{\lfloor n\tau_2\rfloor - \lfloor n\tau_1\rfloor}{n} - (\tau_1-\tau_2)| \leq 2n^{-1}$, we have

$$\sup_{(\tau_1,\tau_2)\in R}|T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2})|_1 \leq \sup_{(\tau_1,\tau_2)\in R}|V(\bar{\theta}_{n,\tau_1,\tau_2}) - V(\theta^*)|_1 + O_p(\log(n)^{3/2}n^{-1/2}). \qquad \text{(B.16)}$$

By Lemma B.6(ii) applied to $p = q$ and $\bar{\Theta}$, we obtain $\iota > 0$, $C > 0$ and $\rho \in (0,1)$ such that for $|\theta - \theta^*|_1 < \iota$,

$$\begin{aligned}|V(\theta) - V(\theta^*)|_1 &\leq C(1 + \||Y_i|_{(\rho^j)_j,q}^q\|_1)(1+\mathbb{E}\zeta_0^2)\cdot|\theta-\theta^*|_1\\ &\leq C(1 + \frac{D^q}{1-\rho})(1+\mathbb{E}\zeta_0^2)\cdot|\theta-\theta^*|_1 =: \tilde{C}\cdot|\theta-\theta^*|_1. \qquad \text{(B.17)}\end{aligned}$$

Since $\mathbb{E}\nabla_\theta\ell(Y_i, X_i^2, \theta^*) = 0$ (cf. Proposition B.2(iii)), Lemma B.2 and Lemma B.5, we have

$$\sup_{(\tau_1,\tau_2)\in R}|\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*)|_1 = O_p(\log(n)^{3/2}n^{-1/2}). \qquad \text{(B.18)}$$

Inserting (B.14) into (B.17), we obtain $\sup_{(\tau_1,\tau_2)\in R}|T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2})|_2 = o_p(1)$. From (B.15) and (B.18) we obtain

$$\sup_{(\tau_1,\tau_2)\in R}|\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*|_1 = O_p(\log(n)^{3/2}n^{-1/2}). \qquad \text{(B.19)}$$

By (B.15), we have

$$\begin{aligned}&\left|\hat{\theta}_{n,\tau_1,\tau_2} - \theta^* + ((\tau_2-\tau_1)V(\theta^*))^{-1}\cdot\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*)\right|_2\\ &\leq |((\tau_2-\tau_1)V(\theta^*) + T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2}))^{-1}|_1\cdot|T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2})|_1\\ &\quad\times|((\tau_2-\tau_1)V(\theta^*))^{-1}\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*)|_1. \qquad \text{(B.20)}\end{aligned}$$

Using (B.16), (B.17) and (B.19), we have $\sup_{(\tau_1,\tau_2)\in R}|T_{n,\tau_1,\tau_2}(\bar{\theta}_{n,\tau_1,\tau_2})|_1 = O_p(\log(n)^{3/2}n^{-1/2})$. Inserting this and (B.18) into (B.20), we obtain

$$\sup_{(\tau_1,\tau_2)\in R}\left|\hat{\theta}_{n,\tau_1,\tau_2} - \theta^* + ((\tau_2-\tau_1)V(\theta^*))^{-1}\cdot\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*)\right|_1 = O_p(\log(n)^3 n^{-1}).$$

Since $\sup_{(\tau_1,\tau_2)\in R}|\nabla_\theta L_{n,\tau_1,\tau_2}^c(\theta^*) - \nabla_\theta L_{n,\tau_1,\tau_2}(\theta^*)|_1 = O_p(n^{-1})$ by Lemma B.2, the proof is finished.

PROOF. (PROOF OF THEOREM 3.3) Let $R := \{(\tau_1,\tau_2)\in[0,1]^2 : \tau_1 < \tau_2, |\tau_1-\tau_2| \geq \kappa\}$. By Lemma B.4 (applied with $M = 2 + \frac{a'}{4}$, $a = \frac{a'}{4}$), we obtain $C > 0$, $\rho \in (0,1)$ and $\iota > 0$ such that

$$\delta_M^{\nabla_\theta\ell(Z,\theta^*)}(k) \leq C\rho^k,$$

and thus (component-wise) $\Delta_M^{\nabla_\theta\ell(Z,\theta^*)}(k) \leq \frac{C}{1-\rho}\rho^k$. Let $W_i := -V(\theta^*)^{-1}\nabla_\theta\ell(X_i^2, Y_i, \theta^*)$ and $S(j) := \sum_{i=1}^j W_i$. By Proposition B.2(iii), $\mathbb{E}W_i = 0$ and

$$\Sigma = \mathrm{Cov}(W_i) = V(\theta^*)^{-1}I(\theta^*)V(\theta^*)^{-1}.$$

By Corollary 1 in Wu and Zhou (2011), there exists a richer probability space and i.i.d. $V_1, V_2, \ldots \sim N(0, I_{(r+s+1)\times(r+s+1)})$, a process $(\hat{S}(i))_{i=1,\ldots,n}$ and $S^0(i) = \sum_{j=1}^{i} V_j$ such that $(S(i))_{i=1,\ldots,n} \stackrel{d}{=} (\hat{S}(i))_{i=1,\ldots,n}$ and

$$\max_{i=1,\ldots,n} |\hat{S}(i) - \Sigma^{1/2} S^0(i)| = O_p(n^{1/\min\{M,4\}} \log(n)^{3/2}).$$

With Theorem 3.2 we obtain:

$$\sup_{(\tau_1,\tau_2)\in R} \big|\sqrt{n(\tau_2-\tau_1)}(\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*) - (n(\tau_2-\tau_1))^{-1/2}(S(\lfloor n\tau_2\rfloor) - S(\lfloor n\tau_1\rfloor))\big|$$
$$= O_p(\log(n)^3 n^{-1/2}). \tag{B.21}$$

By the Gaussian approximation result above, on $D(R)^{r+s+1}$

$$(n(\tau_2-\tau_1))^{-1/2}(S(\lfloor n\tau_2\rfloor) - S(\lfloor n\tau_1\rfloor)) \stackrel{d}{=} (n(\tau_2-\tau_1))^{-1/2}(\hat{S}(\lfloor n\tau_2\rfloor) - \hat{S}(\lfloor n\tau_1\rfloor)) \tag{B.22}$$

and

$$\sup_{(\tau_1,\tau_2)\in R} \big|(n(\tau_2-\tau_1))^{-1/2}(\hat{S}(\lfloor n\tau_2\rfloor) - \hat{S}(\lfloor n\tau_1\rfloor))$$
$$-\Sigma^{1/2}\cdot(n(\tau_2-\tau_1))^{-1/2}(S^0(\lfloor n\tau_2\rfloor) - S^0(\lfloor n\tau_1\rfloor))\big|$$
$$= O_p(n^{\frac{1}{\min\{M,4\}}-\frac{1}{2}} \log(n)^{3/2}). \tag{B.23}$$

By Donsker's theorem, it holds in $D[0,1]^{r+s+1}$ that $n^{-1/2}S^0(\lfloor nr\rfloor) \xrightarrow{d} B(r)$ with some standard $(r+s+1)$-dimensional Brownian motion B. We now apply the continuous mapping theorem to

$$\Phi : D[0,1]^{r+s+1} \to D(R)^{r+s+1}, \quad f \mapsto \Big[(\tau_1,\tau_2) \mapsto \frac{\Sigma^{1/2}(f(\tau_2) - f(\tau_1))}{\sqrt{\tau_2-\tau_1}}\Big].$$

Clearly, Φ is continuous with respect to the $\|\cdot\|_\infty$-norm since

$$\|\Phi(f_1) - \Phi(f_2)\|_\infty = \sup_{(\tau_1,\tau_2)\in R}\Big|\frac{\Sigma^{1/2}(f_1(\tau_2) - f_1(\tau_1))}{\sqrt{\tau_2-\tau_1}} - \frac{\Sigma^{1/2}(f_1(\tau_2) - f_1(\tau_1))}{\sqrt{\tau_2-\tau_1}}\Big|$$
$$\leq \frac{2(r+s+1)^2 \max_{i,j}(\Sigma^{1/2})_{ij}}{\sqrt{\kappa}}\|f_1 - f_2\|_\infty.$$

We obtain on $D(R)^{r+s+1}$ that

$$(n(\tau_2-\tau_1))^{-1/2}\Sigma^{1/2}\big\{S^0(\lfloor n\tau_2\rfloor) - S^0(\lfloor n\tau_1\rfloor)\big\} = \Phi(S^0(\lfloor n\cdot\rfloor) \xrightarrow{d} \Phi(B(\cdot))$$
$$= \Sigma^{1/2}\frac{\big\{B(\tau_2) - B(\tau_1)\big\}}{\sqrt{\tau_2-\tau_1}}. \tag{B.24}$$

Combining (B.21), (B.22), (B.23) and (B.24), we obtain the result.

Proof. (Proof of Proposition 3.2) Using similar arguments as in the proof of Theorem 3.1 (now with $\tau_1 \geq \kappa'$ instead of $\tau_2 - \tau_1 \geq \kappa$), we obtain

$$\sup_{\tau_1\geq\kappa'} |\hat{\theta}_{n,0,\tau_1} - \theta^*| \xrightarrow{p} 0. \tag{B.25}$$

(i) By Lemma B.2, Lemma B.5 and $|\frac{\lfloor n\tau_1\rfloor}{n\tau_1}-1|\leq n^{-1}$, we have for fixed $\iota>0$,

$$\sup_{|\theta-\theta^*|_2\leq\iota}\sup_{\tau_1\geq\kappa'}|\bar{V}_{n,\tau_1}(\theta)-V(\theta)|_1\xrightarrow{p}0.$$

Plugging in $\theta=\bar{\theta}_{n,\tau_1}$, we obtain

$$\sup_{\tau_1\geq\kappa'}|\bar{V}_{n,\tau_1}(\hat{\theta}_{n,0,\tau_1})-V(\hat{\theta}_{n,0,\tau_1})|_1\xrightarrow{p}0$$

By Lipschitz continuity of $V(\cdot)$ (cf. (B.17)) and (B.25), we obtain the result.
(ii) Define $g(x,y,\theta):=\nabla_\theta\ell(x,y,\theta)\cdot\nabla_\theta\ell(x,y,\theta)'$ and $\tilde{g}_{\tilde{\theta}}(\zeta,y,\theta):=g(R_\zeta(y,\tilde{\theta}),y,\theta)$, where $R_\zeta(y,\theta):=\zeta^2\sigma^2(y,\theta)$. Let $\bar{\Theta}\subset int(\Theta)$ be some compact set. Using Lemma B.6(ii) and (B.55), it is easy to see that for any $p>0$, one can find $\iota>0$, $C>0$ and $\rho\in(0,1)$ such that (component-wise),

$$\begin{aligned}&\sup_{\theta,\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota}|\tilde{g}_{\tilde{\theta}}(\zeta,y,\theta)-\tilde{g}_{\tilde{\theta}}(\zeta,y',\theta)|\\&\leq C(1+|y|^{2p}_{(\rho^j)_j,2p}+|y'|^{2p}_{(\rho^j)_j,2p})|y-y'|^p_{(\rho^j)_j,p}(1+\zeta^2)^2.\end{aligned}\tag{B.26}$$

and

$$\sup_{\theta,\theta',\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota,|\theta'-\tilde{\theta}|_1<\iota}\frac{|\tilde{g}_{\tilde{\theta}}(\zeta,y,\theta)-\tilde{g}_{\tilde{\theta}}(\zeta,y,\theta')|}{|\theta-\theta'|_1}\leq C(1+|y|^p_{(\rho^j)_j,p})(1+\zeta^2)^2.\tag{B.27}$$

In the following we will enlarge C,ρ and reduce ι if necessary without further notice. Note that $\sup_{\theta\in\Theta}|\nabla_\theta(\sigma^2(0,\theta))|<\infty$ and thus (component-wise) $\sup_{\theta\in\Theta}|\nabla_\theta\ell(x,0,\theta)|\leq C(1+|x|)$. With Lemma B.6(iii) we conclude that (component-wise) $\sup_{\theta\in\Theta}|\nabla_\theta\ell(x,y,\theta)|\leq C(1+|y|^2_{(\rho^j)_j,1})(1+|x|)\cdot|y|_{(\rho^j)_j,1}$. Using again Lemma B.6(iii), we obtain

$$\sup_{\theta\in\Theta}|g(x,y,\theta)-g(x,y',\theta)|\leq C(1+|y|^5_{(\rho^j)_j,1}+|y'|^5_{(\rho^j)_j,1})(1+|x|)^2\cdot|y-y'|_{(\rho^j)_j,1}.$$

This shows that g has similar properties as $\nabla^2_\theta\ell$, but with factors $(1+\zeta^2)^2$ in (B.26), (B.27) instead of $(1+\zeta^2)$. Therefore, we obtain the same result as in (i) under the stated moment condition.

Proof. (Proof of Corollary 3.3) By Proposition 3.1, we have

$$|\hat{\theta}_{n,\tau_1^*,\tau_2^*}-(\theta^*+H\Delta^*)|_1\xrightarrow{p}0.\tag{B.28}$$

Furthermore, from Proposition 3.2 we have

$$\bar{\Sigma}_{n,\tau_1^*}\xrightarrow{p}\Sigma.\tag{B.29}$$

Insertion into $\hat{B}_n$ yields

$$\begin{aligned}\hat{B}_n&\geq\hat{B}_n(\tau_1^*,\tau_2^*)\\&=\sqrt{n(\tau_2^*-\tau_1^*)}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2}\{H'\hat{\theta}_{n,\tau_1^*,\tau_2^*}-H'\theta^*\}\\&=\sqrt{n(\tau_2^*-\tau_1^*)}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2}\{H'(\hat{\theta}_{n,\tau_1^*,\tau_2^*}-(\theta^*+H\Delta^*))\}\\&\qquad+\sqrt{n(\tau_2^*-\tau_1^*)}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2}\cdot H'H\Delta^*.\end{aligned}$$

Due to (B.28), the first summand is of smaller order than the second. By (B.29) and since $H'H\Delta^*>0$, the second summand converges to ∞. This proves the assertion.

Proof. (Proof of Theorem 3.4) We first consider the case under H_0. We use the same notation as in the proof of Theorem 3.3, in particular, we use a Gaussian approximation of $S(j) := \sum_{i=1}^{j} W_i$ by some process $\Sigma_H^{1/2} S^0$, where $W_i = -H'V(\theta^*)^{-1}\nabla_\theta \ell(X_i^2, Y_i, \theta^*)$ and $\Sigma_H = H'\Sigma H$. We apply the Bahadur representation from Theorem 3.2 to

$$\hat{\theta}_{n,\tau_1,\tau_2} - \hat{\theta}_{n,0,\tau_1} = \{\hat{\theta}_{n,\tau_1,\tau_2} - \theta^*\} - \{\hat{\theta}_{n,0,\tau_1} - \theta^*\},$$

which shows that

$$\sup_{(\tau_1,\tau_2)\in R_{\kappa,\kappa'}} \Big| \sqrt{n(\tau_2-\tau_1)\frac{\tau_1}{\tau_2}} H'(\hat{\theta}_{n,\tau_1,\tau_2} - \hat{\theta}_{n,0,\tau_1}) - \Big\{ \sqrt{\frac{\tau_1}{n\tau_2(\tau_2-\tau_1)}}(S(\lfloor n\tau_2\rfloor) - S(\lfloor n\tau_1\rfloor)) - \sqrt{\frac{\tau_2-\tau_1}{n\tau_2\tau_1}} S(\lfloor n\tau_1\rfloor)\Big\}\Big| = O_p\Big(\frac{\log(n)^3}{n^{1/2}}\Big). \tag{B.30}$$

We now define

$$\Phi : D[0,1]^{r+s+1} \to D(R_{\kappa,\kappa'})^{r+s+1},$$
$$f \mapsto \Big[(\tau_1,\tau_2) \mapsto \Sigma^{1/2}\Big\{\sqrt{\frac{\tau_1}{\tau_2(\tau_2-\tau_1)}}(f(\tau_2) - f(\tau_1)) - \sqrt{\frac{\tau_2-\tau_1}{\tau_2\tau_1}} f(\tau_1)\Big\}\Big],$$

which is Lipschitz continuous with respect to the $\|\cdot\|_\infty$-norm. Thus we obtain on $D(R_{\kappa,\kappa'})^{r+s+1}$ that

$$\begin{aligned}&\Sigma_H^{1/2}\Big[\sqrt{\frac{\tau_1}{n\tau_2(\tau_2-\tau_1)}}(S^0(\lfloor n\tau_2\rfloor) - S^0(\lfloor n\tau_1\rfloor)) - \sqrt{\frac{\tau_2-\tau_1}{n\tau_2\tau_1}} S^0(\lfloor n\tau_1\rfloor)\Big]\\ &= \Phi(S^0(\lfloor n\cdot\rfloor) \xrightarrow{d} \Phi(B(\cdot)) = \frac{\Sigma_H^{1/2}}{\sqrt{\tau_2}}\Big\{\sqrt{\frac{\tau_1}{\tau_2-\tau_1}}\{B(\tau_2) - B(\tau_1)\} - \sqrt{\frac{\tau_2-\tau_1}{\tau_1}} B(\tau_1)\Big\}.\end{aligned} \tag{B.31}$$

From (B.30) and (B.31) and the continuous mapping theorem, we conclude that

$$\begin{aligned}&\sup_{(\tau_1,\tau_2)\in R_{\kappa,\kappa'}} \sqrt{n(\tau_2-\tau_1)\frac{\tau_1}{\tau_2}} \cdot \Sigma_H^{-1/2} \cdot H'(\hat{\theta}_{n,\tau_1,\tau_2} - \hat{\theta}_{n,0,\tau_1})\\ &\xrightarrow{d} \sup_{(\tau_1,\tau_2)\in R_{\kappa,\kappa'}} \frac{1}{\sqrt{\tau_2}}\Big\{\sqrt{\frac{\tau_1}{\tau_2-\tau_1}}\{B(\tau_2) - B(\tau_1)\} - \sqrt{\frac{\tau_2-\tau_1}{\tau_1}} B(\tau_1)\Big\}.\end{aligned}$$

By Proposition 3.2, the result follows with Slutzky's lemma.

We now consider the situation under the alternative H_1. By Proposition 3.1, we have

$$|\hat{\theta}_{n,\tau_1^*,\tau_2^*} - (\theta^* + H\Delta^*)|_1 \xrightarrow{p} 0, \tag{B.32}$$

and by Theorem 3.1 (note that the process X_i, $i = 1, ..., \lfloor n\tau_1^*\rfloor$ is based on the evolution with parameter θ^*):

$$|\hat{\theta}_{n,0,\tau_1^*} - \theta^*|_1 \xrightarrow{p} 0, \tag{B.33}$$

Furthermore, from Proposition 3.2 we have

$$\bar{\Sigma}_{n,\tau_1^*} \xrightarrow{p} \Sigma. \tag{B.34}$$

Insertion into $\hat{B}_n^{cp}$ yields

$$\begin{aligned}\hat{B}_n^{cp} &\geq \hat{B}_n^{cp}(\tau_1^*, \tau_2^*) \\ &= \sqrt{n(\tau_2^* - \tau_1^*)\frac{\tau_1^*}{\tau_2^*}}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2}\{H'\hat{\theta}_{n,\tau_1^*,\tau_2^*} - H'\hat{\theta}_{n,0,\tau_1^*}\} \\ &= \sqrt{n(\tau_2^* - \tau_1^*)\frac{\tau_1^*}{\tau_2^*}}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2}\{H'(\hat{\theta}_{n,\tau_1^*,\tau_2^*} - (\theta^* + H\Delta^*) + \theta^* - \hat{\theta}_{n,0,\tau_1^*})\} \\ &\quad + \sqrt{n(\tau_2^* - \tau_1^*)\frac{\tau_1^*}{\tau_2^*}}(H'\bar{\Sigma}_{n,\tau_1^*}H)^{-1/2} \cdot H'H\Delta^*.\end{aligned}$$

Due to (B.32) and (B.33), the first summand is of smaller order than the second. By (B.34) and since $H'H\Delta^* > 0$, the second summand converges to ∞. This proves the assertion.

B.3. Technical lemmata

From Francq and Zakoïan (2004) (the proof of Theorem 2.2, part (ii) therein), we directly obtain (ii),(iii) of the following Proposition.

PROPOSITION B.2. *(Properties of $I(\theta)$, $V(\theta)$) Let Assumption 3.1 hold. Assume that $\mu_4 := \mathbb{E}\zeta_0^4 < \infty$. Then the following statements hold true.*

(i) There exists $\iota > 0$ such that for all $\theta \in \Theta$ with $|\theta - \theta^*| < \iota$, $V(\theta)$ and $I(\theta)$ are finite.
(ii) $I(\theta^*)$ is nonsingular. It holds that $I(\theta^*) = \frac{\mu_4 - 1}{2}V(\theta^*)$.
(iii) $\mathbb{E}\nabla_\theta \ell(X_i^2, Y_i, \theta^*) = 0$.

PROOF. (PROOF OF PROPOSITION B.2) (i) By Proposition B.1, there exists $q > 0$ with $\|X_0^2\|_q < \infty$. From the bounds (B.55) (applied with $p = q$) we conclude that $V(\theta)$, $I(\theta)$ are finite as long as $|\theta - \theta^*|$ is small enough.
(ii),(iii) This was already shown in Francq and Zakoïan (2004), see the proof step (ii) of Theorem 2.2 (the missing $\frac{1}{2}$ is due to the different formulation of the likelihood).

LEMMA B.2. (NEGLIGIBILITY OF TRUNCATION) *Let Assumption 3.1 and H_0 hold. Then for $g = \nabla_\theta^l \ell$, $l = 0, 1, 2$ it holds that*

(i)

$$\sup_{r\in[0,1]}\sup_{\theta\in\Theta}\Big|\sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i^c, \theta) - \sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i, \theta)\Big| = O_p(1).$$

(ii)

$$\sup_{r\in[0,1]}\sup_{\theta\in\Theta}\frac{1}{n}\Big|\sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i^c, \theta) - \sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i, \theta)\Big| \to 0 \quad a.s.$$

PROOF. (PROOF OF LEMMA B.2) Note that for arbitrary $0 < \tilde{q} \leq \min\{q, 1\}$ and random variables $Z_i = (Z_{i1}, Z_{i2}, ...)$ with $\|Z_i\|_q \le D$ it holds that

$$\||Z_i|_{(\rho^j)_j,1}\|_{\tilde{q}} \le \Big(\sum_{j=1}^{\infty} \rho^{\tilde{q}j}\|Z_{ij}\|_{\tilde{q}}^{\tilde{q}}\Big)^{1/\tilde{q}} \le D(\frac{1}{1-\rho^{\tilde{q}}})^{1/\tilde{q}} =: \tilde{D}(\tilde{q}).$$

Let

$$W_i := \sup_{\theta\in\Theta} |g(X_i^2, Y_i^c, \theta) - g(X_i^2, Y_i, \theta)|.$$

By Lemma B.6(iii), we have with Hölder's inequality for $0 < q' \le q$ chosen such that $0 < q'(l+3) \le 1$:

$$\begin{aligned}
&\|W_i\|_{q'} \\
&= \big\| \sup_{\theta\in\Theta} |g(X_i^2, Y_i^c, \theta) - g(X_i^2, Y_i, \theta)| \big\|_{q'} \\
&\le C(1 + \||Y_i|_{(\rho^j)_j,1}\|_{q'(l+3)}^{l+1} + \||Y_i^c|_{(\rho^j)_j,1}\|_{q'(l+3)}^{l+1}) \cdot (1 + \|X_i^2\|_{q'(l+3)})\||Y_i - Y_i^c|_{(\rho^j)_j,1}\|_{q'(l+3)} \\
&\le C(1 + 2\tilde{D}(q')^{l+1})(1 + D) \cdot \Big(\sum_{j=i}^{\infty} \rho^{q'(l+3)j}\|X_j^2\|_{q'(l+3)}^{q'(l+3)}\Big)^{1/(q'(l+3))} \\
&\le C(1 + 2\tilde{D}(q')^{l+1})(1 + D)\tilde{D}(q'(l+3))\rho^i =: \tilde{C} \cdot \rho^i. \qquad \text{(B.35)}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\Big\| \sup_{r\in[0,1]} \sup_{\theta\in\Theta} \Big| \sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i^c, \theta) - \sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i, \theta)\Big|\Big\|_{q'} &\le \Big\|\sum_{i=1}^{n} W_i\Big\|_{q'} \\
&\le \Big(\sum_{i=1}^{n} \|W_i\|_{q'}^{q'}\Big)^{1/q'} \\
&\le \tilde{C}\Big(\sum_{i=1}^{n} (\rho^{q'})^i\Big)^{1/q'} < \infty,
\end{aligned}$$

giving the result.
(ii) It holds that

$$\sup_{r\in[0,1]} \sup_{\theta\in\Theta} \frac{1}{n}\Big|\sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i^c, \theta) - \sum_{i=1}^{\lfloor nr\rfloor} g(X_i^2, Y_i, \theta)\Big| \le \frac{1}{n}\sum_{i=1}^{n} W_i.$$

In the following we show that $W_i \to 0$ $(i \to \infty)$ a.s.. Then the assertion follows with a Cesaro sum argument. Let $\varepsilon > 0$ be arbitrary. Then with Markov's inequality and (B.35),

$$\sum_{i=1}^{\infty} \mathbb{P}(|W_i| > \varepsilon) \le \sum_{i=1}^{\infty} \frac{\tilde{C}^{q'}}{\varepsilon^{q'}} \rho^{q'i} < \infty,$$

showing $W_i \to 0$ with Borel-Cantelli's lemma.

Lemma B.3. (Negligibility/Replacement of parameter change) *Let Assumption 3.1 hold for $\theta^*, \theta^* + H'\Delta^* \in \Theta$. Let H_1 hold. Let $\tilde{X}_i, \tilde{\sigma}_i$, $i \le n$ denote a GARCH(r, s) process following (1.1) with constant parameters $\theta(i) \equiv \theta^* + H'\Delta^*$ for all $i \le n$. Put $\tilde{Y}_i = (\tilde{X}_j^2 : j \le i-1)$. Then it holds that*

(i)

$$\sup_{\theta\in\Theta}\Big|\sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor}\ell(X_i^2,Y_i,\theta)-\sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor}\ell(\tilde{X}_i^2,\tilde{Y}_i,\theta)\Big|=O_p(1).$$

(ii)

$$\sup_{\theta\in\Theta}\frac{1}{n}\Big|\sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor}\ell(X_i^2,Y_i,\theta)-\sum_{i=\lfloor n\tau_1^*\rfloor+1}^{\lfloor n\tau_2^*\rfloor}\ell(\tilde{X}_i^2,\tilde{Y}_i,\theta)\Big|\to 0 \quad a.s.$$

Proof. (Proof of Lemma B.3) By Proposition B.1 applied to $\bar{\theta}^* := \theta^* + H'\Delta^* \in \Theta$, $\tilde{X}_i$, $i \le n$ is a stationary process and fulfills

$$\tilde{P}_i = A_i(\bar{\theta}^*)\tilde{P}_{i-1} + a_i(\bar{\theta}^*), \qquad i \le n,$$

where $\tilde{P}_i = (\tilde{X}_i^2, ..., \tilde{X}_{i-r+1}^2, \tilde{\sigma}_i^2, ..., \tilde{\sigma}_{i-s+1}^2)'$. Furthermore, for some $q > 0$ small enough, $\||\tilde{P}_0|_2\|_q < \infty$.

We now show that $\|X_i^2\|_q$ is uniformly bounded for all i, and afterwards draw the connection to $\tilde{X}_i^2$. Under H_1 the process X_i fulfills

$$P_i = A_i(\bar{\theta}^*)P_{i-1} + a_i(\bar{\theta}^*), \qquad i \in I := \{i_0 := \lfloor n\tau_1^*\rfloor + 1, ..., \lfloor n\tau_2^*\rfloor\}, \tag{B.36}$$

where $P_i = (X_i^2, ..., X_{i-r+1}^2, \sigma_i^2, ..., \sigma_{i-s+1}^2)'$. Until $i = \lfloor n\tau_1^*\rfloor$, X_i^2 is stationary with parameter θ^*. By Proposition B.1, there exists $q > 0$ small enough such that $\||P_{i_0-1}|_2\|_q < \infty$. Furthermore, from (B.36) we obtain for $i \ge i_0$ that

$$P_i = \sum_{k=0}^{i-i_0}\big[A_i(\bar{\theta}^*)\cdot ... \cdot A_{i-k+1}(\bar{\theta}^*)\big]a_{i-k}(\bar{\theta}^*) + \big[A_i(\bar{\theta}^*)\cdot ... \cdot A_{i_0}(\bar{\theta}^*)\big]P_{i_0-1}.$$

Since $\bar{\theta}^* \in \Theta$, we conclude as in Proposition B.1 that for some (possibly small) $q > 0$ and some constant $C > 0$,

$$\begin{aligned}\||P_i|_2\|_q &\le \sum_{k=0}^{i-i_0}\big\|\,\big|A_i(\bar{\theta}^*)\cdot ... \cdot A_{i-k+1}(\bar{\theta}^*)\big|_2\big\|_q\big\|\,|a_{i-k}(\bar{\theta}^*)|_2\big\|_q + \big\|\,\big|A_i(\bar{\theta}^*)\cdot ... \cdot A_{i_0}(\bar{\theta}^*)\big|_2\big\|_q\big\||P_{i_0-1}|_2\big\|_q\\ &\le C\big\|\,|a_0(\bar{\theta}^*)|_2\big\|_q\sum_{k=0}^{i-i_0}\big(s(q,\bar{\theta}^*)^{q/m}\big)^k + \big(s(q,\bar{\theta}^*)^{q/m}\big)^{i-i_0}\big\|\,|P_{i_0-1}|_2\big\|_q,\end{aligned}$$

which shows that $\sup_{i\ge i_0}\|X_i^2\|_q \le \sup_{i\ge i_0}\|\,|P_i|_2\|_q < \infty$ is uniformly bounded for all i.

Connection to $\tilde{X}_i^2$: We conclude that for $i \in I$,

$$P_i - \tilde{P}_i = A_i(\bar{\theta}^*)\cdot ... \cdot A_{i_0}(\bar{\theta}^*)\cdot(P_{i_0-1} - \tilde{P}_{i_0-1}).$$

Since $\bar{\theta}^* \in \Theta$, we conclude as in Proposition B.1 that for some (possibly small) $q > 0$ and some constant $C > 0$,

$$\begin{aligned}\|\,|P_i - \tilde{P}_i|_2\|_q &\le \|\,|A_i(\bar{\theta}^*)\cdot ... \cdot A_{i_0}(\bar{\theta}^*)|_2\|_q \cdot \||P_{i_0-1} - \tilde{P}_{i_0-1}|_2\|_q\\ &\le C\cdot(s(q,\bar{\theta}^*)^{q/m})^{i-i_0}\||P_{i_0-1} - \tilde{P}_{i_0-1}|_2\|_q.\end{aligned}$$

Since $\||P_{i_0-1} - \tilde{P}_{i_0-1}|_2\|_q < \infty$, we conclude that

$$\|X_i^2 - \tilde{X}_i^2\|_q \le \|\,|P_i - \tilde{P}_i|_2\|_q \le C'\tilde{\rho}^{i-i_0}, \qquad i \in \{i_0, ..., \lfloor n\tau_2^*\rfloor\}, \tag{B.37}$$

where $\tilde{\rho} := s(q, \bar{\theta}^*)^{q/m}$, $C' = C\||P_{i_0-1} - \tilde{P}_{i_0-1}|_2\|_q$.

Define

$$W_i := \sup_{\theta \in \Theta} \big|\ell(X_i^2, Y_i, \theta) - \ell(\tilde{X}_i^2, \tilde{Y}_i, \theta)\big|.$$

We have shown that $\sup_{i \le n} \|X_i^2\|_q, \sup_{i \le n} \|\tilde{X}_i^2\|_q \le D < \infty$ with some constant D, thus we obtain from Lemma B.6(iii) that for $0 \le 3q' \le q < 1$,

$$\begin{aligned}
\|W_i\|_{q'} &\le C\big(1 + \||Y_i|_{(\rho^j)_j,1}\|_{3q'} + \||\tilde{Y}_i|_{(\rho^j)_j,1}\|_{3q'}\big) \cdot (1 + \|X_i^2\|_{3q'})\||Y_i - \tilde{Y}_i|_{(\rho^j)_j,1}\|_{3q'} \\
&\quad + C\|X_i^2 - \tilde{X}_i^2\|_{q'} \\
&\le C(1 + 2\tilde{D}(q')^{l+1})(1 + D) \cdot \Big(\sum_{j=1}^{\infty} \rho^{3q'j} \|X_{i-j}^2 - \tilde{X}_{i-j}^2\|_{3q'}^{3q'}\Big)^{1/(3q')} \\
&\quad + C\|X_i^2 - \tilde{X}_i^2\|_{q'},
\end{aligned} \tag{B.38}$$

where $\tilde{D}(\tilde{q}) := D(\frac{1}{1-\rho^{\tilde{q}}})^{1/\tilde{q}}$ (cf. also the proof of Lemma B.2). Note that by (B.37), for $i \ge i_0$,

$$\|X_i^2 - \tilde{X}_i^2\|_{q'} \le C' \tilde{\rho}^{i-i_0},$$

and thus

$$\begin{aligned}
\sum_{j=1}^{\infty} \rho^{3q'j} \|X_{i-j}^2 - \tilde{X}_{i-j}^2\|_{3q'}^{3q'} &\le \sum_{j=1}^{i-i_0} \rho^{3q'j} \|X_{i-j}^2 - \tilde{X}_{i-j}^2\|_{3q'}^{3q'} + (2D)^{3q'} \sum_{j=i-i_0+1}^{\infty} \rho^{3q'j} \\
&\le C' \sum_{j=1}^{i-i_0} \rho^{3q'j} \tilde{\rho}^{3q'(i-j-i_0)} + (2D)^{3q'} \sum_{j=i-i_0+1}^{\infty} \rho^{3q'j} \\
&\le C' \max\{\rho, \tilde{\rho}\}^{3q'(i-i_0)} + (2D)^{3q'} \frac{\rho^{3q'(i-i_0)}}{1 - \rho^{3q'}}.
\end{aligned}$$

Insertion of these results into (B.38) yields that there exists some constant $C'' > 0$ such that

$$\|W_i\|_{q'} \le C'' \max\{\rho, \tilde{\rho}\}^{i-i_0}. \tag{B.39}$$

Therefore, we have

$$\begin{aligned}
&\Big\| \sup_{\theta \in \Theta} \Big| \sum_{i=\lfloor n\tau_1^* \rfloor + 1}^{\lfloor n\tau_2^* \rfloor} \ell(X_i^2, Y_i, \theta) - \sum_{i=\lfloor n\tau_1^* \rfloor + 1}^{\lfloor n\tau_2^* \rfloor} \ell(\tilde{X}_i^2, \tilde{Y}_i, \theta) \Big| \Big\|_{q'} \\
&\le \Big\| \sum_{i=i_0}^{n} W_i \Big\|_{q'} \le \Big(\sum_{i=i_0}^{n} \|W_i\|_{q'}^{q'} \Big)^{1/q'} \le C'' \Big(\sum_{i=i_0}^{n} (\max\{\rho, \tilde{\rho}\}^{q'})^{i-i_0} \Big)^{1/q'} < \infty,
\end{aligned}$$

which proves the result.

(ii) The proof is similar to the proof of Lemma B.2(ii). Note that

$$\sup_{\theta \in \Theta} \frac{1}{n} \Big| \sum_{i=\lfloor n\tau_1^* \rfloor + 1}^{\lfloor n\tau_2^* \rfloor} \ell(X_i^2, Y_i, \theta) - \sum_{i=\lfloor n\tau_1^* \rfloor + 1}^{\lfloor n\tau_2^* \rfloor} \ell(\tilde{X}_i^2, \tilde{Y}_i, \theta) \Big| \le \frac{1}{n} \sum_{i=i_0}^{n} W_i.$$

By (B.39),

$$\sum_{i=i_0}^{\infty} \mathbb{P}(|W_i| > \varepsilon) \le \sum_{i=i_0}^{\infty} \frac{(C'')^{q'}}{\varepsilon^{q'}} \rho^{q'(i-i_0)} < \infty.$$

This shows that $W_{i+i_0-1} \to 0$ a.s. for $i \to \infty$. A Cesaro sum argument yields $\frac{1}{n}\sum_{i=i_0}^{n} W_i \to 0$.

Let us use the abbreviation $Z_i := (X_i^2, Y_i)$. We now state results about the dependence measure of the stationary processes $g(Z_i,\theta)$, where $g \in \{\nabla_\theta \ell, \nabla_\theta^2 \ell\}$.

LEMMA B.4. (DEPENDENCE MEASURES OF $\nabla_\theta\ell$, $\nabla_\theta^2\ell$) *Let Assumption 3.1 hold. Let $M \geq 1$. Assume that $\mathbb{E}|\zeta_0|^{2(M+a)} < \infty$ for some $a > 0$. Let $g \in \{\nabla_\theta \ell, \nabla_\theta^2 \ell\}$. Then there exists some $C > 0$, $\rho \in (0,1)$, $\iota > 0$ such that*

$$\sup_{|\theta-\theta^*|_1<\iota} \delta_M^{g(Z,\theta)}(k) \leq C\rho^k, \qquad \delta_M^{\sup_{|\theta-\theta^*|_1<\iota} g(Z,\theta)}(k) \leq C\rho^k.$$

PROOF. (PROOF OF LEMMA B.4) We only prove the second assertion, the first is nearly the same. Let $(X_i^2)^* = H(\mathcal{F}_i^*)$ and $Z_i^* := ((X_i^2)^*, Y_i^*)$. Let $\kappa = \frac{a}{3}$. Choose $p > 0$ small enough such that $\frac{(M+3\kappa)}{\kappa} p \leq q$. By Hoelder's inequality ($\frac{M}{M+3\kappa} + \frac{\kappa}{M+3\kappa} + \frac{2\kappa}{M+3\kappa} = 1$) and Lemma B.6(ii) there exists $\iota > 0$, $C > 0$, $\rho \in (0,1)$ such that

$$\begin{aligned}
&\delta_M^{\sup_{|\theta-\theta^*|_1<\iota} |g(Z,\theta)|}(i)| \\
&= \Big\| \sup_{|\theta-\theta^*|_1<\iota} |g(Z_i,\theta)| - \sup_{|\theta-\theta^*|_1<\iota} |g(Z_i^*,\theta)| \Big\|_M \\
&\leq \Big\| \sup_{|\theta-\theta^*|_1<\iota} |g(Z_i,\theta) - g(Z_i^*,\theta)| \Big\|_M \\
&= \Big\| \sup_{|\theta-\theta^*|_1<\iota} |\tilde{g}_{\theta^*}(\zeta_i, Y_i, \theta) - \tilde{g}_{\theta^*}(\zeta_i, Y_i^*, \theta)| \Big\|_M \\
&\leq C(1 + \| |Y_i|_{(\rho^j)_j,2p}^{2p} \|_{\frac{M+3\kappa}{2\kappa}} + \| |Y_i^*|_{(\rho^j)_j,2p}^{2p} \|_{\frac{M+3\kappa}{2\kappa}}) \| |Y_i - Y_i^*|_{(\rho^j)_j,p}^{p} \|_{\frac{M+3\kappa}{\kappa}} (1 + \|\zeta^2\|_{M+3\kappa}) \\
&\leq C(1 + 2\frac{D^{2p}}{1-\rho})(1 + \|\zeta^2\|_{M+3\kappa}) \cdot \sum_{j=1}^{\infty} \rho^j \|X_{i-j}^2 - (X_{i-j}^*)^2\|_{\frac{(M+3\kappa)p}{\kappa}}^p \\
&\leq \tilde{C} \cdot \sum_{j=1}^{i} \rho^j [\delta_q^{X^2}(i-j)]^p,
\end{aligned}$$

where $\tilde{C} := C(1 + 2\frac{D^{2p}}{1-\rho})(1 + \|\zeta^2\|_{M+3\kappa})$. By Proposition B.1(ii), it holds that $\delta_q^{X^2}(k) = O(c^k)$, which finishes the proof.

In the following we make use of results from Zhang and Wu (2017). Therefore we have to define $\Delta_q^Z(m) := \sum_{k=m}^{\infty} \delta_q^Z(k)$ and $\|Z\|_{q,\alpha} := \sup_{m\geq 0}(m+1)^\alpha \Delta_q^Z(m)$.

LEMMA B.5. *Let Assumption 3.1 hold. Additionally, assume that for some $a' > 0$, $\mathbb{E}|\zeta_0|^{4+a'} < \infty$. Then there exists $\iota > 0$ such that for $g = \nabla_\theta^l \ell$, $l = 1, 2$, it holds that*

$$\sup_{|\theta-\theta^*|_1<\iota} \sup_{r\in[0,1]} \Big| \frac{1}{n} \sum_{i=1}^{\lfloor nr \rfloor} \big\{ g(X_i^2, Y_i, \theta) - \mathbb{E} g(X_i^2, Y_i, \theta) \big\} \Big| = O_p\Big(\Big(\frac{\log(n)^3}{n} \Big)^{1/2} \Big).$$

PROOF. (PROOF OF LEMMA B.5) Let $\iota > 0$ (is chosen below). Let $S_j(\theta) := \sum_{i=1}^{j} \big\{ g(X_i^2, Y_i, \theta) -$

$\mathbb{E}g(X_i^2, Y_i, \theta)\big\}$, $j = 1, ..., n$. For fixed $n \in \mathbb{N}$, choose $d \in \mathbb{N}$ such that $2^{d-1} \le n \le 2^d$. For $i = 0, 1, ..., d-1$, define

$$\Phi_i(\theta) := \max_{1 \le k \le 2^{d-i}} |S_{2^i \cdot k}(\theta) - S_{2^i(k-1)}|.$$

By a dyadic expansion of $j \in \{1, ..., n\}$ we obtain

$$\max_{j=0,\dots,n} |S_j(\theta)| \le \sum_{i=0}^{d-1} \Phi_i(\theta).$$

Note that

$$\sup_{|\theta-\theta^*|_1<\iota} \sup_{r\in[0,1]} \Big| \sum_{i=1}^{\lfloor nr \rfloor} \big\{g(X_i^2, Y_i, \theta) - \mathbb{E}g(X_i^2, Y_i, \theta)\big\}\Big| \le \sum_{i=0}^{d-1} \sup_{|\theta-\theta^*|_1<\iota} \Phi_i(\theta).$$

Thus, for $Q > 0$, by stationarity,

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{|\theta-\theta^*|_1<\iota} \sup_{r\in[0,1]} \Big| \sum_{i=1}^{\lfloor nr \rfloor} \big\{g(X_i^2, Y_i, \theta) - \mathbb{E}g(X_i^2, Y_i, \theta)\big\}\Big| > Q(n\log(n)^3)^{1/2}\Big) \\
&\le \sum_{i=0}^{d-1} \mathbb{P}\Big(\sup_{|\theta-\theta^*|_1<\iota} \Phi_i(\theta) > \frac{Q(n\log(n)^3)^{1/2}}{d}\Big) \\
&\le \sum_{i=0}^{d-1} 2^{d-i} \cdot \mathbb{P}\Big(\sup_{|\theta-\theta^*|_1<\iota} |S_{2^i}(\theta)| > \frac{Q(n\log(n)^3)^{1/2}}{d}\Big). \qquad \text{(B.40)}
\end{aligned}$$

Since $\theta^* \in int(\Theta)$, there exists $\iota_1 > 0$ such that $\bar{\Theta} := \{\theta \in \Theta : |\theta - \theta^*|_1 \le \iota_1\} \subset int(\Theta)$.

Apply Lemma B.6(ii) with $p = q$ and Lemma B.4 applied to $M = 2 + \frac{a'}{4}$, $a = \frac{a'}{4}$, we obtain corresponding $C > 0$, $\rho \in (0,1)$, $0 < \iota < \iota_1$ such that the statements of the Lemmata hold true.

We now use a simple chaining argument. Let Θ_n be a discretization of $\Theta \subset \mathbb{R}^{r+s+1}$ such that for each $\theta \in \Theta$ there exists some $\theta' \in \Theta_n$ with $|\theta - \theta'|_1 \le n^{-1}$.

We conclude that for $1 \le m \le n$, it holds that

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{|\theta-\theta^*|_1<\iota} |S_m(\theta)| > \frac{Q(n\log(n)^3)^{1/2}}{d}\Big) \\
&\le \mathbb{P}\Big(\sup_{\theta\in\Theta_n, |\theta-\theta^*|_1<\iota} |S_m(\theta)| > \frac{Q(n\log(n)^3)^{1/2}}{2d}\Big) \\
&\qquad + \mathbb{P}\Big(\sup_{\theta,\theta'\in\Theta, |\theta-\theta'|_1\le n^{-1}, |\theta-\theta^*|_1<\iota, |\theta'-\theta^*|_1<\iota} |S_m(\theta) - S_m(\theta')| > \frac{Q(n\log(n)^3)^{1/2}}{2d}\Big) \\
&=: I_n + II_n. \qquad \text{(B.41)}
\end{aligned}$$

By Lemma B.4 applied to $M = 2 + \frac{a'}{4}$, $a = \frac{a'}{4}$, we have $\Delta_M^{\sup_{|\theta-\theta^*|_1<\iota} |g(Z,\theta)|}(k) = O(\rho^k)$ and $\sup_{|\theta-\theta^*|_1<\iota} \Delta_M^{g(Z,\theta)}(k) = O(\rho^k)$. Let $\alpha = \frac{1}{2}$. Then

$$W_{M,\alpha} := \| \sup_{|\theta-\theta^*|_1<\iota} |g(Z,\theta)| \|_{M,\alpha} = \sup_{m\ge 0} (m+1)^\alpha \Delta_M^{\sup_{|\theta-\theta^*|_1<\iota} |g(Z,\theta)|}(m) < \infty,$$

and

$$W_{2,\alpha} := \sup_{|\theta-\theta^*|_1<\iota} \|g(Z,\theta)\|_{2,\alpha} = \sup_{m\geq 0}(m+1)^\alpha \sup_{|\theta-\theta^*|_1<\iota} \Delta_2^{g(Z,\theta)}(m) < \infty.$$

Note that $l = 1 \wedge \log \#\Theta_n \leq (r+s+1)\log(n)$ and $Qn^{1/2}\log(n)^{3/2} \geq \sqrt{ml}W_{2,\alpha} + m^{1/M}l^{3/2}W_{M,\alpha} \gtrsim m^{1/2}\log(m)^{1/2} + m^{1/M}\log(m)^{3/2}$ for Q large enough.

By applying Theorem 6.2 of Zhang and Wu (2017) with $q = M$ to $(g(Z_i,\theta))_{\theta\in\tilde{\Theta}_n,|\theta-\theta^*|_1<\iota}$, we have with some constants $C_\alpha > 0$:

$$\begin{aligned} I_n &= \mathbb{P}\Big(\sup_{\theta\in\Theta_n,|\theta-\theta^*|_1<\iota} |S_m(\theta)| > \frac{Q(n\log(n)^3)^{1/2}}{2d}\Big) \\ &\leq \frac{C_\alpha m\cdot l^{M/2}W_{M,\alpha}^M}{(Q/2d)^M(n^{1/2}\log(n)^{3/2})^M} + C_\alpha \exp\Big(-\frac{C_\alpha(Q/2d)^2 n\log(n)^3}{mW_{2,\alpha}^2}\Big) \\ &\leq O(m\cdot n^{-\frac{M}{2}} + n^{-2}), \end{aligned} \tag{B.42}$$

for Q large enough, since $d \leq \log_2(n)+1$ and $m \leq n$.

Since $g(Z_i,\theta) = \tilde{g}_{\theta^*}(\zeta_i,Y_i,\theta)$ and $g(Z_i,\theta') = \tilde{g}_{\theta^*}(\zeta_i,Y_i,\theta')$, we have with Lemma B.6(ii):

$$\begin{aligned} &\sup_{\theta,\theta'\in\tilde{\Theta},|\theta-\theta'|_1\leq n^{-1},|\theta-\theta^*|_1<\iota,|\theta'-\theta^*|_1<\iota} |g(Z_i,\theta) - g(Z_i,\theta')| \\ &\leq C(1+|Y_i|^p_{(\rho^j)_j,p})(1+\zeta_i^2)n^{-1}. \end{aligned}$$

Thus

$$\begin{aligned} &\Big\|\sup_{\theta,\theta'\in\Theta,|\theta-\theta'|_1\leq n^{-1},|\theta-\theta^*|_1<\iota,|\theta'-\theta^*|_1<\iota}\Big|\sum_{i=1}^m\big\{\mathbb{E}_0 g(Z_i,\theta) - \mathbb{E}_0 g(Z_i,\theta')\big\}\Big|\Big\|_1 \\ &\leq 2C(1+\||Y_0|^p_{(\rho^j)_j,p}\|_1)(1+\mathbb{E}\zeta_0^2)\frac{m}{n} \\ &\leq 2C(1+\frac{D^p}{1-\rho})(1+\mathbb{E}\zeta_0^2)\frac{m}{n} = O(\frac{m}{n}). \end{aligned}$$

With Markov's inequality, we therefore obtain

$$II_n \leq \frac{2\tilde{C}m}{(Q/2d)n^{3/2}\log(n)^{3/2}}. \tag{B.43}$$

Inserting (B.42) and (B.43) into (B.41) and then into (B.40), we obtain with some constant $\tilde{C} > 0$:

$$\begin{aligned} &\mathbb{P}\Big(\sup_{|\theta-\theta^*|_1<\iota}\sup_{r\in[0,1]}\Big|\sum_{i=1}^{\lfloor nr\rfloor}\big\{g(X_i^2,Y_i,\theta) - \mathbb{E}g(X_i^2,Y_i,\theta)\big\}\Big| > Q(n\log(n)^3)^{1/2}\Big) \\ &\leq \tilde{C}\sum_{i=0}^{d-1} 2^{d-i}\cdot\Big(2^i\cdot n^{-M/2} + n^{-2} + 2^i\cdot n^{-3/2}\log(n)^{1/2}\Big) \\ &\leq \tilde{C}dn\cdot\Big(n^{-M/2} + n^{-2} + n^{-3/2}\log(n)^{1/2}\Big) \to 0, \end{aligned}$$

showing the assertion.

B.4. Analytical properties of the likelihood

For the following results, we derive some analytical properties of the likelihood we use. This allows us to separate analytical and stochastic treatment. For $p > 0$, some sequence $(y_j)_{j\in\mathbb{N}}$ of real numbers and some sequence $(\chi_j)_{j\in\mathbb{N}}$ of nonnegative real numbers, define the weighted seminorm

$$|y|_{\chi,p} := \Big(\sum_{j=1}^{\infty} \chi_j |y_j|^p\Big)^{1/p}.$$

Later, we will plug in $x = X_i$ and $y = Y_i$ into $\ell(x, y, \theta)$ and its derivatives. To make use of all connections between x,y, define $R_\zeta(y, \theta) := \zeta^2\sigma^2(y, \theta)$, and

$$\tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) := \ell(R_\zeta(y, \tilde{\theta}), y, \theta).$$

In the following Lemma B.6(ii), we collect some analytical properties of $\tilde{\ell}_{\tilde{\theta}}$ to calculate functional dependence measures of $\ell(X_i^2, Y_i, \theta)$. The bounds in (iii) will be used to show that the truncated likelihood $\ell(X_i^2, Y_i^c, \theta)$ is near to $\ell(X_i^2, Y_i, \theta)$; for this argument we cannot use the connection between X_i^2 and Y_i.

Lemma B.6. *$\theta \mapsto \sigma^2(y, \theta)$ and $\theta \mapsto \ell(x, y, \theta)$ are three times continuously differentiable. Let $\bar{\Theta} \subset int(\Theta)$ be a compact subset. Then for any $p > 0$, there exists $\iota > 0$ and $C > 0$, $\rho \in (0, 1)$ such that (component-wise),*

(i) for $l = 0, 1, 2, 3$:

$$\sup_{\theta\in\bar{\Theta}} \frac{|\nabla_\theta^l(\sigma^2(y,\theta))|}{\sigma^2(y,\theta)} \le C(1+|y|^p_{(\rho^j)_j,p}), \qquad \sup_{\theta,\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota} \frac{\sigma^2(y,\tilde{\theta})}{\sigma^2(y,\theta)} \le C(1+|y|^p_{(\rho^j)_j,p}).$$

(ii) for $l = 0, 1, 2$,

$$\sup_{\theta,\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota} |\nabla_\theta^l \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) - \nabla_\theta^l \tilde{\ell}_{\tilde{\theta}}(\zeta, y', \theta)|$$
$$\le C(1 + |y|^{2p}_{(\rho^j)_j,2p} + |y'|^{2p}_{(\rho^j)_j,2p})|y - y'|^p_{(\rho^j)_j,p}(1 + \zeta^2).$$

and

$$\sup_{\theta,\theta',\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota,|\theta'-\tilde{\theta}|_1<\iota} \frac{|\nabla_\theta^l \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) - \nabla_\theta^l \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta')|}{|\theta-\theta'|_1} \le C(1 + |y|^p_{(\rho^j)_j,p})(1 + \zeta^2).$$

(iii) for $l = 0, 1, 2$,

$$\sup_{\theta\in\Theta} |\nabla_\theta^l \ell(x, y, \theta) - \nabla_\theta \ell(x, y', \theta)| \le C(1+|y|^{l+1}_{(\rho^j)_j,1} + |y'|^{l+1}_{(\rho^j)_j,1})(1+|x|)\cdot|y-y'|_{(\rho^j)_j,1},$$

and

$$|\ell(x, y, \theta) - \ell(x', y, \theta)| \le C|x - x'|.$$

Proof. (Proof of Lemma B.6) (i) From Proposition (B.1)(iii) we obtain that the following explicit representation holds, where $F(y, \theta) := (\alpha_0 + \sum_{j=1}^r \alpha_j y_j, 0, ..., 0)'$:

$$\sigma^2(y, \theta) = \sum_{k=0}^{\infty} \big(B(\theta)^k F(y_{k\to}, \theta)\big)_1, \tag{B.44}$$

where $y_{k\to} = (y_{k+1}, y_{k+2}, ...)$. We conclude that

$$\begin{aligned}
\sigma^2(y,\theta) &= \alpha_0 \sum_{k=0}^{\infty} (B(\theta))_{11} + \sum_{j=1}^{r} \alpha_j \sum_{k=0}^{\infty} (B(\theta)^k)_{11} y_{k+j} \\
&\overset{k'=k+j}{=} \alpha_0 \sum_{k=0}^{\infty} (B(\theta))_{11} + \sum_{k'=1}^{\infty} \Big(\sum_{j=1}^{r} \alpha_j (B(\theta)^{k'-j})_{11} \mathbb{1}_{k' \geq j} \Big) y_{k'} \\
&=: c_0(\theta) + \sum_{k'=1}^{\infty} c_{k'}(\theta) y_j. \qquad \text{(B.45)}
\end{aligned}$$

From Proposition (B.1)(iii) we obtain that $c_j(\theta) \geq 0$ satisfies

$$\sup_{\theta\in\Theta} |c_k(\theta)| \leq C \cdot \rho^k \qquad \text{(B.46)}$$

with some $\rho \in (0,1)$, $C > 0$ and $c_0(\theta) \geq \sigma_{min}^2 > 0$ (due to $\alpha_0 \geq \alpha_{min} > 0$). Furthermore we conclude that $\sigma^2(y,\theta)$ is three times continuously differentiable w.r.t. θ with

$$\nabla_\theta^k (\sigma^2(y,\theta)) = \nabla_\theta^k c_0(\theta) + \sum_{k=1}^{\infty} \nabla_\theta^k c_k(\theta) \cdot y_k, \qquad k \in \{0,1,2,3\}, \qquad \text{(B.47)}$$

where $(\nabla_\theta^k c_k(\theta))_k$ is still geometrically decaying with $\sup_{\theta\in\Theta} |\nabla_\theta^l c_l(\theta)|_\infty \leq C \cdot \rho^k$, say (enlarge $C > 0, \rho \in (0,1)$ if necessary).

In the following we make use of some arguments that were already used in Francq and Zakoïan (2004). Note that for $j = 0, ..., r$, we have $\partial_{\alpha_j} F(y,\theta) \leq \frac{1}{\alpha_j} F(y,\theta)$ and thus

$$\partial_{\alpha_j} c_k(\theta) \leq \frac{1}{\alpha_j} c_k(\theta). \qquad \text{(B.48)}$$

For $j = 1, ..., s$, we have ('$\leq$' is meant component-wise)

$$\partial_{\beta_j} (B(\theta)^k) = \sum_{i=1}^{k} B(\theta)^{i-1} (\partial_{\beta_j} B(\theta)) B(\theta)^{k-i} \leq \frac{1}{\beta_j} k B(\theta)^k.$$

since $\partial_{\beta_j} B(\theta) \leq \frac{1}{\beta_j} B(\theta)$. We therefore obtain

$$\partial_{\beta_j} c_k(\theta) \leq \frac{1}{\beta_j} k \cdot c_k(\theta). \qquad \text{(B.49)}$$

From (B.48) and (B.49) we obtain the inequalities

$$\partial_{\theta_j} c_k(\theta) \leq \frac{k+1}{\theta_j} c_k(\theta).$$

Similar argumentations lead to the bounds for higher order derivatives (cf. also Francq and Zakoïan (2004)):

$$\partial_{\theta_{j_1}} \partial_{\theta_{j_2}} c_k(\theta) \leq \frac{(k+1)^2}{\theta_{j_1}\theta_{j_2}} c_k(\theta), \qquad \partial_{\theta_{j_1}} \partial_{\theta_{j_2}} \partial_{\theta_{j_3}} c_k(\theta) \leq \frac{(k+1)^3}{\theta_{j_1}\theta_{j_2}\theta_{j_3}} c_k(\theta).$$

If $\bar\Theta \subset int(\tilde\Theta)$ is some compact subspace, we therefore obtain with $C_1 := \max\{\frac{1}{\theta_j} : j =$

$1, ..., r + s + 1, \theta \in \bar{\Theta}\}$ for arbitrary small $p > 0$:

$$\begin{aligned}
\frac{\partial_{\theta_j}(\sigma^2(y,\theta))}{\sigma^2(y,\theta)} &\le C_1 \frac{\sum_{k=0}^{\infty}(k+1)c_k(\theta)}{\sum_{k=0}^{\infty} c_k(\theta)} \\
&\le \frac{C_1 c_0(\theta)}{\sigma^2_{min}} + C_1 \sum_{k=1}^{\infty}(k+1)\frac{c_k(\theta)y_k}{c_0(\theta)+c_k(\theta)y_k} \\
&\le \frac{C_1 c_0(\theta)}{\sigma^2_{min}} + \sum_{k=1}^{\infty}(k+1)\Big(\frac{c_k(\theta)}{c_0(\theta)}\Big)^p y_k^p,
\end{aligned}$$

where we have used $\frac{x}{1+x} \le x^s$ in the last inequality. Since $c_k(\theta)^s \le C^s(\rho^s)^k$, we can find $\tilde{C} > 0, \tilde{\rho} \in (0,1)$ such that

$$\sup_{\theta\in\bar{\Theta}} \frac{|\partial_{\theta_j}(\sigma^2(y,\theta))|}{\sigma^2(y,\theta)} \le \tilde{C}(1 + |y|^p_{(\tilde{\rho}^j)_j,p}),$$

and similarly for the higher order derivatives (component-wise):

$$\sup_{\theta\in\bar{\Theta}} \frac{|\nabla^l_\theta(\sigma^2(y,\theta))|}{\sigma^2(y,\theta)} \le \tilde{C}(1 + |y|^p_{(\tilde{\rho}^j)_j,p}), \quad l = 1,2,3.$$

For $\theta, \tilde{\theta} \in \bar{\Theta}$ and arbitrary small $p > 0$, choose $\delta > 0$ such that $\bar{\rho} := (1+\delta)\rho^p < 1$. Then choose $\iota > 0$ such that $|\theta - \tilde{\theta}|_1 < \iota$ implies (component-wise) $B(\tilde{\theta}) \le (1+\delta)B(\theta)$. For $|\theta - \tilde{\theta}| < \iota$, it then holds that $c_k(\tilde{\theta}) \le (1+\delta)^k c_k(\theta)$. We conclude that

$$\begin{aligned}
\frac{\sigma^2(y,\tilde{\theta})}{\sigma^2(y,\theta)} &\le \frac{c_0(\tilde{\theta})}{\sigma^2_{min}} + \sum_{k=1}^{\infty}\frac{c_k(\tilde{\theta})y_k}{c_0(\theta)+c_k(\theta)y_k} \\
&\le \frac{c_0(\tilde{\theta})}{\sigma^2_{min}} + \sum_{k=1}^{\infty}\frac{c_k(\tilde{\theta})}{c_k(\theta)}\cdot\Big(\frac{c_k(\theta)}{c_0(\theta)}\Big)^p y_k^p \\
&\le \frac{c_0(\tilde{\theta})}{\sigma^2_{min}} + \frac{C^p}{\sigma^{2p}_{min}}\sum_{k=1}^{\infty}((1+\delta)\rho^p)^k y_k^p.
\end{aligned}$$

We conclude that there exists $\bar{C} > 0$, $\bar{\rho} \in (0,1)$ such that

$$\sup_{\theta,\tilde{\theta}\in\bar{\Theta},|\theta-\tilde{\theta}|_1<\iota} \frac{\sigma^2(y,\tilde{\theta})}{\sigma^2(y,\theta)} \le \bar{C}(1 + |y|^p_{(\bar{\rho}^j)_j,p}).$$

(ii) From the differentiability of $\theta \mapsto \sigma^2(y,\theta)$ we obtain that $\theta \mapsto \ell(x,y,\theta)$ is three times continuously differentiable and

$$\ell(x,y,\theta) = \frac{1}{2}\Big(\frac{x}{\sigma^2(y,\theta)} + \log(\sigma^2(y,\theta))\Big), \tag{B.50}$$

$$\nabla_\theta \ell(x,y,\theta) = \frac{\nabla_\theta(\sigma^2(y,\theta))}{2\sigma^2(y,\theta)}\Big(1 - \frac{x}{\sigma^2(y,\theta)}\Big), \tag{B.51}$$

$$\begin{aligned}
\nabla^2_\theta \ell(x,y,\theta) &= \Big[-\frac{\nabla_\theta(\sigma^2(y,\theta))\nabla_\theta(\sigma^2(y,\theta))'}{2(\sigma^2(y,\theta))^2} + \frac{\nabla^2_\theta(\sigma^2(y,\theta))}{2\sigma^2(y,\theta)}\Big]\Big(1 - \frac{x}{\sigma^2(y,\theta)}\Big) \\
&\quad + \frac{\nabla_\theta(\sigma^2(y,\theta))\nabla_\theta(\sigma^2(y,\theta))'}{2(\sigma^2(y,\theta))^2}\cdot\frac{x}{\sigma^2(y,\theta)}.
\end{aligned} \tag{B.52}$$

For the corresponding quantity $\tilde{\ell}_{\tilde{\theta}}$ we obtain

$$\nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) = \frac{\nabla_\theta(\sigma^2(y,\theta))}{2\sigma^2(y,\theta)}\Big(1 - \frac{\sigma^2(y,\tilde{\theta})}{\sigma^2(y,\theta)}\zeta^2\Big).$$

By (i), we obtain that for $p > 0$, there exist constants $\iota > 0$, $C_2 > 0$, $\rho_2 \in (0,1)$ such that (component-wise):

$$\sup_{|\theta-\tilde{\theta}|_1<\iota} |\nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta)| \le C_2(1 + |y|^{p/2}_{(\rho_2^j)_j, p/2})(1 + (1 + |y|^{p/2}_{(\rho_2^j)_j, p/2})\zeta^2). \tag{B.53}$$

By using

$$\begin{aligned} |y|^{p/2}_{(\rho_2^j)_j, p/2} &\le \sum_{j=1}^\infty \rho_2^{j/2} \cdot \rho_2^{j/2} y_j^p \le (\sum_{j=1}^\infty \rho_2^j)^{1/2} (\sum_{j=1}^\infty \rho_2^j y_j^p)^{1/2} \\ &= (1-\rho_2)^{-1/2} |y|^{p/2}_{(\rho_2^j)_j, p}, \end{aligned} \tag{B.54}$$

we can obtain the more compact form

$$\sup_{|\theta-\tilde{\theta}|_1<\iota} |\nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta)| \le C_3(1 + |y|^{p}_{(\rho_2^j)_j, p})(1 + \zeta^2).$$

with some new constant $C_3 > 0$. Due to the similar structure, we can use similar techniques to obtain (component-wise):

$$\sup_{|\theta-\tilde{\theta}|_1<\iota} |\nabla^l_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta)| \le C_3(1 + |y|^{p}_{(\rho_2^j)_j, p})(1 + \zeta^2) =: M_p(y,\zeta), \qquad l = 1,2,3. \tag{B.55}$$

From (B.47) we deduce that (component-wise) for $l = 0,1,2$ with some constant $C_4 > 0$, uniformly in $\theta, \theta' \in \bar{\Theta}$:

$$|\nabla^l_\theta(\sigma^2(y,\theta)) - \nabla^l_\theta(\sigma^2(y',\theta))| \le C_4 |y - y'|_{(\rho^j)_j, 1}, \tag{B.56}$$

By using $|\frac{1}{\sigma^2(y,\theta)} - \frac{1}{\sigma^2(y,\theta)}| \le \frac{1}{\sigma^4_{min}}|\sigma^2(y,\theta) - \sigma^2(y',\theta)|$ and the very rough bounds $\sigma^2(y,\theta) \ge \sigma^2_{min}$, (B.56) and (B.47), we obtain (component-wise) with some constant $C_5 > 0$:

$$\sup_{\theta\in\bar{\Theta}} |\nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) - \nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y', \theta)| \le C_5(1 + |y|_{(\rho^j)_j,1} + |y'|_{(\rho^j)_j,1})^2 |y - y'|_{(\rho^j)_j,1}(1 + \zeta^2)$$

Similar results can be obtained for higher derivatives (component-wise), $l = 1,2$:

$$\begin{aligned} &\sup_{\theta\in\bar{\Theta}} |\nabla^l_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y, \theta) - \nabla_\theta \tilde{\ell}_{\tilde{\theta}}(\zeta, y', \theta)| \\ &\le C_5(1 + |y|_{(\rho^j)_j,1} + |y'|_{(\rho^j)_j,1})^{1+l} |y - y'|_{(\rho^j)_j,1}(1 + \zeta^2) =: N_l(y, y', \zeta). \end{aligned} \tag{B.57}$$

Using (B.55) and (B.57), we have for $l = 1,2$ and arbitrary small $p' > 0$ (use $\min\{1, x\} \le$

$x^{p'}$):

$$\begin{aligned}
&\sup_{\theta\in\bar\Theta}|\nabla_\theta^l\tilde\ell_{\tilde\theta}(\zeta,y,\theta)-\nabla_\theta\tilde\ell_{\tilde\theta}(\zeta,y',\theta)|\\
&\le \min\{M_p(y,\zeta)+M_p(y',\zeta),N_l(y,y',\zeta)\}\\
&=\{M_p(y,\zeta)+M_p(y',\zeta)\}\min\{1,\frac{N_l(y,y',\zeta)}{M_p(y,\zeta)+M_p(y',\zeta)}\}\\
&\le \{M_p(y,\zeta)+M_p(y',\zeta)\}\Big(\frac{N_l(y,\zeta)}{M_p(y,\zeta)+M_p(y',\zeta)}\Big)^{p'}\\
&=\{M_p(y,\zeta)+M_p(y',\zeta)\}^{1-p'}N_l(y,\zeta)^{p'}.
\end{aligned}$$

Choosing $p'\in(0,\min\{1,p'(1+l)\}$, we obtain

$$\begin{aligned}
\{M_p(y,\zeta)+M_p(y',\zeta)\}^{1-p'} &\le C_3^{1-p'}(1+|y|^p_{(\rho_2^j)_j,p}+|y'|^p_{(\rho_2^j)_j,p})(1+\zeta^2)^{1-p'},\\
N_l(y,y',\zeta)^{p'} &\le C_5^{p'}(1+|y|^p_{(\rho^{pj}),p}+|y'|^p_{(\rho^{pj}),p})|y-y'|^p_{(\rho^{pj})_j,p}(1+\zeta^2)^{p'}.
\end{aligned}$$

With (B.54), $\rho_3:=\max\{\rho_2,\rho^p\}$ and some constant $C_6>0$

$$\sup_{\theta,\tilde\theta\in\bar\Theta,|\theta-\tilde\theta|_1<\iota}|\nabla_\theta^l\tilde\ell_{\tilde\theta}(\zeta,y,\theta)-\nabla_\theta^l\tilde\ell_{\tilde\theta}(\zeta,y',\theta)|\le C_6(1+|y|^{2p}_{(\rho_3^j)_j,2p}+|y'|^{2p}_{(\rho_3^j)_j,2p})|y-y'|^p_{(\rho_3^j)_j,p}(1+\zeta^2).$$

By using (B.55) and the mean value theorem, we obtain for $l=1,2$:

$$\begin{aligned}
&\sup_{\theta,\theta',\tilde\theta\in\bar\Theta,|\theta-\tilde\theta|_1<\iota,|\theta'-\tilde\theta|_1<\iota}\frac{|\nabla_\theta^l\tilde\ell_{\tilde\theta}(\zeta,y,\theta)-\nabla_\theta^l\tilde\ell_{\tilde\theta}(\zeta,y,\theta')|}{|\theta-\theta'|_1}\\
&\le \sup_{|\bar\theta-\tilde\theta|_1<\iota}|\nabla_\theta^{l+1}\tilde\ell_{\tilde\theta}(\zeta,y,\bar\theta)|_\infty\le M_p(y,\zeta),
\end{aligned}$$

giving the result.

(iii) Using the representations (B.51), (B.52) and the inequalities (B.56), (B.47) and $\sigma^2(y,\theta)\ge\sigma^2_{min}$, the first inequality is an immediate consequence. The second inequality follows since

$$|\ell(x,y,\theta)-\ell(x',y,\theta)|\le\frac{1}{2\sigma^2(y,\theta)}|x-x'|.$$

References

Berkes, I., L. Horváth, and P. Kokoszka (2003). GARCH processes: structure and estimation. Bernoulli 9(2), 201–227.

Chen, B. and Y. Hong (2016). Detecting for smooth structural changes in garch models. Econometric Theory 32(3), 740–791.

Francq, C. and J.-M. Zakoïan (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. Bernoulli 10(4), 605–637.

Wu, W. B. and W. Min (2005). On linear processes with dependent innovations. Stochastic Processes and Their Applications 115(6), 939 – 958.

Wu, W. B. and X. Shao (2004). Limit theorems for iterated random functions. Journal Applied Probabability 41(2), 425–436.

Wu, W. B. and Z. Zhou (2011). Gaussian approximations for non-stationary multiple time series. Statistica Sinica 21(3), 1397–1413.

Zhang, D. and W. B. Wu (2017). Gaussian Approximation for High Dimensional Time Series. Annals of Statistics 45(5), 1895–1919.