0.1 Theoretical properties of the test

To start with, we introduce the auxiliary statistic

$$\widehat{\Phi}_{n,T} = \max_{1 \le i < j \le n} \widehat{\Phi}_{ij,T}, \tag{0.1}$$

where

$$\widehat{\Phi}_{ij,T} = \max_{(u,h) \in \mathcal{G}_T} \Big\{ \Big| \frac{\widehat{\phi}_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \Big| - \lambda(h) \Big\}$$

and $\widehat{\phi}_{ij,T}(u,h) = \sum_{t=1}^T w_{t,T}(u,h) \big\{ (\varepsilon_{it} - \bar{\varepsilon}_i) + \boldsymbol{\beta}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - (\varepsilon_{jt} - \bar{\varepsilon}_j) - \boldsymbol{\beta}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \big\}$ with $\bar{\varepsilon}_i = \bar{\varepsilon}_{i,T} = T^{-1} \sum_{t=1}^T \varepsilon_{it}$ and $\bar{\mathbf{X}}_i = \bar{\mathbf{X}}_{i,T} = T^{-1} \sum_{t=1}^T \mathbf{X}_{it}$ respectively. Our first theoretical result characterizes the asymptotic behaviour of the statistic $\widehat{\Phi}_{n,T}$.

Theorem 0.1. *Suppose that the error processes $\mathcal{E}_i = \{\varepsilon_{it} : 1 \le t \le T\}$ are independent across i and satisfy* **??**–**??** *for each i. Moreover, let* **??**–**??** *be fulfilled and assume that $\widehat{\sigma}_i^2 = \sigma_i^2 + o_p(\rho_T)$ with $\rho_T = o(1/\log T)$ for each i. Then*

$$\mathbb{P}\big(\widehat{\Phi}_{n,T} \le q_{n,T}(\alpha) \big| \{\mathbf{X}_{it} : 1 \le t \le T, 1 \le i \le n\} \big) = (1 - \alpha) + o(1) \ \ a.s.$$

1 Proof of the Theorem 0.1

In this section, we prove the theoretical results from Section **??**. We use the following notation: The symbol C denotes a universal real constant which may take a different value on each occurrence. For $a, b \in \mathbb{R}$, we write $a_+ = \max\{0, a\}$ and $a \vee b = \max\{a, b\}$. For any set A, the symbol $|A|$ denotes the cardinality of A. The notation $X \stackrel{\mathcal{D}}{=} Y$ means that the two random variables X and Y have the same distribution. Finally, $f_0(\cdot)$ and $F_0(\cdot)$ denote the density and distribution function of the standard normal distribution, respectively.

Auxiliary results using strong approximation theory

The main purpose of this section is to prove that there is a version of the multiscale statistic $\widehat{\Phi}_{n,T}$ defined in (0.1) which is close to a Gaussian statistic whose distribution is known. More specifically, we prove the following result.

Proposition A.1. *Under the conditions of Theorem 0.1, there exist statistics $\widetilde{\Phi}_{n,T}$ for $T = 1, 2, \ldots$ with the following two properties: (i) $\widetilde{\Phi}_{n,T}$ has the same distribution as $\widehat{\Phi}_{n,T}$ for any T, and (ii)*

$$\big| \widetilde{\Phi}_{n,T} - \Phi_{n,T} \big| = o_p\Big(\frac{T^{1/q}}{\sqrt{T h_{\min}}} + \rho_T \sqrt{\log T} + \rho_T \max_{1 \le i \le n} \Big| \sum_{s=1}^T \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \\ + \frac{T^{-1/2}}{\sqrt{T h_{min}}} \max_{1 \le i \le n} \max_{1 \le t \le T} \Big| \sum_{s=1}^t (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| \Big),$$

where $\Phi_{n,T}$ is a Gaussian statistic as defined in (**??**).

Proof of Proposition A.1. For the proof, we draw on strong approximation theory for each stationary process $\mathcal{E}_i = \{\varepsilon_{it} : 1 \le t \le T\}$ that fulfill the conditions **??**–**??**. By Theorem 2.1 and Corollary 2.1 in Berkes et al. (2014), the following strong approximation result holds true: On a richer probability space, there exist a standard Brownian motion $\mathbb{B}$ and a sequence $\{\widetilde{\varepsilon}_t : t \in \mathbb{N}\}$ such that $[\widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_T] \stackrel{\mathcal{D}}{=} [\varepsilon_1, \dots, \varepsilon_T]$ for each T and

$$\max_{1 \le t \le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_s - \sigma \mathbb{B}(t) \Big| = o\big(T^{1/q}\big) \quad \text{a.s.,} \tag{1.1}$$

where $\sigma^2 = \sum_{k \in \mathbb{Z}} \mathrm{Cov}(\varepsilon_0, \varepsilon_k)$ denotes the long-run error variance.
To apply this result, we define

$$\widetilde{\Phi}_{n,T} = \max_{1 \le i < j \le n} \widetilde{\Phi}_{ij,T}, \tag{1.2}$$

where

$$\widetilde{\Phi}_{ij,T} = \max_{(u,h) \in \mathcal{G}_T} \Big\{ \Big| \frac{\widetilde{\phi}_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \Big| - \lambda(h) \Big\},$$

where $\widetilde{\phi}_{ij,T}(u,h) = \sum_{t=1}^T w_{t,T}(u,h) \big\{ (\widetilde{\varepsilon}_{it} - \bar{\widetilde{\varepsilon}}_i) + \boldsymbol{\beta}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - (\widetilde{\varepsilon}_{jt} - \bar{\widetilde{\varepsilon}}_j) - \boldsymbol{\beta}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \big\}$ and $\widetilde{\sigma}_i^2$ are the same estimators as $\widehat{\sigma}_i^2$ with $Y_{it} = (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top \mathbf{X}_{it} + m_i(t/T) + \big(\alpha_i - \widehat{\alpha}_i\big) + \varepsilon_{it}$ replaced by $\widetilde{Y}_{t,T} = (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top \mathbf{X}_{it} + m_i(t/T) + \big(\alpha_i - \widehat{\alpha}_i\big) + \widetilde{\varepsilon}_{it}$ for $1 \le t \le T$. In addition, we let

$$\Phi_{n,T} = \max_{1 \le i < j \le n} \Phi_{ij,T} = \max_{1 \le i < j \le n} \max_{(u,h) \in \mathcal{G}_T} \Big\{ \Big| \frac{\phi_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \Big| - \lambda(h) \Big\}$$

$$\Phi^{\diamond}_{n,T} = \max_{1 \le i < j \le n} \Phi^{\diamond}_{ij,T} = \max_{1 \le i < j \le n} \max_{(u,h) \in \mathcal{G}_T} \Big\{ \Big| \frac{\phi_{ij,T}(u,h)}{\{\widetilde{\sigma}_i^2 + \widetilde{\sigma}_j^2\}^{1/2}} \Big| - \lambda(h) \Big\}$$

with

$$\phi_{ij,T}(u,h) = \sum\nolimits_{t=1}^{T} w_{t,T}(u,h) \big\{ \widehat{\sigma}_i (Z_{it} - \bar{Z}_i) + \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - \widehat{\sigma}_j (Z_{jt} - \bar{Z}_j) - \widehat{\boldsymbol{\beta}}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \big\}$$

and $Z_{it} = \mathbb{B}_i(t) - \mathbb{B}_i(t-1)$. With this notation, we can write

$$\big| \widetilde{\Phi}_{n,T} - \Phi_{n,T} \big| \le \big| \widetilde{\Phi}_{n,T} - \Phi^{\diamond}_{n,T} \big| + \big| \Phi^{\diamond}_{n,T} - \Phi_{n,T} \big|. \tag{1.3}$$

First consider $|\widetilde{\Phi}_{n,T} - \Phi^{\diamond}_{n,T}|$. Straightforward calculations yield that

$$\big| \widetilde{\Phi}_{n,T} - \Phi^{\diamond}_{n,T} \big| \le \{\widetilde{\sigma}_i^2 + \widetilde{\sigma}_j^2\}^{-1/2} \max_{1 \le i < j \le n} \max_{(u,h) \in \mathcal{G}_T} \big| \widetilde{\phi}_{ij,T}(u,h) - \phi_{ij,T}(u,h) \big|. \tag{1.4}$$

Using summation by parts, ($\sum_{i=1}^n a_i b_i = \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}) + A_n b_n$ with $A_j = \sum_{j=1}^i a_j$) we further obtain that

$$\begin{aligned}
&\big| \widetilde{\phi}_{ij,T}(u,h) - \phi_{ij,T}(u,h) \big| = \\
&\quad = \Big| \sum_{t=1}^{T} w_{t,T}(u,h) \big\{ (\widetilde{\varepsilon}_{it} - \bar{\widetilde{\varepsilon}}_i) + \boldsymbol{\beta}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - (\widetilde{\varepsilon}_{jt} - \bar{\widetilde{\varepsilon}}_j) - \boldsymbol{\beta}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) - \\
&\qquad - \widehat{\sigma}_i (Z_{it} - \bar{Z}_i) - \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - \widehat{\sigma}_j (Z_{jt} - \bar{Z}_j) + \widehat{\boldsymbol{\beta}}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \big\} \Big| = \\
&\quad = \Big| \sum_{t=1}^{T-1} A_{ij,t} \big(w_{t,T}(u,h) - w_{t+1,T}(u,h) \big) + A_{ij,T} w_{t,T}(u,h) \Big|,
\end{aligned}$$

where

$$A_{ij,t} = \sum_{s=1}^{t} \big\{ (\widetilde{\varepsilon}_{is} - \bar{\widetilde{\varepsilon}}_i) + \boldsymbol{\beta}_i^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) - (\widetilde{\varepsilon}_{js} - \bar{\widetilde{\varepsilon}}_j) - \boldsymbol{\beta}_j^\top (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) -$$

$$- \widehat{\sigma}_i (Z_{it} - \bar{Z}_i) - \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - \widehat{\sigma}_j (Z_{jt} - \bar{Z}_j) + \widehat{\boldsymbol{\beta}}_j^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \big\}.$$

Note that by construction $A_{ij,T} = 0$ for all pairs (i, j). Denoting

$$W_T(u,h) = \sum_{t=1}^{T-1} |w_{t+1,T}(u,h) - w_{t,T}(u,h)|,$$

we have

$$\big|\widetilde{\phi}_{ij,T}(u,h) - \phi_{ij,T}(u,h)\big| = \Big| \sum_{t=1}^{T-1} A_{ij,t} \big(w_{t,T}(u,h) - w_{t+1,T}(u,h)\big) \Big| \le W_T(u,h) \max_{1\le t\le T} |A_{ij,t}|. \tag{1.5}$$

Now consider $\max_{1\le t\le T} |A_{ij,t}|$:

$$\begin{aligned}
\max_{1\le t\le T} |A_{ij,t}| \le & \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\widetilde{\varepsilon}_{is} - \bar{\widetilde{\varepsilon}}_i) - \sigma_i \sum_{s=1}^{t} \big\{ Z_{is} - \bar{Z}_i \big\} \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\widetilde{\varepsilon}_{js} - \bar{\widetilde{\varepsilon}}_j) - \sigma_j \sum_{s=1}^{t} \big\{ Z_{js} - \bar{Z}_j \big\} \Big| + \\
& + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_j - \widehat{\boldsymbol{\beta}}_j)^\top (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big| \le \\
\le & \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{is} - \sigma_i \sum_{s=1}^{t} Z_{is} \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{js} - \sigma_j \sum_{s=1}^{t} Z_{js} \Big| + \\
& + \max_{1\le t\le T} \Big| t(\bar{\widetilde{\varepsilon}}_i - \sigma_i \bar{Z}_i) \Big| + \max_{1\le t\le T} \Big| t(\bar{\widetilde{\varepsilon}}_j - \sigma_j \bar{Z}_j) \Big| + \\
& + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_j - \widehat{\boldsymbol{\beta}}_j)^\top (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big| \le \\
\le & 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{is} - \sigma_i \sum_{s=1}^{t} Z_{is} \Big| + 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{js} - \sigma_j \sum_{s=1}^{t} Z_{js} \Big| + \\
& + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_j - \widehat{\boldsymbol{\beta}}_j)^\top (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big| = \\
= & 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{is} - \sigma_i \sum_{s=1}^{t} \big(\mathbb{B}_i(s) - \mathbb{B}_i(s-1)\big) \Big| + \\
& + 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{js} - \sigma_j \sum_{s=1}^{t} \big(\mathbb{B}_j(s) - \mathbb{B}_j(s-1)\big) \Big| + \\
& + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \max_{1\le t\le T} \Big| \sum_{s=1}^{t} (\boldsymbol{\beta}_j - \widehat{\boldsymbol{\beta}}_j)^\top (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big| = \\
= & 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{is} - \sigma_i \mathbb{B}_i(t) \Big| + 2 \max_{1\le t\le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_{js} - \sigma_j \mathbb{B}_j(t) \Big| +
\end{aligned}$$

$$+ \max_{1\le t\le T} \Big| (\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)^\top \sum_{s=1}^t (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \max_{1\le t\le T} \Big| (\boldsymbol{\beta}_j - \widehat{\boldsymbol{\beta}}_j)^\top \sum_{s=1}^t (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big|$$

Now suppose that for all $1 \le i \le n$ we have $\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i = o_P(T^{-1/2})$ almost surely. Then applying the strong approximation result (1.1), we can infer that

$$\max_{1\le t\le T} |A_{ij,t}| = o(T^{1/q}) + o_P(T^{-1/2}) \max_{1\le t\le T} \Big| \sum_{s=1}^t (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + o_P(T^{-1/2}) \max_{1\le t\le T} \Big| \sum_{s=1}^t (\mathbf{X}_{js} - \bar{\mathbf{X}}_j) \Big| \tag{1.6}$$

Standard arguments show that $\max_{(u,h)\in\mathcal{G}_T} W_T(u,h) = O(1/\sqrt{Th_{\min}})$. Plugging (1.6) in (1.5) and then in (1.4), we can thus infer that

$$\begin{aligned} \big|\widetilde{\Phi}_{n,T} - \Phi^\diamond_{n,T}\big| &\le \{\widetilde{\sigma}_i^2 + \widetilde{\sigma}_j^2\}^{-1/2} \max_{(u,h)\in\mathcal{G}_T} W_T(u,h) \max_{1\le i<j\le n} \max_{1\le t\le T} |A_{ij,t}| \le \\ &\le o\Big(\frac{T^{1/q}}{\sqrt{Th_{\min}}}\Big) + o_P\Big(\frac{T^{-1/2}}{\sqrt{Th_{\min}}}\Big) \max_{1\le i\le n} \max_{1\le t\le T} \Big| \sum_{s=1}^t (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big|. \end{aligned} \tag{1.7}$$

Now consider $|\Phi^\diamond_{n,T} - \Phi_{n,T}|$. Since $\phi_{ij,T}(u,h)$ conditionally on $\{\mathbf{X}_{it}\}$ is distributed as $N(\widehat{\boldsymbol{\beta}}_i^\top(\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - \widehat{\boldsymbol{\beta}}_j^\top(\mathbf{X}_{jt} - \bar{\mathbf{X}}_j), \sigma_i^2 + \sigma_j^2)$ for all $(u,h) \in \mathcal{G}_T$ and all $1 \le i < j \le n$, $|\mathcal{G}_T| = O(T^\theta)$ for some large but fixed constant θ by Assumption (??), n is fixed and $\widehat{\sigma}_i^2 = \sigma_i^2 + o_p(\rho_T)$ as well as $\widetilde{\sigma}_i^2 = \sigma_i^2 + o_p(\rho_T)$, we can establish that

$$\begin{aligned} \big|\Phi^\diamond_{n,T} - \Phi_{n,T}\big| &\le \max_{1\le i<j\le n} \max_{(u,h)\in\mathcal{G}_T} \Big| \frac{\phi_{ij,T}(u,h)}{\{\widetilde{\sigma}_i^2 + \widetilde{\sigma}_j^2\}^{1/2}} - \frac{\phi_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \Big| = \\ &= o_P(\rho_T\sqrt{\log T}) + o_P(\rho_T) \max_{1\le i\le n} \Big| \sum_{s=1}^T \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big|. \end{aligned} \tag{1.8}$$

Plugging (1.7) and (1.8) in (1.3) completes the proof. □

Auxiliary results using anti-concentration bounds

In this section, we establish some properties of the Gaussian statistic $\Phi_{n,T}$ defined in (**??**). We in particular show that $\Phi_{n,T}$ does not concentrate too strongly in small regions of the form $[x - \delta_T, x + \delta_T]$ with δ_T converging to zero.

Proposition A.2. *Under the conditions of Theorem 0.1, it holds that*

$$\sup_{x\in\mathbb{R}} \mathbb{P}\Big(|\Phi_{n,T} - x| \le \delta_T \Big| \{\mathbf{X}_{it} : 1 \le i \le n, 1 \le t \le T\} \Big) = o(1) \ \ a.s.,$$

where $\delta_T = T^{1/q}/\sqrt{Th_{\min}} + \rho_T\sqrt{\log T}$.

Proof of Proposition A.2. We write $x = (u,h)$ along with $\mathcal{G}_T = \{x : x \in \mathcal{G}_T\} = \{x_1, \dots, x_p\}$, where $p := |\mathcal{G}_T| \le O(T^\theta)$ for some large but fixed $\theta > 0$ by our assumptions. Moreover, for $k = 1, \dots, p$, we set

$$U_{ij,2k-1} = \frac{\phi_{ij,T}(x_{k1}, x_{k2})}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} - \lambda(x_{k2})$$

$$U_{ij,2k} = -\frac{\phi_{ij,T}(x_{k1}, x_{k2})}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} - \lambda(x_{k2})$$

with $x_k = (x_{k1}, x_{k2})$. This notation allows us to write

$$\Phi_{n,T} = \max_{1 \le i < j \le n} \max_{1 \le k \le 2p} U_{ij,k},$$

where $(U_{12,1}, \dots, U_{(n-1)n,2p})^\top \in \mathbb{R}^{n(n-1)p}$ is a Gaussian random vector.
Here we need proposition with conditional mean and variances!
The main technical tool for proving Proposition A.2 are anti-concentration bounds for Gaussian random vectors. The following proposition slightly generalizes anti-concentration results derived in Chernozhukov et al. (2015), in particular Theorem 3 therein.

Proposition A.3. *Let $(X_1, \dots, X_p)^\top$ be a Gaussian random vector in $\mathbb{R}^p$ with $\mathbb{E}[X_j] = \mu_j$ and $\mathrm{Var}(X_j) = \sigma_j^2 > 0$ for $1 \le j \le p$. Define $\overline{\mu} = \max_{1 \le j \le p} |\mu_j|$ together with $\underline{\sigma} = \min_{1 \le j \le p} \sigma_j$ and $\overline{\sigma} = \max_{1 \le j \le p} \sigma_j$. Moreover, set $a_p = \mathbb{E}[\max_{1 \le j \le p}(X_j - \mu_j)/\sigma_j]$ and $b_p = \mathbb{E}[\max_{1 \le j \le p}(X_j - \mu_j)]$. For every $\delta > 0$, it holds that*

$$\sup_{x \in \mathbb{R}} \mathbb{P}\Big(\big| \max_{1 \le j \le p} X_j - x \big| \le \delta \Big) \le C\delta \big\{ \overline{\mu} + a_p + b_p + \sqrt{1 \vee \log(\underline{\sigma}/\delta)} \big\},$$

where $C > 0$ depends only on $\underline{\sigma}$ and $\overline{\sigma}$.

The proof of Proposition A.3 is provided in Khismatullina and Vogt (2018). $\square$

Proof of Theorem 0.1

To prove Theorem 0.1, we make use of the two auxiliary results derived above. By Proposition A.1, there exist statistics $\widetilde{\Phi}_{n,T}$ for $T = 1, 2, \dots$ which are distributed as $\widehat{\Phi}_T$ for any $T \ge 1$ and which have the property that

$$\begin{aligned} \big| \widetilde{\Phi}_{n,T} - \Phi_{n,T} \big| = o_p\Big(& \frac{T^{1/q}}{\sqrt{T h_{\min}}} + \rho_T \sqrt{\log T} + \rho_T \max_{1 \le i \le n} \Big| \sum_{s=1}^{T} \widehat{\boldsymbol{\beta}}_i^\top (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| + \\ & + \frac{T^{-1/2}}{\sqrt{T h_{min}}} \max_{1 \le i \le n} \max_{1 \le t \le T} \Big| \sum_{s=1}^{t} (\mathbf{X}_{is} - \bar{\mathbf{X}}_i) \Big| \Big), \end{aligned} \tag{1.9}$$

where $\Phi_{n,T}$ is a Gaussian statistic as defined in (**??**). The approximation result (1.9) allows us to replace the multiscale statistic $\widehat{\Phi}_{n,T}$ by an identically distributed version $\widetilde{\Phi}_{n,T}$ which is close to the Gaussian statistic $\Phi_{n,T}$.
In the next step, we show that

$$\sup_{x \in \mathbb{R}} \big| \mathbb{P}(\widetilde{\Phi}_{n,T} \le x) - \mathbb{P}(\Phi_{n,T} \le x) \big| = o(1) \text{ a.s.}, \tag{1.10}$$

which immediately implies the statement of Theorem 0.1.

References

Berkes, I., Liu, W. and Wu, W. B. (2014). Komlós-Major-Tusnády approximation under dependence. *Annals of Probability*, **42** 794–817.

Chernozhukov, V., Chetverikov, D. and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probability Theory and Related Fields*, **162** 47–70.

Khismatullina, M. and Vogt, M. (2018). Multiscale inference and long-run variance estimation in nonparametric regression with time series errors. *Preprint*.