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*

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

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)*

$$\big| \widetilde{\Phi}_{n,T} - \Phi_{n,T} \big| = o_p\Big(\frac{T^{1/q}}{\sqrt{T h_{\min}}} + \rho_T \sqrt{\log T} \Big),$$

*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 $[\widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_T] \overset{\mathcal{D}}{=} [\varepsilon_1, \dots, \varepsilon_T]$ 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 $Y_{it} = (\widehat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i)^\top \mathbf{X}_{it} + m_i(t/T) + \big(\alpha_i - \widehat{\alpha}_i\big) + \varepsilon_{it}$ replaced by $\widetilde{Y}_{t,T} = (\widehat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i)^\top \mathbf{X}_{it} + m_i(t/T) + \big(\alpha_i - \widehat{\alpha}_i\big) + \widetilde{\varepsilon}_{it}$ 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 $\phi_{ij,T}(u,h) = \sum_{t=1}^{T} w_{t,T}(u,h) \big\{ \widehat{\sigma}_i (Z_{it} - \bar{Z}_i) - \widehat{\sigma}_j (Z_{jt} - \bar{Z}_j) \big\}$ 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 $|\Phi^\diamond_{n,T} - \Phi_{n,T}|$. Since $\phi_{ij,T}(u,h) \sim N(0, \sigma_i^2 + \sigma_j^2)$ 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

$$\big| \Phi^\diamond_{n,T} - \Phi_{n,T} \big| \le \max_{1 \le i < j \le n} \max_{(u,h) \in \mathcal{G}_T} \Big| \frac{\phi_{ij,T}(u,h)}{\{\widetilde{\sigma}_i^2 + \widetilde{\sigma}_j^2\}^{1/2}} - \frac{\phi_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \Big| = o_P(\rho_T \sqrt{\log T}). \tag{1.4}$$

Plugging (1.4) in (1.3), we get

$$\big| \widetilde{\Phi}_{n,T} - \Phi_{n,T} \big| \le \big| \widetilde{\Phi}_{n,T} - \Phi^\diamond_{n,T} \big| + o_P(\rho_T \sqrt{\log T}).$$

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

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

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

$$\begin{aligned} \big| \widetilde{\phi}_{ij,T}(u,h) - \phi_{ij,T}(u,h) \big| &\le W_T(u,h) \max_{1 \le t \le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_s - \sigma \sum_{s=1}^{t} \big\{ \mathbb{B}(s) - \mathbb{B}(s-1) \big\} \Big| \\ &= W_T(u,h) \max_{1 \le t \le T} \Big| \sum_{s=1}^{t} \widetilde{\varepsilon}_s - \sigma \mathbb{B}(t) \Big|, \end{aligned}$$

where

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

Standard arguments show that $\max_{(u,h)\in\mathcal{G}_T} W_T(u,h) = O(1/\sqrt{Th_{\min}})$. Applying the strong approximation result (1.1), we can thus infer that

$$\begin{aligned}\big|\widetilde{\Phi}_T - \Phi_T^{\diamond}\big| &\le \widetilde{\sigma}^{-1} \max_{(u,h)\in\mathcal{G}_T} \big|\widetilde{\phi}_T(u,h) - \phi_T(u,h)\big| \\ &\le \widetilde{\sigma}^{-1} \max_{(u,h)\in\mathcal{G}_T} W_T(u,h) \max_{1\le t\le T} \Big|\sum_{s=1}^{t} \widetilde{\varepsilon}_s - \sigma\mathbb{B}(t)\Big| = o_p\Big(\frac{T^{1/q}}{\sqrt{Th_{\min}}}\Big). \qquad (1.5)\end{aligned}$$

Plugging (1.5) into (1.3) completes the proof. $\square$

Auxiliary results using anti-concentration bounds

In this section, we establish some properties of the Gaussian statistic Φ_T defined in (**??**). We in particular show that Φ_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* **??**, *it holds that*

$$\sup_{x\in\mathbb{R}} \mathbb{P}\Big(|\Phi_T - x| \le \delta_T\Big) = o(1),$$

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

Proof of Proposition A.2. 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 at the end of this section for completeness. To apply Proposition A.3 to our setting at hand, we introduce the following notation: 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 $j = 1,\dots,p$, we set

$$\begin{aligned}X_{2j-1} &= \frac{\phi_T(x_{j1},x_{j2})}{\sigma} - \lambda(x_{j2}) \\ X_{2j} &= -\frac{\phi_T(x_{j1},x_{j2})}{\sigma} - \lambda(x_{j2})\end{aligned}$$

with $x_j = (x_{j1}, x_{j2})$. This notation allows us to write

$$\Phi_T = \max_{1 \le j \le 2p} X_j,$$

where $(X_1, \dots, X_{2p})^\top$ is a Gaussian random vector with the following properties: (i) $\mu_j := \mathbb{E}[X_j] = -\lambda(x_{j2})$ and thus $\overline{\mu} = \max_{1 \le j \le 2p} |\mu_j| \le C\sqrt{\log T}$, and (ii) $\sigma_j^2 := \operatorname{Var}(X_j) = 1$ for all j. Since $\sigma_j = 1$ for all j, it holds that $a_{2p} = b_{2p}$. Moreover, as the variables $(X_j - \mu_j)/\sigma_j$ are standard normal, we have that $a_{2p} = b_{2p} \le \sqrt{2\log(2p)} \le C\sqrt{\log T}$. With this notation at hand, we can apply Proposition A.3 to obtain that

$$\sup_{x \in \mathbb{R}} \mathbb{P}\Big(\big|\Phi_T - x\big| \le \delta_T\Big) \le C\delta_T\Big[\sqrt{\log T} + \sqrt{\log(1/\delta_T)}\Big] = o(1)$$

with $\delta_T = T^{1/q}/\sqrt{Th_{\min}} + \rho_T\sqrt{\log T}$, which is the statement of Proposition A.2. $\square$

Proof of Theorem ??

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

$$\big|\widetilde{\Phi}_T - \Phi_T\big| = o_p\Big(\frac{T^{1/q}}{\sqrt{Th_{\min}}} + \rho_T\sqrt{\log T}\Big), \tag{1.6}$$

where Φ_T is a Gaussian statistic as defined in (**??**). The approximation result (1.6) allows us to replace the multiscale statistic $\widehat{\Phi}_T$ by an identically distributed version $\widetilde{\Phi}_T$ which is close to the Gaussian statistic Φ_T. In the next step, we show that

$$\sup_{x \in \mathbb{R}} \big|\mathbb{P}(\widetilde{\Phi}_T \le x) - \mathbb{P}(\Phi_T \le x)\big| = o(1), \tag{1.7}$$

which immediately implies the statement of Theorem **??**. For the proof of (1.7), we use the following simple lemma:

Lemma A.4. *Let V_T and W_T be real-valued random variables for $T = 1, 2, \dots$ such that $V_T - W_T = o_p(\delta_T)$ with some $\delta_T = o(1)$. If*

$$\sup_{x \in \mathbb{R}} \mathbb{P}(|V_T - x| \le \delta_T) = o(1), \tag{1.8}$$

then

$$\sup_{x \in \mathbb{R}} \big|\mathbb{P}(V_T \le x) - \mathbb{P}(W_T \le x)\big| = o(1). \tag{1.9}$$

The statement of Lemma A.4 can be summarized as follows: If W_T can be approximated by V_T in the sense that $V_T - W_T = o_p(\delta_T)$ and if V_T does not concentrate too strongly in small regions of the form $[x - \delta_T, x + \delta_T]$ as assumed in (1.8), then the distribution of W_T can be approximated by that of V_T in the sense of (1.9).

Proof of Lemma A.4. It holds that

$$\begin{aligned}
&\big|\mathbb{P}(V_T \le x) - \mathbb{P}(W_T \le x)\big| \\
&= \big|\mathbb{E}\big[1(V_T \le x) - 1(W_T \le x)\big]\big| \\
&\le \big|\mathbb{E}\big[\big\{1(V_T \le x) - 1(W_T \le x)\big\}1(|V_T - W_T| \le \delta_T)\big]\big| + \big|\mathbb{E}\big[1(|V_T - W_T| > \delta_T)\big]\big| \\
&\le \mathbb{E}\big[1(|V_T - x| \le \delta_T, |V_T - W_T| \le \delta_T)\big] + o(1) \\
&\le \mathbb{P}(|V_T - x| \le \delta_T) + o(1).
\end{aligned}$$

$\square$

We now apply this lemma with $V_T = \Phi_T$, $W_T = \widetilde{\Phi}_T$ and $\delta_T = T^{1/q}/\sqrt{Th_{\min}} + \rho_T\sqrt{\log T}$: From (1.6), we already know that $\widetilde{\Phi}_T - \Phi_T = o_p(\delta_T)$. Moreover, by Proposition A.2, it holds that

$$\sup_{x \in \mathbb{R}} \mathbb{P}\Big(|\Phi_T - x| \le \delta_T \Big) = o(1). \tag{1.10}$$

Hence, the conditions of Lemma A.4 are satisfied. Applying the lemma, we obtain (1.7), which completes the proof of Theorem **??**.

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.