

**Proposition A.1.** *There exists a sequence of random numbers  $\{\gamma_{n,T}\}_T$ , that converges to 0 as  $T \rightarrow \infty$ , such that*

$$\mathbb{P}\left(\left|\widehat{\Phi}_{n,T} - \widehat{\Phi}_{n,T}\right| > \gamma_{n,T}\right) = o(1). \quad (1)$$

**Proof of Proposition A.1.** Straightforward calculations yield that

$$\begin{aligned} \left|\widehat{\Phi}_{n,T} - \widehat{\Phi}_{n,T}\right| &\leq \max_{1 \leq i < j \leq n} \max_{(u,h) \in \mathcal{G}_T} \left( \left| \frac{\widehat{\phi}_{ij,T}(u,h)}{(\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2)^{1/2}} - \frac{\widehat{\phi}_{ij,T}(u,h)}{(\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2)^{1/2}} \right| + \left| \frac{\widehat{\phi}_{ij,T}(u,h)}{(\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2)^{1/2}} - \frac{\widehat{\phi}_{ij,T}(u,h)}{\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{1/2}} \right| \right) \leq \\ &\leq \max_{1 \leq i < j \leq n} \left( \left| \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} - \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \right| \max_{(u,h) \in \mathcal{G}_T} \left| \widehat{\phi}_{ij,T}(u,h) \right| \right) + \\ &\quad + \max_{1 \leq i < j \leq n} \left( \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \max_{(u,h) \in \mathcal{G}_T} \left| \widehat{\phi}_{ij,T}(u,h) - \widehat{\phi}_{ij,T}(u,h) \right| \right). \end{aligned}$$

Then, consider the difference of the kernel averages:

$$\begin{aligned} \left| \widehat{\phi}_{ij,T}(u,h) - \widehat{\phi}_{ij,T}(u,h) \right| &= \left| \sum_{t=1}^T w_{t,T}(u,h) \{ (\beta_i - \widehat{\beta}_i)^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) - (\beta_j - \widehat{\beta}_j)^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \} \right| \leq \\ &\leq \left| \sum_{t=1}^T w_{t,T}(u,h) (\beta_i - \widehat{\beta}_i)^\top (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) \right| + \left| \sum_{t=1}^T w_{t,T}(u,h) (\beta_j - \widehat{\beta}_j)^\top (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \right| = \\ &= |\beta_i - \widehat{\beta}_i|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) \right| + |\beta_j - \widehat{\beta}_j|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \right| = \\ &= |\beta_i - \widehat{\beta}_i|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) \right| + |\beta_j - \widehat{\beta}_j|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) (\mathbf{X}_{jt} - \bar{\mathbf{X}}_j) \right| \leq \\ &\leq |\beta_i - \widehat{\beta}_i|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) \mathbf{X}_{it} \right| + |(\beta_i - \widehat{\beta}_i)^\top \bar{\mathbf{X}}_i| \left| \sum_{t=1}^T w_{t,T}(u,h) \right| + \\ &\quad + |\beta_j - \widehat{\beta}_j|^\top \left| \sum_{t=1}^T w_{t,T}(u,h) \mathbf{X}_{jt} \right| + |(\beta_j - \widehat{\beta}_j)^\top \bar{\mathbf{X}}_j| \left| \sum_{t=1}^T w_{t,T}(u,h) \right| \end{aligned}$$

Hence,

$$\begin{aligned} \left|\widehat{\Phi}_{n,T} - \widehat{\Phi}_{n,T}\right| &\leq \max_{1 \leq i < j \leq n} \left| \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} - \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \right| \max_{1 \leq i < j \leq n} \max_{(u,h) \in \mathcal{G}_T} \left| \widehat{\phi}_{ij,T}(u,h) \right| + \\ &\quad + 2 \max_{1 \leq i < j \leq n} \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \max_{1 \leq i \leq n} \left( |\beta_i - \widehat{\beta}_i|^\top \max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=1}^T w_{t,T}(u,h) \mathbf{X}_{it} \right| \right) + \\ &\quad + 2 \max_{1 \leq i < j \leq n} \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \max_{1 \leq i \leq n} |(\beta_i - \widehat{\beta}_i)^\top \bar{\mathbf{X}}_i| \max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=1}^T w_{t,T}(u,h) \right| \end{aligned} \quad (2)$$

We start by evaluating the second summand in (2).

First, by our assumptions  $\widehat{\sigma}_i^2 = \sigma_i^2 + o_P(\rho_T)$ . Moreover, for all  $i \in \{1, \dots, n\}$  we know  $\sigma_i^2 \neq 0$ . Hence,

$$\max_{1 \leq i < j \leq n} \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} = O_P(1). \quad (3)$$

Then, by Theorem ??, we know that

$$|\beta_i - \hat{\beta}_i| = O_P(1/\sqrt{T}). \quad (4)$$

Now consider the term  $\left| \sum_{t=1}^T w_{t,T}(u, h) \mathbf{X}_{it} \right|$ . Without loss of generality, we can regard the covariates  $\mathbf{X}_{it}$  to be scalars  $X_{it}$ , not vectors. The proof in case of vectors proceeds analogously.

By construction the weights  $w_{t,T}(u, h)$  are not equal to 0 if and only if  $T(u - h) \leq t \leq T(u + h)$ . We can use this fact to rewrite

$$\left| \sum_{t=1}^T w_{t,T}(u, h) X_{it} \right| = \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u, h) X_{it} \right|.$$

Note that

$$\begin{aligned} \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}^2(u, h) &= \sum_{t=1}^T w_{t,T}^2(u, h) = \\ &= \sum_{t=1}^T \frac{K^2 \left( \frac{t}{h} - u \right) \left[ S_{T,2}(u, h) - \left( \frac{t}{h} - u \right) S_{T,1}(u, h) \right]^2}{\left\{ \sum_{s=1}^T K^2 \left( \frac{s}{h} - u \right) \left[ S_{T,2}(u, h) - \left( \frac{s}{h} - u \right) S_{T,1}(u, h) \right]^2 \right\}} = \\ &= 1. \end{aligned}$$

Denoting by  $D_{T,u,h}$  the number of integers between  $\lfloor T(u-h) \rfloor$  and  $\lceil T(u+h) \rceil$  incl. (with obvious bounds  $2Th \leq D_{T,u,h} \leq 2Th + 2$ ), we can normalize the weights as follows:

$$\sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} (\sqrt{D_{T,u,h}} \cdot w_{t,T}(u, h))^2 = D_{T,u,h}.$$

According to Theorem ?? (Theorem 2(ii) in ?), if we denote the weights from the theorem as  $a_t = \sqrt{D_{T,u,h}} \cdot w_{t,T}(u, h)$ , we can bound the following probability:

$$\begin{aligned} \mathbb{P} \left( \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} \sqrt{D_{T,u,h}} \cdot w_{t,T}(u, h) X_{it} \right| \geq x \right) &\leq \\ &\leq C_1 \frac{\left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |\sqrt{D_{T,u,h}} \cdot w_{t,T}(u, h)|^{q'} \right) \|X_{i\cdot}\|_{q',\alpha}^{q'}}{x^{q'}} + C_2 \exp \left( -\frac{C_3 x^2}{D_{T,u,h} \|X_{i\cdot}\|_{2,\alpha}^2} \right) = \\ &= C_1 \frac{(\sqrt{D_{T,u,h}})^{q'} \left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |w_{t,T}(u, h)|^{q'} \right) \|X_{i\cdot}\|_{q',\alpha}^{q'}}{x^{q'}} + C_2 \exp \left( -\frac{C_3 x^2}{D_{T,u,h} \|X_{i\cdot}\|_{2,\alpha}^2} \right) \end{aligned}$$

Now take any  $\delta > 0$ :

$$\begin{aligned}
& \mathbb{P} \left( \frac{\max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) X_{it} \right|}{\sqrt{T}} \geq \delta \right) = \\
& = \mathbb{P} \left( \max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) X_{it} \right| \geq \delta \sqrt{T} \right) \leq \\
& \stackrel{\text{Boole's inequality}}{\leq} \sum_{(u,h) \in \mathcal{G}_T} \mathbb{P} \left( \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) X_{it} \right| \geq \delta \sqrt{T} \right) = \\
& \stackrel{\text{"normalisation"}}{=} \sum_{(u,h) \in \mathcal{G}_T} \mathbb{P} \left( \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} \sqrt{D_{T,u,h}} \cdot w_{t,T}(u,h) X_{it} \right| \geq \delta \sqrt{D_{T,u,h} T} \right) \leq \\
& \stackrel{\text{Wu's Theorem}}{\leq} \sum_{(u,h) \in \mathcal{G}_T} \left[ C_1 \frac{(\sqrt{D_{T,u,h}})^{q'} (\sum |w_{t,T}(u,h)|^{q'}) \|X_{i\cdot}\|_{q',\alpha}^{q'}}{(\delta \sqrt{D_{T,u,h} T})^{q'}} + C_2 \exp \left( -\frac{C_3 (\delta \sqrt{D_{T,u,h} T})^2}{D_{T,u,h} \|X_{i\cdot}\|_{2,\alpha}^2} \right) \right] = \\
& \stackrel{\text{simplification}}{=} \sum_{(u,h) \in \mathcal{G}_T} \left[ C_1 \frac{(\sum |w_{t,T}(u,h)|^{q'}) \|X_{i\cdot}\|_{q',\alpha}^{q'}}{(\delta \sqrt{T})^{q'}} + C_2 \exp \left( -\frac{C_3 \delta^2 T}{\|X_{i\cdot}\|_{2,\alpha}^2} \right) \right] \leq \\
& \leq C_1 \frac{T^\theta \|X_{i\cdot}\|_{q',\alpha}^{q'}}{T^{q'/2} \cdot \delta^{q'}} \max_{(u,h) \in \mathcal{G}_T} \left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |w_{t,T}(u,h)|^{q'} \right) + C_2 T^\theta \exp \left( -\frac{C_3 \delta^2 T}{\|X_{i\cdot}\|_{2,\alpha}^2} \right) = \\
& = C \frac{T^{\theta-q'/2}}{\delta^{q'}} + C T^\theta \exp(-CT\delta^2).
\end{aligned}$$

where the symbol  $C$  denotes a universal real constant that does not depend neither on  $T$  nor on  $\delta$  and that takes a different value on each occurrence. Here in the last equality we used the following facts:

1.  $\|X_{i\cdot}\|_{q',\alpha}^{q'} = \sup_{t \geq 0} (t+1)^\alpha \sum_{s=t}^\infty \delta_{q'}(H_i, s) < \infty$  holds true since  $\sum_{s=t}^\infty \delta_{q'}(H_i, s) = O(t^{-\alpha})$  by Assumption ??;
2.  $\max_{(u,h) \in \mathcal{G}_T} \left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |w_{t,T}(u,h)|^{q'} \right) < \infty$  because for every  $x \in [0, 1]$  we have  $0 \leq |x|^{q'/2} \leq x \leq 1$ . Thus, since  $\sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}^2(u,h) = 1$ , we have  $0 \leq w_{t,T}^2(u,h) \leq 1$  for all  $t$  and

$$0 \leq |w_{t,T}(u,h)|^{q'} = |w_{t,T}^2(u,h)|^{q'/2} \leq w_{t,T}^2(u,h) \leq 1.$$

This leads us to a bound:

$$\max_{(u,h) \in \mathcal{G}_T} \left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |w_{t,T}(u,h)|^{q'} \right) \leq \max_{(u,h) \in \mathcal{G}_T} \left( \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} |w_{t,T}(u,h)|^2 \right) = 1 < \infty.$$

3.  $\|X_{i\cdot}\|_{2,\alpha}^2 < \infty$  (follows from 1).

By Assumption ??,  $\theta - q'/2 < 0$  and the term on the RHS of the above inequality is converging to 0 as  $T \rightarrow \infty$  for any fixed  $\delta > 0$ . Hence,

$$\max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) X_{it} \right| = o_P(\sqrt{T}),$$

and similarly

$$\max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) \mathbf{X}_{it} \right| = o_P(\sqrt{T}). \quad (5)$$

Combining (3), (4) and (5), we get the following:

$$\begin{aligned} & 2 \max_{1 \leq i < j \leq n} \{\hat{\sigma}_i^2 + \hat{\sigma}_j^2\}^{-1/2} \max_{1 \leq i \leq n} \left( |\beta_i - \hat{\beta}_i|^\top \max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=1}^T w_{t,T}(u,h) \mathbf{X}_{it} \right| \right) = \\ & = O_P(1) \cdot O_P(1/\sqrt{T}) \cdot o_P(\sqrt{T}) = o_P(1). \end{aligned} \quad (6)$$

Now, consider the third summand in (2).

Similarly as before,

$$\max_{1 \leq i < j \leq n} \{\hat{\sigma}_i^2 + \hat{\sigma}_j^2\}^{-1/2} = O_P(1) \quad (7)$$

and

$$|\beta_i - \hat{\beta}_i| = O_P(1/\sqrt{T}). \quad (8)$$

Then, by Proposition ??  $\bar{\mathbf{X}}_i = o_P(1)$ .

Finally, consider the local linear kernel weights  $w_{t,T}(u,h)$  defined in (?). Again, by construction the weights  $w_{t,T}(u,h)$  are not equal to 0 if and only if  $T(u-h) \leq t \leq T(u+h)$ . We can use this fact to bound  $\left| \sum_{t=1}^T w_{t,T}(u,h) \right|$  for all  $(u,h) \in \mathcal{G}_T$  using the Cauchy-Schwarz inequality:

$$\begin{aligned} \left| \sum_{t=1}^T w_{t,T}(u,h) \right| &= \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) \right| = \left| \sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}(u,h) \cdot 1 \right| \leq \\ &\leq \sqrt{\sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} w_{t,T}^2(u,h)} \sqrt{\sum_{t=\lfloor T(u-h) \rfloor}^{\lceil T(u+h) \rceil} 1^2} = \\ &= \sqrt{1} \cdot \sqrt{D_{T,u,h}} \leq \sqrt{2Th} + 2 \leq \sqrt{2Th_{\max}} + 2 \leq \sqrt{T} + 2. \end{aligned}$$

Hence,

$$\max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=1}^T w_{t,T}(u,h) \right| = O(\sqrt{T}). \quad (9)$$

Combining (7), (8), Proposition ?? and (9), we get the following:

$$\begin{aligned} & 2 \max_{1 \leq i < j \leq n} \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \max_{1 \leq i \leq n} |(\beta_i - \widehat{\beta}_i)^\top \bar{\mathbf{X}}_i| \max_{(u,h) \in \mathcal{G}_T} \left| \sum_{t=1}^T w_{t,T}(u, h) \right| = \\ & = O_P(1) \cdot O_P(1/\sqrt{T}) \cdot o_P(1) \cdot O(\sqrt{T}) = o_P(1). \end{aligned} \quad (10)$$

Lastly, we look at the first summand in (2). Since  $\widehat{\sigma}_i^2 = \sigma_i^2 + o_P(\rho_T)$  and  $\widehat{\sigma}_i^2 = \sigma_i^2 + o_P(\rho_T)$  by our assumptions, we have that

$$\max_{1 \leq i < j \leq n} |\{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} - \{\sigma_i^2 + \sigma_j^2\}^{-1/2}| = o_P(\rho_T). \quad (11)$$

Then since  $\widehat{\phi}_{ij,T}(u, h)$  has the same distribution as  $\widetilde{\phi}_{ij,T}(u, h)$  for each  $1 \leq i < j \leq n$  and each  $(u, h) \in \mathcal{G}_T$ , we can look at  $\max_{(u,h) \in \mathcal{G}_T} |\widetilde{\phi}_{ij,T}(u, h)|$  instead:

$$\begin{aligned} \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} \widehat{\phi}_{ij,T}(u, h) \right| \geq c_T \right) &= \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} \widetilde{\phi}_{ij,T}(u, h) \right| \geq c_T \right) = \\ &= \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} \widetilde{\phi}_{ij,T}(u, h) - \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| + \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| \right| \geq c_T \right) \leq \\ &\leq \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} \widetilde{\phi}_{ij,T}(u, h) - \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| \right| \geq c_T/2 \right) + \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| \right| \geq c_T/2 \right) \leq \\ &\leq \mathbb{P} \left( \max_{(u,h) \in \mathcal{G}_T} |\widetilde{\phi}_{ij,T}(u, h) - \phi'_{ij,T}(u, h)| \geq c_T/2 \right) + \mathbb{P} \left( \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| \geq c_T/2 \right). \end{aligned} \quad (12)$$

Here we will need one result that we will prove further: by (??) we have

$$\max_{(u,h) \in \mathcal{G}_T} |\widetilde{\phi}_{ij,T}(u, h) - \phi'_{ij,T}(u, h)| = o_P \left( \frac{T^{1/q}}{\sqrt{T} h_{\min}} \right).$$

Furthermore,  $\phi'_{ij,T}(u, h)$  is distributed as  $N(0, \sigma_i^2 + \sigma_j^2)$  for all  $(u, h) \in \mathcal{G}_T$  and all  $1 \leq i < j \leq n$  and  $|\mathcal{G}_T| = O(T^\theta)$  for some large but fixed constant  $\theta$  by Assumption ??. By the standard results from the probability theory, we know that

$$\max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| = O_P(\sqrt{\log T}).$$

Hence, if we take  $c_T = o(\sqrt{\log T})$  in (12), we will get the following:

$$\begin{aligned} \mathbb{P} \left( \left| \max_{(u,h) \in \mathcal{G}_T} \widehat{\phi}_{ij,T}(u, h) \right| \geq c_T \right) &\leq \\ &\leq \mathbb{P} \left( \max_{(u,h) \in \mathcal{G}_T} |\widetilde{\phi}_{ij,T}(u, h) - \phi'_{ij,T}(u, h)| \geq c_T/2 \right) + \mathbb{P} \left( \max_{(u,h) \in \mathcal{G}_T} |\phi'_{ij,T}(u, h)| \geq c_T/2 \right) = \\ &= o(1) + o(1) = o(1), \end{aligned}$$

which means that

$$\left| \max_{(u,h) \in \mathcal{G}_T} \widehat{\phi}_{ij,T}(u, h) \right| = o_P(\sqrt{\log T}) \quad (13)$$

Combining (11) and (13), we get the following:

$$\begin{aligned}
& \max_{1 \leq i < j \leq n} \left| \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} - \{\widehat{\sigma}_i^2 + \widehat{\sigma}_j^2\}^{-1/2} \right| \max_{1 \leq i < j \leq n} \max_{(u,h) \in \mathcal{G}_T} \left| \widehat{\phi}_{ij,T}(u, h) \right| = \\
& = o_P(\rho_T) \cdot o_P(\sqrt{\log T}) = \\
& = o_P(1).
\end{aligned} \tag{14}$$

Plugging (6), (10) and (14) in (2), we get that  $|\widehat{\Phi}_{n,T} - \widehat{\Phi}_{n,T}| = o_P(1)$  and the statement of the theorem follows.  $\square$