Supplement to 'Sieve-Based Inference for Infinite-Variance Linear Processes'

Giuseppe Cavaliere, Iliyan Georgiev
and A.M. Robert Taylor

S.1. Introduction. This supplement contains additional theoretical results and proofs for the theory stated in Cavaliere, Georgiev and Taylor (2015), CGT hereafter. The supplement is organized as follows. Section S.2 provides a lemma with two tail inequalities regarding the series of the coefficients from the AR(∞) representations. Section S.3 reports the proof of Lemma 2 and corollaries from Section 6 in CGT. Section S.4 contains proofs of the results given in Section 7.3.1 of CGT. Finally, Section S.5 discusses the case of multiple restrictions.

S.2. A Tail Inequality. We first establish two inequalities between the tails of the series of autoregressive coefficients and their powers.

Lemma S.1. *Under Assumption 1, let $k^2/T + 1/k \to 0$ as $T \to \infty$. Then for large T, for η in a sufficiently small left neighborhood of $\alpha \wedge 1 := \min\{\alpha, 1\}$ and for $\zeta > 0$ sufficiently small, it holds that*

$$\text{(S.2.1)} \qquad \tilde{a}_T a_T^{-2}(k \sum_{j=k+1}^{\infty} |\beta_j|^{\eta})^{1/\eta} \leq \tilde{a}_T a_T^{-\zeta-2} + \sum_{j=k+1}^{\infty} |\beta_j|.$$

If $k^3/T + 1/k \to 0$ as $T \to \infty$, then also

$$\text{(S.2.2)} \qquad \tilde{a}_T a_T^{-2}(k \sum_{j=k+1}^{\infty} |\beta_j|^{\eta})^{1/\eta} \leq \tilde{a}_T a_T^{-9/4} + a_T^{-1/5} \sum_{j=k+1}^{\infty} |\beta_j|.$$

Proof. In the case of a finite-order AR representation the inequality is obvious, so we discuss the opposite case.

From $\sum_{j=1}^{\infty} j^{2/\delta}|\beta_j| < \infty$ it follows that $|\beta_j| \leq j^{-2/\delta}$ for large j. For fixed k, the expression $(\sum_{j=k+1}^{\infty} |\beta_j|)^{-\eta} \sum_{j=k+1}^{\infty} |\beta_j|^{\eta}$ cannot be prolonged by continuity to the zero sequence in ℓ_2, so we consider separately the sets

$$\mathbb{B}_l := \{\{|\beta_j|\}_{j=k+1}^{\infty} \quad : \quad 0 \leq |\beta_j| \leq a_T^{-\zeta} j^{-2/\delta}, j = k+1, ...\},$$
$$\mathbb{B}_u := \{\{|\beta_j|\}_{j=k+1}^{\infty} \quad : \quad a_T^{-\zeta} j^{-2/\delta} \leq |\beta_j| \leq j^{-2/\delta}, j = k+1, ...\}.$$

Using the upper bound in the standard estimate

$$\frac{(K+1)^{1-s}}{s-1} \leq \sum_{j=K+1}^{\infty} j^{-s} \leq \frac{K^{1-s}}{s-1} \text{ for } s>1, \tag{S.2.3}$$

we find that on $\mathbb{B}_l$,

$$\tilde{a}_T a_T^{-2}(k \sum_{j=k+1}^{\infty} |\beta_j|^{\eta})^{1/\eta} \leq C\tilde{a}_T a_T^{-\zeta-2}(k+1)^{\frac{2(\delta-\eta)}{\delta\eta}} = o(\tilde{a}_T a_T^{-\zeta-2})$$

for $T \geq 2$, $\delta < \eta$ and sufficiently small $\zeta > 0$. If $\mathbb{B}_u$ is equipped with the ℓ_2 metric, it becomes a compact in ℓ_2, for it is closed, bounded and for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $\{|\beta_j|\}_{j=k+1}^{\infty} \in \mathbb{B}_u$ it holds that $\sum_{j=N}^{\infty} \beta_j^2 \leq \sum_{j=N}^{\infty} j^{-4/\delta} < \epsilon$. The expression $(\sum_{j=k+1}^{\infty} |\beta_j|)^{-\eta} \sum_{j=k+1}^{\infty} |\beta_j|^{\eta}$ defines a continuous real function on $\mathbb{B}_u$ and, hence, attains a maximum there. Let $\{|\beta_j^{\sharp}|\}_{j=k+1}^{\infty}$ denote a maximizing sequence; by examining directional derivatives, it follows that $|\beta_j^{\sharp}|$ $(j = k+1, ...)$ satisfy

$$|\beta_j^{\sharp}| = \begin{cases} a_T^{-\zeta} j^{-2/\delta} & \text{if } B^{\sharp} < a_T^{-\zeta} j^{-2/\delta} \\ B^{\sharp} & \text{if } a_T^{-\zeta} j^{-2/\delta} \leq B^{\sharp} < j^{-2/\delta} \\ j^{-2/\delta} & \text{if } j^{-2/\delta} \leq B^{\sharp} \end{cases}$$

with $B^{\sharp} = \left(\sum_{j=k+1}^{\infty} |\beta_j^{\sharp}| \big/ \sum_{j=k+1}^{\infty} |\beta_j^{\sharp}|^{\eta}\right)^{\frac{1}{1-\eta}} \neq 0$. We examine this condition without attempting to find all $|\beta_j^{\sharp}|$ exactly.

As $B^{\sharp} > 0$ and $j^{-2/\delta}$ is decreasing in j, $|\beta_j^{\sharp}| = j^{-2/\delta}$ necessarily holds from some index onwards. Let $K_2 \geq k$ be the smallest natural $\geq k$ such that $|\beta_j^{\sharp}| = j^{-2/\delta}$ for $j \geq K_2 + 1$. Then $(K_2+1)^{-2/\delta} \leq B^{\sharp}$ and, if $K_2 > k$, then $B^{\sharp} < K_2^{-2/\delta}$ and $|\beta_j^{\sharp}| \neq j^{-2/\delta}$ for $j = k+1, ..., K_2$.

Still if $K_2 > k$, then either $|\beta_{k+1}^{\sharp}| = B^{\sharp}$ or $|\beta_{k+1}^{\sharp}| = a_T^{-\zeta} j^{-2/\delta} > B^{\sharp}$. In the former case it must be that $a_T^{-\zeta}(k+1)^{-2/\delta} \leq B^{\sharp}$, so $a_T^{-\zeta} j^{-2/\delta} \leq B^{\sharp}$ for all $j \geq k+1$, the value $a_T^{-\zeta} j^{-2/\delta}$ is never taken by $|\beta_j^{\sharp}|$ and at K_2 a switch between $B^{\sharp}$ and $j^{-2/\delta}$ takes place; define $K_1 = k$ in this case. On the other hand, if $|\beta_{k+1}^{\sharp}| = a_T^{-\zeta}(k+1)^{-2/\delta} > B^{\sharp}$, let $K_1 < K_2$ be the largest natural j such that $a_T^{-\zeta} j^{-2/\delta} > B^{\sharp}$ for $j = k+1, ..., K_1$. Then at K_1 a switch between $a_T^{-\zeta} j^{-2/\delta}$ and $B^{\sharp}$ or $j^{-2/\delta}$ takes place.

Summarizing,

$$|\beta_j^{\sharp}| = \begin{cases} a_T^{-\zeta} j^{-2/\delta} & k+1 \leq j \leq K_1 \\ B^{\sharp} & K_1+1 \leq j \leq K_2 \\ j^{-2/\delta} & j \geq K_2+1, ..., \end{cases}$$

with

$$B^\sharp = \left(\frac{a_T^{-\zeta} \sum_{j=k+1}^{K_1} j^{-2/\delta} + (K_2 - K_1)\, B^\sharp + \sum_{j=K_2+1}^{\infty} j^{-2/\delta}}{a_T^{-\zeta\eta} \sum_{j=k+1}^{K_1} j^{-2\eta/\delta} + (K_2 - K_1)\, B^{\sharp\eta} + \sum_{j=K_2+1}^{\infty} j^{-2\eta/\delta}} \right)^{\frac{1}{1-\eta}},$$

where the first two conditions may be satisfied by an empty set of j's. If switches do occur, then $a_T^{-\zeta}(K_1+1)^{-2/\delta} \le B^\sharp \le a_T^{-\zeta} K_1^{-2/\delta}$ holds at a switch away from the $a_T^{-\zeta} j^{-2/\delta}$ branch, and $(1+K_2)^{-2/\delta} \le B^\sharp \le K_2^{-2/\delta}$ at the start of the $j^{-2/\delta}$ branch.

Solving for $B^\sharp$ in its defining equation gives

$$B^\sharp = \left(\frac{a_T^{-\zeta} \sum_{j=k+1}^{K_1} j^{-2/\delta} + \sum_{j=K_2+1}^{\infty} j^{-2/\delta}}{a_T^{-\zeta\eta} \sum_{j=k+1}^{K_1} j^{-2\eta/\delta} + \sum_{j=K_2+1}^{\infty} j^{-2\eta/\delta}} \right)^{\frac{1}{1-\eta}}$$

and using (S.2.3), it follows that $B^\sharp$ satisfies

$$\text{(S.2.4)} \quad \frac{a_T^{-\zeta}\{(k+1)^{1-2/\delta} - K_1^{1-2/\delta}\} + (K_2+1)^{1-2/\delta}}{a_T^{-\zeta\eta}(k^{1-2\eta/\delta} - (K_1+1)^{1-2\eta/\delta}) + K_2^{1-2\eta/\delta}} \le \frac{2-\delta}{2\eta-\delta}(B^\sharp)^{1-\eta} \le \frac{a_T^{-\zeta} k^{1-2/\delta} + K_2^{1-2/\delta}}{a_T^{-\zeta\eta}\{(k+1)^{1-2\eta/\delta} - K_1^{1-2\eta/\delta}\} + (K_2+1)^{1-2\eta/\delta}}$$

We examine the implications of this inequality and the switching conditions for subsequential limits. Two cases are possible.

1. If two switches occur, then $K_2/K_1 \sim a_T^{\delta\zeta/2}$ from the switching conditions. Let $a_T^{-\zeta}(k/K_2)^{1-2/\delta} \to c$ as $T \to \infty$, possibly along a subsequence; we are looking for the values of c that can occur. Passing to the (subsequential) limit in (S.2.4), it follows that

$$c + 1 \le \frac{2-\delta}{2\eta-\delta}(\lim K_2^{2/\delta} B^\sharp)^{1-\eta} \le c+1,$$

and since $K_2^{2/\delta} B^\sharp \to 1$, the unique subsequential limit is $c = 2\,(1-\eta)\,/(2\eta-\delta)$, and thus, it is the limit of $a_T^{-\zeta}(k/K_2)^{1-2/\delta}$ as $T \to \infty$. Further, since $a_T^2 K_1^{2/\delta} B_\sharp \to 1$, we find that $K_1 \sim c^{\frac{\delta}{2-\delta}} a_T^{\frac{\zeta\delta^2}{2(2-\delta)}} k$ and $K_2 \sim c^{\frac{\delta}{2-\delta}} a_T^{\frac{\zeta\delta}{2-\delta}} k$. Then

$$\frac{(\sum_{j=k+1}^{\infty} |\beta_j|^\eta)^{1/\eta}}{\sum_{j=k+1}^{\infty} |\beta_j|} = \frac{(a_T^{-\zeta\eta} \sum_{j=k+1}^{K_1} j^{-2\eta/\delta} + (K_2-K_1) B^{\sharp\eta} + \sum_{j=K_2+1}^{\infty} j^{-2\eta/\delta})^{1/\eta}}{a_T^{-\zeta} \sum_{j=k+1}^{K_1} j^{-2/\delta} + (K_2-K_1) B^\sharp + \sum_{j=K_2+1}^{\infty} j^{-2/\delta}}$$

$$\sim \frac{(c_1 a_T^{-\zeta\eta} k^{1-2\eta/\delta} + c_2 a_T^{-\frac{\zeta(2\eta-\delta)}{2-\delta}} k^{1-2\eta/\delta})^{1/\eta}}{c_3 a_T^{-\zeta} k^{1-2/\delta}} \sim c_4 (a_T^{\frac{\zeta\delta}{2-\delta}} k)^{\frac{1-\eta}{\eta}}$$

for some positive $c_{1,2,3,4}$, and

$$\tilde{a}_T a_T^{-2} \frac{(k\sum_{j=k+1}^{\infty}|\beta_j|^{\eta})^{1/\eta}}{\sum_{j=k+1}^{\infty}|\beta_j|} \sim c_4 \tilde{a}_T a_T^{-2} k^{1/\eta}(a_T^{\frac{\zeta\delta}{2-\delta}}k)^{\frac{1-\eta}{\eta}} = o(1)$$

for $k^2/T \to 0$, η in a sufficiently small left neighbourhood of $\alpha \wedge 1$ and $\zeta > 0$ sufficiently small. The same expression is $o(a_T^{-1/5})$ for $k^3/T \to 0$ and $\zeta = \frac{1}{4}$.

2. Alternatively, if the value of $a_T^{-\zeta} j^{-2/\delta}$ is not taken by any $|\beta_j^{\sharp}|$, we are left with

$$\frac{(K_2+1)^{1-2/\delta}}{K_2^{1-2\eta/\delta}} \leq \frac{2-\delta}{2\eta-\delta}(B^{\sharp})^{1-\eta} \leq \frac{K_2^{1-2/\delta}}{(K_2+1)^{1-2\eta/\delta}}$$

or equivalently,

$$\left(\tfrac{2\eta-\delta}{2-\delta}\right)^{\frac{1}{1-\eta}} \left(1+\tfrac{1}{K_2}\right)^{\frac{\delta-2}{\delta(1-\eta)}} K_2^{-\frac{2}{\delta}} \leq B^{\sharp} \leq \left(\tfrac{2\eta-\delta}{2-\delta}\right)^{\frac{1}{1-\eta}} \left(1+\tfrac{1}{K_2}\right)^{\frac{2-\delta}{\delta(1-\eta)}} (K_2+1)^{-\frac{2}{\delta}}.$$

As $(K_2+1)^{-2/\delta} \leq B^{\sharp} \leq K_2^{-2/\delta}$ should also hold, it follows that

$$\frac{2\eta-\delta}{2-\delta}\left(1+\frac{1}{K_2}\right)^{\frac{\delta-2}{\delta}} \leq 1 \leq \frac{2\eta-\delta}{2-\delta}\left(1+\frac{1}{K_2}\right)^{\frac{2-\delta}{\delta}},$$

which is equivalent to $K_2 \leq \left[\{(2-\delta)/(2\eta-\delta)\}^{\frac{\delta}{2-\delta}} - 1\right]^{-1}$ for $\eta \in (\frac{\delta}{2}, 1)$. As this is inconsistent with $K_2 > k \to \infty$ as $T \to \infty$, for large T the maximizing sequence is $|\beta_j^{\sharp}| = j^{-2/\delta}$, $j \geq k+1$. Therefore, $\sum_{j=1}^{\infty} j^{2/\delta}|\beta_j| < \infty$ implies for large T that

$$\frac{\sum_{j=k+1}^{T}|\beta_j|^{\eta}}{(\sum_{j=k+1}^{T}|\beta_j|)^{\eta}} \leq \frac{\sum_{j=k+1}^{T} j^{-2\eta/\delta}}{(\sum_{j=k+1}^{T} j^{-2/\delta})^{\eta}} \leq \frac{\frac{\delta}{2\eta-\delta}k^{1-2\eta/\delta}}{(\frac{\delta}{2-\delta})^{\eta}(k+1)^{\eta-2\eta/\delta}} \leq ck^{1-\eta}$$

for obvious choices of c, so

$$\tilde{a}_T a_T^{-2} \frac{(k\sum_{j=k+1}^{\infty}|\beta_j|^{\eta})^{1/\eta}}{\sum_{j=k+1}^{\infty}|\beta_j|} \leq c^{1/\eta}\tilde{a}_T a_T^{-2} k^{2/\eta-1} = \begin{cases} o(1) & \text{if } k^2/T \to 0 \\ o(a_T^{-1/5}) & \text{if } k^3/T \to 0 \end{cases}$$

for η smaller than but close to $\alpha \wedge 1$.

This proves the lemma. □

S.3. Additional proofs. Together with the matrix norms $\|\cdot\|_2$ and $\|\cdot\|$ employed in the the paper, here we also use the linear space matrix norms $\|\cdot\|_1 := \sup_{\|x\|_1=1}\|(\cdot)x\|_1$ and $\|\cdot\|_\infty := \sup_{\|x\|_\infty=1}\|(\cdot)x\|_\infty$ induced respectively by the 1 and max vector norms.

S.3.1. *Proof of Lemma 2.* Regarding S_{00}^k of part (a), we argue first that $\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2 = O_P(l_T\tilde{a}_T a_T^{-2})\max\{k^\epsilon a_k, k\} = o_P(1)$ when $k^2/T \to 0$. Then $\lambda_{\min}(S_{00}^k\sigma_T^{-2}) \geq \lambda_{\min}(\Sigma_k) - \|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2 = \lambda_{\min}(\Sigma_k) + o_P(1)$ by Weyl's inequality (Seber, 2008, p.117), so $\lambda_{\min}(S_{00}^k\sigma_T^{-2})$ is bounded away from zero in probability and $(S_{00}^k)^{-1}$ exists with probability approaching one. Further we use the fact that

$$\|(S_{00}^k)^{-1}\sigma_T^2 - \Sigma_k^{-1}\|_2 \leq \frac{\|\Sigma_k^{-1}\|_2^2\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2}{1 - \|\Sigma_k^{-1}\|_2\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2}$$

if $\|\Sigma_k^{-1}\|_2\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2 < 1$. The latter inequality holds with probability approaching one since $\|\Sigma_k^{-1}\|_2$ is bounded as $k \to \infty$ and $\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2 = o_P(1)$, so we can conclude that also

$$\|(S_{00}^k)^{-1}\sigma_T^2 - \Sigma_k^{-1}\|_2 = O_P(l_T\tilde{a}_T a_T^{-2})\max\{k^\epsilon a_k, k\}.$$

The proof of part (a) is completed observing that $a_T^{-2}\sigma_T^2$ is bounded away from zero in probability as it converges in distribution to an a.s. positive ($\alpha/2$-stable) r.v.

We present now the evaluation of $\|S_{00}^k\sigma_T^{-2} - \Sigma_k\|_2$. A generic element of S_{00}^k is $\sum_{t=k}^{T-1} X_{t-i}X_{t-j} = c_{ij}^k + \xi_{ij}^{\leq} + \xi_{ij}^{>}$ (for $0 \leq i,j \leq k-1$), where $c_{ij}^k := \sum_{t=k}^{T-1}\sum_{v=0}^{\infty}\varepsilon_{t-\max(i,j)-v}^2\gamma_v\gamma_{v+|j-i|}$ and

$$\xi_{ij}^R := \textstyle\sum_{t=k}^{T-1}\sum_{u,v=0}^{\infty}\mathbb{I}_{\{u\neq v+j-i\}}\gamma_u\gamma_v\varepsilon_{t-i-u}\varepsilon_{t-j-v}\mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|R\tilde{a}_T},$$

$R \in \{>, \leq\}$. With $C^k := (c_{ij}^k)_{i,j=0}^{k-1}$, it holds that

$$\begin{aligned}\|S_{00}^k - C^k\|_2 &\leq \|(\xi_{ij})_{i,j}^{\leq}\|_2 + \|(\xi_{0,|i-j|}^{>})_{i,j}\|_2 + \|(\xi_{0,|i-j|}^{>} - \xi_{ij}^{>})_{i,j}\|_2 \\ &\leq \|(\xi_{ij})_{i,j}^{\leq}\|_2 + \max_{i=0,\ldots,k-1}\sum_{j=0}^{k-1}|\xi_{0,|i-j|}^{>}| + \sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}|\end{aligned}$$

since $\|(\xi_{0,|i-j|}^{>})_{i,j}\|_2 \leq \max_{i=0,\ldots,k-1}\sum_{j=0}^{k-1}|\xi_{0,|i-j|}^{>}|$ as $(\xi_{0,|i-j|}^{>})_{i,j}$ is symmetric (in general, $\|\cdot\|_2 \leq \|\cdot\|_1^{1/2}\|\cdot\|_\infty^{1/2}$). Let first $\mathrm{E}\,|\varepsilon_1| = \infty$ (so $\alpha \in (0,1]$). Since $\|\cdot\|_2 \leq \|\cdot\|$, the inequalities can be continued as

$$\|S_{00}^k - C^k\|_2 \leq (\Xi_1^{\leq} + \Xi_2^{\leq})^{1/2} + 2\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| + \sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}|,$$

where

$$\begin{aligned}\Xi_1^{\leq} &:= \textstyle\sum_{i,j}\sum_{s,t=k}^{T-1}\sum_{a,b,u,v=0}^{\infty}\mathbb{I}_A\gamma_a\gamma_b\gamma_u\gamma_v\varepsilon_{s-i-a}\varepsilon_{s-j-b}\varepsilon_{t-i-u}\varepsilon_{t-j-v} \\ &\quad\times\mathbb{I}_{|\varepsilon_{s-i-a}\varepsilon_{s-j-b}|\leq\tilde{a}_T}\mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|\leq\tilde{a}_T}, \\ \Xi_2^{\leq} &:= \textstyle\sum_{i,j}\sum_{s,t=k}^{T-1}\sum_{a,b,u,v=0}^{\infty}\mathbb{I}_{A^c}\gamma_a\gamma_b\gamma_u\gamma_v\varepsilon_{s-i-a}\varepsilon_{s-j-b}\varepsilon_{t-i-u}\varepsilon_{t-j-v} \\ &\quad\times\mathbb{I}_{|\varepsilon_{s-i-a}\varepsilon_{s-j-b}|\leq\tilde{a}_T}\mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|\leq\tilde{a}_T},\end{aligned}$$

$A := \{\#(\{s-i-a, s-j-b, t-i-u, t-j-v\}) = 4\}$, $A^c := \{\#(\{s-i-a, s-j-b, t-i-u, t-j-v\}) = 2 \text{ or } 3\}$. Further,

$$\mathrm{E}\,|\Xi_1^{\leq}| \leq k^2T^2\,\mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})^2(\textstyle\sum_{u=0}^{\infty}|\gamma_u|)^4 = O(k^2\tilde{a}_T^2)$$

for $\alpha \in (0,1)$, as $\tilde{a}_T^{-1}T\,\mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T}) \to \alpha/(1-\alpha)$ by Karamata's theorem [KT], and $\mathrm{E}\,|\Xi_1^{\leq}| = O(k^2l_T^2\tilde{a}_T^2)$ for $\alpha = 1$, as $\tilde{a}_T^{-1}T\,\mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})$ is slowly varying in this case. Similarly,

$$\mathrm{E}\,|\Xi_2^{\leq}| \leq 4k^2T\,\mathrm{E}(|\varepsilon_1^2\varepsilon_2^2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})(\textstyle\sum_{u=0}^{\infty}|\gamma_u|)^4 = O(k^2\tilde{a}_T^2)$$

by KT, so $\Xi_1^{\leq} + \Xi_2^{\leq} = O_P(k^2l_T^2\tilde{a}_T^2)$. Also, for every $\eta \in (\delta, \alpha)$,

$$\mathrm{E}\,|\sum_{j=0}^{k-1}|\xi_{0,j}^{>}||^{\eta} \leq kT\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^{\eta}\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde{a}_T})(\sum_{u=0}^{\infty}|\gamma_u|^{\eta})^2 = O(k\tilde{a}_T^{\eta}) \tag{S.3.1}$$

by KT with $\mathrm{E}(|\varepsilon_1\varepsilon_2|^{\eta}) < \infty$, so $\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| = O_P(k^{1/\eta}\tilde{a}_T)$ by Markov's inequality, and by letting $\eta \uparrow \alpha$, $\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| = o_P(k^{\epsilon}a_k\tilde{a}_T)$ for every $\epsilon > 0$. Similarly, since $|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}|$ does not exceed

$$(\sum_{t=k-i\wedge j}^{k-1} + \sum_{t=T-i\wedge j}^{T-1})\sum_{u,v=0}^{\infty}\mathbb{I}_{\{u\neq v+|j-i|\}}|\gamma_u||\gamma_v||\varepsilon_{t-u}\varepsilon_{t-|j-i|-v}|\mathbb{I}_{|\varepsilon_{t-u}\varepsilon_{t-|j-i|-v}|>\tilde{a}_T},$$

with $i \wedge j := \min(i,j)$, it follows that

$$\mathrm{E}\,|\sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}||^{\eta} \leq (k+1)^3\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^{\eta}\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde{a}_T})(\sum_{u=0}^{\infty}|\gamma_u|^{\eta})^2 \tag{S.3.2}$$

is $O((k^3/T)\tilde{a}_T^{\eta}) = o(k\tilde{a}_T^{\eta})$, so $\sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}| = o_P(k^{\epsilon}a_k\tilde{a}_T)$ for every $\epsilon > 0$. By combining these results, also $\|S_{00}^k - C^k\|_2 = o_P(k^{\epsilon}a_kl_T\tilde{a}_T)$.

Instead, for $\mathrm{E}\,|\varepsilon_1| < \infty$ (so $\alpha \in [1,2)$), we write $\|(\xi_{ij})_{i,j}^{\leq}\|_2 \leq \sqrt{2}(\Xi_3^{\leq} + \Xi_4^{\leq})^{1/2}$ with

$$\Xi_3^{\leq} := \textstyle\sum_{i,j}\{\sum_{t=k}^{T-1}\sum_{u,v=0}^{\infty}\mathbb{I}_{\{u\neq v+j-i\}}\gamma_u\gamma_v(\varepsilon_{t-i-u}\varepsilon_{t-j-v}\mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|\leq\tilde{a}_T} - \mu_T)\}^2,$$
$$\Xi_4^{\leq} := \mu_T^2\textstyle\sum_{i,j}\{\sum_{t=k}^{T-1}\sum_{u,v=0}^{\infty}\mathbb{I}_{\{u\neq v+j-i\}}\gamma_u\gamma_v\}^2$$

and $\mu_T := \mathrm{E}(\varepsilon_1\varepsilon_2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})$, so

$$\|S_{00}^k - C^k\|_2 \leq \sqrt{2}(\Xi_3^{\leq} + \Xi_4^{\leq})^{1/2} + 2\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| + \sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}|.$$

The terms in the upper bound satisfy: (i) $\Xi_3^{\leq} = O_P(k^2\tilde{a}_T^2)$, as

$$\mathrm{E}\,|\Xi_3^{\leq}| \leq 4k^2T\,\mathrm{E}(\varepsilon_1\varepsilon_2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T} - \mu_T)^2(\sum|\gamma_u|)^4 \leq O(k^2T)\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})$$

is $O(k^2\tilde{a}_T^2)$ by independence and KT; (ii) $|\Xi_4^{\leq}| \leq \mu_T^2k^2T^2(\sum|\gamma_u|)^4 = O(k^2\tilde{a}_T^2)$ for $\alpha \in (1,2)$ since

$$\text{(S.3.3)}\qquad \begin{aligned} |\mu_T| &= \left|\mathrm{E}(\varepsilon_1\varepsilon_2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})\right| = |-\mathrm{E}(\varepsilon_1\varepsilon_2\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde{a}_T})| \\ &\leq \mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde{a}_T}) = O(T^{-1}\tilde{a}_T) \end{aligned}$$

using $\mathrm{E}\,(\varepsilon_1\varepsilon_2) = 0$ and KT, whereas $|\Xi_4^{\leq}| = O_P(k^2T^2) = O_P(k^2l_T^2a_T^2)$ for $\alpha = 1$ as $\mu_T = O\,(1)$; (iii) $2\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| + \sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}| = O_P(k\tilde{a}_T)$ for $\alpha \in (1,2)$ by (S.3.1)-(S.3.2) with $\eta = 1$, and $2\sum_{j=0}^{k-1}|\xi_{0,j}^{>}| + \sum_{i,j=0}^{k-1}|\xi_{0,|i-j|}^{>} - \xi_{ij}^{>}| = O_P(kT) = O_P(kl_T\tilde{a}_T)$ for $\alpha = 1$ using the same displays. Thus, $\|S_{00}^k - C^k\|_2 = O_P\,(kl_T\tilde{a}_T)$ in the case $\mathrm{E}\,|\varepsilon_1| < \infty$, and by the earlier argument, $\|S_{00}^k - C^k\|_2 = O_P(l_T\tilde{a}_T)\max\{k^\epsilon a_k, k\}$ for all $\alpha \in (0,2)$ and $\epsilon > 0$.

In its turn,

$$\text{(S.3.4)}\quad c_{ij}^k = r_{|i-j|}\sum_{t=k}^{T-1}\varepsilon_{t-\max(i,j)}^2 + \sum_{v=0}^{\infty}\rho_v^{ij}[\varepsilon_{k-\max(i,j)-v}^2 - \varepsilon_{T-\max(i,j)-1-v}^2],$$

where $\rho_v^{ij} := \sum_{u=v+1}^{\infty}\gamma_u\gamma_{u+|j-i|}$ has $|\rho_v^{ij}| \leq \sum_{u=v+1}^{\infty}\gamma_u^2 := \tilde{\gamma}_v^2$. For δ of Assumption 1(b) it follows that $\sum_{v=0}^{\infty}|\rho_v^{ij}|^{\delta/2} \leq \sum_{v=0}^{\infty}\sum_{u=v+1}^{\infty}|\gamma_u|^\delta = \sum_{u=1}^{\infty}u|\gamma_u|^\delta < \infty$, so the series in (S.3.4) are a.s. convergent because ε_t^2 has tail index $\alpha/2$. Further, as $r_{|i-j|}^2 \leq r_0^2$,

$$\text{(S.3.5)}\qquad \begin{aligned} &\|\Sigma_k\sigma_T^2 - C^k\|_2^2 \leq \|\Sigma_k\sigma_T^2 - C^k\|^2 \\ \leq\;& 3(\sum_{t=1}^{k}\varepsilon_t^2 + \sum_{t=T-k}^{T}\varepsilon_t^2)^2\sum_{i,j}r_{ij}^2 + 3\sum_{i,j}(\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{T-\max(i,j)-1-v}^2)^2 \\ &+3\sum_{i,j}(\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{k-\max(i,j)-v}^2)^2 \leq 3k^2[O_P(a_k^4)r_{00}^2 \\ &+ \max_{i=1,\dots,k}(\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{T-i-v}^2)^2 + \max_{i=1,\dots,k}(\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{k-i-v}^2)^2] = O_P(k^2a_k^4) \end{aligned}$$

using that $\max_{i=1,\dots,k}|a_k^{-2}\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{T-i-v}^2|$, $\max_{i=1,\dots,k}|a_k^{-2}\sum_{v=0}^{\infty}\tilde{\gamma}_v^2\varepsilon_{k-i-v}^2|$, $a_k^{-2}\sum_{t=1}^{k}\varepsilon_t^2$ and $a_k^{-2}\sum_{t=T-k}^{T}\varepsilon_t^2$ converge weakly to a.s. finite r.v.'s (see Theorem 3.2 of Davis and Resnick (1985a) for the former two, as ε_t^2 are in the

$\alpha/2$-stable domain of attraction with normalisation a_T^2). From the triangle inequality and the condition $k^2/T \to 0$, we conclude that, for every $\epsilon > 0$,

$$\|S_{00}^k - \Sigma_k \sigma_T^2\|_2 = O_P(l_T \tilde{a}_T) \max\{k^\epsilon a_k, k\}. \tag{S.3.6}$$

Regarding $S_{0\varepsilon}^k$ in part (b), first,

$$\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \varepsilon_t\|^2 = \sum_{i=1}^k (\sum_{t=k+1}^T \sum_{j=0}^\infty \gamma_j \varepsilon_{t-j-i} \varepsilon_t)^2.$$

For $\{\varepsilon_t\}$ with $\mathrm{E}\,|\varepsilon_1| = \infty$ (and hence, $\alpha \le 1$), we write $\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \varepsilon_t\|^2 \le 2\{(\Lambda_1^>)^2 + \Lambda_1 + \Lambda_2\}$ with

$$\begin{aligned}
\Lambda_1^> &:= \textstyle\sum_{i=1}^k \sum_{t=k+1}^T \sum_{j=0}^\infty |\gamma_j||\varepsilon_{t-j-i}\varepsilon_t| \mathbb{I}_{|\varepsilon_{t-j-i}\varepsilon_t| > \tilde{a}_T} \\
\Lambda_1 &:= \textstyle\sum_{i=1}^k \sum_{s,t=k+1}^T \sum_{h,j=0}^\infty \mathbb{I}_B \gamma_h \gamma_j \varepsilon_{s-h-i}\varepsilon_s \varepsilon_{t-j-i}\varepsilon_t \mathbb{I}_{|\varepsilon_{s-h-i}\varepsilon_s| \le \tilde{a}_T} \mathbb{I}_{|\varepsilon_{t-j-i}\varepsilon_t| \le \tilde{a}_T}, \\
\Lambda_2 &:= \textstyle\sum_{i=1}^k \sum_{s,t=k+1}^T \sum_{h,j=0}^\infty \mathbb{I}_{B^c} \gamma_h \gamma_j \varepsilon_{s-h-i}\varepsilon_s \varepsilon_{t-j-i}\varepsilon_t \mathbb{I}_{|\varepsilon_{s-h-i}\varepsilon_s| \le \tilde{a}_T} \mathbb{I}_{|\varepsilon_{t-j-i}\varepsilon_t| \le \tilde{a}_T},
\end{aligned}$$

$B := \{\#(\{s, t, s-h-i, t-j-i\}) = 4\}$, $B^c := \{\#(\{s, t, s-h-i, t-j-i\}) = 2$ or $3\}$. Similarly to the evaluations of $(\Xi^>)^2, \Xi_{1,2}$, we find for $\eta \in (\delta, \alpha)$ that $\mathrm{E}\,|\Lambda_1^>|^\eta = O(k\tilde{a}_T^\eta)$, $\mathrm{E}\,|\Lambda_1| = O(kl_T^2\tilde{a}_T^2)$ and $\mathrm{E}\,|\Lambda_2| = O(k\tilde{a}_T^2)$, so $\Lambda_1^> = o_P(k^\epsilon a_k \tilde{a}_T)$ for every $\epsilon > 0$, $\Lambda_1 = O_P(kl_T^2\tilde{a}_T^2)$ for l_T as in part (a), and $\Lambda_2 = O_P(k\tilde{a}_T^2)$, giving $\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \varepsilon_t\| = o_P(k^\epsilon a_k l_T \tilde{a}_T)$ for every $\epsilon > 0$. On the other hand, in the case $\mathrm{E}\,|\varepsilon_1| < \infty$ (where $\alpha \ge 1$), $\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \varepsilon_t\|^2 \le 3(\Lambda_2^> + \Lambda_3 + \Lambda_4)$ with

$$\Lambda_2^> := \textstyle\sum_{i=1}^k (\sum_{t=k+1}^T \sum_{j=0}^\infty |\gamma_j||\varepsilon_{t-j-i}\varepsilon_t| \mathbb{I}_{|\varepsilon_{t-j-i}\varepsilon_t| > \tilde{a}_T})^2$$

satisfying by KT $\mathrm{E}\,|\Lambda_2^>|^{\eta/2} \le kT\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^\eta \mathbb{I}_{|\varepsilon_1\varepsilon_2| > \tilde{a}_T})(\sum |\gamma_u|^\eta) = O(k\tilde{a}_T^\eta)$ for $\eta \in [1, \alpha)$, $\alpha > 1$, whereas $\mathrm{E}\,|\Lambda_2^>|^{1/2} = O(kT) = O(kl_T\tilde{a}_T)$ for $\alpha = 1$, so $\Lambda_2^> = O_P(k^\epsilon a_k^2 l_T^2 \tilde{a}_T^2)$ by Markov's inequality, and

$$\begin{aligned}
\Lambda_3 &:= \textstyle\sum_{i=1}^k \{\sum_{t=k+1}^T \sum_{j=0}^\infty \gamma_j (\varepsilon_{t-j-i}\varepsilon_t \mathbb{I}_{|\varepsilon_{t-j-i}\varepsilon_t| \le \tilde{a}_T} - \mu_T)\}^2 = O_P(k\tilde{a}_T^2), \\
\Lambda_4 &:= \mu_T^2 \textstyle\sum_{i=1}^k (\sum_{t=k+1}^T \sum_{j=0}^\infty \gamma_j)^2 = O(kl_T^2\tilde{a}_T^2)
\end{aligned}$$

as $\Xi_{3,4}$ earlier. Thus, $\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \varepsilon_t\| = O_P(k^\epsilon a_k l_T \tilde{a}_T)$ for every $\epsilon > 0$ in the case $\mathrm{E}\,|\varepsilon_1| < \infty$, and by the previous argument, for all $\alpha \in (0, 2)$.

Second,

$$\begin{aligned}
\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k \rho_{t,k}\| &= \{\sum_{i=1}^k (\sum_{j=k+1}^\infty \beta_j \sum_{t=k+1}^T X_{t-i}X_{t-j})^2\}^{1/2} \\
&\le \sqrt{2} \sum_{i=1}^k \sum_{j=k+1}^\infty |\beta_j||c_{ij}^k| + \sqrt{2}\{\sum_{i=1}^k (\sum_{j=k+1}^\infty \beta_j \xi_{ij})^2\}^{1/2}
\end{aligned}$$

where $c_{ij}^k := \sum_{t=k}^{T-1} \sum_{v=0}^{\infty} \varepsilon_{t-j-v}^2 \gamma_v \gamma_{v+j-i}$ is

$$c_{ij}^k = r_{j-i} \sum_{t=k}^{T-1} \varepsilon_{t-j}^2 - \sum_{v=0}^{\infty} \rho_v^{ij} \varepsilon_{T-j-v-1}^2 + \sum_{v=0}^{\infty} \rho_v^{ij} \varepsilon_{k-j-v}^2$$

and $\xi_{ij} := \sum_{t=k}^{T-1} \sum_{u=0}^{\infty} \sum_{v=0}^{\infty} \mathbb{I}_{\{u \neq v+j-i\}} \gamma_u \gamma_v \varepsilon_{t-i-u} \varepsilon_{t-j-v}$; cf. (S.3.4) with $i < j$. We find that (i)

$$\sum_{i=1}^{k} \sum_{j=k+1}^{\infty} |\beta_j||r_{j-i}| \sum_{t=k}^{T-1} \varepsilon_{t-j}^2 \leq \{\sum_{j=k+1}^{\infty} |\beta_j|\}\{\sum_{t=k}^{T-1} \sum_{j=k+1}^{\infty} (\sum_{i=1}^{k} |r_{j-i}|)\varepsilon_{t-j}^2\},$$

where $\sum_{t=k}^{T-1} \sum_{j=k+1}^{\infty} (\sum_{i=1}^{k} |r_{j-i}|)\varepsilon_{t-j}^2$ is distributed like

$$\begin{aligned}\sum_{t=0}^{T-k-1} \sum_{j=1}^{\infty} (\sum_{i=j}^{j+k-1} |r_i|)\varepsilon_{t-j}^2 \ &\leq \ T\,\mathrm{E}[\varepsilon_1^2 \mathbb{I}_{|\varepsilon_1| \leq a_T}] \sum_{j=1}^{\infty} \sum_{i=j}^{\infty} |r_i| \\ &+ \sum_{t=0}^{T-1} \sum_{j=1}^{\infty} (\sum_{i=j}^{\infty} |r_i|)(\varepsilon_{t-j}^2 - \mathrm{E}[\varepsilon_{t-j}^2 \mathbb{I}_{|\varepsilon_{t-j}| \leq a_T}]) = O_P(a_T^2)\end{aligned}$$

since ε_t^2 have tail index $\alpha/2$ and

$$\sum_{j=1}^{\infty} (\sum_{i=j}^{\infty} |r_i|)^{\delta/2} \leq \sum_{s=0}^{\infty} |\gamma_s|^{\delta/2} \sum_{j=1}^{\infty} j|\gamma_{s+j}|^{\delta/2} \leq \sum_{s=0}^{\infty} |\gamma_s|^{\delta/2} \sum_{s=0}^{\infty} s|\gamma_s|^{\delta/2} < \infty$$

by Assumption 1(b), so Theorem 4.1 of Davis and Resnick (1985a) applies (with their $c_j = \sum_{i=j}^{\infty} |r_i|$) jointly with KT; (ii)

$$\sum_{i=1}^{k} \sum_{j=k+1}^{\infty} |\beta_j| \sum_{v=0}^{\infty} |\rho_v^{ij}| \varepsilon_{T-j-v-1}^2 \leq \{\sum_{j=k+1}^{\infty} |\beta_j|\}\{\sum_{i=1}^{k} \sum_{j=k+1}^{\infty} \sum_{v=0}^{\infty} |\rho_v^{ij}| \varepsilon_{T-j-v-1}^2\}$$

with

$$\begin{aligned}\mathrm{E}(\sum_{i=1}^{k} \sum_{j=k+1}^{\infty} \sum_{v=0}^{\infty} |\rho_v^{ij}| \varepsilon_{T-j-v-1}^2 \mathbb{I}_{|\varepsilon_{T-j-v-1}| \leq a_T}) &\leq \mathrm{E}(\varepsilon_1^2 \mathbb{I}_{|\varepsilon_1| \leq a_T}) \sum_{i=1}^{k} \sum_{j=k+1}^{\infty} \sum_{v=0}^{\infty} |\rho_v^{ij}| \\ &\leq O(T^{-1} a_T^2) \sum_{i=1}^{k} \sum_{j=k+1}^{\infty} \sum_{v=0}^{\infty} \sum_{u=v+1}^{\infty} |\gamma_u||\gamma_{u+|i-j|}| \leq O(T^{-1} a_T^2)(\sum_{u=1}^{\infty} u|\gamma_u|)^2\end{aligned}$$

and similarly, for $\eta \in (\delta, \alpha)$,

$$\mathrm{E}(\sum_{i=1}^{k} \sum_{j=k+1}^{\infty} \sum_{v=0}^{\infty} |\rho_v^{ij}| \varepsilon_{T-j-v-1}^2 \mathbb{I}_{|\varepsilon_{T-j-v-1}| > a_T})^{\eta/2} \leq \mathrm{E}(|\varepsilon_1|^{\eta} \mathbb{I}_{|\varepsilon_1| \leq a_T})(\sum_{u=1}^{\infty} u|\gamma_u|^{\eta/2})^2$$

is $O(T^{-1}a_T^\eta)$ by KT, so $\sum_{i=1}^k\sum_{j=k+1}^\infty\sum_{v=0}^\infty|\rho_v^{ij}|\varepsilon_{T-j-v-1}^2=o_P(a_T^2)$; (iii),

$$\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j|\sum_{v=0}^\infty|\rho_v^{ij}|\varepsilon_{k-j-v}^2=o_P(a_T^2)\sum_{j=k+1}^\infty|\beta_j|$$

likewise. Thus, $\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j||c_{ij}^k| = O_P(a_T^2)\sum_{j=k+1}^\infty|\beta_j|$ by combining the previous estimates.

Further, we split $\xi_{ij}=\xi_{ij}^{\le}+\xi_{ij}^{>}$ as in the proof of part (a):

$$\xi_{ij}^R:=\sum_{t=k}^{T-1}\sum_{u,v=0}^\infty\mathbb{I}_{\{u\neq v+j-i\}}\gamma_u\gamma_v\varepsilon_{t-i-u}\varepsilon_{t-j-v}\mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|R\tilde a_T},\ \ R\in\{\le,>\},$$

and for $\{\varepsilon_t\}$ with $\mathrm{E}\,|\varepsilon_1|=\infty$ we find that,

$$\mathrm{E}(\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j||\xi_{ij}^{\le}|)\le Tk\,\mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|\le\tilde a_T})(\sum_{u=0}^\infty|\gamma_u|)^2\sum_{j=k+1}^\infty|\beta_j|$$

is $o(a_T^2)\sum_{j=k+1}^\infty|\beta_j|$ by KT, and similarly, for $\eta\in(\delta,\alpha)$,

$$\mathrm{E}(\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j||\xi_{ij}^{>}|)^\eta\le Tk\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^\eta\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde a_T})(\sum_{u=0}^\infty|\gamma_u|^\eta)^2\sum_{j=k+1}^\infty|\beta_j|^\eta$$

is $O(k\tilde a_T^\eta)\sum_{j=k+1}^\infty|\beta_j|^\eta$, so by using (S.2.1), $\{\sum_{i=1}^k(\sum_{j=k+1}^\infty\beta_j\xi_{ij})^2\}^{1/2}\le\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j||\xi_{ij}|=o_P(a_T^{1-\zeta})+O_P(a_T^2)\sum_{j=k+1}^\infty|\beta_j|$ for $\zeta>0$ sufficiently small, and $\|\sum_{t=k+1}^T\mathbf{X}_{t-1}^k\rho_{t,k}\|=o_P(a_T^{1-\zeta})+O_P(a_T^2)\sum_{j=k+1}^\infty|\beta_j|$ in this case. In the case $\mathrm{E}\,|\varepsilon_1|<\infty$, as in the proof of part (a), we find that, (i),

$$\begin{aligned}\text{(S.3.7)}\qquad \mathrm{E}(\sum_{i=1}^k(\sum_{j=k+1}^\infty\beta_j\xi_{ij}^{\le})^2)&\le[\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j|\,\mathrm{E}\{(\xi_{ij}^{\le})^2\}](\sum_{j=k+1}^\infty|\beta_j|)\\ &\le k\left[4T\,\mathrm{E}(\varepsilon_1^2\varepsilon_2^2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\le\tilde a_T})+T^2\mu_T^2\right](\sum_{u=0}^\infty|\gamma_u|)^4(\sum_{j=k+1}^\infty|\beta_j|)^2\end{aligned}$$

is $O(kl_T^2\tilde a_T^2)(\sum_{j=k+1}^\infty|\beta_j|)^2$, the inequalities respectively from Cauchy-Schwartz and by separating products where some ε is squared from those where all ε's are distinct, and the magnitude order from KT and (S.3.3), and (ii),

$$\mathrm{E}(\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j||\xi_{ij}^{>}|)\le kT\,\mathrm{E}(|\varepsilon_1\varepsilon_2|\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde a_T})(\sum_{u=0}^\infty|\gamma_u|)^2\sum_{j=k+1}^\infty|\beta_j|$$

is $O\left(kl_T\tilde{a}_T\right)\sum_{j=k+1}^{\infty}|\beta_j|$ by KT. Thus,

$$\{\sum_{i=1}^{k}(\sum_{j=k+1}^{\infty}\beta_j\xi_{ij})^2\}^{1/2}\leq\sqrt{2}\{\sum_{i=1}^{k}(\sum_{j=k+1}^{\infty}\beta_j\xi_{ij}^{\leq})^2\}^{1/2}+\sqrt{2}\sum_{i=1}^{k}\sum_{j=k+1}^{\infty}|\beta_j||\xi_{ij}^{>}|$$

is $O_P\left(kl_T\tilde{a}_T\right)\sum_{j=k+1}^{\infty}|\beta_j|$ with $kl_T\tilde{a}_T=o(a_T^2)$ when $k^2/T\to 0$. Finally, $\|\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^{k}\rho_{t,k}\|=O_P(a_T^2)\sum_{j=k+1}^{\infty}|\beta_j|$ when $\mathrm{E}\,|\varepsilon_1|<\infty$.

The magnitude order of $S_{0\varepsilon}^{k}$ is obtained by combining the magnitude orders of $\|\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^{k}\varepsilon_t\|$ and $\|\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^{k}\rho_{t,k}\|$. □

S.3.2. *Proof of Corollary 4.* Using the fact that $\hat{\Gamma}_k^{-1}$ is the lower triangular Toeplitz matrix with first column $(1:-\hat{\boldsymbol{\beta}}_{k-1}')'$, it can be checked directly that $\hat{\Gamma}_k^{-1}\boldsymbol{\gamma}_k=(I_k-\boldsymbol{\Gamma}_k)\hat{\boldsymbol{\beta}}_k+\boldsymbol{\gamma}_k$, from where

$$\begin{aligned}\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k &= \hat{\Gamma}_k\hat{\boldsymbol{\beta}}_k-\boldsymbol{\gamma}_k=\hat{\Gamma}_k(\hat{\boldsymbol{\beta}}_k-\hat{\Gamma}_k^{-1}\boldsymbol{\gamma}_k)=\hat{\Gamma}_k(\boldsymbol{\Gamma}_k\hat{\boldsymbol{\beta}}_k-\boldsymbol{\gamma}_k)\\ &= \hat{\Gamma}_k(\boldsymbol{\Gamma}_k\hat{\boldsymbol{\beta}}_k-\boldsymbol{\Gamma}_k\boldsymbol{\beta}_k)=\hat{\Gamma}_k\Gamma_k(\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k).\end{aligned}$$

Hence, $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1\leq\|\hat{\Gamma}_k\|_1\|\Gamma_k\|_1\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1$, where $\|\cdot\|_1$ equals the maximum absolute column sum. As $\|\Gamma_k\|_1\leq\|\boldsymbol{\gamma}\|_1:=\sum_{i=0}^{\infty}|\gamma_i|<\infty$, and thus,

$$\|\hat{\Gamma}_k\|_1\leq\|\Gamma_k\|_1+\|\hat{\Gamma}_k-\Gamma_k\|_1\leq\|\boldsymbol{\gamma}\|_1+\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1,$$

it holds further that $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1\leq(\|\boldsymbol{\gamma}\|_1+\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1)\|\boldsymbol{\gamma}\|_1\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1$ and, for small $\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1$, $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1\leq\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1\|\boldsymbol{\gamma}\|_1^2/(1-\|\boldsymbol{\gamma}\|_1\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1)$. Hence, $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1=O_P(\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1)=O_P(k^{1/2}\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|)=o_P(1)$ by (7.1) with $k^2/T\to 0$ and $k^{1/2}\sum_{j=k+1}^{\infty}|\beta_j|\leq\sum_{j=k+1}^{\infty}j|\beta_j|=o\left(1\right)$.

Similarly, $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|\leq\|\hat{\Gamma}_k\|_2\|\Gamma_k\|_2\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|$, with

$$\begin{aligned}\|\hat{\Gamma}_k\|_2 &\leq \|\Gamma_k\|_2+\|\hat{\Gamma}_k-\Gamma_k\|_2\leq\|\Gamma_k\|_2+\|\hat{\Gamma}_k-\Gamma_k\|_1^{1/2}\|\hat{\Gamma}_k-\Gamma_k\|_\infty^{1/2}\\ &\leq \|\Gamma_k\|_2+\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1,\end{aligned}$$

so $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|\leq(\|\Gamma_k\|_2+\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1)\|\Gamma_k\|_2\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|$. Since $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|_1=o_P(1)$ and $\|\Gamma_k\|_2\leq\|\Gamma_k\|_1^{1/2}\|\Gamma_k\|_\infty^{1/2}\leq\|\boldsymbol{\gamma}\|_1<\infty$, it follows that $\|\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_k\|=O_P(\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|)=o_P(1)$. □

S.3.3. *Proof of Corollary 5.* It holds that $|\hat{C}_T\left(\lambda\right)-C\left(\lambda\right)|\leq R_T\left(\lambda\right)+I_T\left(\lambda\right)$ with

$$\begin{aligned}R_T\left(\lambda\right) &:= \left||1+\sum_{j=1}^{k}\hat{\gamma}_j\cos\left(\lambda j\right)|^2-|1+\sum_{j=1}^{\infty}\gamma_j\cos\left(\lambda j\right)|^2\right|\\ &\leq \Big(2+\sum_{j=1}^{\infty}|\gamma_j|+\sum_{j=1}^{k}|\hat{\gamma}_j|\Big)\Big(\sum_{j=1}^{k}|\hat{\gamma}_j-\gamma_j|+\sum_{j=k+1}^{\infty}|\gamma_j|\Big)\end{aligned}$$

$$\leq \Big(2 + 2\sum_{j=1}^{\infty}|\gamma_j| + \|\hat{\boldsymbol{\gamma}}_k - \boldsymbol{\gamma}_k\|_1\Big)\Big(\|\hat{\boldsymbol{\gamma}}_k - \boldsymbol{\gamma}_k\|_1 + \sum_{j=k+1}^{\infty}|\gamma_j|\Big) = o_P(1)$$

since $\|\hat{\gamma}_k - \boldsymbol{\gamma}_k\|_1 = o_P(1)$ by the proof of Corollary 4 and $\sum_{j=1}^{\infty}|\gamma_j| < \infty$. Similarly, $I_T(\lambda) := \Big||\sum_{j=1}^{\infty}\gamma_j \sin(\lambda j)|^2 - |\sum_{j=1}^{k}\hat{\gamma}_j \sin(\lambda j)|^2\Big| = o_P(1)$ using the same upper bounds. As these bounds are independent of λ, convergence is uniform in λ. □

S.4. Proofs from Section 7.3.1 of CGT. Similarly to Lemma 8.3 of Kreiss (1997), the following bounds can be established for $\hat{\gamma}_j$.

LEMMA S.2. *There exist constants $b_{jk} \geq 0$ and C such that, for large k and uniformly in $j \in \mathbb{N}$, it holds that*

$$|\hat{\gamma}_j - \gamma_j| \leq \left(1 + \frac{1}{k}\right)^{-j} O_P(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|_1 + \sum_{j=k+1}^{\infty}|\beta_j|) + b_{jk} \tag{S.4.1}$$

and $\sum_{j=0}^{\infty} b_{jk} \leq C\sum_{j=k+1}^{\infty}|\beta_j|$.

S.4.1. *Proof of Lemma 3.*

S.4.1.1. *Preparation.* From $\hat{\varepsilon}_t - \varepsilon_t = (\boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_k)'\mathbf{X}_{t-1}^k + \rho_{t,k}$ it follows that

$$\sum_{t=k+1}^{T}(\hat{\varepsilon}_t - \varepsilon_t) = \sum_{t=k+1}^{T}(\mathbf{X}_{t-1}^k)'(\boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_k) + \sum_{t=k+1}^{T}\rho_{t,k}, \tag{S.4.2}$$

where, (i), $\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^k = O_P(k^{1/2}a_T)$ is implied by the Beveridge-Nelson decomposition of X_t ($X_t = \gamma(1)\varepsilon_t - \Delta Z_t$, $Z_t := \sum_{j=0}^{\infty}\varepsilon_{t-j}\sum_{i=j+1}^{\infty}\gamma_i$), which yields

$$\begin{aligned}\|\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^k\| &\leq k^{1/2}|\gamma(1)||\sum_{t=k}^{T-1}\varepsilon_t| + k\max_{t\in\{0,\ldots,k-1\}\cup\{T-k,\ldots,T-1\}}(|\varepsilon_t| + |Z_t|)\\ &= O_P(k^{1/2}a_T l_T + k a_k) = O_P(k^{1/2}a_T l_T)\end{aligned}$$

as $a_T^{-1}\sum_{t=k}^{T-1}\varepsilon_t - a_T^{-1}T\,\mathrm{E}(\varepsilon_1\mathbb{I}_{|\varepsilon_1|\leq a_T}) = O_P(1)$, $a_T^{-1}T\,\mathrm{E}(\varepsilon_1\mathbb{I}_{|\varepsilon_1|\leq a_T}) = l_T$, and (ii), $\sum_{t=k+1}^{T}\rho_{t,k} = o_P(l_T)$ by Markov's inequality. Indeed, for $\mathrm{E}|\varepsilon_1| < \infty$,

$$\begin{aligned}(\sum_{t=k+1}^{T}\rho_{t,k})^2 &\leq 2(\sum_{t=k+1}^{T}\sum_{i=k+1}^{\infty}\beta_i\sum_{j=0}^{\infty}\gamma_j\varepsilon_{t-i-j}\mathbb{I}_{|\varepsilon_{t-i-j}|\leq a_T})^2\\ &\quad+2(\sum_{t=k+1}^{T}\sum_{i=k+1}^{\infty}\beta_i\sum_{j=0}^{\infty}\gamma_j\varepsilon_{t-i-j}\mathbb{I}_{|\varepsilon_{t-i-j}|> a_T})^2 := 2R_k^{\leq} + 2R_k^{>}\end{aligned}$$

with

$$\begin{aligned} \mathrm{E}\, R_k^{\leq} &\leq \mathrm{E}(\varepsilon_1^2 \mathbb{I}_{|\varepsilon_1|\leq a_T}) \sum_{t=k+1}^{T} \sum_{i,m=k+1}^{\infty} |\beta_i||\beta_m| \sum_{j,n=0}^{\infty} |\gamma_j||\gamma_n| \\ &\quad + \{\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1|\leq a_T})\}^2 \sum_{t,s=k+1}^{T} \sum_{i,m=k+1}^{\infty} |\beta_i||\beta_m| \sum_{j,n=0}^{\infty} |\gamma_j||\gamma_n| \\ &= O(1)\, \|\boldsymbol{\gamma}\|_1^2 (a_T \sum_{i=k+1}^{\infty} |\beta_i|)^2 = o(1) \end{aligned}$$

by using $\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1|\leq a_T}) = -\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1|> a_T})$ and KT ($\|\boldsymbol{\gamma}\|_1 := \sum_{i=0}^{\infty} |\gamma_i|$), and

$$\mathrm{E}(R_k^{>})^{1/2} \leq \mathrm{E}(|\varepsilon_1| \mathbb{I}_{|\varepsilon_1|>a_T}) T \sum_{i=k+1}^{\infty} |\beta_i| \sum_{j=0}^{\infty} |\gamma_j| = O(l_T)(a_T \sum_{i=k+1}^{\infty} |\beta_i|) = o(l_T),$$

so $\sum_{t=k+1}^{T} \rho_{t,k} = o_P(l_T)$ in this case, whereas for $\mathrm{E}\,|\varepsilon_1| = \infty$ it holds that

$$\begin{aligned} \sum_{t=k+1}^{T} |\rho_{t,k}| &\leq \sum_{t=k+1}^{T} \sum_{i=k+1}^{\infty} |\beta_i||X_{t-i}| \mathbb{I}_{|X_{t-i}|\leq a_T} + \sum_{t=k+1}^{T} \sum_{i=k+1}^{\infty} |\beta_i||X_{t-i}| \mathbb{I}_{|X_{t-i}|> a_T} \\ &=: C_k^{\leq} + C_k^{>} \end{aligned}$$

with

$$\begin{aligned} \mathrm{E}\, C_k^{\leq} &\leq T\, \mathrm{E}(|X_1| \mathbb{I}_{|X_1|\leq a_T}) \sum_{i=k+1}^{\infty} |\beta_i| = O(l_T)(a_T \sum_{i=k+1}^{\infty} |\beta_i|) = o(l_T), \\ \mathrm{E}(C_k^{>})^{\eta} &\leq \mathrm{E}(\sum_{i=k+1}^{2T-k-1} |X_{T-i}| \mathbb{I}_{|X_{T-i}|>a_T} \sum_{j=k+1}^{\infty} |\beta_i| + \sum_{i=T-k}^{\infty} |X_{-i}| \mathbb{I}_{|X_{-i}|>a_T} \sum_{j=k+1+i}^{T+i} |\beta_j|)^{\eta} \\ &\leq T\, \mathrm{E}(|X_1|^{\eta} \mathbb{I}_{|X_1|>a_T}) \{2(\sum_{j=k+1}^{\infty} |\beta_i|)^{\eta} + T^{-1} \sum_{i=T-k}^{\infty} \sum_{j=k+1+i}^{T+i} |\beta_j|^{\eta}\} \\ &\leq O(a_T^{\eta}) \{2(\sum_{j=k+1}^{\infty} |\beta_i|)^{\eta} + \sum_{j=T+1}^{\infty} |\beta_j|^{\eta}\} \end{aligned}$$

by KT for $\eta \in [\delta, \alpha)$, $\alpha \leq 1$, so from $a_T \sum_{j=k+1}^{\infty} |\beta_i| \to 0$ and $\sum_{j=T+1}^{\infty} |\beta_j|^{\eta} = O(T^{1-2\eta/\delta})$ (under $\sum_{j=1}^{\infty} j^{2/\delta} |\beta_j| < \infty$), it follows that $C_k^{>} = o_P(1) + O_P(a_T T^{1/\eta - 2/\delta}) = o_P(1)$ as $\eta \in [\delta, \alpha)$ can be chosen such that $1/\alpha + 1/\eta < 2/\delta$; eventually $\sum_{t=k+1}^{T} |\rho_{t,k}| = o_P(l_T)$ for $\mathrm{E}\,|\varepsilon_1| = \infty$. Returning to (S.4.2) and using (7.1), it follows that for all $\epsilon > 0$,

$$\sum_{t=k+1}^{T} (\hat{\varepsilon}_t - \varepsilon_t) = O_P(k^{1/2} a_k a_T^{\epsilon} + l_T) = O_P(k^{1/2} a_k a_T^{\epsilon}). \tag{S.4.3}$$

As $(\hat{\varepsilon}_t - \varepsilon_{t,k})^2 = \{(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k)'\mathbf{X}_{t-1}^k\}^2$ and π is a.s. bijective, we find

$$\sum_{t=k+1}^{T} (\hat{\varepsilon}_t - \varepsilon_{t,k})^2 \;=\; (\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k)' S_{00}^k (\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k) \;\leq\; \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 \|S_{00}^k\|_2$$
$$\leq\; \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 (\sigma_T^2 \|\Sigma_k\|_2 + o_P(a_T^2)) \;=\; O_P(a_T^2 \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2)$$

by Lemma 2(a), and since $\sigma_T^2 = O_P(a_T^2)$ and $\|\Sigma_k\|_2 = O(1)$. Next, from $(\hat{\varepsilon}_t - \varepsilon_t)^2 \leq 2(\hat{\varepsilon}_t - \varepsilon_{t,k})^2 + 2\rho_{t,k}^2$ and the a.s. bijectivity of π, it follows that P^π-a.s. (i.e., conditional on the data and $\{w_t\}_{t=k+1}^T$, with randomness stemming from π alone),

$$\text{(S.4.4)} \qquad \sum_{t=k+1}^{T} (\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(t)})^2 = \sum_{t=k+1}^{T} (\hat{\varepsilon}_t - \varepsilon_t)^2 \;=:\; \|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2$$
$$\leq 2\{\sum_{t=k+1}^{T} (\hat{\varepsilon}_t - \varepsilon_{t,k})^2 + S_{\rho\rho}^k\} = O_P(a_T^2 \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2) + 2S_{\rho\rho}^k,$$

where $S_{\rho\rho}^k := \sum_{t=k+1}^T \rho_{t,k}^2 = o_P(l_T)$ by Markov's inequality. In fact, for $\mathrm{E}\,|\varepsilon_1| < \infty$ it holds that

$$S_{\rho\rho}^k \;\leq\; 2\sum_{t=k+1}^{T} (\sum_{i=k+1}^{\infty} \beta_i \sum_{j=0}^{\infty} \gamma_j \varepsilon_{t-i-j} \mathbb{I}_{|\varepsilon_{t-i-j}| \leq a_T})^2$$
$$+2\sum_{t=k+1}^{T} (\sum_{i=k+1}^{\infty} \beta_i \sum_{j=0}^{\infty} \gamma_j \varepsilon_{t-i-j} \mathbb{I}_{|\varepsilon_{t-i-j}| > a_T})^2 \;=:\; 2S_{\rho\rho}^{\leq} + 2S_{\rho\rho}^{>}$$

with (i)

$$\mathrm{E}\, S_{\rho\rho}^{\leq} \;=\; \sum_{t=k+1}^{T} \mathrm{E}(\sum_{i=k+1}^{\infty} \beta_i \sum_{j=0}^{\infty} \gamma_j \varepsilon_{t-i-j} \mathbb{I}_{|\varepsilon_{t-i-j}| \leq a_T})^2$$
$$=\; \mathrm{E}(\varepsilon_1^2 \mathbb{I}_{|\varepsilon_1| \leq a_T}) \sum_{t=k+1}^{T} \sum_{i=k+1}^{\infty} (\sum_{j=k+1}^{i} \beta_j \gamma_{i-j})^2$$

$$+\{\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1| \leq a_T})\}^2 \sum_{t=k+1}^{T} \sum_{i,m=k+1}^{\infty} \sum_{j,n=0}^{\infty} \mathbb{I}_{\{i+j \neq m+n\}} \beta_j \beta_m \gamma_j \gamma_n$$
$$\leq\; O(a_T^2)(\sum_{i=k+1}^{\infty} \sum_{j=k+1}^{i} |\beta_j||\gamma_{i-j}|)^2 + O(T^{-1} l_T a_T^2)(\sum_{j=k+1}^{\infty} |\beta_j|)^2 \|\boldsymbol{\gamma}\|_1^2$$
$$\leq\; O(1)\,(a_T^2 \sum_{j=k+1}^{\infty} |\beta_j|)^2 \|\boldsymbol{\gamma}\|_1^2 = o(1)$$

by using $\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1|\le a_T}) = -\mathrm{E}(\varepsilon_1 \mathbb{I}_{|\varepsilon_1|>a_T})$ and KT, and (ii),

$$\mathrm{E}(S^{>}_{\rho\rho})^{\frac{1}{2}} \le \mathrm{E}(|\varepsilon_1|\mathbb{I}_{|\varepsilon_1|>a_T})T\sum_{i=k+1}^{\infty}|\beta_i|\sum_{j=0}^{\infty}|\gamma_j| = O(l_T)(a_T\sum_{i=k+1}^{\infty}|\beta_i|) = o(l_T),$$

whereas for $\mathrm{E}\,|\varepsilon_1| = \infty$ it holds that $(S^k_{\rho\rho})^{1/2} \le \sum_{t=k+1}^{T}|\rho_{t,k}| = o_P\,(l_T)$ by the earlier argument for (S.4.2), so $S^k_{\rho\rho} = o_P(l_T)$ independently of $\mathrm{E}\,|\varepsilon_1|$ (as the square of slowly varying is slowly varying). Thus, continuing (S.4.4),

$$\|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2 = O_P(a_T^2\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2) + o_P(l_T) = o_P(a_T^2). \tag{S.4.5}$$

Further, as $\hat{\varepsilon}_t^2 \le 2\varepsilon_t^2 + 2(\hat{\varepsilon}_t - \varepsilon_t)^2$, it holds P^π-a.s. that

$$\sum_{t=k+1}^{T}\hat{\varepsilon}^2_{\pi(t)} = \sum_{t=k+1}^{T}\hat{\varepsilon}^2_t =: \hat{\sigma}^2_{Tk} \le 2\sigma_T^2 + 2\,\|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2 = O_P(a_T^2).$$

S.4.1.2. *Proof of part (a).* After this preparation, we turn to

$$\begin{aligned} S^{*k}_{00} - S^{\dagger k}_{00} &= \sum_{t=k+2}^{T}\{\sum_{j=0}^{t-k-2}\hat{\boldsymbol{\gamma}}_{j:k}\varepsilon^*_{t-j-1}\sum_{j=0}^{t-k-2}\hat{\boldsymbol{\gamma}}'_{j:k}\varepsilon^*_{t-j-1} \\ &\quad - \sum_{j=0}^{t-k-2}\boldsymbol{\gamma}_{j:k}\varepsilon^\dagger_{t-j-1}\sum_{j=0}^{t-k-2}\boldsymbol{\gamma}'_{j:k}\varepsilon^\dagger_{t-j-1}\} \\ &= \sum_{t=k+2}^{T}\sum_{j,i=0}^{t-k-2}w_{t-j-1}w_{t-i-1}\{\hat{\boldsymbol{\gamma}}_{j:k}\hat{\boldsymbol{\gamma}}'_{i:k}\hat{\varepsilon}_{\pi(t-j-1)}\hat{\varepsilon}_{\pi(t-i-1)} \\ &\quad -\boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}'_{i:k}\varepsilon_{\pi(t-j-1)}\varepsilon_{\pi(t-i-1)}\}. \end{aligned}$$

Let $\mathbf{G}_{ji:k} := vec(\boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}'_{i:k})$ and $\hat{\mathbf{G}}_{ji:k} := vec(\hat{\boldsymbol{\gamma}}_{j:k}\hat{\boldsymbol{\gamma}}'_{i:k})$, such that

$$\begin{aligned} \Delta_{00} := vec(S^{*k}_{00} - S^{\dagger k}_{00}) &= \sum_{t=k+2}^{T}\sum_{j,i=0}^{t-k-2}w_{t-j-1}w_{t-i-1}\big(\hat{\mathbf{G}}_{ji:k}\hat{\varepsilon}_{\pi(t-j-1)}\hat{\varepsilon}_{\pi(t-i-1)} \\ &\quad -\mathbf{G}_{ji:k}\varepsilon_{\pi(t-j-1)}\varepsilon_{\pi(t-i-1)}\big). \end{aligned}$$

We split Δ_{00} into $\Delta_{00} = \Delta^{(1)}_{00} + \Delta^{(2)}_{00} + \Delta^{(3)}_{00}$ with $\Delta^{(1)}_{00} := \sum_{t=k+1}^{T-1}\varepsilon^2_{\pi(t)}c_{t,t}$,

$$\Delta^{(2)}_{00} := \sum_{s,t=k+1}^{T-1}\mathbb{I}_{s\ne t}w_s w_t\varepsilon_{\pi(s)}\varepsilon_{\pi(t)}c_{s,t}$$

$$\Delta^{(3)}_{00} := \sum_{t=k+1}^{T-1}\sum_{j,i=0}^{t-k-2}\hat{\mathbf{G}}_{ji:k}w_{t-j}w_{t-i}\big(\hat{\varepsilon}_{\pi(t-j)}\hat{\varepsilon}_{\pi(t-i)} - \varepsilon_{\pi(t-j)}\varepsilon_{\pi(t-i)}\big)$$

$$= \sum_{s,t=k+1}^{T-1} w_s w_t (\hat{\varepsilon}_{\pi(s)}\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(s)}\varepsilon_{\pi(t)})\hat{d}_{s,t},$$

where

$$c_{s,t} := \sum_{i=\max\{0,s-t\}}^{T-t-1} (\hat{\mathbf{G}}_{t-s+i,i:k} - \mathbf{G}_{t-s+i,i:k}),\ \hat{d}_{s,t} := \sum_{i=\max\{0,s-t\}}^{T-t-1} \hat{\mathbf{G}}_{t-s+i,i:k}.$$

For an r.p. π, with $\mathrm{E}^\dagger$ denoting expectation under $P^\dagger$, it holds that, first, $\mathrm{E}^\dagger \|\Delta_{00}^{(1)}\| \leq \sigma_T^2 \max_{t=k+1,\dots,T} \|c_{t,t}\|$, where $c_{t,t}$ remain to be evaluated.

Second, regarding $\Delta_{00}^{(2)}$, for Rademacher w_t it holds that

$$\text{(S.4.6)}\quad \begin{aligned} \mathrm{E}^\dagger \|\Delta_{00}^{(2)}\|^2 &= \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}\, \mathrm{E}^\dagger(\varepsilon_{\pi(s)}^2 \varepsilon_{\pi(t)}^2)\{c_{s,t}' c_{s,t} + c_{s,t}' c_{t,s}\} \\ &= \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)}^2) \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}\{c_{s,t}' c_{s,t} + c_{s,t}' c_{t,s}\} \\ &= O_P(T^{-2} a_T^4) \sum_{s,t=k+1}^{T-1} |c_{s,t}' c_{s,t} + c_{s,t}' c_{t,s}| \end{aligned}$$

because

$$\begin{aligned} \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)}^2) &= \sum_{u,v=k+1}^{T} \mathbb{I}_{u\neq v} P\{\pi(k+1) = u, \pi(k+2) = v\} \varepsilon_u^2 \varepsilon_v^2 \\ &= O(T^{-2})\{\sigma_T^4 - \sum_{t=k+1}^{T} \varepsilon_t^4\} = O_P(T^{-2} a_T^4), \end{aligned} \quad \text{(S.4.7)}$$

whereas for $w_t = 1$ a.s. (all t),

$$\begin{aligned} \mathrm{E}^\dagger \|\Delta_{00}^{(2)}\|^2 = {} & \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)}^2) \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}\{c_{s,t}' c_{s,t} + c_{s,t}' c_{t,s}\} \\ & + \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T)}) \\ & \times \sum_{s,t,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,v\}=3}\{c_{s,t}' c_{s,v} + c_{s,t}' c_{v,s} + c_{t,s}' c_{s,v} + c_{t,s}' c_{v,s}\} \\ & + \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)} \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T-1)} \varepsilon_{\pi(T)}) \sum_{s,t,u,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,u,v\}=4} c_{s,t}' c_{u,v}, \end{aligned}$$

where $\mathrm{E}^{\dagger}(\varepsilon^2_{\pi(k+1)}\varepsilon^2_{\pi(k+2)}) = O_P(T^{-2}a_T^4)$ as before,

$$\text{(S.4.8)}\quad \mathrm{E}^{\dagger}(\varepsilon^2_{\pi(k+1)}\varepsilon_{\pi(k+2)}\varepsilon_{\pi(T)})$$
$$= O(T^{-3})\{\sigma_T^2[(\sum_{t=k+1}^{T}\varepsilon_t)^2 - \sigma_T^2] - 2[(\sum_{t=k+1}^{T}\varepsilon_t^3)(\sum_{t=k+1}^{T}\varepsilon_t) - \sum_{t=k+1}^{T}\varepsilon_t^4]\}$$

is $O_P(T^{-3}a_T^4 l_T)$ as powers of slowly varying functions vary slowly, and

$$\mathrm{E}^{\dagger}(\varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)}\varepsilon_{\pi(T-1)}\varepsilon_{\pi(T)})$$
$$\text{(S.4.9)}\leq O(T^{-4})\{(\sum_{t=k+1}^{T}\varepsilon_t)^4 + \sum_{t=k+1}^{T}\varepsilon_t^4 + 4\sigma_T^4 + 8(\sum_{t=k+1}^{T}\varepsilon_t^3)(\sum_{t=k+1}^{T}\varepsilon_t)\}$$

is $O_P(T^{-4}a_T^4 l_T)$ because $\sum_{t=k+1}^{T}\varepsilon_t = O_P(a_T l_T)$ and $\sum_{t=k+1}^{T}\varepsilon_t^i = O_P(a_T^i)$ $(i = 2, ..., 4)$, so

$$\text{(S.4.10)}\quad \mathrm{E}^{\dagger}\,\|\Delta_{00}^{(2)}\|^2 = O_P(T^{-2}a_T^4)\sum_{s,t=k+1}^{T-1}\mathbb{I}_{s\neq t}|c'_{s,t}c_{s,t} + c'_{s,t}c_{t,s}|$$
$$+O_P(T^{-3}a_T^4 l_T)\sum_{s,t,v=k+1}^{T-1}\mathbb{I}_{\#\{s,t,v\}=3}|c'_{s,t}c_{s,v} + c'_{s,t}c_{v,s} + c'_{t,s}c_{s,v} + c'_{t,s}c_{v,s}|$$
$$+O_P(T^{-4}a_T^4 l_T)\sum_{s,t,u,v=k+1}^{T-1}\mathbb{I}_{\#\{s,t,u,v\}=4}|c'_{s,t}c_{u,v}|,$$

and $c_{s,t}$ remain to be evaluated.

Third, regarding $\Delta_{00}^{(3)}$, for Rademacher w_t it holds that

$$\mathrm{E}^{\dagger}\,\|\Delta_{00}^{(3)}\|^2 = \sum_{s,t=k+1}^{T-1}\mathrm{E}^{\dagger}(\hat{\varepsilon}_{\pi(s)}\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(s)}\varepsilon_{\pi(t)})^2\{\hat{d}'_{s,t}\hat{d}_{s,t} + \mathbb{I}_{s\neq t}\hat{d}'_{s,t}\hat{d}_{t,s}\}$$
$$\leq \mathrm{E}^{\dagger}(\hat{\varepsilon}^2_{\pi(k+1)} - \varepsilon^2_{\pi(k+1)})^2\sum_{s=k+1}^{T-1}\|\hat{d}_{s,s}\|^2$$
$$+\,\mathrm{E}^{\dagger}(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})^2\sum_{s,t=k+1}^{T-1}\mathbb{I}_{s\neq t}|\hat{d}'_{s,t}\hat{d}_{s,t} + \hat{d}'_{s,t}\hat{d}_{t,s}|,$$

where $\mathrm{E}^{\dagger}(\hat{\varepsilon}^2_{\pi(k+1)} - \varepsilon^2_{\pi(k+1)})^2$ equals

$$O(T^{-1})\sum_{s=k+1}^{T}(\hat{\varepsilon}_s^2 - \varepsilon_s^2)^2 \leq O(T^{-1})\,\|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2\,(\sigma_T^2 + \|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2)$$
$$= O_P(T^{-1}a_T^4\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2) + o_P(T^{-1}l_T a_T^2)$$

and is $O_P(T^{-1}a_k^2 a_T^{2+\epsilon})$ for all $\epsilon > 0$, using (S.4.5) and (7.1), and also

$$\begin{aligned}
\text{(S.4.11)} \quad & \mathrm{E}^\dagger (\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})^2 \\
&= O(T^{-2}) \sum_{s,t=k+1}^{T} \mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)^2 \leq O(T^{-2})(\hat{\sigma}_{Tk}^2 + \sigma_T^2)\,\|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2 \\
&= O_P(T^{-2}a_T^4\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2) + o_P(T^{-2}l_T a_T^2)
\end{aligned}$$

is $O_P(T^{-2}a_k^2 a_T^{2+\epsilon})$ for all $\epsilon > 0$, so

$$\begin{aligned}
\text{(S.4.12)} \qquad \mathrm{E}^\dagger\,\|\Delta_{00}^{(3)}\|^2 &= O_P(T^{-1}a_k^2 a_T^{2+\epsilon}) \sum_{s=k+1}^{T-1} \|\hat{d}_{s,s}\|^2 \\
&\quad + O_P(T^{-2}a_k^2 a_T^{2+\epsilon}) \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}|\hat{d}'_{s,t}\hat{d}_{s,t} + \hat{d}'_{s,t}\hat{d}_{t,s}|
\end{aligned}$$

for all $\epsilon > 0$, where $\hat{d}_{s,t}$ remain to be evaluated. If $w_t = 1$ a.s. (all t),

$$\begin{aligned}
\mathrm{E}^\dagger\,\|\Delta_{00}^{(3)}\|^2 &= \sum_{s,t=k+1}^{T-1} \mathrm{E}^\dagger (\hat{\varepsilon}_{\pi(s)}\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(s)}\varepsilon_{\pi(t)})^2 \{\hat{d}'_{s,t}\hat{d}_{s,t} + \mathbb{I}_{s\neq t}\hat{d}'_{s,t}\hat{d}_{t,s}\} \\
&+ \mathrm{E}^\dagger\{(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(T)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(T)})\} \\
&\quad \times \sum_{s,t,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,v\}=3}(\hat{d}'_{s,t} + \hat{d}'_{s,t})(\hat{d}_{s,v} + \hat{d}_{v,s}) \\
&+ \mathrm{E}^\dagger\{(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(T-1)}\hat{\varepsilon}_{\pi(T)} - \varepsilon_{\pi(T-1)}\varepsilon_{\pi(T)})\} \\
&\quad \times \sum_{s,t,u,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,u,v\}=4}\hat{d}'_{s,t}\hat{d}_{u,v},
\end{aligned}$$

where $|\,\mathrm{E}^\dagger\{(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(T)} - \varepsilon_{\pi(k+1)}\varepsilon_{\pi(T)})\}|$ equals

$$\begin{aligned}
& O(T^{-3})|\sum_{s,t,v=k+1}^{T} \mathbb{I}_{\#\{s,t,v\}=3}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)(\hat{\varepsilon}_s\hat{\varepsilon}_v - \varepsilon_s\varepsilon_v)| \\
= \; & O(T^{-3}) \sum_{s=k+1}^{T} [\{\sum_{t=k+1}^{T} \mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)\}^2 - \sum_{t=k+1}^{T} \mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)^2] \\
\leq \; & O(T^{-3}) \sum_{s=k+1}^{T} \{\sum_{t=k+1}^{T} \mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)\}^2 \\
= \; & O(T^{-3}) \sum_{s=k+1}^{T} [\hat{\varepsilon}_s^2\{\sum_{t=k+1}^{T} \mathbb{I}_{s\neq t}(\hat{\varepsilon}_t - \varepsilon_t)\}^2 + (\hat{\varepsilon}_s - \varepsilon_s)^2(\sum_{t=k+1}^{T} \mathbb{I}_{s\neq t}\varepsilon_t)^2]
\end{aligned}$$

$$\leq O(T^{-3})[\hat{\sigma}_{Tk}^2\{\sum_{t=k+1}^{T}(\hat{\varepsilon}_t-\varepsilon_t)\}^2+\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2\{\hat{\sigma}_{Tk}^2+\sigma_T^2+(\sum_{t=k+1}^{T}\varepsilon_t)^2\}]$$

is $O_P(T^{-3}ka_k^2a_T^{2+\epsilon})$ for all $\epsilon>0$, using (S.4.3), (S.4.5) and (7.1), and similarly, $|\,\mathrm{E}^{\dagger}\{(\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}-\varepsilon_{\pi(k+1)}\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(T-1)}\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T-1)}\varepsilon_{\pi(T)})\}|$ equals

$$O(T^{-4})|\sum_{s,t,u,v=k+1}^{T}\mathbb{I}_{\#\{s,t,u,v\}=4}(\hat{\varepsilon}_s\hat{\varepsilon}_t-\varepsilon_s\varepsilon_t)(\hat{\varepsilon}_u\hat{\varepsilon}_v-\varepsilon_u\varepsilon_v)|$$

$$=O(T^{-4})|\{\sum_{s,t=k+1}^{T}\mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t-\varepsilon_s\varepsilon_t)\}^2$$

$$-4\sum_{s,t,v=k+1}^{T}\mathbb{I}_{\#\{s,t,v\}=3}(\hat{\varepsilon}_s\hat{\varepsilon}_t-\varepsilon_s\varepsilon_t)(\hat{\varepsilon}_s\hat{\varepsilon}_v-\varepsilon_s\varepsilon_v)-2\sum_{s,t=k+1}^{T}\mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t-\varepsilon_s\varepsilon_t)^2\}|$$

$$\leq O(T^{-4})\{\sum_{s,t=k+1}^{T}\mathbb{I}_{s\neq t}(\hat{\varepsilon}_s\hat{\varepsilon}_t-\varepsilon_s\varepsilon_t)\}^2+O_P(T^{-4}ka_k^2a_T^{2+\epsilon})$$

using previous evaluations, so further

$$\begin{aligned}=\;& O(T^{-4})\{[\sum_{t=k+1}^{T}(\hat{\varepsilon}_t-\varepsilon_t)]^2+2\sum_{t=k+1}^{T}\varepsilon_t\sum_{s=k+1}^{T}(\hat{\varepsilon}_s-\varepsilon_s)+\sum_{t=k+1}^{T}(\varepsilon_t^2-\hat{\varepsilon}_t^2)\}^2\\ &+O_P(T^{-4}ka_k^2a_T^{2+\epsilon})\;=\;O(T^{-4})\{(k^{1/2}a_ka_T^{\epsilon}+a_Tl_T)k^{1/2}a_ka_T^{\epsilon}\\ &+\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2+2\,\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|\,\hat{\sigma}_{Tk}\}^2+O_P(T^{-4}ka_k^2a_T^{2+\epsilon})\end{aligned}$$

which is $O_P(T^{-4}ka_k^2a_T^{2+\epsilon})$ for all $\epsilon>0$. Hence,

$$\begin{aligned}\mathrm{E}^{\dagger}\,\|\Delta_{00}^{(3)}\|^2\;=\;& O_P(T^{-1}a_k^2a_T^{2+\epsilon})\sum_{s=k+1}^{T-1}\|\hat{d}_{s,s}\|^2\\ &+O_P(T^{-2}a_k^2a_T^{2+\epsilon})\sum_{s,t=k+1}^{T-1}\mathbb{I}_{s\neq t}|\hat{d}_{s,t}'\hat{d}_{s,t}+\hat{d}_{s,t}'\hat{d}_{t,s}|\\ &+O_P(T^{-3}ka_k^2a_T^{2+\epsilon})\sum_{s,t,v=k+1}^{T-1}\mathbb{I}_{\#\{s,t,v\}=3}|(\hat{d}_{s,t}'+\hat{d}_{s,t}')(\hat{d}_{s,v}+\hat{d}_{v,s})|\\ &+O_P(T^{-4}ka_k^2a_T^{2+\epsilon})\sum_{s,t,u,v=k+1}^{T-1}\mathbb{I}_{\#\{s,t,u,v\}=4}|\hat{d}_{s,t}'\hat{d}_{u,v}|.\end{aligned}\tag{S.4.13}$$

We now turn to $c_{s,t}$ and $\hat{d}_{s,t}$. As in the proof of Corollary 4,

$$\text{(S.4.14)} \qquad \begin{aligned} \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|_\infty &\leq \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|_1 \\ := \sum_{j=1}^{T-k-2} |\hat{\gamma}_j - \gamma_j| &= O_P(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|_1 + \sum_{j=k+1}^{\infty} |\beta_j|), \end{aligned}$$

$$\text{(S.4.15)} \qquad \begin{aligned} \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\| &:= (\sum_{j=1}^{T-k-2} |\hat{\gamma}_j - \gamma_j|^2)^{1/2} \\ &= O_P(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\| + \sum_{j=k+1}^{\infty} |\beta_j|). \end{aligned}$$

Using also the identity $\|vec(\mathbf{ab}')\| = \|\mathbf{a}\| \, \|\mathbf{b}\|$, the triangle inequality and (7.1), we obtain, for all $\epsilon > 0$ and s, t that

$$\begin{aligned} \|c_{s,t}\| &= \|vec\{\sum_{i=\max\{0,s-t\}}^{T-t-1} (\hat{\boldsymbol{\gamma}}_{t-s+i:k} - \boldsymbol{\gamma}_{t-s+i:k})\hat{\boldsymbol{\gamma}}'_{i:k} + \boldsymbol{\gamma}_{t-s+i:k}(\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k})'\}\| \\ &\leq (2\|\boldsymbol{\gamma}\| + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T+k-2}\|) \\ &\quad \times \sum_{i=\max\{0,s-t\}}^{T-t-1} (\|\hat{\boldsymbol{\gamma}}_{t-s+i:k} - \boldsymbol{\gamma}_{t-s+i:k}\| + \|\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k}\|) \\ &= O_P(k)\|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\| = O_P(k a_k a_T^{\epsilon-1}) \end{aligned}$$

uniformly in s, t. Thus, $\mathrm{E}^\dagger \|\Delta_{00}^{(1)}\| = O_P(k a_k a_T^{1+\epsilon})$ and, returning to (S.4.6) and (S.4.10), $\mathrm{E}^\dagger \|\Delta_{00}^{(2)}\|^2 = O_P(k^2 a_k^2 a_T^{2+\epsilon})$ for all $\epsilon > 0$.

Further, using (S.4.15), (i),

$$\begin{aligned} \sum_{s=k+1}^{T-1} \|\hat{d}_{s,s}\|^2 &\leq \sum_{s=k+1}^{T-1} (\sum_{i=0}^{T-s-1} \|\hat{\boldsymbol{\gamma}}_{i:k}\|^2)^2 \leq \sum_{s=k+1}^{T-1} (k \sum_{i=0}^{T-s-1} |\hat{\gamma}_i|^2)^2 \\ &\leq 2Tk^2(\|\boldsymbol{\gamma}\|^4 + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|^4) = O_P(Tk^2), \end{aligned}$$

and (ii),

$$\begin{aligned} \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t} |\hat{d}'_{s,t}\hat{d}_{s,t} + \hat{d}'_{s,t}\hat{d}_{t,s}| &\leq \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}(\|\hat{d}_{s,t}\|^2 + \|\hat{d}_{s,t}\|\|\hat{d}_{t,s}\|) \\ &\leq 2 \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}(\sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\|\|\hat{\boldsymbol{\gamma}}_{i:k}\|)^2 \\ &\leq \|\hat{\boldsymbol{\gamma}}_{T-k-2}\|^2 \sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t}(\sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\|)^2 \end{aligned}$$

$$
\begin{aligned}
&= O(T^2)(\|\boldsymbol{\gamma}\| + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|)^2 (\sum_{i=0}^{T-k-2} \|\hat{\boldsymbol{\gamma}}_{i:k}\|)^2 \\
&= O(k^2T^2)(\|\boldsymbol{\gamma}\| + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|)^4 = O_P(k^2T^2).
\end{aligned}
$$

Before (iii), observe that, by Assumption 1(b), Lemma S.2 and (7.1),

$$
\begin{aligned}
\sum_{i=1}^{T-k-2} i\|\hat{\boldsymbol{\gamma}}_{i:k}\| &\leq \sum_{i=1}^{T-k-2} i\|\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k}\| + \sum_{i=1}^{T-k-2} i\|\boldsymbol{\gamma}_{i:k}\| \\
&\leq \sum_{i=1}^{T-k-2} i\{\sum_{j=0}^{k-1}(\hat{\gamma}_{i-j} - \gamma_{i-j})^2\}^{1/2} + k\sum_{i=1}^{\infty} i|\gamma_i| + k^2\|\boldsymbol{\gamma}\|_1 \\
&= O_P(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|_1 + \sum_{j=k+1}^{\infty}|\beta_j|)k^{1/2}\sum_{i=1}^{T-k-2} i(1+\frac{1}{k})^{i\wedge k-i-1} \\
&+\sqrt{2}\sum_{i=1}^{T-k-2} i \sum_{j=0}^{(k-1)\wedge i} |b_{i-j,k}| + O(k^2) \\
&= O_P(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|_1 + \sum_{j=k+1}^{\infty}|\beta_j|)k^{5/2} + O(kT)\sum_{j=k+1}^{\infty}|\beta_j| \\
&+O(k^2) = O_P(k^3a_k a_T^{\epsilon-1} + kTa_T^{-1} + k^2)
\end{aligned} \tag{S.4.16}
$$

for all $\epsilon > 0$, so using also (S.4.15) and (7.1),

$$
\begin{aligned}
&\sum_{s,t,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,v\}=3}|(\hat{d}_{s,t} + \hat{d}_{t,s})'(\hat{d}_{s,v} + \hat{d}_{v,s})| \\
\leq\ & \sum_{s=k+1}^{T-1}(\sum_{t=k+1}^{T-1} \mathbb{I}_{s\neq t}\|\hat{d}_{s,t} + \hat{d}_{t,s}\|)^2 \\
\leq\ & 4\sum_{s=k+1}^{T-1}(\sum_{t=k+1}^{T-1} \mathbb{I}_{s\neq t} \sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\|\|\hat{\boldsymbol{\gamma}}_{i:k}\|)^2 \\
\leq\ & 4\|\hat{\boldsymbol{\gamma}}_{T-k-2}\|^2 \sum_{s=k+1}^{T-1}(\sum_{t=k+1}^{T-1} \mathbb{I}_{s\neq t} \sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\|)^2 \\
\leq\ & 4(\|\boldsymbol{\gamma}\| + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|)^2 \sum_{s=k+1}^{T-1}(2\sum_{i=1}^{T-k-2} i\|\hat{\boldsymbol{\gamma}}_{i:k}\|)^2 \\
=\ & O_P(k^4T + k^6a_k^2Ta_T^{\epsilon-2} + k^2T^3a_T^{-2})
\end{aligned}
$$

for all $\epsilon > 0$, and (iv), similarly,

$$\begin{aligned}\sum_{s,t,u,v=k+1}^{T-1} \mathbb{I}_{\#\{s,t,u,v\}=4} |\hat{d}_{s,t}' \hat{d}_{u,v}| &\le (\sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t} \|\hat{d}_{s,t}\|)^2 \\ &\le (\sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t} \sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\| \|\hat{\boldsymbol{\gamma}}_{i:k}\|)^2 \\ &\le \|\hat{\boldsymbol{\gamma}}_{T-k-2}\|^2 (\sum_{s,t=k+1}^{T-1} \mathbb{I}_{s\neq t} \sum_{i=0}^{T-1-\max\{t,s\}} \|\hat{\boldsymbol{\gamma}}_{|t-s|+i:k}\|)^2 \\ &\le (\|\boldsymbol{\gamma}\| + \|\hat{\boldsymbol{\gamma}}_{T-k-2} - \boldsymbol{\gamma}_{T-k-2}\|)^2 (2 \sum_{s=k+1}^{T-1} \sum_{i=1}^{T-k-2} i \|\hat{\boldsymbol{\gamma}}_i\|)^2 \\ &= O_P(k^4T^2 + k^6 a_k^2 T^2 a_T^{\epsilon-2} + k^2 T^4 a_T^{-2}).\end{aligned}$$

Thus, returning to (S.4.12) and (S.4.13),

$$\begin{aligned}\mathrm{E}^\dagger \|\Delta_{00}^{(3)}\|^2 &= O_P(k^2 a_k^2 a_T^{2+\epsilon}) \\ &\quad + O_P(T^{-4} k a_k^2 a_T^{2+\epsilon}) O_P(k^4 T^2 + k^6 a_k^2 T^2 a_T^{\epsilon-2} + k^2 T^4 a_T^{-2})\end{aligned}$$

is $O_P(k^2 a_k^2 a_T^{2+\epsilon})$ for all $\epsilon > 0$, if $k^3/T \to 0$. As also $\mathrm{E}^\dagger \|\Delta_{00}^{(1)}\| = O_P(k a_k a_T^{1+\epsilon})$ and $\mathrm{E}^\dagger \|\Delta_{00}^{(2)}\|^2 = O_P(k^2 a_k^2 a_T^{2+\epsilon})$ were found to hold, it follows that $\|\Delta_{00}\| = O_{P^\dagger}(k a_k a_T^{1+\epsilon})$ in P-probability for all $\epsilon > 0$, in the case where π is an r.p.

For $\pi = id$ (and Rademacher w_t), $\mathrm{E}^\dagger \|\Delta_{00}^{(1)}\| \le \sigma_T^2 \max_{t=k+1,\dots,T} \|c_{t,t}\|$ is $O_P(k a_k a_T^{1+\epsilon})$ as previously,

$$\mathrm{E}^\dagger \|\Delta_{00}^{(2)}\|^2 \le \sum_{s,t=k+1}^{T-1} \varepsilon_s^2 \varepsilon_t^2 \mathbb{I}_{s\neq t} |c_{s,t}' c_{s,t} + c_{s,t}' c_{t,s}| \le \sigma_T^4 O_P(k^2 a_k^2 a_T^{\epsilon-2})$$

is $O_P(k^2 a_k^2 a_T^{2+\epsilon})$ using the previous uniform estimate of $\|c_{s,t}\|$, and finally,

$$\begin{aligned}\mathrm{E}^\dagger \|\Delta_{00}^{(3)}\|^2 &= \mathrm{E}^\dagger \|\sum_{s,t=k+1}^{T-1} w_s w_t (\hat{\varepsilon}_s \hat{\varepsilon}_t - \varepsilon_s \varepsilon_t) \hat{d}_{s,t}\|^2 \\ &\le \sum_{s,t=k+1}^{T-1} (\hat{\varepsilon}_s \hat{\varepsilon}_t - \varepsilon_s \varepsilon_t)^2 \{\hat{d}_{s,t}' \hat{d}_{s,t} + \hat{d}_{s,t}' \hat{d}_{t,s}\} \\ &\le \{\sum_{s,t=k+1}^{T-1} (\hat{\varepsilon}_s \hat{\varepsilon}_t - \varepsilon_s \varepsilon_t)^2\} (\max_{s,t=k+1,\dots,T-1} \|\hat{d}_{s,t}\|)^2,\end{aligned}$$

where $\sum_{s,t=k+1}^{T-1}(\hat{\varepsilon}_s\hat{\varepsilon}_t - \varepsilon_s\varepsilon_t)^2 = O_P(a_k^2 a_T^{2+\epsilon})$ for $\epsilon > 0$ as in (S.4.11) and

$$\begin{aligned}\|\hat{d}_{s,t}\| &\leq \sum_{i=\max\{0,s-t\}}^{T-t-1} \|\mathbf{G}_{t-s+i,i:k}\| + \|c_{s,t}\| \\ &\leq \sum_{i=\max\{0,s-t\}}^{T-t-1} \|\boldsymbol{\gamma}_{t-s+i:k}\|\|\boldsymbol{\gamma}_{i:k}\| + \|c_{s,t}\| \leq k\|\boldsymbol{\gamma}\|^2 + O_P(ka_k a_T^{\epsilon-1})\end{aligned}$$

is $O_P(k)$ uniformly in s, t, so also $\mathrm{E}^\dagger \|\Delta_{00}^{(3)}\|^2 = O_P(k^2 a_k^2 a_T^{2+\epsilon})$ for every $\epsilon > 0$. By combining the evaluations of $\|\Delta_{00}^{(i)}\|$ $(i = 1, 2, 3)$ and applying Markov's inequality, the first statement in part (a) is proved also for the wild bootstrap scheme.

Regarding the lower bound for $S_{00}^{\dagger k}$, let

$$\begin{aligned}\tau &:= \min\left\{t : k+1 \leq t \leq T, |\varepsilon_t| = \max_{k+1\leq s\leq T} |\varepsilon_s|\right\}, \\ \mathcal{T} &:= \left\{\tau = \min\{t : k+1 \leq t \leq T, |\varepsilon_t| = \max_{k+1\leq s\leq T-1} |\varepsilon_{\pi(s)}|\right\};\end{aligned}$$

then $P(\mathcal{T}) \to 1$. By considerations of positive semi-definiteness, for outcomes in $\mathcal{T}$ it holds that

$$\lambda_{\min}(S_{00}^{\dagger k}) \geq \lambda_{\min}\Big(\varepsilon_\tau^2 \sum_{j=0}^{T-\pi^{-1}(\tau)-1} \boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}' + \varepsilon_\tau w_{\pi^{-1}(\tau)}\Delta_\lambda\Big), \quad \text{(S.4.17)}$$

where the right-hand side matrix collects the terms of $S_{00}^{\dagger k}$ involving ε_τ, with

$$\Delta_\lambda := \sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq\pi(t)} w_t \varepsilon_{\pi(t)}(d_{\pi^{-1}(\tau),t} + d_{t,\pi^{-1}(\tau)}),$$

$d_{s,t} := \sum_{i=\max\{0,s-t\}}^{T-t-1} \boldsymbol{\gamma}_{t-s+i:k}\boldsymbol{\gamma}_{i:k}'$. We evaluate Δ_λ for the three bootstrap schemes.

If π is an r.p., then

$$\begin{aligned}\mathrm{E}^\dagger \|\Delta_\lambda\|^2 &= \sum_{t=k+1}^{T-1} \mathrm{E}^\dagger(\mathbb{I}_{\tau\neq\pi(t)}\varepsilon_{\pi(t)}^2 \|d_{\pi^{-1}(\tau),t} + d_{t,\pi^{-1}(\tau)}\|^2) \\ &\quad + \sum_{t,s=k+1}^{T-1} \mathrm{E}^\dagger[\mathbb{I}_{\pi(t)\neq\tau\neq\pi(s),t\neq s} w_t w_s \varepsilon_{\pi(t)}\varepsilon_{\pi(s)} \\ &\quad \times \mathrm{tr}\{(d_{\pi^{-1}(\tau),t} + d_{t,\pi^{-1}(\tau)})'(d_{\pi^{-1}(\tau),s} + d_{s,\pi^{-1}(\tau)})\}].\end{aligned}$$

Next, if further w_t are Rademacher, this reduces to

$$\begin{aligned}
\mathrm{E}^{\dagger}\,\|\Delta_\lambda\|^2 &= \sum_{t=k+1}^{T-1} \mathrm{E}^{\dagger}(\mathbb{I}_{\tau\neq\pi(t)}\varepsilon_{\pi(t)}^2\|d_{\pi^{-1}(\tau),t}+d_{t,\pi^{-1}(\tau)}\|^2)\\
&= \sum_{t=k+1}^{T-1} \mathrm{E}^{\dagger}(\sum_{u,v=k+1}^{T}\mathbb{I}_{t\neq v}\mathbb{I}_{\pi(t)=u,\pi(v)=\tau}\varepsilon_u^2\|d_{v,t}+d_{t,v}\|^2)\\
&= \sum_{t=k+1}^{T-1}\sum_{u,v=k+1}^{T}\mathbb{I}_{t\neq v}\,\mathrm{P}^{\dagger}(\pi\,(t)=u,\pi(v)=\tau)\varepsilon_u^2\|d_{v,t}+d_{t,v}\|^2\\
&= O(T^{-2})\sum_{t=k+1}^{T-1}\sum_{u,v=k+1}^{T}\mathbb{I}_{t\neq v}\mathbb{I}_{u\neq\tau}\varepsilon_u^2\|d_{v,t}+d_{t,v}\|^2\\
&= O(T^{-2})\sigma_T^2\sum_{v=k+1}^{T}\sum_{t=k+1}^{T-1}\mathbb{I}_{t\neq v}\|d_{v,t}+d_{t,v}\|^2
\end{aligned}$$

because $\mathrm{P}^{\dagger}(\pi\,(t)=u,\pi(v)=\tau)=(T-k)^{-1}\,(T-k-1)^{-1}$ for $t\neq v, u\neq\tau$ (τ is †-measurable). Further,

$$\begin{aligned}
\sum_{t=k+1}^{T-1}\mathbb{I}_{t\neq v}\|d_{v,t}\|^2 &\leq \sum_{t=k+1}^{T-1}\mathbb{I}_{t\neq v}\|\sum_{i=\max\{0,v-t\}}^{T-t-1}\boldsymbol{\gamma}_{t-v+i:k}\boldsymbol{\gamma}_{i:k}'\|^2\\
&\leq 2k\|\boldsymbol{\gamma}\|^2(k^2\|\boldsymbol{\gamma}\|^2+k\sum_{i=1}^{\infty}i\gamma_i^2)=O(k^3) \qquad \text{(S.4.18)}
\end{aligned}$$

and similarly for $\sum_{t=k+1}^{T-1}\mathbb{I}_{t\neq v}\|d_{t,v}\|^2$, so for the permuted wild bootstrap, $\mathrm{E}^{\dagger}\,\|\Delta_\lambda\|^2=O_P(T^{-1}k^3)\sigma_T^2=O_P(T^{-1}k^3a_T^2)=o_P(a_T^2)$ as $k^3/T\to 0$, and $\|\Delta_\lambda\|=o_{P^\dagger}(a_T)$ in P-probability. If $w_t=1$ a.s. (all t), the estimate of $\sum_{t=k+1}^{T-1}\mathrm{E}^{\dagger}(\mathbb{I}_{\tau\neq\pi(t)}\varepsilon_{\pi(t)}^2\|d_{\pi^{-1}(\tau),t}+d_{t,\pi^{-1}(\tau)}\|^2)$ remains valid. Additionally,

$$\begin{aligned}
&\mathrm{E}^{\dagger}[\mathbb{I}_{\pi(t)\neq\tau\neq\pi(s),t\neq s}\varepsilon_{\pi(t)}\varepsilon_{\pi(s)}\,\mathrm{tr}\{(d_{\pi^{-1}(\tau),t}+d_{t,\pi^{-1}(\tau)})'(d_{\pi^{-1}(\tau),s}+d_{s,\pi^{-1}(\tau)})\}]\\
&\quad=\mathrm{E}^{\dagger}\sum_{u,v,w=k+1}^{T}\mathbb{I}_{\#\{t,s,w\}=3}\mathbb{I}_{\pi(t)=u,\pi(s)=v,\pi(w)=\tau}\varepsilon_u\varepsilon_v\\
&\qquad\times\mathrm{tr}\{(d_{w,t}+d_{t,w})'(d_{w,s}+d_{s,w})\}\\
&\quad=\sum_{u,v,w=k+1}^{T}\mathbb{I}_{\#\{t,s,w\}=3}\,\mathrm{P}^{\dagger}(\pi(t)=u,\pi(s)=v,\pi(w)=\tau)\\
&\qquad\times\varepsilon_u\varepsilon_v\,\mathrm{tr}\{(d_{w,t}+d_{t,w})'(d_{w,s}+d_{s,w})\}
\end{aligned}$$

$$= O(T^{-3}) \sum_{u,v=k+1}^{T} \varepsilon_u \varepsilon_v \mathbb{I}_{\#\{u,v,\tau\}=3} \sum_{w=k+1}^{T} \mathbb{I}_{\#\{t,s,w\}=3} \times \operatorname{tr}\{(d_{w,t}+d_{t,w})'(d_{w,s}+d_{s,w})\}$$

uniformly in t, s because $\mathrm{P}^\dagger(\pi(t)=u, \pi(s)=v, \pi(w)=\tau) = O(T^{-3})\mathbb{I}_{\#\{u,v,\tau\}=3}$ for $\#\{t,s,w\}=3$. Hence,

$$|\sum_{t,s=k+1}^{T-1} \mathrm{E}^\dagger[\mathbb{I}_{\pi(t)\neq\tau\neq\pi(s),t\neq s}\varepsilon_{\pi(t)}\varepsilon_{\pi(s)} \operatorname{tr}\{(d_{\pi^{-1}(\tau),t}+d_{t,\pi^{-1}(\tau)})'(d_{\pi^{-1}(\tau),s}+d_{s,\pi^{-1}(\tau)})\}]|$$

$$\leq O(T^{-3})[(\sum_{u=k+1}^{T}\varepsilon_u)^2+\sigma_T^2] \sum_{w=k+1}^{T} |\sum_{t,s=k+1}^{T-1} \mathbb{I}_{\#\{t,s,w\}=3} \times \operatorname{tr}\{(d_{w,t}+d_{t,w})'(d_{w,s}+d_{s,w})\}|$$

is $O_P(T^{-2}k^4a_T^2 l_T) = o_P(a_T^2)$ for $k^3/T \to 0$, since

$$|\sum_{t,s=k+1}^{T-1} \mathbb{I}_{\#\{t,s,w\}=3} \operatorname{tr}\{(d_{w,t}+d_{t,w})'(d_{w,s}+d_{s,w})\}| \leq 2^4\|\boldsymbol{\gamma}\|^2(k^2\|\boldsymbol{\gamma}\|_1 + k\sum_{i=1}^{\infty} i|\gamma_i|)^2 = O(k^4).$$

Therefore, $\mathrm{E}^\dagger\|\Delta_\lambda\|^2 = o_P(a_T^2)$, and $\|\Delta_\lambda\| = o_{P^\dagger}(a_T)$ in P-probability, also for the permutation bootstrap.

For π equal to the identity (wild bootstrap), it holds that

$$\mathrm{E}^\dagger\|\Delta_\lambda\|^2 = \sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\varepsilon_t^2\|d_{\tau,t}+d_{t,\tau}\|^2, \tag{S.4.19}$$

where

$$\begin{aligned} &\mathrm{E}\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2\|d_{\tau,t}+d_{t,\tau}\|^2 \\ =\;& \mathrm{E}\,\mathrm{E}(\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2\|d_{\tau,t}+d_{t,\tau}\|^2|\tau) \\ =\;& \mathrm{E}\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\,\mathrm{E}(\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2|\tau)\|d_{\tau,t}+d_{t,\tau}\|^2 \\ \leq\;& \mathrm{E}(\mathbb{I}_{|\varepsilon_1|\leq a_T}\varepsilon_1^2)\,\mathrm{E}\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\|d_{\tau,t}+d_{t,\tau}\|^2 \end{aligned} \tag{S.4.20}$$

is $O(T^{-1}k^3 a_T^2) = o(a_T^2)$ because $\mathbb{I}_{\tau\neq t}\,\mathrm{E}(\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2|\tau) \leq \mathbb{I}_{\tau\neq t}\,\mathrm{E}(\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2) = \mathbb{I}_{\tau\neq t}\,\mathrm{E}(\mathbb{I}_{|\varepsilon_1|\leq a_T}\varepsilon_1^2)$ a.s. by the maximizing property of τ, $\mathrm{E}(\mathbb{I}_{|\varepsilon_t|\leq a_T}\varepsilon_t^2) = O(T^{-1}a_T^2)$ by KT,

$$\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\|d_{\tau,t} + d_{t,\tau}\|^2 \leq 4k\|\boldsymbol{\gamma}\|^2(k^2\|\boldsymbol{\gamma}\|^2 + k\sum_{i=1}^{\infty} i\gamma_i^2) = O(k^3)$$

with a deterministic bound (see equation (S.4.18)), and $k^3/T \to 0$; similarly,

$$\begin{aligned} & \mathrm{E}(\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\mathbb{I}_{|\varepsilon_t|>a_T}\varepsilon_t^2\|d_{\tau,t} + d_{t,\tau}\|^2)^{\eta/2} \\ \leq\; & \mathrm{E}(\mathbb{I}_{|\varepsilon_1|>a_T}|\varepsilon_1|^\eta)\,\mathrm{E}\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\|d_{\tau,t} + d_{t,\tau}\|^\eta \end{aligned}$$

is $O(T^{-1}k^2 a_T^\eta) = o(a_T^\eta)$ for $\eta \in [\delta, \alpha)$ and δ from Assumption 1(b), by KT and since, with a deterministic upper bound and $k^2/T \to 0$,

$$\sum_{t=k+1}^{T-1} \mathbb{I}_{\tau\neq t}\|d_{\tau,t} + d_{t,\tau}\|^\eta \leq O\,(1)\,(k^2\sum_{i=0}^{\infty}|\gamma_i|^\eta + k\sum_{i=1}^{\infty} i|\gamma_i|^\eta) = O(k^2),$$

so $\sum_{t=k+1}^{T-1}\mathbb{I}_{\tau\neq t}\mathbb{I}_{|\varepsilon_t|>a_T}\varepsilon_t^2\|d_{\tau,t} + d_{t,\tau}\|^2 = o_P(a_T^2)$. Recalling also (S.4.19) and (S.4.20), it follows that $\|\Delta_\lambda\| = o_{P^\dagger}(a_T)$ in P-probability also for the wild bootstrap.

By Weyl's inequality (Seber, 2008, p.117), the estimate of $\|\Delta_\lambda\|$ yields

$$\begin{aligned} \text{(S.4.21)}\qquad & |\lambda_{\min}\Big(\varepsilon_\tau^2 \sum_{j=0}^{T-\pi^{-1}(\tau)-1} \boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}' + \varepsilon_\tau w_{\pi^{-1}(\tau)}\Delta_\lambda\Big) \\ & -\varepsilon_\tau^2\lambda_{\min}\Big(\sum_{j=0}^{T-\pi^{-1}(\tau)-1} \boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}'\Big)| \leq |\varepsilon_\tau|\|\Delta_\lambda\| = o_{P^\dagger}(a_T^2) \end{aligned}$$

in P-probability, because $a_T^{-1}\varepsilon_\tau$ converges in distribution under P. Again by Weyl's inequality and matrix symmetry,

$$\begin{aligned} \left|\lambda_{\min}\Big(\sum_{j=0}^{T-\pi^{-1}(\tau)-1} \boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}'\Big) - \lambda_{\min}(\Sigma_k)\right| \;&\leq\; \|\sum_{j=T-\pi^{-1}(\tau)}^{\infty} \boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}'\|_2 \\ \leq\; \sum_{j=T-\pi^{-1}(\tau)}^{\infty} \|\boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}'\|_1 \;&\leq\; k(\max_i|\gamma_i|) \sum_{j=\max\{0,T-\pi^{-1}(\tau)-k+1\}}^{\infty} |\gamma_j| \end{aligned}$$

is $o_{P^\dagger}(1)$ in P-probability, the magnitude order because $P^\dagger(T-\pi^{-1}(\tau)-k+1\geq k)\overset{P}{\to}1$ and $k\sum_{j=k}^{\infty}|\gamma_j|\leq\sum_{j=k}^{\infty}j|\gamma_j|\to 0$ as $k\to\infty$ as the tail of a convergent series. Since $\lambda_{\min}(\Sigma_k)$ is bounded away from zero, it follows that $\lambda_{\min}(\sum_{j=0}^{T-\pi^{-1}(\tau)-1}\boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}')$ is bounded away from zero in $P^\dagger$-probability, and as further $a_T^{-1}\varepsilon_\tau$ converges weakly under P to a distribution with no atom at zero, it follows that $\lambda_{\min}(a_T^{-2}\varepsilon_\tau^2\sum_{j=0}^{T-\pi^{-1}(\tau)-1}\boldsymbol{\gamma}_{j:k}\boldsymbol{\gamma}_{j:k}')$ is bounded away from zero in $P^\dagger$, then P, probability. Recalling (S.4.17) and (S.4.21), we can conclude that also $\lambda_{\min}(a_T^{-2}S_{00}^{\dagger k})$ is bounded away from zero in $P^\dagger$, then P, probability.

S.4.1.3. *Proof of part (b).* Write $S_{0\varepsilon}^{*k}-S_{0\varepsilon}^{\dagger k}=\sigma_1+\sigma_2^*-\sigma_2^\dagger+\sigma_3$ with $\sigma_2^*:=\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^{*k}\rho_{\pi(t),k}w_t$, $\sigma_2^\dagger:=\sum_{t=k+1}^{T}\mathbf{X}_{t-1}^{\dagger k}\rho_{\pi(t),k}^\dagger$ and

$$\begin{aligned}\sigma_3 &:= \sum_{t=k+1}^{T}(\mathbf{X}_{t-1}^{*k}-\mathbf{X}_{t-1}^{\dagger k})\varepsilon_{\pi(t)}w_t\\ &= \sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}(\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k})\hat{\varepsilon}_{\pi(t-j-1)}\varepsilon_{\pi(t)}w_{t-j-1}w_t\\ &\quad+\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\boldsymbol{\gamma}_{j:k}(\hat{\varepsilon}_{\pi(t-j-1)}-\varepsilon_{\pi(t-j-1)})\varepsilon_{\pi(t)}w_{t-j-1}w_t.\end{aligned}$$

If π is an r.p., we discuss $\sigma_1+\sigma_2^*$, $\sigma_2^\dagger$ and σ_3, no matter how w_t are specified, whereas if π is the identity, we evaluate σ_2^*, $\sigma_2^\dagger$ and σ_3.

Let π **be the identity** (and w_t be Rademacher). For the discussion of σ_2^*, define modified $\rho_{t,kj}:=\rho_{t,k}-\varepsilon_{t-j-1}\sum_{i=k+1}^{j+1}\beta_i\gamma_{j+1-i}$, $(j=k,...,T-k-2)$, as $\rho_{t,k}$ 'cleaned' from the contribution of ε_{t-j-1}, and $\rho_{t,kj}:=\rho_{t,k}$ for $j=0,...,k-1$. It holds that

$$\begin{aligned}\mathrm{E}^\dagger\|\sigma_2^*\|^2 &= \sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}\|^2\hat{\varepsilon}_{t-j-1}^2\rho_{t,k}^2\\ &\leq 4\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}(\|\boldsymbol{\gamma}_{j:k}\|^2+\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2)(\varepsilon_{t-j-1}^2+(\hat{\varepsilon}_{t-j-1}-\varepsilon_{t-j-1})^2)\rho_{t,k}^2\\ &\leq 8(\sum_{t=k+2}^{T}\sum_{j=k}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|\varepsilon_{t-j-1}^2|\sum_{i=k+1}^{j+1}\beta_i\gamma_{j+1-i}|)^2\\ &+8\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\rho_{t,kj}^2+4\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\rho_{t,k}^2\end{aligned}$$

$$+4\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}(\|\boldsymbol{\gamma}_{j:k}\|^2+\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2)(\hat{\varepsilon}_{t-j-1}-\varepsilon_{t-j-1})^2\rho_{t,k}^2$$
$$=:\ 8(\varsigma^2+\sigma_{21}^*)+4(\sigma_{22}^*+\sigma_{23}^*).$$

Here

$$\begin{aligned}\varsigma &= \sum_{t=k+2}^{T}\sum_{j=k}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|\varepsilon_{t-j-1}^2(\mathbb{I}_{\{|\varepsilon_{t-j-1}|\le a_T\}}+\mathbb{I}_{\{|\varepsilon_{t-j-1}|>a_T\}})|\sum_{i=k+1}^{j+1}\beta_i\gamma_{j+1-i}| \\ &:= \varsigma^{\le}+\varsigma^{>}=o_P(a_T)\end{aligned}$$

by Markov's inequality, since

$$\begin{aligned}\mathrm{E}(\varsigma^{\le}) &\le T\,\mathrm{E}(\varepsilon_1^2\mathbb{I}_{\{|\varepsilon_1|\le a_T\}})\|\boldsymbol{\gamma}\|\sum_{j=k}^{T-k-2}\|\boldsymbol{\gamma}_{j:k}\|\sum_{i=k+1}^{\infty}|\beta_i| \\ &\le T\,\mathrm{E}(\varepsilon_1^2\mathbb{I}_{\{|\varepsilon_1|\le a_T\}})\|\boldsymbol{\gamma}\|(\sum_{i=1}^{\infty}i|\gamma_i|+k\sum_{i=k+1}^{\infty}|\gamma_i|)\sum_{i=k+1}^{\infty}|\beta_i|=o_P(a_T)\end{aligned}$$

by KT, Assumption 1(b) and the condition $\sum_{i=k+1}^{\infty}|\beta_i|=o(a_T^{-1})$ and, for $\eta\in[\delta,\alpha)$, by using the same facts,

$$\begin{aligned}\mathrm{E}\,|\varsigma^{>}|^{\frac{\eta}{2}} &\le T\,\mathrm{E}(|\varepsilon_1|^{\eta}\mathbb{I}_{\{|\varepsilon_1|>a_T\}})\|\boldsymbol{\gamma}\|^{\frac{\eta}{2}}\sum_{j=k}^{T-k-2}\|\boldsymbol{\gamma}_{j:k}\|^{\frac{\eta}{2}}(\sum_{i=k+1}^{\infty}|\beta_i|)^{\frac{\eta}{2}} \\ &\le T\,\mathrm{E}(|\varepsilon_1|^{\eta}\mathbb{I}_{\{|\varepsilon_1|>a_T\}})\|\boldsymbol{\gamma}\|^{\frac{\eta}{2}}(\sum_{i=1}^{\infty}i|\gamma_i|^{\frac{\eta}{2}}+k\sum_{i=k+1}^{\infty}|\gamma_i|^{\frac{\eta}{2}})(\sum_{i=k+1}^{\infty}|\beta_i|)^{\frac{\eta}{2}}\end{aligned}$$

is $o_P(a_T^{\eta/2})$. Further, for $\eta\in[\delta,\alpha)$ and all t, by independence of ε_{t-j-1} and $\rho_{t,kj}$,

$$\mathrm{E}\,(\sigma_{21}^*)^{\frac{\eta}{2}}\le\mathrm{E}\,|\varepsilon_1|^{\eta}(\sup_{t,j}\mathrm{E}\,|\rho_{t,kj}|^{\eta})\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^{\eta},$$

where the sup is over $t=k+2,...,T$ and $j=0,...,t-k-2$. By Hölder's inequality, for $\alpha>1$ and $\eta\in[1,\alpha)$ it holds that

$$\begin{aligned}\mathrm{E}\,|\rho_{t,kj}|^{\eta} &= \mathrm{E}\,|\rho_{k+1,k}|^{\eta}\le(\sum_{i=k+1}^{\infty}|\beta_i|)^{\eta-1}\sum_{i=k+1}^{\infty}|\beta_i|\,\mathrm{E}\,|X_{k+1-i}|^{\eta} \\ &= (\sum_{i=k+1}^{\infty}|\beta_i|)^{\eta}\,\mathrm{E}\,|X_1|^{\eta}=o(a_T^{-\eta})\end{aligned}$$

if $j = 0, ..., k-1$, and similarly,

$$\begin{aligned} \mathrm{E}\,|\rho_{t,kj}|^{\eta} &\leq (\sum_{i=k+1}^{\infty}|\beta_i|)^{\eta-1}\sum_{i=k+1}^{\infty}|\beta_i|\,\mathrm{E}\,|\sum_{l=0}^{\infty}\mathbb{I}_{l\neq j+1}\gamma_l\varepsilon_{t-l}|^{\eta} \\ &\leq (\sum_{i=k+1}^{\infty}|\beta_i|)^{\eta}\|\boldsymbol{\gamma}\|^{\eta}\,\mathrm{E}\,|\varepsilon_1|^{\eta} = o(a_T^{-\eta}) \end{aligned}$$

for $j = k, ..., t-k-2$. Hence, a common, in t and j, $o(a_T^{-\eta})$ upper bound exists for $\mathrm{E}\,|\rho_{t,kj}|^{\eta}$, yielding

$$\mathrm{E}\,(\sigma_{21}^*)^{\frac{\eta}{2}} \leq o(kTa_T^{-\eta})(\sum_{j=0}^{\infty}|\gamma_j|^{\eta}) = o(kTa_T^{-\eta}),\ \eta\in[1,\alpha).$$

As $k^2/T \to 0$, this yields $\sigma_{21}^* = o_P(a_T^2)$. On the other hand, for $\alpha \leq 1$ and $\eta \in [\delta, \alpha)$ it holds that $\mathrm{E}\,|\rho_{t,kj}|^{\eta} = \mathrm{E}\,|\rho_{k+1,k}|^{\eta} \leq \mathrm{E}\,|X_1|^{\eta}\sum_{i=k+1}^{\infty}|\beta_i|^{\eta}$ if $j = 0, ..., k-1$, and

$$\mathrm{E}\,|\rho_{t,kj}|^{\eta} \leq \sum_{i=k+1}^{\infty}|\beta_i|^{\eta}\,\mathrm{E}\,|\sum_{l=0}^{\infty}\mathbb{I}_{l\neq j+1}\gamma_l\varepsilon_{t-l}|^{\eta} \leq \mathrm{E}\,|\varepsilon_1|^{\eta}\sum_{i=0}^{\infty}|\gamma_i|^{\eta}\sum_{i=k+1}^{\infty}|\beta_i|^{\eta}$$

if $j = k, ..., t-k-2$, so $\mathrm{E}\,(\sigma_{21}^*)^{\eta/2} = O(kT\sum_{i=k+1}^{\infty}|\beta_i|^{\eta})\sum_{j=k+1}^{\infty}|\gamma_j|^{\eta}$ and, by (S.2.2) and Markov's inequality, $\sigma_{21}^* = o_P(\tilde{a}_T^2)$ under $k^3/T \to 0$.

Next,

$$\sigma_{22}^* \leq \sigma_T^2 S_{\rho\rho}^k\sum_{j=0}^{T-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2 = o_P(l_T a_T^2)\sum_{j=0}^{T-k-2}\sum_{i=0}^{k-1}|\hat{\gamma}_{j-i}-\gamma_{j-i}|^2$$

since $S_{\rho\rho}^k = o_P(l_T)$ was proved in the preparation, so from (S.4.1) and (7.1), for all $\epsilon > 0$,

$$\begin{aligned} \sigma_{22}^* &= o_P(kl_T a_T^2)\{[\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1^2 + (\sum_{j=k+1}^{\infty}|\beta_j|)^2]\sum_{j=0}^{\infty}\left(1+\frac{1}{k}\right)^{-2j} + \sum_{j=0}^{\infty}b_{jk}^2\} \\ &= o_P(kl_T a_T^2)\{k^2a_k^2a_T^{\epsilon-2} + (\sum_{j=k+1}^{\infty}|\beta_j|)^2\} = o_P(k^3a_k^2T^{\epsilon}) = o_P(\tilde{a}_T^2) \end{aligned}$$

if $k^4/T \to 0$ by choosing small $\epsilon > 0$, and similarly, from (S.4.5) and (7.1),

$$\begin{aligned} \sigma_{23}^* &\leq \|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2\,S_{\rho\rho}^k\sum_{j=0}^{T-k-2}(\|\boldsymbol{\gamma}_{j:k}\|^2+\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2) \\ &= o_P(kl_T a_T^2\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|^2)(\|\boldsymbol{\gamma}\|^2+o_P(1)) = o_P(\tilde{a}_T^2) \end{aligned}$$

if $k^2/T \to 0$. We conclude that $\mathrm{E}^\dagger\|\sigma_2^*\|^2 = o_P(\tilde a_T^2)$ if $k^4/T \to 0$ and, hence, $\|\sigma_2^*\| = o_{P^\dagger}(\tilde a_T)$ in P-probability if π is the identity.

Regarding $\sigma_2^\dagger = \sum_{t=k+1}^T \mathbf{X}_{t-1}^{\dagger k}\rho_{t,k}^\dagger$ ($\rho_{t,k}^\dagger = \sum_{i=k+1}^{t-k-1}\beta_i X_{t-i}^\dagger$), we reuse several steps of the evaluation of $\|\sum_{t=k+1}^T \mathbf{X}_{t-1}^k\rho_{t,k}\|$ in the proof of Lemma 2. Namely, all the evaluations of expressions in $|\varepsilon_t|$ and ε_t^2 (equal to $|\varepsilon_t^\dagger|$ and $(\varepsilon_t^\dagger)^2$, resp.) can be used as there, upon replacement of ε_t ($t \le k$) by zeroes. A minor modification is needed only for (what is now)

$$\xi_{ij}^{\le,\dagger} := \sum_{t=k+1+\max\{i,j\}}^{T-1}\sum_{u=0}^{t-i-k-1}\sum_{v=0}^{t-j-k-1}\mathbb{I}_{\{u\neq v+j-i\}} \times \mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|\le\tilde a_T}\varepsilon_{t-i-u}^\dagger\varepsilon_{t-j-v}^\dagger\gamma_u\gamma_v$$

with

$$\mathrm{E}^\dagger(\xi_{ij}^{\le,\dagger})^2 \le 2\|\boldsymbol{\gamma}\|^2\sum_{t=k+1+\max\{i,j\}}^{T-1}\sum_{u=0}^{t-i-k-1}\sum_{v=0}^{t-j-k-1}\mathbb{I}_{\{u\neq v+j-i\}} \times \mathbb{I}_{|\varepsilon_{t-i-u}\varepsilon_{t-j-v}|\le\tilde a_T}\varepsilon_{t-i-u}^2\varepsilon_{t-j-v}^2|\gamma_u||\gamma_v|$$

possessing, by KT, $\mathrm{E}\,\mathrm{E}^\dagger(\xi_{ij}^{\le,\dagger})^2 \le O(\tilde a_T^2)\|\boldsymbol{\gamma}\|^4$ uniformly in i,j, so in place of (S.3.7),

$$\mathrm{E}\,\mathrm{E}^\dagger(\sum_{i=1}^k(\sum_{j=k+1}^\infty\beta_j\xi_{ij}^{\le,\dagger})^2) \le [\sum_{i=1}^k\sum_{j=k+1}^\infty|\beta_j|\,\mathrm{E}\,\mathrm{E}^\dagger\{(\xi_{ij}^{\le,\dagger})^2\}](\sum_{j=k+1}^\infty|\beta_j|)$$

is $O(k\tilde a_T^2)(\sum_{j=k+1}^\infty|\beta_j|)^2$. Thus, as in the discussion of $\|\sum_{t=k+1}^T\mathbf{X}_{t-1}^k\rho_{t,k}\|$ in Lemma 2, we conclude that $\|\sum_{t=k+1}^T\mathbf{X}_{t-1}^{\dagger k}\rho_{t,k}^\dagger\| = O_{P^\dagger}(a_T^2)\sum_{j=k+1}^\infty|\beta_j| = o_{P^\dagger}(a_T)$ in P-probability.

Further,

$$\mathrm{E}^\dagger\|\sigma_3\|^2 \le 4\sum_{t=k+2}^T\sum_{j=0}^{t-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2\underbrace{\{\varepsilon_{t-j-1,k}^2+(\hat\varepsilon_{t-j-1}-\varepsilon_{t-j-1,k})^2\}\varepsilon_t^2}_{=:e_{tj}}$$

$$+4\sum_{t=k+2}^T\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\{(\hat\varepsilon_{t-j-1}-\varepsilon_{t-j-1,k})^2\overset{\downarrow}{+}\rho_{t-j-1,k}^2\}\varepsilon_t^2 := 4(\sigma_{31}+\sigma_{32}+\sigma_{33}).$$

First,

$$\sigma_{31} = \sum_{t=k+2}^T\sum_{j=0}^{t-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}-\boldsymbol{\gamma}_{j:k}\|^2 e_{tj} = \sum_{t=k+2}^T\sum_{j=0}^{t-k-2}\sum_{i=0}^{k-1}(\hat\gamma_{j-i}-\gamma_{j-i})^2 e_{tj}$$

$$\leq \quad kO_P(\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|_1^2+(\sum_{j=k}^{\infty}|\beta_j|)^2)\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}e_{tj}$$
$$+2\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\sum_{i=0}^{(k-1)\wedge j}b_{j-i,k}^2 e_{tj}$$

contains, (i),

$$\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}e_{tj} \;\leq\; 2\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\{\varepsilon_{t-j-1}^2+\rho_{t-j-1,k}^2\}\varepsilon_t^2$$
$$+\,2\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|^2\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\|\mathbf{X}_{t-j-2}^k\|^2\varepsilon_t^2$$

is $O_P(a_k^{2+\epsilon}\tilde{a}_T^2)$ for all $\epsilon>0$ since, (i.i), $\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}(1+\frac{1}{k})^{-2j}\varepsilon_{t-j-1}^2\varepsilon_t^2 = O_P(a_k^{2+\epsilon}\tilde{a}_T^2)$ for all $\epsilon>0$, as

$$\mathrm{E}(\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\varepsilon_{t-j-1}^2\varepsilon_t^2\mathbb{I}_{\{|\varepsilon_{t-j-1}\varepsilon_t|\leq\tilde{a}_T\}})$$
$$\leq \quad T\,\mathrm{E}(\varepsilon_1^2\varepsilon_2^2\mathbb{I}_{\{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T\}})\sum_{j=0}^{\infty}\left(1+\frac{1}{k}\right)^{-2j} = O(k\tilde{a}_T^2),$$
$$\mathrm{E}(\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\varepsilon_{t-j-1}^2\varepsilon_t^2\mathbb{I}_{\{|\varepsilon_{t-j-1}\varepsilon_t|>\tilde{a}_T\}})^{\frac{\eta}{2}}$$
$$\leq \quad T\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^{\eta}\mathbb{I}_{\{|\varepsilon_1\varepsilon_2|>\tilde{a}_T\}})\sum_{j=0}^{\infty}\left(1+\frac{1}{k}\right)^{-j\eta} = O(k\tilde{a}_T^{\eta})$$

by KT for all $\eta\in(0,\alpha)$, (i.ii), similarly,

$$\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\|\mathbf{X}_{t-j-2}^k\|^2\varepsilon_t^2 = O_P(a_k^{4+\epsilon}\tilde{a}_T^2)$$

for all $\epsilon>0$, and is multiplied by $\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|^2 = O_P(a_k^2 a_T^{\epsilon-2}) = O_P(a_k^{-2})$ for sufficiently small $\epsilon>0$ by (7.1) and $k^3/T\to 0$, (i.iii),

$$\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\rho_{t-j-1,k}^2\varepsilon_t^2 = o_P(\tilde{a}_T^2)$$

since, for $\eta\in[\delta,\alpha)$,

$$\mathrm{E}(\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\left(1+\frac{1}{k}\right)^{-2j}\rho_{t-j-1,k}^2\varepsilon_t^2)^{\frac{\eta}{2}} \leq$$

$$\leq \; T \,\mathrm{E}\,|\varepsilon_1|^\eta \,\mathrm{E}\,|\rho_{k+1,k}|^\eta \sum_{j=0}^{\infty} \left(1+\frac{1}{k}\right)^{-j\eta} \;=\; O(kT)\,\mathrm{E}\,|\rho_{k+1,k}|^\eta$$

is $O_P(kTa_T^{-\eta})$ as $\mathrm{E}\,|\rho_{k+1,k}|^\eta$ was evaluated in the discussion of σ_{21}^* and $k^2/T \to 0$, and (ii),

$$\begin{aligned}
&\sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \sum_{i=0}^{(k-1)\wedge j} b_{j-i,k}^2 e_{tj} \\
\leq\;& 2 \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \sum_{i=0}^{(k-1)\wedge j} b_{j-i,k}^2 \{\varepsilon_{t-j-1}^2 + \rho_{t-j-1,k}^2\}\varepsilon_t^2 \\
&+2\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \sum_{i=0}^{(k-1)\wedge j} b_{j-i,k}^2 \|\mathbf{X}_{t-j-2}^k\|^2 \varepsilon_t^2 \\
=\;& O_P(kT^{-1}\tilde{a}_T^{4+\epsilon} + \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 a_k^2 kT^{-1}\tilde{a}_T^{4+\epsilon})(\sum_{i=k+1}^{\infty} |\beta_i|)^2 = o_P(a_T^2)
\end{aligned}$$

for all $\epsilon > 0$, by taking expectations as in (i) and using $\sum_{j=0}^{t-k-2}\sum_{i=0}^{k-1} b_{j-i,k}^2 \leq k\sum_{j=0}^{t-k-2} b_{jk}^2 \leq k(\sum_{j=0}^{\infty} b_{jk})^2 \leq Ck(\sum_{i=k+1}^{\infty}|\beta_i|)^2$, (7.1) and the conditions $k^3/T \to 0$ and $a_T \sum_{i=k+1}^{\infty} |\beta_i| \to 0$; for example,

$$\begin{aligned}
&\mathrm{E}(\sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \sum_{i=0}^{k-1} b_{j-i,k}^2 \varepsilon_{t-j-1}^2 \varepsilon_t^2 \mathbb{I}_{\{|\varepsilon_{t-j-1}\varepsilon_t| \leq \tilde{a}_T\}}) \\
\leq\;& T\,\mathrm{E}(\varepsilon_1^2\varepsilon_2^2 \mathbb{I}_{\{|\varepsilon_1\varepsilon_2|\leq \tilde{a}_T\}}) \sum_{j=0}^{T-k-2}\sum_{i=0}^{k-1} b_{j-i,k}^2 = O(k\tilde{a}_T^2)(\sum_{i=k+1}^{\infty} |\beta_i|)^2
\end{aligned}$$

and, by Hölder's inequality for $\eta \in (0,\alpha)$,

$$\begin{aligned}
&\mathrm{E}(\sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \sum_{i=0}^{k-1} b_{j-i,k}^2 \varepsilon_{t-j-1}^2 \varepsilon_t^2 \mathbb{I}_{\{|\varepsilon_{t-j-1}\varepsilon_t| > \tilde{a}_T\}})^{\frac{\eta}{2}} \\
&\leq\; T\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^\eta \mathbb{I}_{\{|\varepsilon_1\varepsilon_2| > \tilde{a}_T\}}) \times \sum_{j=0}^{T-k-2} (\sum_{i=0}^{k-1} b_{j-i,k}^2)^{\frac{\eta}{2}} \\
&\leq\; O(\tilde{a}_T^\eta) T^{1-\frac{\eta}{2}} (\sum_{j=0}^{T-k-2}\sum_{i=0}^{k-1} b_{j-i,k}^2)^{\frac{\eta}{2}} \;=\; \{O(kT^{\frac{2}{\eta}-1}\tilde{a}_T^2)(\sum_{i=k+1}^{\infty}|\beta_i|)^2\}^{\frac{\eta}{2}}.
\end{aligned}$$

By combining (i), (ii) and (7.1), for all $\epsilon > 0$,

$$\sigma_{31} = O_P(ka_k^{2+\epsilon}\tilde{a}_T^2)(\|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|_1^2 + (\sum_{j=k}^{\infty}|\beta_j|)^2) + 2\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\sum_{i=0}^{k-1} b_{j-i,k}^2 e_{tj}$$

is $O_P(k^2 a_k^4 a_T^{\epsilon}) + o_P(\tilde{a}_T^2) = o_P(\tilde{a}_T^2)$ since $k^4/T \to 0$ and $a_T \sum_{i=k+1}^{\infty} |\beta_i| \to 0$.

Second,

$$\sigma_{32} \leq \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\boldsymbol{\gamma}_{j:k}\|^2 \|\mathbf{X}_{t-j-2}^k\|^2 \varepsilon_t^2 = \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 O_P(a_k^{4+\epsilon} \tilde{a}_T^2)$$

is $O_P(a_k^6 a_T^{\epsilon}) = o_P(\tilde{a}_T^2)$ using (7.1) and $k^4/T \to 0$, again by taking expectations and replacing the geometric series in powers of $1 + \frac{1}{k}$ by

$$\sum_{j=0}^{T-k-2} \|\boldsymbol{\gamma}_{j:k}\|^2 \leq k \sum_{j=0}^{\infty} \gamma_j^2 = O\left(k\right), \quad \sum_{j=0}^{T-k-2} \|\boldsymbol{\gamma}_{j:k}\|^{\eta} \leq k \sum_{j=0}^{\infty} |\gamma_j|^{\eta} = O\left(k\right)$$

for $\eta \geq \delta$. Third, $\sigma_{33} := \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\boldsymbol{\gamma}_{j:k}\|^2 \rho_{t-j-1,k}^2 \varepsilon_t^2 = o_P(\tilde{a}_T^2)$, similarly to σ_{21}^*.

Returning to the initial decomposition of $\mathrm{E}^{\dagger} \|\sigma_3\|^2$, we can conclude that $\mathrm{E}^{\dagger} \|\sigma_3\|^2 = o_P(\tilde{a}_T^2)$ for $k^4/T \to 0$. By combining it with the evaluations of σ_2^* and $\sigma_2^{\dagger}$, we complete the proof of part (b) for π equal to the identity.

For an r.p. π, we start from

$$\sigma_1 + \sigma_2^* = \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \hat{\boldsymbol{\gamma}}_{j:k} \hat{\varepsilon}_{\pi(t-j-1)} (\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(t)}) w_{t-j-1} w_t.$$

If w_t are Rademacher, then

$$\begin{aligned} \mathrm{E}^{\dagger} \|\sigma_1 + \sigma_2^*\|^2 &= \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\hat{\boldsymbol{\gamma}}_{j:k}\|^2 \, \mathrm{E}^{\dagger} \{\hat{\varepsilon}_{\pi(t-j-1)}^2 (\hat{\varepsilon}_{\pi(t)} - \varepsilon_{\pi(t)})^2\} \\ &= \mathrm{E}^{\dagger} \{\hat{\varepsilon}_{\pi(k+1)}^2 (\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+2)})^2\} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\hat{\boldsymbol{\gamma}}_{j:k}\|^2, \end{aligned}$$

with $\mathrm{E}^{\dagger} \{\hat{\varepsilon}_{\pi(k+1)}^2 (\hat{\varepsilon}_{\pi(k+2)} - \varepsilon_{\pi(k+2)})^2\} \leq O(T^{-2}) \hat{\sigma}_{Tk}^2 \|\hat{\boldsymbol{\varepsilon}}_T - \boldsymbol{\varepsilon}_T\|^2 = O_P(T^{\epsilon-2} a_k^2 a_T^2)$, so

$$\begin{aligned} \mathrm{E}^{\dagger} \|\sigma_1 + \sigma_2^*\|^2 &= O_P(T^{\epsilon-1} a_k^2 a_T^2) \sum_{j=0}^{T-k-2} \|\hat{\boldsymbol{\gamma}}_{j:k}\|^2 \\ &= O_P(T^{\epsilon-1} k a_k^2 a_T^2)(\|\boldsymbol{\gamma}\|^2 + \|\hat{\boldsymbol{\gamma}}_{T+k-2} - \boldsymbol{\gamma}_{T+k-2}\|^2) \end{aligned}$$

is $O_P(T^{\epsilon-1} k a_k^2 a_T^2)$ for all $\epsilon > 0$, using equation (S.4.5) and (7.1). On the other hand, if $w_t = 1$ $P^{\dagger}$-a.s. (all t), then

$$\mathrm{E}^{\dagger} \|\sigma_1 + \sigma_2^*\|^2 = \sum_{s,t=k+2}^{T} \sum_{i=0}^{s-k-2} \sum_{j=0}^{t-k-2} \hat{\boldsymbol{\gamma}}_{i:k}' \hat{\boldsymbol{\gamma}}_{j:k} \times$$

$$\times \mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(s-i-1)}\hat{\varepsilon}_{\pi(t-j-1)}(\hat{\varepsilon}_{\pi(s)}-\varepsilon_{\pi(s)})(\hat{\varepsilon}_{\pi(t)}-\varepsilon_{\pi(t)})\}$$

$$= \mathrm{E}^{\dagger}\{\hat{\varepsilon}^2_{\pi(k+1)}(\hat{\varepsilon}_{\pi(k+2)}-\varepsilon_{\pi(k+2)})^2\}\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}\|^2$$

$$+ \mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})^2\}\sum_{t=k+2}^{T}\sum_{i,j=0}^{t-k-2}\mathbb{I}_{\{i\neq j\}}\hat{\boldsymbol{\gamma}}'_{i:k}\hat{\boldsymbol{\gamma}}_{j:k}$$

$$+ \mathrm{E}^{\dagger}\{\hat{\varepsilon}^2_{\pi(k+1)}(\hat{\varepsilon}_{\pi(k+2)}-\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})\}$$

$$\times \sum_{s,t=k+2}^{T}\sum_{i=\max\{0,s-t\}}^{s-k-2}\hat{\boldsymbol{\gamma}}'_{i:k}\hat{\boldsymbol{\gamma}}_{t-s+i:k}\mathbb{I}_{\#\{s,t,s-i-1\}=3}$$

$$+ \mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}(\hat{\varepsilon}_{\pi(k+3)}-\varepsilon_{\pi(k+3)})(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})\}$$

$$\times \sum_{s,t=k+2}^{T}\sum_{i=0}^{s-k-2}\sum_{j=0}^{t-k-2}\mathbb{I}_{\#\{s-i-1,t-j-1,s,t\}=4}\hat{\boldsymbol{\gamma}}'_{i:k}\hat{\boldsymbol{\gamma}}_{j:k}$$

by separating according to the possible subscript repetitions. Here, first, $\mathrm{E}^{\dagger}\{\hat{\varepsilon}^2_{\pi(k+1)}(\hat{\varepsilon}_{\pi(k+2)}-\varepsilon_{\pi(k+2)})^2\} = O_P(T^{\epsilon-2}a_k^2a_T^2)$ for all $\epsilon > 0$, as found earlier. Second,

$$\mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})^2\} = O(T^{-3})\sum_{u,v,s=k+1}^{T}\mathbb{I}_{\#\{u,v,s\}=3}\hat{\varepsilon}_u\hat{\varepsilon}_v(\hat{\varepsilon}_s-\varepsilon_s)^2$$

$$\leq O(T^{-3})\{(\sum_{u=k+1}^{T}\hat{\varepsilon}_u)^2\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2 - 2(\sum_{u=k+1}^{T}\hat{\varepsilon}_u)\sum_{s=k+1}^{T}\hat{\varepsilon}_s(\hat{\varepsilon}_s-\varepsilon_s)^2$$

$$+ \sum_{u=k+1}^{T}\hat{\varepsilon}^2_u(\hat{\varepsilon}_u-\varepsilon_u)^2\}$$

$$\leq O(T^{-3})\{(\sum_{u=k+1}^{T}\hat{\varepsilon}_u)^2 + 2\hat{\sigma}_{Tk}|\sum_{u=k+1}^{T}\hat{\varepsilon}_u| + \hat{\sigma}^2_{Tk}\}\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2$$

Hence, using $\sum_{t=k+1}^{T}\varepsilon_t = O_P(a_Tl_T)$, equation (S.4.3) and $k^3/T \to 0$,

$$\mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})^2\} = O_P(T^{-3})\{a_T^2l_T + ka_k^2T^{\epsilon} + k^{1/2}a_ka_TT^{\epsilon}\}$$
$$\times \|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2 = O_P(T^{-3}a_T^2l_T)\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|^2 = O_P(T^{\epsilon-3}a_k^2a_T^2)$$

for all $\epsilon > 0$. Third, with $|\sum_{u=k+1}^{T}\hat{\varepsilon}^2_u(\hat{\varepsilon}_u-\varepsilon_u)| \leq \hat{\sigma}^2_{Tk}\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\|$, we find

$$\mathrm{E}^{\dagger}\{\hat{\varepsilon}^2_{\pi(k+1)}(\hat{\varepsilon}_{\pi(k+2)}-\varepsilon_{\pi(k+2)})(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})\}$$

$$= O(T^{-3})\sum_{u,v,s=k+1}^{T}\mathbb{I}_{\#\{u,v,s\}=3}\hat{\varepsilon}^2_u(\hat{\varepsilon}_v-\varepsilon_v)(\hat{\varepsilon}_s-\varepsilon_s)$$

$$\begin{aligned}
&\leq O(T^{-3})[\hat{\sigma}_{Tk}^2\{\sum_{v=k+1}^{T}(\hat{\varepsilon}_v-\varepsilon_v)\}^2 \\
&\quad -2\sum_{u=k+1}^{T}\hat{\varepsilon}_u^2(\hat{\varepsilon}_u-\varepsilon_u)\sum_{u=k+1}^{T}(\hat{\varepsilon}_u-\varepsilon_u)+\sum_{u=k+1}^{T}\hat{\varepsilon}_u^2(\hat{\varepsilon}_u-\varepsilon_u)^2] \\
&= O_P(T^{\epsilon-3}ka_k^2a_T^2+T^{\epsilon-3}k^{1/2}a_ka_T^2\left\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\right\|+T^{-3}\hat{\sigma}_{Tk}^2\left\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\right\|^2) \\
&= O_P(T^{\epsilon-3}ka_k^2a_T^2+T^{\epsilon-3}k^{1/2}a_k^2a_T+T^{\epsilon-3}a_k^2)=O_P(T^{\epsilon-3}ka_k^2a_T^2)
\end{aligned}$$

for all $\epsilon>0$. Fourth,

$$\begin{aligned}
&\mathrm{E}^{\dagger}\{\hat{\varepsilon}_{\pi(k+1)}\hat{\varepsilon}_{\pi(k+2)}(\hat{\varepsilon}_{\pi(k+3)}-\varepsilon_{\pi(k+3)})(\hat{\varepsilon}_{\pi(T)}-\varepsilon_{\pi(T)})\} \\
&= O(T^{-4})\sum_{u,v,s,t=k+1}^{T}\mathbb{I}_{\#\{u,v,s,t\}=4}\hat{\varepsilon}_u\hat{\varepsilon}_t(\hat{\varepsilon}_v-\varepsilon_v)(\hat{\varepsilon}_s-\varepsilon_s) \\
&= O(T^{-4})\{\sum_{u,v=k+1}^{T}\mathbb{I}_{u\neq v}\hat{\varepsilon}_u(\hat{\varepsilon}_v-\varepsilon_v)\}^2-\sum_{u,v,s=k+1}^{T}\mathbb{I}_{\#\{u,v,s\}=3}\hat{\varepsilon}_u\hat{\varepsilon}_v(\hat{\varepsilon}_s-\varepsilon_s)^2 \\
&\quad -\sum_{u,v,s=k+1}^{T}\mathbb{I}_{\#\{u,v,s\}=3}\hat{\varepsilon}_u^2(\hat{\varepsilon}_v-\varepsilon_v)(\hat{\varepsilon}_s-\varepsilon_s)-\sum_{u,v=k+1}^{T}\mathbb{I}_{u\neq v}\hat{\varepsilon}_u^2(\hat{\varepsilon}_v-\varepsilon_v)^2\},
\end{aligned}$$

where the magnitude order of the last three sums was determined above, so

$$\begin{aligned}
&= O(T^{-4})\{\sum_{u=k+1}^{T}\hat{\varepsilon}_u\sum_{v=k+1}^{T}(\hat{\varepsilon}_v-\varepsilon_v)-\sum_{u=k+1}^{T}\hat{\varepsilon}_u(\hat{\varepsilon}_u-\varepsilon_u)\}^2+O_P(T^{\epsilon-4}ka_k^2a_T^2) \\
&\leq O(T^{-4})[(\sum_{u=k+1}^{T}\hat{\varepsilon}_u)^2\{\sum_{v=k+1}^{T}(\hat{\varepsilon}_v-\varepsilon_v)\}^2+\hat{\sigma}_{T,k}^2\left\|\hat{\boldsymbol{\varepsilon}}_T-\boldsymbol{\varepsilon}_T\right\|^2]+O_P(T^{\epsilon-4}ka_k^2a_T^2) \\
&= O_P(T^{-4}(a_T^2l_T+ka_k^2T^{\epsilon})ka_k^2T^{\epsilon})+O_P(T^{\epsilon-4}ka_k^2a_T^2)=O_P(T^{\epsilon-4}ka_k^2a_T^2)
\end{aligned}$$

for all $\epsilon>0$ if $k^3/T\to 0$. Returning to $\|\sigma_1+\sigma_2^*\|$,

$$\begin{aligned}
\mathrm{E}^{\dagger}\left\|\sigma_1+\sigma_2^*\right\|^2 &= O_P(T^{\epsilon-1}a_k^2a_T^2)\sum_{j=0}^{T-k-2}\|\hat{\boldsymbol{\gamma}}_{j:k}\|^2 \\
&\quad +O_P(T^{\epsilon-3}ka_k^2a_T^2)\sum_{t=k+2}^{T}\sum_{i,j=0}^{t-k-2}\mathbb{I}_{\{i\neq j\}}\hat{\boldsymbol{\gamma}}_{i:k}'\hat{\boldsymbol{\gamma}}_{j:k} \\
&\quad +O_P(T^{\epsilon-3}ka_k^2a_T^2)\sum_{s,t=k+2}^{T}\sum_{i=\max\{0,s-t\}}^{s-k-2}\hat{\boldsymbol{\gamma}}_{i:k}'\hat{\boldsymbol{\gamma}}_{t-s+i:k}\mathbb{I}_{\#\{s,t,s-i-1\}=3} \\
&\quad +O_P(T^{\epsilon-4}ka_k^2a_T^2)\sum_{s,t=k+2}^{T}\sum_{i=0}^{s-k-2}\sum_{j=0}^{t-k-2}\mathbb{I}_{\#\{s-i-1,t-j-1,s,t\}=4}\hat{\boldsymbol{\gamma}}_{i:k}'\hat{\boldsymbol{\gamma}}_{j:k}
\end{aligned}$$

$$= O_P(T^{\epsilon-1}ka_k^2a_T^2)\|\hat{\boldsymbol{\gamma}}\|^2 + O_P(T^{\epsilon-2}k^2a_k^2a_T^2)\|\hat{\boldsymbol{\gamma}}\|_1^2 + O_P(T^{\epsilon-2}k^3a_k^2a_T^2 + T^{\epsilon-1}k^2a_k^2a_T) + O_P(T^{\epsilon-2}k^3a_k^2a_T^2)\|\hat{\boldsymbol{\gamma}}\|_1^2$$

is $O_P(T^{\epsilon-1}ka_k^2a_T^2)$ for all $\epsilon > 0$ if $k^3/T \to 0$, the magnitude orders using (S.4.14), (S.4.16) and reasoning applied previously. Hence, $\|\sigma_1 + \sigma_2^*\| = o_{P^\dagger}(T^{\epsilon-1/2}k^{1/2}a_ka_T)$ in P-probability for all $\epsilon > 0$ and an r.p. π.

Next, with $\varpi_{kj} := \sum_{m=k+1}^{j}\beta_m\gamma_{j-m}$ $(j = k+1, ..., T-k-2)$, $\sigma_2^\dagger$ can be written as

$$\begin{aligned}
\sigma_2^\dagger &= \sum_{t=2k+2}^{T}\left[\sum_{i=1}^{t-k-1}\boldsymbol{\gamma}_{i-1:k}\varepsilon_{t-i}^\dagger\right]\left[\sum_{j=k+1}^{t-k-1}\varpi_{kj}\varepsilon_{t-j}^\dagger\right] \\
&= \sum_{t=2k+2}^{T}\sum_{i=k+1}^{t-k-1}\boldsymbol{\gamma}_{i-1:k}\varpi_{ki}\varepsilon_{\pi(t-i)}^2 + \sum_{t=2k+2}^{T}\sum_{i=1}^{t-k-1}\sum_{j=k+1}^{t-k-1}\mathbb{I}_{i\neq j}\boldsymbol{\gamma}_{i-1:k}\varpi_{kj}\varepsilon_{t-i}^\dagger\varepsilon_{t-j}^\dagger,
\end{aligned}$$

where, independently of how w_t are specified,

$$\begin{aligned}
\mathrm{E}^\dagger\|\sum_{t=2k+2}^{T}\sum_{i=k+1}^{t-k-1}\boldsymbol{\gamma}_{i-1:k}\varpi_{ki}\varepsilon_{\pi(t-i)}^2\| &\leq \mathrm{E}^\dagger\{\varepsilon_{\pi(T)}^2\}\sum_{t=2k+2}^{T}\sum_{i=k+1}^{t-k-1}\|\boldsymbol{\gamma}_{i-1:k}\||\varpi_{ki}| \\
&= \sigma_T^2\|\boldsymbol{\gamma}\|_1^2\sum_{m=k+1}^{\infty}|\beta_m| = o_P(a_T)
\end{aligned}$$

under Assumption 1(b) and the condition $\sum_{m=k+1}^{\infty}|\beta_m| = o(a_T^{-1})$. Further, with $f_{s,t} := \sum_{j=1+\max\{k,s-t\}}^{T-t}\boldsymbol{\gamma}_{j+t-s-1:k}\varpi_{kj}$, the term

$$\sigma_{2\times}^\dagger := \sum_{t=2k+2}^{T}\sum_{i=1}^{t-k-1}\sum_{j=k+1}^{t-k-1}\mathbb{I}_{i\neq j}\boldsymbol{\gamma}_{i-1:k}\varpi_{kj}\varepsilon_{t-i}^\dagger\varepsilon_{t-j}^\dagger = \sum_{s=2k+1}^{T-1}\sum_{t=k+1}^{T-k-1}\mathbb{I}_{s\neq t}\varepsilon_s^\dagger\varepsilon_t^\dagger f_{s,t}$$

has, for Rademacher w_t,

$$\mathrm{E}^\dagger\|\sigma_{2\times}^\dagger\|^2 = \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2\varepsilon_{\pi(k+2)}^2)(\sum_{s=2k+1}^{T-1}\sum_{t=k+1}^{T-k-1}\mathbb{I}_{s\neq t}\|f_{s,t}\|^2 + \sum_{s,t=2k+1}^{T-k-1}f_{s,t}'f_{t,s}) := \mathbf{e}_{2\times}^\dagger,$$

where

$$\begin{aligned}
\sum_{s=2k+1}^{T-1}\sum_{t=k+1}^{T-k-1}\mathbb{I}_{s\neq t}\|f_{s,t}\|^2 &\leq \|\boldsymbol{\gamma}\|^2\sum_{s=2k+1}^{T-1}\sum_{t=k+1}^{T-k-1}\mathbb{I}_{s\neq t}(\sum_{j=1+\max\{k,s-t\}}^{T-t}|\varpi_{kj}|)^2 \\
&\leq T^2\|\boldsymbol{\gamma}\|^2(\sum_{j=k+1}^{T-k-1}\sum_{m=k+1}^{j}|\beta_m||\gamma_{j-m}|)^2 \\
&\leq T^2\|\boldsymbol{\gamma}\|_1^4(\sum_{m=k+1}^{\infty}|\beta_m|)^2 = o(T^2a_T^{-2}),
\end{aligned}$$

and similarly, also $\sum_{s,t=2k+1}^{T-k-1} f_{s,t}' f_{t,s} = o(T^2 a_T^{-2})$, so

$$\mathrm{E}^\dagger \|\sigma_{2\times}^\dagger\|^2 = \mathbf{e}_{2\times}^\dagger = o(T^2 a_T^{-2})\, \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)}^2) = o_P(a_T^2)$$

using equation (S.4.7). On the other hand, for $w_t = 1$ a.s. (all t),

$$\begin{aligned}
&\mathrm{E}^\dagger \|\sigma_{2\times}^\dagger\|^2 \;=\; \mathbf{e}_{2\times}^\dagger + \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T)}) \\
&\times\Big(\sum_{s=2k+1}^{T-1} \sum_{t,u=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{su} + \sum_{s=2k+1}^{T-k-1} \sum_{t=k+1}^{T-k-1} \sum_{u=2k+1}^{T-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{us} \\
&+ \sum_{s=2k+1}^{T-1} \sum_{t=2k+1}^{T-k-1} \sum_{u=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{tu} + \sum_{s,u=2k+1}^{T-1} \sum_{t=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{ut}\Big) \\
&+ \mathrm{E}^\dagger(\varepsilon_{\pi(k+1)} \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T-1)} \varepsilon_{\pi(T)}) \sum_{s,u=2k+1}^{T-1} \sum_{t,v=k+1}^{T-k-1} \mathbb{I}_{\#\{u,v,s,t\}=4} f_{st}' f_{uv}
\end{aligned}$$

is $o_P(a_T^2)$ as

$$\begin{aligned}
&\Big| \sum_{s=2k+1}^{T-1} \sum_{t,u=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{su} \Big| \le \sum_{s=2k+1}^{T-1} \Big(\sum_{t=k+1}^{T-k-1} \mathbb{I}_{s\neq t} \|f_{st}\|\Big)^2 \\
\le\;& \sum_{s=2k+1}^{T-1} \Big(\sum_{t=k+1}^{T-k-1} \mathbb{I}_{s\neq t} \sum_{j=1+\max\{k,s-t\}}^{T-t} \|\boldsymbol{\gamma}_{j+t-s-1:k}\| |\varpi_{kj}|\Big)^2 \\
\le\;& T^3 \|\boldsymbol{\gamma}\|^2 \Big(\sum_{j=k+1}^{T-k-1} \sum_{m=k+1}^{j} |\beta_m| |\gamma_{j-m}|\Big)^2 \le T^3 \|\boldsymbol{\gamma}\|_1^4 \Big(\sum_{j=k+1}^{\infty} |\beta_j|\Big)^2 = o(T^3 a_T^{-2}),
\end{aligned}$$

and similarly $\sum_{s=2k+1}^{T-k-1} \sum_{t=k+1}^{T-k-1} \sum_{u=2k+1}^{T-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{us} = o(T^3 a_T^{-2})$, also $\sum_{s=2k+1}^{T-1} \sum_{t=2k+1}^{T-k-1} \sum_{u=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{tu} = o(T^3 a_T^{-2})$, and eventually, likewise, $\sum_{s,u=2k+1}^{T-1} \sum_{t=k+1}^{T-k-1} \mathbb{I}_{\#\{u,s,t\}=3} f_{st}' f_{ut} = o(T^3 a_T^{-2})$, whereas

$$\begin{aligned}
&\Big| \sum_{s,u=2k+1}^{T-1} \sum_{t,v=k+1}^{T-k-1} \mathbb{I}_{\#\{u,v,s,t\}=4} f_{st}' f_{uv} \Big| \le \Big(\sum_{s=2k+1}^{T-1} \sum_{t=k+1}^{T-k-1} \mathbb{I}_{s\neq t} \|f_{st}\|\Big)^2 \\
\le\;& T^4 \|\boldsymbol{\gamma}\|^2 \Big(\sum_{j=k+1}^{T-k-1} \sum_{m=k+1}^{j} |\beta_m| |\gamma_{j-m}|\Big)^2 = o(T^4 a_T^{-2})
\end{aligned}$$

and the expectations were evaluated in (S.4.8) and (S.4.9). By combining the above results with Markov's inequality, it follows that $\sigma_2^\dagger = o_{P^\dagger}(a_T)$ in P-probability.

Finally, we consider σ_3, still in the case where π is an r.p.:

$$
\begin{aligned}
\|\sigma_3\| \quad \leq \quad & \| \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} (\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k}) \hat{\varepsilon}_{\pi(t-j-1)} \varepsilon_{\pi(t)} w_{t-j-1} w_t \| \\
& + \| \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \boldsymbol{\gamma}_{j:k} (\hat{\varepsilon}_{\pi(t-j-1)} - \varepsilon_{\pi(t-j-1)}) \varepsilon_{\pi(t)} w_{t-j-1} w_t \|,
\end{aligned}
$$

where the second norm on the right-hand side is of the same form as $\|\sigma_1 + \sigma_2^*\|$, with $\boldsymbol{\gamma}_{j:k}$ in place of $\hat{\boldsymbol{\gamma}}_{j:k}$, and is $o_{P^\dagger}(T^{\epsilon-1/2}k^{1/2}a_k a_T)$ in P-probability, for all $\epsilon > 0$, by a similar argument as for $\sigma_1 + \sigma_2^*$. Regarding the other norm, say $\|\sigma_{31}\|$, for Rademacher $\{w_t\}$ it holds that

$$
\begin{aligned}
\mathrm{E}^\dagger \|\sigma_{31}\|^2 \quad &= \quad \mathrm{E}^\dagger \{\hat{\varepsilon}^2_{\pi(k+1)} \varepsilon^2_{\pi(k+2)}\} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k}\|^2 \\
&\leq \quad O(T^{-1}k) \sigma_T^2 \hat{\sigma}_{Tk}^2 \|\hat{\boldsymbol{\gamma}}_{T+k-2} - \boldsymbol{\gamma}_{T+k-2}\|^2 \\
&= \quad O(T^{-1}k a_T^4) \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k\|^2 = O_P(T^{\epsilon-1}k a_k^2 a_T^2)
\end{aligned}
$$

using equations (S.4.14), (S.4.5) and (7.1), for all $\epsilon > 0$. Similarly, for $w_t = 1$ $P^\dagger$-a.s. (all t),

$$
\begin{aligned}
\mathrm{E}^\dagger \sigma_{31} \quad &= \quad \sum_{s,t=k+2}^{T} \sum_{i=0}^{s-k-2} \sum_{j=0}^{t-k-2} (\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k})' (\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k}) \\
& \qquad \times \mathrm{E}^\dagger \{\hat{\varepsilon}_{\pi(s-i-1)} \hat{\varepsilon}_{\pi(t-j-1)} \varepsilon_{\pi(s)} \varepsilon_{\pi(t)}\} \\
&= \quad \mathrm{E}^\dagger \{\hat{\varepsilon}^2_{\pi(k+1)} \varepsilon^2_{\pi(k+2)}\} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k}\|^2
\end{aligned}
$$

$$
\begin{aligned}
& + \mathrm{E}^\dagger \{\hat{\varepsilon}_{\pi(k+1)} \hat{\varepsilon}_{\pi(k+2)} \varepsilon^2_{\pi(T)}\} \sum_{t=k+2}^{T} \sum_{i,j=0}^{t-k-2} \mathbb{I}_{\{i \neq j\}} (\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k})' (\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k}) \\
& + \mathrm{E}^\dagger \{\hat{\varepsilon}^2_{\pi(k+1)} \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T)}\} \\
& \times \sum_{s,t=k+2}^{T} \sum_{i=\max\{0,s-t\}}^{s-k-2} (\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k})' (\hat{\boldsymbol{\gamma}}_{t-s+i:k} - \boldsymbol{\gamma}_{t-s+i:k}) \mathbb{I}_{\#\{s,t,s-i-1\}=3} \\
& + \mathrm{E}^\dagger \{\hat{\varepsilon}_{\pi(k+1)} \hat{\varepsilon}_{\pi(k+2)} \varepsilon_{\pi(k+3)} \varepsilon_{\pi(T)}\} \\
& \times \sum_{s,t=k+2}^{T} \sum_{i=0}^{s-k-2} \sum_{j=0}^{t-k-2} \mathbb{I}_{\#\{s-i-1,t-j-1,s,t\}=4} (\hat{\boldsymbol{\gamma}}_{i:k} - \boldsymbol{\gamma}_{i:k})' (\hat{\boldsymbol{\gamma}}_{j:k} - \boldsymbol{\gamma}_{j:k})
\end{aligned}
$$

implying that

$$\begin{aligned}\mathrm{E}^{\dagger}\,\sigma_{31} &= O_P(T^{\epsilon-1}ka_k^2a_T^2)+O_P(T^{-2}k^2a_T^4l_T)\|\hat{\boldsymbol{\gamma}}_{T+k-2}-\boldsymbol{\gamma}_{T+k-2}\|^2\\ &\quad +O_P(T^{-2}k^3a_ka_T^{3+\epsilon}l_T+T^{-1}ka_T^3l_T)\|\hat{\boldsymbol{\gamma}}_{T+k-2}-\boldsymbol{\gamma}_{T+k-2}\|\\ &= O_P(T^{\epsilon-1}ka_k^2a_T^2)+O_P(T^{-2}k^2a_T^4l_T)\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\|^2\\ &\quad +O_P(T^{-2}k^3a_ka_T^{3+\epsilon}l_T+T^{-1}ka_T^3l_T)\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_k\| = O_P(T^{\epsilon-1}ka_k^2a_T^2),\end{aligned}$$

for all $\epsilon>0$, as

$$\begin{aligned}&\sum_{s,t=k+2}^{T}\sum_{i=\max\{0,s-t\}}^{s-k-2}\|\hat{\boldsymbol{\gamma}}_{i:k}-\boldsymbol{\gamma}_{i:k}\|\|\hat{\boldsymbol{\gamma}}_{t-s+i:k}-\boldsymbol{\gamma}_{t-s+i:k}\|\mathbb{I}_{\#\{s,t,s-i-1\}=3}\\ \leq\;& 2T\|\hat{\boldsymbol{\gamma}}_{T+k-2}-\boldsymbol{\gamma}_{T+k-2}\|\sum_{i=0}^{T-k-2}i\|\hat{\boldsymbol{\gamma}}_{i:k}-\boldsymbol{\gamma}_{i:k}\|\\ =\;& \|\hat{\boldsymbol{\gamma}}_{T+k-2}-\boldsymbol{\gamma}_{T+k-2}\|O_P(k^3a_kTa_T^{\epsilon-1}+kT^2a_T^{-1})\end{aligned}$$

by (S.4.16) and $k^3/T\to 0$, so $\|\sigma_3\|=o_{P^\dagger}(T^{\epsilon-1/2}k^{1/2}a_ka_T)$ in P-probability.

Combined with the evaluations of $\|\sigma_1+\sigma_2^*\|$ and $\sigma_2^\dagger$, this proves part (b) in the r.p. case.

S.4.1.4. *Proof of part (c).* We consider the bootstrap schemes separately for $S_{0\varepsilon}^{k\dagger}-\sigma_2^\dagger=\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\boldsymbol{\gamma}_{j:k}\varepsilon_{\pi(t-j-1)}\varepsilon_{\pi(t)}w_{t-j-1}w_t$, and we use previous evaluations for $\|\sigma_2^\dagger\|=o_{P^\dagger}(a_T)$ in P-probability.

For π equal to the identity it holds that

$$\begin{aligned}\mathrm{E}^{\dagger}\,\|S_{0\varepsilon}^{k\dagger}-\sigma_2^\dagger\|^2 &= \sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\varepsilon_t^2\\ &= \sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\varepsilon_t^2(\mathbb{I}_{|\varepsilon_{t-j-1}\varepsilon_t|\leq\tilde{a}_T}+\mathbb{I}_{|\varepsilon_{t-j-1}\varepsilon_t|>\tilde{a}_T})\end{aligned}$$

with

$$\mathrm{E}\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\varepsilon_t^2\mathbb{I}_{|\varepsilon_{t-j-1}\varepsilon_t|\leq\tilde{a}_T}\leq Tk\,\mathrm{E}(\varepsilon_1^2\varepsilon_2^2\mathbb{I}_{|\varepsilon_1\varepsilon_2|\leq\tilde{a}_T})\sum_{j=0}^{\infty}\gamma_j^2,$$

which is $O(\tilde{a}_T^2k)$, and for $\eta\in[\delta,\alpha)$,

$$\mathrm{E}(\sum_{t=k+2}^{T}\sum_{j=0}^{t-k-2}\|\boldsymbol{\gamma}_{j:k}\|^2\varepsilon_{t-j-1}^2\varepsilon_t^2\mathbb{I}_{|\varepsilon_{t-j-1}\varepsilon_t|>\tilde{a}_T})^{\frac{\eta}{2}}\leq Tk\,\mathrm{E}(|\varepsilon_1\varepsilon_2|^\eta\mathbb{I}_{|\varepsilon_1\varepsilon_2|>\tilde{a}_T})\sum_{j=0}^{\infty}|\gamma_j|^\eta$$

which is $O(\tilde{a}_T^{\eta} k)$ by KT, so

$$\mathrm{E}^{\dagger}\, \|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\|^2 = O_P(\tilde{a}_T^2 k^{2/\eta}) \text{ and } \|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\| = o_{P^{\dagger}}(\tilde{a}_T a_k^{1+\epsilon})$$

in P-probability, for $\epsilon > 0$, by Markov's inequality. Adding $\|\sigma_2^{\dagger}\| = o_{P^{\dagger}}(a_T)$ completes the case π equal to the identity.

If π is an r.p., for $S_{0\varepsilon}^{k\dagger}$ it holds that

$$\begin{aligned} \mathrm{E}^{\dagger}\, \|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\| &\leq \mathrm{E}^{\dagger} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\boldsymbol{\gamma}_{j:k}\| |\varepsilon_{\pi(t-j-1)}| |\varepsilon_{\pi(t)}| \\ &\leq O_P(T^{-1}k)(\sum_{t=k+1}^{T} |\varepsilon_t|)^2 = O_P(T^{-1}k \max\{T^2, a_T^2 l_T^2\}). \end{aligned}$$

Hence, by Markov's inequality, $\|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\| = O_{P^{\dagger}}(T^{-1}k \max\{T^2, a_T^2 l_T^2\})$ in P-probability. For large α this can be sharpened slightly by evaluating the conditional variance. For Rademacher $\{w_t\}$,

$$\mathrm{E}^{\dagger}\, \|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\|^2 = \mathrm{E}^{\dagger} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\boldsymbol{\gamma}_{j:k}\|^2 \varepsilon_{\pi(t-j-1)}^2 \varepsilon_{\pi(t)}^2 \leq O_P(T^{-1}k) \sigma_T^4$$

is $O_P(T^{-1}k a_T^4)$, whereas for constant w_t,

$$\begin{aligned} \mathrm{E}^{\dagger}\, \|S_{0\varepsilon}^{k\dagger} - \sigma_2^{\dagger}\|^2 &= \sum_{s,t=k+2}^{T} \sum_{i=0}^{s-k-2} \sum_{j=0}^{t-k-2} \boldsymbol{\gamma}_{i:k}' \boldsymbol{\gamma}_{j:k}\, \mathrm{E}^{\dagger}\{\varepsilon_{\pi(s-i-1)} \varepsilon_{\pi(t-j-1)} \varepsilon_{\pi(s)} \varepsilon_{\pi(t)}\} \\ &= \mathrm{E}^{\dagger}\{\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)}^2\} \sum_{t=k+2}^{T} \sum_{j=0}^{t-k-2} \|\boldsymbol{\gamma}_{j:k}\|^2 \\ &\quad + \mathrm{E}^{\dagger}\{\varepsilon_{\pi(k+1)} \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T)}^2\} \sum_{t=k+2}^{T} \sum_{i,j=0}^{t-k-2} \mathbb{I}_{\{i \neq j\}} \boldsymbol{\gamma}_{i:k}' \boldsymbol{\gamma}_{j:k} \\ &\quad + \mathrm{E}^{\dagger}\{\varepsilon_{\pi(k+1)}^2 \varepsilon_{\pi(k+2)} \varepsilon_{\pi(T)}\} \\ &\quad \times \sum_{s,t=k+2}^{T} \sum_{i=\max\{0,s-t\}}^{s-k-2} \boldsymbol{\gamma}_{i:k}' \boldsymbol{\gamma}_{t-s+i:k} \mathbb{I}_{\#\{s,t,s-i-1\}=3} \\ &\quad + \mathrm{E}^{\dagger}\{\varepsilon_{\pi(k+1)} \varepsilon_{\pi(k+2)} \varepsilon_{\pi(k+3)} \varepsilon_{\pi(T)}\} \\ &\quad \times \sum_{s,t=k+2}^{T} \sum_{i=0}^{s-k-2} \sum_{j=0}^{t-k-2} \mathbb{I}_{\#\{s-i-1,t-j-1,s,t\}=4} \boldsymbol{\gamma}_{i:k}' \boldsymbol{\gamma}_{j:k} \\ &= O_P(T^{-1}k a_T^4) \|\boldsymbol{\gamma}\|^2 + O_P(T^{-2}k^2 a_T^4 l_T) = O_P(T^{-1}k a_T^4) \end{aligned}$$

using $\sum_{i=1}^{\infty} i\|\boldsymbol{\gamma}_{i:k}\| \leq k^2\|\boldsymbol{\gamma}\|_1 + k\sum_{i=1}^{\infty} i|\gamma_i| = O(k^2)$, so $\|S_{0\varepsilon}^{k\dagger} - \sigma_2^\dagger\| = O_{P^\dagger}(T^{-1/2}k^{1/2}a_T^2)$ in P-probability. Adding $\|\sigma_2^\dagger\| = o_{P^\dagger}(a_T)$ completes the proof in the case where π is equal to the identity. □

S.5. Multiple Restrictions. Consider the Wald statistic W and its bootstrap counterparts W^* and W_Δ^* defined in Remark 4.2(ix). As in the proof of Theorem 3, and using repeatedly the notation introduced there in what follows, define the r.v.'s

$$\begin{aligned} W^\dagger &:= T\sigma_T^{-2}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k}[L_k(S_{00}^{\dagger k})^{-1}L_k']^{-1}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k}, \\ \breve{W} &:= T\sigma_T^{-2}L_k(\breve{S}_{00}^{k})^{-1}\breve{S}_{0\varepsilon}^{k}[L_k(\breve{S}_{00}^{k})^{-1}L_k']^{-1}L_k(\breve{S}_{00}^{k})^{-1}\breve{S}_{0\varepsilon}^{k}, \\ W^{\pi k} &:= T\sigma_T^{-2}L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k}[L_k(S_{00}^{\pi k})^{-1}L_k']^{-1}L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k}. \end{aligned}$$

The following sequential distances can then be evaluated in place of those in the proof of Theorem 3.

1. As in step 1 of that proof, it holds that $\rho_L(\mathcal{L}^*(W^*), \mathcal{L}^\dagger(W^*)) = 0$ and $\rho_L(\mathcal{L}^*(W_\Delta^*), \mathcal{L}^\dagger(W_\Delta^*)) = 0$.

2. We argue below that, for the wild bootstrap, $W_\Delta^* = W^\dagger + o_{P^\dagger}(T\tilde{a}_T^2 a_T^{-4})$ in P-probability, so

$$\rho_L(\mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}W_\Delta^*), \mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}W^\dagger)) = o_P(1), \tag{S.5.1}$$

whereas for an r.p. π, $W^* = W^\dagger + o_{P^\dagger}(T\tilde{a}_T^2 a_T^{-4})$ in P-probability, so

$$\rho_L(\mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}W^*), \mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}W^\dagger)) = o_P(1). \tag{S.5.2}$$

3(a). As in step 3 of Theorem 3's proof, under bootstrap schemes $\mathbf{w}_R$ it holds that $\rho_L(\mathcal{L}^{\boldsymbol{\varepsilon}}(S_{00}^{\dagger k}), \mathcal{L}^{|\boldsymbol{\varepsilon}|}(\breve{S}_{00}^{k})) = 0$ and $\rho_L(\mathcal{L}^{\boldsymbol{\varepsilon}}(S_{0\varepsilon}^{\dagger k}), \mathcal{L}^{|\boldsymbol{\varepsilon}|}(\breve{S}_{0\varepsilon}^{k})) = 0$ for symmetric ε_t's. As a result, $\rho_L(\mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}W^\dagger), \mathcal{L}^{|\boldsymbol{\varepsilon}|}(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W})) = 0$.

3(b). Under scheme $(\boldsymbol{\pi}_R, \mathbf{w}_1)$, it holds that $W^\dagger = \breve{W}$ algebraically.

4(a). Under symmetry of the distribution of ε_t, similarly to step 2,

$$L_k(\breve{S}_{00}^{k})^{-1}\breve{S}_{0\varepsilon}^{k} = L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k} + o_{P^{|\boldsymbol{\varepsilon}|}}(a_T^{-2}\tilde{a}_T + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T^{-1}k^{1/2}a_k),$$

$$\begin{aligned} [L_k(\breve{S}_{00}^{k})^{-1}L_k']^{-1}L_k(\breve{S}_{00}^{k})^{-1}\breve{S}_{0\varepsilon}^{k} &= [L_k(S_{00}^{\pi k})^{-1}L_k']^{-1}L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k} \\ &\quad + o_{P^{|\boldsymbol{\varepsilon}|}}(\tilde{a}_T + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T k^{1/2}a_k), \end{aligned}$$

in P-probability, so $\rho_L(\mathcal{L}^{|\boldsymbol{\varepsilon}|}(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{|\boldsymbol{\varepsilon}|}(a_T^4\tilde{a}_T^{-2}T^{-1}W^{\pi k}) = o_P(1)$. In the case where π is the identity, using

$$\begin{aligned} \hat{\sigma}_{Tk}^{-2}\sigma_T^2 &= 1 + \hat{\sigma}_{Tk}^{-2}(\|\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}\|^2 + 2\sum \hat{\varepsilon}_t(\varepsilon_t - \hat{\varepsilon}_t)) \\ &\leq 1 + \hat{\sigma}_{Tk}^{-2}(\|\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}\|^2 + 2\hat{\sigma}_{Tk}\|\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}\|) \\ &= 1 + O_P(a_T^{-2}(a_k^2 a_T^{\epsilon} + a_k a_T^{1+\epsilon})) = 1 + O_P(a_k a_T^{\epsilon-1}) \end{aligned}$$

for $\epsilon > 0$, we can conclude that $W^{\pi k}$ equals $\hat{\sigma}_{Tk}^{-2}\sigma_T^2 W = W + o_{P^{|\varepsilon|}}(W) = W + o_{P^{|\varepsilon|}}(a_T^{-4}\tilde{a}_T^2 T)$, resulting in

$$\rho_L(\mathcal{L}^{|\varepsilon|}(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{|\varepsilon|}(a_T^4\tilde{a}_T^{-2}T^{-1}W)) = o_P(1).$$

In the case of an r.p. π, $\hat{\sigma}_{Tk}^{-2}\sigma_T^2 W$ conditional on $\{|\varepsilon_t|\}_{t=-\infty}^{k}$ and the order statistics of $\{|\varepsilon_t|\}_{t=k+1}^{T}$ is distributed like $W^{\pi k}$ conditional on $\{|\varepsilon_i|\}_{i=-\infty}^{T}$, so

$$\rho_L(\mathcal{L}^{|\varepsilon|}(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{|\mathbf{e}|}(a_T^4\tilde{a}_T^{-2}T^{-1}W) = o_P(1).$$

As $\hat{\sigma}_{Tk}^{-2}\sigma_T^2 TW = TW + o_{P^{|\varepsilon|}}(W)$, also

$$\rho_L(\mathcal{L}^{|\varepsilon|}(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{|\mathbf{e}|}(a_T^4\tilde{a}_T^{-2}T^{-1}W) = o_P(1).$$

4(b). Under scheme $(\boldsymbol{\pi}_R, \mathbf{w}_1)$, similarly to step 2,

$$L_k(\breve{S}_{00}^k)^{-1}\breve{S}_{0\varepsilon}^k = L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k} + o_{P^\dagger}(a_T^{-2}\tilde{a}_T + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T^{-1}k^{1/2}a_k),$$

$$\begin{aligned}[L_k(\breve{S}_{00}^k)^{-1}L_k']^{-1}L_k(\breve{S}_{00}^k)^{-1}\breve{S}_{0\varepsilon}^k \quad = \quad & [L_k(S_{00}^{\pi k})^{-1}L_k']^{-1}L_k(S_{00}^{\pi k})^{-1}S_{0\varepsilon}^{\pi k} \\ & + o_{P^\dagger}(\tilde{a}_T + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T k^{1/2}a_k)\end{aligned}$$

in P-probability. As $\hat{\sigma}_{Tk}^{-2}\sigma_T^2 W$ conditional on $\{\varepsilon_t\}_{t=-\infty}^{k}$ and the order statistics of $\{\varepsilon_i\}_{i=k+1}^{T}$ is distributed like $W^{\pi k}$ under $P^\dagger$, it follows that

$$\rho_L(\mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{\mathbf{e}}(a_T^4\tilde{a}_T^{-2}T^{-1}W) = o_P(1).$$

Finally, $\hat{\sigma}_{Tk}^{-2}\sigma_T^2 W = W + o_{P^{\mathbf{e}}}(W)$, so also

$$\rho_L(\mathcal{L}^\dagger(a_T^4\tilde{a}_T^{-2}T^{-1}\breve{W}), \mathcal{L}^{\mathbf{e}}(a_T^4\tilde{a}_T^{-2}T^{-1}W) = o_P(1).$$

Then we combine the previous distances using the triangle inequality to conclude that

$$\begin{aligned} \text{for } \boldsymbol{\pi}_{id} \quad &: \quad \rho_L(\mathcal{L}^*(a_T^4\tilde{a}_T^{-2}T^{-1}W_\Delta^*), \mathcal{L}^{|\varepsilon|}(a_T^4\tilde{a}_T^{-2}T^{-1}W)) = o_P(1), \\ \text{for } \boldsymbol{\pi}_R, \mathbf{w}_R \quad &: \quad \rho_L(\mathcal{L}^*(a_T^4\tilde{a}_T^{-2}T^{-1}W^*), \mathcal{L}^{|\mathbf{e}|}(a_T^4\tilde{a}_T^{-2}T^{-1}W)) = o_P(1), \\ \text{for } \boldsymbol{\pi}_R, \mathbf{w}_{id} \quad &: \quad \rho_L(\mathcal{L}^*(a_T^4\tilde{a}_T^{-2}T^{-1}W^*), \mathcal{L}^{\mathbf{e}}(a_T^4\tilde{a}_T^{-2}T^{-1}W)) = o_P(1). \end{aligned}$$

These are equivalent to the convergence asserted in Remark 4.2.(ix).

The argument for Step 2 above can be structured as follows. First, the following refinement of Lemma 3(c) can be proved similarly to that lemma (we skip the details).

LEMMA S.3. *Let $k^4/T + 1/k \to 0$ and Assumption 1 hold. Then it holds that:*

(a) For an r.p. π, $\|S_{0\varepsilon}^{k\dagger}\| = O_{P^\dagger}((k + a_k^{1+\epsilon})\tilde{a}_T)$ in P probability, for every $\epsilon > 0$.

Moreover, as in Lemma 1, let the selection matrix L has δ'-summable rows under linear weighting (i.e. such that $\sum_{j=1}^{\infty} j|l_{ij}|^{\delta'} < \infty$, $i = 1, ..., m$) for some $\delta' \in (\delta, \frac{2\alpha}{2+\alpha})$, with δ as defined in Assumption 1. Then:

(b) For an r.p. π, $\|(L\Sigma^{-1})_k S_{0\varepsilon}^{\dagger k}\| = O_{P^\dagger}(\tilde{a}_T l_T)$ in P probability, where $l_T = 1$ for $\alpha \neq 1$ and l_T is slowly varying for $\alpha = 1$.

(c) For π equal to the identity, $\|(L\Sigma^{-1})_k S_{0\varepsilon}^{\dagger k}\| = O_{P^\dagger}(a_T^{1+\epsilon})$ in P probability, for every $\epsilon > 0$.

Next, it can be used to derive the following expansions.

LEMMA S.4. *Under Assumption 1 and the δ'-summability assumption of Lemma 1, it holds in P probability that, for small $\epsilon > 0$,*

$$L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} = \sigma_T^{-2}(L\Sigma^{-1})_k S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{-\epsilon-1}) = o_{P^\dagger}(a_T^{\epsilon-1})$$

if π is the identity and $k^4/T + 1/k \to 0$, whereas

$$L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} = \sigma_T^{-2}(L\Sigma^{-1})_k S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{-2}\tilde{a}_T) = o_{P^\dagger}(a_T^{-2}\tilde{a}_T l_T)$$

if π is an r.p., $k^5/T + 1/k \to 0$ and for $\alpha \leq 1$, also $k^{3+2/\alpha+\zeta}/T \to 0$ for some $\zeta > 0$.

Finally, (S.5.1) and (S.5.2) can be obtained as follows. Similarly to the argument for Step 2 in the proof of Theorem 3, it can be shown that

$$\text{(S.5.3)}\quad L_k(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k + \mathbb{I}_{\pi=id}\Delta\hat{\boldsymbol{\beta}}_k^*) = L_k(S_{00}^{*k})^{-1}(S_{0\varepsilon}^{*k} - \sigma_1^{id})$$
$$= L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{-1-\epsilon} + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T^{-1}k^{1/2}a_k),$$

$$\text{(S.5.4)}\quad [L_k(S_{00}^{*k})^{-1}L_k']^{-1}L_k(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k + \mathbb{I}_{\pi=id}\Delta\hat{\boldsymbol{\beta}}_k^*)$$
$$= [L_k(S_{00}^{\dagger k})^{-1}L_k']^{-1}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{1-\epsilon} + \mathbb{I}_{\pi=r.p.}T^{\epsilon-1/2}a_T k^{1/2}a_k),$$

It also holds that

$$\begin{aligned}|\sigma_{Tk}^{*2} - \hat{\sigma}_{Tk}^2| &\leq (\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k)'S_{00}^{*k}(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k) + 2\hat{\sigma}_{Tk}[(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k)'S_{00}^{*k}(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k)]^{1/2} \\ &= S_{\varepsilon 0}^{*k}(S_{00}^{*k})^{-1}S_{0\varepsilon}^{*k} + 2\hat{\sigma}_{Tk}[S_{\varepsilon 0}^{*k}(S_{00}^{*k})^{-1}S_{0\varepsilon}^{*k}]^{1/2}\end{aligned}$$

so $|\sigma_{Tk}^{*2} - \hat{\sigma}_{Tk}^2| = o_{P^\dagger}(a_k^{1+\epsilon}\tilde{a}_T)$ for $\pi = id$ and $|\sigma_k^{*2} - \hat{\sigma}_k^2| = O_{P^\dagger}((k + a_k)\tilde{a}_T)$ for an r.p. π, in P-probability, for all $\epsilon > 0$. As $|\hat{\sigma}_{Tk}^2 - \sigma_T^2| = o_P(a_k a_T^{1+\epsilon})$, all $\epsilon > 0$ (see equation (7.1) and (S.4.5)), it follows that

$$\text{(S.5.5)}\qquad \sigma_k^{*2} = \sigma_T^2 + o_{P^\dagger}(a_k a_T^{1+\epsilon} + \mathbb{I}_{\pi=r.p.}k a_T^{1+\epsilon}).$$

From (S.5.3) and (S.5.4), using that the eigenvalues of $S_{00}^{\dagger k}$ have exact magnitude order a_T^2 in $P^\dagger$,P probability, and $L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} = O_{P^\dagger}(a_T^{\epsilon-1})$ for $\pi = id$ and $\epsilon > 0$ (by Lemma S.4), it follows that for small $\epsilon > 0$,

$$T^{-1}\sigma_{Tk}^{*2}W_\Delta^* = L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k}[L_k(S_{00}^{\dagger k})^{-1}L_k']^{-1}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k}+o_{P^\dagger}(a_T^{-\epsilon}) = o_{P^\dagger}(a_T^{\epsilon})$$

for the wild bootstrap, if $k^4/T \to 0$. Upon division by σ_k^{*2} and its approximation by σ_T^2 according to (S.5.5), it is obtained that $W_\Delta^* = W^\dagger + o_{P^\dagger}(T\tilde{a}_T^2 a_T^{-4})$ and, as a consequence, equation (S.5.1) holds.

If π is an r.p., independently of the specification of $\{w_t\}_{t=k+1}^T$, equations (S.5.3) and (S.5.4) specialize to

$$L_k(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k) = L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{-1-\epsilon} + T^{\epsilon-1/2}a_T^{-1}k^{1/2}a_k),$$

$$\begin{aligned}[L_k(S_{00}^{*k})^{-1}L_k']^{-1}L_k(\hat{\boldsymbol{\beta}}_k^* - \hat{\boldsymbol{\beta}}_k) &= [L_k(S_{00}^{\dagger k})^{-1}L_k']^{-1}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} \\ &\quad + o_{P^\dagger}(a_T^{1-\epsilon} + T^{\epsilon-1/2}k^{1/2}a_k a_T).\end{aligned}$$

As now $L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} = O_{P^\dagger}(a_T^{-2}\tilde{a}_T l_T)$ under the hypotheses of Lemma S.4, it follows that, if $k^5/T \to 0$ (and for $\alpha \le 1$, also $k^{3+2/\alpha+\zeta}/T \to 0$ for some $\zeta > 0$), then for sufficiently small $\epsilon > 0$,

$$T^{-1}\sigma_{Tk}^{*2}W^* = L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k}[L_k(S_{00}^{\dagger k})^{-1}L_k']^{-1}L_k(S_{00}^{\dagger k})^{-1}S_{0\varepsilon}^{\dagger k} + o_{P^\dagger}(a_T^{-\epsilon})$$

is $o_{P^\dagger}(a_T^{\epsilon})$ in P probability. As previously, jointly with (S.5.5) this leads to $W^* = W^\dagger + o_{P^\dagger}(T\tilde{a}_T^2 a_T^{-4})$ and, hence, (S.5.2) follows.

Additional References.

Cavaliere G., I. Georgiev and A.M.R. Taylor (2015). Sieve-based inference for infinite-variance stationary linear processes.

Seber G. (2008). *A Matrix Handbook for Statisticians*, John Wiley and Sons. NY.

Department of Statistics
Via Belle Arti 41
I-40126 Bologna, Italy
E-mail: giuseppe.cavaliere@unibo.it

Department of Statistics
Via Belle Arti 41
I-40126 Bologna, Italy
and
Nova School of Business and Economics
Campus de Campolide
1099-032 Lisbon, Portugal
E-mail: igeorgiev@fe.unl.pt

Essex Business School
Wivenhoe Park
CO4 3SQ Colchester, UK
E-mail: rtaylor@essex.ac.uk