Appendix A. Supplementary appendix

In this appendix, we prove the main propositions and theorems in the main text. Auxiliary technical lemmas that require significant derivation will be proved in a subsequent appendix.

Appendix A.1. Low-dimensional

We begin first with covariance inequalities for τ-mixing variables which will be used throughout our proofs and the existence of the long-run variance.

Lemma A.1. *Under the assumptions A.2 and A.3, for every h, $i, j = 1, ..., m$, $k \geq 0$,*

(i) we have,

$$[\Gamma_{k,h}(t/T)]_{i,j} = |Cov(X_{ti}\nu_t^h, X_{t+k,j}\nu_{t+k}^h)| \leq 2^{\frac{1}{R-1}} \tau_k^{*\frac{R-2}{R-1}} \|X_{ti}\nu_t^h\|_R^{\frac{R-1}{R}} \|X_{t+k,j}\nu_{t+k}^h\|_R < \infty,$$

where $R > 2$ and τ_k^* are defined in assumption A.3;

(ii) Furthermore, set $\tilde{R} \in (2, q/2]$ for $q > 4$, then we have

$$|Cov(X_{ti}X_{tj}, X_{t+k,i}X_{t+k,j})| \leq 2^{\frac{1}{\tilde{R}-1}} \tau_k^{\frac{\tilde{R}-2}{\tilde{R}-1}} \|X_{ti}X_{tj}\|_{\tilde{R}}^{\frac{\tilde{R}-1}{\tilde{R}}} \|X_{t+k,i}X_{t+k,j}\|_{\tilde{R}} < \infty;$$

(iii) Additionally, $\Omega_h(\tau) = \sum_{k=-\infty}^{\infty} \Gamma_{k,h}(\tau) < \infty$.

Proof. Both parts (i) and (ii) correspond to Lemma C.3 and C.4 of Chen and Maung (2025) while part (iii) is similar to Lemma A.5 of the aforementioned paper and relies on part (i). □

Proof of Proposition 1

The following proof strategy is similar to the low-dimensional case in Chen and Maung (2023) but we deviate on at least two fronts: we do not use the reflection method which changes the bounds of our summations, and we rely on τ-mixing instead of β-mixing. The decision not to use the reflection approach is context-specific. Here, we are not interested in real-time out-of-sample forecasting as in the mentioned paper, but rather accurate in-sample estimation of impulse responses hence we use local information available prior to and after time t. This is similar to a symmetric rolling window approach centered on t that is common in estimating time-varying coefficients in macroeconomics and finance.

To continue, we establish a convenient representation of the local linear estimator. Note that we evaluate our estimator at a fixed given horizon h. Rewrite:

$$\underset{(2m\times 1)}{\hat{\theta}_t^h} = \begin{bmatrix} S_0(t/T) & S_1^\top(t/T) \\ S_1(t/T) & S_2(t/T) \end{bmatrix}^{-1} \begin{bmatrix} R_0^h(t/T) \\ R_1^h(t/T) \end{bmatrix} \equiv S(t/T)^{-1} R^h(t/T)$$

where

$$\underset{(m\times m)}{S_j(t/T)} = T^{-1}\sum_{s=1}^{T-h} X_s X_s^\top \left(\frac{s-t}{T}\right)^j k_{s,t},$$
$$\underset{(m\times 1)}{R_j^h(t/T)} = T^{-1}\sum_{s=1}^{T-h} X_s y_{s+h} \left(\frac{s-t}{T}\right)^j k_{s,t}.$$

Note that we do not index $S(t/T)$ by h as even though the horizon appears in the summation, it is non-asymptotic under our framework. As a general rule, we index a quantity with the horizon if it contains y_{t+h}, v_t^h or the local projection parameters, which are indeed objects that vary with the horizon. Define the following quantities:

$$\underset{(m\times 1)}{r_j^h(t/T)} = T^{-1}\sum_{s=1}^{T-h} X_s \nu_s^h \left(\frac{s-t}{T}\right)^j k_{s,t}$$
$$\underset{(m\times 1)}{Q_{s,t}^h} = \gamma^h\left(\frac{s}{T}\right) - \gamma^h\left(\frac{t}{T}\right) - \left(\frac{s-t}{T}\right)\gamma^{h\prime}\left(\frac{t}{T}\right) - \frac{1}{2}\left(\frac{s-t}{T}\right)^2 \gamma^{h\prime\prime}\left(\frac{t}{T}\right),$$
$$\underset{(m\times 1)}{D_j^h(t/T)} = T^{-1}\sum_{s=1}^{T-h} X_s X_s^\top \left(\frac{s-t}{T}\right)^j k_{s,t} Q_{s,t}^h,$$
$$\underset{(m\times 1)}{B_j^h(t/T)} = \frac{1}{2} S_{j+2}(t/T)\gamma^{h\prime\prime}(t/T),$$

then by substituting in y_{s+h} with the local projection and the Taylor remainder $Q_{s,t}^h$, we have the following expansion:

$$\hat{\theta}_t^h - \theta_t^h = S(t/T)^{-1}\{r^h(t/T) + B^h(t/T) + D^h(t/T)\}, \tag{A.1}$$

where

$$r^h(t/T) = (r_0^h(t/T)^\top, r_1^h(t/T)^\top)^\top$$
$$B^h(t/T) = (B_0^h(t/T)^\top, B_1^h(t/T)^\top)^\top$$
$$D^h(t/T) = (D_0^h(t/T)^\top, D_1^h(t/T)^\top)^\top.$$

The proof is facilitated with the following lemmas applied to (A.1). Their derivations are postponed to Section B.

Lemma A.2. *Under the conditions of Proposition* 1*, we have for all* t *and* h*:*

$$b^{-j} S_j(t/T) = \mu_j M(t/T)\{1 + o_p(1)\},$$

and

$$b^{-j} D_j^h(t/T) = o_p(b^2).$$

Lemma A.3. *Under the conditions of Proposition* 1*, for all* t *and* h *we have*

$$TbVar(\boldsymbol{B}^{-1}r^h(t/T)) = \tilde{\Omega}_h(t/T) + o(1),$$

where $\tilde{\Omega}_h(t/T) = diag\{\nu_0\Omega_h(t/T), \nu_2\Omega_h(t/T)\}$*,* $\boldsymbol{B} = diag\{I_{(m\times m)}, bI_{(m\times m)}\}$ *and* $\nu_j = \int u^j K(u)du$*.*

Lemma A.4. *Under the conditions of Proposition* 1*, we have for all* t *and* h*,*

$$\sqrt{Tb}\boldsymbol{B}^{-1}r^h(t/T) \to^d N(0, \tilde{\Omega}_h(t/T)).$$

We are now ready to complete the proof. Firstly, by Lemma $A.2$ and noting that $\mu_1 = 0$ by the symmetry of the kernel,

$$S(t/T)^{-1} \to^p \begin{bmatrix} M(t/T) & \mathbf{0} \\ \mathbf{0} & b^2\mu_2 M(t/T) \end{bmatrix}^{-1} = \mathbf{B}^{-1}\tilde{M}^{-1}(t/T)\mathbf{B}^{-1},$$

where $\tilde{M}(t/T) = diag\{M(t/T), \mu_2 M(t/T)\}$. Next, we also have

$$B^h(t/T) = \begin{bmatrix} \frac{b^2}{2}\mu_2 M(t/T)\gamma^{h''}(t/T) \\ \frac{b^3}{2}\mu_3 M(t/T)\gamma^{h''}(t/T) \end{bmatrix}\{1 + o_p(1)\} = \begin{bmatrix} \frac{b^2}{2}\mu_2 M(t/T)\gamma^{h''}(t/T) \\ 0 \end{bmatrix}\{1 + o_p(1)\},$$

where $\mu_3 = 0$. Therefore

$$\mathbf{B}^{-1}B^h(t/T) = \begin{bmatrix} \frac{b^2}{2}\mu_2 M(t/T)\gamma^{h''}(t/T) \\ 0 \end{bmatrix} + o_p(b^2).$$

Next, note that we have the same rate for $\mathbf{B}^{-1}D^h(t/T) = o_p(b^2)$. Therefore,

$$\mathbf{B}(\hat{\theta}_t^h - \theta_t^h) - \begin{bmatrix} \frac{b^2}{2}\mu_2\gamma^{h''}(t/T) \\ 0 \end{bmatrix} + o_p(b^2) \approx \tilde{M}^{-1}(t/T)\mathbf{B}^{-1}r^h(t/T).$$

The proof is then complete by multiplying throughout by $\sqrt{Tb}$ and applying Lemma $A.4$ to the right hand side. □

To rigorously prove Proposition 2, we first, define

$$Z_{t_1,t_2,t_3,t_4} \equiv X_{t_1,i}X_{t_1,i'}X_{t_2,j}X_{t_2,j'}X_{t_3,a}\nu_{t_3}^h X_{t_4,b}\nu_{t_4}^h,$$

where $X_{t_i,a}$ refers to the ath element of X_{t_i}, and we require the following conditions:

Assumption V: Let $R = 2(1+\varrho) > 2$ for $\varrho > 0$, then (i) for all $i, j, i', j' = 1, ..., m$ and $s, t = 1, ..., T$, $\|X_{s,i}X_{s,j}X_{t,i'}X_{t,j'}\|_{8(1+\varrho)} < \infty$ and $\|X_{s,i}X_{t,j}\nu_s^h\nu_t^h\|_{8(1+\varrho)} < \infty$; and (ii) For $t_1, t_2, t_3, t_4 \in \mathbb{Z}$, $\{Z_{t_1,t_2,t_3,t_4}\}$ is τ-mixing with coefficients given by $\widetilde{\tau}_k = O(k^{-\theta})$ and $\theta > 5(R-1)/(R-2)$.

We remark that the moment assumptions are similar to assumption T2 of Cai et al. (2022) albeit stronger as the local projection error is serially correlated and not independent of the regressors. Hence, the expectations of their products need to be adequately controlled. Additionally, note that since X_t can include an intercept term, the above condition can be reduced to simpler combinations of X_t and ν_t^h. In fact, Assumptions A.3 and A.4 can be nested in here. The assumption on mixing is not restrictive and can be replaced with geometric mixing. We now begin with the proof:

Proof of Proposition 2

Firstly,

$$
\begin{aligned}
\hat{\Omega}_h(t/T) - &\Omega_h(t/T) \\
&= \left[\frac{b}{T\nu_0} \left(\sum_{s=1}^{T-h} X_s \hat{\nu}_s^h k_{s,t} \right) \left(\sum_{s=1}^{T-h} X_s \hat{\nu}_s^h k_{s,t} \right)^\top - \frac{b}{T\nu_0} \left(\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t} \right) \left(\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t} \right)^\top \right] \\
&+ \left[\frac{b}{T\nu_0} \left(\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t} \right) \left(\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t} \right)^\top - \Omega_h(t/T) \right] \\
&\equiv \Omega_{1t} + \Omega_{2t}.
\end{aligned}
$$

We start with the first term which represents the estimation error of ν_s^h. We can further decompose this into the following three terms

$$
\begin{aligned}
\Omega_{1t} = &\frac{b}{Tv_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} [k_{s,t} k_{r,t} (\hat{\nu}_s^h - \nu_s^h) \nu_r^h X_s X_r^\top] + \frac{b}{Tv_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} [k_{s,t} k_{r,t} \nu_s^h (\hat{\nu}_r^h - \nu_r^h) X_s X_r^\top] \\
&+ \frac{b}{Tv_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} [k_{s,t} k_{r,t} (\hat{\nu}_s^h - \nu_s^h)(\hat{\nu}_r^h - \nu_r^h) X_s X_r^\top] \\
&\equiv \Omega_{11t} + \Omega_{12t} + \Omega_{13t}.
\end{aligned}
$$

Next, from the proof of Proposition 1

$$
\hat{\gamma}_t^h - \gamma_t^h = M_t^{-1} \left[r_{0,t}^h + \frac{1}{2} b^2 \mu_2 M_t \gamma_t^{h''} \right] + o_p \left((Tb)^{-1/2} + b^2 \right).
$$

where we have labeled $M(t/T) = M_t$ for convenience and similarly for $r_{0,t}^h$ and $\gamma_t^{h''}$. Next, we focus on the (i,j)th element of Ω_{11t} and its leading term is given by

$$
\frac{b}{Tv_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} \left[k_{s,t} k_{r,t} \left(-X_s^\top \left[M_s^{-1} \left\{ T^{-1} \sum_{l=1}^{T-h} X_l \nu_l^h k_{l,s} + \frac{1}{2} b^2 M_s \gamma_s'' \right\} \right] \right) \nu_r^h X_{s,i} X_{r,j} \right] \equiv \tilde{\Omega}_{11t,A} + \tilde{\Omega}_{11t,B}.
$$

We analyze the term related to the stochastic error (ignoring the -1 multiple):

$$
\begin{aligned}
[\tilde{\Omega}_{11t,A}]_{(i,j)} &= \frac{b}{T^2 v_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} \sum_{l=1}^{T-h} k_{s,t} k_{r,t} k_{l,s} (X_s^\top M_s^{-1} X_l) \nu_l^h \nu_r^h X_{s,i} X_{r,j} \\
&= \frac{b}{T^2 v_0} \sum_{s=1}^{T-h} \sum_{r=1}^{T-h} \sum_{l=1}^{T-h} k_{s,t} k_{r,t} k_{l,s} (\sum_{a=1}^{m} \sum_{b=1}^{m} m_{s,(a,b)} X_{s,a} X_{l,b}) \nu_l^h \nu_r^h X_{s,i} X_{r,j},
\end{aligned}
$$

where $m_{s,(a,b)}$ is the (a,b)th element of M_s^{-1}. Since m is finite, we focus on a specific (a,b) pair for simplicity (i.e. ignore the summation over a and b).

We consider the following cases:

$$
\begin{aligned}
[\tilde{\Omega}_{11t,A}]_{(i,j)} &= \frac{b}{T^2 \nu_0} \sum_{s=1}^{T-h} k_{s,t}^2 k_{s,s} m_{s,(a,b)} X_{s,a} X_{s,b} X_{s,i} X_{s,j} \nu_s^{h2} \qquad \text{(s = l =r)} \\
&+ \frac{b}{T^2 \nu_0} \sum_{\substack{s=1 \\ s \neq r}}^{T-h} \sum_{r=1}^{T-h} k_{s,t} k_{r,t} k_{s,s} m_{s,(a,b)} X_{s,a} X_{s,b} X_{s,i} X_{r,j} \nu_s^h \nu_r^h \qquad \text{(s = l} \neq \text{r)} \\
&+ \frac{b}{T^2 \nu_0} \sum_{\substack{s=1 \\ s \neq l}}^{T-h} \sum_{l=1}^{T-h} k_{s,t}^2 k_{l,s} m_{s,(a,b)} X_{s,a} X_{l,b} X_{s,i} X_{s,j} \nu_s^h \nu_l^h \qquad \text{(s = r} \neq \text{l)} \\
&+ \frac{b}{T^2 \nu_0} \sum_{\substack{s=1 \\ s \neq l}}^{T-h} \sum_{l=1}^{T-h} k_{s,t} k_{l,t} k_{l,s} m_{s,(a,b)} X_{s,a} X_{l,b} X_{s,i} X_{l,j} \nu_l^{h2} \qquad \text{(l = r} \neq \text{s)} \\
&+ \frac{b}{T^2 v_0} \sum_{s \neq l \neq r} k_{s,t} k_{r,t} k_{l,s} m_{s,(a,b)} X_{s,a} X_{l,b} X_{s,i} X_{r,j} \nu_l^h \nu_r^h \\
&\equiv \Omega^{(1)} + \Omega^{(2a)} + \Omega^{(2b)} + \Omega^{(2c)} + \Omega^{(3)}.
\end{aligned}
$$

We start with $\Omega^{(1)}$,

$$
E[|\Omega^{(1)}|] \leq \frac{C}{T^2 b^2} \sum_{s=1}^{T-h} E[|X_{s,a} X_{s,b} X_{s,i} X_{s,j} \nu_s^{h2}|] = O\left(\frac{1}{Tb}\right),
$$

where the expectation exists via Cauchy–Schwarz inequality and Assumption V above. Hence, $\Omega^{(1)} \to^p 0$.

We next focus on $\Omega^{(3)}$ as the approach for the other terms are similar. We start with the second moment and for cleaner notation, we re-define the indices as such:

$$
\begin{aligned}
E\left(\{\Omega^{(3)}\}^2\right) =& \frac{b^2}{T^4 \nu_0^2} \sum_{s_1 \neq s_2 \neq s_3} \sum_{s_4 \neq s_5 \neq s_6} \underbrace{k_{s_1,t} k_{s_3,t} k_{s_2,s_1} k_{s_4,t} k_{s_6,t} k_{s_5,s_4}}_{\equiv k_{s_1,\ldots,s_6}} \underbrace{m_{s_1,(a,b)} m_{s_4,(a,b)}}_{\leq C} \\
&\times E[X_{s_1,a} X_{s_1,i} X_{s_2,b} \nu_{s_2}^h X_{s_3,j} \nu_{s_3}^h \cdot X_{s_4,a} X_{s_4,i} X_{s_5,b} \nu_{s_5}^h X_{s_6,j} \nu_{s_6}^h].
\end{aligned}
$$

Without loss of generality, consider the case $s_1 < \ldots < s_6$, and let d_1 represent the first largest distance among $\Delta s_{z+1} = s_{z+1} - s_z$ for $z = 1, \ldots, 5$. Similar to the strategy in Atak et al. (2025), we consider the subcase where $d_1 = \Delta s_2$ (i.e. the first gap is the largest), then by an application of Lemma $A.1$ and Assumption V, we have for a $\varrho > 0$ and $R = 2(1+\varrho)$:

$$\left|E\left(\{\Omega^{(3)}\}^2\right)\right| \leq \frac{Cb^2}{T^4\nu_0^2} \sum_{\substack{s_1<\ldots<s_6 \\ d_1=\Delta s_2}} k_{s_1,\ldots,s_6} \left|E[X_{s_1,a}X_{s_1,i}]E[X_{s_2,b}\nu_{s_2}^h X_{s_3,j}\nu_{s_3}^h \cdot X_{s_4,a}X_{s_4,i}X_{s_5,b}\nu_{s_5}^h X_{s_6,j}\nu_{s_6}^h]\right|$$

$$+ \frac{2^{\frac{1}{R-1}}Cb^2}{T^4\nu_0^2} \sum_{s_1=1}^{T-5} \sum_{d_1=1}^{T-5-s_1+1} \sum_{s_3=s_1+d_1+1}^{s_1+2d_1} \sum_{s_4=s_3+1}^{s_3+d_1} \sum_{s_5=s_4+1}^{s_4+d_1} \sum_{s_6=s_5+1}^{\min\{s_5+d_1,T\}} k_{s_1,\ldots,s_6} \widetilde{\tau}_{d_1}^{\frac{R-2}{R-1}} \|X_{s_1,a}X_{s_1,i}\|_R^{\frac{R-1}{R}}$$

$$\times \|X_{s_2,b}\nu_{s_2}^h X_{s_3,j}\nu_{s_3}^h X_{s_4,a}X_{s_4,i}X_{s_5,b}\nu_{s_5}^h X_{s_6,j}\nu_{s_6}^h\|_R$$

$$\equiv O_1 + O_2$$

Even though we do not have an m.d.s. assumption on the error terms, we can show (although tediously) that the O_1 is $o(1)$ through repeated use of the mixing inequality. For O_2 note that by Assumption V, $\|X_{s_1,a}X_{s_1,i}\|_R$ is bounded and by repeated Hölder's inequality we have:

$$\|X_{s_2,b}\nu_{s_2}^h X_{s_3,j}\nu_{s_3}^h X_{s_4,a}X_{s_4,i}X_{s_5,b}\nu_{s_5}^h X_{s_6,j}\nu_{s_6}^h\|_{2(1+\varrho)}$$

$$\leq \|X_{s_2,b}\nu_{s_2}^h X_{s_3,j}\nu_{s_3}^h\|_{4(1+\varrho)} \cdot \|X_{s_4,a}X_{s_4,i}\|_{8(1+\varrho)}\|X_{s_5,b}\nu_{s_5}^h X_{s_6,j}\nu_{s_6}^h\|_{8(1+\varrho)} < \infty.$$

Next, $k_{s_1,\ldots,s_6} = O(1/b^6)\cdot K(\frac{s_1-t}{Tb})$, and hence for some constants $c_1, c_2 > 0$

$$O_2 \leq \frac{c_1}{(Tb)^4\nu_0^2} \sum_{s=1}^{T} K\left(\frac{s-t}{Tb}\right) \sum_{d_1=1}^{\infty} d_1^4 \widetilde{\tau}_{d_1}^{\frac{R-2}{R-1}} \leq \frac{c_1}{(Tb)^4\nu_0^2} \sum_{s=1}^{T} K\left(\frac{s-t}{Tb}\right) c_2 = O\left(\frac{1}{(Tb)^3}\right) = o(1),$$

where the second inequality follows from the condition on the mixing coefficient.

This can be generalized to the other subcases where $d_1 = \Delta s_{z+1}$ for $z = 2, ..., 5$. For example, when $z = 2$, the analogous summation for O_2 can be written as:

$$\frac{2^{\frac{1}{R-1}}Cb^2}{T^4\nu_0^2} \sum_{s_2=2}^{T-4} \sum_{d_1=2}^{T-s_2-3} \sum_{s_1=\max\{s_2-d_1+1,1\}}^{s_2-1} \sum_{s_4=s_2+d_1+1}^{s_2+2d_1} \sum_{s_5=s_4+1}^{s_4+d_1} \sum_{s_6=s_5+1}^{\min\{s_5+d_1,T\}} k_{s_1,\ldots,s_6} \widetilde{\tau}_{d_1}^{\frac{R-2}{R-1}} \|X_{s_1,a}X_{s_1,i}X_{s_2,b}\nu_{s_2}^h\|_R^{\frac{R-1}{R}}$$

$$\times \|X_{s_3,j}\nu_{s_3}^h X_{s_4,a}X_{s_4,i}X_{s_5,b}\nu_{s_5}^h X_{s_6,j}\nu_{s_6}^h\|_R,$$

which can also be shown to be $O(1/(Tb)^3)$ as previously.

Hence, we conclude that $\Omega_{1t} = o_p(1)$. For Ω_{2t}, in light of the result in Lemma $A.3$, we just need to show that $E(\Omega_{2t}^2) \to 0$ to invoke Chebyshev's inequality. The proof strategy for this is repetitive and very similar to our derivation of $E((\Omega^{(3)})^2)$. □

Appendix A.2. High-dimensional

Before starting the proof of Theorem 1, we require the following lemmas (whose proofs are postponed to Appendix B):

Lemma A.5. *Under the conditions of Theorem* 1*, with probability at least* $1 - Q_T^*$ *where* $Q_T^* \to 0$ *as* $T \to \infty$*, we have for any* $v \in \mathbb{R}^{2(m-1)}$ *such that the* $|\{j : v_j \neq 0\}| \leq s_T$*,*

$$\kappa^* \|v\|_2^2 \leq \frac{s_T}{Tb} \|Z_t v\|_{K_t}^2,$$

where $\kappa^* > 0$ *and* $\|\cdot\|_{K_t}, Z_t$ *and* K_t *are defined in Assumption H.2(ii).*

Lemma A.6. *Under the conditions of Theorem* 1*,*

(i) we have

$$P\left(\max_{1\leq j\leq 2(m-1)} \left|\frac{1}{\sqrt{Tb}} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) \tilde{z}_{s,t,j} e_s^\star\right| \leq c\sqrt{\log\left(\frac{8m_T}{\delta_1}\right)}\right) \geq 1 - \delta_1$$

where $\tilde{z}_{s,t,j}$ *is the jth element of* $\tilde{z}_{s,t} = (z_s^\top, z_s^\top(s-t)/Tb)^\top$*,* $e_s^\star$ *refers to either* e_{1s} *or* ν_s^h*,* $0 < \delta < 1$ *and* $c > 0$*;*

(ii) Furthermore,

$$P\Bigg(\max_{1\leq j\leq 2(m-1)} \left|\frac{1}{Tb} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) \left(Y_s^{(i)} - a_{0,t}^{o,h(i)\top} z_s - b a_{1,t}^{o,h(i)\top} \left(\frac{s-t}{Tb}\right) z_s\right) \tilde{z}_{s,t,j}\right| \leq c\left[\left(\frac{m_T}{\delta_2 (Tb)^{\varkappa-1}}\right)^{1/\varkappa} \vee \sqrt{\frac{\log(24m_T/\delta_2)}{Tb}} \vee \frac{m_T}{\delta_2 (Tb)^{1/2}}\right]\Bigg) \geq 1 - \delta_2,$$

where $Y_s^{(i)}$ *is a placeholder that refers to either* $Y_s^{(1)} = \varepsilon_s$ *or* $Y_s^{(2)} = y_{s+h}$ *and* $a_{j,t}^{o,h(1)}$ *refers to the true parameters in the unpenalized regression model in* (14) *for* $j = 0, 1$ *while* $a_{j,t}^{o,h(2)}$ *refers to the corresponding parameters in* (15)*. Here,* $\varkappa = ((\varphi^* + 1)R^* - 1)/(\varphi^* + R^* - 1) > 2$ *and* $R^* > 2$*.*

Lemma A.7. *Let* $\check{\theta}_t^{h(i)} \in \mathbb{R}^{2(m_T-1)}$ *be the local linear Lasso estimator to* (14) *if* $i = 1$ *and to* (15) *if* $i = 2$*. Then we have, under the conditions of Theorem* 1*, and with probability at least* $1 - Q_T - \delta_2$*:*

$$\|\check{\theta}_t^{h(i)} - \theta_t^{o,h(i)}\|_2 \leq \frac{C}{\kappa}\sqrt{s_T}\lambda, \text{ and } (Tb)^{-1/2}\|Z_t(\check{\theta}_t^{h(i)} - \theta_t^{o,h(i)})\|_{K_t} \leq \frac{C}{\sqrt{\kappa}}\sqrt{s_T}\lambda,$$

where $\theta_t^{o,h(i)}$ *refers to the corresponding true parameters,* κ *is from assumption H.2(ii),* λ *is from Assumption H.1(ii), and* Q_T *is from Assumption H.2.*

Note that we will let both $\delta_1 = \delta_1(T) \to 0$ and $\delta_2 \equiv \delta_2(T) \to 0$ as $T \to \infty$ slowly. This can be done for example by setting it to be $1/\log(Tb)$.

Proof of Theorem 1

Our proof strategy is similar to Belloni et al. (2014) and Hecq et al. (2023) but is more complicated due to the nonparametric estimation. Recall that our model is (11) and the post-double selection estimator is given in (16).

We first begin with some definitions. Let $\widetilde{\varepsilon}_{s,t} = (\varepsilon_s, ((s-t)/Tb)\varepsilon_s)^\top$ and $\widetilde{z}_{s,t} = (z_s^\top, ((s-t)/Tb)z_s^\top)^\top$. We obtain the data matrices $\widetilde{\varepsilon}_t$ and Z_t by stacking the vectors over the time sample which results in a $(T-h)\times 2$ vector and $(T-h)\times 2(m_T-1)$ matrix respectively. Let $\mathcal{S}$ be the index set of the variables selected in the double selection procedure (i.e. $I_1 \cup I_2$) and their gradient terms (the interaction with time). Then, label the sub-matrix of selected variables (corresponding to the columns of Z_t whose indices are in $\mathcal{S}$) as $Z_t^{\mathcal{S}}$. For an arbitrary matrix Z, denote the (weighted) projection matrix as $\mathcal{P}^W(Z) = Z(Z^\top W Z)^{-1}Z^\top W$ and the (weighted) annihilator matrix $\mathcal{M}^W(Z) = I - \mathcal{P}^W(Z)$. Note that $\mathcal{M}^W(\cdot)$ is no longer symmetric, but still idempotent. Construct the following $1 \times (T-h)$ vector:

$$A_t = e_1^{0\top}(\widetilde{\varepsilon}_t^\top b^{-1} K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}})\widetilde{\varepsilon}_t)^{-1}\widetilde{\varepsilon}_t^\top b^{-1} K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}}),$$

where e_1^0 is a 2×1 vector with 1 in the first position and 0 in its second and K_t is a $(T-h)\times(T-h)$ diagonal matrix with $\{K((s-t)/(Tb))\}_{s=1}^{T-h}$ as the diagonal elements.

Then our local linear (partially) partitioned regression estimator of the impulse response at time t for horizon h is given by:

$$\sqrt{Tb}(\check{\beta}_t^h - \beta_t^h) = \sqrt{Tb}A_t[Z_t\theta_{-e,t}^h + \nu^h] + \sqrt{Tb}\frac{b^2}{2}A_t Q_t \theta''^h(t/T) + \sqrt{Tb}A_t r_{s,t}$$

where $\theta_{-e,t}^h = (\widetilde{\vartheta}^h(t/T)^\top, b\widetilde{\vartheta}'^h(t/T)^\top)^\top$, Q_t is the stacked matrix of $(\varepsilon_s(\frac{(s-t)}{Tb})^2, z_s^\top(\frac{(s-t)}{Tb})^2)$, $\theta''^h(t/T)$ is the second derivative of all the coefficients of the model, and $r_{s,t}$ is the Taylor remainder. Rearranging terms around,

$$\sqrt{Tb}\left(\check{\beta}_t - \beta_t^h - \frac{b^2}{2}A_t Q_t \theta''^h(t/T)\right) = \sqrt{Tb}A_t[Z_t\theta_{-e,t}^h + \nu^h] + \sqrt{Tb}A_t r_{s,t} \equiv I + \sqrt{Tb}r_{s,t}.$$

The remainder term is of smaller order ($r_{s,t} = o(b^2)$) and thus we focus on the leading terms. We have:

$$I = e_1^\top \underbrace{(\widetilde{\varepsilon}_t^\top K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}})\widetilde{\varepsilon}_t/Tb)^{-1}}_{\equiv I_N^{-1}} \underbrace{\widetilde{\varepsilon}_t^\top K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}})[Z_t\theta_{-e,t}^h + \nu^h]/\sqrt{Tb}}_{\equiv I_D}.$$

We start with I_D:

$$I_D = \widetilde{\varepsilon}_t^\top K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}})Z_t\theta_{-e,t}^h/\sqrt{Tb} + \widetilde{\varepsilon}_t^\top K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}})\nu^h/\sqrt{Tb} \equiv I_{D1} + I_{D2}.$$

Note that

$$\underset{(2\times 1)}{I_{D1}} = \begin{bmatrix} \varepsilon^\top K_t \mathcal{M}^{K_t}\left(Z_t^S\right) Z_t\theta_{-e,t}^h/\sqrt{Tb} \\ (d_t \circ \varepsilon)^\top K_t \mathcal{M}^{K_t}\left(Z_t^S\right) Z_t\theta_{-e,t}^h/\sqrt{Tb} \end{bmatrix}, \tag{A.2}$$

where $\varepsilon = (\varepsilon_1, ..., \varepsilon_{T-h})^\top$, $d_t = (\frac{1-t}{Tb}, ..., \frac{T-h-t}{Tb})^\top$, and $\circ$ denotes element-wise multiplication.

Now, for the reduced form equation,

$$\varepsilon_s = \vartheta^{(1)\top}(t/T)z_s + \vartheta'^{(1)\top}(t/T)(\frac{s-t}{T})z_s + \frac{b^2}{2}\vartheta''^{(1)\top}(t/T)(\frac{s-t}{Tb})^2 z_s + \tilde{r}_{s,t} + e_{1t}, \tag{A.3}$$

where $\tilde{r}_{s,t}$ is again the $o(b^2)$ Taylor remainder. Plug (A.3) into (A.2). We focus on the first element since the approach for the second element is the same:

$$\begin{aligned}&(Tb)^{-1/2}[\gamma_t^\top Z_t^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\theta^h_{-e,t} + \eta_t^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\theta^h_{-e,t} + e_1^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\theta^h_{-e,t}]\\ &\equiv I_{D11} + I_{D12} + I_{D13}.\end{aligned}$$

where $\gamma_t = (\vartheta^{(1)\top}(t/T), b\vartheta'^{(1)\top}(t/T))^\top$ and η_t contains terms related to the second-order derivative and the remainder which are smaller by an order of b^2. Hence, we focus on the first and last terms. For I_{D11} by idempotence of $\mathcal{M}^{K_t}(\cdot)$:

$$|I_{D11}| \le \sqrt{Tb}\|\mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\gamma_t/\sqrt{Tb}\|_{K_t}\|\mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\theta^h_{-e,t}/\sqrt{Tb}\|_{K_t} \equiv \sqrt{Tb}(I_{D11,1}\cdot I_{D11,2}),$$

where $\|v\|_{K_t} = \sqrt{v^\top K_t v}$. Let γ^* be the solution to the unweighted noiseless problem $\min_{\gamma:\gamma_j=0 \text{ for } j\notin\mathcal{S}} \|Z_t\gamma_t - Z_t\gamma\|_2$. Furthermore, recall that I_1 is the index set of selected (level) terms from (14). Let $I_1^\star$ to be index set containing I_1 and their associated gradient terms. Note that $I_1^\star \subseteq \mathcal{S}$, then for conformable vectors v we have $\|\mathcal{M}^{K_t}(Z_t^S)v\|_{K_t} \le \|\mathcal{M}^{K_t}(Z_t^{I_1^\star})v\|_{K_t}$. By construction of the weighted annihilator matrix,

$$\begin{aligned}I_{D11,1} &\le \|\mathcal{M}^{K_t}\left(Z_t^{I_1^\star}\right) Z_t\gamma_t/\sqrt{Tb}\|_{K_t} = \min_{\gamma:\gamma_j=0 \text{ for } j\notin\mathcal{S}} \|Z_t\gamma_t - Z_t\gamma\|_{K_t}/\sqrt{Tb} \le \|Z_t\gamma_t - Z_t\gamma^*\|_{K_t}/\sqrt{Tb}\\ &\le \|Z_t(\gamma_t - \hat{\gamma})\|_{K_t}/\sqrt{Tb} \le \frac{C}{\sqrt{\kappa}}\sqrt{s_T}\lambda,\end{aligned}$$

where $\hat{\gamma}$ is the Lasso estimator from (14). The penultimate inequality is due to the construction of γ^* and the last inequality is from Lemma $A.7$ which holds with probability $1-\delta_2-Q_T$.

For $I_{D11,2}$, note that $\vartheta^{h(2)}(t/T) = \beta^h(t/T)\vartheta^{(1)}(t/T) + \widetilde{\vartheta}^h(t/T)$, then

$$\theta^h_{-e,t} = \begin{bmatrix}\widetilde{\vartheta}^h(t/T)\\ b\widetilde{\vartheta}'^h(t/T)\end{bmatrix} = \begin{bmatrix}\vartheta^{h(2)(t/T)}\\ b\vartheta'^{h(2)(t/T)}\end{bmatrix} - \begin{bmatrix}\beta^h(t/T)\vartheta^{(1)}(t/T)\\ \beta'^h(t/T)b\vartheta'^{(1)}(t/T)\end{bmatrix} \equiv a_t^h - (b_t^h \circ \gamma_t). \tag{A.4}$$

So,

$$I_{D11,2} \le \|\mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t a_t^h/\sqrt{Tb}\|_{K_t} + \|\mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t\gamma_t/\sqrt{Tb}\|_{K_t}\|b_t^h\|_{K_t}.$$

The second term is $I_{D11,1}$ while the derivation of the first term is analogous to the derivation of $I_{D11,1}$ with the main difference being the use of the index set $I_2^\star$ instead. Hence, the first term is upper bounded by $\|Z_t(a_t^h - \hat{a}_t^h)\|_{K_t}/\sqrt{Tb}$ where $\hat{a}_t^h$ is the Lasso estimator of (15) which is bounded with high probability by $\frac{C}{\sqrt{\kappa}}\sqrt{s_T}\lambda$.

Next we consider I_{D13}, by (A.4):

$$|I_{D13}| \leq |e_1^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t a_t^h/\sqrt{Tb}| + |e_1^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right) Z_t (b_t^h \circ \gamma_t)/\sqrt{Tb}| \equiv I_{D13,1} + I_{D13,2}$$

Since the monotonicity of the annihilator matrix does not necessarily carry over to the ℓ_1 case, define $\check{a}_{\mathcal{S}} = \operatorname{argmin}_{a:a_j=0 \text{ for } j\notin\mathcal{S}} \|Z_t a_t^h - Z_t a\|_{K_t}^2$ and likewise $\check{\gamma}_{\mathcal{S}} = \operatorname{argmin}_{\gamma:\gamma_j=0 \text{ for } j\notin\mathcal{S}} \|Z_t\gamma_t - Z_t\gamma\|_{K_t}^2$. Then,

$$I_{D13,1} = |e_1^\top K_t Z_t(\check{a}_{\mathcal{S}} - a_t^h)/\sqrt{Tb}| \leq \|\check{a}_{\mathcal{S}} - a_t^h\|_1 \|e_1^\top K_t Z_t/\sqrt{Tb}\|_\infty.$$

Note that by Lemma $A.5$, with probability $1 - Q_T^*$ we get $\|\check{a}_{\mathcal{S}} - a_t^h\|_1 = \|\check{a}_{\mathcal{S}} - a_t^h\|_2 \leq \frac{\sqrt{s_T}}{\sqrt{\kappa^*}\sqrt{Tb}}\|Z_t(\check{a}_{\mathcal{S}} - a_t^h)\|_{K_t} \leq \frac{\sqrt{s_T}}{\sqrt{\kappa^*}\sqrt{Tb}}\|Z_t(\hat{a}_t^h - a_t^h)\|_{K_t}$ where $\hat{a}_t^h$ is the lasso estimator. And by Lemma $A.7$ we conclude that $\|\check{a}_{\mathcal{S}} - a_t^h\|_1 \leq C s_T \lambda$. By Lemma $A.6$ with probability at least $1-\delta_1$, $\|e_1^\top K_t Z_t/\sqrt{Tb}\|_\infty \leq \varsigma_T$ where $\varsigma_T = c\sqrt{\frac{\log(8m_T)}{\delta_1}}$. Hence, with high probability,

$$I_{D13,1} \leq C\varsigma_T s_T \lambda \equiv \varrho_T.$$

Similarly,

$$I_{D13,2} \leq \left\|b_t^h\right\|_\infty \cdot \left\|Z_t^\top K_t e_1/\sqrt{Tb}\right\|_\infty \cdot \|\gamma_t - \check{\gamma}_{\mathcal{S}}\|_1 \leq \varrho_T.$$

Now we study I_{D2} which has a similar expression:

$$\underset{(2\times1)}{I_{D2}} = \begin{bmatrix} \varepsilon^\top K_t \mathcal{M}\left(Z_t^{\mathcal{S}}\right)\nu^h/\sqrt{Tb} \\ (d_t\circ\varepsilon)^\top K_t \mathcal{M}\left(Z_t^{\mathcal{S}}\right)\nu^h/\sqrt{Tb} \end{bmatrix}.$$

Again we focus on the first element:

$$\begin{aligned} &(Tb)^{-1/2}[\gamma_t^\top Z_t^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right)\nu^h + \eta_t^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right)\nu^h + e_1^\top K_t \mathcal{M}^{K_t}\left(Z_t^{\mathcal{S}}\right)\nu^h] \\ &\equiv I_{D21} + I_{D22} + I_{D23}. \end{aligned}$$

We start with I_{D21},

$$|I_{D21}| = |(\check{\gamma}_{\mathcal{S}} - \gamma_t)^\top Z_t^\top K_t \nu^h/\sqrt{Tb}| \leq \|\check{\gamma}_{\mathcal{S}} - \gamma_t\|_1 \|Z_t^\top K_t \nu^h/\sqrt{Tb}\|_\infty \leq \varrho_T.$$

For I_{D23}, we have:

$$I_{D23} = e_1^\top K_t \nu^h/\sqrt{Tb} - e_1^\top K_t \mathcal{P}^{K_t}(Z_t^{\mathcal{S}})\nu^h/\sqrt{Tb} \equiv I_{D23,1} + I_{D23,2}.$$

We start with the second term:

$$\begin{aligned} |I_{D23,2}| &\leq |e_1^\top K_t Z_t^{\mathcal{S}}(Z_t^{\mathcal{S}\top} K_t Z_t^{\mathcal{S}})^{-1} Z_t^{\mathcal{S}\top} K_t \nu^h/\sqrt{Tb}| \\ &\leq \|Z_t^{\mathcal{S}\top} K_t \nu^h/\sqrt{Tb}\|_\infty \cdot s_T \|(Z_t^{\mathcal{S}\top} K_t Z_t^{\mathcal{S}}/Tb)^{-1}\|_2 \|Z_t^{\mathcal{S}\top} K_t e_1/\sqrt{Tb}\|_\infty/\sqrt{Tb}. \end{aligned}$$

Here we can apply Lemma $A.6$ twice since $\|Z_t^{\mathcal{S}\top} K_t \nu^h\|_\infty \leq \|Z_t^\top K_t \nu^h\|_\infty$ and likewise for the term with e_1. Furthermore, with high probability, $\|(Z_t^{\mathcal{S}\top} K_t Z_t^{\mathcal{S}}/Tb)^{-1}\|_2$ is bounded. So $|I_{D23,2}| \leq C\varsigma_T^2 s_T/\sqrt{Tb}$, which goes to 0 based on our rate assumptions.

Hence our leading term is $I_{D23,1}$.

For I_N^{-1} we can apply the same approach to show that the leading term is given by $(\tilde{\varepsilon}_t^\top K_t \tilde{\varepsilon}_t/Tb)^{-1}$. The procedure is repetitive and we thus omit it.

For the bias term, we write

$$\frac{b^2}{2} A_t Q_t \theta''^h(t/T) = \frac{b^2}{2} e_1^{0\top} I_N^{-1} \underbrace{[\tilde{\varepsilon}_t^\top K_t \mathcal{M}^{K_t}(Z_t^{\mathcal{S}}) Q_t/Tb]}_{I_B} \theta''^h(t/T).$$

For notational convenience, let $\mathcal{M}^{K_t}(Z_t^{\mathcal{S}}) \equiv \mathcal{M}^{K_t}$, $D_t = diag(d_{s,t})$ where $d_s = \frac{s-t}{Tb}$, then $\tilde{\varepsilon}_t = \begin{bmatrix} \varepsilon & D\varepsilon \end{bmatrix}$ and $Q = \begin{bmatrix} D^2\varepsilon & D^2 z \end{bmatrix}$. The matrix expression is then given as

$$I_B = \frac{1}{Tb} \begin{bmatrix} \varepsilon^\top K_t \mathcal{M}^{K_t} D^2 \varepsilon & \varepsilon^\top K_t \mathcal{M}^{K_t} D^2 z \\ (D\varepsilon)^\top K_t \mathcal{M}^{K_t} D^2 \varepsilon & (D\varepsilon)^\top K_t \mathcal{M}^{K_t} D^2 z \end{bmatrix}.$$

We focus on the $(1,1)$ block. Since $\mathcal{M}^{K_t} = I - \mathcal{P}^{K_t}$, we have

$$\frac{1}{Tb}\varepsilon^\top K_t \mathcal{M}^{K_t} D^2 \varepsilon = \frac{1}{Tb}\varepsilon^\top K_t D^2 \varepsilon - \left(\frac{1}{Tb}\varepsilon^\top K_t Z_t^{\mathcal{S}}\right)\left(\frac{1}{Tb}\left(Z_t^{\mathcal{S}}\right)^\top K_t Z_t^{\mathcal{S}}\right)^{-1}\left(\frac{1}{Tb}\left(Z_t^{\mathcal{S}}\right)^\top K_t D^2 \varepsilon\right).$$

By Lemma $A.2$, $\frac{1}{Tb}\varepsilon^\top K_t D^2 \varepsilon = \mu_2 E[\varepsilon_t^2] + o_p(1)$. Let the variables selected in $\mathcal{S}$ be $z_t^{\mathcal{S}}$, then the second term converges to $\mu_2 E(\varepsilon_t z_t^{\mathcal{S}}) E(z_t^{\mathcal{S}} z_t^{\mathcal{S}\top}) E(z_t^{\mathcal{S}\top}\varepsilon_t)$. The derivation for $(1,2)$ is similar. For the $(2,1)$ and $(2,2)$ block, they converges to 0 since $\mu_3 = 0$ given the symmetric kernels. As explained above the leading order of I_N^{-1} converges to

$$\begin{bmatrix} E(\varepsilon_t^2) & 0 \\ 0 & \mu_2 E(\varepsilon_t^2) \end{bmatrix}.$$

Therefore, the entire bias term is $O(b^2)$.

Finally, we can then apply a variant of Lemma $A.4$ to $I_{D23,1}$ to obtain our result. $\square$

Appendix B. Proofs of Lemmas

Proof of Lemma $A.2$ We focus on $j=0$, as the extension to $j>0$ is similar.

$$E\left[T^{-1}\sum_{s=1}^{T-h}X_sX_s^\top k_{s,t}\right]=T^{-1}\sum_{s=1}^{T-h}M(t/T)k_{s,t}+o(1)\quad\text{(by Lipschitzness of } M(\cdot))$$
$$=M(t/T)\underbrace{\int K(u)du}_{=1}+o(1)\quad\text{(by Riemann sum approximation).}$$

Note that the second term is $o(1)$ because it is smaller by an order of b. For variance, consider the (i,j) element of S_0:

$$Var([S_0(t/T)]_{(i,j)})=(Tb)^{-2}\sum_{s=1}^{T-h}Var(X_{si}X_{sj})K^2\left(\frac{s-t}{Tb}\right)$$
$$+2(Tb)^{-2}\sum_{1\le s<l\le T-h}Cov\left(X_{si}X_{sj}K(\frac{s-t}{Tb}),X_{li}X_{lj}K(\frac{l-t}{Tb})\right).$$

For the first term,

$$(Tb)^{-2}\sum_{s=1}^{T-h}Var(X_{si}X_{sj})K^2\left(\frac{s-t}{Tb}\right)\le E[(X_{ti}X_{tj})^2](Tb)^{-2}\sum_{s=1}^{T-h}K^2\left(\frac{s-t}{Tb}\right)+o(1)=O\left(\frac{1}{Tb}\right),$$

where we have again used the Lipschitz condition and the Riemann sum approximation to get $\int K^2(u)du\le C<\infty$ for some $C>0$. Next, let $k=l-s$, then the second term simplifies to

$$\frac{2}{(Tb)^2}\sum_{s=1}^{T-h}\sum_{k=1}^{T-s-h}K\left(\frac{s-t}{Tb}\right)K\left(\frac{s+k-t}{Tb}\right)\cdot\text{Cov}\left(X_{si}X_{sj},X_{s+k,i}X_{s+k,j}\right)$$
$$\le\frac{2C}{(Tb)^2}\sum_{s=1}^{T-h}K\left(\frac{s-t}{Tb}\right)\sum_{k=1}^{T-s-h}K\left(\frac{s+k-t}{Tb}\right)\tau_k^{\frac{\tilde{R}-2}{\tilde{R}-1}}\quad\text{(by Lemma } A.1),$$

where $\tilde{R}\in(2,q/2]$. Note that $K((s-t)/Tb)$ is non-zero if and only if $|s-t|\le Tb$, let $\mathcal{S}_t$ represent the set of these indices and notice that the cardinality of $\mathcal{S}_t$ is $O(Tb)$. Next, we use the finite bounds on the kernels to arrive at

$$\frac{2C}{Tb}\sum_{k=1}^{T-s-h}\tau_k^{\frac{\tilde{R}-2}{\tilde{R}-1}}\le\frac{2C}{Tb}\sum_{k=1}^{\infty}k^{-\varphi\frac{\tilde{R}-2}{\tilde{R}-1}}=O(1/(Tb)).$$

The final equality requires greater exposition. By assumption A.3, $\tau_k=O(k^{-\varphi})$ where $\varphi>(q-2)/(q-4)$ and $q>4$. For summability, we require $\varphi\frac{\tilde{R}-2}{\tilde{R}-1}>1$. The worst case is for when $\frac{\tilde{R}-2}{\tilde{R}-1}$ is the smallest, which occurs when $\tilde{R}=q/2$, hence φ needs to satisfy $\varphi>\frac{q-2}{q-4}$, which is guaranteed by assumption A.3(i). Hence the infinite sum converges. Therefore, we conclude that $Var([S_0(t/T)]_{(i,j)})=o(1)$ and we get our result for the first part. For the second result on $b^{-j}D_j^h(t/T)$, the steps are almost exactly the same but with the additional consideration that $Q_{s,t}^h=o(b^2)$. □

Proof of Lemma $A.3$ Since we have a block diagonal setup, we focus on each block individually. Recall that $r^h(t/T) = (r_0^h(t/T)^\top, r_1^h(t/T)^\top)^\top$ and $\Gamma_{j,h}(t/T) = Cov(X_t\nu_t^h, X_{t+j}\nu_{t+j}^h)$. Additionally, we define $\Gamma_{s,l,h} = \Gamma(s/T, l/T) = Cov_h(X_s\nu_s^h, X_l\nu_l^h)$. Then,

$$TbVar(r_0^h(t/T)) = T^{-1}b\sum_{s=1}^{T-h}\Gamma_{0,h}(s/T)k_{s,t}^2 + 2T^{-1}b\sum_{1\le s<l\le T-h}\Gamma_{s,l,h}k_{s,t}k_{l,t} \equiv I_1 + I_2.$$

Likewise, by Lipschitz continuity and the Riemann sum approximation we have

$$I_1 = \Gamma_{0,h}(t/T)\int K^2(u)du + o(1).$$

Moving on to the covariances, we can capture pairs of indices that are asymptotically 'close'. Consider a sequence $g_T \to \infty$ such that $g_T/(Tb) \to 0$ and $g_T/\sqrt{T} \to 0$. We consider the following index sets: $\Theta_1 = \{(s,l) : 1 \le s-l \le g_T \text{ for } 1 \le s < l \le T-h\}$ and Θ_2 contains the pairs of indices (s,l) that obey $1 \le s < l \le T-h$ but are not in Θ_1. Then, we have

$$I_2 = 2T^{-1}b\sum_{(s,l)\in\Theta_1}\Gamma_{s,l,h}k_{s,t}k_{l,t} + 2T^{-1}b\sum_{(s,l)\in\Theta_2}\Gamma_{s,l,h}k_{s,t}k_{l,t} \equiv I_{21} + I_{22}.$$

We focus on I_{22}. Consider the (i,j) element of $\Gamma_{s,l,h}$, then we have

$$\begin{aligned}
[I_{22}]_{(i,j)} &= 2T^{-1}b\sum_{\substack{1\le s<l\le T-h\\ l-s>g_T}} Cov(X_{si}\nu_s^h, X_{lj}\nu_l^h)k_{s,t}k_{l,t} \\
&= 2T^{-1}b\sum_{k=g_T+1}^{T-h}\sum_{s=1}^{T-h-k} Cov\left(X_{si}\nu_s^h, X_{(s+k)j}\nu_{s+k}^h\right)k_{s,t}k_{s+k,t} \\
&\le 2T^{-1}b\sum_{k=g_T+1}^{T-h} C_0\tau_k^{*\frac{R-2}{R-1}}\sum_{s=1}^{T-h-k} k_{s,t}k_{s+k,t}, \qquad \text{(B.1)}
\end{aligned}$$

where we have applied Lemma $A.1$ with some $C_0 > 0$. Next, we use the boundedness of the kernel to get $k_{s+k,t} \le C_1/b$ for some $C_1 > 0$, and $T^{-1}\sum_{s=1}^{T-h-k} k_{s,t} \le T^{-1}\sum_{s=1}^{T-h} k_{s,t} \le C_2$. This yields

$$\text{(B.1)} \le C\sum_{k=g_T+1}^{T-h} k^{-\varphi^*\frac{R-2}{R-1}} \le C\frac{1}{a-1}g_T^{-(a-1)} = o(1)$$

where we have used assumption A.3 for the first inequality and an integral bound for the second. Here, $a = \varphi^*\frac{R-2}{R-1}$ and since $\varphi^* > \frac{R-1}{R-2}$ by construction, we have $a > 1$ and thus the sum converges to 0 as $g_T \to \infty$.

Next, for the (i,j) element of I_{21}:

$$
\begin{aligned}
[I_{21}]_{(i,j)} &= 2T^{-1}b \sum_{\substack{1\le s<l\le T-h \\ 1\le l-s\le g_T}} Cov(X_{si}\nu_s^h, X_{lj}\nu_l^h)k_{s,t}k_{l,t} \\
&= 2T^{-1}b \sum_{s=1}^{T-h-1} \sum_{1\le l-s\le g_T} Cov(X_{si}\nu_s^h, X_{lj}\nu_l^h)k_{s,t}k_{l,t} \\
&= 2T^{-1}b^2 \sum_{s=1}^{T-h-1} k_{s,t}^2 \sum_{1\le l-s\le g_T} Cov(X_{si}\nu_s^h, X_{lj}\nu_l^h) \\
&+ 2T^{-1}b^2 \sum_{s=1}^{T-h-1} \sum_{1\le l-s\le g_T} Cov(X_{si}\nu_s^h, X_{lj}\nu_l^h)(k_{s,t}k_{l,t} - k_{s,t}^2) \\
&= I_{211} + I_{212}.
\end{aligned}
$$

Note that as $T\to\infty$, we have $I_{211} \to 2\int K^2(u)du \sum_{k=1}^{\infty}[\Gamma_{k,h}(t/T)]_{(i,j)}$. Next, notice that by the Lipschitz requirement on the kernel, $|k_{l,t}-k_{s,t}| \le L\frac{1}{b}|\frac{l-s}{Tb}| \le L\frac{g_T}{Tb^2}$, then

$$
|I_{212}| \le Lg_T T^{-2} b^{-1} \sum_{s=1}^{T-h-1} k_{s,t} \left\{ \sum_{k=1}^{\infty} |[\Gamma_{k,h}(t/T)]_{(i,j)}| + g_T b \right\} = O\left(\frac{g_T}{Tb} + \frac{g_T^2}{T}\right) = o(1)
$$

where we have used Lemma $A.1$ to yield $\sum_{k=1}^{\infty}|[\Gamma_{k,h}(t/T)]_{(i,j)}| = O(1)$, and the Riemann sum approximation for $k_{s,t}$. The $o(1)$ comes from the construction of g_T. Therefore, $[I_{21}]_{(i,j)} \to 2\nu_0 \sum_{k=1}^{\infty}[\Gamma_{k,h}(t/T)]_{(i,j)}$. Together with the result for I_1 above, we conclude that

$$
TbVar(r_0^h(t/T)) \to \nu_0 \left\{ \Gamma_{0,h}(t/T) + 2\sum_{k=1}^{\infty} \Gamma_{k,h}(t/T) \right\} = \nu_0 \Omega_h(t/T).
$$

The same steps can be taken to show that

$$
TbVar(b^{-1}r_1^h(t/T)) \to \nu_2\Omega_h(t/T) \text{ and } TbCov(r_0^h(t/T), b^{-1}r_1^h(t/T)) \to \nu_1\Omega_h(t/T) = 0_{(m\times m)},
$$

since $\nu_1 = \int uK^2(u)du = 0$ due to symmetry. □

Proof of Lemma $A.4$

Before proceeding, we reproduce the necessary central limit theorem by Neumann (2013) here:

Theorem 2.1 of Neumann (2013) Let $(\xi_{T,t})_{t=1,\dots,T}$, $T\in\mathbb{N}$ be a triangular array of random variables with $E[\xi_{T,t}]=0$ and $\sum_{t=1}^{T} E[\xi_{T,t}^2] \le c_0$ for all T,t and some $c_0<\infty$. Assume that as $T\to\infty$, we have

$$
\sigma_T^2 \equiv Var(\xi_{T,1} + \dots + \xi_{T,T}) \to \sigma^2 \in [0,\infty) \tag{B.2}
$$

and

$$
\sum_{t=1}^{T} E[\xi_{T,t}^2 1_{\{|\xi_{T,t}|>\epsilon\}}] \to 0 \tag{B.3}
$$

for all $\epsilon > 0$. Next, assume that there exists a summable sequence $(\tau_k)_{k\in\mathbb{N}}$ such that for all $u \in \mathbb{N}$ and all indices $1 \leq s_1 \leq s_2 < ... < s_u < s_u + k = j_1 \leq j_2 \leq T$, the following covariances are upper-bounded as such: for all measurable functions $g : \mathbb{R}^u \to \mathbb{R}$ with $\|g\|_\infty = sup_{x\in\mathbb{R}^u}|g(x)| \leq 1$,

$$|Cov(g(\xi_{T,s_1}, ..., \xi_{T,s_u})\xi_{T,s_u}, \xi_{T,j_1})| \leq (E[\xi^2_{T,s_u}] + E[\xi^2_{T,j_1}] + T^{-1})\tau_k \quad \text{(B.4)}$$

and

$$|Cov(g(\xi_{T,s_1}, ..., \xi_{T,s_u}), \xi_{T,j_1}\xi_{T,j_2})| \leq (E[\xi^2_{T,j_1}] + E[\xi^2_{T,j_2}] + T^{-1})\tau_k. \quad \text{(B.5)}$$

Then we conclude that

$$\xi_{T,1} + ... + \xi_{T,T} \to^d N(0, \sigma^2).$$

□

The proof of Lemma $A.4$ therefore involves verifying the 4 conditions in (B.2)-(B.5). Note that we can rewrite

$$v^\top \sqrt{Th}\mathbf{B}^{-1} r^h(t/T) = v^\top \frac{1}{\sqrt{Tb}} \sum_{s=1}^{T-h} \nu_s^h K\left(\frac{s-t}{Tb}\right) \underbrace{\begin{bmatrix} X_S \\ \left(\frac{s-t}{Tb}\right) X_s \end{bmatrix}}_{\equiv Z_{s,t}},$$

where we have pre-multiplied the vector by a conformable unit vector v to apply Crámer-Wold. Note that $E[v^\top \frac{1}{Tb} \sum_{s=1}^{T-h} \nu_s^h K(\frac{s-t}{Tb}) Z_{s,t}] = 0$ by exogeneity.

For (B.2), (B.2) is guaranteed by the existence of the long-run variance in Lemma $A.1$.

Next to verify (B.3), we will show that

$$\lim_{T\to\infty} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) E\left[\left|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}\right|^2 1_{\{|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}| > \epsilon\}}\right] = 0.$$

Note that this is easy to verify given our moment conditions of X_s and ν_s^h:

$$\begin{aligned}
E\left[\left|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}\right|^2 1_{\{|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}| > \epsilon\}}\right] &= E\left[\left|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}\right|^2 \left(\frac{\epsilon}{\epsilon}\right)^{R-2} 1_{\{|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}| > \epsilon\}}\right] \\
&< E\left[\left|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}\right|^2 \left(\frac{|v^\top Z_{s,t} v_s^h / \sqrt{Tb}|}{\epsilon}\right)^{R-2} 1_{\{|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}| > \epsilon\}}\right] \\
&\leq E\left[\frac{|v^\top Z_{s,t} v_s^h|^R}{Tb} \left(\frac{|1/\sqrt{Tb}|}{\epsilon}\right)^{R-2}\right] \\
&= \frac{1}{Tb} Tb^{-(R-2)/2} \epsilon^{-(R-2)} E[|v^\top Z_{s,t} v_s^h|^R],
\end{aligned}$$

where $R = pq/(p+q)$ from assumption A.3. By Hölder's inequality:

$$E[|v^\top Z_{s,t} \nu_s^h|^R] \leq \|\nu_s^h\|_p^R \|Z_{s,t}\|_q^R \leq C < \infty$$

where the bound comes from their moment conditions. Therefore,

$$\sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) E\left[\left|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}\right|^2 1_{\{|\frac{v^\top Z_{s,t} v_s^h}{\sqrt{Tb}}|>\epsilon\}}\right] \leq \frac{(Tb)^{-(R-2)/2}}{\epsilon^{(R-2)}} \frac{C}{Tb} \sum_{s=1}^{T-h} K(\frac{s-t}{Tb})$$
$$= O((Tb)^{-(R-2)/2}) = o(1),$$

since $R > 2$.

Next, we need only to verify (B.4) since (B.5) follows similarly but relying on assumption A.3(ii). From Theorem 2.1 of Neumann (2013), we require the following notation: for any $u \in \mathbb{N}$, the indices $1 \leq s_1 < s_2 < ... < s_u < s_u + k \leq s_u + k' < T - h$ and $g : \mathbb{R}^u \to R$ with $\sup_{v \in \mathbb{R}^u} \|g(v)\| \leq 1$. Next, define the following notations:

$$Y_{s_u,t} = g(v^\top Z_{s_1,t}\nu_{s_1}^h/\sqrt{Tb}, v^\top Z_{s_2,t}\nu_{s_2}^h/\sqrt{Tb}, ..., v^\top Z_{s_u,t}\nu_{s_u}^h/\sqrt{Tb}) v^\top Z_{s_u,t}\nu_{s_u}^h$$
$$Z_{s_u+k,t} = v^\top Z_{s_u+k,t}\nu_{s_u+k}^h,$$

$Q_{|Y_{s_u,t}|}$ is the quantile function of $|Y_{s_u,t}|$, $G_{|Z_{s_u+k,t}|}$ is the generalized inverse of $z \mapsto \int_0^z Q_{|Z_{s_u+k,t}|}(x)dx$, the information set $\mathcal{F}_{-\infty}^{s_u} = \sigma(Z_{s_u,t}\nu_{s_u}^h, Z_{s_u-1,t}\nu_{s_u-1}^h, Z_{s_u-2,t}\nu_{s_u-2}^h, ...)$, and the L_1-mixingale type coefficient $\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t}) = \|E[Z_{s_u+k,t}|\mathcal{F}_{-\infty}^{s_u}] - E[Z_{s_u+k,t}]\|_1$. Hence, our equivalent of (B.4) is given by:

$$\begin{aligned}
\frac{1}{Tb}|Cov(Y_{s_u,t}, Z_{s_u+k,t})| &\leq \frac{2}{Tb}\int_0^{\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t})/2} Q_{|Y_{s_u,t}|} \circ G_{|Z_{s_u+k,t}|}(x)dx \\
&\leq \frac{2}{Tb}\int_0^{\|Z_{s_u+k,t}\|_1} 1_{\{x<\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t})/2\}} Q_{|Y_{s_u,t}|} \circ G_{|Z_{s_u+k,t}|}(x)dx \\
&\leq \frac{1}{Tb}\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t})^{\frac{R-2}{R-1}} \left(\int_0^{\|Z_{s_u+k,t}\|_1} [Q_{|Y_{s_u,t}|} \circ G_{|Z_{s_u+k,t}|}(x)]^{R-1} dx\right)^{1/(R-1)} \\
&\leq \frac{c}{Tb}\tau_k^{*\frac{R-2}{R-1}} \left[\left(\int_0^1 Q_{|Y_{s_u,t}|}^R(y)dy\right)^{\frac{R-1}{R}} \left(\int_0^1 Q_{|Z_{s_u+k,t}|}^R(y)dy\right)^{1/R}\right]^{1/(R-1)} \\
&= \frac{c}{Tb}\tau_k^{*\frac{R-2}{R-1}} \|Y_{s_u,t}\|_R \|Z_{s_u+k,t}\|_R^{\frac{1}{R-1}}.
\end{aligned}$$

We discuss this derivation in detail here. The first inequality is due to the covariance inequality (Proposition 5) of Dedecker and Doukhan (2003). For the second inequality, note that

$$\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t}) \leq \|E[Z_{s_u+k,t}|\mathcal{F}_{-\infty}^{s_u}]\|_1 + \|E[Z_{s_u+k,t}]\|_1 \leq 2\|Z_{s_u+k,t}\|_1$$

by the triangle inequality, and the following inequality is valid by Jensen's inequality and the law of iterated expectations for $\|E[Z_{s_u+k,t}|\mathcal{F}_{-\infty}^{s_u}]\|_1 \leq \|Z_{s_u+k,t}\|_1$ and $\|E[Z_{s_u+k,t}]\|_1 \leq \|Z_{s_u+k,t}\|_1$. The third inequality follows from Hölder's inequality for $\frac{R-2}{R-1} + \frac{1}{R-1} = 1$:

$$2\int_0^{\|Z_{s_u+k,t}\|_1} 1_{\{x<\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t})/2\}} Q_{|Y_{s_u,t}|} \circ G_{|Z_{s_u+k,t}|}(x)dx$$
$$\leq 2\left(\int_0^{\gamma(\mathcal{F}_{-\infty}^{s_u}, Z_{s_u+k,t})/2} 1dx\right)^{\frac{R-2}{R-1}} \left(\int_0^{\|Z_{s_u+k,t}\|_1} [Q_{|Y_{s_u,t}|} \circ G_{|Z_{s_u+k,t}|}(x)]^{R-1}dx\right)^{\frac{1}{R-1}}.$$

The first term integrates to the upper limit and we note that the relations (2.2.13) and (2.2.18) in Dedecker et al. (2007) imply that the L_1- mixingale coefficient is smaller than the τ-mixing coefficient. For the second term, it follows from a change of variables where we set $x = \int_0^y Q_{|Y_{s_u,t}|}(e)de$ such that $du/dy = Q_{|Y_{s_u,t}|}(y)$, and at the upper limit of the integral ($\|Z_{s_u+k,t}\|_1$) we have $y = 1$. Additionally, by construction of G, we have $G_{Z_{s_u+k,t}}(\int_0^y Q_{|Y_{s_u,t}|}(e)de) = y$. Subsequently, we apply Hölder's inequality again to split the second term. The final equality follows from the property of the quantile function: $\int_0^1 Q^R_{|Y_{s_u,t}|}(y)dy = E|Y^R_{s_u,t}|$ and $\int_0^1 Q^R_{|Z_{s_u+k,t}|}(y)dy = E|Z^R_{s_u+k,t}|$.

Next, recall that $\sup_{v\in\mathbb{R}^u} \|g(v)\| \leq 1$, then

$$\|Y_{s_u,t}\|_R \leq \|v^\top Z_{s_u,t}\nu^h_{s_u}\|_R = O(1),$$

where we have shown this bound earlier in the derivation of condition (B.3). The same can be said for $\|Z_{s_u+k,t}\|_2$. Then by assumption A.3, we have $\tau^*_k = O(k^{-\varphi^*})$ where $\varphi^* > (R-1)/(R-2)$ so $\varphi^* \cdot \frac{R-2}{R-1} > 1$ and $\tau_k^{*\frac{R-2}{R-1}}$ is summable and (B.4) is satisfied. The approach for (B.5) is more convenient because the first term in the covariance is given solely by $g(\cdot)$ which is uniformly bounded by 1 and therefore the quantile function will also be bounded by 1. With this observation, we can show that it will be $O((k'-k)^{-\tilde{\varphi}}/Tb)$ where $\tilde{\varphi} > 1$ is from assumption A.3(ii). The conditions are therefore satisfied and we can appeal to the central limit theorem. □

Proof of Lemma $A.5$

The proof relies on assumption H.2(i) and the proof strategy follows that of Lemma A.1 in Hecq et al. (2023). □

Proof of Lemma $A.6$

We start with part (i). Since $\{\tilde{z}_{s,t,j}e^\star_s\}$ is τ-mixing, by Theorem 3.1 of Babii et al. (2024), there exists constants $c_1, c_2 > 0$ such that we get

$$P\left(\max_{1\leq j\leq 2(m-1)}\left|\frac{1}{\sqrt{Tb}}\sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right)\tilde{z}_{s,t,j}e^\star_s\right| > u\right) \leq 2c_1 m_T \frac{(Tb)^{1-\kappa/2}}{u^\kappa} + 8m_T \exp\left(-\frac{c_2 u^2(Tb)}{B_T^2}\right),$$

where $B_T^2 = \max_j \sum_s \sum_k Cov(\tilde{z}_{s,t,j}e^\star_s, \tilde{z}_{k,t,j}e^\star_k)$. By Lemma C.2 of Chen and Maung (2025), $B_T^2 = O(Tb)$. Next we invert the probability, so that for any $\delta_1 \in (0,1)$,

$$2c_1 m_T \frac{(Tb)^{1-\kappa/2}}{u^\kappa} \leq \frac{\delta_1}{2},$$

and

$$u \geq C_1 \left(\frac{m_T}{\delta_1 (Tb)^{\kappa/2-1}} \right).$$

Given our assumptions on κ, it is possible to show that it is > 2, hence this fraction goes to 0 as $Tb \to \infty$. The dominating order comes from the second term because $B_T^2 = O(Tb)$:

$$u \geq C_2 \sqrt{\frac{\log(8m_T)}{\delta_1}},$$

which concludes the proof of part (i).

For part (ii),

$$\begin{aligned}
&P\left(\max_{1 \leq j \leq 2(m-1)} \left| \frac{1}{Tb} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) \left(Y_s^{(i)} - a_{0,t}^{o,h(i)\top} z_s - b a_{1,t}^{o,h(i)\top} \left(\frac{s-t}{Tb}\right) z_s \right) \tilde{z}_{s,t,j} \right| > u \right) \\
&\leq \sum_{j=1}^{2(m_T-1)} P\left(\left| \frac{1}{Tb} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) \underbrace{\left(Y_s^{(i)} - a_{0,s}^{o,h(i)\top} z_s \right)}_{=e_{1s} \text{ or } e_{2s}^h} \tilde{z}_{s,t,j} > u \right) \\
&+ \sum_{j=1}^{2(m_T-1)} P\left(\left| \frac{1}{Tb} \sum_{s=1}^{T-h} K\left(\frac{s-t}{Tb}\right) \underbrace{\left(a_{0,s}^{o,h(i)} - a_{0,t}^{o,h(i)} - b a_{1,t}^{o,h(i)} \left(\frac{s-t}{Tb}\right) \right)^\top z_s}_{\equiv R_{s,t,i}(z_s)} \tilde{z}_{s,t,j} \right| > u \right) \qquad \equiv P_1 + P_2.
\end{aligned}$$

By a second order Taylor expansion, $R_{s,t,i}(z_s)$ contains b^2. Furthermore, due to sparsity, we can show via Markov's inequality that the P_2 is of order $O(m_T s_T b^2)$. By assumption H.1(i) $s_T b^2 \propto (Tb)^{-1/2}$ which yields the third term in the probability in part (ii) of Lemma $A.6$ (i.e. this part is due to the local linear approximation).

For P_1, we can repeat the process by applying Theorem 3.1 of Babii et al. (2024) again. The rates are different now since we have $1/Tb$ instead of $1/\sqrt{Tb}$ as was the case in part (i) of the proof.

The proof is complete by inverting the probability, this time by setting the rates to be $\leq \delta_2/3$ for some $\delta_2 \in (0, 1)$. $\square$

Proof of Lemma $A.7$

For ease of notation, we focus on the estimator from (15) since the derivation carries over to the case with (14).

First, define the difference between the estimator $\check{\theta}_t^{h(2)} = (\check{a}_{0,t}^{h(2)\top}, b\check{a}_{1,t}^{h(2)\top})^\top$ and the true value for the level and gradient terms as:

$$d_{jt} = \check{a}_{0,t,j}^{h(2)} - a_{0,t,j}^{o,h(2)} \text{ and } \dot{d}_{jt} = b[\check{a}_{1,t,j}^{h(2)} - a_{1,t,j}^{o,h(2)}],$$

for $j = 1, ..., m_T - 1$.

It can then be shown that the local linear lasso error satisfies the following constraint:

$$\max\left\{\sum_{j=s_{(2),T}+1}^{m_T-1}|d_{jt}|,\sum_{j=s_{(2),T}+1}^{m_T-1}|\dot{d}_{jt}|\right\}\leq b\left(\sum_{j=1}^{s_{(2),T}}|d_{jt}|+\sum_{j=1}^{s_{(2),T}}|\dot{d}_{jt}|\right),$$

where $b=\max\{\lambda_1^{(2)}/\lambda_2^{(2)},\lambda_2^{(2)}/\lambda_1^{(2)}\}+\tilde{\eta}$ and $\tilde{\eta}>0$. This was shown by Li et al. (2015) for the varying-coefficient model and Chen and Maung (2023) for nonparametric time-varying forecast combinations. Hence we omit its derivation for brevity.

Define $d_t=(d_{1t},...,d_{m_T-1,t},\dot{d}_{1t},...,\dot{d}_{m_T-1,t})^\top$, then a key implication of this result is the following cone constraint:

$$\|d_t\|_1=\|d_t^{s(2),T}\|_1+\|d_t^{\neg s(2),T}\|_1\leq\|d_t^{s(2),T}\|_1+b(\|d_t^{s(2),T}\|_1)=(1+b)\|d_t^{s(2),T}\|_1.$$

where $d_t^{s(2),T}$ refers to the difference of the non-zero coefficients and $d_t^{\neg s(2),T}$ is the complement of that index set.

Next, by definition of the lasso estimator (as a minimizer) to (15) we have:

$$(Tb)^{-1}\|y^h-Z_t\check{\theta}_t^{h(2)}\|_{K_t}^2+\lambda_1^{(2)}\|\check{a}_{0,t}^{h(2)}\|_1+\lambda_2^{(2)}\|\check{a}_{1,t}^{h(2)}\|_1\leq(Tb)^{-1}\|y^h+Z_t\theta_t^{o,h(2)}\|_{K_t}^2+\lambda_1^{(2)}\|a_{0,t}^{o,h(2)}\|_1+\lambda_2^{(2)}\|a_{1,t}^{o,h(2)}\|_1,$$

where y^h is the vector of $\{y_{s+h}\}$, and Z_t and $\|\cdot\|_{K_t}$ are defined in assumption H.2(ii). Then, note that

$$\|y^h-Z_t\check{\theta}_t^{h(2)}\|_{K_t}^2-\|y^h+Z_t\theta_t^{o,h(2)}\|_{K_t}^2=\|Z_t(\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)})\|_{K_t}^2-2(y^h-Z_t\theta_t^{o,h(2)})^\top K_tZ_t(\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)}),$$

hence together with assumption H.1(ii) $\lambda_1^{(2)},\lambda_2^{(2)}\propto\lambda$:

$$(Tb)^{-1}\|Z_t(\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)})\|_{K_t}^2\leq2(Tb)^{-1}(y^h-Z_t\theta_t^{o,h(2)})^\top K_tZ_t(\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)})+\lambda[\|\check{\theta}_t^{h(2)}\|_1-\|\theta_t^{o,h(2)}\|_1].$$

Note that we have

$$(y^h-Z_t\theta_t^{o,h(2)})^\top K_tZ_t(\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)})\leq\|Z_t^\top K_t(y^h-Z_t\theta_t^{o,h(2)})\|_\infty\|\check{\theta}_t^{h(2)}-\theta_t^{o,h(2)}\|_1.$$

Define the event $\Lambda_T=\{(Tb)^{-1}\|Z_t^\top K_t(y^h-Z_t\theta_t^{o,h(2)})\|_\infty\leq\lambda\}$. Note that by Lemma $A.6$, we have $P(\Lambda_T)\geq1-\delta_2$. We will now assume that Λ_T holds and so:

$$(Tb)^{-1}\|Z_td_t\|_{K_t}^2\leq2\lambda\|d_t\|_1+\lambda\|d_t\|_1,\tag{B.6}$$

where we have used the reverse triangle inequality. By the cone constraint, we have $3\lambda\|d_t\|_1\leq3(1+b)\lambda\|d_t^{s_{(2),T}}\|_1$. Then by the restricted eigenvalue assumption in H.2(ii), we have

$$\kappa\|d_t\|_2^2\leq(Tb)^{-1}\|Z_td_t\|_{K_t}^2,$$

which holds with probability $1 - Q_T$. Hence

$$\kappa \|d_t\|_2^2 \leq 3(1+b)\lambda \|d_t^{s_{(2),T}}\|_1 \leq 3(1+b)\lambda \sqrt{s_{(2),T}} \|d_t\|_2.$$

So, we have $\|d_t\|_2 \leq \frac{3(1+b)}{\kappa}\sqrt{s_T}\lambda$.

Finally from (B.6), $(Tb)^{-1/2}\|Z_t d_t\|_{K_t} \leq \frac{C}{\kappa}\lambda\sqrt{s_T}$. □