Supplementary Material on "Testing Homogeneity in Panel Data Models with Interactive Fixed Effects"

Liangjun Su[a], Qihui Chen[b]

[a] School of Economics, Singapore Management University, Singapore

[b] Department of Economics, University of California, San Diego

January 11, 2013

This appendix provides proofs for all technical lemmas and Theorem 3.2 in the paper titled "Testing Homogeneity in Panel Data Models with Interactive Fixed Effects" by Liangjun Su and Qihui Chen

C Proofs of Technical Lemmas

Proof of Lemma A.2. We follow closely the proofs of Theorems 2.1 and 3.1 in Moon and Weidner (2010a, **MWa** hereafter) to study the asymptotic property of $\hat{\boldsymbol{\beta}}$ under $\mathbb{H}_1(\gamma_{NT})$. By allowing local deviations from the homogenous panel data models, the consistency of $\hat{\boldsymbol{\beta}}$ can be demonstrated as in **MWa.** Let $\mathbf{X}_0 = (\sqrt{NT}/\left\|\mathbf{e}\right\|)\mathbf{e}$, $\epsilon_0 \equiv \left\|\mathbf{e}\right\|/\sqrt{NT}$, and $\epsilon_k \equiv \beta_k^0 - \beta_k$ for $k = 1, ..., K$. Note that under $\mathbb{H}_1(\gamma_{NT})$, conditions (A.6) and (A.7) in **MWa** continue to hold for sufficiently large (N,T) as

$$\upsilon_{1NT} \equiv \sum_{k=1}^{K} \left|\beta_k^0 - \beta_k\right| \frac{\left\|\mathbf{X}_k\right\|}{\sqrt{NT}} + \frac{\left\|\mathbf{e}\right\|}{\sqrt{NT}} = o_P(1) + O_P\left(\delta_{NT}^{-1} + \gamma_{NT}\right) = o_P(1)$$

under Assumptions A.1(iii) and (vi) provided $\left\|\boldsymbol{\beta}^0 - \boldsymbol{\beta}\right\| = o_p(1)$. This enables us to apply Lemma A.1(iii) of **MWa** to obtain

$$\begin{aligned}
\mathcal{L}_{NT}(\boldsymbol{\beta}) &= \frac{1}{NT}\sum_{k_1=0}^{K}\sum_{k_2=0}^{K}\epsilon_{k_1}\epsilon_{k_2}L^{(2)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}\right) \\
&\quad + \frac{1}{NT}\sum_{k_1=0}^{K}\sum_{k_2=0}^{K}\sum_{k_3=0}^{K}\epsilon_{k_1}\epsilon_{k_2}\epsilon_{k_3}L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{X}_{k_3}\right) + O_P\left(\upsilon_{1NT}^4\right),
\end{aligned}$$

where for any integer $g \geq 1$,

$$L^{(g)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, ..., \mathbf{X}_{k_g}\right) = \frac{1}{g!} \sum_{\text{all } g! \text{ permutations of } (k_1,...,k_g)} \tilde{L}^{(g)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, ..., \mathbf{X}_{k_g}\right),$$

$$\tilde{L}^{(g)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, ..., \mathbf{X}_{k_g}\right) = \sum_{l=1}^{g} (-1)^{l+1} \sum_{\substack{v_1+v_2+\cdots+v_l=g \\ m_1+\cdots+m_{l+1}=l-1 \\ 2\geq v_j>1,\ m_j\geq 0}} \operatorname{tr}\left\{S^{(m_1)}\mathcal{T}_{k_1\ldots}^{(v_1)}S^{(m_2)}\cdots S^{(m_l)}\mathcal{T}_{\cdots k_g}^{(v_l)}S^{(m_{l+1})}\right\},$$

$S^{(0)} = -M_{\lambda^0}$, $S^{(m)} = \Phi_3^m$, $\mathcal{T}_k^{(1)} = \lambda^0 F^{0\prime}\mathbf{X}_k' + \mathbf{X}_k F^0 \lambda^{0\prime}$ for $k = 0, 1, ..., K$, and $\mathcal{T}_{k_1 k_2}^{(2)} = \mathbf{X}_{k_1}\mathbf{X}_{k_2}'$ for $k_1, k_2 = 0, 1, ..., K$. [The subscript indices in $\mathcal{T}_{k_1\ldots}^{(v_1)}$ or $\mathcal{T}_{\cdots k_g}^{(v_l)}$ may contain either one (e.g., k_1 or k_g) or two elements (e.g., (k_1, k_2) or (k_{g-1}, k_g)) depending on whether v_1 or v_l takes value 1 or 2.] By straightforward calculations, one verifies that $\tilde{L}^{(2)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}\right)$ =tr$(M_{\lambda^0}\mathbf{X}_{k_1}\ M_{F^0}\mathbf{X}_{k_2}' M_{\lambda^0})$ and $\tilde{L}^{(3)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{X}_{k_3}\right) = -\text{tr}(\Phi M_{\lambda^0}\mathbf{X}_{k_1}\Phi_1'\mathbf{X}_{k_2}M_{F^0}\mathbf{X}_{k_3}' + M_{\lambda^0}\mathbf{X}_{k_1}M_{F^0}\mathbf{X}_{k_2}'\Phi_1\mathbf{X}_{k_3}')$. It follows that

$$L^{(2)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}\right) = \operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_{k_1}M_{F^0}\mathbf{X}_{k_2}'M_{\lambda^0}\right) = \operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_{k_1}M_{F^0}\mathbf{X}_{k_2}'\right), \text{ and}$$

$$L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{X}_{k_3}\right) = -\frac{1}{3} \sum_{\text{all 6 permutations of } (k_1,k_2,k_3)} \operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_{k_1}M_{F^0}\mathbf{X}_{k_2}'\Phi_1\mathbf{X}_{k_3}'M_{\lambda^0}\right).$$

Furthermore, we have

$$\mathcal{L}_{NT}(\boldsymbol{\beta}) = \mathcal{L}_{NT}\left(\boldsymbol{\beta}^0\right) + L_{1NT}(\boldsymbol{\beta}) + L_{2NT}(\boldsymbol{\beta}) + R_{NT} + O_P\left(\upsilon_{1NT}^4\right) - O_P\left(\epsilon_0^4\right)$$

where

$$L_{1NT}(\boldsymbol{\beta}) \equiv \frac{2}{NT}\sum_{k=1}^{K}\epsilon_k\epsilon_0 L^{(2)}\left(\lambda^0, F^0, \mathbf{X}_k, \mathbf{X}_0\right) + \frac{3}{NT}\sum_{k=1}^{K}\epsilon_k\epsilon_0\epsilon_0 L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_k, \mathbf{X}_0, \mathbf{X}_0\right),$$

$$L_{2NT}(\boldsymbol{\beta}) \equiv \frac{1}{NT}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\epsilon_{k_1}\epsilon_{k_2}L^{(2)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}\right),$$

$$R_{NT}(\boldsymbol{\beta}) \equiv \frac{3}{NT}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\epsilon_{k_1}\epsilon_{k_2}\epsilon_0 L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{X}_0\right) + \frac{1}{NT}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\sum_{k_3=1}^{K}\epsilon_{k_1}\epsilon_{k_2}\epsilon_{k_3}L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{X}_{k_3}\right).$$

Clearly, L_{1NT} and L_{2NT} are linear and quadratic in $\epsilon_k = \beta_k^0 - \beta_k$, $k = 1, ..., K$, respectively, and R_{NT} reflects the terms in the third order likelihood expansion that are asymptotically negligible (argued below). Noting that $L^{(g)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, \mathbf{X}_{k_2}, ..., \mathbf{X}_{k_g}\right)$ is linear in the last g elements and $\epsilon_0\mathbf{X}_0 = \mathbf{e}$, we have

$$\begin{aligned}
L_{1NT}(\boldsymbol{\beta}) &\equiv \frac{2}{NT}\sum_{k=1}^{K}\epsilon_k\left[L^{(2)}\left(\lambda^0, F^0, \mathbf{X}_k, \mathbf{e}\right) + \frac{3}{2}L^{(3)}\left(\lambda^0, F^0, \mathbf{X}_k, \mathbf{e}, \mathbf{e}\right)\right] \\
&= \frac{2}{NT}\sum_{k=1}^{K}\epsilon_k\left[\operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_k M_{F^0}\mathbf{e}'\right) - \frac{1}{2}\sum_{\text{all 6 permutations of } (\mathbf{X}_k,\mathbf{e},\mathbf{e})}\operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_k M_{F^0}\mathbf{e}'\Phi_1\mathbf{e}'\right)\right] \\
&= -2\gamma_{NT}\left(\boldsymbol{\beta} - \boldsymbol{\beta}^0\right)'\left(C_{NT}^{(1)} + C_{NT}^{(2)}\right)
\end{aligned}$$

where the $K \times 1$ vectors $C_{NT}^{(1)}$ and $C_{NT}^{(2)}$ are defined in (A.2) and (A.3), respectively. Next,

$$L_{2NT}(\boldsymbol{\beta}) = \frac{1}{NT}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\epsilon_{k_1}\epsilon_{k_2}\operatorname{tr}\left(M_{\lambda^0}\mathbf{X}_{k_1}M_{F^0}\mathbf{X}_{k_2}'\right) = \left(\boldsymbol{\beta}-\boldsymbol{\beta}^0\right)' D_{NT}\left(\boldsymbol{\beta}-\boldsymbol{\beta}^0\right)$$

where D_{NT} is defined in (3.6). As in **MWa**, noticing that

$$\frac{1}{NT}\left(\epsilon_0\right)^{g-r} L^{(g)}\left(\lambda^0, F^0, \mathbf{X}_{k_1}, ..., \mathbf{X}_{k_r}, \mathbf{X}_0..., \mathbf{X}_0\right) = O_P\left(\left(\|\mathbf{e}\|/\sqrt{NT}\right)^{g-r}\right) = O_P\left(\left(\delta_{NT}^{-1}+\gamma_{NT}\right)^{g-r}\right),$$

we can readily determine the probability order of R_{NT} as $R_{NT} \equiv O_P\left(\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^0\right\|^2\left(\delta_{NT}^{-1}+\gamma_{NT}\right)+\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^0\right\|^3\right)$. It follows that

$$\mathcal{L}_{NT}(\boldsymbol{\beta}) = \mathcal{L}_{NT}\left(\boldsymbol{\beta}^0\right) - 2\gamma_{NT}\left(\boldsymbol{\beta}-\boldsymbol{\beta}^0\right)'\left(C_{NT}^{(1)}+C_{NT}^{(2)}\right) + \left(\boldsymbol{\beta}-\boldsymbol{\beta}^0\right)' D_{NT}\left(\boldsymbol{\beta}-\boldsymbol{\beta}^0\right) + \tilde{R}_{NT}(\boldsymbol{\beta}) \quad \text{(C.1)}$$

where

$$\tilde{R}_{NT}(\boldsymbol{\beta}) = O_P\left\{\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^0\right\|^2\left(\delta_{NT}^{-1}+\gamma_{NT}\right)+\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^0\right\|^3+\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^0\right\|\left(\delta_{NT}^{-3}+\gamma_{NT}^3\right)\right\}. \quad \text{(C.2)}$$

Under Assumptions A.1-A.3, we can readily show that $D_{NT}^{-1} = O_P(1)$, $C_{NT}^{(1)} = O_P(1)$, and $C_{NT}^{(2)} = O_P\left(\delta_{NT}^{-2}/\gamma_{NT}\right)$. Let $\vartheta_{NT} \equiv \gamma_{NT}D_{NT}^{-1}(C_{NT}^{(1)}+C_{NT}^{(2)})$. In view of the fact that $\mathcal{L}_{NT}(\hat{\boldsymbol{\beta}}) \leq \mathcal{L}_{NT}\left(\boldsymbol{\beta}^0+\vartheta_{NT}\right)$, we apply (C.1) to the objects on both sides to obtain

$$\begin{aligned}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0-\vartheta_{NT}\right)' D_{NT}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0-\vartheta_{NT}\right) &\leq \tilde{R}_{NT}\left(\boldsymbol{\beta}^0+\vartheta_{NT}\right) - \tilde{R}_{NT}\left(\hat{\boldsymbol{\beta}}\right) \\ &= O_P\left[\gamma_{NT}^2\left(\delta_{NT}^{-1}+\gamma_{NT}\right)+\gamma_{NT}\delta_{NT}^{-3}\right] - \tilde{R}_{NT}\left(\hat{\boldsymbol{\beta}}\right)\end{aligned}$$

where the last line follows from (C.2) and Assumption A.3. Consequently,

$$\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0 = \vartheta_{NT} + O_P\{\left[\gamma_{NT}^2\left(\delta_{NT}^{-1}+\gamma_{NT}\right)+\gamma_{NT}\delta_{NT}^{-3}\right]^{1/2}\},$$

and the result follows. ■

Proof of Lemma A.3. The proof follows from that of Corollary A.2 in Hall and Heyde (1980) subject to obvious modifications on notation. ■

Proof of Lemma A.4. The proof follows from that of Lemma 2.2 in Sun and Chiang (1997) by using Lemma A.5 to replace their Lemma 2.1. ■

Proof of Lemma A.5. The proof is analogous to that of Lemma 2.1 of Sun and Chiang (1997) with obvious modifications. ■

Proof of Lemma A.6. Without loss of generality (wlog) we consider the case $m = 3$. It suffices to show that

$$EH \equiv \sum_{1\leq t_1<...<t_3\leq T}\sum_{1\leq t_4<...<t_6\leq T} E_{\mathcal{D}}\left[h^{(3)}\left(\xi_{t_1},\xi_{t_2},\xi_{t_3}\right)h^{(3)}\left(\xi_{t_4},\xi_{t_5},\xi_{t_6}\right)\right] = O_P\left(T^3\right).$$

We consider four cases for the time indices in the set $S \equiv \{t_1, t_2, ..., t_6\}$ inside the summation: (a) $\#S = 6$, (b) $\#S = 5$, (c) $\#S = 4$ and (d) $\#S = 3$. We use EH_a, EH_b, EH_c and EH_d to denote EH when the time indices are restricted to cases (a)-(d), respectively. By Cauchy-Schwarz's and Jensen's inequalities, we have

$$
\begin{aligned}
& E_{\mathcal{D}} \left| h^{(3)}\left(\xi_{t_1}, \xi_{t_2}, \xi_{t_3}\right) h^{(3)}\left(\xi_{t_4}, \xi_{t_5}, \xi_{t_6}\right) \right| \\
\leq\ & \frac{1}{2} E_{\mathcal{D}} \left[\left| h^{(3)}\left(\xi_{t_1}, \xi_{t_2}, \xi_{t_3}\right) \right|^2 + \left| h^{(3)}\left(\xi_{t_4}, \xi_{t_5}, \xi_{t_6}\right) \right|^2 \right] \\
\leq\ & \frac{1}{2} \left\{ \left[E_{\mathcal{D}} \left| h^{(3)}\left(\xi_{t_1}, \xi_{t_2}, \xi_{t_3}\right) \right|^{2+\sigma} \right]^{2/(2+\sigma)} + \left[E_{\mathcal{D}} \left| h^{(3)}\left(\xi_{t_1}, \xi_{t_2}, \xi_{t_3}\right) \right|^{2+\sigma} \right]^{2/(2+\sigma)} \right\} \leq C \sum_{q=1}^{6} \bar{C}_{\mathcal{D}}\left(t_q\right)
\end{aligned}
$$

where $\bar{C}_{\mathcal{D}}(t) \equiv 1 + \max_{1 \leq q \leq m} C_{q\mathcal{D}}(t)$. With this, one can readily show that $EH_d = O_P\left(T^3\right)$.

To bound EH_a, let $1 \leq s_1 < ... < s_6 \leq T$ be the permutation of $t_1, ..., t_6$ in ascending order and let d_c be the c-th largest difference among $s_{j+1} - s_j$, $j = 1, ..., 5$. Define $H(s_1, ..., s_6) = h^{(3)}\left(\xi_{t_1}, \xi_{t_2}, \xi_{t_3}\right) h^{(3)}\left(\xi_{t_4}, \xi_{t_5}, \xi_{t_6}\right)$. For any $1 \leq j \leq 5$, put $P_{0\mathcal{D}}^{(6)}\left(E^{(6)}\right) = P_{\mathcal{D}}\left(\left(\xi_{t_1}, .., \xi_{t_6}\right) \in E^{(6)}\right)$, and $P_{j\mathcal{D}}^{(6)}\left(E^{(j)} \times E^{(6-j)}\right) = P_{\mathcal{D}}((\xi_{t_1}, .., \xi_{t_j}) \in E^{(j)}) P_{\mathcal{D}}((\xi_{t_{j+1}}, .., \xi_{t_6}) \in E^{(6-j)})$, where $E^{(j)}$ is a Borel set in $\mathbb{R}^{jl}$. It is easy to verify that for any $0 \leq j \leq 5$, $\int \left| H(s_1, ..., s_6) \right|^{1+\sigma/2} dP_j^{(6)} \leq C \sum_{q=1}^{6} \bar{C}_{\mathcal{D}}(t_q)$. By Lemma A.5 with $\tilde{\sigma} = \sigma/2$, we have

$$
\left| E\left[H(s_1, ..., s_6)\right] \right| \leq \begin{cases} C \sum_{q=1}^{6} \bar{C}_{\mathcal{D}}(t_q)\, \alpha^{\mathcal{D}}(s_2 - s_1)^{\sigma/(2+\sigma)} & \text{if } s_2 - s_1 = d_1 \\ C \sum_{q=1}^{6} \bar{C}_{\mathcal{D}}(t_q)\, \alpha^{\mathcal{D}}(s_6 - s_5)^{\sigma/(2+\sigma)} & \text{if } s_6 - s_5 = d_1 \end{cases}.
$$

Therefore

$$
\begin{aligned}
& \sum_{\substack{1 \leq s_1 < ... < s_6 \leq T \\ s_2 - s_1 = d_1}} \left| E\left[H(s_1, ..., s_6)\right] \right| \\
& \leq C \sum_{s_1=1}^{T-5} \sum_{s_2 = s_1 + \max_{j \geq 3}\{s_j - s_{j-1}\}}^{T-4} \sum_{s_3 = s_2+1}^{T-3} \cdots \sum_{s_6 = s_5+1}^{T} \sum_{q=1}^{6} \bar{C}_{\mathcal{D}}(t_q)\, \alpha^{\mathcal{D}}(s_2 - s_1)^{\sigma/(2+\sigma)} \\
& \leq 3C \sum_{s_1=1}^{T-1} \sum_{s_2=s_1+1}^{T} \sum_{q=1}^{2} \bar{C}_{\mathcal{D}}(t_q)(s_2 - s_1)^4 \alpha^{\mathcal{D}}(s_2 - s_1)^{\sigma/(2+\sigma)} \leq O_P(T) \sum_{\tau=1}^{T} \tau^4 \alpha^{\mathcal{D}}(\tau)^{\sigma/(2+\sigma)}.
\end{aligned}
$$

By the same token, $\sum_{\substack{1 \leq s_1 < ... < s_6 \leq T \\ s_6 - s_5 = d_1}} \left| E\left[H(s_1, ..., s_6)\right] \right| \leq O_P(T) \sum_{\tau=1}^{T} \tau^4 \alpha^{\mathcal{D}}(\tau)^{\sigma/(2+\sigma)}$. Analogously, we can show that

$$
\begin{aligned}
\sum_{\substack{1 \leq s_1 < ... < s_6 \leq T \\ s_2 - s_1 = d_2 \text{ or } s_6 - s_5 = d_2}} \left| E\left[H(s_1, ..., s_6)\right] \right| & \leq O_P\left(T^2\right) \sum_{\tau=1}^{T} \tau^3 \alpha^{\mathcal{D}}(\tau)^{\sigma/(2+\sigma)}, \text{ and} \\
\sum_{\substack{1 \leq s_1 < ... < s_6 \leq T \\ s_2 - s_1 = d_3 \text{ or } s_6 - s_5 = d_3}} \left| E\left[H(s_1, ..., s_6)\right] \right| & \leq O_P\left(T^3\right) \sum_{\tau=1}^{T} \tau^2 \alpha^{\mathcal{D}}(\tau)^{\sigma/(2+\sigma)}.
\end{aligned}
$$

It follows that $EH_a \leq O_P\left(T^3\right) \sum_{\tau=1}^{T} \tau^2 \alpha^{\mathcal{D}}(\tau)^{\sigma/(2+\sigma)} = O_P\left(T^3\right)$. Analogously, we can show that $EH_s = O_P\left(T^3\right)$ for $s = b, c$. ■

Proof of Lemma A.7. (i) Following the proof of Lemma A.1(iv) in Su and Jin (2012), we have

$$\mu_1(\hat{\Omega}_i) \leq \mu_1(\Omega_i) + \left\|\hat{\Omega}_i - \Omega_i\right\|_F = \mu_1(\Omega_i) + O_P\left(T^{-1/2}\right) \text{ for each } i.$$

(ii) Following the proof of Lemma A.1(v) in Su and Jin (2012), we have

$$\mu_{\min}(\hat{\Omega}_i) \geq \mu_{\min}(\Omega_i) - \left\|\hat{\Omega}_i - \Omega_i\right\|_F = \mu_{\min}(\Omega_i) - O_P\left(T^{-1/2}\right) \text{ for each } i.$$

(iii) Let $\eta_{NT} \equiv (NT)^{1/(4+2\sigma)}$. To obtain a uniform result, we consider the (j,k)'th element of $\hat{\Omega}_i - \Omega_i$:

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}[X_{it,j}X_{it,k} - E_{\mathcal{D}}(X_{it,j}X_{it,k})] &= T^{-1}\sum_{t=1}^{T}[X_{it,j}X_{it,k} - E_{\mathcal{D}}(X_{it,j}X_{it,k})]\,1\{|X_{it,j}X_{it,k}| \leq \eta_{NT}\} \\ &\quad + T^{-1}\sum_{t=1}^{T}[X_{it,j}X_{it,k} - E_{\mathcal{D}}(X_{it,j}X_{it,k})]\,1\{|X_{it,j}X_{it,k}| > \eta_{NT}\} \\ &\equiv s_{1i} + s_{2i}, \text{ say.} \end{aligned}$$

Note that $\mathrm{Var}_{\mathcal{D}}[X_{it,j}X_{it,k}1\{|X_{it,j}X_{it,k}| \leq \eta_{NT}\}] \leq E_{\mathcal{D}}(X_{it,j}^2 X_{it,k}^2)$ which is bounded in probability by Assumption A.1(iii). By Boole's inequality, Lemma A.4, and Assumption A.2(i), we have

$$\begin{aligned} & P_{\mathcal{D}}\left(\max_{1\leq i\leq N}|s_{1i}| \geq Ca_{NT}\right) \\ \leq & \sum_{i=1}^{N} P_{\mathcal{D}}\left(T^{-1}\sum_{t=1}^{T}[X_{it,j}X_{it,k} - E_{\mathcal{D}}(X_{it,j}X_{it,k})]\,1\{|X_{it,j}X_{it,k}| \leq \eta_{NT}\} \geq Ca_{NT}\right) \\ \leq & \; 2\tau_0\sum_{i=1}^{N}\exp\left(-\frac{TC^2a_{NT}^2}{2E_{\mathcal{D}}\left(X_{it,j}^2X_{it,k}^2\right) + 4Ca_{NT}\eta_{NT}/3}\right) + 2NT\alpha_{\mathcal{D}}(\tau_0) = o_P(1) \text{ for sufficiently large } C. \end{aligned}$$

By Boole's and Chebyshev's inequalities, Assumption A.1(iii), and the dominated convergence theorem, as $(N,T) \to \infty$ we have

$$\begin{aligned} P_{\mathcal{D}}\left(\max_{1\leq i\leq N}|s_{2i}| \geq Ca_{NT}\right) &\leq P_{\mathcal{D}}\left(\max_{1\leq i\leq N}\max_{1\leq t\leq T}|X_{it,j}X_{it,k}| > \eta_{NT}\right) \leq \sum_{i=1}^{N}\sum_{t=1}^{T}P_{\mathcal{D}}(|X_{it,j}X_{it,k}| > \eta_{NT}) \\ &\leq \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E_{\mathcal{D}}\left[|X_{it,j}X_{it,k}|^{4+2\sigma}\,1\left(|X_{it,j}X_{it,k}|^{4+2\sigma} > \eta_{NT}^{4+2\sigma}\right)\right] \\ &= o_P(1). \end{aligned}$$

Consequently, $\max_{1\leq i\leq N}|T^{-1}\sum_{t=1}^{T}[X_{it,j}X_{it,k} - E_{\lambda_i^0}(X_{it,j}X_{it,k})]| = O_P(a_{NT})$ for each (j,k). This completes the proof of (iii).

(iv) By (i)-(iii), $\max_{1\leq i\leq N}\left\|\Omega_i^{-1}\right\|_F \leq K^{1/2}\max_{1\leq i\leq N}\left\|\Omega_i^{-1}\right\| = K^{1/2}[\min_{1\leq i\leq N}\mu_{\min}(\Omega_i)]^{-1} = O_P(1)$ and $\max_{1\leq i\leq N}||\hat{\Omega}_i^{-1}||_F \leq K^{1/2}[\min_{1\leq i\leq N}\mu_{\min}(\hat{\Omega}_i)]^{-1} = O_P(1)$. It follows that $\max_{1\leq i\leq N}||\hat{\Omega}_i^{-1} - \Omega_i^{-1}||_F = \max_{1\leq i\leq N}||\Omega_i^{-1}(\Omega_i - \hat{\Omega}_i)\hat{\Omega}_i^{-1}||_F \leq \max_{1\leq i\leq N}||\Omega_i^{-1}||_F \max_{1\leq i\leq N}||\Omega_i - \hat{\Omega}_i||_F \max_{1\leq i\leq N}||\hat{\Omega}_i^{-1}||_F = O_P(a_{NT})$. ■

Proof of Lemma A.8. Using $\eta_{tr} = 1_{tr} - T^{-1}F_{tr}$ where $1_{tr} \equiv 1\{t=r\}$ and $F_{tr} \equiv F_t^{0\prime}\left(F^{0\prime}F^0/T\right)^{-1}F_r^0,$ we have

$$\begin{aligned} D_{1NT} &= \frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{r=1}^{T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}\eta_{tr}X_{ir}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)X_{iq}\eta_{qs} \\ &= \frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}X_{it}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)X_{is} \\ &\quad+\frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}X_{it}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)X_{iq}F_{qs} \\ &\quad+\frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{r=1}^{T}\varepsilon_{it}\varepsilon_{is}F_{tr}X_{ir}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)X_{is} \\ &\quad+\frac{1}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{r=1}^{T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}F_{tr}X_{ir}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)X_{iq}F_{qs} \\ &\equiv D_{1NT,1}+D_{1NT,2}+D_{1NT,3}+D_{1NT,4},\ \text{say.} \end{aligned}$$

We prove the lemma by showing that $D_{1NT,l} = o_P(1)$ for $l = 1,2,3,4.$

First, we prove that $D_{1NT,1} = o_P(1)$. We decompose $D_{1NT,1}$ as follows

$$\begin{aligned} D_{1NT,1} &= \frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}X_{it}'\Omega_i^{-1}\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}X_{is} \\ &\quad+\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}X_{it}'\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}X_{is} \\ &\equiv D_{1NT,11}+D_{1NT,12},\ \text{say.} \end{aligned}$$

By the fact that $\text{tr}(AB) \leq \|A\|_F\|B\|_F$, the submultiplicative property of $\|\cdot\|_F$, and Lemma A.7,

$$\begin{aligned} |D_{1NT,12}| &= \left|\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\text{tr}\left[\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}\sum_{1\leq s\neq t\leq T}X_{is}X_{it}'\varepsilon_{it}\varepsilon_{is}\right]\right| \\ &\leq \max_{1\leq i\leq N}\left\{\left\|\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right\|_F\left\|\Omega_i-\hat{\Omega}_i\right\|_F\|\Omega_i^{-1}\|_F\right\}\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\left\|\sum_{1\leq s\neq t\leq T}X_{is}X_{it}'\varepsilon_{it}\varepsilon_{is}\right\|_F \\ &= O_P\left(a_{NT}^2\right)O_P\left(N^{1/2}\right)=O_P\left(N^{1/2}a_{NT}^2\right)=o_P(1) \end{aligned}$$

where the second equality follows because $E_{\mathcal{D}}||\sum_{1\leq s\neq t\leq T}X_{is}X_{it}'\varepsilon_{it}\varepsilon_{is}||_F \leq \{E_{\mathcal{D}}||\sum_{1\leq s\neq t\leq T}X_{is}X_{it}'\varepsilon_{it}\varepsilon_{is}||_F^2\}^{1/2}$ $= O_P(T)$ by Jensen's and Davydov's inequalities.

Let $\bar{X}_{it} \equiv \Omega_i^{-1/2}X_{it}$ and $\bar{Z}_{ir} \equiv \bar{X}_{ir}\bar{X}_{ir}' - E_{\mathcal{D}}(\bar{X}_{ir}\bar{X}_{ir}')$. Then

$$\begin{aligned} -D_{1NT,11} &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq r\neq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}_{it}'\bar{Z}_{ir}\bar{X}_{is}+\frac{2}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}_{it}'\bar{Z}_{it}\bar{X}_{is} \\ &\equiv D_{1NT,111}+2D_{1NT,112}. \end{aligned}$$

Noting that $E_{\mathcal{D}}(D_{1NT,112}) = 0$ and $\text{Var}_{\mathcal{D}}(D_{1NT,112}) = O(T^{-2})$ by Davydov's inequality, we have $D_{1NT,112} = o_P(1)$ by Chebyshev's inequality.

Let $\vartheta_i^0(\xi_{it}, \xi_{is}, \xi_{ir}) \equiv \varepsilon_{it}\varepsilon_{is}\bar{X}_{it}'\bar{Z}_{ir}\bar{X}_{is}$ where $\xi_{it} \equiv (\varepsilon_{it}, X_{it}')'$. Define the symmetric version of ϑ_i^0 as $\vartheta_i(\xi_{it}, \xi_{is}, \xi_{ir}) = [\vartheta_i^0(\xi_{it}, \xi_{is}, \xi_{ir}) + \vartheta_i^0(\xi_{it}, \xi_{ir}, \xi_{is}) + \vartheta_i^0(\xi_{ir}, \xi_{it}, \xi_{is})]/3$. Then we have

$$D_{1NT,111} = \frac{6}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{1\leq t<s<r\leq T}\vartheta_i(\xi_{it}, \xi_{is}, \xi_{ir}) = \frac{(T-1)(T-2)}{TN^{1/2}}\sum_{i=1}^{N}\mathcal{U}_{1iT}$$

where $\mathcal{U}_{1iT} \equiv \frac{6}{T(T-1)(T-2)}\sum_{1\leq t<s<r\leq T}\vartheta_i(\xi_{it}, \xi_{is}, \xi_{ir})$. To study $D_{1NT,111}$, we assume that $K = 1$ below to simplify notation as the general case follows from the Slutsky lemma. Let $\vartheta_{i(s)}$ and $\mathcal{H}_{iT}^{(s)}$ be analogously defined as $\vartheta_{(s)}$ and $\mathcal{H}_T^{(s)}$ for $s = 0, 1, 2, 3$, defined before Lemma A.5 with the kernel ϑ_i in place of ϑ. Then we have $\vartheta_{i(0)} = 0$, $\vartheta_{i(1)}(\xi_1) = 0$, and $\vartheta_{i(2)}(\xi_1, \xi_2) = 0$. This implies that $\mathcal{U}_{1iT}$ is a third order degenerate U-statistic with $\mathcal{H}_{iT}^{(c)} = 0$ for $c = 1, 2$ and $\mathcal{H}_{iT}^{(3)} = \mathcal{U}_{1iT}$. By the repeated use of Cauchy-Schwarz's and Jensen's inequalities,

$$\begin{aligned} E_{\mathcal{D}}\left|\varepsilon_{it_1}\varepsilon_{it_2}\bar{X}_{it_1}'\bar{Z}_{it_3}\bar{X}_{it_2}\right|^{2+\sigma} &\leq \frac{1}{2}\left[E_{\mathcal{D}}\left(\left|\varepsilon_{it_1}\bar{X}_{it_1}\varepsilon_{it_2}\bar{X}_{it_2}\right|^{4+2\sigma} + \left|\bar{Z}_{it_3}\right|^{4+2\sigma}\right)\right] \\ &\leq \frac{1}{4}\left[E_{\mathcal{D}}\left|\varepsilon_{it_1}\bar{X}_{it_1}\right|^{8+4\sigma} + E_{\mathcal{D}}\left|\varepsilon_{it_2}\bar{X}_{it_2}\right|^{8+4\sigma} + 2E_{\mathcal{D}}\left|\bar{Z}_{it_3}\right|^{4+2\sigma}\right] \\ &\leq \sum_{q=1}^{3}\bar{C}_{1i\mathcal{D}}(t_q) \end{aligned}$$

where $\bar{C}_{1i\mathcal{D}}(t) \equiv \frac{1}{2}(E_{\mathcal{D}}|\varepsilon_{it}\bar{X}_{it}|^{8+4\sigma} + E_{\mathcal{D}}|\bar{Z}_{it}|^{4+2\sigma})$. It is easy to see that $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{1i\mathcal{D}}(t) = O_P(1)$ by Assumption A.1 and Markov inequality. In view of the fact

$$L_{1iT} \equiv \max\left\{\int|\vartheta_i(v_{t_1}, v_{t_2}, v_{t_3})|^{2+\sigma}\ \Pi_{s=1}^{3}dF_{it_s}(v_{t_s}|\mathcal{D}),\ E_{\mathcal{D}}\left|\vartheta_i\left(\xi_{it_1}, \xi_{it_2}, \xi_{it_3}\right)\right|^{2+\sigma}\right\} \leq \sum_{q=1}^{3}\bar{C}_{1i\mathcal{D}}(t_q)$$

we apply Lemma A.4 to obtain $\sum_{i=1}^{N}E_{\mathcal{D}}(\mathcal{U}_{1iT}^2) = NO_P(T^{-3})$. It follows that

$$\text{Var}_{\mathcal{D}}(D_{1NT,111}) = \frac{(T-1)^2(T-2)^2}{T^2N}\sum_{i=1}^{N}\text{Var}_{\mathcal{D}}(\mathcal{U}_{1iT}) \leq O_P(T^2N^{-2})\sum_{i=1}^{N}E_{\mathcal{D}}(\mathcal{U}_{1iT}^2) = O_P(T^{-1}).$$

In addition, note that $E_{\mathcal{D}}(D_{1NT,111}) = 2T^{-2}N^{-1/2}\sum_{i=1}^{N}\sum_{1\leq t<s<r\leq T}E_{\mathcal{D}}(\varepsilon_{it}\varepsilon_{is}\bar{X}_{it}'\bar{Z}_{ir}\bar{X}_{is})$. We consider two cases for the indices in $\{t, s, r\}$: (a) either $r - s > \tau_*$ or $s - t > \tau_*$, and (b) $r - s \leq \tau_*$ and $s - t \leq \tau_*$. In case (a) we have

$$\left|E_{\mathcal{D}}\left(\varepsilon_{it}\varepsilon_{is}\bar{X}_{it}'\bar{Z}_{ir}\bar{X}_{is}\right)\right| \leq \begin{cases} 8\alpha_{\mathcal{D}}(\tau_*)^{(1+\sigma)/(2+\sigma)}C_{2i\mathcal{D}}(t, s, r) & \text{if } r - s > \tau_* \\ 8\alpha_{\mathcal{D}}(\tau_*)^{(1+\sigma)/(2+\sigma)}C_{3i\mathcal{D}}(t, s, r) & \text{if } s - t > \tau_* \end{cases},$$

where $C_{2i\mathcal{D}}(t, s, r) \equiv \left\|\varepsilon_{it}\varepsilon_{is}\bar{X}_{it}\bar{X}_{is}\right\|_{4+2\sigma,\mathcal{D}}\left\|\bar{Z}_{ir}\right\|_{4+2\sigma,\mathcal{D}}$ and $C_{3i\mathcal{D}}(t, s, r) \equiv ||\varepsilon_{is}\bar{Z}_{ir}\bar{X}_{is}||_{(8+4\sigma)/3,\mathcal{D}}\left\|\varepsilon_{it}\bar{X}_{it}\right\|_{8+4\sigma,\mathcal{D}}$. Note that $C_{2i\mathcal{D}}(t_1, t_2, t_3) \leq \sum_{q=1}^{3}[1 + \bar{C}_{1i\mathcal{D}}(t_q)]$ and

$$\begin{aligned} C_{3i\mathcal{D}}(t_1, t_2, t_3) &\leq \frac{1}{2}\left\{||\varepsilon_{it_2}\bar{Z}_{it_3}\bar{X}_{it_2}||_{(8+4\sigma)/3,\mathcal{D}}^2 + \left|\left|\varepsilon_{it_1}\bar{X}_{it_1}\right|\right|_{8+4\sigma,\mathcal{D}}^2\right\} \\ &\leq \frac{1}{4}\left\{||\varepsilon_{it_2}\bar{X}_{it_2}||_{8+4\sigma,\mathcal{D}}^2 + ||\bar{Z}_{it_3}||_{4+2\sigma,\mathcal{D}}^2 + 2\left\|\varepsilon_{it_1}\bar{X}_{it_1}\right\|_{8+4\sigma,\mathcal{D}}^2\right\} \leq \sum_{q=1}^{3}\bar{C}_{3i\mathcal{D}}(t_q) \end{aligned}$$

where $\bar{C}_{3i\mathcal{D}}(t) \equiv \frac{1}{4}(||\bar{Z}_{it}||^2_{4+2\sigma,\mathcal{D}}+3\left\|\varepsilon_{it}\bar{X}_{it}\right\|^2_{8+4\sigma,\mathcal{D}})$. It is easy to verify that $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{3i\mathcal{D}}(t) = O_P(1)$. This, together with the fact that the total number of terms in the summation is of order $O\left(NTm^2\right)$ in case (b) implies that

$$\begin{aligned} E_{\mathcal{D}}(D_{1NT,111}) &\leq 16N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[2+\bar{C}_{1i\mathcal{D}}(t)+\bar{C}_{3i\mathcal{D}}(t)\right]\alpha_{\mathcal{D}}(\tau_*)^{(1+2\sigma)/(4+2\sigma)}+T^{-2}N^{-1/2}O_P\left(NT\tau_*^2\right) \\ &= O_P\left(N^{1/2}T\alpha_{\mathcal{D}}(\tau_*)^{(1+\sigma)/(2+\sigma)}+N^{1/2}T^{-1}\tau_*^2\right)=o_P(1) \text{ by Assumption A.2}(i). \end{aligned}$$

Consequently $E_{\mathcal{D}}\left(D^2_{1NT,111}\right) = \mathrm{Var}_{\mathcal{D}}(D_{1NT,111})+[E_{\mathcal{D}}(D_{1NT,111})]^2 = o_P(1)$ and $D_{1NT,111}=o_P(1)$ by Chebyshev's inequality.

Next, we prove that $D_{1NT,2}=o_P(1)$. We decompose $D_{1NT,2}$ as follows

$$\begin{aligned} D_{1NT,2} &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}X'_{it}\hat{\Omega}_i^{-1}\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}X_{iq}F_{qs} \\ &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}X'_{it}\Omega_i^{-1}\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}X_{iq}F_{qs} \\ &\quad +\frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{q=1}^{T}\varepsilon_{it}\varepsilon_{is}X'_{it}\left(\hat{\Omega}_i^{-1}-\Omega_i^{-1}\right)\left(\Omega_i-\hat{\Omega}_i\right)\Omega_i^{-1}X_{iq}F_{qs} \\ &\equiv D_{1NT,21}+D_{1NT,22}. \end{aligned}$$

The second term can be bounded analogously to $D_{1NT,12}$. For $D_{1NT,21}$, one can readily show that

$$\begin{aligned} -D_{1NT,21} &= \frac{1}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s\neq t\leq T}\sum_{1\leq r,q\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}'_{it}\bar{Z}_{ir}\bar{X}_{iq}F_{qs} \\ &= \frac{1}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq q\neq r\neq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}'_{it}\bar{Z}_{ir}\bar{X}_{iq}F_{qs}+o_P(1) \\ &= \frac{1}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq q\neq r\neq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}'_{it}\bar{Z}_{ir}\left[\bar{X}_{iq}-E_{\mathcal{D}}\left(\bar{X}_{iq}\right)\right]F_{qs} \\ &\quad +\frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq r\neq s\neq t\leq T}\varepsilon_{it}\varepsilon_{is}\bar{X}'_{it}\bar{Z}_{ir}\zeta_{is}+o_P(1)\equiv D_{1NT,211}+D_{1NT,212}+o_P(1), \end{aligned}$$

where $\zeta_{is}\equiv T^{-1}\sum_{q=1}^{T}E_{\mathcal{D}}\left(\bar{X}_{iq}\right)F_{qs}$ is measurable with respect to $\mathcal{D}$. The study of $D_{1NT,212}$ is analogous to that of $D_{1NT,111}$. To study $D_{1NT,211}$, let $\vartheta_i^0\left(\xi_{it},\xi_{is},\xi_{ir},\xi_{iq}\right)\equiv\varepsilon_{it}\varepsilon_{is}\bar{X}'_{it}\bar{Z}_{ir}\left[\bar{X}_{iq}-E_{\mathcal{D}}\left(\bar{X}_{iq}\right)\right]F_{qs}$. Define the symmetric version of ϑ_i^0 as $\vartheta_i(\xi_{it},\xi_{is},\xi_{ir},\xi_{iq})=\sum_{\text{all 24 permutations of }(t,s,r,q)}\vartheta_i^0(\xi_{it},\xi_{is},\xi_{ir},\xi_{iq})/24$. Then we have

$$D_{1NT,211}=\frac{24}{T^3N^{1/2}}\sum_{i=1}^{N}\sum_{1\leq t<s<r<q\leq T}\vartheta_i(\xi_{it},\xi_{is},\xi_{ir},\xi_{iq})=\frac{(T-1)(T-2)(T-3)}{T^2N^{1/2}}\sum_{i=1}^{N}\mathcal{U}_{2iT}$$

where $\mathcal{U}_{1iT}\equiv\frac{24}{T(T-1)(T-2)(T-3)}\sum_{1\leq t<s<r<q\leq T}\vartheta_i(\xi_{it},\xi_{is},\xi_{ir},\xi_{iq})$. For this ϑ_i, using analogous notations defined before Lemma A.6, we have $\vartheta_{i(0)}=0$, $\vartheta_{i(1)}(v_1)=0$, $\vartheta_{i(2)}(v_1,v_2)=0$, $\vartheta_{i(3)}(v_1,v_2,v_3)=0$,

$\mathcal{H}_{iT}^{(c)} = 0$ for $c = 1, 2, 3$, and $\mathcal{H}_{iT}^{(4)} = \mathcal{U}_{2iT}$. This implies that $\mathcal{U}_{2iT}$ is a fourth order degenerate U-statistic and by Lemma A.6, $\sum_{i=1}^{N} E_{\mathcal{D}}\left(\mathcal{U}_{2iT}^{2}\right) = NO_P\left(T^{-3}\right)$. It follows that $\mathrm{Var}_{\mathcal{D}}\left(D_{1NT,211}\right) \leq O\left(T^{2}N^{-1}\right)\sum_{i=1}^{N} E_{\mathcal{D}}\left(\mathcal{U}_{2iT}^{2}\right) = O_P\left(T^{-1}\right)$. In addition, we can show that $E_{\mathcal{D}}\left(D_{1NT,211}\right) = o_P\left(1\right)$. Consequently, $D_{1NT,211} = o_P\left(1\right)$ by Chebyshev's inequality.

Analogously, we can show that $D_{1NT,3} = o_P\left(1\right)$ and $D_{1NT,4} = o_P\left(1\right)$. ■

Proof of Lemma A.9. (i) Let $\xi_{i,ts} \equiv \left[X_{it} - E_{\mathcal{D}}\left(X_{it}\right)\right]' \Omega_i^{-1} X_{is}$. We have

$$\begin{aligned} D_{2NT,1} &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s<t\leq T}\sum_{r=1,r\neq t,s}^{T}\varepsilon_{it}\varepsilon_{is}F_{tr}\xi_{i,rs} + \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s<t\leq T}\varepsilon_{it}\varepsilon_{is}F_{tt}\xi_{i,ts} \\ &\quad + \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s<t\leq T}\varepsilon_{it}\varepsilon_{is}F_{ts}\xi_{i,ss} \equiv D_{2NT,11} + D_{2NT,12} + D_{2NT,13}, \text{ say.} \end{aligned}$$

As in the analysis of $D_{1NT,111}$, we can apply Lemma A.6 to show $D_{2NT,11} = o_P\left(1\right)$. Noting that $E_{\mathcal{D}}\left(D_{2NT,12}\right) = 0$ and $E_{\mathcal{D}}\left(D_{2NT,12}^2\right) = \frac{1}{T^4 N}\sum_{i=1}^{N}\sum_{1\leq s,r<t\leq T} F_{tt}^2\ E_{\mathcal{D}}(\varepsilon_{it}^2\varepsilon_{is}\varepsilon_{ir}\xi_{i,ts}\xi_{i,tr}) = O_P\left(T^{-1}\right)$, we have $D_{2NT,12} = O_P\left(T^{-1/2}\right) = o_P\left(1\right)$. Similarly, $D_{2NT,13} = o_P\left(1\right)$.

(ii) We can show that $D_{2NT,2} = \bar{D}_{2NT,2} + o_P\left(1\right)$ where

$$\bar{D}_{2NT,2} \equiv \frac{1}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s<t\leq T}\sum_{r=1,r\neq t,s}^{T}\sum_{q=1,q\neq t,s,r}^{T}\varepsilon_{it}\varepsilon_{is}F_{tr}\left[X_{ir} - E_{\mathcal{D}}\left(X_{ir}\right)\right]'\Omega_i^{-1}\left[X_{iq} - E_{\mathcal{D}}\left(X_{iq}\right)\right]F_{qs}.$$

As in the analysis of $D_{1NT,211}$, we can apply Lemma A.6 to show $\bar{D}_{2NT,2} = o_P\left(1\right)$.

(iii) Recall $\zeta_{is} \equiv T^{-1}\sum_{q=1}^{T} E_{\mathcal{D}}\left(X_{iq}\right)F_{qs}$. We can readily show that $D_{2NT,3} = \bar{D}_{2NT,3} + o_P\left(1\right)$ where

$$\bar{D}_{2NT,3} = \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\leq s<t\leq T}\sum_{r=1,r\neq t,s}^{T}\varepsilon_{it}\varepsilon_{is}F_{tr}\left[X_{ir} - E_{\mathcal{D}}\left(X_{ir}\right)\right]'\Omega_i^{-1}\zeta_{is}.$$

As in the analysis of $D_{1NT,111}$, we can apply Lemma A.6 to show $\bar{D}_{2NT,3} = o_P\left(1\right)$. ■

Proof of Lemma A.10. Using $M_{F^0} = I_T - P_{F^0}$ we can decompose D_{3NT} as follows

$$\begin{aligned} D_{3NT} &= \gamma_{NT}N^{-1/2}\sum_{i=1}^{N}\left(\varepsilon_i'P_{X_i}X_i\delta_i - \varepsilon_i'P_{F^0}P_{X_i}X_i\delta_i - \varepsilon_i'P_{X_i}P_{F^0}X_i\delta_i + \varepsilon_i'P_{F^0}P_{X_i}P_{F^0}X_i\delta_i\right) \\ &\equiv D_{3NT,1} - D_{3NT,2} - D_{3NT,3} + D_{3NT,4}, \text{ say.} \end{aligned}$$

For $D_{3NT,1}$, we further decompose it as follows

$$\begin{aligned} D_{3NT,1} &= \gamma_{NT}T^{-1}N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'X_i\Omega_i^{-1}X_i'X_i\delta_i + \gamma_{NT}T^{-1}N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'X_i(\Omega_i^{-1} - \hat{\Omega}_i^{-1})X_i'X_i\delta_i \\ &\equiv D_{3NT,11} + D_{3NT,12} \text{ say.} \end{aligned} \tag{C.3}$$

We dispense with $D_{3NT,12}$ first. By Lemma A.7,

$$\begin{aligned} \left|D_{3NT,12}\right| &\leq \gamma_{NT}T^{-1}\max_{1\leq i\leq N}\left\|\Omega_i^{-1} - \hat{\Omega}_i^{-1}\right\|\left\{N^{-1/2}\sum_{i=1}^{N}\left\|\varepsilon_i'X_i\right\|\left\|X_i'X_i\delta_i\right\|\right\} \\ &= \gamma_{NT}T^{-1}O_P\left(a_{NT}\right)O_P\left(N^{1/2}T^{3/2}\right) = O_P\left(N^{1/4}a_{NT}\right) = o_P\left(1\right). \end{aligned} \tag{C.4}$$

For $D_{3NT,11}$, we have

$$\begin{aligned} E_{\mathcal{D}}\left(D_{3NT,11}^{2}\right) &= T^{-3}N^{-3/2}\sum_{i=1}^{N}\sum_{1\leq t_1,t_2,t_3,t_4\leq T} E_{\mathcal{D}}\left[\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}'\bar{X}_{it_2}X_{it_2}'\delta_i\ \bar{X}_{it_3}'\bar{X}_{it_4}X_{it_4}'\delta_i\right] \\ &\quad +T^{-3}N^{-3/2}\sum_{1\leq i\neq j\leq N}\sum_{1\leq t_1,t_2,t_3,t_4\leq T} E_{\mathcal{D}}\left[\varepsilon_{it_1}\varepsilon_{jt_3}\bar{X}_{it_1}'\bar{X}_{it_2}X_{it_2}'\delta_i\ \bar{X}_{jt_3}'\bar{X}_{jt_4}X_{jt_4}'\delta_j\right] \\ &\equiv ED_{11}+ED_{12},\ \text{say}, \end{aligned}$$

where recall $\bar{X}_{it}\equiv\Omega_i^{-1/2}X_{it}$. For notational simplicity, we assume that $K=1$. For ED_{11}, we consider two cases for the time indices $\{t_1,t_2,t_3,t_4\}$ inside the summation: (a) $\#\{t_1,t_2,t_3,t_4\}=4$, (b) $\#\{t_1,t_2,t_3,t_4\}\leq 3$. We use ED_{11a} and ED_{11b} to denote ED_{11} when the individual indices in the summation are restricted to be cases (a) and (b), respectively. By direct moment calculations, $ED_{11b}=T^{-3}N^{-3/2}O_P\left(NT^3\right)=O_P\left(N^{-1/2}\right)$. For case (a), we consider two subcases: $(a1)$ either t_1 or t_3 is largest in the set $\{t_1,t_2,t_3,t_4\}$, $(a2)$ $t_1<t_2<t_3<t_4$, $t_2<t_1<t_3<t_4$, or $(t_1\leftrightarrow t_3)$, or $(t_2\leftrightarrow t_4)$, and $(a3)$ $t_1<t_3<t_2<t_4$, or $(t_1\leftrightarrow t_3)$, or $(t_2\leftrightarrow t_4)$, where, e.g., $(t_1\leftrightarrow t_3)$ means that t_1 and t_3 can exchange positions. We use $ED_{11a}(s)$ to denote the corresponding summation when the individual indices are restricted to be subcases (as) for $s=1,2,3$, respectively. In subcase $(a1)$, $ED_{11a}(1)=0$. In subcase $(a2)$, wolg consider $t_1<t_2<t_3<t_4$. Then we can separate variables indexed by t_4 from those indexed by (t_1,t_2,t_3) and apply the fact that $E(\varepsilon_{it_3}|\mathcal{F}_{NT,t_3-1})=0$ and Davydov's inequality to obtain

$$\left|E_{\mathcal{D}}\left[\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i\ \bar{X}_{it_3}\bar{X}_{it_4}X_{it_4}\delta_i\right]\right|\leq 8C_{4i\mathcal{D}}(t_1,t_2,t_3,t_4)\,\alpha_{\mathcal{D}}(t_4-t_3)^{(1+2\sigma)/(4+2\sigma)},$$

where $C_{4i\mathcal{D}}(t_1,t_2,t_3,t_4)\equiv\left\|\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i\ \bar{X}_{it_3}\right\|_{2+\sigma,\mathcal{D}}\left\|\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{4+2\sigma,\mathcal{D}}$. Note that

$$\begin{aligned} C_{4i\mathcal{D}} &\leq \frac{1}{2}\left\{\left\|\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i\ \bar{X}_{it_3}\right\|_{2+\sigma,\mathcal{D}}^2+\left\|\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2\right\} \\ &\leq \frac{1}{2}\left\{\left\|\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_3}\right\|_{4+2\sigma,\mathcal{D}}\left\|\bar{X}_{it_2}X_{it_2}\delta_i\right\|_{4+2\sigma,\mathcal{D}}+\left\|\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2\right\} \\ &\leq \frac{1}{4}\left\{\left\|\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_3}\right\|_{4+2\sigma,\mathcal{D}}^2+\left\|\bar{X}_{it_2}X_{it_2}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2+2\left\|\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2\right\} \\ &\leq \frac{1}{8}\left\{\left\|\varepsilon_{it_1}\bar{X}_{it_1}\right\|_{8+4\sigma,\mathcal{D}}^2+\left\|\varepsilon_{it_3}\bar{X}_{it_3}\right\|_{8+4\sigma,\mathcal{D}}^2+2\left\|\bar{X}_{it_2}X_{it_2}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2+4\left\|\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2\right\} \end{aligned}$$

and $C_{4i\mathcal{D}}(t_1,t_2,t_3,t_4)\leq\sum_{q=1}^{4}\bar{C}_{4i\mathcal{D}}(t_q)$, where $\bar{C}_{4i\mathcal{D}}(t)\equiv\frac{1}{4}[\left\|\varepsilon_{it}\bar{X}_{it}\right\|_{8+4\sigma,\mathcal{D}}^2+3\left\|\bar{X}_{it}X_{it}\delta_i\right\|_{4+2\sigma,\mathcal{D}}^2]$ satisfies $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{4i\mathcal{D}}(t)=O_P(1)$. Then

$$\begin{aligned} ED_{11a}(2) &\leq 8T^{-4}N^{-3/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{1\leq t_1,t_2,t_3,t_4\leq T}\sum_{q=1}^{4}\bar{C}_{4i\mathcal{D}}(t_q)\,\alpha_{\mathcal{D}}(|t_4-t_3|)^{(1+2\sigma)/(4+2\sigma)} \\ &\leq 8\left\{T^{-1}N^{-3/2}\sum_{i=1}^{N}\sum_{t=1}^{T}C_{4i\mathcal{D}}(t)\right\}\left\{\sum_{s=1}^{T}\alpha_{\mathcal{D}}(s)^{(1+2\sigma)/(4+2\sigma)}\right\}=O_P\left(N^{-1/2}\right). \end{aligned}$$

In subcase $(a3)$, we consider $t_1<t_3<t_2<t_4$. Using $E(\varepsilon_{it_3}|\mathcal{F}_{NT,t_3-1})=0$ and Lemma A.3 yields

$$\left|E_{\mathcal{D}}\left[\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i\ \bar{X}_{it_3}\bar{X}_{it_4}X_{it_4}\delta_i\right]\right|\leq 8C_{5i\mathcal{D}}(t_1,t_2,t_3,t_4)\,\alpha_{\mathcal{D}}(t_2-t_3)^{(1+2\sigma)/(4+2\sigma)}.$$

where $C_{5i\mathcal{D}}(t_1,t_2,t_3,t_4) \equiv \left\|\varepsilon_{it_1}\varepsilon_{it_3}\bar{X}_{it_1}\bar{X}_{it_3}\right\|_{4+2\sigma,\mathcal{D}}\left\|\bar{X}_{it_2}X_{it_2}\delta_i\bar{X}_{it_4}X_{it_4}\delta_i\right\|_{2+\sigma,\mathcal{D}}$ satisfies $C_{5i\mathcal{D}}(t_1,t_2,t_3,t_4)$ $\le \sum_{q=1}^{4}\bar{C}_{5i\mathcal{D}}(t_q)$ where $\bar{C}_{5i\mathcal{D}}(t)=\frac{1}{2}[\left\|\varepsilon_{it}\bar{X}_{it}\right\|^2_{8+4\sigma,\mathcal{D}}+\left\|\bar{X}_{it}X_{it}\delta_i\right\|^2_{4+2\sigma,\mathcal{D}}]$. Then $ED_{11a}(3)\le CT^{-1}N^{-3/2}$ $\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{5i\mathcal{D}}(t)\sum_{s=1}^{T}\alpha_{\mathcal{D}}(s)^{(1+2\sigma)/(4+2\sigma)}=O_P\left(N^{-1/2}\right)$. Consequently, $D_{11}=O_P(N^{-1/2})=o_P(1)$.

Note that if either $t_1>t_2$ or $t_3>t_4$, the term inside the summation for ED_{12} is 0 by Assumption A.2(iii). This, in conjunction with the fact that $(\varepsilon_{it},X_{it})$ are independent across i conditional on $\mathcal{D}$ by Assumption A.2(ii), leads to

$$ED_{12}=T^{-3}N^{-3/2}\sum_{1\le i<j\le N}\sum_{1\le t_1<t_2\le T}E_{\mathcal{D}}(\varepsilon_{it_1}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i)\sum_{1\le t_3<t_4\le T}E_{\mathcal{D}}(\varepsilon_{jt_3}\bar{X}'_{jt_3}\bar{X}_{jt_4}X'_{jt_4}\delta_j).$$

By Davydov's inequality,

$$\left|E_{\mathcal{D}}(\varepsilon_{it_1}\bar{X}_{it_1}\bar{X}_{it_2}X_{it_2}\delta_i)\right|\le 8C_{6i\mathcal{D}}(t_1,t_2)\,\alpha_{\mathcal{D}}(|t_2-t_1|)^{(5+4\sigma)/(8+4\sigma)}$$

where $C_{6i\mathcal{D}}(t_1,t_2)\equiv\left\|\varepsilon_{it_1}\bar{X}_{it_1}\right\|_{8+4\sigma,\mathcal{D}}\left\|\bar{X}_{it_2}X'_{it_2}\delta_i\right\|_{4+2\sigma,\mathcal{D}}$ satisfies $C_{6i\mathcal{D}}(t_1,t_2)\le\sum_{q=1}^{2}\bar{C}_{6i\mathcal{D}}(t_q)$ where $\bar{C}_{6i\mathcal{D}}(t)=\frac{1}{2}\{\left\|\varepsilon_{it}\bar{X}_{it}\right\|^2_{8+4\sigma,\mathcal{D}}+\left\|\bar{X}_{it}X_{it}\delta_i\right\|^2_{4+2\sigma,\mathcal{D}}\}$, and similar result holds for $E_{\mathcal{D}}(\varepsilon_{jt_3}\bar{X}'_{jt_3}\bar{X}_{jt_4}X'_{jt_4}\delta_j)$. It follows that

$$ED_{12}\le CT^{-1}N^{1/2}\left\{N^{-1}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{6i\mathcal{D}}(t)\sum_{s=1}^{T}\alpha_{\mathcal{D}}(s)^{(5+4\sigma)/(8+4\sigma)}\right\}^2=O_P\left(T^{-1}N^{1/2}\right).$$

In sum we have shown that $E_{\mathcal{D}}\left(D^2_{3NT,11}\right)=O_P\left(N^{-1/2}+T^{-1}N^{1/2}\right)=o_P(1)$, implying that $D_{3NT,11}=o_P(1)$. Combining this with (C.3)-(C.4) yields $D_{3NT,1}=o_P(1)$.

For $D_{3NT,2}$, we further decompose it as follows

$$\begin{aligned}D_{3NT,2} &= \gamma_{NT}T^{-1}N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'P_{F^0}X_i\Omega_i^{-1}X_i'X_i\delta_i+\gamma_{NT}T^{-1}N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'P_{F^0}X_i(\Omega_i^{-1}-\hat{\Omega}_i^{-1})X_i'X_i\delta_i\\ &\equiv D_{3NT,21}+D_{3NT,22}\text{ say.}\end{aligned}\tag{C.5}$$

As in the study of $D_{3NT,12}$, we can readily determine that $D_{3NT,22}=O_P\left(N^{1/4}a_{NT}\right)=o_P(1)$.

$$\begin{aligned}E_{\mathcal{D}}\left(D^2_{3NT,21}\right) &= T^{-5}N^{-3/2}\sum_{i=1}^{N}\sum_{1\le t_1,t_2,\ldots,t_6\le T}F_{t_1t_2}F_{t_4t_5}E_{\mathcal{D}}\left(\varepsilon_{it_1}\varepsilon_{it_4}\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i\;\bar{X}'_{it_5}\bar{X}_{it_6}X'_{it_6}\delta_i\right)\\ &\quad+T^{-5}N^{-3/2}\sum_{1\le i\ne j\le N}\sum_{1\le t_1,t_2,\ldots,t_6\le T}F_{t_1t_2}F_{t_4t_5}E_{\mathcal{D}}\left(\varepsilon_{it_1}\varepsilon_{jt_4}\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i\;\bar{X}'_{jt_5}\bar{X}_{jt_6}X'_{jt_6}\delta_j\right)\\ &\equiv ED_{21}+ED_{22},\text{ say.}\end{aligned}$$

For ED_{21}, we consider two cases for the time indices $\{t_1,t_2,...,t_6\}$ inside the summation: (a) $\#\{t_1,t_2,...,t_6\}$ $=6$ and (b) $\#\{t_1,t_2,...,t_6\}\le 5$. We use ED_{21a} and ED_{21b} to denote ED_{21} when the individual indices in the summation are restricted to be cases (a) and (b), respectively. It is easy to see that $ED_{21b}=O_P\left(N^{-1/2}\right)$. For case (a), if either t_1 or t_4 is largest in the set $\{t_1,t_2,...,t_6\}$, then $E_{\mathcal{D}}(\varepsilon_{it_1}\varepsilon_{it_4}\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i$

$\bar{X}'_{it_5}\bar{X}_{it_6}X'_{it_6}\delta_i) = 0$; otherwise, one can use the m.d.s. property of $\{\varepsilon_{it}, \mathcal{F}_{NT,t}\}$ and Davydov's inequality as above to bound the last expectation and obtain

$$ED_{21a} \leq CT^{-1}N^{-3/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{7i\mathcal{D}}(t)\left\{T^{-2}\sum_{i_1=1}^{T}\sum_{t_2=1}^{T}F_{t_1t_2}\right\}^2 = O_P\left(N^{-1/2}\right) = o_P(1) \qquad \text{(C.6)}$$

for some $\bar{C}_{7i\mathcal{D}}(t)$ with $N^{-1}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\bar{C}_{7i\mathcal{D}}(t) = O_P(1)$.

Note that if either $t_1 > (t_2 \vee t_3)$ or $t_4 > (t_5 \vee t_6)$ where $a \vee b \equiv \max(a,b)$, the term inside the summation for D_{22} is 0. This, in conjunction with the fact that $(\varepsilon_{it}, X_{it})$ are independent across i conditional on $\mathcal{D}$, leads to

$$\begin{aligned} ED_{22} &= T^{-5}N^{-3/2}\sum_{1\leq i\neq j\leq N}\sum_{1\leq t_1<(t_2\vee t_3)\leq T}\sum_{1\leq t_4<(t_5\vee t_6)\leq T}F_{t_1t_2}F_{t_4t_5}E_{\mathcal{D}}(\varepsilon_{it_1}\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i) \\ &\quad \times E_{\mathcal{D}}(\varepsilon_{jt_4}\bar{X}'_{jt_5}\bar{X}_{jt_6}X'_{jt_6}\delta_j). \end{aligned}$$

Note that

$$\begin{aligned} &\left|E_{\mathcal{D}}(\varepsilon_{it_1}\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i)\right| \\ &\leq \begin{cases} 8\alpha_{\mathcal{D}}\left((t_2\wedge t_3)-t_1\right)^{(1+\sigma)/(2+\sigma)}\|\varepsilon_{it_1}\|_{8+4\sigma,\mathcal{D}}\left\|\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{(8+4\sigma)/3,\mathcal{D}} & \text{if } t_1<t_2 \text{ and } t_1<t_3 \\ 8\alpha_{\mathcal{D}}\left(t_3-t_1\right)^{(1+\sigma)/(2+\sigma)}\left\|\varepsilon_{it_1}\bar{X}_{it_2}\right\|_{4+2\sigma,\mathcal{D}}\left\|\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{4+2\sigma,\mathcal{D}} & \text{if } t_2<t_1<t_3 \\ 8\alpha_{\mathcal{D}}\left(t_2-t_1\right)^{(1+\sigma)/(2+\sigma)}\left\|\varepsilon_{it_1}\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{(8+4\sigma)/3,\mathcal{D}}\left\|\bar{X}_{it_2}\right\|_{8+4\sigma,\mathcal{D}} & \text{if } t_3<t_1<t_2 \end{cases} \end{aligned}$$

where $a \wedge b \equiv \min(a,b)$. Similar result holds for $E_{\mathcal{D}}(\varepsilon_{jt_4}\bar{X}'_{jt_5}\bar{X}_{jt_6}X'_{jt_6}\delta_j)$. It is easy to see that

$$\begin{aligned} \|\varepsilon_{it_1}\|_{8+4\sigma,\mathcal{D}}\left\|\bar{X}'_{it_2}\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{(8+4\sigma)/3,\mathcal{D}} &\leq \frac{1}{4}\{2\|\varepsilon_{it_1}\|^2_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it_2}\right\|_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{4+2\sigma,\mathcal{D}}\} \\ \left\|\varepsilon_{it_1}\bar{X}_{it_2}\right\|_{4+2\sigma,\mathcal{D}}\left\|\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{4+2\sigma,\mathcal{D}} &\leq \frac{1}{4}\{\|\varepsilon_{it_1}\|_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it_2}\right\|_{8+4\sigma,\mathcal{D}} + 2\left\|\bar{X}_{it_3}X'_{it_3}\delta_i\right\|^2_{4+2\sigma,\mathcal{D}}\} \\ \left\|\varepsilon_{it_1}\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{(8+4\sigma)/3,\mathcal{D}}\left\|\bar{X}_{it_2}\right\|_{8+4\sigma,\mathcal{D}} &\leq \frac{1}{4}\{\|\varepsilon_{it_1}\|_{8+4\sigma,\mathcal{D}} + 2\left\|\bar{X}_{it_2}\right\|^2_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it_3}X'_{it_3}\delta_i\right\|_{4+2\sigma,\mathcal{D}}\}. \end{aligned}$$

Let $\bar{C}_{8i\mathcal{D}}(t) \equiv \|\varepsilon_{it}\|^2_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it}\right\|^2_{8+4\sigma,\mathcal{D}} + \left\|\bar{X}_{it}X'_{it}\delta_i\right\|^2_{4+2\sigma,\mathcal{D}} + 1$. Then

$$D_{22} \leq CT^{-1}N^{1/2}\left\{T^{-2}N^{-1}\sum_{i=1}^{N}\sum_{1\leq t_1,t_2\leq T}\sum_{q=1}^{2}\bar{C}_{8i\mathcal{D}}(t_q)F_{t_1t_2}\sum_{s=1}^{T}\alpha_{\mathcal{D}}(s)^{(1+\sigma)/(2+\sigma)}\right\}^2 = O_P\left(T^{-1}N^{1/2}\right).$$

This, together with (C.5) and (C.6), implies that $E_{\mathcal{D}}\left(D^2_{3NT,21}\right) = o_P(1)$. Then $D_{3NT,2} = o_P(1)$.

The proofs of $D_{3NT,3} = o_P(1)$ and $D_{3NT,4} = o_P(1)$ are analogous to those of $D_{3NT,1} = o_P(1)$ and $D_{3NT,3} = o_P(1)$, respectively, and thus are omitted. This completes the proof of (i). ■

D Proof of Theorem 3.2

By Theorem 3.1 and the Slutsky lemma, it suffices to prove the first two parts of the theorem. In fact, we prove a slightly stronger result, i.e., under $\mathbb{H}_1(\gamma_{NT})$ in (3.5), (i) $\hat{B}_{NT} = B_{NT} + o_P(1)$ and (ii)

$\hat{V}_{NT} = V_{NT} + o_P(1)$.

Step 1. We prove (i) $\hat{B}_{NT} = B_{NT} + o_P(1)$ **under** $\mathbb{H}_1(\gamma_{NT})$. Let $\iota_s \equiv \iota_{Ts}$ denote a $T \times 1$ vector with 1 in the t'th position and zeros everywhere else. Then $h_{i,ts} = \sum_{r=1}^{T}\sum_{q=1}^{T}\eta_{tr}X_{ir}'(X_iX_i)^{-1}X_{iq}\eta_{qs} = \iota_t' M_{F^0}P_{X_i}M_{F^0}\iota_s$ and $\hat{h}_{i,ts} = \iota_t' M_{\hat{F}}P_{X_i}M_{\hat{F}}\iota_s$. We decompose $\hat{B}_{NT} - B_{NT}$ as follows:

$$\begin{aligned}
\hat{B}_{NT} - B_{NT} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{it}^2\hat{h}_{i,tt} - \varepsilon_{it}^2 h_{i,tt}\right) \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{it}^2 - \varepsilon_{it}^2\right)h_{i,tt} + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{h}_{i,tt} - h_{i,tt}\right)\varepsilon_{it}^2 \\
&\quad + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\varepsilon}_{it}^2 - \varepsilon_{it}^2\right)\left(\hat{h}_{i,tt} - h_{i,tt}\right) \\
&\equiv \hat{B}_{NT,1} + \hat{B}_{NT,2} + \hat{B}_{NT,3}, \text{ say.}
\end{aligned}$$

We prove (i) by showing $(i1)$ $\hat{B}_{NT,1} = o_P(1)$, $(i2)$ $\hat{B}_{NT,2} = o_P(1)$, and $(i3)$ $\hat{B}_{NT,3} = o_P(1)$.

First, we prove $(i1)$ $\hat{B}_{NT,1} = o_P(1)$ **under** $\mathbb{H}_1(\gamma_{NT})$. Using $\hat{\varepsilon}_{it}^2 - \varepsilon_{it}^2 = (\hat{\varepsilon}_{it} - \varepsilon_{it})^2 + 2(\hat{\varepsilon}_{it} - \varepsilon_{it})\varepsilon_{it}$, we have

$$\hat{B}_{NT,1} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{\varepsilon}_{it} - \varepsilon_{it})^2 h_{i,tt} + \frac{2}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{\varepsilon}_{it} - \varepsilon_{it})\varepsilon_{it}h_{i,tt} \equiv \hat{B}_{NT,11} + 2\hat{B}_{NT,12}, \text{ say.}$$

Noting that diag(H_i) is p.s.d., we have by Cauchy-Schwarz's inequality

$$\begin{aligned}
\hat{B}_{1NT,11} &= N^{-1/2}\sum_{i=1}^{N}(\hat{\varepsilon}_i - \varepsilon_i)'\operatorname{diag}(H_i)(\hat{\varepsilon}_i - \varepsilon_i) \\
&\leq 3N^{-1/2}\sum_{i=1}^{N}(d_{1i} - \varepsilon_i)'\operatorname{diag}(H_i)(d_{1i} - \varepsilon_i) + 3N^{-1/2}\sum_{i=1}^{N}d_{2i}'\operatorname{diag}(H_i)d_{2i} \\
&\quad + 3N^{-1/2}\sum_{i=1}^{N}d_{3i}'\operatorname{diag}(H_i)d_{3i} \\
&\equiv 3\hat{B}_{NT,111} + 3\hat{B}_{NT,112} + 3\hat{B}_{NT,113}, \text{ say.}
\end{aligned}$$

We will show that $\hat{B}_{NT,11s} = o_P(1)$ for $s = 1, 2, 3$.

Using $d_{1i} = M_{F^0}(\varepsilon_i + c_i) + \sum_{k=1}^{K}(\beta_k^0 - \hat{\beta}_k)M_k^{(0)}\left(\varepsilon_i + F^0\lambda_i^0 + c_i\right)$ and $M_{F^0} = I_T - P_{F^0}$ we have

$$\begin{aligned}
\hat{B}_{NT,111} &\leq 3N^{-1/2}\sum_{i=1}^{N}\varepsilon_i' P_{F^0}\operatorname{diag}(H_i)P_{F^0}\varepsilon_i + 3N^{-1/2}\sum_{i=1}^{N}c_i' M_{F^0}\operatorname{diag}(H_i)M_{F^0}c_i \\
&\quad + 3N^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{K}(\beta_k^0 - \hat{\beta}_k)(\beta_l^0 - \hat{\beta}_l)\left(\varepsilon_i + F^0\lambda_i^0 + c_i\right)' M_k^{(0)}\operatorname{diag}(H_i)M_l^{(0)}\left(\varepsilon_i + F^0\lambda_i^0 + c_i\right) \\
&\equiv 3\hat{B}_{NT,111a} + 3\hat{B}_{NT,111b} + 3\hat{B}_{NT,111c}, \text{ say.}
\end{aligned}$$

By the fact that tr$(AB) \leq$tr(A)tr(B) for p.s.d. matrices A and B, tr$(AB) \leq$tr$(A)\,\mu_1(B)$ for symmetric matrix B and p.s.d. matrix A, $\|M_{F^0}P_{X_i}M_{F^0}\|_F \leq K^{1/2}$, tr$(M_{F^0}P_{X_i}M_{F^0}) \leq K$, and $\sum_{t=1}^{T}\iota_t\iota_t' = I_T$ we have

$$
\begin{aligned}
& \max_{1\le i\le N}\left\|F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\|_F \\
= & \max_{1\le i\le N}\mathrm{tr}\left[F^{0\prime}\mathrm{diag}\left(H_i\right)F^0F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right]^{1/2}\le \max_{1\le i\le N}\mathrm{tr}\left[F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right] \\
= & \max_{1\le i\le N}\sum_{t=1}^T\mathrm{tr}\left(F_t^0\iota_t^{\prime}M_{F^0}P_{X_i}M_{F^0}\iota_tF_t^{0\prime}\right)=\max_{1\le i\le N}\mathrm{tr}\left(M_{F^0}P_{X_i}M_{F^0}\sum_{t=1}^T\iota_t\iota_t^{\prime}F_t^{0\prime}F_t^0\right) \\
\le & K\mu_1\left(\sum_{t=1}^T\iota_t\iota_t^{\prime}\left\|F_t^0\right\|_F^2\right)\le K\max_{1\le t\le T}\left\|F_t^0\right\|_F^2=o_P\left(T^{1/(4+2\sigma)}\right)
\end{aligned}
$$

where the last equality follows because by Boole's and Chebyshev's inequalities, Assumption A.1(i), and the dominated convergence theorem, for any $C>0$

$$
\begin{aligned}
P\left(\max_{1\le t\le T}\left\|F_t^0\right\|_F^2\ge (CT)^{1/(4+2\sigma)}\right) & \le \sum_{t=1}^T P\left(\left\|F_t^0\right\|_F^2\ge (CT)^{1/(4+2\sigma)}\right) \\
& \le \frac{1}{CT}\sum_{t=1}^T E\left[\left\|F_t^0\right\|_F^{8+4\sigma}1\left\{\left\|F_t^0\right\|_F^{8+4\sigma}\ge CT\right\}\right]=o(1).
\end{aligned}
$$

Similarly, $\max_{1\le i\le N}\mathrm{tr}[X_i^{\prime}\mathrm{diag}\left(H_i\right)X_i]=\max_{1\le i\le N}\sum_{t=1}^T\mathrm{tr}(X_{it}\iota_t^{\prime}M_{F^0}P_{X_i}M_{F^0}\iota_tX_{it}^{\prime})\le K\max_{1\le i\le N}\max_{1\le t\le T}$ $\left\|X_{it}\right\|_F^2=o_P((NT)^{1/(4+2\sigma)})$. It follows that

$$
\begin{aligned}
\left|\hat{B}_{NT,111a}\right| & = T^{-2}N^{-1/2}\sum_{i=1}^N\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\left(F^{0\prime}F^0/T\right)^{-1}F^{0\prime}\varepsilon_i\varepsilon_i^{\prime}F^0\left(F^{0\prime}F^0/T\right)^{-1}\right\} \\
& \le T^{-2}N^{-1/2}\sum_{i=1}^N\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\}\mathrm{tr}\left\{\left(F^{0\prime}F^0/T\right)^{-1}F^{0\prime}\varepsilon_i\varepsilon_i^{\prime}F^0\left(F^{0\prime}F^0/T\right)^{-1}\right\} \\
& \le T^{-2}N^{-1/2}\max_{1\le i\le N}\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\}\left[\mu_{\min}\left(\left(F^{0\prime}F^0/T\right)\right)\right]^{-2}\sum_{i=1}^N\left\|F^{0\prime}\varepsilon_i\right\|_F^2 \\
& = T^{-2}N^{-1/2}o_P\left(T^{1/(4+2\sigma)}\right)O_P(NT)=o_P\left(T^{-1}N^{1/2}T^{1/(4+2\sigma)}\right)=o_P(1)\text{ by Assumption A.3}(i).
\end{aligned}
$$

For $\hat{B}_{NT,111b}$, we use $M_{F^0}=I_T-P_{F^0}$ to decompose it as follows:

$$
\begin{aligned}
\left|\hat{B}_{NT,111b}\right| & = N^{-1/2}\sum_{i=1}^N c_i^{\prime}M_{F^0}\mathrm{diag}\left(H_i\right)M_{F^0}c_i \\
& = N^{-1/2}\sum_{i=1}^N c_i^{\prime}\mathrm{diag}\left(H_i\right)c_i+N^{-1/2}\sum_{i=1}^N c_i^{\prime}P_{F^0}\mathrm{diag}\left(H_i\right)P_{F^0}c_i-2N^{-1/2}\sum_{i=1}^N c_i^{\prime}\mathrm{diag}\left(H_i\right)P_{F^0}c_i \\
& \equiv \hat{B}_{1NT,111b}(1)+\hat{B}_{1NT,111b}(2)-2\hat{B}_{1NT,111b}(3),\text{ say.}
\end{aligned}
$$

Then

$$
\begin{aligned}
\hat{B}_{NT,111b}(1) &\leq N^{-1/2}\sum_{i=1}^{N}\left[\left(\beta^0-\hat{\beta}\right)+\gamma_{BT}\delta_i\right]' X_i'\mathrm{diag}\left(H_i\right)X_i\left[\left(\beta^0-\hat{\beta}\right)+\gamma_{BT}\delta_i\right] \\
&\leq O_P\left(\gamma_{NT}^2\right)N^{-1/2}\sum_{i=1}^{N}\mathrm{tr}\left(X_i'\mathrm{diag}\left(H_i\right)X_i\right)=O_P\left(T^{-1}\right)\max_{1\leq i\leq N}\mathrm{tr}\left[X_i'\mathrm{diag}\left(H_i\right)X_i\right] \\
&= O_P\left(T^{-1}\left(NT\right)^{1/(4+2\sigma)}\right)=o_P\left(1\right), \\
\hat{B}_{NT,111b}(2) &= N^{-1/2}\sum_{i=1}^{N}\mathrm{tr}\left(F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\left(F^{0\prime}F^0\right)^{-1}F^{0\prime}c_ic_i'F^0\left(F^{0\prime}F^0\right)^{-1}\right) \\
&\leq \max_{1\leq i\leq N}\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\}N^{-1/2}\sum_{i=1}^{N}\mathrm{tr}\left(c_ic_i'F^0\left(F^{0\prime}F^0\right)^{-1}\left(F^{0\prime}F^0\right)^{-1}F^{0\prime}\right) \\
&\leq \max_{1\leq i\leq N}\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\}N^{-1/2}\sum_{i=1}^{N}\mathrm{tr}\left(c_ic_i'\right)\mathrm{tr}\left(F^0\left(F^{0\prime}F^0\right)^{-1}\left(F^{0\prime}F^0\right)^{-1}F^{0\prime}\right) \\
&\leq \max_{1\leq i\leq N}\mathrm{tr}\left\{F^{0\prime}\mathrm{diag}\left(H_i\right)F^0\right\}T^{-1}\gamma_{NT}^2N^{-1/2}\sum_{i=1}^{N}\left\|X_i\right\|_F^2\mathrm{tr}\left(\left(F^{0\prime}F^0/T\right)^{-1}\right) \\
&= o_P\left(T^{1/(4+2\sigma)}\right)T^{-1}\gamma_{NT}^2O_P\left(N^{1/2}T\right)=T^{-1}O_P\left(T^{1/(4+2\sigma)}\right)=o_P\left(1\right),
\end{aligned}
$$

and $\hat{B}_{NT,111b}(3)\leq\{\hat{B}_{NT,111b}(1)\}^{1/2}\{\hat{B}_{NT,111b}(2)\}^{1/2}=o_P(1)$ by Cauchy-Schwarz's inequality. It follows that $\hat{B}_{NT,111b}=o_P(1)$. For $\hat{B}_{NT,111c}$, we have

$$
\begin{aligned}
\hat{B}_{NT,111c} &= N^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\left(\beta_l^0-\hat{\beta}_l\right)\left(\varepsilon_i+F^0\lambda_i^0+c_i\right)'M_k^{(0)}\mathrm{diag}\left(H_i\right)M_l^{(0)}\left(\varepsilon_i+F^0\lambda_i^0+c_i\right) \\
&\leq 3N^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\left(\beta_l^0-\hat{\beta}_l\right)\varepsilon_i'M_k^{(0)}\mathrm{diag}\left(H_i\right)M_l^{(0)}\varepsilon_i \\
&\quad +3N^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\left(\beta_l^0-\hat{\beta}_l\right)\lambda_i^{0\prime}FM_k^{(0)}\mathrm{diag}\left(H_i\right)M_l^{(0)}F^0\lambda_i^0 \\
&\quad +3N^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\left(\beta_l^0-\hat{\beta}_l\right)c_i'M_k^{(0)}\mathrm{diag}\left(H_i\right)M_l^{(0)}c_i \\
&\equiv 3\hat{B}_{NT,111c}(1)+3\hat{B}_{NT,111c}(2)+3\hat{B}_{NT,111c}(3),\ \text{say.}
\end{aligned}
$$

By Cauchy-Schwarz's inequality and the fact that $\mu_1(\mathrm{diag}(H_i))\leq 1$, we have

$$
\begin{aligned}
\hat{B}_{NT,111c}(1) &\leq KN^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)^2\varepsilon_i'M_k^{(0)}\mathrm{diag}\left(H_i\right)M_k^{(0)}\varepsilon_i \\
&\leq KN^{1/2}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)^2\mathrm{tr}\left(M_k^{(0)}M_k^{(0)}N^{-1}\sum_{i=1}^{N}\varepsilon_i\varepsilon_i'\right) \\
&\leq KN^{1/2}\left\{\sum_{k=1}^{K}(\beta_k^0-\hat{\beta}_k)^2\left\|M_k^{(0)}\right\|_F^2\right\}\mu_1\left(N^{-1}\sum_{i=1}^{N}\varepsilon_i\varepsilon_i'\right) \\
&= N^{1/2}O_P\left(\gamma_{NT}^2\right)O_P\left(1+T/N\right)=O_P\left(T^{-1}+N^{-1}\right)=o_P\left(1\right)
\end{aligned}
$$

where we have used the fact that

$$\mu_1\left(N^{-1}\sum_{i=1}^{N}\varepsilon_i\varepsilon_i'\right) \leq \mu_1\left(N^{-1}\sum_{i=1}^{N}E\left(\varepsilon_i\varepsilon_i'\right)\right)+\left\|N^{-1}\sum_{i=1}^{N}\left[\varepsilon_i\varepsilon_i'-E\left(\varepsilon_i\varepsilon_i'\right)\right]\right\|_F = O(1)+O_P(T/N).$$

Similarly, using the fact that $\left\|M_k^{(0)}F^0\right\|_F=O_P(1)$ we have

$$\begin{aligned}
\hat{B}_{NT,111c}(2) &\leq KN^{-1/2}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)^2\lambda_i^{0\prime}FM_k^{(0)}M_k^{(0)}F^0\lambda_i^0\\
&\leq KN^{-1/2}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)^2\left\|M_k^{(0)}F^0\right\|_F^2\mu_1\left(N^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\prime}\right)\\
&= N^{1/2}O_P\left(\gamma_{NT}^2\right)O_P(1)=O_P\left(T^{-1}\right)=o_P(1), \text{ and}\\
\hat{B}_{NT,111c}(3) &\leq pN^{1/2}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)^2\left\|M_k^{(0)}F^0\right\|_F^2 N^{-1}\sum_{i=1}^{N}\|c_i\|^2\\
&= N^{1/2}O_P\left(\gamma_{NT}^2\right)O_P\left(\gamma_{NT}^2T\right)=O_P\left(N^{-1/2}T^{-1}\right)=o_P(1).
\end{aligned}$$

It follows that $\hat{B}_{NT,111c}=o_P(1)$. In sum, $\hat{B}_{NT,111}=o_P(1)$.

For $\hat{B}_{NT,112}$, we have

$$\begin{aligned}
\hat{B}_{NT,112} &\leq 3N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'M^{(1)}\text{diag}(H_i)M^{(1)}\varepsilon_i+3N^{-1/2}\sum_{i=1}^{N}\lambda_i^{0\prime}F^{0\prime}M^{(1)}\text{diag}(H_i)M^{(1)}F^0\lambda_i^0\\
&\quad+3N^{-1/2}\sum_{i=1}^{N}c_i'M^{(1)}\text{diag}(H_i)M^{(1)}c_i\\
&\equiv 3\hat{B}_{NT,112a}+3\hat{B}_{NT,112b}+3\hat{B}_{NT,112c}, \text{ say.}
\end{aligned}$$

Observe that

$$\begin{aligned}
\hat{B}_{NT,112a} &\leq N^{-1/2}\sum_{i=1}^{N}\varepsilon_i'M^{(1)}M^{(1)}\varepsilon_i=N^{1/2}\text{tr}\left(M^{(1)}M^{(1)}N^{-1}\sum_{i=1}^{N}\varepsilon_i\varepsilon_i'\right)\\
&\leq N^{1/2}\left\|M^{(1)}\right\|_F^2\mu_1\left(N^{-1}\sum_{i=1}^{N}\varepsilon_i\varepsilon_i'\right)=N^{1/2}O_P\left(N^{-1}\right)O_P\left(1+TN^{-1}\right)\\
&= o_P(1) \text{ by Assumption A.3}(i).
\end{aligned}$$

Similarly, we can show that $\hat{B}_{NT,112b}=o_P(1)$ and $\hat{B}_{NT,112c}=o_P(1)$. Thus $\hat{B}_{NT,112}=o_P(1)$.

By the fact that $\mu_1(\text{diag}(H_i))\leq\mu_1(H_i)\leq 1$, (B.2)-(B.3) and Assumption A.3(i), we have

$$\begin{aligned}
\hat{B}_{NT,113} &\leq N^{-1/2}\sum_{i=1}^{N}\|d_{3i}\|^2\leq\left\|M^{(2)}+M^{(rem)}\right\|^2N^{-1/2}\sum_{i=1}^{N}\left\|\varepsilon_i+F^0\lambda_i^0+c_i\right\|^2\\
&= O_P\left(\delta_{NT}^{-4}\right)O_P\left(N^{1/2}T\right)=o_P(1).
\end{aligned}$$

In sum, we have shown that $\hat{B}_{NT,11}=o_P(1)$.

Now, we consider $\hat{B}_{NT,12}$. We first decompose it as follows

$$\begin{aligned}
\hat{B}_{NT,12} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\hat{\varepsilon}_i-\varepsilon_i)'\operatorname{diag}(H_i)\varepsilon_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(d_{1i}-\varepsilon_i)'\operatorname{diag}(H_i)\varepsilon_i+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}d_{2i}'\operatorname{diag}(H_i)\varepsilon_i+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}d_{3i}'\operatorname{diag}(H_i)\varepsilon_i \\
&\equiv \hat{B}_{NT,121}+\hat{B}_{NT,122}+\hat{B}_{NT,123},\ \text{say.}
\end{aligned}$$

We only show that $\hat{B}_{NT,121}=o_P(1)$ since the proof for $\hat{B}_{NT,12s}=o_P(1)$ for $s=2,3$ is similar. We further decompose $\hat{B}_{NT,121}$ as follows:

$$\begin{aligned}
\hat{B}_{NT,121} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(P_{F^0}\varepsilon_i+M_{F^0}c_i+\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)M_k^{(0)}\left(\varepsilon_i+F^0\lambda_i^0+c_i\right)\right)'\operatorname{diag}(H_i)\varepsilon_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varepsilon_i'P_{F^0}\operatorname{diag}(H_i)\varepsilon_i+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}c_i'M_{F^0}\operatorname{diag}(H_i)\varepsilon_i \\
&\quad +\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)M_k^{(0)}\left(\varepsilon_i+F^0\lambda_i^0+c_i\right)'\operatorname{diag}(H_i)\varepsilon_i \\
&\equiv \hat{B}_{NT,121a}+\hat{B}_{NT,121b}+\hat{B}_{NT,121c},\ \text{say.}
\end{aligned}$$

Observe that

$$\begin{aligned}
\hat{B}_{NT,121a} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0\iota_s'M_{F^0}P_{X_i}M_{F^0}\iota_s\varepsilon_{is} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0\iota_s'P_{X_i}\iota_s\varepsilon_{is} \\
&\quad +\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0\iota_s'P_{F^0}P_{X_i}P_{F^0}\iota_s\varepsilon_{is} \\
&\quad -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0\iota_s'P_{F^0}P_{X_i}\iota_s\varepsilon_{is} \\
&\quad -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0\iota_s'P_{X_i}P_{F^0}\iota_s\varepsilon_{is} \\
&\equiv \hat{B}_{NT,121a}(1)+\hat{B}_{NT,121a}(2)-\hat{B}_{NT,121a}(3)-\hat{B}_{NT,121a}(4),\ \text{say.}
\end{aligned}$$

As in the proof of Lemma A.8, replacing $(X_i'X_i/T)^{-1}$ by Ω_i^{-1} and then applying Chebyshev's inequality yields

$$\begin{aligned}
\hat{B}_{NT,121a}(1) &= T^{-1}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0X_{is}'(X_i'X_i/T)^{-1}X_{is}\varepsilon_{is} \\
&= T^{-1}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F_s^0X_{is}'\Omega_i^{-1}X_{is}\varepsilon_{is}+o_P(1) \\
&= O_P\left(T^{-1}N^{1/2}\right)+o_P(1)=o_P(1).
\end{aligned}$$

Similarly, we can show that $\hat{B}_{NT,121a}(s) = o_P(1)$ for $s = 2,3,4$. It follows that $\hat{B}_{NT,121a} = o_P(1)$. For $\hat{B}_{NT,121b}$, we have

$$\begin{aligned}
\hat{B}_{NT,121b} &= N^{-1/2}\sum_{i=1}^{N}\left[(\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}})+\gamma_{NT}\delta_i\right]' X_i' M_{F^0}\mathrm{diag}(H_i)\varepsilon_i \\
&= N^{-1/2}(\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}})'\sum_{i=1}^{N} X_i' M_{F^0}\mathrm{diag}(H_i)\varepsilon_i + N^{-1/2}\gamma_{NT}\sum_{i=1}^{N}\delta_i' X_i' M_{F^0}\mathrm{diag}(H_i)\varepsilon_i \\
&= O_P\left(\gamma_{NT}N^{1/2}\right) = O_P\left(N^{1/4}T^{-1/2}\right) = o_P(1) \text{ by Assumption A.3}\,(i)
\end{aligned}$$

as we can readily show that $\left\| N^{-1/2}\sum_{i=1}^{N} X_i' M_{F^0}\mathrm{diag}(H_i)\varepsilon_i \right\|_F = O_P\left(N^{1/2}\right)$. Similarly, we can show that $\hat{B}_{NT,121c} = o_P(1)$. It follows that $\hat{B}_{NT,121} = o_P(1)$. Analogously, we can $\hat{B}_{NT,12s} = o_P(1)$ for $s = 2,3$ and thus we have $\hat{B}_{NT,12} = o_P(1)$. This completes the proof of $(i1)$ $\hat{B}_{NT,1} = o_P(1)$ under $\mathbb{H}_1(\gamma_{NT})$.

Next, we prove $(i2)$ $\hat{B}_{NT,2} = o_P(1)$ **under** $\mathbb{H}_1(\gamma_{NT})$. Observing that

$$\begin{aligned}
\hat{H}_i - H_i &= M_{\hat{F}}P_{X_i}M_{\hat{F}} - M_{F^0}P_{X_i}M_{F^0} \\
&= (M_{\hat{F}}-M_{F^0})P_{X_i}(M_{\hat{F}}-M_{F^0}) + M_{F^0}P_{X_i}(M_{\hat{F}}-M_{F^0}) + (M_{\hat{F}}-M_{F^0})P_{X_i}M_{F^0} \quad \text{(D.1)}
\end{aligned}$$

we have

$$\begin{aligned}
\hat{B}_{NT,2} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t'\left(M_{\hat{F}}P_{X_i}M_{\hat{F}} - M_{F^0}P_{X_i}M_{F^0}\right)\iota_t\varepsilon_{it}^2 \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t'(M_{\hat{F}}-M_{F^0})P_{X_i}(M_{\hat{F}}-M_{F^0})\iota_t\varepsilon_{it}^2 + \frac{2}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' M_{F^0}P_{X_i}(M_{\hat{F}}-M_{F^0})\iota_t\varepsilon_{it}^2 \\
&\equiv \hat{B}_{NT,21} + 2\hat{B}_{NT,22}, \text{ say.} \quad \text{(D.2)}
\end{aligned}$$

By the fact that $\left\|M_{\hat{F}}-M_{F^0}\right\|_F = O_P(N^{-1/2})$ by (B.1)-(B.3) and noticing that

$$\begin{aligned}
\mu_1\left(N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t\iota_t'\varepsilon_{it}^2\right) &\leq \mu_1\left(\max_{1\leq s\leq T}\left[N^{-1}\sum_{i=1}^{N}E\left(\varepsilon_{is}^2\right)\right]\sum_{t=1}^{T}\iota_t\iota_t'\right) + \left\|N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t\iota_t'\left[\varepsilon_{it}^2 - E\left(\varepsilon_{it}^2\right)\right]\right\|_F \\
&= \max_{1\leq s\leq T}\left[N^{-1}\sum_{i=1}^{N}E\left(\varepsilon_{is}^2\right)\right] + O_P(T/N), \quad \text{(D.3)}
\end{aligned}$$

we have by Assumption A.3(i)

$$\begin{aligned}
\left|\hat{B}_{NT,21}\right| &\leq \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t'(M_{\hat{F}}-M_{F^0})(M_{\hat{F}}-M_{F^0})\iota_t\varepsilon_{it}^2 \\
&= \mathrm{tr}\left[(M_{\hat{F}}-M_{F^0})(M_{\hat{F}}-M_{F^0})\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t\iota_t'\varepsilon_{it}^2\right)\right] \\
&\leq N^{1/2}\left\|M_{\hat{F}}-M_{F^0}\right\|_F^2\mu_1\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t\iota_t'\varepsilon_{it}^2\right) \\
&= N^{1/2}O_P\left(N^{-1}\right)O_P(1+T/N) = O_P(N^{-1/2}+TN^{-3/2}) = o_P(1). \quad \text{(D.4)}
\end{aligned}$$

Using (B.1) we decompose $\hat{B}_{NT,22}$ as follows

$$\begin{aligned}\hat{B}_{NT,22} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' M_{F^0} P_{X_i}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)M_k^{(0)}\iota_t\varepsilon_{it}^2 \\ &\quad +\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' M_{F^0} P_{X_i} M^{(1)}\iota_t\varepsilon_{it}^2+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' M_{F^0} P_{X_i}(M^{(2)}+M^{(rem)})\iota_t\varepsilon_{it}^2 \\ &\equiv \hat{B}_{NT,221}+\hat{B}_{NT,222}+\hat{B}_{NT,223}, \text{ say.}\end{aligned}$$

Using $M_{F^0}=I_T-P_{F^0}$ we further decompose $\hat{B}_{1NT,221}$ as follows

$$\begin{aligned}\hat{B}_{1NT,221} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' P_{X_i}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)M_k^{(0)}\iota_t\varepsilon_{it}^2 \\ &\quad -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' P_{F^0} P_{X_i}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)M_k^{(0)}\iota_t\varepsilon_{it}^2 \\ &\equiv \hat{B}_{NT,221a}-\hat{B}_{NT,221b}, \text{ say.}\end{aligned}$$

Following the arguments used in the proof of Lemma A.7(iii), we can show that $\max_{1\le t\le T} N^{-1}\sum_{i=1}^{N}\varepsilon_{it}^2 = O_P\left(1+\bar{a}_{NT}\right)$ where $\bar{a}_{NT}$ is defined as a_{NT} with the exchange of N and T in the latter's definition. This, together with the fact that $\|AB\|_F \le \|A\|_F\,\mu_1(B)$ for p.s.d. matrix A and symmetric matrix B and $\|P_{X_i}\|_F \le K^{1/2}$, implies that

$$\begin{aligned}\left|\hat{B}_{NT,221a}\right| &= \left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\text{tr}\left(P_{X_i}M_k^{(0)}\sum_{t=1}^{T}\iota_t\iota_t'\varepsilon_{it}^2\right)\right| \\ &= \left|\frac{1}{\sqrt{N}}\sum_{k=1}^{K}\left(\beta_k^0-\hat{\beta}_k\right)\text{tr}\left(M_k^{(0)}\sum_{t=1}^{T}\iota_t\iota_t'\sum_{i=1}^{N}\varepsilon_{it}^2 P_{X_i}\right)\right| \\ &\le \frac{\left\|\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}\right\|_F}{\sqrt{N}}\sum_{k=1}^{K}\left\|M_k^{(0)}\right\|_F\left\|\sum_{t=1}^{T}\iota_t\iota_t'\max_{1\le t\le T}\sum_{i=1}^{N}\varepsilon_{it}^2 P_{X_i}\right\|_F \\ &\le O_P\left(N^{-1/2}\gamma_{NT}\right)\sum_{k=1}^{K}\left\|M_k^{(0)}\right\|_F\max_{1\le t\le T}\sum_{i=1}^{N}\varepsilon_{it}^2 = O_P\left(N^{1/2}\gamma_{NT}\right)\left(\max_{1\le t\le T}N^{-1}\sum_{i=1}^{N}\varepsilon_{it}^2\right) \\ &= O_P\left(N^{1/4}T^{-1/2}\right)O_P\left(1+\bar{a}_{NT}\right)=o_P(1) \text{ by Assumption A.3.}\end{aligned}$$

Similarly, we can show that $\hat{B}_{NT,221b}=o_P(1)$ by symmetrizing $P_{F^0}P_{X_i}$ by $(P_{F^0}P_{X_i}+P_{X_i}P_{F^0})/2$. Thus $\hat{B}_{NT,221}=o_P(1)$. For $\hat{B}_{NT,222}$, we have $\hat{B}_{NT,222}=N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\iota_t' P_{X_i}M^{(1)}\iota_t\varepsilon_{it}^2 - N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}$

$\iota_t' P_{F^0} P_{X_i} M^{(1)} \iota_t \varepsilon_{it}^2 \equiv \hat{B}_{NT,222a} - \hat{B}_{NT,222b}$. Note that

$$
\begin{aligned}
\left|\hat{B}_{NT,222a}\right| &= \left|\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}' \left(X_i' X_i\right)^{-1} X_i' M^{(1)} \iota_t \varepsilon_{it}^2\right| \\
&= \left|\frac{1}{\sqrt{N}} \operatorname{tr}\left(M^{(1)} \sum_{t=1}^{T} \iota_t \sum_{i=1}^{N} \varepsilon_{it}^2 X_{it}' \left(X_i' X_i\right)^{-1} X_i'\right)\right| \\
&\leq N^{-1/2} \left\|M^{(1)}\right\|_F \left\|\sum_{t=1}^{T} \iota_t \sum_{i=1}^{N} \varepsilon_{it}^2 X_{it}' \left(X_i' X_i\right)^{-1} X_i'\right\|_F \\
&= N^{-1/2} \left\|M^{(1)}\right\|_F \left\{\operatorname{tr}\left[\sum_{t=1}^{T} \iota_t \sum_{i=1}^{N} \varepsilon_{it}^2 X_{it}' \left(X_i' X_i\right)^{-1} X_i' \sum_{s=1}^{T} \sum_{j=1}^{N} \varepsilon_{js}^2 X_j' \left(X_j' X_j\right)^{-1} X_{js} \iota_s'\right]\right\}^{1/2} \\
&= N^{-1/2} \left\|M^{(1)}\right\|_F \left\{\sum_{t=1}^{T} \sum_{i=1}^{N} \varepsilon_{it}^2 X_{it}' \left(X_i' X_i\right)^{-1} X_i' \sum_{j=1}^{N} \varepsilon_{jt}^2 X_j' \left(X_j' X_j\right)^{-1} X_{jt}\right\}^{1/2} \\
&= N^{-1/2} O_P\left(N^{-1/2}\right) O_P\left(N T^{-1/2}\right) = O_P\left(T^{-1/2}\right) = o_P(1).
\end{aligned}
$$

Similarly, we can show that $\hat{B}_{NT,222b} = O_P\left(T^{-1/2}\right) = o_P(1)$. Thus $\hat{B}_{NT,222b} = o_P(1)$. Next, write $\hat{B}_{NT,223} = N^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \iota_t' P_{X_i} (M^{(2)} + M^{(rem)}) \iota_t \varepsilon_{it}^2 - N^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \iota_t' P_{F^0} P_{X_i} (M^{(2)} + M^{(rem)}) \iota_t \varepsilon_{it}^2$ $\equiv \hat{B}_{NT,223a} - \hat{B}_{NT,223b}$. Then

$$
\begin{aligned}
\left|\hat{B}_{NT,223a}\right| &= \left|\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \operatorname{tr}\left\{P_{X_i} (M^{(2)} + M^{(rem)}) \sum_{t=1}^{T} \iota_t \iota_t' \varepsilon_{it}^2\right\}\right| \\
&= \left|\frac{1}{\sqrt{N}} \operatorname{tr}\left\{(M^{(2)} + M^{(rem)}) \sum_{t=1}^{T} \iota_t \iota_t' \sum_{i=1}^{N} \varepsilon_{it}^2 P_{X_i}\right\}\right| \\
&\leq \frac{1}{\sqrt{N}} \left\|(M^{(2)} + M^{(rem)})\right\|_F \left\|\sum_{t=1}^{T} \iota_t \iota_t' \sum_{i=1}^{N} \varepsilon_{it}^2 P_{X_i}\right\|_F \\
&= O_P\left(\sqrt{N} \delta_{NT}^{-2}\right) \max_{1 \leq t \leq T} N^{-1} \sum_{i=1}^{N} \varepsilon_{it}^2 = O_P\left(\sqrt{N} \delta_{NT}^{-2}\right) O_P\left(1 + \bar{a}_{NT}\right) = o_P(1).
\end{aligned}
$$

Similarly, $\hat{B}_{NT,223b} = o_P(1)$. Thus $\hat{B}_{NT,223} = o_P(1)$. In sum, we have shown that $\hat{B}_{NT,22} = o_P(1)$, which, in conjunction with (D.2) and (D.4), implies that $\hat{B}_{NT,2} = o_P(1)$.

Now, we prove $(i3)$ $\hat{B}_{NT,3} = o_P(1)$ **under** $\mathbb{H}_1(\gamma_{NT})$. Using $a^2 - b^2 = (a-b)^2 + 2(a-b)b$ we have

$$
\begin{aligned}
\hat{B}_{NT,3} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{\varepsilon}_{it} - \varepsilon_{it}\right)^2 \left(\hat{h}_{i,tt} - h_{i,tt}\right) + \frac{2}{\sqrt{N}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{\varepsilon}_{it} - \varepsilon_{it}\right) \varepsilon_{it} \left(\hat{h}_{i,tt} - h_{i,tt}\right) \\
&\equiv \hat{B}_{NT,31} + 2\hat{B}_{NT,32}, \text{ say.}
\end{aligned}
$$

Using the arguments as used in the proof of Lemma A.7(iii), we can show that $\max_{1 \leq i \leq N} T^{-1} \|X_i\|_F^2 = O_P(1 + a_{NT}) = O_P(1)$. Then by (B.1) and the fact that $\left\|M_{\hat{F}} - M_{F^0}\right\|_F = O_P\left(N^{-1/2}\right)$, $\max_{1 \leq i \leq N} \max_{1 \leq t \leq T}$

$\|X_{it}\| = o_P((NT)^{1/(8+4\sigma)})$ and $\max_{1\le t\le T}\left\|F_t^0\right\| = o_P(T^{1/(8+4\sigma)})$, we have that uniformly in (i,t)

$$\begin{aligned}
&\left|\iota_t' M_{F^0} P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
\le\ & \left|\iota_t' P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| + \left|\iota_t' P_{F^0} P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
\le\ & \left|X_{it}'\left(X_i'X_i\right)^{-1} X_i(M_{\hat F} - M_{F^0})\iota_t\right| + \left|F_t^{0\prime}\left(F_0'F_0\right)^{-1} F^0 P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
\le\ & T^{-1}\left\|M_{\hat F} - M_{F^0}\right\|_F \left\{\left\|(X_i'X_i/T)^{-1}\right\|_F \|X_i\|_F\|X_{it}\| + \left\|(F_0'F_0/T)^{-1}\right\| \left\|F^0\right\|_F \left\|F_t^0\right\| \|P_{X_i}\|_F\right\} \\
=\ & O_P\left(T^{-1}N^{-1/2}\right)\left\{\|X_i\|_F\|X_{it}\| + \left\|F^0\right\|\left\|F_t^0\right\|\right\} = o_P\left(T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)}\right). \qquad \text{(D.5)}
\end{aligned}$$

By (D.1) and (D.5)

$$\begin{aligned}
\max_{1\le i\le N}\max_{1\le t\le T}\left|\hat h_{i,tt} - h_{i,tt}\right| &\le \max_{1\le i\le N}\max_{1\le t\le T}\left|\iota_t'(M_{\hat F} - M_{F^0})P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
&\quad +2\max_{1\le i\le N}\max_{1\le t\le T}\left|\iota_t' M_{F^0} P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
&\le \left\|M_{\hat F} - M_{F^0}\right\|_F^2 + 2\max_{1\le i\le N}\max_{1\le t\le T}\left|\iota_t' M_{F^0} P_{X_i}(M_{\hat F} - M_{F^0})\iota_t\right| \\
&= O_P\left(N^{-1}\right) + o_P\left(T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)}\right). \qquad \text{(D.6)}
\end{aligned}$$

As in the study of $\hat B_{NT,11}$, we can readily show that

$$\begin{aligned}
N^{-1}\sum_{i=1}^N \|\hat\varepsilon_i - \varepsilon_i\|^2 &\le 3N^{-1}\sum_{i=1}^N \|d_{1i} - \varepsilon_i\|^2 + 3N^{-1}\sum_{i=1}^N \|d_{2i}\|^2 + 3N^{-1}\sum_{i=1}^N \|d_{3i}\|^2 \\
&= O_P(1) + O_P\left(1 + TN^{-1}\right) + o_P(1) = O_P\left(1 + TN^{-1}\right). \qquad \text{(D.7)}
\end{aligned}$$

It follows from (D.5) and (D.7) and Assumptions A.3 (i)-(ii) that

$$\begin{aligned}
\hat B_{NT,31} &\le N^{1/2}\max_{1\le i\le N}\max_{1\le t\le T}\left|\hat h_{i,tt} - h_{i,tt}\right|\left\{N^{-1}\sum_{i=1}^N \|\hat\varepsilon_i - \varepsilon_i\|^2\right\} \\
&= N^{1/2}\left(O_P\left(N^{-1}\right) + o_P\left(T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)}\right)\right) O_P\left(1 + TN^{-1}\right) = o_P(1).
\end{aligned}$$

We now decompose $\hat B_{1NT,32}$ as follows:

$$\begin{aligned}
\hat B_{NT,32} &= \frac{1}{\sqrt N}\sum_{i=1}^N (\hat\varepsilon_i - \varepsilon_i)' \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i \\
&= \frac{1}{\sqrt N}\sum_{i=1}^N \left\{(d_{1i} - \varepsilon_i)' \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i + d_{2i}' \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i + d_{3i}' \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i\right\} \\
&\equiv \hat B_{NT,321} + \hat B_{NT,322} + \hat B_{NT,323}, \text{ say.}
\end{aligned}$$

Using the expression for d_{1i} we further decompose $\hat B_{NT,321}$ as follows:

$$\begin{aligned}
\hat B_{NT,321} &= \frac{1}{\sqrt N}\sum_{i=1}^N \varepsilon_i' P_{F^0} \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i + \frac{1}{\sqrt N}\sum_{i=1}^N c_i' M_{F^0} \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i \\
&\quad + \frac{1}{\sqrt N}\sum_{i=1}^N\sum_{k=1}^K \left(\beta_k^0 - \hat\beta_k\right)\left(\varepsilon_i + \lambda_i^{0\prime} F^{0\prime} + c_i\right) M_k^{(0)} \operatorname{diag}\left(\hat H_i - H_i\right)\varepsilon_i \\
&\equiv \hat B_{NT,321a} + \hat B_{NT,321b} + \hat B_{NT,321c}, \text{ say.}
\end{aligned}$$

To study the terms on the right hand of the last equation, we find that it is convenient to bound $\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\left(\hat{h}_{i,tt}-h_{i,tt}\right)F_t^0\varepsilon_{it}\right\|$ and $\sum_{i=1}^{N}\sum_{t=1}^{T}|\hat{h}_{i,tt}-h_{i,tt}|\,\|X_{it}\varepsilon_{it}\|$. Notice that by (D.1) and the triangle inequality,

$$\begin{aligned}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\left(\hat{h}_{i,tt}-h_{i,tt}\right)F_t^0\varepsilon_{it}\right\| &\leq \sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\iota_t'(M_{\hat{F}}-M_{F^0})P_{X_i}(M_{\hat{F}}-M_{F^0})\iota_tF_t^0\varepsilon_{it}\right\| \\ &\quad +2\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\iota_t'M_{F^0}P_{X_i}(M_{\hat{F}}-M_{F^0})\iota_tF_t^0\varepsilon_{it}\right\| \\ &\equiv S_{1NT}+2S_{2NT},\ \text{say.}\end{aligned}$$

By (B.1), $\left\|(M_{\hat{F}}-M_{F^0})\iota_t\right\|_F\leq\left\|\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}\right\|\sum_{k=1}^{K}\left\|M_k^{(0)}\iota_t\right\|_F+\left\|M^{(1)}\iota_t\right\|_F+\left\|\left(M^{(2)}+M^{(rem)}\right)\iota_t\right\|_F$. It is easy to show that uniformly in t, $\left\|M_k^{(0)}\iota_t\right\|_F=O_P\left(T^{-1/2}\right)\left\|F_t^0\right\|$, $\left\|M^{(1)}\iota_t\right\|_F=O_P\left(N^{-1/2}T^{-1/2}\right)(1+\left\|F_t^0\right\|)=O_P\left(N^{-1/2}T^{-1/2}\right)\left\|F_t^0\right\|$, and $\left\|\left(M^{(2)}+M^{(rem)}\right)\iota_t\right\|_F=O_P\left(\delta_{NT}^{-2}+\delta_{NT}^{-1}\gamma_{NT}\right)=O_P\left(\delta_{NT}^{-2}\right)$. Thus

$$\left\|(M_{\hat{F}}-M_{F^0})\iota_t\right\|_F=O_P\left(N^{-1/2}T^{-1/2}\left\|F_t^0\right\|+\delta_{NT}^{-2}\right). \tag{D.8}$$

It follows that

$$\begin{aligned}S_{1NT} &\leq \sum_{i=1}^{N}\sum_{t=1}^{T}O_P\left(N^{-1/2}T^{-1/2}\left\|F_t^0\right\|+\delta_{NT}^{-2}\right)^2\left\|P_{X_i}\right\|_F\left\|F_t^0\varepsilon_{it}\right\| \\ &= O_P\left(N^{-2}+T^{-2}\right)O_P\left(NT\right)=O_P\left(NT^{-1}+N^{-1}T\right).\end{aligned}$$

Using $M_{F^0}=I_T-P_{F^0}$,

$$\begin{aligned}S_{2NT} &\leq \sum_{i=1}^{N}\left\|\sum_{t=1}^{T}X_{it}'\left(X_i'X_i\right)^{-1}X_i'(M_{\hat{F}}-M_{F^0})\iota_tF_t^0\varepsilon_{it}\right\|_F \\ &\quad +\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}F_t^{0\prime}\left(F_0'F_0\right)^{-1}F^0P_{X_i}(M_{\hat{F}}-M_{F^0})\iota_tF_t^0\varepsilon_{it}\right\|_F \\ &\equiv S_{2NT,1}+S_{2NT,2},\ \text{say.}\end{aligned}$$

Then by (D.8),

$$\begin{aligned}S_{2NT,1} &\leq \sum_{i=1}^{N}\left\|\left(X_i'X_i\right)^{-1}X_i'\right\|\sum_{t=1}^{T}O_P\left(N^{-1/2}T^{-1/2}\left\|F_t^0\right\|+\delta_{NT}^{-2}\right)\left\|X_{it}\right\|\left\|F_t^0\varepsilon_{it}\right\| \\ &= O_P\left(\delta_{NT}^{-2}\right)O_P\left(NT^{1/2}\right),\ \text{and} \\ S_{2NT,2} &\leq \sum_{i=1}^{N}\left\|\left(F_0'F_0\right)^{-1}F^0P_{X_i}\right\|_F\sum_{t=1}^{T}O_P\left(N^{-1/2}T^{-1/2}\left\|F_t^0\right\|+\delta_{NT}^{-2}\right)\left\|F_t^0\right\|\left\|F_t^0\varepsilon_{it}\right\| \\ &= O_P\left(\delta_{NT}^{-2}\right)O_P\left(NT^{1/2}\right).\end{aligned}$$

Consequently,

$$\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\left(\hat{h}_{i,tt}-h_{i,tt}\right)F_t^0\varepsilon_{it}\right\|=O_P\left(\delta_{NT}^{-2}NT^{1/2}\right). \tag{D.9}$$

Similarly, we can show that

$$\sum_{i=1}^{N}\sum_{s=1}^{T}\left|\hat{h}_{i,ss}-h_{i,ss}\right|\|X_{is}\varepsilon_{is}\|=O_P\left(\delta_{NT}^{-2}NT^{1/2}\right). \tag{D.10}$$

By (D.9) and Assumptions A.3(i)-(ii)

$$\begin{aligned}
\hat{B}_{NT,321a} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varepsilon_i' P_{F^0}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}F_t^0(F^{0\prime}F^0)^{-1}\sum_{s=1}^{T}F_s^0\left(\hat{h}_{i,ss}-h_{i,ss}\right)\varepsilon_{is} \\
&\leq \frac{1}{\sqrt{N}}\left\|(F^{0\prime}F^0/T)^{-1}\right\|\max_{1\leq i\leq N}\left\|T^{-1}\sum_{t=1}^{T}\varepsilon_{it}F_t^0\right\|\sum_{i=1}^{N}\left\|\sum_{s=1}^{T}F_s^0\left(\hat{h}_{i,ss}-h_{i,ss}\right)\varepsilon_{is}\right\| \\
&= O_P\left(N^{-1/2}\right)O_P\left(a_{NT}\right)O_P\left(\delta_{NT}^{-2}NT^{1/2}\right)=O_P\left(N^{1/2}T^{1/2}a_{NT}\delta_{NT}^{-2}\right)=o_P(1),
\end{aligned}$$

where we have used the fact that $\max_{1\leq i\leq N}\left\|T^{-1}\sum_{t=1}^{T}\varepsilon_{it}F_t^0\right\|=O_P\left(a_{NT}\right)$ which can be proved by following the arguments used in the proof of Lemma A.7(iii). Observe that

$$\begin{aligned}
\left|\hat{B}_{NT,321b}\right| &\leq \left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}c_i'\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\right|+\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}c_i'P_{F^0}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\right| \\
&\equiv \hat{B}_{NT,321b}(1)+\hat{B}_{NT,321b}(2).
\end{aligned}$$

By (D.10) and Assumption A.3(i),

$$\begin{aligned}
\hat{B}_{NT,321b}(1) &= \left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\left(\boldsymbol{\beta}_0-\hat{\boldsymbol{\beta}}\right)+\gamma_{NT}\delta_i\right]' X_{it}\varepsilon_{it}\left(\hat{h}_{i,tt}-h_{i,tt}\right)\right| \\
&\leq O_P\left(\gamma_{NT}\right)\sum_{i=1}^{N}\sum_{t=1}^{T}\|X_{it}\varepsilon_{it}\|\left|\hat{h}_{i,tt}-h_{i,tt}\right| \\
&\leq O_P\left(\gamma_{NT}\right)O_P\left(\delta_{NT}^{-2}NT^{1/2}\right)=O_P\left(\delta_{NT}^{-2}N^{3/4}\right)=o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
&\hat{B}_{NT,321b}(2) \\
&= \left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\left(\boldsymbol{\beta}_0-\hat{\boldsymbol{\beta}}\right)+\gamma_{NT}\delta_i\right]' X_{it}F_t^{0\prime}\left(F^{0\prime}F^0\right)^{-1}F^{0\prime}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\right| \\
&\leq O_P\left(\gamma_{NT}\right)\left\|\left(F^{0\prime}F^0\right)^{-1}\right\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\|X_{it}\|\left\|F_t^0\right\|\left\|\sum_{s=1}^{T}\left(\hat{h}_{i,ss}-h_{i,ss}\right)F_s^0\varepsilon_{is}\right\| \\
&\leq O_P\left(\gamma_{NT}N^{-1/2}\right)\left\{\max_{1\leq i\leq N}T^{-1}\sum_{t=1}^{T}\|X_{it}\|\left\|F_t^0\right\|\right\}\sum_{i=1}^{N}\left\|\sum_{s=1}^{T}\left(\hat{h}_{i,ss}-h_{i,ss}\right)F_s^0\varepsilon_{is}\right\| \\
&= O_P\left(T^{-1/2}N^{-3/4}\right)O_P\left(1+a_{NT}\right)O_P\left(\delta_{NT}^{-2}NT^{1/2}\right)=O_P\left(\delta_{NT}^{-2}N^{1/4}\right)=o_P(1),
\end{aligned}$$

where we use the fact that $\max_{1\leq i\leq N}T^{-1}\sum_{t=1}^{T}\|X_{it}\|\left\|F_t^0\right\|=O_P\left(1+a_{NT}\right)$ which can be proved by using arguments as used in the proof of Lemma A.7(iii). It follows that $\hat{B}_{NT,321b}=o_P(1)$. Using the fact that $\left\|M_k^{(0)}(\varepsilon_i+F^0\lambda_i^0+c_i)\right\|=O_P(1)\left\{\left\|\lambda_i^0\right\|+(NT)^{-1/2}\left\|\mathbf{X}_k\varepsilon_i\right\|_F+\gamma_{NT}\left\|X_i\right\|_F\right\}$ and $\left\|\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}\right\|=O_P(\gamma_{NT})$ and Assumption A.3, we have

$$\begin{aligned}\left|\hat{B}_{NT,321c}\right| &= \left|\frac{1}{\sqrt{N}}\sum_{i=1}^N\sum_{k=1}^K\left(\beta_k^0-\hat{\beta}_k\right)\left(\varepsilon_i'+\lambda_i^{0\prime}F^{0\prime}+c_i\right)M_k^{(0)}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\right| \\ &\le O_P(\gamma_{NT})\max_{1\le i\le N}\max_{1\le t\le T}\left|\hat{h}_{i,tt}-h_{i,tt}\right|\frac{1}{\sqrt{N}}\sum_{i=1}^N\left(\left\|\lambda_i^0\right\|+(NT)^{-1/2}\left\|\mathbf{X}_k\varepsilon_i\right\|_F+\gamma_{NT}\left\|X_i\right\|_F\right)\left\|\varepsilon_i\right\| \\ &= O_P(\gamma_{NT})\,O_P\left(N^{-1}+T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)}\right)N^{1/2}T^{1/2}=o_P(1).\end{aligned}$$

In sum, we have shown that $\hat{B}_{NT,321}=o_P(1)$.

Now, we study $\hat{B}_{NT,322}$. We first decompose $\hat{B}_{NT,322}$ as follows:

$$\begin{aligned}\hat{B}_{NT,322} &= \frac{1}{\sqrt{N}}\sum_{i=1}^N\left(\varepsilon_i'+\lambda_i^{0\prime}F^{0\prime}+c_i'\right)M^{(1)}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^N\{\varepsilon_i'M^{(1)}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i+\lambda_i^{0\prime}F^{0\prime}M^{(1)}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i+c_i'M^{(1)}\text{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\} \\ &\equiv \hat{B}_{NT,322a}+\hat{B}_{NT,322b}+\hat{B}_{NT,322c},\ \text{say.}\end{aligned}$$

Note that $\left|\iota_t'M^{(1)}\iota_s\right|=O_P\left(N^{-1}T^{-1}\right)\left\{\left\|\sum_{i=1}^N\varepsilon_{it}\lambda_t^0\right\|\left\|F_s^0\right\|+\left\|\sum_{i=1}^N\varepsilon_{is}\lambda_s^0\right\|\left\|F_t^0\right\|\right\}+O_P\left(N^{-1/2}T^{-3/2}\right)$ $\left\|F_s^0\right\|\left\|F_t^0\right\|$. By Boole's inequality, Doob's inequality (e.g., Hall and Heyde (1980, pp.14-15)) for m.d.s., and then Davydov's inequality, for any $\epsilon>0$ we have

$$\begin{aligned}P_{\mathcal{D}}\left(\max_{1\le t\le T}\left\|N^{-1/2}\sum_{i=1}^N\varepsilon_{it}\lambda_t^0\right\|_F>T^{1/8}\epsilon\right) &\le \sum_{t=1}^T P_{\mathcal{D}}\left(\left\|N^{-1/2}\sum_{i=1}^N\varepsilon_{it}\lambda_t^0\right\|_F>T^{1/8}\epsilon\right) \\ &\le \frac{1}{TN^4\epsilon^8}\sum_{t=1}^T E_{\mathcal{D}}\left\|\sum_{i=1}^N\varepsilon_{it}\lambda_t^0\right\|_F^8=O_P(1).\end{aligned}$$

It follows that $\max_{1\le t\le T}\left\|\sum_{i=1}^N\varepsilon_{it}\lambda_t^0\right\|=O_P\left(N^{1/2}T^{1/8}\right)$. So uniformly in (t,s) we have

$$\left|\iota_t'M^{(1)}\iota_s\right|=O_P\left(N^{-1/2}T^{-7/8}\right)\left(\left\|F_s^0\right\|+\left\|F_t^0\right\|\right)+O_P\left(N^{-1/2}T^{-3/2}\right)\left\|F_s^0\right\|\left\|F_t^0\right\|.$$

Then

$$\begin{aligned}\left|\hat{B}_{NT,322a}\right| &= \frac{1}{\sqrt{N}}\left|\sum_{i=1}^N\sum_{t=1}^T\varepsilon_{it}\sum_{s=1}^T\iota_t'M^{(1)}\iota_s\left(\hat{h}_{i,ss}-h_{i,ss}\right)\varepsilon_{is}\right| \\ &\le N^{-1/2}\sum_{i=1}^N\sum_{t=1}^T\left\|\varepsilon_{it}\right\|\sum_{s=1}^T\left\|\varepsilon_{is}\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right| \\ &\quad\times\left\{O_P\left(N^{-1/2}T^{-7/8}\right)\left(\left\|F_s^0\right\|+\left\|F_t^0\right\|\right)+O_P\left(N^{-1/2}T^{-3/2}\right)\left\|F_s^0\right\|\left\|F_t^0\right\|\right\} \\ &\le O_P\left(N^{-1/2}T^{-7/8}\right)N^{-1/2}\sum_{i=1}^N\sum_{t=1}^T\left\|\varepsilon_{it}\right\|\sum_{s=1}^T\left\|\varepsilon_{is}\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right|\left\|F_s^0\right\| \\ &\quad+O_P\left(N^{-1/2}T^{-7/8}\right)N^{-1/2}\sum_{i=1}^N\sum_{t=1}^T\left\|\varepsilon_{it}\right\|\left\|F_t^0\right\|\sum_{s=1}^T\left\|\varepsilon_{is}\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right|\left\|F_s^0\right\|\end{aligned}$$

$$
\begin{aligned}
&+O_P\left(N^{-1/2}T^{-3/2}\right)N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\varepsilon_{it}\|\left\|F_t^0\right\|\sum_{s=1}^{T}\|\varepsilon_{is}\|\left|\hat{h}_{i,ss}-h_{i,ss}\right|\left\|F_s^0\right\| \\
\equiv\ & \hat{B}_{NT,322a}(1)+\hat{B}_{NT,322a}(2)+\hat{B}_{NT,322a}(3)\,,\text{ say.}
\end{aligned}
$$

By Assumption A.3(i),

$$
\begin{aligned}
\hat{B}_{NT,322a}(1) &\leq O_P\left(N^{-1}T^{1/8}\right)\left\{\max_{1\leq i\leq N}T^{-1}\sum_{t=1}^{T}\|\varepsilon_{it}\|\right\}\sum_{i=1}^{N}\sum_{s=1}^{T}\|\varepsilon_{is}\|\left|\hat{h}_{i,ss}-h_{i,ss}\right|\left\|F_s^0\right\| \\
&= O_P\left(N^{-1}T^{1/8}\right)O_P(1)\,O_P\left(\delta_{NT}^{-2}NT^{1/2}\right)=O_P\left(\delta_{NT}^{-2}T^{5/8}\right)=o_P(1)\,.
\end{aligned}
$$

Similarly, $\hat{B}_{NT,322a}(s)=o_P(1)$ for $s=2,3$. So $\hat{B}_{1NT,322a}=o_P(1)$. By the same token, we can show that $\hat{B}_{NT,322b}=o_P(1)$ and $\hat{B}_{NT,322c}=o_P(1)$. It follows that $\hat{B}_{NT,322}=o_P(1)$.

For $\hat{B}_{NT,323}$, we have

$$
\begin{aligned}
\hat{B}_{NT,323} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}d_{3i}'\operatorname{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\{\varepsilon_i'[M^{(2)}+M^{(rem)}]\operatorname{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i+\lambda_i^{0\prime}F^{0\prime}[M^{(2)}+M^{(rem)}]\operatorname{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i \\
&\quad +c_i'[M^{(2)}+M^{(rem)}]\operatorname{diag}\left(\hat{H}_i-H_i\right)\varepsilon_i\} \\
&\equiv \hat{B}_{NT,323a}+\hat{B}_{NT,323b}+\hat{B}_{NT,323c},\text{ say.}
\end{aligned}
$$

Noting that $\iota_t'[M^{(2)}+M^{(rem)}]\iota_s=O_P\left(\delta_{NT}^{-2}T^{-7/8}\right)\left(\left\|F_s^0\right\|+\left\|F_t^0\right\|\right)+O_P\left(\delta_{NT}^{-2}T^{-3/2}\right)\left\|F_s^0\right\|\left\|F_t^0\right\|+O_P\left(\delta_{NT}^{-1}\gamma_{NT}\right)$, we have

$$
\begin{aligned}
\left|\hat{B}_{NT,323a}\right| &= \left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}\iota_t'[M^{(2)}+M^{(rem)}]\iota_s\left(\hat{h}_{i,ss}-h_{i,ss}\right)\varepsilon_{is}\right| \\
&\leq O_P\left(\delta_{NT}^{-2}T^{-7/8}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}|\varepsilon_{it}|\left\|F_s^0\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right||\varepsilon_{is}| \\
&\quad +O_P\left(\delta_{NT}^{-2}T^{-7/8}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}|\varepsilon_{it}|\left\|F_t^0\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right||\varepsilon_{is}| \\
&\quad +O_P\left(\delta_{NT}^{-2}T^{-3/2}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}|\varepsilon_{it}|\left\|F_t^0\right\|\left\|F_s^0\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right||\varepsilon_{is}| \\
&\quad +O_P\left(\delta_{NT}^{-1}\gamma_{NT}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}|\varepsilon_{it}|\left|\hat{h}_{i,ss}-h_{i,ss}\right||\varepsilon_{is}| \\
&\equiv \hat{B}_{NT,323a}(1)+\hat{B}_{NT,323a}(2)+\hat{B}_{NT,323a}(3)+\hat{B}_{NT,323a}(4)\,.
\end{aligned}
$$

By Assumption A.3(i),

$$
\begin{aligned}
\hat{B}_{NT,323a}(1) &= O_P\left(N^{-1/2}\delta_{NT}^{-2}T^{1/8}\right)\left(\max_{1\leq i\leq N}T^{-1}\sum_{t=1}^{T}|\varepsilon_{it}|\right)\sum_{i=1}^{N}\sum_{s=1}^{T}\left\|F_s^0\right\|\left|\hat{h}_{i,ss}-h_{i,ss}\right||\varepsilon_{is}| \\
&= O_P\left(N^{-1/2}\delta_{NT}^{-2}T^{1/8}\right)O_P(1)\,O_P\left(\delta_{NT}^{-2}NT^{1/2}\right)=O_P\left(\delta_{NT}^{-4}N^{1/2}T^{5/8}\right)=o_P(1)\,.
\end{aligned}
$$

Similarly, $\hat{B}_{NT,323a}(s) = o_P(1)$ for $s = 2, 3, 4$. So $\hat{B}_{NT,323a} = o_P(1)$. By the same token, we can show that $\hat{B}_{NT,323b} = o_P(1)$ and $\hat{B}_{NT,323c} = o_P(1)$. It follows that $\hat{B}_{NT,323} = o_P(1)$.

In sum, we have shown that $\hat{B}_{NT,32} = o_P(1)$. This completes the proof of (i).

Step 2. We prove (ii) $\hat{V}_{NT} = V_{NT} + o_P(1)$ **under** $\mathbb{H}_1(\gamma_{NT})$. Recall $b_{it} = X_{it} - T^{-1}\sum_{r=1}^{T} F_{tr} E_{\mathcal{D}}(X_{ir})$, $\bar{b}_{it} = \Omega_i^{-1/2} b_{it}$, and $\hat{b}_{it} = \hat{\Omega}_i^{-1/2}[X_{it} - T^{-1}\sum_{r=1}^{T} \hat{F}_{tr} X_{ir}]$. It follows that

$$\begin{aligned}
\hat{V}_{NT} - V_{NT} &= 4T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\{\left[\hat{\varepsilon}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{\varepsilon}_{is}\right]^2 - \left[\varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2\right\} \\
&\quad + 4T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\{\left[\varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 - E_{\mathcal{D}}\left[\varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2\right\} \\
&\equiv 4V_{NT,1} + 4V_{NT,2} \text{ say.}
\end{aligned}$$

Noting that $E_{\mathcal{D}}(V_{NT,2}) = 0$ and $\text{Var}_{\mathcal{D}}(V_{NT,2}) = o(1)$ by direct moment calculations, we have $V_{NT,2} = o_P(1)$ by Chebyshev's inequality. Thus we are left to show that $V_{NT,1} = o_P(1)$. Again, using $a^2 - b^2 = (a-b)^2 + 2(a-b)b$ yields

$$\begin{aligned}
V_{NT,1} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\hat{\varepsilon}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{\varepsilon}_{is} - \varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\
&\quad + 2T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\hat{\varepsilon}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{\varepsilon}_{is} - \varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]\varepsilon_{it}\bar{b}_{it}'\sum_{r=1}^{t-1}\bar{b}_{ir}\varepsilon_{ir} \\
&\equiv V_{NT,11} + 2V_{NT,12}.
\end{aligned}$$

Let $\bar{V}_{NT,12} \equiv T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}[\varepsilon_{it}\bar{b}_{it}'\sum_{r=1}^{t-1}\bar{b}_{ir}\varepsilon_{ir}]^2$. By Cauchy-Schwarz's inequality $V_{NT,12} \leq \{V_{NT,11}\}^{1/2}$ $\{\bar{V}_{NT,12}\}^{1/2}$. It is straightforward to show that $\bar{V}_{NT,12} = O_P(1)$, implying that it suffices to prove that $V_{NT,1} = o_P(1)$ by showing that $V_{NT,11} = o_P(1)$. Using $\hat{\varepsilon}_{it}\hat{b}_{it} = (\hat{\varepsilon}_{it}\hat{b}_{it} - \varepsilon_{it}\bar{b}_{it}) + \varepsilon_{it}\bar{b}_{it}$ and Cauchy-Schwarz inequality,

$$\begin{aligned}
V_{NT,11} &\leq 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\left(\hat{\varepsilon}_{it}\hat{b}_{it} - \varepsilon_{it}\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\
&\quad + 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\varepsilon_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\left(\hat{b}_{is}\hat{\varepsilon}_{is} - \bar{b}_{is}\varepsilon_{is}\right)\right]^2 \\
&\quad + 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\left(\hat{\varepsilon}_{it}\hat{b}_{it} - \varepsilon_{it}\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\left(\hat{b}_{is}\hat{\varepsilon}_{is} - \bar{b}_{is}\varepsilon_{is}\right)\right]^2 \\
&\equiv 3V_{NT,111} + 3V_{NT,112} + 3V_{NT,113}.
\end{aligned}$$

We complete the proof of (ii) by showing that $(ii1)$ $V_{NT,111} = o_P(1)$, $(ii2)$ $V_{NT,112} = o_P(1)$, and $(ii3)$ $V_{NT,113} = o_P(1)$.

We first show $(ii1)$ $V_{NT,111}=o_P(1)$. Using $\hat{\varepsilon}_{it}\hat{b}_{it}-\varepsilon_{it}\bar{b}_{it}=(\hat{\varepsilon}_{it}-\varepsilon_{it})\bar{b}_{it}+\varepsilon_{it}(\hat{b}_{it}-\bar{b}_{it})+(\hat{\varepsilon}_{it}-\varepsilon_{it})(\hat{b}_{it}-\bar{b}_{it})$ and Cauchy-Schwarz's inequality,

$$\begin{aligned} V_{NT,111} &\leq 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{\varepsilon}_{it}-\varepsilon_{it})\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2+3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\varepsilon_{it}\left(\hat{b}_{it}-\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\ &\quad +3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{\varepsilon}_{it}-\varepsilon_{it})\left(\hat{b}_{it}-\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\ &\equiv 3V_{NT,111a}+3V_{NT,111b}+3V_{NT,111c}. \end{aligned}$$

By Markov's and Davydov's inequalities we can show that $T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}[\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}]^2=O_P(1)$. By Boole's inequality, Doob's inequality (e.g., Hall and Heyde (1980, pp.14-15)) for m.d.s., and then Davydov's inequality, for any $\epsilon>0$ we have

$$\begin{aligned} P_{\mathcal{D}}\left(\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|T^{-1/2}\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right\|_F>N^{1/8}\epsilon\right) &\leq \sum_{i=1}^{N}P_{\mathcal{D}}\left(\max_{2\leq t\leq T}\left\|T^{-1/2}\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right\|_F>N^{1/8}\epsilon\right) \\ &\leq \frac{1}{NT^4\epsilon^8}\sum_{i=1}^{N}E_{\mathcal{D}}\left\|\sum_{s=1}^{T-1}\bar{b}_{is}\varepsilon_{is}\right\|_F^8=O_P(1). \end{aligned}$$

It follows that $\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right\|=O_P\left(T^{1/2}N^{1/8}\right)$. Using this, the fact that $\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|\bar{b}_{it}\right\|_F=o_P((NT)^{1/(8+4\sigma)})$, and Assumption A.3 yields

$$\begin{aligned} V_{NT,111a} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{\varepsilon}_{it}-\varepsilon_{it})^2\left[\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\ &\leq T^{-2}\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|T^{-1/2}\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right\|_F^2\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|\bar{b}_{it}\right\|_F^2\left\{N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{\varepsilon}_{it}-\varepsilon_{it})^2\right\} \\ &= T^{-2}O_P\left(TN^{1/4}\right)o_P\left((NT)^{1/(4+2\sigma)}\right)O_P\left(1+TN^{-1}\right)=o_P(1). \end{aligned}$$

To determine the probability order of $V_{NT,111b}$ and $V_{NT,111c}$, we use the uniform probability order of $\hat{b}_{it}-\bar{b}_{it}$. We decompose $\hat{b}_{it}-\bar{b}_{it}$ as follows

$$\begin{aligned} \hat{b}_{it}-\bar{b}_{it} &= \hat{\Omega}_i^{-1/2}\left[X_{it}-T^{-1}\sum_{r=1}^{T}\hat{F}_{tr}X_{ir}\right]-\Omega_i^{-1/2}\left[X_{it}-T^{-1}\sum_{r=1}^{T}F_{tr}E_{\mathcal{D}}(X_{ir})\right] \\ &= \left(\hat{\Omega}_i^{-1/2}-\Omega_i^{-1/2}\right)X_{it}-\left(\hat{\Omega}_i^{-1/2}-\Omega_i^{-1/2}\right)T^{-1}\sum_{r=1}^{T}\left(\hat{F}_{tr}-F_{tr}\right)X_{ir} \\ &\quad -\left(\hat{\Omega}_i^{-1/2}-\Omega_i^{-1/2}\right)T^{-1}\sum_{r=1}^{T}F_{tr}X_{ir}-\Omega_i^{-1/2}T^{-1}\sum_{r=1}^{T}\left(\hat{F}_{tr}-F_{tr}\right)X_{ir} \\ &\quad -\Omega_i^{-1/2}T^{-1}\sum_{r=1}^{T}F_{tr}\left[X_{ir}-E_{\mathcal{D}}(X_{ir})\right] \\ &\equiv b_{1it}-b_{2it}-b_{3it}-b_{4it}-b_{5it},\ \text{say.} \end{aligned} \tag{D.11}$$

By Lemma A.7(iii) and the fact that $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|X_{it}\| = o_P((NT)^{1/(8+4\sigma)})$ by Boole's and Markov's inequalities, we have $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|b_{1it}\| = o_P(a_{NT}(NT)^{1/(8+4\sigma)})$. Similarly $\max_{1\leq i\leq N}$ $\max_{1\leq t\leq T}\|b_{3it}\| = o_P(a_{NT}T^{1/(8+4\sigma)})$ and $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|b_{5it}\| = o_P(a_{NT}T^{1/(8+4\sigma)})$. For b_{2it}, using Cauchy-Schwarz's inequality for matrices, the fact that $\sum_{r=1}^{T}\iota_r\iota_r' = I_T$, and (B.1)-(B.3) we have

$$
\begin{aligned}
&\left\|T^{-1}\sum_{r=1}^{T}\left(\hat{F}_{tr}-F_{tr}\right)X_{ir}\right\|_F^2 \\
=\ & T^{-2}\sum_{r=1}^{T}\sum_{s=1}^{T}\operatorname{tr}\left\{\iota_r'\left(M_{\hat{F}}-M_F^0\right)\iota_t\iota_t'\left(M_{\hat{F}}-M_F^0\right)\iota_s X_{is}'X_{ir}\right\} \\
\leq\ & T^{-2}\sum_{r=1}^{T}\sum_{s=1}^{T}\operatorname{tr}\left\{X_{ir}\iota_r'\left(M_{\hat{F}}-M_F^0\right)\left(M_{\hat{F}}-M_F^0\right)\iota_s X_{is}'\right\} \\
\leq\ & T^{-1}\sum_{r=1}^{T}\operatorname{tr}\left\{X_{ir}\iota_r'\left(M_{\hat{F}}-M_F^0\right)\left(M_{\hat{F}}-M_F^0\right)\iota_r X_{ir}'\right\} \\
=\ & T^{-1}\operatorname{tr}\left\{\left(M_{\hat{F}}-M_F^0\right)\left(M_{\hat{F}}-M_F^0\right)\sum_{r=1}^{T}\iota_r\iota_r'X_{ir}'X_{ir}\right\} \leq T^{-1}\left\|M_{\hat{F}}-M_F^0\right\|_F^2\mu_1\left(\sum_{r=1}^{T}\iota_r\iota_r'X_{ir}'X_{ir}\right) \\
\leq\ & T^{-1}\left\|M_{\hat{F}}-M_F^0\right\|_F^2\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|X_{it}\|^2 = o_P\left(T^{-1}N^{-1}(NT)^{1/(4+2\delta)}\right).
\end{aligned}
$$

Thus $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|b_{4it}\| = o_P(T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)})$ and $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\|b_{2it}\| = o_P(a_{NT}T^{-1/2}N^{-1/2}(NT)^{1/(8+4\sigma)})$. In sum, we have

$$
\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|\hat{b}_{it}-\bar{b}_{it}\right\| = o_P(a_{NT}(NT)^{1/(8+4\sigma)}).
$$

Then by Assumption A.3,

$$
\begin{aligned}
V_{NT,111b} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\varepsilon_{it}\left(\hat{b}_{it}-\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\
&\leq T^{-1}\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|\hat{b}_{it}-\bar{b}_{it}\right\|^2\ \max_{1\leq i\leq N}\max_{1\leq t\leq T}\left[\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2\left[T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\varepsilon_{it}^2\right] \\
&= T^{-1}o_P(a_{NT}^2(NT)^{1/(4+2\sigma)})O_P\left(TN^{1/4}\right)O_P(1) = o_P(N^{1/4}a_{NT}^2(NT)^{1/(4+2\sigma)}) = o_P(1),
\end{aligned}
$$

and

$$
\begin{aligned}
V_{NT,111c} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{\varepsilon}_{it}-\varepsilon_{it})\left(\hat{b}_{it}-\bar{b}_{it}\right)'\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right]^2 \\
&= T^{-2}\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|\hat{b}_{it}-\bar{b}_{it}\right\|^2\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|\sum_{s=1}^{t-1}\bar{b}_{is}\varepsilon_{is}\right\|^2 N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{\varepsilon}_{it}-\varepsilon_{it})^2 \\
&= T^{-2}o_P(a_{NT}^2(NT)^{1/(4+2\sigma)})\ O_P\left(TN^{1/4}\right)O_P\left(1+N^{-1}T\right) \\
&= o_P(T^{-1}N^{1/4}a_{NT}^2(NT)^{1/(4+2\sigma)}+N^{-3/4}a_{NT}^2(NT)^{1/(4+2\sigma)}) = o_P(1).
\end{aligned}
$$

It follows that $V_{NT,111} = o_P(1)$.

To show $(ii2)$ and $(ii3)$, we find that it is convenient to bound $S_{it} \equiv T^{-1/2} \sum_{s=1}^{t-1} (\hat{b}_{is}\hat{\varepsilon}_{is} - \bar{b}_{is}\varepsilon_{is})$. Using $\hat{\varepsilon}_{it}\hat{b}_{it} - \varepsilon_{it}\bar{b}_{it} = \varepsilon_{it}(\hat{b}_{it} - \bar{b}_{it}) + (\hat{\varepsilon}_{it} - \varepsilon_{it})\,\bar{b}_{it} + (\hat{\varepsilon}_{it} - \varepsilon_{it})\,(\hat{b}_{it} - \bar{b}_{it})$, we have

$$
\begin{aligned}
S_{it} &\equiv T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is}(\hat{b}_{is} - \bar{b}_{is}) + T^{-1/2} \sum_{s=1}^{t-1} (\hat{\varepsilon}_{is} - \varepsilon_{is})\,\bar{b}_{is} + T^{-1/2} \sum_{s=1}^{t-1} (\hat{\varepsilon}_{is} - \varepsilon_{is})\,(\hat{b}_{is} - \bar{b}_{is}) \\
&\equiv S_{1it} + S_{2it} + S_{3it}, \text{ say.}
\end{aligned}
$$

Using (D.11), we have $S_{1it} = T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is}\,(b_{1is} - b_{2is} - b_{3is} - b_{4is} - b_{5is}) \equiv S_{1it,1} - S_{1it,2} - S_{1it,3} - S_{1it,4} - S_{1it,5}$, say. Observe that

$$
\begin{aligned}
\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,1}\| &\le \max_{1\le i\le N} \left\| \hat{\Omega}_i^{-1/2} - \Omega_i^{-1/2} \right\| \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} X_{is}\varepsilon_{is} \right\| \\
&= O_P\left(a_{NT}\right) O_P\left(N^{1/8}\right) = o_P\left(1\right), \\
\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,2}\| &\le \max_{1\le i\le N} \max_{1\le s\le T} \|b_{2is}\| \max_{1\le i\le N} T^{-1/2} \sum_{s=1}^{T} |\varepsilon_{is}| \\
&= o_P\left(a_{NT} T^{-1/2} N^{-1/2} \left(NT\right)^{1/(8+4\sigma)}\right) O_P\left(T^{1/2}\right) = o_P\left(1\right), \\
\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,4}\| &\le \max_{1\le i\le N} \max_{1\le s\le T} \|b_{4is}\| \max_{1\le i\le N} T^{-1/2} \sum_{s=1}^{T} |\varepsilon_{is}| \\
&= o_P\left(T^{-1/2} N^{-1/2} \left(NT\right)^{1/(8+4\sigma)}\right) O_P\left(T^{1/2}\right) = o_P\left(1\right).
\end{aligned}
$$

In addition,

$$
\begin{aligned}
\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,3}\| &\le \max_{1\le i\le N} \max_{1\le t\le T} \left\| \hat{\Omega}_i^{-1/2} - \Omega_i^{-1/2} \right\| \left\| T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is} F_s^{0\prime} \left(F^{0\prime}F^0/T\right)^{-1} T^{-1} \sum_{r=1}^{T} F_r^0 X_{ir} \right\| \\
&\le \left\| \left(F^{0\prime}F^0/T\right)^{-1} \right\| \max_{1\le i\le N} \left\| \hat{\Omega}_i^{-1/2} - \Omega_i^{-1/2} \right\| \max_{1\le i\le N} \left\| T^{-1} \sum_{r=1}^{T} F_r^0 X_{ir} \right\| \\
&\quad \times \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is} F_s^{0\prime} \right\| \\
&= O_P\left(1\right) O_P\left(a_{NT}\right) O_P\left(1\right) O_P\left(N^{1/8}\right) = o_P\left(1\right),
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,5}\| &= \max_{1\le i\le N} \max_{1\le t\le T} \left\| \Omega_i^{-1/2} T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is} F_s^{0\prime} \left(F^{0\prime}F^0/T\right)^{-1} T^{-1} \sum_{r=1}^{T} F_r^0 \left[X_{ir} - E_{\mathcal{D}}\left(X_{ir}\right)\right] \right\| \\
&\le \left\| \left(F^{0\prime}F^0/T\right)^{-1} \right\| \max_{1\le i\le N} \left\| \Omega_i^{-1/2} \right\| \max_{1\le i\le N} \left\| T^{-1} \sum_{r=1}^{T} F_r^0 \left[X_{ir} - E_{\mathcal{D}}\left(X_{ir}\right)\right] \right\| \\
&\quad \times \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} \varepsilon_{is} F_s^{0\prime} \right\| \\
&= O_P\left(1\right) O_P\left(1\right) O_P\left(a_{NT}\right) O_P\left(N^{1/8}\right) = o_P\left(1\right),
\end{aligned}
$$

as we can show that $\max_{1\leq i\leq N}\left\|T^{-1}\sum_{r=1}^{T}F_{r}^{0}\left[X_{ir}-E_{\mathcal{D}}\left(X_{ir}\right)\right]\right\|=O_{P}\left(a_{NT}\right)$ by following the arguments used in the proof of Lemma A.7(iii). It follows that $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{1it}\right\|=o_{P}\left(1\right)$. Similarly, using the expression for $\hat{\varepsilon}_{i}$ in (B.4) and (D.11), we can show that $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{2it}\right\|=o_{P}\left(1\right)$ and $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{3it}\right\|=o_{P}\left(1\right)$ after tedious calculations. Hence $\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{it}\right\|=o_{P}\left(1\right)$. It follows that

$$\begin{aligned}V_{NT,112} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\varepsilon_{it}\bar{b}_{it}^{\prime}\sum_{s=1}^{t-1}\left(\hat{b}_{is}\hat{\varepsilon}_{is}-\bar{b}_{is}\varepsilon_{is}\right)\right]^{2}\\ &\leq \left\{\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{it}\right\|^{2}\right\}\left\{T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\|\varepsilon_{it}\bar{b}_{it}^{\prime}\right\|^{2}\right\}=o_{P}\left(1\right)O_{P}\left(1\right)=o_{P}\left(1\right),\end{aligned}$$

and

$$\begin{aligned}V_{NT,113} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\left(\hat{\varepsilon}_{it}\hat{b}_{it}-\varepsilon_{it}\bar{b}_{it}\right)^{\prime}\sum_{s=1}^{t-1}\left(\hat{b}_{is}\hat{\varepsilon}_{is}-\bar{b}_{is}\varepsilon_{is}\right)\right]^{2}\\ &\leq \left\{\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left\|S_{it}\right\|^{2}\right\}\left\{T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\|\hat{\varepsilon}_{it}\hat{b}_{it}-\varepsilon_{it}\bar{b}_{it}\right\|^{2}\right\}=o_{P}\left(1\right)o_{P}\left(1\right)=o_{P}\left(1\right)\end{aligned}$$

as one can readily show that $T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\|\hat{\varepsilon}_{it}\hat{b}_{it}-\varepsilon_{it}\bar{b}_{it}\right\|^{2}=o_{P}\left(1\right)$. Thus $V_{NT,11}=o_{P}\left(1\right)$. This completes the proof of (ii). ■

REFERENCES

Hall, P. & C. C. Heyde (1980) *Martingale Limit Theory and Its Applications.* Academic Press, New York.

Moon, H. & M. Weidner (2010a) Linear regression for panel with unknown number of factors as interactive fixed effects. *Working paper*, University of Southern California.

Su, L. & S. Jin (2012) Sieve estimation of panel data models with cross section dependence. *Journal of Econometrics* 169, 34-47.

Sun, S. & C-Y. Chiang (1997) Limiting behavior of the perturbed empirical distribution functions evaluated at U-statistics for strongly mixing sequences of random variables. *Journal of Applied Mathematics and Stochastic Analysis* 10, 3-20.