Supplement to "Nonparametric Time–Varying Panel Data Models with Heterogeneity"

Fei Liu
Nankai University, China

This supplement contains three appendices: Appendix B provides some additional discussions on the estimation method, Appendix C presents the proofs of the main theoretical results, and the proofs of the main lemmas are provided in Appendix D.

Appendix B Extra models and estimation methods

B.1 Time–constant panel data model

This appendix studies the panel data model with heterogeneous time-constant regression coefficients:

$$y_{it} = x_{it}^{\top}\beta_i^0 + \lambda_i^{0\top} f_t^0 + \varepsilon_{it}. \tag{B.1}$$

The DLS iterative algorithm can be used to estimate model (B.1) after some necessary modifications:

Step 1. Find initial estimators $\widetilde{B}^{(0)} = (\widetilde{\beta}_1^{(0)}, \cdots, \widetilde{\beta}_N^{(0)})^{\top}$ via minimizing the following nuclear-norm regularization objective function:

$$\left(\widetilde{B}^{(0)}, \widetilde{\Gamma}\right) = \arg \min_{B\in\mathbb{R}^{N\times p}, \Gamma\in\mathbb{R}^{N\times T}} \left\{ \widetilde{Q}(B,\Gamma) + \frac{\phi_{NT}}{\sqrt{NT}} \|\Gamma\|_* \right\},$$

where $B = (\beta_1, \cdots, \beta_N)^{\top}$ and

$$\widetilde{Q}(B,\Gamma) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - x_{it}^{\top}\beta_i - \gamma_{it}\right)^2.$$

With $\widetilde{B}^{(0)}$, we can compute the residuals $\widetilde{R}_{it} = y_{it} - x_{it}^{\top}\widetilde{\beta}_i^{(0)}$ and construct the initial estimators $\widetilde{F}^{(0)}$ and $\widetilde{\Lambda}^{(0)}$ through the principal component analysis of $\widetilde{R}_{it}$'s covariance matrix.

Step 2. With $\widetilde{F}^{(n-1)} = (\widetilde{f}_1^{(n-1)}, \cdots, \widetilde{f}_T^{(n-1)})^{\top}$, β_i^0 and λ_i^0 can be estimated jointly using the least squares method. $\widetilde{\beta}_i^{(n)}$ and $\widetilde{\lambda}_i^{(n)}$ denote the estimators of β_i and λ_i^0, respectively.

Step 3. With $\widetilde{B}^{(n)} = (\widetilde{\beta}_1^{(n)}, \cdots, \widetilde{\beta}_N^{(n)})^{\top}$ and $\widetilde{\Lambda}^{(n)} = (\widetilde{\lambda}_1^{(n)}, \cdots, \widetilde{\lambda}_N^{(n)})^{\top}$, $\widetilde{f}_t^{(n-1)}$ can be updated using the least squares method.

Step 4. Repeat Steps 2–3 until numerical convergence.

B.2 Estimation by sieve method

This appendix offers an alternative method to estimate model (2.1), which is based on sieve approximations. Let $\{\varphi_1(\tau), \varphi_2(\tau), \cdots\}$ be a set of basis functions (e.g., B-spline functions). By sieve method, we have

$$\beta_i(\tau) = \sum_{k=1}^{\infty} c_{ik}\varphi_k(\tau) = \sum_{k=1}^{K_0} c_{ik}\varphi_k(\tau) + \Delta_{K_0}(\tau),$$

where $\Delta_{K_0}(\tau) = \sum_{k=K_0+1}^{\infty} c_{ik}\varphi_k(\tau)$, c_{ik} is a $p\times 1$ vector of unknown sieve coefficients, and K_0 is the truncation parameter. Denote $\widetilde{X}_{it} = \left(\varphi_1(\tau_t)x_{it}^{\top}, \cdots, \varphi_{K_0}(\tau_t)x_{it}^{\top}\right)^{\top}$, $\mathcal{C}_i = \left(c_{i1}^{\top}, \cdots, c_{iK_0}^{\top}\right)^{\top}$ and

Table B.1: AMSEs and SCCs for sieve estimates

	$\text{AMSE}_{\beta,0}$			$\text{SCC}_{\lambda,0}$			$\text{SCC}_{f,0}$		
N/T	40	80	120	40	80	120	40	80	120
40	0.6715	0.2933	0.2033	0.4304	0.7706	0.8502	0.4337	0.7515	0.8319
80	0.6652	0.2832	0.1923	0.6405	0.8359	0.8753	0.7136	0.8881	0.9235
120	0.6594	0.2745	0.1811	0.6921	0.8512	0.8814	0.7952	0.9278	0.9428
	$\text{AMSE}_{\beta,n}$			$\text{SCC}_{\lambda,n}$			$\text{SCC}_{f,n}$		
N/T	40	80	120	40	80	120	40	80	120
40	0.5253	0.1420	0.0753	0.7569	0.9705	0.9838	0.7974	0.9682	0.9760
80	0.4388	0.1346	0.0730	0.9267	0.9786	0.9867	0.9649	0.9875	0.9890
120	0.4265	0.1318	0.0701	0.9440	0.9799	0.9886	0.9842	0.9918	0.9984

$\widetilde{\Delta}_{K_0,it} = x_{it}^\top \Delta_{K_0}(\tau_t)$. Model (2.1) can be rewritten as: $y_{it} = \widetilde{X}_{it}^\top \mathcal{C}_i + \lambda_i^{0\top} f_t^0 + \widetilde{\Delta}_{K_0,it} + \varepsilon_{it}$. For this model, we can use the following iteration algorithm to construct estimators for $\mathcal{C}_i$, f_t^0 and λ_i^0:

Step 1. Obtain the initial estimators $\widehat{C}^{(0)} = (\widehat{\mathcal{C}}_1^{(0)}, \cdots, \widehat{\mathcal{C}}_N^{(0)})^\top$:

$$\left(\widehat{C}^{(0)}, \widehat{\Gamma}^{(0)}\right) = \arg \min_{C \in \mathbb{R}^{N\times pK_0}, \Gamma \in \mathbb{R}^{N\times T}} \left\{ \widetilde{Q}^*(C,\Gamma) + \frac{\phi_{NT}}{\sqrt{NT}} \|\Gamma\|_* \right\},$$

where $\widetilde{Q}^*(C,\Gamma) = \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \left(y_{it} - \widetilde{X}_{it}^\top \mathcal{C}_i - \gamma_{it}\right)^2$. Then, we compute regression residuals and obtain the initial estimators $\widehat{\Lambda}^{(0)}$ and $\widehat{F}^{(0)}$ using the PCA method.

Step 2. With $\widehat{F}^{(n-1)}$ and $\widehat{\Lambda}^{(n-1)}$, update $\widehat{C}^{(n-1)}$ by minimizing $\widetilde{Q}^*(C, \widehat{\Gamma}^{(n-1)})$, where $\widehat{\Gamma}^{(n-1)} = \widehat{\Lambda}^{(n-1)} \widehat{F}^{(n-1)\top}$.

Step 3. With $\widehat{C}^{(n)}$, we estimate F^0 and Λ^0 by the PCA method.

Step 4. Repeat Steps 2–3 until numerical convergence.

With $\widehat{C}^{(n)}$, we construct the sieve estimator: $\widehat{\beta}_i^{(n)}(\tau) = \sum_{k=1}^{K_0} \widehat{c}_{ik}^{(n)} \varphi_k(\tau)$. We re-conduct the simulation study of Example 5.1 in Section 5 to assess the sieve estimators' finite sample performance. We follow Su et al. (2019) to use B-spline polynomials as the basis functions and employ their method to choose interior knots for estimation. AMSEs and SCCs are computed after replicating the experiments for 1000 times. Simulation results are reported in Table B.1.

Appendix C Proofs of the main results

Denote $K_{t,m}(\tau) = K\left(\frac{t-\tau T}{Th}\right)\left(\frac{t-\tau T}{Th}\right)^m$, for $m = 0,1,2,3$. Let $\widetilde{F} = F^0 H$, $\widetilde{\Lambda} = \Lambda^0 H^{-1\top}$, $\widetilde{f}_t = H^\top f_t^0$, $\widetilde{\lambda}_i = H^{-1}\lambda_i^0$, $R_{f,t}^{(n)} = \widehat{f}_t^{(n)} - \widetilde{f}_t$, and $R_f^{(n)} = \widehat{F}^{(n)} - \widetilde{F}$, where the rotation matrix H is defined in Theorem 3.1. Let $\delta_{f,n}$ be $R_{f,t}^{(n)}$'s rate of convergence and $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}^{-1}$. C denotes a finite positive constant and its value can be different at each appearance.

C.1 Proofs of the main theorems

C.1.1 Proof of Theorem 3.1

(1) Let $\Gamma = \Lambda F^\top$, $\Gamma^0 = \Lambda^0 F^{0\top}$, $\gamma_{it} = \lambda_i^\top f_t$ and $\gamma_{it}^0 = \lambda_i^{0\top} f_t^0$. The initial estimator $\widehat{B}^{(0)}(\tau)$ satisfies

$$\left(\widehat{B}^{(0)}(\tau), \widehat{B}'^{(0)}(\tau), \widehat{\Gamma}_\tau\right) = \arg \min_{A,C \in \mathbb{R}^{N\times p}, \Gamma \in \mathbb{R}^{N\times T}} \left\{ Q_\tau(A,C,\Gamma) + \frac{\phi_{NT}}{\sqrt{NT}} \|\Gamma\|_* \right\},$$

where $Q_\tau(A,C,\Gamma) = (NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T \left(y_{it} - x_{it}^\top(a_i + (\tau_t-\tau)c_i) - \gamma_{it}\right)^2 K\left(\frac{t-\tau T}{Th}\right)$, and a_i and c_i are the i-th columns of $A^\top$ and $C^\top$, respectively. For any given τ, define $\widetilde{x}_{it,\tau} = (x_{it}^\top, x_{it}^\top(\tau_t-\tau))^\top$, $d_{i,\tau}^0 = (\beta_i^\top(\tau), \beta_i'^\top(\tau))^\top$ and $d_i = (a_i^\top, c_i^\top)^\top$. For any $A, C \in \mathbb{R}^{N\times p}$ and $\Gamma \in \mathbb{R}^{N\times T}$,

$$
\begin{aligned}
&\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(y_{it} - x_{it}^\top(a_i + (\tau_t-\tau)c_i) - \gamma_{it}\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) \\
&= \frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(\varepsilon_{it} - x_{it}^\top(a_i + (\tau_t-\tau)c_i - \beta_i(\tau_t)) - (\gamma_{it}-\gamma_{it}^0)\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) \\
&= \frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(\varepsilon_{it} - \widetilde{x}_{it,\tau}^\top(d_i - d_{i,\tau}^0) - (\gamma_{it}-\gamma_{it}^0)\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) + O_P(h^4) \\
&= \frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}^2 K\left(\frac{\tau_t-\tau}{h}\right) + \frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{x}_{it,\tau}^\top(d_i - d_{i,\tau}^0) + (\gamma_{it}-\gamma_{it}^0)\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) \\
&\quad - \frac{2}{NTh}\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}\left(\widetilde{x}_{it,\tau}^\top(d_i - d_{i,\tau}^0) + (\gamma_{it}-\gamma_{it}^0)\right) K\left(\frac{\tau_t-\tau}{h}\right) + O_P(h^4), \qquad \text{(C.1)}
\end{aligned}
$$

where the second equality holds by the local linear approximation and its proof is analogous to that of Lemma C.1(1). Let $\widehat{d}_{i,\tau} = (\widehat{\beta}_i^{(0)\top}(\tau), \widehat{\beta}_i'^{(0)\top}(\tau))^\top$. Since $Q_\tau(\widehat{B}^{(0)}(\tau), \widehat{B}'^{(0)}(\tau), \widehat{\Gamma}_\tau) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\widehat{\Gamma}_\tau\right\|_* \le Q_\tau(B(\tau), B'(\tau), \Gamma^0) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\Gamma^0\right\|_*$, (C.1) implies

$$
\begin{aligned}
&\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau} - d_{i,\tau}^0) + (\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\widehat{\Gamma}_\tau\right\|_* \\
&\le 2\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}\left(\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau} - d_{i,\tau}^0) + (\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)\right) K\left(\frac{\tau_t-\tau}{h}\right) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\Gamma^0\right\|_* + O_P(h^4).
\end{aligned}
$$

Without loss of generality, we assume that Γ^0 has the following SVD decomposition: $\Gamma^0 = U\Omega_\gamma V^\top$, where $U_{N\times\min\{N,T\}}$ and $V_{T\times\min\{N,T\}}$ consist of the singular vectors and Ω_γ is a diagonal matrix containing the singular values of Γ^0. Since $\operatorname{rank}(\Gamma^0) = r_0$, the first r_0 singular values of Γ^0 are nonzero and the remaining singular values are zero. Additionally, let $U = (U_1, U_2)$ and $V = (V_1, V_2)$, where (U_1, V_1) and (U_2, V_2) contain the singular vectors corresponding to nonzero and zero singular values, respectively. Define $\mathcal{P}_{\phi,\tau} = U_1U_1^\top\left(\widehat{\Gamma}_\tau - \Gamma^0\right)V_1V_1^\top$ and $\mathcal{M}_{\phi,\tau} = \widehat{\Gamma}_\tau - \Gamma^0 - \mathcal{P}_{\phi,\tau}$. Then,

$$
\left\|\widehat{\Gamma}_\tau\right\|_* = \left\|\Gamma^0 + \mathcal{P}_{\phi,\tau} + \mathcal{M}_{\phi,\tau}\right\|_* \ge \left\|\Gamma^0\right\|_* + \left\|\mathcal{P}_{\phi,\tau}\right\|_* - \left\|\mathcal{M}_{\phi,\tau}\right\|_*,
$$

where the inequality holds by Lemma C.2 of Chernozhukov et al. (2018). Therefore, it follows that

$$
\begin{aligned}
&\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau} - d_{i,\tau}^0) + (\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)\right)^2 K\left(\frac{\tau_t-\tau}{h}\right) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\mathcal{P}_{\phi,\tau}\right\|_* \\
&\le 2\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}\left(\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau} - d_{i,\tau}^0) + (\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)\right) K\left(\frac{\tau_t-\tau}{h}\right) + \frac{\phi_{NT}}{\sqrt{NT}}\left\|\mathcal{M}_{\phi,\tau}\right\|_* + O_P(h^4),
\end{aligned}
$$

where $\widehat{\gamma}_{it,\tau}$ is the (i,t)-th element of $\widehat{\Gamma}_\tau$. Let $D_\tau = (B(\tau), B'(\tau))$ and $\widehat{D}_\tau = (\widehat{B}^{(0)}(\tau), \widehat{B}'^{(0)}(\tau))$. By Cauchy-Schwarz inequality,

$$
\frac{1}{NTh}\left|\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau} - d_{i,\tau}^0) K\left(\frac{\tau_t-\tau}{h}\right)\right| \le \frac{1}{NTh}\left(\sum_{i=1}^N\left\|\sum_{t=1}^T \varepsilon_{it}\widetilde{x}_{it} K\left(\frac{\tau_t-\tau}{h}\right)\right\|^2\right)^{\frac{1}{2}}\left\|\widehat{D}_\tau - D_\tau\right\|_F
$$

$$= O_P\left(\frac{1}{\sqrt{NTh}}\right)\cdot\left\|\widehat{D}_\tau - D_\tau\right\|_F, \tag{C.2}$$

where the last equality holds by the fact $\sum_{i=1}^N \left\|\sum_{t=1}^T \varepsilon_{it} x_{it} K\left(\frac{\tau_t-\tau}{h}\right)\right\|^2 = O_P(NTh)$. It can be proved using Lemma D.3. For $\widehat{\Gamma}_\tau$, by Holder inequality,

$$\begin{aligned}
\frac{1}{NTh}\left|\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}(\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)K\left(\frac{\tau_t-\tau}{h}\right)\right| &= \frac{1}{NTh}\left|\mathrm{tr}\left(\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}(\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)K\left(\frac{\tau_t-\tau}{h}\right)\right)\right| \\
&\le C\frac{1}{NTh}\left\|\mathcal{E}\right\|_\infty\left\|\widehat{\Gamma}_\tau-\Gamma^0\right\|_* \\
&= O_P\left(\frac{\max\{\sqrt{N},\sqrt{T}\}}{NTh}\right)\left(\|\mathcal{P}_{\phi,\tau}\|_* + \|\mathcal{M}_{\phi,\tau}\|_*\right). \qquad \text{(C.3)}
\end{aligned}$$

Moreover, by Lemma C.2 of Chernozhukov et al. (2018),

$$\|\mathcal{M}_{\phi,\tau}\|_*^2 \le \|\mathcal{M}_{\phi,\tau}\|_F^2\,\mathrm{rank}(\mathcal{M}_{\phi,\tau}) \le \left\|\widehat{\Gamma}_\tau-\Gamma^0\right\|_F^2\mathrm{rank}(\mathcal{M}_{\phi,\tau}). \tag{C.4}$$

If $\min\{\sqrt{N},\sqrt{T}\}h\phi_{NT}\to\infty$, it follows that

$$\left(\frac{\phi_{NT}}{\sqrt{NT}}-\frac{\max\{\sqrt{N},\sqrt{T}\}}{NTh}\right)\|\mathcal{P}_{\phi,\tau}\|_*\ge 0, \tag{C.5}$$

with probability approaching 1. Let $\Delta_{D,\tau}=\frac{1}{\sqrt{N}}\|\widehat{D}_\tau-D_\tau\|_F$ and $\Delta_{\Gamma,\tau}=\frac{1}{\sqrt{NT}}\|\widehat{\Gamma}_\tau-\Gamma^0\|_F$. Combining (C.2), (C.3), (C.4), (C.5), and Assumption 1, we obtain

$$\begin{aligned}
&\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T\left(\widetilde{x}_{it,\tau}^\top(\widehat{d}_{i,\tau}-d_{i,\tau}^0)+(\widehat{\gamma}_{it,\tau}-\gamma_{it}^0)\right)^2K\left(\frac{\tau_t-\tau}{h}\right)\\
&\le O_P\left(\frac{1}{\sqrt{Th}}\right)\Delta_{D,\tau}+O_P\left(\frac{1}{\min\{\sqrt{N},\sqrt{T}\}h}\right)\Delta_{\Gamma,\tau}+C\phi_{NT}\Delta_{\Gamma,\tau}+O_P(h^4)\\
&\le O_P\left(\max\left\{\frac{1}{\min\{\sqrt{N},\sqrt{T}\}\sqrt{h}},\phi_{NT}\sqrt{h}\right\}\right)\left(\Delta_{D,\tau}+\frac{\Delta_{\Gamma,\tau}}{\sqrt{h}}\right)+O_P(h^4). \qquad \text{(C.6)}
\end{aligned}$$

By (C.6) and the restricted strong convexity condition in Assumption 1, we can use arguments that are closely related to those in the proof of Theorem 2 of Moon and Weidner (2018) and show that $(\widehat{\Gamma}_\tau-\Gamma^0)W(\tau)^{1/2}\in\mathcal{C}(c)$ with $c=1$, $\Delta_{D,\tau}=O_P\left(\frac{\Delta_{\Gamma,\tau}}{\sqrt{h}}\right)$, and

$$\left(\Delta_{D,\tau}+\frac{\Delta_{\Gamma,\tau}}{\sqrt{h}}\right)^2\le O_P\left(\max\left\{\frac{1}{\min\{\sqrt{N},\sqrt{T}\}\sqrt{h}},\phi_{NT}\sqrt{h}\right\}\right)\left(\Delta_{D,\tau}+\frac{\Delta_{\Gamma,\tau}}{\sqrt{h}}\right)+O_P(h^4),$$

which immediately yields $\Delta_{D,\tau}=O_P\left(\max\left\{h^{3/2},\phi_{NT}\right\}h^{1/2}\right)$ and $\Delta_{\Gamma,\tau}=O_P\left(\max\left\{h^{3/2},\phi_{NT}\right\}h\right)$ under the condition $\min\{\sqrt{N},\sqrt{T}\}h\phi_{NT}\to\infty$. Therefore, we complete the proof of Theorem 3.1(1).
(2) We write

$$\begin{aligned}
\widehat{F}^{(0)}V_{NT}-\frac{1}{NT}\sum_{i=1}^N F^0\lambda_i^0\lambda_i^{0\top}F^{0\top}\widehat{F}^{(0)} &= \frac{1}{NT}\sum_{i=1}^N\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}\lambda_i^{0\top}F^{0\top}\widehat{F}^{(0)}\\
&+\frac{1}{NT}\sum_{i=1}^N F^0\lambda_i^0\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}^\top\widehat{F}^{(0)}+\frac{1}{NT}\sum_{i=1}^N\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}\widetilde{\varepsilon}_i^\top\widehat{F}^{(0)}
\end{aligned}$$

$$+\frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}^\top \widehat{F}^{(0)}$$
$$+\frac{1}{NT}\sum_{i=1}^{N}\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}\begin{pmatrix} x_{i1}^\top\left(\beta_i(\tau_1)-\widehat{\beta}_i^{(0)}(\tau_1)\right)\\ \vdots \\ x_{iT}^\top\left(\beta_i(\tau_T)-\widehat{\beta}_i^{(0)}(\tau_T)\right)\end{pmatrix}^\top \widehat{F}^{(0)}$$
$$+\frac{1}{NT}\sum_{i=1}^{N}F^0\lambda_i^0\widetilde{\varepsilon}_i^\top\widehat{F}^{(0)}+\frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i\lambda_i^{0\top}F^{0\top}\widehat{F}^{(0)}+\frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i\widetilde{\varepsilon}_i^\top\widehat{F}^{(0)}$$
$$:= J_{NT,1}+\cdots+J_{NT,8}, \tag{C.7}$$

where $\widetilde{\varepsilon}_i=(\varepsilon_{i1},\ldots,\varepsilon_{iT})^\top$. Using the arguments that are analogous to those in the proof of Proposition A.1 of Bai (2009), we obtain

$$\frac{1}{\sqrt{T}}\left\|J_{NT,6}\right\|_F=O_P\left(\delta_{NT}\right),\quad \frac{1}{\sqrt{T}}\left\|J_{NT,7}\right\|_F=O_P\left(\delta_{NT}\right),\quad \frac{1}{\sqrt{T}}\left\|J_{NT,8}\right\|_F=O_P\left(\delta_{NT}\right), \tag{C.8}$$

where $\delta_{NT}=\min\{\sqrt{N},\sqrt{T}\}^{-1}$. We proceed with the computation of $J_{NT,1},\ldots,J_{NT,5}$'s convergence rates. For $J_{NT,1}$, by Cauchy-Schwarz inequality,

$$\frac{1}{N}\left\|\sum_{i=1}^{N}\lambda_i^0x_{it}^\top\left(\beta_i(\tau_t)-\widehat{\beta}_i^{(0)}(\tau_t)\right)\right\|\le\frac{1}{N}\left(\sum_{i=1}^{N}\|\lambda_i^0x_{it}^\top\|_F^2\right)^{\frac{1}{2}}\left\|\widehat{B}^{(0)}(\tau_t)-B(\tau_t)\right\|_F=O_P\left(h^{1/2}\max\left\{h^{3/2},\phi_{NT}\right\}\right),$$

where the equality holds by Theorem 3.1(1). It implies

$$\frac{1}{\sqrt{T}}\left\|J_{NT,1}\right\|_F=O_P\left(h^{1/2}\max\left\{h^{3/2},\phi_{NT}\right\}\right). \tag{C.9}$$

Analogously, we have $\frac{1}{\sqrt{T}}\|J_{NT,2}\|_F=O_P(h^{1/2}\max\{h^{3/2},\phi_{NT}\})$, $\frac{1}{\sqrt{T}}\|J_{NT,3}\|_F=O_P(h^{1/2}\max\{h^{3/2},\phi_{NT}\})$, $\frac{1}{\sqrt{T}}\|J_{NT,4}\|_F=O_P(h^{1/2}\max\{h^{3/2},\phi_{NT}\})$ and $\frac{1}{\sqrt{T}}\|J_{NT,5}\|_F=O_P(h\max\{h^3,\phi_{NT}^2\})$. Together with (C.7), (C.8) and (C.9), these results yield

$$\frac{1}{\sqrt{T}}\left\|\widehat{F}^{(0)}V_{NT}-\frac{1}{NT}\sum_{i=1}^{N}F^0\lambda_i^0\lambda_i^{0\top}F^{0\top}\widehat{F}^{(0)}\right\|_F=O_P\left(h^{1/2}\max\left\{h^{3/2},\phi_{NT}\right\}\right). \tag{C.10}$$

Considering that V_{NT} is invertible, we complete the proof of Theorem 3.1(2).
(3) For the loading initial estimator, write

$$\begin{aligned}\widehat{\Lambda}^{(0)}-\Lambda^0H^{-1\top} &= \frac{1}{T}R\widehat{F}^{(0)}-\Lambda^0H^{-1\top}\\ &= \frac{1}{T}\Lambda^0F^{0\top}\widehat{F}^{(0)}+\frac{1}{T}\mathcal{E}\widehat{F}^{(0)}-\frac{1}{T}\widetilde{\Delta}_B\widehat{F}^{(0)}-\Lambda^0H^{-1\top}\\ &= \frac{1}{T}\Lambda^0\left(F^0-\widehat{F}^{(0)}H^{-1}\right)^\top\widehat{F}^{(0)}+\frac{1}{T}\mathcal{E}\left(\widehat{F}^{(0)}-\widetilde{F}\right)-\frac{1}{T}\widetilde{\Delta}_B\widehat{F}^{(0)}+\frac{1}{T}\mathcal{E}F^0H\\ &:= J_{NT,9}+\cdots+J_{NT,12},\end{aligned}$$

where $\widetilde{\Delta}_B$ is an $N\times T$ matrix with its (i,t)-th element as $x_{it}^\top(\widehat{\beta}_i(\tau_t)-\beta_i(\tau_t))$. We now study these terms one by one. For $J_{NT,9}$, it holds by Theorem 3.1(2) that

$$\frac{1}{\sqrt{N}}\left\|J_{NT,9}\right\|_F\le\left(\frac{1}{\sqrt{N}}\left\|\widetilde{\Lambda}\right\|_F\right)\left(\frac{1}{\sqrt{T}}\left\|\widehat{F}^{(0)}-\widetilde{F}\right\|_F\right)\left(\frac{1}{\sqrt{T}}\left\|\widehat{F}^{(0)}\right\|_F\right)=O_P\left(\max\left\{h^{3/2},\phi_{NT}\right\}h^{1/2}\right).$$

Analogously, we have

$$\frac{1}{\sqrt{N}}\left\|J_{NT,10}\right\|_F \leq \left(\frac{1}{\sqrt{NT}}\left\|\mathcal{E}\right\|_F\right)\left(\frac{1}{\sqrt{T}}\left\|\widehat{F}^{(0)}-\widetilde{F}\right\|_F\right) = O_P\left(\max\left\{h^{3/2},\phi_{NT}\right\}h^{1/2}\right).$$

By Theorem 3.1(1) and Cauchy-Schwarz inequality, we obtain $\frac{1}{\sqrt{NT}}\|\widetilde{\Delta}_B\|_F = O_P(\max\{h^{3/2},\phi_{NT}\}h^{1/2})$, which yields

$$\frac{1}{\sqrt{N}}\left\|J_{NT,11}\right\|_F = \left(\frac{1}{\sqrt{NT}}\left\|\widetilde{\Delta}_B\right\|_F\right)\left(\frac{1}{\sqrt{T}}\left\|\widehat{F}^{(0)}\right\|_F\right) = O_P\left(\max\left\{h^{3/2},\phi_{NT}\right\}h^{1/2}\right).$$

For $J_{NT,12}$, note that $\mathcal{E}F^0 = \sum_{t=1}^T \varepsilon_t f_t^{0\top}$. Since $\varepsilon_t f_t^{0\top}$ is conditionally α-mixing, Lemma D.2 implies $\frac{1}{\sqrt{N}}\left\|J_{NT,12}\right\|_F = O_P\left(\frac{1}{\sqrt{T}}\right) = o_P\left(h^{1/2}\max\left\{h^{3/2},\phi_{NT}\right\}\right)$. Therefore, Theorem 3.1(3) holds. ■

C.1.2 Proof of Theorem 3.2

We propose two propositions that can jointly lead to the desired results in Theorem 3.2. Define

$$\begin{aligned}
\Omega_1^{\mathcal{D}}(t,s) &= (\Lambda^{0\top}\Lambda^0)^{-1}\sum_{i=1}^N \lambda_i^0\lambda_i^{0\top}\mathbb{E}_{\mathcal{D}}\left(x_{it}^\top\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}\right),\\
\Omega_2^{\mathcal{D}}(t,s) &= (\Lambda^{0\top}\Lambda^0)^{-1}\sum_{i=1}^N \lambda_i^0\lambda_i^{0\top}\mathbb{E}_{\mathcal{D}}\left(z_{it}^\top\Omega_{f,i}^{\mathcal{D}-1}z_{is}\right),\\
b_{f,t}^{\dagger(n)} &= (Th)^{-n}H^\top\sum_{s_1,s_2,\ldots,s_n=1}^T K_{s_1,0}(\tau_t)\Omega_1^{\mathcal{D}}(t,s_1)\prod_{j=1}^{n-1}(K_{s_{j+1},0}(\tau_{s_j})\Omega_1^{\mathcal{D}}(s_j,s_{j+1}))H^{-1\top}R_{f,s_n}^{(0)}\\
&\quad +T^{-n}H^\top\sum_{s_1,s_2,\ldots,s_n=1}^T \Omega_2^{\mathcal{D}}(t,s_1)\prod_{j=1}^{n-1}\Omega_2^{\mathcal{D}}(s_j,s_{j+1}))H^{-1\top}R_{f,s_n}^{(0)}.
\end{aligned}$$

Proposition C.1. *Let Assumptions 1-2 hold. For $n\geq 1$, as $N,T\to\infty$ simultaneously,*

(1) $N^{-1/2}\left\|\widehat{\Lambda}^{(n)}-\widetilde{\Lambda}\right\|_F = O_P\left(\max\{\delta_{f,n-1},\delta_{NT}\}\right)$;

(2) $T^{-1/2}\left\|\widehat{F}^{(n)}-\widetilde{F}\right\|_F = O_P\left(\max\{\delta_{f,n-1},\delta_{NT}\}\right)$;

(3) $(Th)^{-1}\left\|\left(\widehat{F}^{(n)}-\widetilde{F}\right)^\top W(\tau)\left(\widehat{F}^{(n)}-\widetilde{F}\right)\right\|_F = O_P\left(\max\{\delta_{f,n-1},\delta_{NT}\}^2\right)$.

Proposition C.2. *Let Assumptions 1-2 hold. For $n\geq 1$, as $N,T\to\infty$ simultaneously,*

(1) for any given t,

$$\begin{aligned}
\sqrt{N}\left(\widehat{f}_t^{(n)}-\widetilde{f}_t\right) &= \sqrt{N}H^\top\left(\Lambda^{0\top}\Lambda^0\right)^{-1}\left(\sum_{i=1}^N\lambda_i^0\varepsilon_{it}\right)+O_P\left(\sqrt{N}\|b_{f,t}^{\dagger(n)}\|\right)\\
&\quad +O_P\left(\sqrt{\frac{N}{T^2}}\right)+O_P(h^2)+o_P(1),
\end{aligned}$$

(2) for any given i,

$$\begin{aligned}
\sqrt{T}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right) &= \frac{1}{\sqrt{T}}H^{-1}\Omega_{f,i}^{\mathcal{D}-1}\sum_{t=1}^T z_{it}\varepsilon_{it}+O_P\left(\left(\sum_{t=1}^T\|b_{f,t}^{\dagger(n-1)}\|^2\right)^{1/2}\right)\\
&\quad +O_P\left(\sqrt{\frac{T}{N^2}}\right)+O_P(h^2)+o_P(1),
\end{aligned}$$

(3) for any given i and τ,

$$\begin{aligned}\sqrt{Th}\left(\widehat{\beta}_i^{(n)}(\tau)-\beta_i(\tau)\right) &= \frac{1}{\sqrt{Th}}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\sum_{t=1}^{T}K_{t,0}(\tau)x_{it}\varepsilon_{it}+O_P\left(\sqrt{h}\left(\sum_{t=1}^{T}\|b_{f,t}^{\dagger(n-1)}\|^2\right)^{1/2}\right)\\ &\quad +O_P\left(\sqrt{Th^5}\right)+o_P(1).\end{aligned}$$

Proof of Proposition C.1

(1) Recall that $\Omega_{S,i}=(I-S_i)^\top(I-S_i)$ and $\widehat{\Omega}_{f,i}^{(n)}=T^{-1}\widehat{F}^{(n)\top}(I-S_i)^\top(I-S_i)\widehat{F}^{(n)}$, where S_i is defined in Section 2.2. For $\widehat{\lambda}_i^{(n)}$, we have

$$\begin{aligned}\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i &= \frac{1}{T}\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top-\frac{1}{T}\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i\\ &\quad +\frac{1}{T}\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i\\ &:= \Phi_{NT,1,i}+\Phi_{NT,2,i}+\Phi_{NT,3,i}.\end{aligned}\tag{C.11}$$

Then,

$$\sum_{i=1}^{N}\left\|\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right\|^2=\sum_{i=1}^{N}\|\Phi_{NT,1,i}\|^2+\sum_{i=1}^{N}\|\Phi_{NT,2,i}\|^2+\sum_{i=1}^{N}\|\Phi_{NT,3,i}\|^2+\text{interaction terms}.\tag{C.12}$$

It suffices only to show the first three terms on the right-hand side of (C.12) can at most have the probability order $O_P(N\max\left\{\delta_{f,n-1}^2,T^{-1}\right\})$. Then the convergence of the remaining terms can be proved using the Cauchy-Schwarz inequality. By Lemma C.5(1), we have

$$\sum_{i=1}^{N}\|\Phi_{NT,1,i}\|^2=O_P\left(\frac{Nh^3}{T}\max\left\{1,Th\delta_{f,n-1}^2\right\}\right).\tag{C.13}$$

For $\Phi_{NT,2,i}$,

$$\begin{aligned}\sum_{i=1}^{N}\|\Phi_{NT,2,i}\|^2 &= \frac{1}{T^2}\sum_{i=1}^{N}\left\|\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i\right\|^2+\sum_{i=1}^{N}\left\|\Phi_{NT,2,i}-\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i\right\|^2\\ &\quad +\text{interaction term}\\ &= O_P\left(N\delta_{f,n-1}^2\right),\end{aligned}\tag{C.14}$$

where $\Omega_{f,i}^H=H^\top\Omega_{f,i}^{\mathcal{D}}H$, $\widetilde{z}_{it}=H^\top z_{it}$, and the second equality holds by Lemma C.5(2) and the following result:

$$\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\widetilde{z}_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i\right\|^2 \le \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\widetilde{\lambda}_i\widetilde{z}_{it}^\top\right\|_F^2\right)\left(\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\right)=O_P\left(NT^2\delta_{f,n-1}^2\right).$$

For $\sum_{i=1}^{N}\|\Phi_{NT,3,i}\|^2$, Lemma D.2(1) implies $\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\widetilde{z}_{it}\varepsilon_{it}\right\|^2=O_P(NT)$ under the α-mixing condition in Assumption 1. By Cauchy-Schwarz inequality,

$$\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}R_{f,t}^{(n-1)}\varepsilon_{it}\right\|^2\le\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\|\varepsilon_{it}\|^2\right)\left(\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\right)=O_P\left(NT^2\delta_{f,n-1}^2\right).$$

Together with Lemma C.5(3), it yields

$$\sum_{i=1}^{N}\|\Phi_{NT,3,i}\|^2 = \frac{1}{T^2}\sum_{i=1}^{N}\left\|\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}\varepsilon_{it}\right\|^2+\frac{1}{T^2}\sum_{i=1}^{N}\left\|\Omega_{f,i}^{H-1}\sum_{t=1}^{T}R_{f,t}^{(n-1)}\varepsilon_{it}\right\|^2$$

$$+\sum_{i=1}^{N}\left\|\Phi_{NT,3,i}-\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\left(\widetilde{z}_{it}+R_{f,t}^{(n-1)}\right)\varepsilon_{it}\right\|^2+\text{interaction terms}$$
$$= O_P\left(\frac{N}{T}\max\left\{1,T\delta_{f,n-1}^2\right\}\right). \tag{C.15}$$

Combining (C.13), (C.14) and (C.15), we obtain

$$\frac{1}{\sqrt{N}}\left\|\widehat{\Lambda}^{(n)}-\widetilde{\Lambda}\right\|_F=O_P\left(\max\left\{\delta_{f,n-1},\frac{1}{\sqrt{T}}\right\}\right). \tag{C.16}$$

(2) We outline the main steps in the proof of Proposition C.1(2):

(2.a) we first show the convergence of $N^{-1}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\widehat{\lambda}_i^{(n)\top}$.

(2.b) we formulate the bias terms $\widehat{b}_{f,t}^{(n)}$ and $\widehat{b}_{\varepsilon,t}^{(n)}$ in $\widehat{f}_t^{(n)}-\widetilde{f}_t$ and compute the convergence rate of $\widehat{f}_t^{(n)}-\widetilde{f}_t-\widehat{b}_{f,t}^{(n)}-\widehat{b}_{\varepsilon,t}^{(n)}$, where

$$\widehat{b}_{f,t}^{(n)}=\frac{1}{N}\left(\widehat{\Lambda}^{(n)\top}\widehat{\Lambda}^{(n)}\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\widetilde{\lambda}_i-\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i\widetilde{z}_{it}^\top\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\right),$$
$$\widehat{b}_{\varepsilon,t}^{(n)}=\left(\frac{1}{N}\widehat{\Lambda}^{(n)\top}\widehat{\Lambda}^{(n)}\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i\varepsilon_{it}\right).$$

(2.c) we finally prove Proposition C.1(2) by computing the convergence rates of $\widehat{b}_{f,t}^{(n)}$ and $\widehat{b}_{\varepsilon,t}^{(n)}$.

Proof of (2.a): Write

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\widehat{\lambda}_i^{(n)\top}-H^{-1}\Sigma_\lambda H^{-1\top} &= H^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}-\Sigma_\lambda\right)H^{-1\top}+\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\widetilde{\lambda}_i^\top\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^\top+\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^\top\\ &:= \Phi_{NT,4}+\cdots+\Phi_{NT,7}.\end{aligned}$$

By Assumption 2, $\|\Phi_{NT,4}\|_F=O_P\left(N^{-1/2}\right)$. By Cauchy-Schwarz inequality and Proposition C.1(1),

$$\|\Phi_{NT,5}\|_F\leq\frac{1}{N}\left(\sum_{i=1}^{N}\left\|\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right\|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{N}\left\|\widetilde{\lambda}_i\right\|^2\right)^{\frac{1}{2}}=O_P\left(\max\left\{\delta_{f,n-1},\frac{1}{\sqrt{T}}\right\}\right).$$

Analogously, we have $\|\Phi_{NT,6}\|_F=O_P(\max\{\delta_{f,n-1},T^{-1/2}\})$ and $\|\Phi_{NT,7}\|_F=O_P(\max\{\delta_{f,n-1}^2,T^{-1/2}\})$. Let $\Sigma_\lambda^H=H^{-1}\Sigma_\lambda H^{-1\top}$. We obtain

$$\left\|N^{-1}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\widehat{\lambda}_i^{(n)\top}-\Sigma_\lambda^H\right\|_F=O_P(\max\{\delta_{f,n-1},\delta_{NT}\}). \tag{C.17}$$

Proof of (2.b): For $\widehat{f}_t^{(n)}$, write

$$\begin{aligned}&\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\widehat{\lambda}_i^{(n)\top}\right)\left(\widehat{f}_t^{(n)}-\widetilde{f}_t\right)\\ &=-\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}x_{it}^\top\left(\widehat{\beta}_i^{(n)}(\tau)-\beta_i(\tau)\right)-\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^\top\widetilde{f}_t+\frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i^{(n)}\varepsilon_{it}\\ &:=\Phi_{NT,8,t}+\cdots+\Phi_{NT,10,t}.\end{aligned} \tag{C.18}$$

We now study these terms one by one. For $\Phi_{NT,8,t}$,

$$\begin{aligned}\Phi_{NT,8,t} &= -\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\left(\widehat{\beta}_i^{(n)}(\tau_t)-\beta_i(\tau_t)\right)-\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)x_{it}^\top\left(\widehat{\beta}_i^{(n)}(\tau_t)-\beta_i(\tau_t)\right)\\ &:= \Phi_{NT,8,1,t}+\Phi_{NT,8,2,t}.\end{aligned}$$

Recall that $s_i(\tau)=[I_p,0_p][M_i(\tau)^\top W(\tau)M_i(\tau)]^{-1}M_i(\tau)^\top W(\tau)$, where $M_i(\tau)$ and $W(\tau)$ are defined in Section 2. For $\widehat{\beta}_i^{(n)}(\tau)$, write

$$\begin{aligned}\widehat{\beta}_i^{(n)}(\tau)-\beta_i(\tau) &= s_i(\tau)N_i(\tau)\left(\frac{1}{2}\beta_i''(\tau)h^2+o(h^2)\right)+s_i(\tau)\left(\widetilde{F}\widetilde{\lambda}_i-\widehat{F}^{(n-1)}\widehat{\lambda}_i^{(n)}\right)+s_i(\tau)\widetilde{\varepsilon}_i\\ &:= \Phi_{NT,11,i}(\tau)+\cdots+\Phi_{NT,13,i}(\tau),\end{aligned}\tag{C.19}$$

where $N_i(\tau)=\left((\frac{1-\tau T}{Th})^2x_{i1},(\frac{2-\tau T}{Th})^2x_{i2},\cdots,(\frac{T-\tau T}{Th})^2x_{iT}\right)^\top$. We now plug (C.19) into $\Phi_{NT,8,1,t}$ and compute its convergence rate. By Lemma C.2 and Cauchy-Schwarz inequality,

$$\frac{1}{N^2}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\Phi_{NT,11,i}(\tau_t)\right\|^2 \le \frac{1}{N^2}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\left\|\widetilde{\lambda}_i x_{it}^\top\right\|_F^2\right)\left(\sum_{i=1}^{N}\|\Phi_{NT,11,i}(\tau_t)\|^2\right)=O_P\left(Th^4\right).$$

Let $R_{Sf,i}(\tau)=s_i(\tau)F^0-\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)$. Write

$$\begin{aligned}\Phi_{NT,12,i}(\tau) &= -s_i(\tau)R_f^{(n-1)}\widetilde{\lambda}_i-s_i(\tau)\widetilde{F}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)-s_i(\tau)R_f^{(n-1)}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\\ &= -s_i(\tau)R_f^{(n-1)}\widetilde{\lambda}_i-\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)H\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)-R_{Sf,i}(\tau)H\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\\ &\quad -s_i(\tau)R_f^{(n-1)}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\\ &:= \Phi_{NT,12,1,i}(\tau)+\cdots+\Phi_{NT,12,4,i}(\tau).\end{aligned}\tag{C.20}$$

For $N^{-2}\sum_{t=1}^{T}\|\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\Phi_{NT,12,3,i}(\tau_t)\|^2$, by Lemma C.2(1) and Cauchy-Schwarz inequality,

$$\begin{aligned}\frac{1}{N^2}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top R_{Sf,i}(\tau)H\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\right\|_F^2 &\le \frac{1}{N^2}\left(\sum_{t=1}^{T}\sum_{i=1}^{N}\left\|\widetilde{\lambda}_i x_{it}^\top R_{Sf,i}(\tau_t)\right\|_F^2\right)\|H\|_F^2\left(\sum_{i=1}^{N}\left\|\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right\|^2\right)\\ &= O_P\left(\frac{1}{h}\max\left\{\delta_{f,n-1}^2,\frac{1}{T}\right\}\right).\end{aligned}\tag{C.21}$$

Analogously, we can use Lemma C.2(4) and Cauchy-Schwarz inequality to show

$$\frac{1}{N^2}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\Phi_{NT,12,4,i}(\tau_t)\right\|^2=O_P\left(T\delta_{f,n-1}^2\max\left\{\delta_{f,n-1}^2,\frac{1}{T}\right\}\right).\tag{C.22}$$

Additionally, Lemma C.2(2) implies $\frac{1}{N^2}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\Phi_{NT,13,i}(\tau_t)\right\|^2=O_P\left(\frac{1}{Nh}\right)$. Together with (C.20), (C.21) and (C.22), it yields

$$\begin{aligned}&\sum_{t=1}^{T}\left\|\Phi_{NT,8,1,t}+\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\left(\Phi_{NT,12,1,i}(\tau_t)+\Phi_{NT,12,2,i}(\tau_t)\right)\right\|^2\\ &= O_P\left(T\max\left\{\delta_{f,n-1}^2,\frac{\delta_{NT}^2}{h}\right\}\max\left\{\delta_{f,n-1}^2,\frac{1}{T}\right\}\right).\end{aligned}\tag{C.23}$$

Using (C.16) and Lemma C.2, we obtain $\sum_{t=1}^{T}\|\Phi_{NT,8,2,t}\|^2=O_P(T\left(h^4+\delta_{f,n-1}^2\right)\max\left\{\delta_{f,n-1},T^{-1/2}\right\}^2)$. Together with (C.23), it yields

$$\sum_{t=1}^{T}\left\|\Phi_{NT,8,t}+\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i x_{it}^\top\left(\Phi_{NT,12,1,i}(\tau_t)+\Phi_{NT,12,2,i}(\tau_t)\right)\right\|^2$$

$$= O_P\left(T\max\left\{\delta_{f,n-1}^2, \frac{\delta_{NT}^2}{h}, h^4\right\}\max\left\{\delta_{f,n-1}^2, \frac{1}{T}\right\}\right). \quad \text{(C.24)}$$

We now proceed with $\Phi_{NT,9,t}$. Write

$$\Phi_{NT,9,t} = -\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^{\top}\widetilde{f}_t - \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^{\top}\widetilde{f}_t.$$

(C.16) implies that the second term in $\Phi_{NT,9,t}$ is $O_P(T\max\{\delta_{f,n-1}^2, T^{-1}\}^2)$. Therefore,

$$\sum_{t=1}^{T}\left\|\Phi_{NT,9,t}+\frac{1}{N}\sum_{i=1}^{N}\widetilde{\lambda}_i\left(\widehat{\lambda}_i^{(n)}-\widetilde{\lambda}_i\right)^{\top}\widetilde{f}_t\right\|^2 = O_P\left(T\max\left\{\delta_{f,n-1}^2, \frac{1}{T}\right\}^2\right). \quad \text{(C.25)}$$

Combining (C.18), (C.24) and (C.25), we obtain

$$\sum_{t=1}^{T}\left\|\widehat{f}_t^{(n)}-\widetilde{f}_t-\widehat{b}_{f,t}^{(n)}-\widehat{b}_{\varepsilon,t}^{(n)}\right\|^2 = O_P\left(T\max\left\{\delta_{f,n-1}^2, \frac{\delta_{NT}^2}{h}, h^4\right\}\max\left\{\delta_{f,n-1}^2, \frac{1}{T}\right\}\right). \quad \text{(C.26)}$$

Proof of (2.c): For notational simplicity, we define $\Omega_{10}^{\mathcal{D}}(t,s) = N^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\mathbb{E}_{\mathcal{D}}(x_{it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is})$ and $\Omega_{20}^{\mathcal{D}}(t,s) = N^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\mathbb{E}_{\mathcal{D}}(z_{it}^{\top}\Omega_{f,i}^{\mathcal{D}-1}z_{is})$. Therefore, $\Omega_1^{\mathcal{D}}(t,s) = (N^{-1}\Lambda^{0\top}\Lambda^0)^{-1}\Omega_{10}^{\mathcal{D}}(t,s)$ and $\Omega_2^{\mathcal{D}}(t,s) = (N^{-1}\Lambda^{0\top}\Lambda^0)^{-1}\Omega_{20}^{\mathcal{D}}(t,s)$. Additionally, denote $\Omega_1^H(t,s) = H^{\top}\Omega_1^{\mathcal{D}}(t,s)H^{-1\top}$, $\Omega_2^H(t,s) = H^{\top}\Omega_2^{\mathcal{D}}(t,s)H^{-1\top}$, $\Omega_{10}^H(t,s) = H^{-1}\Omega_{10}^{\mathcal{D}}(t,s)H^{-1\top}$ and $\Omega_{20}^H(t,s) = H^{-1}\Omega_{20}^{\mathcal{D}}(t,s)H^{-1\top}$. Define

$$b_{f,t}^{(n)} = \sum_{s=1}^{T}\left(h^{-1}K_{s,0}(\tau_t)\Omega_1^H(t,s)+\Omega_2^H(t,s)\right)R_{f,s}^{(n-1)},\ b_{\varepsilon,t} = \frac{1}{T}\left(\sum_{i=1}^{N}\widetilde{\lambda}_i\widetilde{\lambda}_i^{\top}\right)^{-1}\sum_{i=1}^{N}\sum_{s=1}^{T}\left(\widetilde{\lambda}_i+\varepsilon_{is}^*\right)\varepsilon_{it},$$

where $\varepsilon_{is}^* = H^{-1}\Omega_{f,i}^{\mathcal{D}-1}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_s)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_s)\mu_{x,i}^{\mathcal{D}}(\tau_s)\varepsilon_{is}$. Lemma D.2(3) implies $\sum_{t=1}^{T}\|b_{\varepsilon,t}\|^2 = O_P\left(T\delta_{NT}^2\right)$. Additionally, simple algebra yields $\sum_{t=1}^{T}\left\|b_{f,t}^{(n)}\right\|^2 = O_P\left(T\delta_{f,n-1}^2\right)$. We can then use Lemma C.5 and Cauchy-Schwarz inequality to obtain

$$\sum_{t=1}^{T}\left\|\widehat{b}_{f,t}^{(n)}-b_{f,t}^{(n)}\right\|^2 = O_P\left(T\max\left\{\delta_{f,n-1}^2, h^4, \frac{1}{Th}\right\}\max\left\{\delta_{f,n-1}^2, \frac{\delta_{NT}^2}{h}\right\}\right), \quad \text{(C.27)}$$

and

$$\sum_{t=1}^{T}\left\|\widehat{b}_{\varepsilon,t}^{(n)}-b_{\varepsilon,t}\right\|^2 = O_P\left(T\max\left\{\delta_{f,n-1}^2, h^4, \frac{1}{Th}\right\}\max\left\{\delta_{f,n-1}^2, \frac{1}{Th}\right\}\right)+O_P\left(\frac{T}{N}\max\left\{\delta_{f,n-1}, \delta_{NT}\right\}^2\right). \quad \text{(C.28)}$$

The results in (C.27) and (C.28) are obvious and we put the proof in the technical supplement (see Liu, 2023). With (C.27) and (C.28), we have

$$\begin{aligned}\sum_{t=1}^{T}\left\|\widehat{f}_t^{(n)}-\widetilde{f}_t\right\|^2 &= \sum_{t=1}^{T}\left\|\widehat{f}_t^{(n)}-\widetilde{f}_t-\widehat{b}_{f,t}^{(n)}-\widehat{b}_{\varepsilon,t}^{(n)}\right\|^2+\sum_{t=1}^{T}\left\|\widehat{b}_{f,t}^{(n)}-b_{f,t}^{(n)}\right\|^2+\sum_{t=1}^{T}\left\|\widehat{b}_{\varepsilon,t}^{(n)}-b_{\varepsilon,t}\right\|^2\\&\quad+\sum_{t=1}^{T}\left\|b_{f,t}^{(n)}\right\|^2+\sum_{t=1}^{T}\|b_{\varepsilon,t}\|^2+\text{interaction terms}\\&= O_P\left(T\max\left\{\delta_{f,n-1}, \delta_{NT}\right\}^2\right).\end{aligned}$$

Therefore, Proposition C.1(3) holds. ■

Proof of Proposition C.2

(1) Using the arguments that are analogous to those in the proof of Proposition C.1(2), we obtain

$$\sqrt{N}\left(\widehat{f}_t^{(n)}-\widetilde{f}_t\right) = \sqrt{N}b_{f,t}^{(n)}+\sqrt{N}b_{\varepsilon,t}+o_P\left(\frac{1}{\sqrt{N}}\right). \quad \text{(C.29)}$$

With (C.29), to prove Proposition C.2(1), it suffices to show that $b_{f,t}^{(n)} = b_{f,t}^{\dagger(n)} + o_P(N^{-1/2})$ and $b_{\varepsilon,t} = H^\top \left(\Lambda^{0\top}\Lambda^0\right)^{-1}\left(\sum_{i=1}^N \lambda_i^0 \varepsilon_{it}\right) + O_P\left(\frac{1}{T}\right) + o_P\left(\frac{1}{\sqrt{N}}\right)$. The convergence rate of $b_{f,t}^{(n)}$ depends on $R_{f,s}^{(n-1)}$. Write

$$R_{f,s}^{(n-1)} - b_{f,s}^{(n-1)} = \left(\widehat{f}_s^{(n-1)} - \widetilde{f}_s - \widehat{b}_{f,s}^{(n-1)} - \widehat{b}_{\varepsilon,s}^{(n-1)}\right) + \left(\widehat{b}_{f,s}^{(n-1)} - b_{f,s}^{(n-1)}\right) + \left(\widehat{b}_{\varepsilon,s}^{(n-1)} - b_{\varepsilon,s}\right) + b_{\varepsilon,s}.$$

By (C.26) and Cauchy-Schwarz inequality,

$$\begin{aligned}
&\left\|\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)\left(\widehat{f}_s^{(n-1)} - \widetilde{f}_s - \widehat{b}_{f,s}^{(n-1)} - \widehat{b}_{\varepsilon,s}^{(n-1)}\right)\right\| \\
&\le \left(\sum_{s=1}^T K_{s,0}(\tau_t)\|\Omega_{10}^H(t,s)\|_F^2\right)^{\frac{1}{2}} \left(\sum_{s=1}^T K_{s,0}(\tau_t)\left\|\widehat{f}_s^{(n-1)} - \widetilde{f}_s - \widehat{b}_{f,s}^{(n-1)} - \widehat{b}_{\varepsilon,s}^{(n-1)}\right\|^2\right)^{\frac{1}{2}} \\
&= O_P\left(Th\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,n-2}, \frac{1}{\sqrt{T}}\right\}\right).
\end{aligned}$$

Analogously, (C.27) and (C.28) imply

$$\begin{aligned}
\left\|\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)\left(\widehat{b}_{f,s}^{(n-1)} - b_{f,s}^{(n-1)}\right)\right\| &= O_P\left(Th\max\left\{\delta_{f,n-2}, h^2, \frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right), \\
\left\|\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)\left(\widehat{b}_{\varepsilon,s}^{(n-1)} - b_{\varepsilon,s}\right)\right\| &= O_P\left(Th\max\left\{\delta_{f,n-2}, h^2, \frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right).
\end{aligned}$$

By Assumption 2 and (C.17),

$$\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)b_{\varepsilon,s} = \sum_{i=1}^N\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)(\widetilde{\Lambda}^\top\widetilde{\Lambda})^{-1}\widetilde{\lambda}_i\varepsilon_{is} + O_P(1) = O_P\left(\max\left\{\sqrt{\frac{Th}{N}}, 1\right\}\right),$$

where the last equality holds because $\sum_{i=1}^N\sum_{s=1}^T K_{s,0}(\tau)\widetilde{\lambda}_i\varepsilon_{it} = O_P(\sqrt{NTh})$, which can be proved by Lemma D.3(2). Therefore,

$$\left\|\sum_{s=1}^T K_{s,0}(\tau_t)\Omega_{10}^H(t,s)\left(R_{f,s}^{(n-1)} - b_{f,s}^{(n-1)}\right)\right\| = O_P\left(Th\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \tag{C.30}$$

Analogously, we have

$$\left\|\sum_{s=1}^T \Omega_{20}^H(t,s)\left(R_{f,s}^{(n-1)} - b_{f,s}^{(n-1)}\right)\right\| = O_P\left(T\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \tag{C.31}$$

Using (C.30) and (C.31) and conducting substitutions sequentially, we obtain

$$\left\|b_{f,t}^{(n)} - b_{f,t}^{\dagger(n)}\right\| = O_P\left(\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \tag{C.32}$$

For $b_{\varepsilon,t}$, by Assumptions 1 and 2, we have

$$\begin{aligned}
b_{\varepsilon,t} &= H^\top\left(\Lambda^{0\top}\Lambda^0\right)^{-1}\left(\sum_{i=1}^N \lambda_i^0\varepsilon_{it}\right) + \frac{1}{T}d_{f,t}^* + o_P\left(\frac{1}{\sqrt{N}}\right) \\
&= H^\top\left(\Lambda^{0\top}\Lambda^0\right)^{-1}\left(\sum_{i=1}^N \lambda_i^0\varepsilon_{it}\right) + O_P\left(\frac{1}{T}\right) + o_P\left(\frac{1}{\sqrt{N}}\right),
\end{aligned} \tag{C.33}$$

where $d^*_{f,t} = H^\top(\Lambda^{0\top}\Lambda^0)^{-1}\sum_{i=1}^N\sum_{s=1}^T \Omega_{f,i}^{\mathcal{D}-1}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_s)\mu_{x,i}^{\mathcal{D}}(\tau_s)\Sigma_{ii,ts}^{\mathcal{D}}$ and $\Sigma_{ij,ts}^{\mathcal{D}} = \mathbb{E}_{\mathcal{D}}(\varepsilon_{it}\varepsilon_{js})$. The second equality in (C.33) holds by Lemma D.2 and the fact $d^*_{f,t} = O_P(1)$ under the conditionally α-mixing conditions in Assumption 1. By (C.29), (C.32) and (C.33), Proposition C.2(1) holds.

(2) Recall that $\widehat{\lambda}_i^{(n)} - \widetilde{\lambda}_i = \Phi_{NT,1,i} + \Phi_{NT,2,i} + \Phi_{NT,3,i}$, where $\Phi_{NT,1,i}, \ldots, \Phi_{NT,3,i}$ are defined in (C.11). By Lemma C.3, we have $\widehat{\Omega}_{f,i}^{(n-1)} - \Omega_{f,i}^H \xrightarrow{P} 0$. By Lemma C.4(1), we have

$$\|\Phi_{NT,1,i}\| = O_P\left(\sqrt{\frac{h^3}{T}}\max\{1, \sqrt{Th}\delta_{f,n-1}\}\right). \tag{C.34}$$

For $\Phi_{NT,2,i}$, by Lemmas C.3 and C.4(2)(4),

$$\begin{aligned}\Phi_{NT,2,i} &= -\frac{1}{T}\left(\widehat{\Omega}_{f,i}^{(n-1)-1} - \Omega_{f,i}^{H-1}\right)\widehat{F}^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i - \frac{1}{T}\Omega_{f,i}^{H-1}R_f^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i \\ &\quad -\frac{1}{T}\Omega_{f,i}^{H-1}\widetilde{F}^\top\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i \\ &= -\frac{1}{T}\Omega_{f,i}^{H-1}\widetilde{F}^\top\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i + O_P\left(\delta_{f,n-1}\max\left\{\delta_{f,n-1}, \frac{1}{\sqrt{Th}}\right\}\right) \\ &= -\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^T \widetilde{z}_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i + O_P\left(\delta_{f,n-1}\max\left\{\delta_{f,n-1}, \frac{1}{\sqrt{Th}}\right\}\right).\end{aligned} \tag{C.35}$$

By Lemmas C.3, C.4 and Assumption 3,

$$\begin{aligned}\sqrt{T}\Phi_{NT,3,i} - \frac{1}{\sqrt{T}}\Omega_{f,i}^{H-1}\sum_{t=1}^T R_{f,t}^{(n-1)}\varepsilon_{it} &= \frac{1}{\sqrt{T}}\left(\widehat{\Omega}_{f,i}^{(n-1)-1} - \Omega_{f,i}^{H-1}\right)\widehat{F}^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i + \frac{1}{\sqrt{T}}\Omega_{f,i}^{H-1}\widetilde{F}^\top\Omega_{S,i}\widetilde{\varepsilon}_i \\ &\quad + \frac{1}{\sqrt{T}}\Omega_{f,i}^{H-1}\left(R_f^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i - R_f^{(n-1)\top}\widetilde{\varepsilon}_i\right) \\ &= \frac{1}{\sqrt{T}}\Omega_{f,i}^{H-1}\widetilde{F}^\top\Omega_{S,i}\widetilde{\varepsilon}_i + o_P(1) \\ &= \frac{1}{\sqrt{T}}\Omega_{f,i}^{H-1}\sum_{t=1}^T \widetilde{z}_{it}\varepsilon_{it} + o_P(1).\end{aligned} \tag{C.36}$$

For $\sum_{t=1}^T R_{f,t}^{(n-1)}\varepsilon_{it}$, we write

$$\begin{aligned}\sum_{t=1}^T R_{f,t}^{(n-1)}\varepsilon_{it} &= \sum_{t=1}^T \varepsilon_{it}\left(\widehat{F}_t^{(n-1)} - \widetilde{f}_t\right) \\ &= \sum_{t=1}^T \varepsilon_{it}\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right) + \sum_{t=1}^T \varepsilon_{it}\left(\widehat{b}_{f,t}^{(n-1)} - b_{f,t}^{(n-1)}\right) + \sum_{t=1}^T \varepsilon_{it}\left(\widehat{b}_{\varepsilon,t}^{(n-1)} - b_{\varepsilon,t}\right) \\ &\quad + \sum_{t=1}^T \varepsilon_{it}b_{f,t}^{(n-1)} + \sum_{t=1}^T \varepsilon_{it}b_{\varepsilon,t}.\end{aligned}$$

By (C.26) and Cauchy-Schwarz inequality,

$$\begin{aligned}\left\|\sum_{t=1}^T \varepsilon_{it}\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right)\right\| &\le \left(\sum_{t=1}^T |\varepsilon_{it}|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T \left\|\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}} \\ &= O_P\left(T\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,n-2}, \frac{1}{\sqrt{T}}\right\}\right).\end{aligned} \tag{C.37}$$

By (C.27),

$$\left\|\sum_{t=1}^T \varepsilon_{it}\left(\widehat{b}_{f,t}^{(n-1)} - b_{f,t}^{(n-1)}\right)\right\| \le \left(\sum_{t=1}^T |\varepsilon_{it}|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T \left\|\widehat{b}_{f,t}^{(n-1)} - b_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}$$

$$= O_P\left(T\max\left\{\delta_{f,n-2}, h^2, \frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \quad (C.38)$$

Analogously, by (C.28),

$$\left\|\sum_{t=1}^{T}\varepsilon_{it}\left(\widehat{b}_{\varepsilon,t}^{(n-1)} - b_{\varepsilon,t}\right)\right\| = O_P\left(T\max\left\{\delta_{f,n-2}, h^2, \frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right).$$

Since $b_{f,t}^{(n-1)} = T^{-1}\sum_{s=1}^{T}\left(h^{-1}K_{s,0}(\tau_t)\Omega_1^H(t,s) + \Omega_2^H(t,s)\right)R_{f,s}^{(n-2)}$, by Cauchy-Schwarz inequality,

$$\begin{aligned}\left\|\sum_{t=1}^{T}\varepsilon_{it}b_{f,t}^{(n-1)}\right\| &\leq \frac{1}{T}\left(\sum_{s=1}^{T}\left\|R_{f,s}^{(n-2)}\right\|\right)^{\frac{1}{2}}\left(\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\left(h^{-1}K_{s,0}(\tau_t)\Omega_1^H(t,s) + \Omega_2^H(t,s)\right)\varepsilon_{it}\right\|_F\right)^{\frac{1}{2}} \\ &= O_P\left(\sqrt{T}\delta_{f,n-2}\right). \quad (C.39)\end{aligned}$$

Recall that $b_{\varepsilon,t} = T^{-1}(\widetilde{\Lambda}^\top\widetilde{\Lambda})^{-1}\sum_{i=1}^{N}\sum_{s=1}^{T}\left(\widetilde{\lambda}_i + \varepsilon_{is}^*\right)\varepsilon_{it}$. By the weak cross-sectional dependence of ε_{it}, we can use simple algebra to show

$$\sum_{t=1}^{T}\varepsilon_{it}b_{\varepsilon,t} = \frac{1}{T}\left(\widetilde{\Lambda}^\top\widetilde{\Lambda}\right)^{-1}\sum_{j=1}^{N}\sum_{t=1}^{T}\widetilde{\lambda}_j\varepsilon_{jt}\varepsilon_{it} + O_P(1). \quad (C.40)$$

By (C.37), (C.38), (C.39) and (C.40),

$$\begin{aligned}&\left\|\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}R_{f,t}^{(n-1)}\varepsilon_{it} - \frac{1}{T}\left(\widetilde{\Lambda}^\top\widetilde{\Lambda}\right)\sum_{j=1}^{N}\sum_{t=1}^{T}\widetilde{\lambda}_j\varepsilon_{jt}\varepsilon_{it}\right\| \\ &= O_P\left(\max\left\{\delta_{f,n-2}, h^2, \frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \quad (C.41)\end{aligned}$$

By (C.34), (C.35), (C.36) and (C.41), the following result holds

$$\sqrt{T}\left(\widehat{\lambda}_i^{(n)} - \widetilde{\lambda}_i\right) = \sqrt{T}b_{\lambda,i}^{(n)} + \sqrt{T}B_{\varepsilon,i} + o_P\left(\frac{1}{\sqrt{T}}\right), \quad (C.42)$$

where $b_{\lambda,i}^{(n)} = -T^{-1}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top R_{f,t}^{(n-1)}$ and

$$B_{\varepsilon,i} = (NT)^{-1}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\sum_{j=1}^{N}\widetilde{z}_{it}\varepsilon_{it} + T^{-1}\Omega_{f,i}^{H-1}\left(\widetilde{\Lambda}^\top\widetilde{\Lambda}\right)^{-1}\sum_{t=1}^{T}\sum_{j=1}^{N}\widetilde{\lambda}_j\varepsilon_{jt}\varepsilon_{it}.$$

For the term with $\widetilde{z}_{it}\widetilde{\lambda}_i^\top$ in $b_{\lambda,i}^{(n)}$, write

$$\begin{aligned}&\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top R_{f,t}^{(n-1)} = \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t\right) \\ &= \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right) + \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top\left(\widehat{b}_{f,t}^{(n-1)} - b_{f,t}^{(n-1)}\right) + \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top b_{f,t}^{(n-1)} \\ &+ \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top\left(\widehat{b}_{\varepsilon,t}^{(n-1)} - b_{\varepsilon,t}\right) + \sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top b_{\varepsilon,t}.\end{aligned}$$

Similarly to (C.37) and (C.38), by (C.26), (C.27) and (C.28), we obtain

$$\left\|\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right)\right\| = O_P\left(T\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,n-2}, \frac{1}{\sqrt{T}}\right\}\right),$$

$$\left\|\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}\left(\widehat{b}_{f,t}^{(n-1)}-b_{f,t}^{(n-1)}\right)\right\| = O_P\left(T\max\left\{\delta_{f,n-2},h^2,\frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2},\frac{\delta_{NT}}{\sqrt{h}}\right\}\right),$$
$$\left\|\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}\left(\widehat{b}_{\varepsilon,t}^{(n-1)}-b_{\varepsilon,t}\right)\right\| = o_P\left(\sqrt{N}\right)+o_P\left(\sqrt{T}\right). \tag{C.43}$$

Recall that $\lambda_i^{\dagger}(\tau)=\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\mu_{x,i}^{\mathcal{D}}(\tau)\widetilde{\lambda}_i^{\top}$. For $\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}b_{f,t}^{(n-1)}$, since $\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}+H^{\top}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\lambda_i^{\dagger}(\tau_t)$ has zero mean and is α-mixing conditional on $\mathcal{D}$,

$$\begin{aligned}\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}b_{f,t}^{(n-1)} &= -H^{\top}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\lambda_i^{\dagger}(\tau_t)b_{f,t}^{(n-1)}+\sum_{t=1}^{T}\left(\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}+H^{\top}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\lambda_i^{\dagger}(\tau_t)\right)b_{f,t}^{(n-1)}\\ &= -H^{\top}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\lambda_i^{\dagger}(\tau_t)\widehat{b}_{f,t}^{(n-1)}+O_P\left(\sqrt{T}\delta_{f,n-2}\right).\end{aligned}$$

For $\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}b_{\varepsilon,t}$, by Lemma D.2,

$$\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}\widehat{b}_{\varepsilon,t}^{(n-1)}=\sum_{t=1}^{T}\sum_{j=1}^{N}\widetilde{z}_{it}\lambda_i^{0\top}\left(\Lambda^{0\top}\Lambda^{0}\right)^{-1}\lambda_j^{0}\varepsilon_{jt}=O_P\left(\sqrt{\frac{T}{N}}\right). \tag{C.44}$$

By (C.32) and (C.43)-(C.44), we have

$$\left\|\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^{\top}R_{f,t}^{(n-1)}+b_{\lambda,i}^{\dagger(n)}\right\|_F = O_P\left(\max\left\{\delta_{f,n-2},h^2,\frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,n-2},\frac{\delta_{NT}}{\sqrt{h}}\right\}\right), \tag{C.45}$$

where $b_{\lambda,i}^{\dagger(n)}=T^{-1}H^{-1}\Omega_{f,i}^{\mathcal{D}-1}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\sum_{t=1}^{T}\lambda_i^{\dagger}(\tau_t)b_{f,t}^{\dagger(n-1)}$ and it satisfies

$$\left\|b_{\lambda,i}^{\dagger(n)}\right\|^2=O_P\left(T^{-1}\sum_{t=1}^{T}\left\|b_{f,t}^{\dagger(n-1)}\right\|^2\right). \tag{C.46}$$

By (C.41) and (C.45),

$$\left\|b_{\lambda,i}^{(n)}-b_{\lambda,i}^{\dagger(n)}\right\|_F=O_P\left(\max\left\{\delta_{f,n-2},h^2,\frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,n-2},\frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \tag{C.47}$$

For $B_{\varepsilon,i}$, by Lemma D.2, Assumptions 1 and 2

$$\begin{aligned}B_{\varepsilon,i} &= \frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}\varepsilon_{it}+\frac{1}{N}d_{\lambda,i}^{*}+o_P\left(\frac{1}{\sqrt{T}}\right)\\ &= \frac{1}{T}H^{-1}\Omega_{f,i}^{\mathcal{D}-1}\sum_{t=1}^{T}z_{it}\varepsilon_{it}+O_P\left(\frac{1}{N}\right)+o_P\left(\frac{1}{\sqrt{T}}\right),\end{aligned} \tag{C.48}$$

where $d_{\lambda,i}^{*}=NH^{-1}\Omega_{f,i}^{\mathcal{D}-1}(\Lambda^{0\top}\Lambda^{0})^{-1}\sum_{j=1}^{N}\lambda_j^{0}\Sigma_{ij,11}^{\mathcal{D}}$ and it satisfies $\|d_{\lambda,i}^{*}\|=O_P(1)$. By (C.42), (C.47), (C.46) and (C.48), Proposition C.2(2) holds.

(3) Recall that $s_i(\tau)=[I_p,0_p][M_i(\tau)^{\top}W(\tau)M_i(\tau)]^{-1}M_i(\tau)^{\top}W(\tau)$ and $\widehat{\beta}_i^{(n)}(\tau)-\beta_i(\tau)=\Phi_{NT,11,i}(\tau)+\cdots+\Phi_{NT,13,i}(\tau)$, where $\Phi_{NT,11,i}(\tau)$, $\cdots$, $\Phi_{NT,13,i}(\tau)$ are defined in (C.19). By Lemma C.1(1), $\Phi_{NT,11,i}(\tau)$ converges to a biased term with the order $O_P(h^2)$:

$$\left\|\Phi_{NT,11,i}(\tau)-\frac{\mu_2}{2}\beta_i''(\tau)h^2\right\|=o_P(h^2), \tag{C.49}$$

where $\mu_2 = \int u^2 K(u)du$. For $\Phi_{NT,12,i}(\tau)$, recall that $\Phi_{NT,12,i}(\tau) = \Phi_{NT,12,1,i}(\tau) + \cdots + \Phi_{NT,12,4,i}(\tau)$, where $\Phi_{NT,12,1,i}(\tau), \cdots, \Phi_{NT,12,4,i}(\tau)$ are defined in (C.20). By Lemma C.1(4), we have

$$\left\|\Phi_{NT,12,1,i}(\tau) + \frac{1}{Th}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i\right\| = O_P\left(\frac{\delta_{f,n-1}}{\sqrt{Th}}\right). \tag{C.50}$$

Using arguments that are analogous to those in the proof of (C.42), we can show that

$$\left\|\Phi_{NT,12,2,i}(\tau) + \Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)b_{\lambda,i}^{(n)}\right\| \leq O_P(1)\left\|\widehat{\lambda}_i^{(n)} - \widetilde{\lambda}_i - b_{\lambda,i}^{(n)}\right\| = O_P\left(\delta_{NT}\right),$$

where $b_{\lambda,i}^{(n)} = -\frac{1}{T}\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}\widetilde{\lambda}_i^\top R_{f,t}^{(n-1)}$. By Lemma C.1(3) and (C.16),

$$\|\Phi_{NT,12,3,i}(\tau)\| = O_P\left(\frac{1}{\sqrt{Th}}\max\left\{\delta_{f,n-1}, \frac{1}{\sqrt{T}}\right\}\right), \ \|\Phi_{NT,12,4,i}(\tau)\| = O_P\left(\delta_{f,n-1}\max\left\{\delta_{f,n-1}, \frac{1}{\sqrt{T}}\right\}\right). \tag{C.51}$$

By (C.50) and (C.51),

$$\left\|\Phi_{NT,12,i}(\tau) - b_{\beta,i}^{(n)}(\tau)\right\| = O_P\left(\max\left\{\delta_{f,n-1}^2 \frac{1}{\sqrt{T}}\right\}\right), \tag{C.52}$$

where $b_{\beta,i}^{(n)} = -(Th)^{-1}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i - \Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)b_{\lambda,i}^{(n)}$. We then proceed with $b_{\beta,i}^{(n)}$. By (C.26) and Cauchy-Schwarz inequality,

$$\begin{aligned}
&\left\|\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top\left(\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right)\right\| \\
&\leq \left(\sum_{t=1}^{T} K_{t,0}^2(\tau)\left\|x_{it}\widetilde{\lambda}_i^\top\right\|_F^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\left\|\widehat{f}_t^{(n-1)} - \widetilde{f}_t - \widehat{b}_{f,t}^{(n-1)} - \widehat{b}_{\varepsilon,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}} \\
&= O_P\left(\sqrt{T^2h}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}, h^2\right\}\max\left\{\delta_{f,n-2}, \frac{1}{\sqrt{T}}\right\}\right).
\end{aligned} \tag{C.53}$$

Analogously, by (C.27) and (C.32),we have

$$\left\|\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top\left(\widehat{b}_{f,t}^{(n-1)} - b_{f,t}^{(n-1)}\right)\right\| = O_P\left(\sqrt{T^2h}\max\left\{\delta_{f,n-2}, h^2, \frac{1}{\sqrt{Th}}\right\}\max\left\{\delta_{f,n-2}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right),$$

$$\left\|\frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top\left(b_{f,t}^{(n-1)} - b_{f,t}^{\dagger(n-1)}\right)\right\| = O_P\left(\max\left\{\delta_{f,0}, h^2, \frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right).$$

For $\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)}$, since $x_{it} - \mu_{x,i}^{\mathcal{D}}(\tau_t)$ is a conditionally α-mixing process and satisfies $\mathbb{E}_{\mathcal{D}}(x_{it} - \mu_{x,i}^{\mathcal{D}}(\tau_t)) = 0$,

$$\begin{aligned}
\frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)} &= \frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)\mu_{x,i}^{\mathcal{D}}(\tau_t)\widetilde{\lambda}_i^\top b_{f,t}^{(n-1)} \\
&\quad + \frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)\left(x_{it} - \mu_{x,i}^{\mathcal{D}}(\tau_t)\right)\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)} \\
&= \frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)\mu_{x,i}^{\mathcal{D}}(\tau_t)\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)} + O_P\left(\frac{\delta_{f,n-2}}{\sqrt{Th}}\right).
\end{aligned} \tag{C.54}$$

Similarly to (C.44), by (C.17), we have $\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top \widehat{b}_{\varepsilon,t}^{(n-1)} = O_P\left(\sqrt{\frac{Th}{N}}\right)$. Together with (C.53) and (C.54), it yields

$$\begin{aligned}
&\left\|\frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\widetilde{\lambda}_i^\top R_{f,t}^{(n-1)} - \frac{1}{Th}\sum_{t=1}^{T} K_{t,0}(\tau)\mu_{x,i}^{\mathcal{D}}(\tau_t)\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)}\right\| \\
&= O_P\left(\max\left\{\delta_{f,0}, h^2, \frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right). \qquad (C.55)
\end{aligned}$$

By (C.47) and (C.55),

$$\left\|b_{\beta,i}^{(n)}(\tau) - b_{\beta,i}^{\dagger(n)}(\tau)\right\| = O_P\left(\max\left\{\delta_{f,0}, h^2, \frac{\delta_{NT}}{\sqrt{h}}\right\}\max\left\{\delta_{f,0}, \frac{\delta_{NT}}{\sqrt{h}}\right\}\right), \qquad (C.56)$$

where $b_{\beta,i}^{\dagger(n)}(\tau)$ is defined as

$$b_{\beta,i}^{\dagger(n)}(\tau) = -\frac{1}{Th}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\sum_{t=1}^{T} K_{t,0}(\tau)\mu_{x,i}^{\mathcal{D}}(\tau_t)\widetilde{\lambda}_i^\top b_{f,t}^{\dagger(n-1)} - \frac{1}{T}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)\Omega_{f,i}^{\mathcal{D}-1}\sum_{t=1}^{T}\lambda_i^\dagger(\tau_t)b_{f,t}^{\dagger(n-1)}.$$

For each term in $b_{\beta,i}^{\dagger(n)}(\tau)$, we can use Cauchy-Schwarz inequality to show that it can at most have the same probability order as $T^{-1}\sum_{t=1}^{T}\left\|b_{f,t}^{\dagger(n-1)}\right\|^2$. Therefore,

$$\left\|b_{\beta,i}^{\dagger(n)}(\tau)\right\|^2 = O_P\left(T^{-1}\sum_{t=1}^{T}\left\|b_{f,t}^{\dagger(n-1)}\right\|^2\right). \qquad (C.57)$$

By (C.52), (C.56), (C.57) and Assumption 2,

$$\|\Phi_{NT,12,i}(\tau)\| = O_P\left(\left(T^{-1}\sum_{t=1}^{T}\|b_{f,t}^{\dagger(n-1)}\|^2\right)^{1/2}\right) + o_P\left(\frac{1}{\sqrt{Th}}\right) + o_P\left(h^2\right). \qquad (C.58)$$

For $\Phi_{NT,13,i}(\tau)$, directly using Lemma C.1(2), Assumptions 1 and 2, we obtain

$$\begin{aligned}
\sqrt{Th}\Phi_{NT,13,i}(\tau) &= \frac{1}{\sqrt{Th}}\left[I_p, 0_p\right]\Omega_{x,i}^{\mathcal{D}-1}(\tau)M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i + o_P(1) \\
&= \Sigma_{x,i}^{\mathcal{D}-1}(\tau)\left(\frac{1}{\sqrt{Th}}\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\varepsilon_{it}\right) + o_P(1). \qquad (C.59)
\end{aligned}$$

By (C.49), (C.58) and (C.59), Proposition C.2(3) holds. ■

Proof of Theorem 3.2

Lemma D.2 implies $\|\sum_{i=1}^{N}\lambda_i^0\varepsilon_{it}\| = O_P(\sqrt{N})$ and $\|\sum_{t=1}^{T} z_{it}\varepsilon_{it}\| = O_P(\sqrt{T})$. Additionally, directly using Lemma D.3 gives $\|\sum_{t=1}^{T} K_{t,0}(\tau)x_{it}\varepsilon_{it}\| = O_P(\sqrt{Th})$. Together with Propositions C.1 and C.2, they prove Theorem 3.2. ■

C.1.3 Proof of Theorem 3.3

Assumption 3 and Proposition C.2 can jointly lead to the desired results in Theorem 3.3, if a $\kappa_{NT} \in [0,1)$ exists such that $b_{f,t}^{\dagger(n)} = O_P\left(\kappa_{NT}^{n-1}\delta_{f,0}\right) + o_P\left(\frac{1}{\sqrt{N}}\right) + o_P\left(\frac{1}{\sqrt{T}}\right)$. Recall that $\Omega_{10}^{\mathcal{D}}(t,s) = N^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\mathbb{E}_{\mathcal{D}}(x_{it}^\top\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is})$. For notational simplicity, let $\xi_{x,it} = x_{it} - \mu_{x,i}^{\mathcal{D}}(\tau_t)$. We have

$$\frac{1}{Th}\sum_{s=1}^{T} K_{s,0}(\tau_t)\Omega_{10}^{\mathcal{D}}(t,s) = \frac{1}{NTh}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\sum_{s=1}^{T} K_{s,0}(\tau_t)\mathbb{E}_{\mathcal{D}}\left(\xi_{x,it}^\top\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\xi_{x,is}\right)$$

$$+\frac{1}{NTh}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\sum_{s=1}^{T}K_{s,0}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\mu_{x,i}^{\mathcal{D}}(\tau_s)$$

$$= \frac{1}{NTh}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\sum_{s=1}^{T}K_{s,0}(\tau_t)\mathbb{E}_{\mathcal{D}}\left(\xi_{x,it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\xi_{x,is}\right)$$

$$+\frac{1}{N}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\mu_{x,i}^{\mathcal{D}}(\tau_t)+O_P\left(\frac{1}{Th}\right), \quad \text{(C.60)}$$

where the third inequality holds by the properties of the kernel function and Riemann integral. By the conditionally α–mixing conditions in Assumption 1, we have the following inequality:

$$\frac{1}{NTh}\sum_{i=1}^{N}\left\|\lambda_i^0\lambda_i^{0\top}\right\|_F\left|\sum_{s=1}^{T}K_{s,0}(\tau_t)\mathbb{E}_{\mathcal{D}}\left(\xi_{x,it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\xi_{x,is}\right)\right|$$

$$\leq \frac{1}{NTh}\sum_{i=1}^{N}\left\|\lambda_i^0\lambda_i^{0\top}\right\|_F\sum_{m_1=1}^{p}\sum_{m_2=1}^{p}\sum_{s=t+1}^{T}K_{s,0}(\tau_t)\Sigma_{x,m_1m_2i}^{(\mathcal{D}-1)}(\tau_t)\left|\mathbb{E}_{\mathcal{D}}\left(\xi_{x,m_1it}\xi_{x,m_2is}\right)\right|$$

$$+\frac{1}{NTh}\sum_{i=1}^{N}\left\|\lambda_i^0\lambda_i^{0\top}\right\|_F\sum_{m_1=1}^{p}\sum_{m_2=1}^{p}\sum_{s=1}^{t}K_{s,0}(\tau_t)\Sigma_{x,m_1m_2i}^{(\mathcal{D}-1)}(\tau_t)\left|\mathbb{E}_{\mathcal{D}}\left(\xi_{x,m_1it}\xi_{x,m_2is}\right)\right|+O_P\left(\frac{1}{Th}\right), \quad \text{(C.61)}$$

where $\Sigma_{x,m_1m_2i}^{(\mathcal{D}-1)}(\tau_t)$ is the (m_1,m_2)-th element of $\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)$, $\xi_{x,mit}$ is the m-th element of $\xi_{x,it}$, for $m_1,m_2=1,2,\ldots,p$. Then we can use the conditionally α-mixing version of Davydov's inequality (see pages 19-20 in Bosq (2012)) to compute the orders of these two terms in (C.61), For the first term,

$$\sum_{s=t+1}^{T}\left|\mathbb{E}_{\mathcal{D}}\left[\xi_{x,m_1it}\xi_{x,m_2is}\right]\right| = \sum_{s=1}^{T-t}\left|Cov_{\mathcal{D}}\left(\xi_{x,m_1i1},\xi_{x,m_2i,1+s}\right)\right|$$

$$\leq \sum_{s=1}^{T-t}\alpha_{ii}(s)^{\delta/(4+\delta)}\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_1i1}|^{2+\delta/2}\right)\right)^{2/(4+\delta)}\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_2i,1+s}|^{2+\delta/2}\right)\right)^{2/(4+\delta)},$$

almost surely, where $c_\delta=(4+\delta)/\delta\cdot 2^{(4+2\delta)/(4+\delta)}$. Moreover, using Cauchy-Schwarz inequality sequentially, we have

$$\mathbb{E}\left[\sum_{s=1}^{T-t}\alpha_{ii}(s)^{\delta/(4+\delta)}\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_1i1}|^{2+\delta/2}\right)\right)^{2/(4+\delta)}\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_2i,1+s}|^{2+\delta/2}\right)\right)^{2/(4+\delta)}\right]$$

$$\leq \sum_{s=1}^{T-t}\alpha_{ii}(s)^{\delta/(4+\delta)}\mathbb{E}\left[\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_1i,1}|^{2+\delta/2}\right)\right)^{4/(4+\delta)}\right]^{1/2}\mathbb{E}\left[\left(\mathbb{E}_{\mathcal{D}}\left(|\xi_{x,m_2i,1+s}|^{2+\delta/2}\right)\right)^{4/(4+\delta)}\right]^{1/2}$$

$$\leq \sum_{s=1}^{T-t}\alpha_{ii}(s)^{\delta/(4+\delta)}\mathbb{E}\left[|\xi_{x,m_1i,1}|^{2+\delta/2}\right]^{2/(4+\delta)}\mathbb{E}\left[|\xi_{x,m_2i,1+s}|^{2+\delta/2}\right]^{2/(4+\delta)}=O(1),$$

where the last equality holds by Assumption 1. It yields that $\sum_{s=t+1}^{T}\left|\mathbb{E}_{\mathcal{D}}\left[\xi_{x,m_1it}\xi_{x,m_2is}\right]\right|=O_P(1)$. Analogously, we can also obtain $\sum_{s=1}^{t}\left|\mathbb{E}_{\mathcal{D}}\left[\xi_{x,m_1it}\xi_{x,m_2is}\right]\right|=O_P(1)$. Therefore,

$$\frac{1}{NTh}\sum_{i=1}^{N}\left\|\lambda_i^0\lambda_i^{0\top}\right\|_F\left|\sum_{s=1}^{T}K_{s,0}(\tau_t)\mathbb{E}_{\mathcal{D}}\left[\xi_{x,it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\xi_{x,is}\right]\right|=O_P\left(\frac{1}{Th}\right). \quad \text{(C.62)}$$

By (C.60) and (C.62),

$$\frac{1}{Th}\sum_{s=1}^{T}K_{s,0}(\tau_t)\Omega_1^{\mathcal{D}}(t,s) = \left(\Lambda^{0\top}\Lambda^0\right)^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\mu_{x,i}^{\mathcal{D}}(\tau_t)+O_P\left(\frac{1}{Th}\right)$$

$$:= \quad \Omega_1^{\mathcal{D}*}(t) + O_P\left(\frac{1}{Th}\right). \tag{C.63}$$

Similarly to (C.60) and (C.62),

$$\begin{aligned}\frac{1}{T}\sum_{s=1}^{T}\Omega_2^{\mathcal{D}}(t,s) &= \frac{1}{T}\left(\Lambda^{0\top}\Lambda^0\right)^{-1}\sum_{i=1}^{N}\sum_{s=1}^{T}\lambda_i^0\lambda_i^{0\top}\mu_{z,i}^{\mathcal{D}\top}(\tau_s)\Omega_{f,i}^{\mathcal{D}-1}\mu_{z,i}^{\mathcal{D}}(\tau_t) + O_P\left(\frac{1}{T}\right)\\ &:= \frac{1}{T}\sum_{s=1}^{T}\Omega_2^{\mathcal{D}*}(t,s) + O_P\left(\frac{1}{T}\right), \end{aligned} \tag{C.64}$$

where $\mu_{z,i}^{\mathcal{D}} = \mathbb{E}_{\mathcal{D}}(z_{it})$ with $z_{it} = f_t^0 - \Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}$. By (C.63) and (C.64),

$$\begin{aligned} b_{f,t}^{\dagger(n)} &= \frac{1}{Th}H^\top\sum_{s=1}^{T}K_{s,0}(\tau_t)\Omega_1^{\mathcal{D}*}(t)^{n-1}\Omega_1^{\mathcal{D}}(t,s)H^{-1\top}R_{F,s}^{(0)}\\ &+\frac{1}{T^n}H^\top\sum_{s_1,s_2,\ldots,s_n=1}^{T}\Omega_2^{\mathcal{D}*}(t,s_1)\prod_{j=1}^{n-2}\Omega_2^{\mathcal{D}*}(s_j,s_{j+1}))\Omega_2^{\mathcal{D}}(s_{n-1},s_n))H^{-1\top}R_{f,s_n}^{(0)} + O_P\left(\frac{\delta_{f,0}}{Th}\right). \end{aligned}$$

In what follows, we show $\max_t\|\Omega_1^{\mathcal{D}*}(t)\|_\infty < 1$ and $\max_{s,t}\|\Omega_2^{\mathcal{D}*}(t,s)\|_\infty < 1$. It is clear to see that $\Omega_1^{\mathcal{D}*}(t) = 0$, if $\mu_{x,i}^{\mathcal{D}}(\tau) = 0$, for each i. If $\mu_{x,i}^{\mathcal{D}}(\tau) \neq 0$, since $\operatorname{rank}\left(\mu_{x,i}^{\mathcal{D}}(\tau)\mu_{x,i}^{\mathcal{D}\top}(\tau)\right) = 1$, it holds by Woodbury matrix formula that

$$\Sigma_{x,i}^{\mathcal{D}-1}(\tau) = \Sigma_{\xi,i}^{\mathcal{D}-1} - \frac{1}{1+\operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau)\mu_{x,i}^{\mathcal{D}\top}(\tau)\right)}\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau)\mu_{x,i}^{\mathcal{D}\top}(\tau)\Sigma_{\xi,i}^{\mathcal{D}-1},$$

where $\Sigma_{\xi,i}^{\mathcal{D}} = \mathbb{E}_{\mathcal{D}}\left(\xi_{x,it}\xi_{x,it}^\top\right)$ with $\xi_{x,it} = x_{it} - \mu_{x,i}^{\mathcal{D}}(\tau_t)$. It is clear to see that

$$\begin{aligned} &\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\mu_{x,i}^{\mathcal{D}}(\tau_t) = \operatorname{tr}\left(\left(\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t) + \Sigma_{\xi,i}^{\mathcal{D}}\right)^{-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)\\ &= \operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t) - \frac{1}{1+\operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)^2\right)\\ &= \frac{\operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)}{1+\operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)} := \frac{c_1(i,t)}{1+c_1(i,t)}, \end{aligned} \tag{C.65}$$

where $c_1(i,t) = \operatorname{tr}\left(\Sigma_{\xi,i}^{\mathcal{D}-1}\mu_{x,i}^{\mathcal{D}}(\tau_t)\mu_{x,i}^{\mathcal{D}\top}(\tau_t)\right)$. Let $\phi_{it} = \mu_{x,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\mu_{x,i}^{\mathcal{D}}(\tau_t)$. We have $\phi_{it} \in [0,1)$ for any i and t. Therefore, $\Omega_1^{\mathcal{D}*}(t) = \left(\Lambda^{0\top}\Lambda^0\right)^{-1}\sum_{i=1}^{N}\lambda_i^0\lambda_i^{0\top}\phi_{it}$ and it satisfies $\max_t\|\Omega_1^{\mathcal{D}*}(t)\|_\infty < 1$. Analogously, we obtain $\max_{t,s}\|\Omega_2^{\mathcal{D}*}(t,s)\|_\infty < 1$. Let $\kappa_{NT} = \max\{\max_t\|\Omega_1^{\mathcal{D}*}(t)\|_\infty, \max_{t,s}\|\Omega_2^{\mathcal{D}*}(t,s)\|_\infty\}$. We have $\kappa_{NT} \in [0,1)$ and $b_{f,t}^{\dagger(n)} = O_P\left(\kappa_{NT}^{n-1}\delta_{f,0}\right) + O_P\left(\frac{\delta_{f,0}}{Th}\right)$. Therefore, Theorem 3.3 holds. ■

C.1.4 Proofs of Theorems 3.4, 4.1, 4.2 and 4.3

Proof of Theorem 3.4

We can use arguments that are closely related to those in the proof of Theorem 3.3 to establish the asymptotic properties of the mean group estimator. We provide the proof in full in the technical supplement (see Liu, 2023). ■

Proof of Theorem 4.1

(1) Recall that the test statistics L_{NT} and $\check{L}_{NT}$ are defined as follows:

$$L_{NT} = \frac{1}{NT\sqrt{h}}\sum_{i=1}^{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T}K\left(\frac{\tau_t-\tau_s}{h}\right)\widehat{e}_{it}\widehat{e}_{ns}, \quad \check{L}_{NT} = \frac{1}{\sqrt{\widehat{\sigma}_L^2}}L_{NT},$$

where $\widehat{\sigma}_L^2 = 2v_0\widehat{\overline{\sigma}}_\varepsilon^4$ with $\widehat{\overline{\sigma}}_\varepsilon^2 = (NT)^{-1}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \widehat{e}_{it}\widehat{e}_{jt}$. It suffices to prove the Theorem 4.1(1) by showing (a) $L_{NT} \xrightarrow{D} \mathcal{N}\left(0,\sigma_L^2\right)$ and (b) $\widehat{\sigma}_L^2 \xrightarrow{P} \sigma_L^2$, under $\mathcal{H}_0$.

For (a), we can use the CLT for U-statistic to establish the asymptotic distribution of L_{NT}. Specifically, we write

$$
\begin{aligned}
NT\sqrt{h}L_{NT} &= \sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{it}\varepsilon_{ns} + 2\sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)(\widehat{e}_{it}-\varepsilon_{it})\varepsilon_{ns} \\
&\quad + \sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)(\widehat{e}_{it}-\varepsilon_{it})(\widehat{e}_{ns}-\varepsilon_{ns}) \\
&:= L_{NT,1}+L_{NT,2}+L_{NT,3}.
\end{aligned}
$$

Among these three terms, $L_{NT,1}$ can be written into a U-statistic that generates the CLT. We put the result in Lemma C.7. For the remaining terms with $\widehat{e}_{it}-\varepsilon_{it}$, write

$$
\widehat{e}_{it}-\varepsilon_{it} = x_{it}^\top\left(\beta_i^0-\widetilde{\beta}_i^{(n)}\right)+\widetilde{\lambda}_i^\top\left(\widetilde{f}_t-\widetilde{f}_t^{(n)}\right)+\widetilde{f}_t^\top\left(\widetilde{\lambda}_i-\widetilde{\lambda}_i^{(n)}\right)+x_{it}^\top\left(\beta_i(\tau_t)-\beta_i^0\right) = \Delta_{e,it}+\Delta_{e,it}^c,
$$

where $\Delta_{e,it} = x_{it}^\top\left(\beta_i^0-\widetilde{\beta}_i^{(n)}\right)+\widetilde{\lambda}_i^\top\left(\widetilde{f}_t-\widetilde{f}_t^{(n)}\right)+\widetilde{f}_t^\top\left(\widetilde{\lambda}_i-\widetilde{\lambda}_i^{(n)}\right)$ and $\Delta_{e,it}^c = x_{it}^\top\left(\beta_i(\tau_t)-\beta_i^0\right)$.[1] It holds under $\mathcal{H}_0$ that $\Delta_{e,it}^c=0$. For $\Delta_{e,it}$, we use the following decomposition: $\Delta_{e,it}=\Delta_{e1,it}+\Delta_{e2,it}+\Delta_{e3,it}$, where $\Delta_{e1,it} = \widetilde{\lambda}_i^\top\left(\widetilde{f}_t-\widetilde{f}_t^{(n)}\right)$, $\Delta_{e2,it} = \widetilde{f}_t^\top\left(\widetilde{\lambda}_i-\widetilde{\lambda}_i^{(n)}\right)$, $\Delta_{e3,it} = x_{it}^\top\left(\beta_i^0-\widetilde{\beta}_i^{(n)}\right)$. Using arguments that are closely related to those in the proofs of Proposition C.2(1) and Theorem 3.3, we can readily obtain $\sum_{t=1}^T\left\|\widetilde{f}_t-\widetilde{f}_t^{(n)}-\widetilde{d}_{\varepsilon,t}\right\|^2 = O_P(T\delta_{NT}^4)$, where $\widetilde{d}_{\varepsilon,t} = H^\top\left(\Lambda^{0\top}\Lambda^0\right)^{-1}\left(\sum_{i=1}^N\lambda_i^0\varepsilon_{it}\right)$. By Assumption 5, Lemma D.3 and Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\left|\sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\widetilde{\lambda}_i^\top\left(\widetilde{f}_t-\widetilde{f}_t^{(n)}-\widetilde{d}_{\varepsilon,t}\right)\right| \\
&\leq \left(\sum_{t=1}^T\left\|\sum_{i=1}^N\sum_{n=1}^N\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\widetilde{\lambda}_i\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\left\|\widetilde{f}_t-\widetilde{f}_t^{(n)}-\widetilde{d}_{\varepsilon,t}\right\|^2\right)^{\frac{1}{2}} \\
&= O_P\left(\sqrt{NTh}\max\{N,T\}\right). \qquad (C.66)
\end{aligned}
$$

Additionally, by the conditionally α-mixing conditions in Assumption 1 and the cross-sectional independence conditions in Assumption 5,

$$
\left|\sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\widetilde{\lambda}_i^\top\widetilde{d}_{\varepsilon,t}\right| = O_P\left(NTh\right). \qquad (C.67)
$$

Combining (C.66) and (C.67), we have

$$
\left|\sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\Delta_{e1,it}\right| = O_P\left(\sqrt{NTh}\max\{N,T\}\right). \qquad (C.68)
$$

Using analogous arguments, we can obtain the following results:

$$
\left|\sum_{i=1}^N\sum_{n=1}^N\sum_{t=1}^T\sum_{s=1,\neq t}^T K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\Delta_{e2,it}\right| = O_P\left(\sqrt{NTh}\max\{N,T\}\right), \qquad (C.69)
$$

[1] Here we re-define $\widetilde{f}_t$ and $\widetilde{\lambda}_i$ with the rotation matrix from the parametric initial estimation: $\widetilde{H} = (NT)^{-1}\sum_{i=1}^N\lambda_i^0\lambda_i^{0\top}F^{0\top}\widetilde{F}^{(0)}\widetilde{V}_{NT}^{-1}$, where $\widetilde{F}^{(0)}$ and $\widetilde{V}_{NT}$ are defined in Appendix B.1.

$$\left|\sum_{i=1}^{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{ns}\Delta_{e3,it}\right| = O_P\left(\sqrt{NT}h\max\{N,T\}\right). \tag{C.70}$$

By (C.68), (C.69), (C.70) and Assumption 5,

$$L_{NT,2} = o_P\left(NT\sqrt{h}\right). \tag{C.71}$$

Analogously, we can readily obtain

$$L_{NT,3} = o_P\left(NT\sqrt{h}\right). \tag{C.72}$$

Combining (C.71), (C.72) and Lemma C.7, we complete the proof of (a) $L_{NT} \xrightarrow{D} \mathcal{N}\left(0,\sigma_L^2\right)$. Then, we only need to prove (b) $\widehat{\sigma}_L^2 \xrightarrow{P} \sigma_L^2$. Recall that $\widehat{\sigma}_L^2 = 2v_0\widehat{\overline{\sigma}}_\varepsilon^4$ and $\sigma_L^2 = 2v_0\overline{\sigma}_\varepsilon^4$. Therefore, it suffices to show that $\widehat{\overline{\sigma}}_\varepsilon^2 \xrightarrow{P} \overline{\sigma}_\varepsilon^2$, where $\overline{\sigma}_\varepsilon^2 = \lim_{N\to\infty} N^{-1}\sum_{i=1}^N\sum_{j=1}^N E[\varepsilon_{it}\varepsilon_{jt}]$ and $\widehat{\overline{\sigma}}_\varepsilon^2 = (NT)^{-1}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \widehat{e}_{it}\widehat{e}_{jt}$. We write

$$\begin{aligned}
\widehat{\overline{\sigma}}_\varepsilon^2 &= \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \varepsilon_{it}\varepsilon_{jt} + \frac{2}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \varepsilon_{it}\left(\widehat{e}_{jt}-\varepsilon_{jt}\right) + \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \left(\widehat{e}_{it}-\varepsilon_{it}\right)\left(\widehat{e}_{jt}-\varepsilon_{jt}\right) \\
&:= L_{NT,4} + L_{NT,5} + L_{NT,6}.
\end{aligned}$$

For $L_{NT,4}$, it follows that $E[L_{NT,4}] = \overline{\sigma}_\varepsilon^2 + o(1)$. In addition,

$$\begin{aligned}
\mathbb{E}\left[\left(L_{NT,4}-\overline{\sigma}_\varepsilon^2\right)^2\right] &= \frac{1}{N^2T^2}\mathbb{E}\left(\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T(\varepsilon_{it}\varepsilon_{jt}-E[\varepsilon_{it}\varepsilon_{jt}])^2\right) \\
&= \frac{1}{N^2T^2}\sum_{i_1=1}^N\sum_{i_2=1}^N\sum_{j_1=1}^N\sum_{j_2=1}^N\sum_{t=1}^T\sum_{s=1}^T \mathbb{E}\left((\varepsilon_{i_1t}\varepsilon_{j_1t}-E[\varepsilon_{i_1t}\varepsilon_{j_1t}])(\varepsilon_{i_2s}\varepsilon_{j_2s}-E[\varepsilon_{i_2s}\varepsilon_{j_2s}])\right) \\
&= O\left(\frac{1}{T}\right),
\end{aligned}$$

where the last equality holds by the second condition in Assumption 2.(iii). Then by Chebyshev's inequality, we obtain

$$L_{NT,4} = \overline{\sigma}_\varepsilon^2 + O_P\left(1\right). \tag{C.73}$$

For $L_{NT,5}$, by Cauchy-Schwarz inequality,

$$|L_{NT,5}| \leq \frac{2}{NT}\left(\sum_{t=1}^T\left|\sum_{i=1}^N\varepsilon_{it}\right|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\left|\sum_{j=1}^N(\widehat{e}_{jt}-\varepsilon_{jt})\right|^2\right)^{\frac{1}{2}}.$$

Under $\mathcal{H}_0$, it follows

$$|L_{NT,5}| \leq \frac{2}{NT}\left(\sum_{t=1}^T\left|\sum_{i=1}^N\varepsilon_{it}\right|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\left|\sum_{j=1}^N\Delta_{e,it}\right|^2\right)^{\frac{1}{2}} = o_P\left(1\right). \tag{C.74}$$

Analogously, we can readily obtain $|L_{NT,6}| = o_P\left(1\right)$, under $\mathcal{H}_0$. Together with (C.73) and (C.74), it yields $\widehat{\overline{\sigma}}_\varepsilon^2 = \overline{\sigma}_\varepsilon^2 + o_P(1)$. Therefore, (b) $\widehat{\sigma}_L^2 \xrightarrow{P} \sigma_L^2$ can hold under $\mathcal{H}_0$, which completes the proof of Theorem 4.1(1).

(2) Note that $\Delta_{e,it}^c = 0$ cannot hold for all i under $\mathcal{H}_1$. Without loss of generality, we assume $\beta_i(\tau) = \beta_i^0 + \nu_{NT}\Delta_{\beta,i}(\tau)$, for $i = 1, 2, \ldots, N_a$ and $\beta_i(\tau) = \beta_i^0$ for $i = N_a + 1, \ldots, N$. In this case, $\Delta_{e,it}^c = x_{it}^\top \left(\beta_i(\tau_t) - \beta_i^0\right) = \nu_{NT} x_{it}^\top \Delta_{\beta,i}(\tau_t)$ for $i = 1, 2, \ldots, N_a$, and $\Delta_{e,it}^c = 0$ otherwise. Then,

$$
\begin{aligned}
\sum_{i=1}^{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right)\Delta_{e,it}^c &= N\nu_{NT}\sum_{i=1}^{N_a}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right) x_{it}^\top \Delta_{\beta,i}(\tau_t) \\
&= N\nu_{NT}\sum_{i=1}^{N_a}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right) \mathbb{E}(x_{it})^\top \Delta_{\beta,i}(\tau_t) \\
&\quad + N\nu_{NT}\sum_{i=1}^{N_a}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right) \left(x_{it} - \mathbb{E}(x_{it})\right)^\top \Delta_{\beta,i}(\tau_t) \\
&= O_P\left(N_a N T^2 h \nu_{NT}\right), \qquad \text{(C.75)}
\end{aligned}
$$

where the last equality holds by Lemma D.4(2). Analogously,

$$
\begin{aligned}
\sum_{i=1}^{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right)\Delta_{e,ns}^c\Delta_{e,it}^c &= \nu_{NT}^2\sum_{i=1}^{N_a}\sum_{n=1}^{N_a}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T} K\left(\frac{\tau_t-\tau_s}{h}\right) x_{it}^\top \Delta_{\beta,i}(\tau_t) x_{ns}^\top \Delta_{\beta,n}(\tau_s) \\
&= O_P\left(N_a^2 T^2 h \nu_{NT}^2\right). \qquad \text{(C.76)}
\end{aligned}
$$

Provided that $\nu_{NT} \to 0$, it suffices to require $N_a^2 T^2 h \nu_{NT}^2/(NT\sqrt{h}) \to \infty$ and then it holds automatically that $N_a N T^2 h \nu_{NT}/(NT\sqrt{h}) \to \infty$. Combining these results, if ν_{NT} and N_a satisfy $N_a N^{-\frac{1}{2}} T^{\frac{1}{2}} h^{\frac{1}{4}} \nu_{NT} \to \infty$, L_{NT} diverges under $\mathcal{H}_1$. Therefore, Theorem 4.1(2) holds. ■

Proof of Theorem 4.2

The arguments that are analogous to those in the proof of Theorem 3.4 in Su and Chen (2013) can be used to establish the bootstrap test statistic's asymptotic properties. We provide the proof in full in the technical supplement (see Liu, 2023). ■

Proof of Theorem 4.3

Since the IC method's consistency is well established in the literature of factor number selection (e.g., Bai and Ng, 2002), we put its proof in the technical supplement (see Liu, 2023). ■

C.2 The main lemmas

This appendix lists the main lemmas that are used in the proofs of our theorems and propositions.

Lemma C.1. *Let Assumptions 1-2 hold. For any given i, n and τ, as $N, T \to \infty$,*

(1) $\left\|s_i(\tau)N_i(\tau) - \mu_2 I_p\right\|_F = O_P\left((Th)^{-1/2}\right)$;

(2) $\left\|s_i(\tau)\widetilde{\varepsilon}_i - (Th)^{-1}\left[I_p, 0_p\right]\Omega_{x,i}^{\mathcal{D}-1}(\tau)M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i\right\|_F = O_P\left((Th)^{-1}\right)$;

(3) $\left\|s_i(\tau)F^0 - \Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)\right\|_F = O_P\left((Th)^{-1/2}\right)$;

(4) $\left\|s_i(\tau)R_f^{(n-1)} - (Th)^{-1}\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\sum_{t=1}^T K_{t,0}(\tau)x_{it}R_{f,t}^{(n-1)}\right\|_F = O_P\left(\delta_{f,n-1}(Th)^{-1/2}\right)$.

Lemma C.2. *Let Assumptions 1-2 hold. For any given i, n and τ, as $N, T \to \infty$,*

(1) $\sum_{i=1}^N \left\|R_{Sf,i}(\tau)\right\|_F^2 = O_P\left(N(Th)^{-1}\right)$, $\sum_{i=1}^N\sum_{t=1}^T \left\|R_{Sf,i}(\tau_t)\right\|_F^2 = O_P\left(Nh^{-1}\right)$, $\sum_{t=1}^T \left\|R_{Sf,i}(\tau_t)\right\|_F^2 = O_P\left(h^{-1}\right)$, $\sum_{i=1}^N\sum_{t=1}^T \left\|\lambda_i^0 x_{it}^\top R_{Sf,i}(\tau_t)\right\|_F^2 = O_P\left(Nh^{-1}\right)$;

(2) $\sum_{i=1}^{N}\|R_{S,i}(\tau)\|_F^2=O_P\left(N(Th)^{-1}\right)$, $\sum_{i=1}^{N}\sum_{t=1}^{T}\|R_{S,i}(\tau_t)\|_F^2=O_P\left(Nh^{-1}\right)$, $\sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,0}(\tau)\|R_{S,i}(\tau_t)\|_F^2=O_P(N)$, $\sum_{t=1}^{T}\|R_{S,i}(\tau_t)\|_F^2=O_P\left(h^{-1}\right)$;

(3) $\sum_{i=1}^{N}\sum_{t=1}^{T}\|s_i(\tau_t)\widetilde{\varepsilon}_i\|^2=O_P\left(Nh^{-1}\right)$, $\sum_{t=1}^{T}\|s_i(\tau_t)\widetilde{\varepsilon}\|^2=O_P\left(h^{-1}\right)$, $\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\lambda_i^0x_{it}^\top s_i(\tau_t)\widetilde{\varepsilon}_i\right\|^2=O_P\left(Nh^{-1}\right)$;

(4) $\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(NT\delta_{f,n-1}^2\right)$, $\sum_{t=1}^{T}\left\|s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(T\delta_{f,n-1}^2\right)$, $\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\lambda_i^0x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(NT\delta_{f,n-1}^2\right)$;

where $R_{Sf,i}(\tau)=s_i(\tau)F^0-\Sigma_{x,i}^{\mathcal{D}-1}(\tau)\Sigma_{xf,i}^{\mathcal{D}}(\tau)$ *and* $R_{S,i}(\tau_t)=s_i(\tau_t)N_i(\tau_t)-\mu_2I_p$.

Lemma C.3. *Let Assumptions 1-2 hold. For any given i and n, as $N,T\to\infty$,*

$$\widehat{\Omega}_{f,i}^{(n-1)}-\Omega_{f,i}^H=O_P\left(\max\left\{\delta_{f,n-1},(Th)^{-1/2}\right\}\right),$$

where $\widehat{\Omega}_{f,i}^{(n-1)}=T^{-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\widehat{F}^{(n-1)}$, $\Omega_{f,i}^H=H^\top\Omega_{f,i}^{\mathcal{D}}H$ *with* $\Omega_{S,i}=(I-S_i)^\top(I-S_i)$ *and* $\Omega_{f,i}^{\mathcal{D}}=\Sigma_f^{\mathcal{D}}-\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$.

Lemma C.4. *Let Assumptions 1-2 hold. For any given i and n, as $N,T\to\infty$,*

(1) $\left\|\widehat{F}^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top\right\|_F=O_P\left((Th^3)^{1/2}\max\{1,(Th)^{1/2}\delta_{f,n-1}\}\right)$;

(2) $\left\|R_f^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i\right\|=O_P\left(T\delta_{f,n-1}^2\right)$;

(3) $\left\|R_f^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i-R_f^{(n-1)\top}\widetilde{\varepsilon}_i\right\|=O_P\left(\delta_{f,n-1}\sqrt{T/h}\right)$;

(4) $\left\|F^{0\top}\Omega_{S,i}R_f^{(n-1)}-\sum_{t=1}^{T}z_{it}R_{f,t}^{(n-1)\top}\right\|_F=O_P\left(\delta_{f,n-1}\sqrt{T/h}\right)$;

(5) $\left\|F^{0\top}\Omega_{S,i}\widetilde{\varepsilon}_i-\sum_{t=1}^{T}z_{it}\varepsilon_{it}\right\|=O_P\left(1/\sqrt{h}\right)$.

Lemma C.5. *Let Assumptions 1-2 hold. For any given n, as $N,T\to\infty$,*

(1) $\sum_{i=1}^{N}\left\|\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top\right\|^2=O_P\left(NTh^3\max\left\{1,Th\delta_{f,n-1}^2\right\}\right)$;

(2) $\sum_{i=1}^{N}\left\|\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i-\Omega_{f,i}^{H-1}\sum_{t=1}^{T}\widetilde{z}_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i\right\|^2=O_P\left(NT^2\delta_{f,n-1}^2\max\left\{\delta_{f,n-1}^2,(Th)^{-1}\right\}\right)$;

(3) $\sum_{i=1}^{N}\left\|\widehat{\Omega}_{f,i}^{(n-1)-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i-\Omega_{f,i}^{H-1}\sum_{t=1}^{T}(\widetilde{z}_{it}+R_{f,t}^{(n-1)})\varepsilon_{it}\right\|^2=O_P\left(NTh^{-1}\max\left\{\delta_{f,n-1}^2,T^{-1}\right\}\right)$.

Lemma C.6. *Let Assumptions 1-3 hold. As $N,T\to\infty$,*

(1) $T^{-1}\widehat{F}^{(0)\top}F^0\xrightarrow{P}\mathcal{Q}$;

(2) $H\xrightarrow{P}\Sigma_\lambda\mathcal{Q}^\top V^{-1}$;

where $\mathcal{Q}=V_{\lambda f}^{1/2}U_{\lambda f}^\top\Sigma_\lambda^{-1/2}$, $V_{\lambda f}$ *is a* $r_0\times r_0$ *diagonal matrix with diagonal elements being the* r_0 *eigenvalues of the matrix* $\Sigma_\lambda^{1/2}\Sigma_f\Sigma_\lambda^{1/2}$ *in a descending order, and* $U_{\lambda f}$ *is the corresponding orthogonal eigenvector matrix satisfying* $U_{\lambda f}^\top U_{\lambda f}=I_{r_0}$. Σ_λ *and* Σ_f *are defined in Assumptions 2 and 3, respectively.*

Lemma C.7. *Let Assumptions 1-2 and 5 hold. As $N,T\to\infty$,*

$$\frac{1}{NT\sqrt{h}}\sum_{i=1}^{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{s=1,\neq t}^{T}K\left(\frac{\tau_t-\tau_s}{h}\right)\varepsilon_{it}\varepsilon_{ns}\xrightarrow{D}\mathcal{N}\left(0,\sigma_L^2\right).$$

Appendix D Proofs of the main lemmas

D.1 Proofs of Lemmas C.1-C.7

Proof of Lemma C.1

(1) Recall that $s_i(\tau) = [I_p, 0_p][M_i(\tau)^\top W(\tau) M_i(\tau)]^{-1} M_i(\tau)^\top W(\tau)$, $\Omega^{\mathcal{D}}_{x,i}(\tau) = diag(1,\mu_2)\otimes \Sigma^{\mathcal{D}}_{x,i}(\tau)$, $\Omega^{\mathcal{D}}_{N,i}(\tau) = (\mu_2, 0)^\top \otimes \Sigma^{\mathcal{D}}_{x,i}(\tau)$, $\widehat{\Omega}_{x,i}(\tau) = (Th)^{-1} M_i(\tau)^\top W(\tau) M_i(\tau)$ and $\widehat{\Omega}_{N,i}(\tau) = (Th)^{-1} M_i(\tau)^\top W(\tau) N_i(\tau)$, where $\Sigma^{\mathcal{D}}_{x,i}(\tau_t) = \mathbb{E}_{\mathcal{D}}(x_{it}x_{it}^\top)$. For given $\tau \in (0,1)$, we write

$$
\begin{aligned}
s_i(\tau)N_i(\tau) &= [I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right)\left(\widehat{\Omega}_{N,i}(\tau) - \Omega^{\mathcal{D}}_{N,i}(\tau)\right) + [I_p,0_p]\,\Omega^{\mathcal{D}-1}_{x,i}(\tau)\left(\widehat{\Omega}_{N,i}(\tau) - \Omega^{\mathcal{D}}_{N,i}(\tau)\right)\\
&\quad + [I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right)\Omega^{\mathcal{D}}_{N,i}(\tau) + [I_p,0_p]\,\Omega^{\mathcal{D}-1}_{x,i}(\tau)\Omega^{\mathcal{D}}_{N,i}(\tau)\\
&:= A_{NT,3,i}(\tau) + \cdots + A_{NT,6,i}(\tau). \qquad \text{(D.1)}
\end{aligned}
$$

We now study the convergence of $A_{NT,3,i}(\tau), \cdots, A_{NT,6,i}(\tau)$. For $A_{NT,3,i}(\tau)$, Lemma D.6(1)(2) imply

$$
\|A_{NT,3,i}(\tau)\|_F \le \left\|\widehat{\Omega}^{-1}_{x,i}\right\|_F \left\|\widehat{\Omega}_{N,i}(\tau) - \Omega^{\mathcal{D}}_{N,i}(\tau)\right\|_F \left\|\widehat{\Omega}_{x,i}(\tau) - \Omega^{\mathcal{D}}_{x,i}(\tau)\right\|_F = O_P\left(\frac{1}{Th}\right). \qquad \text{(D.2)}
$$

Analogously, for $A_{NT,4,i}(\tau)$ and $A_{NT,5,i}(\tau)$, directly using Lemma D.6(1)(2) gives

$$
\|A_{NT,4,i}(\tau)\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \quad \|A_{NT,5,i}(\tau)\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right). \qquad \text{(D.3)}
$$

For $A_{NT,6,i}(\tau)$, simple algebra yields

$$
A_{NT,6,i}(\tau) = \frac{\mu_2}{2} I_p, \qquad \text{(D.4)}
$$

where $\mu_2 = \int u^2 K(u)du$. Combining (D.2), (D.3) and (D.4), we obtain the desired result.

(2) We write

$$
\begin{aligned}
s_i(\tau)\widetilde{\varepsilon}_i &= \frac{1}{Th}[I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right) M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i + \frac{1}{Th}[I_p,0_p]\,\Omega^{\mathcal{D}-1}_{x,i}(\tau) M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i\\
&:= A_{NT,7,i}(\tau) + A_{NT,8,i}(\tau). \qquad \text{(D.5)}
\end{aligned}
$$

It suffices only to show $\|A_{NT,7,i}(\tau)\|_F = O_P\left(\frac{1}{Th}\right)$. Using Lemma D.6(1) and the result $\left\|M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i\right\|_F = O_P\left(\sqrt{Th}\right)$ which is implied by Lemma D.3(1), we obtain

$$
\|A_{NT,7,i}(\tau)\|_F \le C\left\|\widehat{\Omega}^{-1}_{x,i}(\tau)\right\|_F \left\|\widehat{\Omega}_{x,i}(\tau) - \Omega^{\mathcal{D}}_{x,i}(\tau)\right\|_F \left\|\frac{1}{Th} M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i\right\|_F = O_P\left(\frac{1}{Th}\right).
$$

(3) Recall that $R_{Sf,i}(\tau) = s_i(\tau)F^0 - \Sigma^{\mathcal{D}-1}_{x,i}(\tau)\Sigma^{\mathcal{D}}_{xf,i}(\tau)$ and $R_{xf,it} = x_{it}f_t^{0\top} - \Sigma^{\mathcal{D}}_{xf,i}(\tau_t)$. Let $\Omega^{\mathcal{D}}_{xf,i}(\tau) = \left(\Sigma^{\mathcal{D}\top}_{xf,i}(\tau), 0^\top\right)^\top$ and $\widehat{\Omega}_{xf,i}(\tau) = (Th)^{-1}M_i^\top(\tau)W(\tau)F^0$. We obtain the following decomposition:

$$
\begin{aligned}
R_{Sf,i}(\tau) &= [I_p,0_p]\,\widehat{\Omega}^{-1}_{x,i}(\tau)\widehat{\Omega}_{xf,i}(\tau) - \Sigma^{\mathcal{D}-1}_{x,i}(\tau)\Sigma^{\mathcal{D}}_{xf,i}(\tau)\\
&= [I_p,0_p]\,\Omega^{\mathcal{D}-1}_{x,i}(\tau)\left(\widehat{\Omega}_{xf,i}(\tau) - \Omega^{\mathcal{D}}_{xf,i}(\tau)\right) + [I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right)\widehat{\Omega}_{xf,i}(\tau)\\
&= [I_p,0_p]\,\Omega^{\mathcal{D}-1}_{x,i}(\tau)\left(\widehat{\Omega}_{xf,i}(\tau) - \Omega^{\mathcal{D}}_{xf,i}(\tau)\right) + [I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right)\Omega^{\mathcal{D}}_{xf,i}(\tau)\\
&\quad + [I_p,0_p]\left(\widehat{\Omega}^{-1}_{x,i}(\tau) - \Omega^{\mathcal{D}-1}_{x,i}(\tau)\right)\left(\widehat{\Omega}_{xf,i}(\tau) - \Omega^{\mathcal{D}}_{xf,i}(\tau)\right)\\
&:= A_{NT,9,i}(\tau) + A_{NT,10,i}(\tau) + A_{NT,11,i}(\tau). \qquad \text{(D.6)}
\end{aligned}
$$

Lemma D.6(1)(3) yields

$$
\|A_{NT,9,i}(\tau)\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \quad \|A_{NT,10,i}(\tau)\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \quad \|A_{NT,11,i}(\tau)\|_F = O_P\left(\frac{1}{Th}\right).
$$

It follows immediately that $\|R_{Sf,i}(\tau)\| = O_P\left(\frac{1}{\sqrt{Th}}\right)$.

(4) Write

$$\begin{aligned} s_i(\tau)R_f^{(n-1)} &= \frac{1}{Th}[I_p, 0_p]\,\Omega_{x,i}^{\mathcal{D}-1}(\tau)M_i(\tau)^\top W(\tau)R_f^{(n-1)} \\ &\quad + \frac{1}{Th}[I_p, 0_p]\left(\widehat{\Omega}_{x,i}^{-1}(\tau) - \Omega_{x,i}^{\mathcal{D}-1}(\tau)\right)M_i(\tau)^\top W(\tau)R_f^{(n-1)} \\ &:= A_{NT,12,i}(\tau) + A_{NT,13,i}(\tau). \end{aligned} \tag{D.7}$$

Simple algebra yields $A_{NT,12,i}(\tau) = \frac{1}{Th}\sum_{t=1}^T K_{t,0}(\tau)\Sigma_{x,i}^{\mathcal{D}-1}(\tau)x_{it}R_{f,t}^{(n-1)\top}$. By Cauchy-Schwarz inequality,

$$\|A_{NT,12,i}(\tau)\|_F \le \frac{1}{Th}\left(\sum_{t=1}^T K_{t,0}(\tau)\|x_{it}\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T K_{t,0}(\tau)\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}} = O_P(\delta_{f,n-1}). \tag{D.8}$$

Using Lemma D.6(1) and Cauchy-Schwarz inequality, we obtain $\|A_{NT,13,i}(\tau)\|_F = O_P\left(\frac{\delta_{f,n-1}}{\sqrt{Th}}\right)$. Together with (D.8), it leads to the desired result in Lemma C.1(4). ■

Proof of Lemma C.2

(1) Recall that $R_{Sf,i}(\tau_t) = A_{NT,9,i}(\tau_t) + A_{NT,10,i}(\tau_t) + A_{NT,11,i}(\tau_t)$, where $A_{NT,9,i}(\tau)$, $A_{NT,10,i}(\tau)$ and $A_{NT,11,i}(\tau)$ are defined in (D.6). For $A_{NT,9,i}(\tau_t)$, by Lemma D.7(2),

$$\sum_{t=1}^T \|A_{NT,9,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right). \tag{D.9}$$

To compute the probability order of $\sum_{t=1}^T \|A_{NT,10,i}(\tau_t)\|_F^2$, we decompose the inverse matrix $\widehat{\Omega}_{x,i}^{-1}(\tau)$ in the following way: for a positive and sufficiently large number M,

$$\widehat{\Omega}_{x,i}^{-1}(\tau) = \Omega_{x,i}^{\mathcal{D}-1}(\tau) + Q_{x,i}(\tau)\Omega_{x,i}^{\mathcal{D}-1}(\tau) = \Omega_{x,i}^{\mathcal{D}-1}(\tau) + \widetilde{Q}_{x,i}(\tau)\Omega_{x,i}^{\mathcal{D}-1}(\tau) + \widetilde{Q}_{x,i}^c(\tau)\Omega_{x,i}^{\mathcal{D}-1}(\tau),$$

where $Q_{x,i}(\tau) = \left(\widehat{\Omega}_{x,i}^{-1}(\tau) - \Omega_{x,i}^{\mathcal{D}-1}(\tau)\right)\Omega_{x,i}^{\mathcal{D}}(\tau)$, $\widetilde{Q}_{x,i}(\tau) = Q_{x,i}(\tau)I\left(\|Q_{x,i}(\tau)\|_F < M\right)$ and $\widetilde{Q}_{x,i}^c(\tau) = Q_{x,i}(\tau) - \widetilde{Q}_{x,i}(\tau)$. With $\widetilde{Q}_{x,i}(\tau)$ and $\widetilde{Q}_{x,i}^c(\tau)$, we write

$$\begin{aligned} A_{NT,10,i}(\tau_t) &= [I_p, 0_p]\,\widetilde{Q}_{x,i}(\tau_t)\Omega_{x,i}^{\mathcal{D}-1}(\tau_t)\Omega_{xf,i}^{\mathcal{D}}(\tau_t) + [I_p, 0_p]\,\widetilde{Q}_{x,i}^c(\tau_t)\Omega_{x,i}^{\mathcal{D}-1}(\tau_t)\Omega_{xf,i}^{\mathcal{D}}(\tau_t) \\ &:= A_{NT,10,1,i}(\tau_t) + A_{NT,10,2,i}(\tau_t). \end{aligned} \tag{D.10}$$

For $A_{NT,10,1,i}(\tau_t)$, it follows by Lemma D.7 that

$$\sum_{t=1}^T \|A_{NT,10,1,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right). \tag{D.11}$$

For $A_{NT,10,2,i}(\tau_t)$ and any given $\varepsilon > 0$, the uniform convergence result in Lemma D.9(3) yields

$$P\left(\sum_{t=1}^T \|A_{NT,10,2,i}(\tau_t)\|_F^2 > \frac{\varepsilon}{h}\right) \le P\left(\sup_{0<\tau<1}\|Q_{x,i}(\tau)\|_F > M\right) = o(1).$$

Then, it follows that $\sum_{t=1}^T \|A_{NT,10,2,i}(\tau_t)\|_F^2 = o_P\left(h^{-1}\right)$. Together with (D.11), it implies

$$\sum_{t=1}^T \|A_{NT,10,i}(\tau_t)\|_F^2 = \sum_{t=1}^T \|A_{NT,10,1,i}(\tau_t)\|_F^2 + \sum_{t=1}^T \|A_{NT,10,2,i}(\tau_t)\|_F^2 + \text{interaction terms} = O_P\left(\frac{1}{h}\right). \tag{D.12}$$

For $A_{NT,11,i}(\tau_t)$, we write

$$A_{NT,11,i} = [I_p, 0_p]\,\widetilde{Q}_{x,i}(\tau_t)\Omega_{x,i}^{\mathcal{D}-1}(\tau_t)\left(\widehat{\Omega}_{xf,i}(\tau_t) - \Omega_{xf,i}^{\mathcal{D}}(\tau_t)\right)$$

$$
\begin{aligned}
&+ [I_p, 0_p]\, \widetilde{Q}^c_{x,i}(\tau_t)\Omega^{\mathcal{D}-1}_{x,i}(\tau_t)\left(\widehat{\Omega}_{xf,i}(\tau_t) - \Omega^{\mathcal{D}}_{xf,i}(\tau_t)\right)\\
:= \;& A_{NT,11,1,i}(\tau_t) + A_{NT,11,2,i}(\tau_t).
\end{aligned}
$$

Since $\widetilde{Q}_{x,i}(\tau)$ is uniformly bounded, Lemma D.7(2) implies

$$
\sum_{t=1}^{T} \|A_{NT,11,1,i}(\tau_t)\|_F^2 \le C \sum_{t=1}^{T} \left\|\widehat{\Omega}_{xf,i}(\tau_t) - \Omega^{\mathcal{D}}_{xf,i}(\tau_t)\right\|_F^2 = O_P\left(\frac{1}{h}\right). \tag{D.13}
$$

Additionally, it holds by Lemma D.9(3) that $\sum_{t=1}^{T} \|A_{NT,11,1,i}(\tau_t)\|_F^2 = o_P(h^{-1})$. Together with (D.13), it implies

$$
\sum_{t=1}^{T} \|A_{NT,11,i}(\tau_t)\|_F^2 = \sum_{t=1}^{T} \|A_{NT,11,1,i}(\tau_t)\|_F^2 + \sum_{t=1}^{T} \|A_{NT,11,2,i}(\tau_t)\|_F^2 + \text{interaction terms} = O_P\left(\frac{1}{h}\right). \tag{D.14}
$$

Combining (D.9), (D.12), and (D.14) gives $\sum_{t=1}^{T} \|R_{Sf,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right)$. Analogously, we can show $\sum_{i=1}^{N} \|R_{Sf,i}(\tau)\|_F^2 = O_P\left(\frac{N}{Th}\right)$ and $\sum_{i=1}^{N} \|R_{Sf,i}(\tau_t)\|_F^2 = O_P\left(\frac{N}{h}\right)$. Write

$$
\lambda_i^0 x_{it}^\top R_{Sf,i}(\tau_t) = \lambda_i^0 x_{it}^\top \left(A_{NT,9,i}(\tau_t) + A_{NT,10,i}(\tau_t) + A_{NT,11,i}(\tau_t)\right),
$$

where $A_{NT,9,i}(\tau)$, $A_{NT,10,i}(\tau)$, and $A_{NT,11,i}(\tau)$ are defined in (D.6). By Lemma D.1 and Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}}\left[\left\|\lambda_i^0 x_{it}^\top A_{NT,9,i}(\tau_t)\right\|_F^2\right]\right]\\
&\le \mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbb{E}_{\mathcal{D}}\left[\left\|\widehat{\Omega}_{xf,i}(\tau_t) - \Omega^{\mathcal{D}}_{xf,i}(\tau_t)\right\|_F^{\frac{2+\delta^*}{2}}\right]\right)^{\frac{4}{4+\delta^*}}\left(\mathbb{E}_{\mathcal{D}}\left[\left\|x_{it}\lambda_i^{0\top}\right\|_F^{\frac{8+2\delta^*}{\delta^*}}\right]\right)^{\frac{\delta^*}{4+\delta^*}}\right]\\
&= O\left(\frac{N}{h}\right).
\end{aligned}
$$

For $A_{NT,10,i}(\tau_t)$, $\lambda_i^0 x_{it}^\top A_{NT,10,i}(\tau_t) = \lambda_i^0 x_{it}^\top A_{NT,10,1,i}(\tau_t) + \lambda_i^0 x_{it}^\top A_{NT,10,2,i}(\tau_t)$, where $A_{NT,10,1,i}(\tau)$ and $A_{NT,10,2,i}(\tau_t)$ are defined in (D.10). Using Lemma D.1 and Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}}\left[\left\|\lambda_i^0 x_{it}^\top A_{NT,10,1,i}(\tau_t)\right\|_F^2\right]\right]\\
&\le \mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbb{E}_{\mathcal{D}}\left[\left\|\widetilde{Q}_{x,i}(\tau_t)\right\|_F^{\frac{2+\delta^*}{2}}\right]\right)^{\frac{4}{4+\delta^*}}\left(\mathbb{E}_{\mathcal{D}}\left[\left\|x_{it}\lambda_i^{0\top}\right\|_F^{\frac{8+2\delta^*}{\delta^*}}\right]\right)^{\frac{\delta^*}{4+\delta^*}}\right]\\
&= O\left(\frac{N}{h}\right).
\end{aligned}
$$

By the uniform convergence result in Lemma D.9(2), we can show that the term with $A_{NT,10,2,i}(\tau_t)$ is negligible. Therefore, $\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{E}\left[\left\|\lambda_i^0 x_{it}^\top A_{NT,10,i}(\tau_t)\right\|_F^2\right] = O\left(\frac{N}{h}\right)$. Using analogous arguments, we can obtain $\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{E}\left[\left\|\lambda_i^0 x_{it}^\top A_{NT,11,i}(\tau_t)\right\|_F^2\right] = O\left(\frac{N}{h}\right)$. Therefore, $\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{E}\left[\left\|\lambda_i^0 x_{it}^\top R_{Sf,i}(\tau_t)\right\|_F^2\right] = O\left(\frac{N}{h}\right)$.

(2) Recall that $R_{S,i}(\tau_t) = s_i(\tau_t)N_i(\tau_t) - \mu_2 I_p = A_{NT,3,i}(\tau_t) + \cdots + A_{NT,5,i}(\tau_t)$, where $A_{NT,3,i}(\tau_t), \cdots, A_{NT,5,i}(\tau_t)$ are defined in (D.1). Using arguments that are analogous to those in the proof of (D.9) and (D.12), we can use Lemmas D.6 and D.9 to show

$$
\sum_{t=1}^{T} \|A_{NT,3,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right), \quad \sum_{t=1}^{T} \|A_{NT,4,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right), \quad \sum_{t=1}^{T} \|A_{NT,5,i}(\tau_t)\|_F^2 = O_P\left(\frac{1}{h}\right).
$$

Therefore,

$$\sum_{t=1}^{T}\|R_{S,i}(\tau_t)\|_F^2=\sum_{t=1}^{T}\|A_{NT,3,i}(\tau_t)\|_F^2+\sum_{t=1}^{T}\|A_{NT,4,i}(\tau_t)\|_F^2+\sum_{t=1}^{T}\|A_{NT,5,i}(\tau_t)\|_F^2+\text{interaction terms}=O_P\left(\frac{1}{h}\right).$$

Analogously, we also have

$$\sum_{i=1}^{N}\|R_{S,i}(\tau)\|_F^2=O_P\left(\frac{N}{Th}\right),\ \sum_{i=1}^{N}\sum_{t=1}^{T}\|R_{S,i}(\tau_t)\|_F^2=O_P\left(\frac{N}{h}\right),\ \sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,0}(\tau)\|R_{S,i}(\tau_t)\|_F^2=O_P\left(N\right).$$

(3) Recall that $s_i(\tau)\widetilde{\varepsilon}_i=A_{NT,7,i}(\tau)+A_{NT,8,i}(\tau)$, where $A_{NT,7,i}(\tau)$ and $A_{NT,8,i}(\tau)$ are defined in (D.5). By Lemma D.4(1), we have $\sum_{t=1}^{T}\|A_{NT,8,i}(\tau_t)\|_F^2=O_P\left(h^{-1}\right)$. For $A_{NT,7,i}(\tau)$, we write

$$\begin{aligned}A_{NT,7,i}(\tau)&=\frac{1}{Th}\left[I_p,0_p\right]\widetilde{Q}_{x,i}(\tau)\Omega_{x,i}^{\mathcal{D}}(\tau_t)M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i+\frac{1}{Th}\left[I_p,0_p\right]\widetilde{Q}_{x,i}^c(\tau)\Omega_{x,i}^{\mathcal{D}}(\tau_t)M_i(\tau)^\top W(\tau)\widetilde{\varepsilon}_i\\&:=A_{NT,7,1,i}(\tau)+A_{NT,7,2,i}(\tau).\end{aligned}$$

Since $\widetilde{Q}_{x,i}(\tau)$ is uniformly bounded, $\sum_{t=1}^{T}\|A_{NT,7,1,i}(\tau_t)\|_F^2\le C\sum_{t=1}^{T}\|A_{NT,8,i}(\tau_t)\|_F^2=O_P\left(h^{-1}\right)$. For any given positive number ε and $A_{NT,7,2,i}(\tau)$, by Lemma D.9(2),

$$P\left(\sum_{t=1}^{T}\|A_{NT,7,2,i}(\tau_t)\|_F^2>\frac{\varepsilon}{h}\right)\le P\left(\sup_{0<\tau<1}\|Q_{x,i}(\tau)\|_F>M\right)=o(1),$$

which yields $\sum_{t=1}^{T}\|A_{NT,7,2,i}(\tau_t)\|_F^2=o_P(h^{-1})$. Therefore, $\sum_{t=1}^{T}\|s_i(\tau_t)\widetilde{\varepsilon}_i\|_F^2=O_P\left(h^{-1}\right)$. Analogously, we can show $\sum_{i=1}^{N}\sum_{t=1}^{T}\|A_{NT,7,i}(\tau_t)\|_F^2=o_P(Nh^{-1})$ and $\sum_{i=1}^{N}\sum_{t=1}^{T}\|A_{NT,8,i}(\tau_t)\|_F^2=o_P(Nh^{-1})$. Therefore, $\sum_{i=1}^{N}\sum_{t=1}^{T}\|s_i(\tau_t)\widetilde{\varepsilon}_i\|_F^2=O_P\left(Nh^{-1}\right)$ and $\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\lambda_i^0x_{it}^\top s_i(\tau_t)\widetilde{\varepsilon}_i\right\|^2=O_P\left(Nh^{-1}\right)$.

(4) Recall that $s_i(\tau)R_f^{(n-1)}=A_{NT,12,i}(\tau_t)+A_{NT,13,i}(\tau_t)$, where $A_{NT,12,i}(\tau)$ and $A_{NT,13,i}(\tau)$ are defined in (D.7). Similarly to (D.8), we can show $\sum_{t=1}^{T}\|A_{NT,12,i}(\tau_t)\|_F^2=O_P\left(T\delta_{f,n-1}^2\right)$. For $A_{NT,13,i}(\tau)$,

$$\begin{aligned}A_{NT,13,i}(\tau)&=\frac{1}{Th}\left[I_p,0_p\right]\widetilde{Q}_{x,i}(\tau)\Omega_{x,i}^{\mathcal{D}}(\tau_t)M_i(\tau)^\top W(\tau)R_f^{(n-1)}\\&\quad+\frac{1}{Th}\left[I_p,0_p\right]\widetilde{Q}_{x,i}^c(\tau)\Omega_{x,i}^{\mathcal{D}}(\tau_t)M_i(\tau)^\top W(\tau)R_f^{(n-1)}\\&:=A_{NT,13,1,i}(\tau)+A_{NT,13,2,i}(\tau).\end{aligned}$$

Since $\widetilde{Q}_{x,i}(\tau)$ is bounded, we have $\sum_{t=1}^{T}\|A_{NT,13,1,i}(\tau_t)\|_F^2\le C\sum_{t=1}^{T}\|A_{NT,12,i}(\tau_t)\|_F^2=O_P\left(T\delta_{f,n-1}^2\right)$. For $A_{NT,13,2,i}(\tau)$, by Lemma D.9(2), we have $\sum_{t=1}^{T}\|A_{NT,13,2,i}(\tau_t)\|_F^2=o_P(T\delta_{f,n-1}^2)$. Consequently, $\sum_{t=1}^{T}\left\|s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(T\delta_{f,n-1}^2\right)$. Analogously, we have

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(NT\delta_{f,n-1}^2\right),\ \sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\lambda_i^0x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right\|_F^2=O_P\left(NT\delta_{f,n-1}^2\right).$$

Therefore, Lemma C.2(4) holds. ■

Proof of Lemma C.3

Recall that $\widehat{\Omega}_{f,i}^{(n-1)}=T^{-1}\widehat{F}^{(n-1)\top}\Omega_{S,i}\widehat{F}^{(n-1)}$, $\Omega_{f,i}^{H}=H^\top\Omega_{f,i}^{\mathcal{D}}H$ with $\Omega_{S,i}=(I-S_i)^\top(I-S_i)$ and $\Omega_{f,i}^{\mathcal{D}}=\Sigma_f^{\mathcal{D}}-\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$. We write

$$\begin{aligned}\widehat{\Omega}_{f,i}^{(n-1)}-\Omega_{f,i}^{H}&=H^\top\left(\frac{1}{T}F^{0\top}\Omega_{S,i}F^0-\Omega_{f,i}^{\mathcal{D}}\right)H+\frac{1}{T}H^\top F^{0\top}\Omega_{S,i}R_f^{(n-1)}+\frac{1}{T}R_f^{(n-1)\top}\Omega_{S,i}\widetilde{F}\\&\quad+\frac{1}{T}R_f^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\end{aligned}$$

$$:= \quad A_{NT,14,i} + \cdots + A_{NT,17,i}.$$

Since $A_{NT,16,i} = A_{NT,15,i}^\top$, it suffices only to show the convergence of $A_{NT,14,i}$, $A_{NT,15,i}$ and $A_{NT,17,i}$. Recall that $R_{x,it} = x_{it}x_{it}^\top - \Sigma_{x,i}^{\mathcal{D}}(\tau_t)$, $R_{xf,it} = x_{it}f_t^{0\top} - \Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$ and $R_{Sf,i}(\tau_t) = s_i(\tau_t)F^0 - \Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$. For $A_{NT,14,i}$, write

$$\begin{aligned}
H^{-1\top}A_{NT,14,i}H^{-1} &= \frac{1}{T}\sum_{t=1}^{T}\left(f_t^0 f_t^{0\top} - \mathbb{E}_{\mathcal{D}}\left[f_t^0 f_t^{0\top}\right]\right) - \frac{1}{T}\sum_{t=1}^{T} R_{xf,it}^\top \Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,\Sigma_{xf,i}^{\mathcal{D}}(\tau_t) \\
&- \frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,R_{xf,it} + \frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,R_{x,it}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,\Sigma_{xf,i}^{\mathcal{D}}(\tau_t) \\
&- \frac{1}{T}\sum_{t=1}^{T} R_{xf,it}^\top R_{Sf,i}(\tau_t) - \frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top R_{xf,it} + \frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top R_{x,it} R_{Sf,i}(\tau_t) \\
&+ \frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,R_{x,it}R_{Sf,i}(\tau_t) + \frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top R_{x,it}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,\Sigma_{xf,i}^{\mathcal{D}}(\tau_t) \\
&+ \frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top \Sigma_{x,i}^{\mathcal{D}}(\tau_t) R_{Sf,i}(\tau_t) + O_P\left(\frac{1}{T}\right) \\
&:= A_{NT,14,1,i} + \cdots + A_{NT,14,10,i} + O_P\left(\frac{1}{T}\right),
\end{aligned} \tag{D.15}$$

where we use the following property of Riemann integral:

$$\frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\,\Sigma_{xf,i}^{\mathcal{D}}(\tau_t) - \int \Sigma_{xf,i}^{\mathcal{D}\top}(v)\Sigma_{x,i}^{\mathcal{D}-1}(v)\,\Sigma_{xf,i}^{\mathcal{D}}(v)dv = O_P\left(\frac{1}{T}\right).$$

Directly using Lemma D.2(1), we can readily obtain

$$\begin{aligned}
\|A_{NT,14,1,i}\|_F = O_P\left(\frac{1}{\sqrt{T}}\right), \quad \|A_{NT,14,2,i}\|_F = O_P\left(\frac{1}{\sqrt{T}}\right), \\
\|A_{NT,14,3,i}\|_F = O_P\left(\frac{1}{\sqrt{T}}\right), \quad \|A_{NT,14,4,i}\|_F = O_P\left(\frac{1}{\sqrt{T}}\right).
\end{aligned} \tag{D.16}$$

For $A_{NT,14,5,i}$, by Lemma C.2(1) and Cauchy-Schwarz inequality,

$$\|A_{NT,14,5,i}\|_F \leq \left(\sum_{t=1}^{T}\|R_{x,it}\|_F^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\|R_{Sf,i}(\tau_t)\|_F^2\right)^{\frac{1}{2}} = O_P\left(\frac{1}{\sqrt{Th}}\right). \tag{D.17}$$

Analogously, Lemma C.2(1) and Cauchy-Schwarz inequality also imply the following results:

$$\begin{aligned}
&\|A_{NT,14,6,i}\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \quad \|A_{NT,14,7,i}\|_F = O_P\left(\frac{1}{Th}\right), \quad \|A_{NT,14,8,i}\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \\
&\|A_{NT,14,9,i}\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right), \quad \|A_{NT,14,10,i}\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right).
\end{aligned}$$

Together with (D.15), (D.16), and (D.17), these results yield

$$\|A_{NT,14,i}\|_F = O_P\left(\frac{1}{\sqrt{Th}}\right). \tag{D.18}$$

We now study the convergence of $A_{NT,15,i}$. Write

$$H^{-1\top}A_{NT,15,i} \quad = \quad \frac{1}{T}\sum_{t=1}^{T} f_t^0 R_{f,t}^{(n-1)\top} - \frac{1}{T}\sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}R_{f,t}^{(n-1)\top}$$

$$
\begin{aligned}
& -\frac{1}{T}\sum_{t=1}^{T} R_{xf,it}^{\top} s_i(\tau_t) R_f^{(n-1)} - \frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^{\top} x_{it} R_{f,t}^{(n-1)\top} \\
& +\frac{1}{T}\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^{\top} x_{it} x_{it}^{\top} s_i(\tau_t) R_f^{(n-1)} + \frac{1}{T}\sum_{t=1}^{T} \Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) R_{x,it} s_i(\tau_t) R_f^{(n-1)} \\
:= \; & A_{NT,15,1,i} + \cdots + A_{NT,15,6,i}. \qquad \text{(D.19)}
\end{aligned}
$$

For $A_{NT,15,1,i}$, by Cauchy-Schwarz inequality,

$$
\|A_{NT,15,1,i}\|_F \leq \frac{1}{T}\left(\sum_{t=1}^{T}\left\|f_t^0\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}} = O_P\left(\delta_{f,n-1}\right).
$$

Analogously, we have $\|A_{NT,15,2,i}\|_F = O_P\left(\delta_{f,n-1}\right)$. For $A_{NT,15,3,i}$, using Lemma C.2(4) and Cauchy-Schwarz inequality gives

$$
\|A_{NT,15,3,i}\|_F \leq \frac{1}{T}\left(\sum_{t=1}^{T}\|R_{xf,it}\|_F^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\left\|s_i(\tau_t)R_f^{(n-1)}\right\|_F\right)^{\frac{1}{2}} = O_P\left(\delta_{f,n-1}\right).
$$

Analogously, we can obtain $\|A_{NT,15,4,i}\|_F = O_P\left(\frac{\delta_{f,n-1}}{\sqrt{Th}}\right)$, $\|A_{NT,15,5,i}\|_F = O_P\left(\frac{\delta_{f,n-1}}{\sqrt{Th}}\right)$, $\|A_{NT,15,6,i}\|_F = O_P\left(\delta_{f,n-1}\right)$. Therefore, the convergence rate of $A_{NT,15,i}$ is

$$
\|A_{NT,15,i}\| = O_P\left(\delta_{f,n-1}\right). \qquad \text{(D.20)}
$$

For $A_{NT,17,i}$, write

$$
\begin{aligned}
A_{NT,17,i} \;=\; & \frac{1}{T}\sum_{t=1}^{T} R_{f,t}^{(n-1)} R_{f,t}^{(n-1)\top} - \frac{1}{T}\sum_{t=1}^{T} R_{f,t}^{(n-1)} x_{it}^{\top} s_i(\tau_t) R_f^{(n-1)} - \frac{1}{T}\sum_{t=1}^{T}\left(s_i(\tau_t)R_f^{(n-1)}\right)^{\top} x_{it} R_{f,t}^{(n-1)\top} \\
& +\frac{1}{T}\sum_{t=1}^{T}\left(s_i(\tau_t)R_f^{(n-1)}\right)^{\top} x_{it}x_{it}^{\top} s_i(\tau_t) R_f^{(n-1)} \\
:= \; & A_{NT,17,1,i} + \cdots + A_{NT,17,4,i}.
\end{aligned}
$$

We now study these terms one by one. For $A_{NT,17,1,i}$, it is clear to see that

$$
\|A_{NT,17,1,i}\|_F \leq \frac{1}{T}\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2 = O_P\left(\delta_{f,n-1}^2\right).
$$

For $A_{NT,17,2,i}$, we write

$$
\begin{aligned}
A_{NT,17,2,i} \;=\; & \frac{1}{T^2 h}\sum_{t=1}^{T}\sum_{s=1}^{T} K_{s,0}(\tau_t) R_{f,t}^{(n-1)} x_{it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{is} R_{f,s}^{(n-1)\top} \\
& +\frac{1}{T^2 h}\sum_{t=1}^{T}\sum_{s=1}^{T} R_{f,t}^{(n-1)} x_{it}^{\top}\widetilde{Q}_{x,i}(\tau_t) K_s(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{is} R_{f,s}^{(n-1)\top} \\
& +\frac{1}{T^2 h}\sum_{t=1}^{T}\sum_{s=1}^{T} R_{f,t}^{(n-1)} x_{it}^{\top}\widetilde{Q}_{x,i}^{c}(\tau_t) K_s(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{is} R_{f,s}^{(n-1)\top} \\
:= \; & A_{NT,17,2,1,i} + A_{NT,17,2,2,i} + A_{NT,17,2,3,i}.
\end{aligned}
$$

By Cauchy-Schwarz inequality,

$$
\|A_{NT,17,2,1,i}\|_F \;\leq\; \frac{1}{T^2 h}\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\left\|R_{f,s}^{(n-1)\top}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|K_{s,0}(\tau_t)x_{it}^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}\right\|^2\right)^{\frac{1}{2}}
$$

$$= \quad O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right). \tag{D.21}$$

Since $\widetilde{Q}_{x,i}(\tau)$ is uniformly bounded,

$$\|A_{NT,17,2,2,i}\|_F \leq C\,\|A_{NT,17,2,1,i}\|_F = O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right). \tag{D.22}$$

For $A_{NT,17,2,3,i}$, by Lemma D.9(2), we obtain $\|A_{NT,17,2,3,i}\|_F = o_P\left(\delta_{f,n-1}^2 h^{-1/2}\right)$. Together with (D.21) and (D.22), it gives

$$\|A_{NT,17,2,i}\|_F = O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right). \tag{D.23}$$

Analogously, we have $\|A_{NT,17,3,i}\|_F = O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right)$ and $\|A_{NT,17,4,i}\|_F = O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right)$. Therefore,

$$\|A_{NT,17,i}\|_F = O_P\left(\frac{\delta_{f,n-1}^2}{\sqrt{h}}\right). \tag{D.24}$$

Since $A_{NT,16,i} = A_{NT,15,i}^\top$, (D.18), (D.20), and (D.24) jointly lead to the desired result in Lemma C.3. ■

Proof of Lemma C.4

(1) Recall that $R_{xf,it} = x_{it}f_t^{0\top} - \Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$, $R_{x,it} = x_{it}x_{it}^\top - \Sigma_{x,i}^{\mathcal{D}}(\tau_t)$, $R_{S,i}(\tau_t) = s_i(\tau_t)N_i(\tau_t) - \mu_2 I$, $R_{Sf,i}(\tau_t) = s_i(\tau_t)F^0 - \Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\Sigma_{xf,i}^{\mathcal{D}}(\tau_t)$. Write

$$\begin{aligned}
&F^{0\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top \\
&= \sum_{t=1}^T \left(x_{it}f_t^{0\top} - x_{it}x_{it}^\top s_i(\tau_t)F^0\right)^\top \left(\beta_{it} - s_i(\tau_t)\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top\right) \\
&= -\sum_{t=1}^T \left(x_{it}f_t^{0\top} - x_{it}x_{it}^\top s_i(\tau_t)F^0\right)^\top s_i(\tau_t)N_i(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&= -\sum_{t=1}^T R_{xf,it}^\top\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) + \sum_{t=1}^T \Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)R_{x,it}\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&+ \sum_{t=1}^T R_{Sf,i}(\tau_t)^\top\Sigma_{x,i}^{\mathcal{D}}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) - \sum_{t=1}^T R_{xf,it}^\top R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&+ \sum_{t=1}^T R_{Sf,i}(\tau_t)^\top R_{x,it}\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) + \sum_{t=1}^T \Sigma_{xf,i}^{\mathcal{D}}(\tau_t)^\top\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)R_{x,it}R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&+ \sum_{t=1}^T R_{Sf,i}(\tau_t)^\top\Sigma_{x,i}^{\mathcal{D}}(\tau_t)R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) + \sum_{t=1}^T R_{Sf,i}(\tau_t)^\top R_{x,it}R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&:= A_{NT,18,i} + \cdots + A_{NT,25,i}.
\end{aligned} \tag{D.25}$$

We now study the convergence of $A_{NT,18,i},\cdots,A_{NT,25,i}$ one by one. Lemmas D.2(1) and C.2(1) imply

$$\|A_{NT,18,i}\|_F = O_P\left(\sqrt{Th^4}\right),\ \|A_{NT,19,i}\|_F = O_P\left(\sqrt{Th^4}\right),\ \|A_{NT,20,i}\|_F = O_P\left(\sqrt{Th^3}\right). \tag{D.26}$$

For $A_{NT,21,i},\cdots,A_{NT,25,i}$, using Lemmas C.2(2), D.2(1) and Cauchy-Schwarz inequality, we obtain

$$\|A_{NT,21,i}\|_F \leq Ch^2\left(\sum_{t=1}^T \|R_{xf,it}\|_F^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T \|R_{S,i}(\tau_t)\|_F^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{Th^3}\right). \tag{D.27}$$

Analogously, we have $\|A_{NT,22,i}\|_F = O_P\left(\sqrt{Th^3}\right)$, $\|A_{NT,23,i}\|_F = O_P\left(\sqrt{Th^3}\right)$, $\|A_{NT,24,i}\|_F = O_P(h)$ and $\|A_{NT,25,i}\|_F = O_P(h)$. Together with (D.26) and (D.27), these results imply

$$\left\|F^{0\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top\right\|_F = O_P\left(\sqrt{Th^3}\right). \tag{D.28}$$

For $R_f^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top$, it is clear to see that

$$\begin{aligned}
& R_f^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top \\
&= -\sum_{t=1}^T\left(R_{f,t}^{(n-1)\top} - x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right)^\top x_{it}^\top R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&-\mu_2\sum_{t=1}^T\left(R_{f,t}^{(n-1)\top} - x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right)^\top x_{it}^\top\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&= -\mu_2\sum_{t=1}^T R_{f,t}^{(n-1)}x_{it}^\top\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) - \sum_{t=1}^T R_{f,t}^{(n-1)}x_{it}^\top R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&-\sum_{t=1}^T R_f^{(n-1)\top}s_i^\top(\tau_t)x_{it}x_{it}^\top R_{S,i}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) - \mu_2\sum_{t=1}^T R_f^{(n-1)\top}s_i^\top(\tau_t)R_{x,it}\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&-\mu_2\sum_{t=1}^T R_f^{(n-1)\top}s_i^\top(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)\left(\frac{1}{2}\beta_i''(\tau_t)h^2 + o(h^2)\right) \\
&:= A_{NT,26,i} + \cdots + A_{NT,30,i}.
\end{aligned} \tag{D.29}$$

By Lemma C.2(3) and Cauchy-Schwarz inequality,

$$\|A_{NT,26,i}\|_F \le Ch^2\left(\sum_{t=1}^T\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\|x_{it}\|^2\right)^{\frac{1}{2}} = O_P\left(Th^2\delta_{f,n-1}\right). \tag{D.30}$$

Analogously, we have $\|A_{NT,27,i}\|_F = O_P\left(\sqrt{Th^3}\delta_{f,n-1}\right)$, $\|A_{NT,28,i}\|_F = O_P\left(\sqrt{Th^3}\delta_{f,n-1}\right)$, $\|A_{NT,29,i}\|_F = O_P\left(Th^2\delta_{f,n-1}\right)$ and $\|A_{NT,30,i}\|_F = O_P\left(Th^2\delta_{f,n-1}\right)$. Together with (D.30), these results imply

$$\left\|R_f^{(n-1)\top}\Omega_{S,i}\left(\beta_{i1}^\top x_{i1},\cdots,\beta_{iT}^\top x_{iT}\right)^\top\right\|_F = O_P\left(\sqrt{Th^2\delta_{f,n-1}}\right). \tag{D.31}$$

By (D.28) and (D.31), Lemma C.4(1) holds.

(2) We write

$$\begin{aligned}
R_f^{(n-1)\top}\Omega_{S,i}R_f^{(n-1)}\widetilde{\lambda}_i &= \sum_{t=1}^T R_{f,t}^{(n-1)}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i - \sum_{t=1}^T R_{f,t}^{(n-1)}x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\widetilde{\lambda}_i \\
&-\sum_{t=1}^T\left(s_i(\tau_t)R_f^{(n-1)}\right)^\top x_{it}R_{f,t}^{(n-1)\top}\widetilde{\lambda}_i + \sum_{t=1}^T\left(s_i(\tau_t)R_f^{(n-1)}\right)^\top x_{it}x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\widetilde{\lambda}_i \\
&:= A_{NT,31,i} + \cdots + A_{NT,34,i}.
\end{aligned}$$

Simple algebra yields that $\|A_{NT,31,i}\|_F = O_P\left(T\delta_{f,n-1}^2\right)$. Using Lemma C.2(4) and Cauchy-Schwarz inequality, we obtain

$$\|A_{NT,32,i}\|_F \le \left(\sum_{t=1}^T\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\left\|x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right\|_F^2\right)^{\frac{1}{2}}\|\widetilde{\lambda}_i\| = O_P\left(T\delta_{f,n-1}^2\right).$$

Analogously, we have $\|A_{NT,33,i}\|_F = O_P\left(T\delta^2_{f,n-1}\right)$ and $\|A_{NT,34,i}\|_F = O_P\left(T\delta^2_{f,n-1}\right)$. Therefore, Lemma C.4(2) holds.

(3) Write

$$
\begin{aligned}
R_f^{(n-1)\top}\Omega_{S,i}\widetilde{\varepsilon}_i &= \sum_{t=1}^T\left(R_{f,t}^{(n-1)\top} - x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right)^\top\left(\varepsilon_{it} - x_{it}^\top s_i(\tau_t)\widetilde{\varepsilon}_i\right) \\
&= \sum_{t=1}^T R_{f,t}^{(n-1)}\varepsilon_{it} - \sum_{t=1}^T R_f^{(n-1)\top}s_i^\top(\tau_t)x_{it}\varepsilon_{it} - \sum_{t=1}^T R_{f,t}^{(n-1)}x_{it}^\top s_i(\tau_t)\varepsilon_i \\
&\quad + \sum_{t=1}^T R_f^{(n-1)\top}s_i^\top(\tau_t)x_{it}x_{it}^\top s_i(\tau_t)\varepsilon_i \\
&:= A_{NT,35,i} + \cdots + A_{NT,38,i}. \qquad \text{(D.32)}
\end{aligned}
$$

It suffices only to consider $A_{NT,36,i}$, $A_{NT,37,i}$ and $A_{NT,38,i}$. Let $K_t^*(\tau) = \left(K_{t,0}(\tau)I_p, \mu_2^{-1}K_{t,1}(\tau)I_p\right)^\top$. For $A_{NT,36,i}$, we write

$$
\begin{aligned}
A_{NT,36,i} &= -\frac{1}{Th}\sum_{t=1}^T\sum_{s=1}^T K_{s,0}(\tau_t)R_{f,s}^{(n-1)}x_{is}^\top\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}\varepsilon_{it} \\
&\quad - \sum_{t=1}^T\left([I_p, 0_p]\,\widetilde{Q}_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^T K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}R_{f,s}^{(n-1)\top}\right)\right)^\top x_{it}\varepsilon_{it} \\
&\quad - \sum_{t=1}^T\left([I_p, 0_p]\,\widetilde{Q}^c_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^T K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}R_{f,s}^{(n-1)\top}\right)\right)^\top x_{it}\varepsilon_{it} \\
&:= A_{NT,36,1,i} + A_{NT,36,2,i} + A_{NT,36,3,i}. \qquad \text{(D.33)}
\end{aligned}
$$

We now study these terms one by one. By Cauchy-Schwarz inequality and Lemma D.3(1),

$$
\|A_{NT,36,1,i}\| \leq \frac{1}{Th}\left(\sum_{s=1}^T\left\|R_{f,s}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{s=1}^T\left\|x_{is}^\top\right\|^2\left\|\sum_{t=1}^T K_{s,0}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}\varepsilon_{it}\right\|^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right).
$$

Since $\widetilde{Q}_{x,i}(\tau)$ is uniformly bounded, it follows that $\|A_{NT,36,2,i}\| \leq C\|A_{NT,36,1,i}\| = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. By Lemma D.9(2), $\|A_{NT,36,3,i}\| = o_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Therefore, $\|A_{NT,36,i}\| = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Using Lemma C.2(3) and Cauchy-Schwarz inequality gives

$$
\|A_{NT,37,i}\| \leq \left(\sum_{t=1}^T\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T\|x_{it}s_i(\tau_t)\widetilde{\varepsilon}_i\|^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right).
$$

Analogously, we can show $\|A_{NT,38,i}\| = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Therefore, Lemma C.4(3) holds.

(4) Write

$$
\begin{aligned}
F^{0\top}\Omega_{S,i}R_f^{(n-1)} &= \sum_{t=1}^T\left(f_t^{0\top} - x_{it}^\top s_i(\tau_t)F^0\right)^\top\left(R_{f,t}^{(n-1)\top} - x_{it}^\top s_i(\tau_t)R_f^{(n-1)}\right) \\
&= \sum_{t=1}^T f_t^0 R_{f,t}^{(n-1)\top} - \sum_{t=1}^T\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}R_{f,t}^{(n-1)\top} - \sum_{t=1}^T R_{Sf,i}(\tau_t)^\top x_{it}R_{f,t}^{(n-1)\top} \\
&\quad - \sum_{t=1}^T R_{xf,it}^\top s_i(\tau_t)R_f^{(n-1)} + \sum_{t=1}^T\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)R_{x,it}s_i(\tau_t)R_f^{(n-1)}
\end{aligned}
$$

$$
\begin{aligned}
& +\sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top \Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) s_i(\tau_t) R_f^{(n-1)} + \sum_{t=1}^{T} R_{Sf,i}(\tau_t)^\top R_{x,it} s_i(\tau_t) R_f^{(n-1)} \\
:= \; & A_{NT,39,i} + \cdots + A_{NT,45,i}. \qquad \text{(D.34)}
\end{aligned}
$$

It suffices only to study $A_{NT,41,i}$, $\cdots$, $A_{NT,45,i}$. Recall that $K_t^*(\tau) = \left(K_{t,0}(\tau)I_p, \mu_2^{-1}K_{t,1}(\tau)I_p\right)^\top$. For $A_{NT,41,i}$, we write

$$
\begin{aligned}
A_{NT,41,i} \;=\; & -\sum_{t=1}^{T}\left(\frac{1}{Th}\sum_{s=1}^{T} K_{s,0}(\tau_t) R_{xf,is}\right)^\top \Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{it} R_{f,t}^{(n-1)\top} \\
& -\sum_{t=1}^{T}\left([I_p, 0_p]\,\widetilde{Q}_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T} K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{is} f_s^{0\top}\right)\right)^\top x_{it} R_{f,t}^{(n-1)\top} \\
& -\sum_{t=1}^{T}\left([I_p, 0_p]\,\widetilde{Q}^c_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T} K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t) x_{is} f_s^{0\top}\right)\right)^\top x_{it} R_{f,t}^{(n-1)\top} + O_P\left(\frac{1}{h}\right) \\
:= \; & A_{NT,41,1,i} + A_{NT,41,2,i} + A_{NT,41,3,i} + O_P\left(\frac{1}{h}\right). \qquad \text{(D.35)}
\end{aligned}
$$

By Cauchy-Schwarz inequality and Lemma D.4(1),

$$
\|A_{NT,41,1,i}\|_F \le \frac{1}{Th}\left(\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\|x_{it}\|_F^2\left\|\sum_{s=1}^{T}K_{s,0}(\tau_t)R_{xf,is}\right\|_F^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right).
$$

For $A_{NT,41,2,i}$, using Cauchy-Schwarz inequality and Lemma D.7(3), we obtain

$$
\|A_{NT,41,2,i}\|_F \le C\left(\sum_{t=1}^{T}\left\|R_{f,t}^{(n-1)}\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\left\|\widetilde{Q}_{x,i}(\tau_t)\right\|_F^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right).
$$

Using Lemma D.9(2), we can show that $A_{NT,41,3,i}$ is negligible. Combining these results gives $\|A_{NT,41,i}\|_F = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. For $A_{NT,42,i}$, write

$$
\begin{aligned}
A_{NT,42,i} \;=\; & -\sum_{t=1}^{T} R_{xf,it}^\top\left(\frac{1}{Th}\sum_{s=1}^{T}K_{s,0}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}R_{f,s}^{(n-1)\top}\right) \\
& -\sum_{t=1}^{T} R_{xf,it}^\top\left([I_p, 0_p]\,\widetilde{Q}_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T}K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}R_{f,s}^{(n-1)\top}\right)\right) \\
& -\sum_{t=1}^{T} R_{xf,it}^\top\left([I_p, 0_p]\,\widetilde{Q}^c_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T}K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{is}R_{f,s}^{(n-1)\top}\right)\right) \\
:= \; & A_{NT,42,1,i} + A_{NT,42,2,i} + A_{NT,42,3,i} + O_P\left(\frac{1}{h}\right). \qquad \text{(D.36)}
\end{aligned}
$$

Simple algebra gives $\|A_{NT,42,1,i}\|_F = \|A_{NT,41,1,i}\|_F = O_P\left(\delta_{f,n-1}\sqrt{T/h}\right)$, $\|A_{NT,42,2,i}\|_F = \|A_{NT,41,2,i}\|_F = O_P\left(\delta_{f,n-1}\sqrt{T/h}\right)$, and $\|A_{NT,42,3,i}\|_F = \|A_{NT,41,3,i}\|_F = o_P\left(\delta_{f,n-1}\sqrt{T/h}\right)$. Therefore, $\|A_{NT,42,i}\|_F = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Analogously, we can show that $\|A_{NT,43,i}\|_F = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Using Cauchy-Schwarz inequality and Lemma C.2(4), we obtain

$$
\|A_{NT,44,i}\|_F \le \left(\sum_{t=1}^{T}\left\|s_i(\tau)R_f^{(n-1)}\right\|_F^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\|R_{Sf,i}(\tau_t)\|_F^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right).
$$

Analogously, $\|A_{NT,45,i}\|_F = O_P\left(\sqrt{\frac{T}{h}}\delta_{f,n-1}\right)$. Therefore, Lemma C.4(4) holds.

(5) Write

$$
\begin{aligned}
F^{0\top}\Omega_{S,i}\widetilde{\varepsilon}_i &= \sum_{t=1}^{T}\left(f_t^{0\top} - x_{it}^{\top}s_i(\tau_t)F^0\right)^{\top}\left(\varepsilon_{it} - x_{it}^{\top}s_i(\tau_t)\widetilde{\varepsilon}_i\right) \\
&= \sum_{t=1}^{T} f_t^0\varepsilon_{it} - \sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}\varepsilon_{it} - \sum_{t=1}^{T}R_{Sf,i}(\tau_t)^{\top}x_{it}\varepsilon_{it} - \sum_{t=1}^{T}R_{xf,it}^{\top}s_i(\tau_t)\varepsilon_i \\
&\quad + \sum_{t=1}^{T}\Sigma_{xf,i}^{\mathcal{D}\top}(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)R_{x,it}s_i(\tau_t)\varepsilon_i + \sum_{t=1}^{T}R_{Sf,i}(\tau_t)^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)s_i(\tau_t)\varepsilon_i \\
&\quad + \sum_{t=1}^{T}R_{Sf,i}(\tau_t)^{\top}R_{x,it}s_i(\tau_t)\varepsilon_i \\
&:= A_{NT,46,i} + \cdots + A_{NT,52,i}.
\end{aligned} \tag{D.37}
$$

It suffices only to study $A_{NT,48,i},\cdots,A_{NT,52,i}$, For $A_{NT,48,i}$, we have

$$
\begin{aligned}
A_{NT,48,i} &= \sum_{t=1}^{T}\left(\frac{1}{Th}\sum_{s=1}^{T}K_{s,0}(\tau_t)R_{xf,is}\right)^{\top}\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}\varepsilon_{it} \\
&\quad + \sum_{t=1}^{T}\left([I_p,0_p]\,\widetilde{Q}_{x,i}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T}K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}f_t^{0\top}\right)\right)^{\top}x_{it}\varepsilon_{it} \\
&\quad + \sum_{t=1}^{T}\left([I_p,0_p]\,\widetilde{Q}_{x,i}^{c}(\tau_t)\left(\frac{1}{Th}\sum_{s=1}^{T}K_s^*(\tau_t)\Sigma_{x,i}^{\mathcal{D}-1}(\tau_t)x_{it}f_t^{0\top}\right)\right)^{\top}x_{it}\varepsilon_{it} + O_P\left(\frac{1}{h}\right) \\
&:= A_{NT,48,1,i} + A_{NT,48,2,i} + A_{NT,48,3,i} + O_P\left(\frac{1}{h}\right).
\end{aligned}
$$

Since $(Th)^{-1}\sum_{t=1}^{T}K_{s,0}(\tau)R_{xf,it}$ and $\widetilde{Q}_{x,i}(\tau)$ are independent with $\{\varepsilon_{it}\}$ conditional on $\mathcal{D}$, Lemmas D.1 and D.5(2) imply $A_{NT,48,1,i} = O_P\left(\frac{1}{\sqrt{h}}\right)$ and $A_{NT,48,2,i} = O_P\left(\frac{1}{\sqrt{h}}\right)$. By Lemma D.9, $A_{NT,48,3,i}$ is also negligible. These results yield $\|A_{NT,48,i}\| = O_P\left(\frac{1}{\sqrt{h}}\right)$. Analogously, $\|A_{NT,49,i}\| = O_P\left(\frac{1}{\sqrt{h}}\right)$, $\|A_{NT,50,i}\| = O_P\left(\frac{1}{\sqrt{h}}\right)$, $\|A_{NT,51,i}\| = O_P\left(\frac{1}{\sqrt{h}}\right)$ and $\|A_{NT,52,i}\| = O_P\left(\frac{1}{\sqrt{h}}\right)$. Therefore, Lemma C.4(5) holds. ■

Proof of Lemmas C.5

Using the arguments that are closely related to those in the proofs of Lemma C.4, we can establish the desired results in Lemma C.5. The proofs are provided in the technical supplement (see Liu, 2023).■.

Proof of Lemma C.6

(1) By (C.7)-(C.10), Lemma C.6(1) is a direct extension of the Proposition 1 of Bai (2003). Therefore, its proof is omitted in this paper.

(2) Recall that $H = \left(N^{-1}\Lambda^{0\top}\Lambda^0\right)\left(T^{-1}F^{0\top}\widehat{F}^{(0)}\right)V_{NT}^{-1}$. By Assumption 3, we have $N^{-1}\Lambda^{0\top}\Lambda^0 \xrightarrow{P} \Sigma_\lambda$. It follows Lemma C.6(1) that $T^{-1}F^{0\top}\widehat{F}^{(0)} \xrightarrow{P} \mathcal{Q}^{\top}$. Additionally, we can show $V_{NT} \xrightarrow{P} V$ by using the arguments that are analogous to those in the proof of Lemma A.3 in Bai (2003). Combining these results, we can conclude that Lemma C.6(2) holds. ■

Proof of Lemma C.7

Since $\mathcal{T}_{NT}$ can be rewritten as a U-statistic, arguments that are closely related to those in the proof of Theorem 2 in Hall (1984) can be used to establish its asymptotic distribution. Complete proofs are provided in the technical supplement to the paper (see Liu, 2023). ■

D.2 Technical lemmas

This appendix introduces the technical lemmas that are standard in the literature of α-mixing processes. The proofs of these lemmas are provided in full in the technical supplement (see Liu, 2023).

Lemma D.1. *Suppose ξ_{it} satisfies the $\alpha^{\mathcal{D}}$-mixing conditions in Assumption 1.*

$$\mathbb{E}_{\mathcal{D}}\left(\left\|\frac{1}{T}\sum_{t=1}^{T}\xi_{it}\right\|_F^{2+\delta^*/2}\right) \leq \frac{C}{T^{1+\delta^*/4}}\mathbb{E}_{\mathcal{D}}\left(\|\xi_{it}\|_F^{2+\delta^*/2}\right),$$

where $0<\delta^*<\delta$, *if* $\mathbb{E}_{\mathcal{D}}(\xi_{it})=0$, $\mathbb{E}\left(\|\xi_{it}\|_F^{2+\delta/2}\right)<\infty$.

Lemma D.2. *Suppose ξ_{it} satisfies the $\alpha^{\mathcal{D}}$-mixing conditions in Assumption 1 and $\zeta_{it}=\xi_{it}c(\tau_t)$, where $c(\tau)$ is a uniformly bounded function. As $N,T\to\infty$ simultaneously,*

(1) $\mathbb{E}\left(\left\|\sum_{t=1}^{T}\xi_{it}\right\|_F^2\right)\leq CT$, $\mathbb{E}\left(\left\|\sum_{t=1}^{T}\zeta_{it}\right\|_F^2\right)\leq CT$;

(2) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\xi_{it}\right\|_F^2\right)\leq CN$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\zeta_{it}\right\|_F^2\right)\leq CN$;

(3) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\xi_{it}\right\|_F^2\right)\leq CNT$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\zeta_{it}\right\|_F^2\right)\leq CNT$;

(4) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}\xi_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}\zeta_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T$.

Lemma D.3. *Suppose ξ_{it} satisfies the $\alpha^{\mathcal{D}}$-mixing conditions in Assumption 1 and $\zeta_{it}=\xi_{it}c(\tau_t)$, where $c(\tau)$ is a uniformly bounded function. For $m=0,1,2,3,$ as $N,T\to\infty$ simultaneously,*

(1) $\mathbb{E}\left(\left\|\sum_{t=1}^{T}K_{t,m}(\tau)\xi_{it}\right\|_F^2\right)\leq CTh$, $\mathbb{E}\left(\left\|\sum_{t=1}^{T}K_{t,m}(\tau)\zeta_{it}\right\|_F^2\right)\leq CTh$;

(2) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,m}(\tau)\xi_{it}\right\|_F^2\right)\leq CNTh$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,m}(\tau)\zeta_{it}\right\|_F^2\right)\leq CNTh$;

(3) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}K_{t,m}(\tau)\xi_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}Th$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}K_{t,m}(\tau)\zeta_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}Th$.

Lemma D.4. *Suppose ξ_{it} satisfies the $\alpha^{\mathcal{D}}$-mixing conditions in Assumption 1 and $\zeta_{it}=\xi_{it}c(\tau_t)$, where $c(\tau)$ is a uniformly bounded function. For $m=0,1,2,3,$ as $N,T\to\infty$ simultaneously,*

(1) $\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{t=1}^{T}K_{t,m}(\tau_s)\xi_{it}\right\|_F^2\right)\leq CT^2h$, $\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{t=1}^{T}K_{t,m}(\tau_s)\zeta_{it}\right\|_F^2\right)\leq CT^2h$;

(2) $\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,m}(\tau_s)\xi_{it}\right\|_F^2\right)\leq CNT^2h$, $\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}K_{t,m}(\tau_s)\zeta_{it}\right\|_F^2\right)\leq CNT^2h$;

(3) $\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}K_{t,m}(\tau_s)\xi_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2h$,
$\sum_{s=1}^{T}\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}w_{N,i}K_{t,m}(\tau_s)\zeta_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2h$.

Lemma D.5. *Suppose ξ_{it} satisfies the $\alpha^{\mathcal{D}}$-mixing conditions in Assumption 1 and $\zeta_{it}=\xi_{it}c(\tau_t)$, where $c(\tau)$ is a uniformly bounded function. Additionally, ξ_{it} is independent with ε_{it} conditional on $\mathcal{D}$,*

(1) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\xi_{is}\varepsilon_{it}\right\|_F^2\right)\leq CNT^2$, $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\zeta_{is}\varepsilon_{it}\right\|_F^2\right)\leq CNT^2$;

(2) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T K_{s,m}(\tau_t)\xi_{is}\varepsilon_{it}\right\|_F^2\right)\leq CNT^2h$,
$\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T K_{s,m}(\tau_t)\zeta_{is}\varepsilon_{it}\right\|_F^2\right)\leq CNT^2h$;

(3) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T w_{N,i}\xi_{is}\varepsilon_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2$, $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T w_{N,w}\zeta_{is}\varepsilon_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2$;

(4) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T w_{N,i}K_{s,m}(\tau_t)\xi_{is}\varepsilon_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2h$,
$\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T w_{N,i}K_{s,m}(\tau_t)\zeta_{is}\varepsilon_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T^2h$.

Lemma D.6. *Let Assumptions 1-2 hold. For given $\tau\in(0,1)$, as $N,T\to\infty$ simultaneously,*

(1) $\mathbb{E}\left(\left\|\widehat{\Omega}_{x,i}(\tau)-\Omega_{x,i}^{\mathcal{D}}(\tau)\right\|_F^2\right)\leq C(Th)^{-1}$;

(2) $\mathbb{E}\left(\left\|\widehat{\Omega}_{N,i}(\tau)-\Omega_{N,i}^{\mathcal{D}}(\tau)\right\|_F^2\right)\leq C(Th)^{-1}$;

(3) $\mathbb{E}\left(\left\|\widehat{\Omega}_{xf,i}(\tau)-\Omega_{xf,i}^{\mathcal{D}}(\tau)\right\|_F^2\right)\leq C(Th)^{-1}$;

(4) $\mathbb{E}\left(\left\|\widetilde{Q}_{x,i}(\tau)\right\|_F^2\right)\leq C(Th)^{-1}$.

Lemma D.7. *Let Assumptions 1-2 hold. As $N,T\to\infty$,*

(1) $\sum_{i=1}^N\mathbb{E}\left(\left\|\widehat{\Omega}_{N,i}(\tau)-\Omega_{N,i}^{\mathcal{D}}(\tau)\right\|_F^2\right)\leq CN(Th)^{-1}$, $\sum_{t=1}^T\mathbb{E}\left(\left\|\widehat{\Omega}_{N,i}(\tau_t)-\Omega_{N,i}^{\mathcal{D}}(\tau_t)\right\|_F^2\right)\leq Ch^{-1}$,
$\sum_{i=1}^N\sum_{t=1}^T\mathbb{E}\left(\left\|\widehat{\Omega}_{N,i}(\tau_t)-\Omega_{N,i}^{\mathcal{D}}(\tau_t)\right\|_F^2\right)\leq CNh^{-1}$;

(2) $\sum_{t=1}^T\mathbb{E}\left(\left\|\widehat{\Omega}_{xf,i}(\tau_t)-\Omega_{xf,i}^{\mathcal{D}}(\tau_t)\right\|_F^2\right)\leq Ch^{-1}$, $\sum_{i=1}^N\mathbb{E}\left(\left\|\widehat{\Omega}_{xf,i}(\tau)-\Omega_{xf,i}^{\mathcal{D}}(\tau)\right\|_F^2\right)\leq CN(Th)^{-1}$,
$\sum_{i=1}^N\sum_{t=1}^T\mathbb{E}\left(\left\|\widehat{\Omega}_{xf,i}(\tau_t)-\Omega_{xf,i}^{\mathcal{D}}(\tau_t)\right\|_F^2\right)\leq CNh^{-1}$;

(3) $\sum_{t=1}^T\mathbb{E}\left(\left\|\widetilde{Q}_{x,i}(\tau_t)\right\|_F^2\right)\leq Ch^{-1}$, $\sum_{i=1}^N\mathbb{E}\left(\left\|\widetilde{Q}_{x,i}(\tau)\right\|_F^2\right)\leq CN(Th)^{-1}$,
$\sum_{i=1}^N\sum_{t=1}^T\mathbb{E}\left(\left\|\widetilde{Q}_{x,i}(\tau_t)\right\|_F^2\right)\leq CNh^{-1}$.

Lemma D.8. *Let Assumptions 1-2 hold. For ζ_{it} independent with ε_{it}, as $N,T\to\infty$ simultaneously,*

(1) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\widetilde{Q}_{x,i}(\tau_t)\varepsilon_{it}\right\|_F^2\right)\leq CNh^{-1}$;

(2) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\widetilde{Q}_{x,i}(\tau_t)\zeta_{it}\varepsilon_{it}\right\|_F^2\right)\leq CNT(Th)^{-\delta/(4+\delta)}$, *if* $\mathbb{E}\left(\|\xi_{it}\|_F^{2+\delta/2}\right)\leq\infty$;

(3) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T\widetilde{Q}_{x,i}(\tau_t)\zeta_{it}\varepsilon_{it}\right\|_F^2\right)\leq CNT(Th)^{-(2+\delta)/(4+\delta)}$, *if* $\mathbb{E}\left(\|\xi_{it}\|_F^{4+\delta}\right)\leq\infty$;

(4) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T w_{N,i}\widetilde{Q}_{x,i}(\tau_t)\varepsilon_{it}\right\|_F^2\right)\leq C(\gamma_{N,w}h)^{-1}$;

(5) $\mathbb{E}\left(\left\|\sum_{i=1}^N\sum_{t=1}^T w_{N,i}\widetilde{Q}_{x,i}(\tau_t)\zeta_{it}\varepsilon_{it}\right\|_F^2\right)\leq C\gamma_{N,w}^{-1}T(Th)^{-\delta/(4+\delta)}$, *if* $\mathbb{E}\left(\|\xi_{it}\|_F^{2+\delta/2}\right)\leq\infty$;

(6) $\mathbb{E}\left(\left\|\sum_{i=1}^{N}\sum_{t=1}^{T} w_{N,i}\widetilde{Q}_{x,i}(\tau_t)\zeta_{it}\varepsilon_{it}\right\|_F^2\right) \le C\gamma_{N,w}^{-1}T(Th)^{-(2+\delta)/(4+\delta)}$, *if* $\mathbb{E}\left(\|\xi_{it}\|_F^{4+\delta}\right) \le \infty$.

Lemma D.9. *Let Assumptions 1-2 hold. For given* $\tau \in (0,1)$*, as* $N, T \to \infty$ *simultaneously,*

(1) for given $\tau \in (0,1)$*, there exists an* $M(\tau) > 0$ *such that* $\sum_{i=1}^{N} P\left(\|Q_{x,i}(\tau)\|_F > M(\tau)\right) = o(1)$*;*

(2) there exists an $M > 0$ *such that* $\sum_{i=1}^{N} P\left(\sup_{0<\tau<1} \|Q_{x,i}(\tau)\|_F > M\right) = o(1)$*;*

(3) there exists an $M > 0$ *such that* $P\left(\max_{1\le i\le N} \|Q_{f,i}\|_F > M\right) = o(1)$.

References

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bosq, D. (2012). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*. Lecture Notes in Statistics. Springer New York.

Chernozhukov, V., Hansen, C., Liao, Y., and Zhu, Y. (2018). Inference for heterogeneous effects using low-rank estimation of factor slopes. Working paper at https://arxiv.org/pdf/1812.08089.pdf.

Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis*, 14(1):1–16.

Liu, F. (2023). Technical supplement to "nonparametric time–varying panel data models with heterogeneity". Online technical supplement available at https://ssrn.com/abstract=4384118.

Moon, H. R. and Weidner, M. (2018). Nuclear norm regularized estimation of panel regression models. Working paper available at https://arxiv.org/pdf/1810.10987.pdf.

Su, L. and Chen, Q. (2013). Testing homogeneity in panel data models with interactive fixed effects. *Econometric Theory*, 29(6):1079–1135.

Su, L., Wang, X., and Jin, S. (2019). Sieve estimation of time-varying panel data models with latent structures. *Journal of Business & Economic Statistics*, 37(2):334–349.