Proofs for "Statistical Inferences Using Large Estimated Covariances for Panel Data and Factor Models "

Jushan Bai* Yuan Liao†

Columbia University University of Maryland

A Proofs for Section 3

It can be shown that the following identity holds:

$$\widehat{f}_t - H_W f_t = \widehat{V}^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s' W_T u_t/N + \frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\widehat{\eta}_{st} + \frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\widehat{\theta}_{st}\right) \tag{A.1}$$

where $\widehat{\eta}_{st} = f_s'\Lambda' W_T u_t/N$, and $\widehat{\theta}_{st} = f_t'\Lambda' W_T u_s/N$.

Let $\eta_{st} = f_s'\Lambda' W u_t/N$, and $\theta_{st} = f_t'\Lambda' W u_s/N$.

A.1 Proof of Theorem 3.1

We first cite the Weyl's theorem:

Lemma A.1. (Weyl's Theorem) *Let $\{\lambda_i\}_{i=1}^N$ be the eigenvalues of Σ in descending order. Correspondingly, let $\{\widehat{\lambda}_i\}_{i=1}^N$ be the eigenvalues of $\widehat{\Sigma}$ in descending order. Then for all $i \leq N$, $|\widehat{\lambda}_i - \lambda_i| \leq \|\widehat{\Sigma} - \Sigma\|$.*

Lemma A.2. *All the eigenvalues of $\widehat{V}^{-1}$ are $O_p(1)$.*

Proof. Let w.p.a.1 be short for "with probability approaching one". It suffices to show the first r largest eigenvalues of the T by T matrix $YW_TY'/(TN)$ are bounded away from zero. Note that these eigenvalues are also the first r largest eigenvalues of the $N \times N$ matrix $W_T^{1/2}Y'YW_T^{1/2}/(TN) = W_T^{1/2}\frac{1}{TN}\sum_{t=1}^T Y_tY_t'W_T^{1/2}$. Let $S = \frac{1}{T}\sum_{t=1}^T Y_tY_t'$. It suffices to show that w.p.a.1, the first r largest eigenvalues of $W_T^{1/2}SW_T^{1/2}/N$ are bounded away from both zero and infinity.

*Department of Economics, Columbia University, New York, NY 10027.

†Department of Mathematics, University of Maryland at College Park, College Park, MD 20742.

Because all the eigenvalues of W are bounded away from both zero and infinity, and W_T consistently estimates W in the operator norm, so w.p.a.1, all the eigenvalues of W_T are bounded away from both zero and infinity. In addition, by the pervasiveness assumption, all the eigenvalues of $\Lambda'\Lambda$ are growing at rate $O(N)$. It follows from

$$\lambda_{\max}(\mathrm{cov}(f_t)^{1/2}\Lambda' W_T \Lambda \mathrm{cov}(f_t)^{1/2}) \leq \lambda_{\max}(W_T)\lambda_{\max}(\Lambda'\Lambda)\lambda_{\max}(\mathrm{cov}(f_t))$$

and

$$\lambda_{\min}(\mathrm{cov}(f_t)^{1/2}\Lambda' W_T \Lambda \mathrm{cov}(f_t)^{1/2}) \geq \lambda_{\min}(W_T)\lambda_{\min}(\Lambda'\Lambda)\lambda_{\min}(\mathrm{cov}(f_t))$$

that w.p.a.1, all the eigenvalues of $\mathrm{cov}(f_t)^{1/2}\Lambda' W_T \Lambda \mathrm{cov}(f_t)^{1/2}/N$ are bounded away from both zero and infinity. This statement also applies to the first r largest eigenvalues of the $N \times N$ matrix $W_T^{1/2}\Lambda \mathrm{cov}(f_t)\Lambda' W_T^{1/2}/N$.

Let $\Sigma_y = \mathrm{cov}(Y_t)$. Because

$$W_T^{1/2}\Sigma_y W_T^{1/2} = W_T^{1/2}\Lambda \mathrm{cov}(f_t)\Lambda' W_T^{1/2} + W_T^{1/2}\Sigma_u W_T^{1/2},$$

and $\|W_T^{1/2}\Sigma_u W_T^{1/2}/N\| = O_p(N^{-1})$. By the Weyl's theorem, w.p.a.1., the first r eigenvalues of $W_T^{1/2}\Sigma_y W_T^{1/2}/N$ are also bounded away from both zero and infinity. Moreover, $\|S - \Sigma_y\| = O_p(N\sqrt{\log N/T})$ (see Lemma 5 of Fan et al. 2013), which implies

$$\|W_T^{1/2}(S - \Sigma_y)W_T^{1/2}/N\| = o_p(1).$$

Still by the Weyl's theorem, w.p.a.1, the first r eigenvalues of $W_T^{1/2} S W_T^{1/2}$ are bounded away from both zero and infinity. $\square$

Lemma A.3. *(i)* $\|H_W\| = O_p(1)$ *and* $\|H_W^{-1}\| = O_p(1)$
(ii) $H_W \mathrm{cov}(f_t) H_W' = I_r + O_p(T^{-1/2} + N^{-1/2} + \|W_T - W\|)$,
$H_W' H_W = \mathrm{cov}(f_t)^{-1} + O_p(T^{-1/2} + N^{-1/2} + \|W_T - W\|)$.

Proof. We have, $\|H_W\| = \|\widehat{V}^{-1}\|\|\widehat{F}\|\|F\|\|\Lambda' W_T \Lambda\|/(NT) = O_p(1)$ since $\|\widehat{F}\| = O_p(\sqrt{T}) = \|F\|$ and $\|\Lambda'\Lambda\| = O(N)$. In addition,

$$I_r = \widehat{F}'\widehat{F}/T = \widehat{F}'(\widehat{F} - FH_W')/T + (\widehat{F} - FH_W')'FH_W'/T$$

$$+ H_W(F'F/T - \mathrm{cov}(f_t))H_W' + H_W \mathrm{cov}(f_t) H_W'.$$

By (A.5) (we prove in the below, which does not depend on H_W^{-1}):

$$\|\widehat{F} - FH_W'\|^2/T = \frac{1}{T}\sum_{t=1}^T \|\widehat{f}_t - H_W f_t\|^2 = O_p(N^{-1} + T^{-1} + \|W_T - W\|^2).$$

It follows from $\|F'F/T - \mathrm{cov}(f_t)\| = O_p(T^{-1/2})$ that

$$H_W \mathrm{cov}(f_t) H_W' = I_r + O_p(T^{-1/2} + N^{-1/2} + \|W_T - W\|).$$

Since $\lambda_{\min}(\mathrm{cov}(f_t)) > c > 0$, we have $\lambda_{\min}(H_W H_W')$ is bounded away from zero, which implies $\|H_W^{-1}\| = O_p(1)$. Right multiplying H_W and left multiplying H_W^{-1} to the identity of $H_W \mathrm{cov}(f_t) H_W'$ yields

$$\mathrm{cov}(f_t) H_W' H_W = I_r + O_p(T^{-1/2} + N^{-1/2} + \|W_T - W\|),$$

which gives the desired result for $H_W' H_W$.

□

A.1.1 Limiting distribution for estimated loadings

Lemma A.4. *For each $j \le N$,*
(i) $\|\frac{1}{T}\sum_{t=1}^T (\widehat{f}_t - H_W f_t) u_{jt}\| = O_p(\|W_T - W\|(\|W_T - W\| + \sqrt{\frac{\log N}{T}} + \sqrt{\frac{1}{N}}) + \frac{1}{T} + \frac{1}{\sqrt{NT}} + \frac{1}{N})$.
(ii) $\|\frac{1}{T}\sum_{t=1}^T \widehat{f}_t (H_W f_t - \widehat{f}_t)' H_W'^{-1} \lambda_j\| = O_p(\|W_T - W\|(\|W_T - W\| + \sqrt{\frac{\log N}{T}} + \frac{1}{\sqrt{N}}) + \frac{\log N}{T} + \sqrt{\frac{\log N}{NT}} + \frac{1}{N})$.
(iii) $\|(H_W - \frac{1}{NT} V^{-1} \widehat{F}' F \Lambda' W \Lambda) \frac{1}{T}\sum_{t=1}^T f_t u_{it}\| = O_p(\|W_T - W\|/\sqrt{T})$.

Proof. (i) By the identity (A.1) and triangular inequality, we have,

$$\begin{aligned}
&\|\frac{1}{T}\sum_{t=1}^T (\widehat{f}_t - H_W f_t) u_{it}\| \le \|\widehat{V}^{-1}\| \Big[\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s (E u_s' W u_t) u_{it}/N\| \\
&+ \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s (u_s' W u_t - E u_s' W u_t) u_{it}/N\| \\
&+ \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s' (W_T - W) u_t u_{it}/N\| + \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s \eta_{st} u_{it}/N\| \\
&+ \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s (\eta_{st} - \widehat{\eta}_{st}) u_{it}/N\| + \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s \theta_{st} u_{it}/N\| + \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s (\widehat{\theta}_{st} - \theta_{st}) u_{it}/N\| \Big].
\end{aligned}$$

Note that $\|\widehat{V}^{-1}\| = O_p(1)$. All the other terms on the right hand side are bounded in Lemmas

A.14 and A.15, which yield the result.
(ii) Let $a = \|H_W \frac{1}{T}\sum_{t=1}^T f_t(H_W f_t - \widehat{f}_t)' H_W'^{-1}\lambda_j\|$. By Lemma A.13,

$$\|\frac{1}{T}\sum_{t=1}^T \widehat{f}_t(H_W f_t - \widehat{f}_t)' H_W'^{-1}\lambda_j\| = a + O_p(\|W_T - W\|^2 + N^{-1} + T^{-1}).$$

We now bound a. Since $\|H_W\| = O_p(1) = \|H_W'^{-1}\|$ and $\|\lambda_j\| = O(1)$, we have $a = O_p(1)\|\frac{1}{T}\sum_{t=1}^T f_t(H_W f_t - \widehat{f}_t)'\|_F$. The triangular inequality implies

$$\begin{aligned}
&\|\frac{1}{T}\sum_{t=1}^T (H_W f_t - \widehat{f}_t) f_t'\|_F \leq \|\widehat{V}^{-1}\| \Big[\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'(W_T - W) u_t f_t'/N\|_F \\
&+\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s' W u_t f_t'/N\|_F \\
&+\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s(\widehat{\eta}_{st} - \eta_{st}) f_t'\|_F + \|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s \eta_{st} f_t'\|_F \\
&+\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s(\widehat{\theta}_{st} - \theta_{st}) f_t'\|_F + \|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s \theta_{st} f_t'\|_F \Big].
\end{aligned}$$

Again, except for $\|\widehat{V}^{-1}\| = O_p(1)$, all the other terms on the right hand side are bounded in Lemmas A.13 and A.15.
(iii) The objective is bounded by

$$\|\frac{1}{NT} V^{-1}\widehat{F}' F \Lambda'(W_T - W)\Lambda \frac{1}{T}\sum_{t=1}^T f_t u_{jt}\|$$

$$\leq O_p(\|\frac{1}{T}\sum_{t=1}^T f_t u_{jt}\|\|W_T - W\|) = O_p(\|W_T - W\|/\sqrt{T}).$$

□

We now derive the limit of H_W using a similar argument of Bai (2003).

Lemma A.5. *$H_W \to^p Q_W'^{-1}$, $\widehat{F}' F/T \to^p Q_W$ and $\widehat{V} \to^p V$ where V is the diagonal matrix consisting of the eigenvalues of $\Sigma_\Lambda \mathrm{cov}(f_t)$.*

Proof. Let $\tilde{Y} = W_T^{1/2} Y$ and $\tilde{\Lambda} = W_T^{1/2}\Lambda$ and $\tilde{u} = W_T^{1/2} u$. Then $\tilde{Y} = \tilde{\Lambda} F' + \tilde{u}$. The columns of $\widehat{F}/\sqrt{T}$ are the eigenvectors corresponding to the largest r eigenvalues of $Y' W_T Y = \tilde{Y}'\tilde{Y}$. In addition, $\|W_T - W\| = o_p(1)$ implies $\tilde{\Lambda}'\tilde{\Lambda}/N = \Lambda' W_T \Lambda/N \to \Sigma_\Lambda$. Hence Proposition 1 of Bai (2003) can be directly applied to $(\widehat{F}, F, \tilde{\Lambda}, \tilde{Y})$, which implies $\|\widehat{F}' F/T - Q_W\| = o_p(1)$, where $Q_W = V^{1/2}\Gamma'\Sigma_\Lambda^{-1/2}$. This then implies $H_W \to^p V^{-1}Q_W\Sigma_\Lambda$. The result follows from

$$V^{-1}Q_W\Sigma_\Lambda = V^{-1}V^{1/2}\Gamma'\Sigma_\Lambda^{-1/2}\Sigma_\Lambda = Q_W'^{-1}.$$

The third convergence follows from applying Lemma A.3 of Bai (2003) to $(\tilde{Y}, F, \tilde{\Lambda})$.

$\square$

Proof of Theorem 3.1: $\widehat{\lambda}_j$ (limiting distribution)

By (A.7) and Lemma A.4,

$$\widehat{\lambda}_j - H_W'^{-1}\lambda_j = H_W\frac{1}{T}\sum_{t=1}^T f_t u_{jt} + O_p(\|W_T - W\|(\|W_T - W\| + \sqrt{\frac{\log N}{T}} + \frac{1}{\sqrt{N}})$$

$$+\frac{\log N}{T} + \sqrt{\frac{\log N}{NT}} + \frac{1}{N}).$$

By the assumptions $\|W_T - W\| = o_p(\min\{T^{-1/4}, \sqrt{\frac{N}{T}}\})$ and $T = o(N^2)$,

$$\sqrt{T}(\widehat{\lambda}_j - H_W'^{-1}\lambda_j) = H_W\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t u_{jt} + o_p(1). \tag{A.2}$$

The desired limiting distribution follows from the assumed central limit theorem and Lemma A.5.

A.1.2 Limiting distributions for estimated factors

We first obtain some lemmas to strengthen the convergence rates.

Lemma A.6. *For each $t \leq T$,*
(i) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(u_s'Wu_t - Eu_s'Wu_t)/N\| = O_p(\|W_T - W\|/\sqrt{N} + 1/N + 1/\sqrt{NT})$.
(ii) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(Eu_s'Wu_t)/N\| = O_p(\|W_T - W\|/\sqrt{T} + 1/\sqrt{NT} + 1/T)$.
(iii) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s\theta_{st}\| = O_p(\|W_T - W\|/\sqrt{N} + 1/N + 1/\sqrt{NT})$.
(iv) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'(W_T - W)u_t/N\| = O_p(\|W_T - W\|(\|W_T - W\| + 1/\sqrt{N} + \sqrt{\log N/T}))$.
(v) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})\| = O_p(\|W_T - W\|(\|W_T - W\| + 1/\sqrt{N} + \sqrt{\log N/T}))$.
(vi) $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})\| = O_p(\|W_T - W\|(\|W_T - W\| + 1/\sqrt{N} + 1/\sqrt{T})) + o_p(1/\sqrt{N})$.

Proof. (i) The objective is bounded by

$$\|\frac{1}{T}\sum_{s=1}^T f_s(u_s'Wu_t - Eu_s'Wu_t)/N\| + O_p(\|W_T - W\|/\sqrt{N} + 1/N + 1/\sqrt{NT}).$$

By assumption 3.4 $E(\frac{1}{\sqrt{NT}}\sum_{s=1}^T f_s(u_s'Wu_t - Eu_s'Wu_t))^2 = O(1)$, the first term is $O_p(1/\sqrt{NT})$.

(ii) Since $\max_{t\leq T}\sum_{s=1}^T |Eu_s'Wu_t/N| = O(1)$, the objective is bounded by $\|\frac{1}{T}\sum_{s=1}^T f_s(Eu_s'Wu_t)/N\| + O_p(\|W_T - W\|/\sqrt{T} + 1/\sqrt{NT} + 1/T)$. Also,

$$E\|\frac{1}{NT}\sum_{s=1}^T f_s(Eu_s'Wu_t)\| \leq \frac{1}{NT}\sum_{s=1}^T E\|f_s\||Eu_s'Wu_t|$$

$$\leq \max_{s\leq T} E\|f_s\|\frac{1}{NT}\sum_{s=1}^T |Eu_s'Wu_t| = O(1/T),$$

since $\max_{s\leq T} E\|f_s\| = E\|f_s\| = O(1)$.

(iii) The objective is bounded by

$$\|\frac{1}{T}\sum_{s=1}^T f_s u_s'W\Lambda f_t/N\| + O_p(\|W_T - W\|/\sqrt{N} + 1/N + 1/\sqrt{NT}).$$

By assumption 3.4 $\|\frac{1}{\sqrt{NT}}\sum_{s=1}^T f_s u_s'W\Lambda\|_F = O_p(1)$, the first term is $O_p(1/\sqrt{NT})$.

The proofs for (iv) and (v) are straightforward based on the triangular inequality and Lemmas A.7, A.9(i)(ii) below.

(vi) By the triangular inequality, the objective is bounded by

$$a + O_p(\|W_T - W\|(\|W_T - W\| + 1/\sqrt{N} + 1/\sqrt{T})),$$

where $a = \|H_W\frac{1}{T}\sum_{s=1}^T f_s f_s'\Lambda'(W_T - W)u_t/N\|$. The desired result then follows from assumption 3.1 $\|\Lambda'(W_T - W)u_t/\sqrt{N}\| = o_p(1)$.

$\square$

Proof of Theorem 3.1: $\widehat{f}_t$ (limiting distribution)

Let

$$d_T = \|W_T - W\|(\|W_T - W\| + 1/\sqrt{N} + \sqrt{\log N/T}) + 1/N + 1/T + 1/\sqrt{NT}.$$

Then by Lemma A.6, $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'W_T u_t\| = O_p(d_T) = \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s\widehat{\theta}_{st}\|$ and $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})\| = O_p(d_T) + o_p(1/\sqrt{N})$. It follows from identity (A.1) and Lemma A.5 that

$$\sqrt{N}(\widehat{f}_t - H_W f_t) = \sqrt{N}\widehat{V}^{-1}\left(\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'W_T u_t/N + \frac{1}{T}\sum_{s=1}^T \widehat{f}_s\widehat{\eta}_{st} + \frac{1}{T}\sum_{s=1}^T \widehat{f}_s\widehat{\theta}_{st}\right)$$

$$= \sqrt{N}\widehat{V}^{-1}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st} + \sqrt{N}\widehat{V}^{-1}\left[\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s' W_T u_t + \frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\widehat{\eta}_{st} - \eta_{st}) + \frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\widehat{\theta}_{st}\right]$$
$$= \sqrt{N}\widehat{V}^{-1}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st} + O_p(\sqrt{N}d_T) + o_p(1) = \widehat{V}^{-1}\frac{\widehat{F}'F}{T}\frac{\Lambda' W u_t}{\sqrt{N}} + o_p(1).$$

Hence

$$\sqrt{N}(\widehat{f}_t - H_W f_t) = V^{-1}Q_W\frac{\Lambda' W u_t}{\sqrt{N}} + o_p(1). \tag{A.3}$$

The desired limiting distribution follows from the assumption that

$$(\Lambda' W \Sigma_u W \Lambda)^{-1/2}\Lambda' W u_t \to^d \mathcal{N}(0, I_r).$$

Proof of Theorem 3.1: common components

Write $C_{it} = \lambda_i' f_t$ and $\widehat{C}_{it} = \widehat{\lambda}_i'\widehat{f}_t$. We have, for each fixed i, t,

$$\widehat{C}_{it} - C_{it} = (\widehat{f}_t - H_W f_t)' H_W'^{-1}\lambda_i + f_t' H_W'(\widehat{\lambda}_i - H_W'^{-1}\lambda_i) + K_T \tag{A.4}$$

where $K_T = (\widehat{f}_t - H_W f_t)'(\widehat{\lambda}_i - H_W'^{-1}\lambda_i) = O_p(T^{-1} + N^{-1} + \|W_T - W\|^2)$. By the definition of H_W,$H_W^{-1}\widehat{V}^{-1}\widehat{F}'F/T = (\Lambda' W_T \Lambda/N)^{-1}$. Also, Lemma A.3 implies $H_W' H_W = \mathrm{cov}(f_t)^{-1} + O_p(T^{-1/2} + N^{-1/2} + \|W_T - W\|)$.

It then follows from (A.2) and (A.3) that

$$\begin{aligned}
\widehat{C}_{it} - C_{it} &= \frac{1}{NT}\lambda_i' H_W^{-1}\widehat{V}^{-1}\widehat{F}' F \Lambda' W u_t + \frac{1}{T} f_t' H_W' H_W \sum_{s=1}^{T} f_s u_{is} + o_p(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}) \\
&= \frac{1}{N}\lambda_i'(\Lambda' W \Lambda/N)^{-1}\Lambda' W u_t + f_t'\mathrm{cov}(f_t)^{-1}\frac{1}{T}\sum_{s=1}^{T} f_s u_{is} + o_p(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}) \\
&= \frac{1}{\sqrt{N}}A_{it} + \frac{1}{\sqrt{T}}B_{it} + o_p(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}),
\end{aligned}$$

where $A_{it} = \lambda_i'(\Lambda' W \Lambda/N)^{-1}\Lambda' W u_t/\sqrt{N}$ and $B_{it} = f_t'\mathrm{cov}(f_t)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T} f_s u_{is}$.

Let $G_W = N(\Lambda' W \Lambda)^{-1}\Lambda' W \Sigma_u W \Lambda(\Lambda' W \Lambda)^{-1}$. We then have

$$(\lambda_i' G_W \lambda_i)^{-1/2} A_{it} \to^d \mathcal{N}(0, 1)$$

and $(f_t'\mathrm{cov}(f_t)^{-1}\Phi_i\mathrm{cov}(f_t)^{-1}f_t)^{-1/2}B_{it} \to^d \mathcal{N}(0, 1)$. The same argument of the proof of Theorem 3 in Bai (2003) then implies

$$\frac{\widehat{C}_{it} - C_{it}}{(\lambda_i' G_W \lambda_i/N + f_t'\mathrm{cov}(f_t)^{-1}\Phi_i\mathrm{cov}(f_t)^{-1}f_t/T)^{1/2}} \to^d \mathcal{N}(0, 1).$$

The result then follows since $\Lambda' W \Lambda / N \to \Sigma_\Lambda$.

A.1.3 Proof of Theorem 3.2

Lemma A.7.

$$\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t - H_W f_t\|^2 = O_p(\|W_T - W\|^2 + N^{-1} + T^{-1}). \tag{A.5}$$

Proof. The triangular inequality and (A.1) imply that for all $t \le T$,

$$\begin{aligned}
\|\widehat{f}_t - H_W f_t\| &\le \|\widehat{V}^{-1}\|\Big[\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s'(W_T - W)u_t/N\| \\
&+\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(u_s'Wu_t - Eu_s'Wu_t)/N\| \\
&+\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s Eu_s'Wu_t/N\| + \|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\widehat{\eta}_{st} - \eta_{st})\| + \|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}\| \\
&+\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\widehat{\theta}_{st} - \theta_{st})\| + \|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}\|\Big] \\
&= \|\widehat{V}^{-1}\|\sum_{i=1}^{7} G_{it}
\end{aligned} \tag{A.6}$$

Term $\|\widehat{V}^{-1}\|$ is bounded by Lemma A.2. Term $\frac{1}{T}\sum_{i=1}^{T} G_{1t}^2$, $\frac{1}{T}\sum_{i=1}^{T} G_{2t}^2$ and $\frac{1}{T}\sum_{i=1}^{T} G_{3t}^2$ are bounded by Lemma A.10(ii)(iii) respectively. Hence

$$\frac{1}{T}\sum_{i=1}^{T} G_{1t}^2 + G_{2t}^2 + G_{3t}^2 = O_p(\|W_T - W\|^2 + \frac{1}{N} + \frac{1}{T}).$$

The remaining terms follow from Lemmas A.11 (ii)(iv), A.12 (ii)(iv)

$$\frac{1}{T}\sum_{i=1}^{T} G_{4t}^2 + G_{5t}^2 + G_{6t}^2 + G_{7t}^2 = O_p(\|W_T - W\|^2 + \frac{1}{N}).$$

This gives the desired result.

□

Proof of Theorem 3.2: uniform convergence

By Lemmas A.10 and A.8, $\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s'(W_T - W)u_t/N\| = O_p((\log T)^{1/r_1}\|W_T -$

$W\|)$. By Lemma A.8,

$$\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(u_s'Wu_t-Eu_s'Wu_t)/N\|=O_p(T^{1/(2\delta)}N^{-1/2}).$$

By Lemma A.10, $\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_sEu_s'Wu_t/N\|=O_p(T^{-1/2})$. By Lemmas A.11 and A.8, $\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(\widehat{\eta}_{st}-\eta_{st})\|=O_p((\log T)^{1/r_1}\|W_T-W\|)$. By Lemma A.8, $\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s\eta_{st}\|=O_p(T^{1/(2\delta)}N^{-1/2})$. By Lemmas A.12 and A.8, $\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(\widehat{\theta}_{st}-\theta_{st})\|=O_p((\log T)^{1/r_2}\|W_T-W\|)$. By Lemma A.12 that $\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s\theta_{st}\|=O_p((\log T)^{1/r_2}N^{-1/2})$. It then follows from Lemma A.2 and inequality (A.6) that

$$\max_{t\le T}\|\widehat{f}_t-H_Wf_t\|=O_p(((\log T)^{1/r_1}+(\log T)^{1/r_2})\|W_T-W\|+T^{1/(2\delta)}N^{-1/2}+T^{-1/2}).$$

Using the fact that $\widehat{\Lambda}'=\widehat{F}'Y'/T$ and $Y=\Lambda F'+u$, we obtain that for each $j\le N$,

$$\widehat{\lambda}_j-H_W'^{-1}\lambda_j=\frac{1}{T}\sum_{t=1}^T\widehat{f}_t(H_Wf_t-\widehat{f}_t)'H_W'^{-1}\lambda_j+\frac{1}{T}\sum_{t=1}^Tu_{jt}(\widehat{f}_t-H_Wf_t)+\frac{1}{T}\sum_{t=1}^TH_Wf_tu_{jt}. \quad (A.7)$$

The uniform convergence rate for $\widehat{\lambda}_j$ then follows from Lemma A.9.

A.2 Technical lemmas

Lemma A.8. *(i)* $\max_{t\le T}\|u_t/\sqrt{N}\|=O_p((\log T)^{1/r_1})$,
$\max_{t\le T}\|f_t\|=O_p((\log T)^{1/r_2})$.
(ii) $\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(u_s'Wu_t-Eu_s'Wu_t)/N\|=O_p(T^{1/(2\delta)}N^{-1/2})$.
(iii) $\max_{t\le T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s\eta_{st}\|=O_p(T^{1/(2\delta)}N^{-1/2})$.

Proof. (i) The results follow immediately from the exponential-tail conditions for (u_t,f_t), and the Bonferroni's method.
(ii) Cauchy-Schwarz inequality implies that

$$\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(u_s'Wu_t-Eu_s'Wu_t)/N\|^2\le\frac{1}{T}\sum_{s=1}^T\|\widehat{f}_s\|^2\frac{1}{T}\sum_{s=1}^T|(u_s'Wu_t-Eu_s'Wu_t)/N|^2.$$

Let $\psi_t=\frac{1}{T}\sum_{s=1}^T|(u_s'Wu_t-Eu_s'Wu_t)/N|^2$. For $\delta\ge1$, Hölder's inequality gives $\psi_t^\delta\le\frac{1}{T}\sum_{s=1}^T|(u_s'Wu_t-Eu_s'Wu_t)/N|^{2\delta}$. Thus

$$E\psi_t^\delta\le E|(u_s'Wu_t-Eu_s'Wu_t)/N|^{2\delta}=O(N^{-\delta}),$$

where $O(N^{-\delta})$ does not depend on either s or t. Then for any $s > 0$, by Bonferroni and Markov inequalities,

$$P(\max_{t\leq T}\psi_t > s) \leq T\max_{t\leq T}P(\psi_t^\delta > s^\delta) \leq \frac{T\max_{t\leq T}E\psi_t^\delta}{s^\delta} = O(\frac{T}{N^\delta s^\delta}),$$

which implies $\max_{t\leq T}\psi_t = O_p(T^{1/\delta}/N)$. Due to $\frac{1}{T}\sum_{s=1}^T\|\widehat{f}_s\|^2 = r$, we have

$$\max_{t\leq T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s(u_s'Wu_t - Eu_s'Wu_t)/N\|^2 \leq r\max_{t\leq T}\psi_t = O_p(T^{1/\delta}/N).$$

(iii) We have $\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_s\eta_{st}\|^2 \leq \frac{r}{T}\sum_{s=1}^T\|f_s\|^2\|\Lambda'Wu_t/N\|^2$. Let $\phi_t = \|\Lambda'Wu_t/N\|^2$. Then for any $s > 0$, Bonferroni and Markov inequalities imply that

$$P(\max_{t\leq T}\phi_t > s) \leq T\max_{t\leq T}P(\phi_t^\delta > s^\delta) \leq \frac{T\max_{t\leq T}E\phi_t^\delta}{s^\delta} = O(\frac{T}{N^\delta s^\delta}),$$

which implies $\max_{t\leq T}\phi_t = O_p(T^{1/\delta}/N)$. The result then follows from $\frac{1}{T}\sum_{s=1}^T\|f_s\|^2 = O_p(1)$. □

Lemma A.9. *(i)* $\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^T f_tu_{jt}\| = O_p(\sqrt{(\log N)/T})$.
(ii) $\max_{i,j\leq N}|\frac{1}{T}\sum_{t=1}^T u_{jt}u_{it} - Eu_{jt}u_{it}| = O_p(\sqrt{\log N/T})$.
(iii)

$$\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^T\widehat{f}_t(H_Wf_t - \widehat{f}_t)'H_W'^{-1}\lambda_j\| = O_p(\|W_T - W\| + \sqrt{\frac{1}{N}} + \sqrt{\frac{1}{T}}).$$

(iv) $\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^T u_{jt}(\widehat{f}_t - H_Wf_t)\| = O_p(\|W_T - W\| + N^{-1/2} + T^{-1/2})$.

Proof. (i) and (ii) are proved in Fan et al. (2013, Lemma C.3).
(iii) It follows from Cauchy-Schwarz inequality and Lemma A.7 that the objective is bounded by,

$$(\frac{1}{T}\sum_{t=1}^T\|\widehat{f}_t\|^2)^{1/2}(\frac{1}{T}\sum_{t=1}^T\|H_Wf_t - \widehat{f}_t\|^2)^{1/2}\|H_W^{-1}\|\max_{j\leq N}\|\lambda_j\|$$
$$= O_p(\|W_T - W\| + N^{-1/2} + T^{-1/2}).$$

(iv) Because $\max_{j\leq N}\frac{1}{T}\sum_{t=1}^T u_{jt}^2 \leq O_p(\sqrt{\log N/T}) + \max_{j\leq N}Eu_{jt}^2$. Hence (iv) follows from Cauchy-Schwarz inequality and Lemma A.7. □

Lemma A.10. *(i)* $\max_{t\leq T}\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_sE(u_s'Wu_t)/N\| = O_p(T^{-1/2})$,
$\frac{1}{T}\sum_{t=1}^T\|\frac{1}{T}\sum_{s=1}^T\widehat{f}_sEu_s'Wu_t/N\|^2 = O_p(T^{-1})$.

(ii) For each t,

$$\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'(W-W_T)u_t/N\| = \|u_t/\sqrt{N}\|O_p(\|W_T-W\|),$$

where the $O_p(\cdot)$ does not depend on t, and

$$\frac{1}{T}\sum_{t=1}^T \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'(W-W_T)u_t/N\|^2 = O_p(\|W_T-W\|^2).$$

(iii) $\frac{1}{T}\sum_{t=1}^T \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(u_s'Wu_t - Eu_s'Wu_t)/N\|^2 = O_p(N^{-1})$
(iv) For each t, $\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s(u_s'Wu_t - Eu_s'Wu_t)/N\| = O_p(N^{-1/2})$.

Proof. (i) By the Cauchy-Schwarz inequality, $\|T^{-1}\sum_{s=1}^T \widehat{f}_s E(u_s'Wu_t)/N\|$ is bounded by

$$\max_{t\leq T}(\frac{1}{T}\sum_{s=1}^T \|\widehat{f}_s\|^2)^{1/2}(\frac{1}{T}\sum_{s=1}^T \frac{(Eu_s'Wu_t)^2}{N^2})^{1/2} = O_p(\frac{1}{\sqrt{T}}).$$

Hence the second statement of (i) follows immediately from

$$\frac{1}{T}\sum_{t=1}^T \|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s Eu_s'Wu_t/N\|^2 \leq \max_{t\leq T}\|\frac{1}{T}\sum_{s=1}^T \widehat{f}_s E(u_s'Wu_t)/N\|^2.$$

(ii) By the Cauchy-Schwarz inequality, $\frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s'(W_T-W)u_t/N$ is bounded by

$$(\frac{1}{T}\sum_{s=1}^T \|\widehat{f}_s\|^2)^{1/2}(\frac{1}{T}\sum_{s=1}^T \frac{1}{N^2}\|u_s\|^2\|u_t\|^2\|W_T-W\|^2)^{1/2} = O_p(1)\|u_t/\sqrt{N}\|\|W_T-W\|.$$

The second statement follows since $\frac{1}{T}\sum_{t=1}^T \|u_t\|^2 = O_p(N)$ and $O_p(1)$ above does not depend on t.
(iii) Since

$$E\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T |(u_s'Wu_t - Eu_s'Wu_t)/N|^2 = E|(u_s'Wu_t - Eu_s'Wu_t)/N|^2 = O(N^{-1})$$

and $\frac{1}{T}\sum_{t=1}^T \|\widehat{f}_t\|^2 = r$, the objective is bounded by

$$\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \|\widehat{f}_s\|^2 \frac{1}{T}\sum_{s=1}^T \frac{1}{N^2}|u_s'Wu_t - Eu_s'Wu_t|^2$$

$$= \frac{r}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\frac{1}{N^2}|u_s'Wu_t - Eu_s'Wu_t|^2 = O_p(\frac{1}{N}).$$

(iv) Note that

$$E\frac{1}{T}\sum_{s=1}^{T}|(u_s'Wu_t - Eu_s'Wu_t)/N|^2 = E|(u_s'Wu_t - Eu_s'Wu_t)/N|^2 = O(N^{-1})$$

and $\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t\|^2 = r$. The objective is then bounded by
$(\frac{1}{T}\sum_{s=1}^{T}|(u_s'Wu_t - Eu_s'Wu_t)/N|^2)^{1/2}$, which is $O_p(N^{-1/2})$. □

Lemma A.11. *(i) For each $t \leq T$, $\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}\| = O_p(N^{-1/2})$.*
(ii) $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}\|^2 = O_p(N^{-1})$.
(iii) For each $t \leq T$, $\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})\| = \|u_t/\sqrt{N}\|O_p(\|W_T - W\|)$, where $O_p(\cdot)$ does not depend on t.
(iv) $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})\|^2 = O_p(\|W_T - W\|^2)$.

Proof. (i) First, $E\|\Lambda'Wu_t/N\|^2 = O(N^{-1})$. By the Cauchy-Schwarz inequality, with the fact $\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t\|^2 = r$, we have
$\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}\|^2 \leq \frac{r}{T}\sum_{s=1}^{T}\|f_s\|^2\|\Lambda'Wu_t/N\|^2 = O_p(N^{-1})$.
(ii) The same argument as above implies that
$\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}\|^2 \leq O_p(1)\frac{1}{T}\sum_{t=1}^{T}\|\Lambda'Wu_t/N\|^2 = O_p(N^{-1})$.
(iii) We have

$$\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})\| \leq \|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s f_s'\Lambda'(W_T - W)u_t/N\|$$
$$\leq O_p(1)\|\Lambda/\sqrt{N}\|\|u_t/\sqrt{N}\|\|W_T - W\|.$$

Parts (iii) and (iv) then follow immediately.

□

Lemma A.12. *(i) For each $t \leq T$, $\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}\| = \|f_t\|O_p(N^{-1/2})$,*
(ii) $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}\|^2 = O_p(N^{-1})$.
(iii) For each $t \leq T$, $\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})\| = \|f_t\|O_p(\|W_T - W\|)$.
(iv) $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})\|^2 = O_p(\|W_T - W\|^2)$.
None of the $O_p(\cdot)$ terms in (i)-(iv) depend on t.

Proof. (i) First, $E\|\Lambda'Wu_t/N\|^2 = O(N^{-1})$. By the Cauchy-Schwarz inequality, with the fact

$\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t\|^2 = r$, we have

$$\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}\|^2 \leq \frac{r}{T}\sum_{s=1}^{T}\|f_t\|^2\|\Lambda' W u_s/N\|^2 = O_p(\|f_t\|^2 N^{-1}),$$

where the last equality follows from $E\frac{1}{T}\sum_{s=1}^{T}\|\Lambda' W u_s/N\|^2 = O(N^{-1})$. Part (i) and (ii) follow immediately.

(iii) We have,

$$\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\theta_{st}-\widehat{\theta}_{st})\| \leq \|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s f_t'\Lambda'(W_T - W)u_s/N\|$$

$$\leq \frac{1}{T}\sum_{s=1}^{T}\|\widehat{f}_s f_t'\Lambda'\|\|u_s/N\|\|W_T - W\| = O(\|W_T - W\|)\frac{1}{T}\sum_{s=1}^{T}\|\widehat{f}_s\|\|u_s/\sqrt{N}\|\|f_t\|.$$

Parts (iii) and (iv) then follow immediately.

$\square$

Lemma A.13. *(i)*

$$\|\frac{1}{T}\sum_{t=1}^{T}\widehat{f}_t(H_W f_t - \widehat{f}_t)' H_W'^{-1}\lambda_j\|$$

$$= \|H_W\frac{1}{T}\sum_{t=1}^{T}f_t(H_W f_t - \widehat{f}_t)' H_W'^{-1}\lambda_j\| + O_p(\|W_T - W\|^2 + N^{-1} + T^{-1}).$$

(ii)

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s' W u_t/N f_t'\|_F = O_p(\|W_T - W\|\sqrt{\frac{\log N}{T}} + \sqrt{\frac{\log N}{NT}} + \frac{\log N}{T}),$$

(iii) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}f_t'\|_F = O_p(\|W_T - W\|\sqrt{\frac{\log N}{T}} + \sqrt{\frac{\log N}{NT}} + \sqrt{\frac{\log N}{T^2}})$, *(iv)* $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}f_t'\|_F = O_p(\|W_T - W\|\frac{1}{\sqrt{N}} + \frac{1}{N} + \frac{1}{\sqrt{TN}})$.

Proof. Write $c_T = \|W_T - W\|\sqrt{\log N/T} + \sqrt{\log N/(NT)} + \sqrt{\log N/T^2}$.

(i) It suffices to find the rate of $a \equiv \|\frac{1}{T}\sum_{t=1}^{T}(\widehat{f}_t - H_W f_t)(H_W f_t - \widehat{f}_t)' H_W'^{-1}\lambda_j\|$.

In fact,

$$a \leq \frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t - H_W f_t\|^2\|H_W'^{-1}\lambda_j\| = O_p(\|W_T - W\|^2 + N^{-1} + T^{-1}).$$

(ii) Using the fact that $\frac{1}{T}\sum_{t=1}^T \|\widehat{f}_t - H_W f_t\|^2 = O_p(\|W_T - W\|^2 + N^{-1} + T^{-1})$, we have

$$\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s u_s' W u_t / N f_t' \|_F \leq \|H_W \frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T f_s u_s' W u_t f_t' / N\|_F + O_p(c_T).$$

The first term on the right hand side is $O_p((\log N)/T)$, which yields the result.

(iii) We have

$$\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s \eta_{st} f_t'\|_F \leq \|H_W \frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T f_s \eta_{st} f_t'\|_F + O_p(c_T)$$

$$= \|H_W \frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T f_s f_s' \Lambda' W u_t f_t' / N\|_F + O_p(c_T) = O_p(\frac{1}{\sqrt{NT}}) + O_p(c_T)$$

where we used the assumption that $\frac{1}{\sqrt{NT}}\sum_{t=1}^T \Lambda' W u_t f_t' = O_p(1)$.

(iv) We have

$$\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s \theta_{st} f_t'\|_F$$

$$\leq \|H_W \frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T f_s \theta_{st} f_t'\|_F + \|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T (\widehat{f}_s - H_W f_s) f_t' \Lambda' W u_s f_t' / N\|_F$$

$$\equiv a + b.$$

By the Cauchy-Schwarz inequality,

$$b \leq (\frac{1}{T}\sum_{s=1}^T \|\widehat{f}_s - H_W f_s\|^2)^{1/2} (\frac{1}{T}\sum_{s=1}^T g_s^2)^{1/2},$$

where $g_s^2 = \|\frac{1}{T}\sum_{t=1}^T f_t' \Lambda' W u_s f_t' / N\|_F^2$. By the assumption that $E\|\Lambda' W u_t / N\|^4 = O(N^{-2})$ and Cauchy-Schwarz inequality, $Eg_s^2 = O(1)(E\|\Lambda' W u_s / N\|^4)^{1/2} = O(N^{-1})$. Thus $b = O_p(\|W_T - W\|/\sqrt{N} + N^{-1} + 1/\sqrt{TN})$. Also,

$$a = \|H_W \frac{1}{T}\sum_{s=1}^T f_s u_s' W \Lambda / N \frac{1}{T}\sum_{t=1}^T f_t f_t'\|_F \leq O_p(1) \|\frac{1}{NT}\sum_{s=1}^T f_s u_s' W \Lambda\|.$$

By the assumption that $\frac{1}{\sqrt{NT}}\sum_{t=1}^T \Lambda' W u_t f_t' = O_p(1)$, $a = O_p(1/\sqrt{NT})$.

$\square$

Lemma A.14. *For each $i \leq N$,*

(i) $\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T \widehat{f}_s (Eu_s' W u_t) u_{it} / N\| = O_p(1/T)$,

(ii)

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(u_s'Wu_t-Eu_s'Wu_t)u_{it}/N\|=O_p(N^{-1/2}\|W_T-W\|+1/N+1/\sqrt{NT}),$$

(iii) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\eta_{st}u_{it}\|=O_p(1/\sqrt{NT}+1/N)$.
(iv) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s\theta_{st}u_{it}\|=O_p(\|W_T-W\|/\sqrt{T}+1/\sqrt{NT}+1/T)$.

Proof. (i)

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(Eu_s'Wu_t)u_{it}/N\|\le O_p(T^{-1/2})\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(Eu_s'Wu_t)/N\|,$$

which can be further bounded using the Cauchy-Schwarz inequality. Hence the right hand side is

$$O_p(T^{-1/2})(\frac{1}{T}\sum_{s=1}^{T}|E(u_s'Wu_t)/N|^2)^{1/2}=O_p(T^{-1/2})(\frac{1}{T}\sum_{s=1}^{T}|E(u_s'Wu_t)/N|)^{1/2}=O_p(T^{-1}).$$

(ii) The objective is bounded by

$$\|H_W\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}f_s(u_s'Wu_t-Eu_s'Wu_t)u_{it}/N\|+O_p(N^{-1/2}\|W_T-W\|+N^{-1}+(NT)^{-1/2}).$$

By the Cauchy-Schwarz inequality, the first term is bounded by
$O_p(1)[\frac{1}{T}\sum_{t=1}^{T}(\frac{1}{T}\sum_{s=1}^{T}f_s(u_s'Wu_t-Eu_s'Wu_t)/N)^2]^{1/2}=O_p(1/\sqrt{NT})$, where the last equality follows from the assumption that
$E(\frac{1}{\sqrt{NT}}\sum_{s=1}^{T}f_s(u_s'Wu_t-Eu_s'Wu_t))^2=O(1)$.
(iii) The objective is bounded by $a+b$, where $a=\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_sf_s'\Lambda'WEu_tu_{it}/N\|$ and $b=\|\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_sf_s'\frac{1}{T}\sum_{t=1}^{T}(\Lambda'Wu_tu_{it}-\Lambda'W_TEu_tu_{it})/N\|$. Because $\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}(\Lambda'Wu_tu_{it}-\Lambda'W_TEu_tu_{it})=O_p(1)$, $b=O_p(1/\sqrt{NT})$. Let $(\Lambda'W)_j$ denote the jth column of $\Lambda'W$. Then

$$a\le O_p(1)\|\frac{1}{N}\sum_{j=1}^{N}(\Lambda'W)_jEu_{jt}u_{it}\|\le O_p(\max_{j\le N}\|(\Lambda'W)_j\|)\frac{1}{N}\sum_{j=1}^{N}|Eu_{jt}u_{it}|=O_p(1/N).$$

(iv) The objective is bounded by

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}(\widehat{f}_s-f_s)\theta_{st}u_{it}\|+\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}f_s\theta_{st}u_{it}\|.$$

For each fixed $i \leq N$, it can be shown that the first term is $O_p(\|W_T - W\|/\sqrt{T} + 1/\sqrt{NT} + 1/T)$ and the second term is bounded by $O_p(1/T)$.

$\square$

Lemma A.15. *For each $i \leq N$,*
(i)

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s'(W_T - W)u_t u_{it}/N\| = O_p(\|W_T - W\|(\|W_T - W\| + \frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}})),$$

(ii) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})u_{it}\| = O_p(\|W_T - W\|(\sqrt{\frac{\log N}{T}} + 1/\sqrt{N}))$.
(iii) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})u_{it}\| = O_p(\|W_T - W\|/\sqrt{T})$.
(iv) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s u_s'(W_T - W)u_t f_t'/N\|_F = O_p(\|W_T - W\|\sqrt{\log N/T})$.
(v) $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\eta_{st} - \widehat{\eta}_{st})f_t'\|_F = O_p(\|W_T - W\|\sqrt{\log N/T})$.
(vi)

$$\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})f_t'\|_F = O_p(\|W_T - W\|(\|W_T - W\| + \sqrt{\log N/T} + 1/\sqrt{N}).$$

Proof. (i) The result follows from the rate of $T^{-1}\sum_{t=1}^{T}\|\widehat{f}_t - f_t\|^2$ and that $\|\frac{1}{T}\sum_{s=1}^{T} f_s u_s'\| = O_p(\sqrt{N(\log N)/T})$.
(ii) The objective is bounded by $a + b$, where

$$a = \|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s f_s'\Lambda'(W_T - W)Eu_t u_{it}/N\|$$

and

$$b = \|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\widehat{f}_s f_s'\Lambda'(W_T - W)(u_t u_{it} - Eu_t u_{it})/N\|.$$

We first bound b. Since for each $i \leq N$, $\|\frac{1}{T}\sum_{t=1}^{T}(u_t u_{it} - Eu_t u_{it})\| = O_p(\sqrt{N\log N/T})$. Thus $b = O_p(\|W_T - W\|\sqrt{\log N/T})$. On the other hand, since $\|\Sigma_u\|_1$ is bounded,

$$a = O_p(\|W_T - W\|\|\frac{Eu_t u_{it}}{\sqrt{N}}\|) \leq O_p(\|W_T - W\|)(\frac{1}{N}\max_{j\leq N}\sum_{j=1}^{N}|Eu_{jt}u_{it}|^2)^{1/2}$$

$$= O_p(\|W_T - W\|/\sqrt{N}).$$

(iii) Since $\|\frac{1}{T}\sum_{t=1}^T f_t u_{it}\| = O_p(1/\sqrt{T})$ for each fixed $i \leq N$, we have,

$$\|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\widehat{f}_s(\theta_{st} - \widehat{\theta}_{st})u_{it}\| = \|\frac{1}{T}\sum_{t=1}^T \frac{1}{T}\widehat{f}_s f_t'\Lambda'(W_T - W)u_s u_{it}/N\|$$

$$= O_p(\|W_T - W\|/\sqrt{T}).$$

(iv)(v) The fact that $\|\frac{1}{T}\sum_{t=1}^T u_t f_t'\|_F = O_p(\sqrt{N\log N/T})$ yields the result.
(vi) By the triangular inequality and the rate for $\frac{1}{T}\sum_{s=1}^T \|\widehat{f}_s - H_W f_s\|^2$, the objective is bounded by

$$\|H_W \frac{1}{T}\sum_{t=1}^T \frac{1}{T}\sum_{s=1}^T f_s f_t'\Lambda'(W_T - W)u_s f_t'/N\|_F + O_p(\|W_T - W\|(\|W_T - W\| + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{T}})).$$

It then follows from $\|\frac{1}{T}\sum_{s=1}^T f_s u_s'\|_F = O_p(\sqrt{N\log N/T})$ that the first term is $O_p(\|W_T - W\|\sqrt{\log N/T})$. $\square$

B Proofs for Section 4

B.1 Proof of Proposition 4.1

Recall that for $i \leq N$, $\xi_i = (\Lambda'\Sigma_u^{-1})_i$, and $e_t = \Sigma_u^{-1}u_t$.

In fact, $\|\frac{1}{\sqrt{N}}\Lambda'(\widehat{\Sigma}_u^{-1} - \Sigma_u^{-1})u_t\| = \|\frac{1}{\sqrt{N}}\Lambda'\widehat{\Sigma}_u^{-1}(\Sigma_u - \widehat{\Sigma}_u)\Sigma_u^{-1}u_t\|$, which is bounded by

$$\|\frac{1}{\sqrt{N}}\Lambda'(\widehat{\Sigma}_u^{-1} - \Sigma_u^{-1})(\Sigma_u - \widehat{\Sigma}_u)\Sigma_u^{-1}u_t\| + \|\frac{1}{\sqrt{N}}\Lambda'\Sigma_u^{-1}(\Sigma_u - \widehat{\Sigma}_u)\Sigma_u^{-1}u_t\| \equiv a + b.$$

It follows from Fan et al. (2013) that $\|\widehat{\Sigma}_u^{-1} - \Sigma_u^{-1}\| = O_p(m_N\omega_T^{1-q}) = \|\widehat{\Sigma}_u - \Sigma_u\|$, and hence $a = O_p(\sqrt{N}m_N^2\omega_T^{2-2q})$.

For $\Lambda'\Sigma_u^{-1} = (\xi_1, ..., \xi_N)$, and $\Sigma_u^{-1}u_t = (e_{1t}, ..., e_{Nt})'$, we have

$$\begin{aligned} b &= \|\frac{1}{\sqrt{N}}\sum_{i,j}\xi_i(\Sigma_{u,ij} - \widehat{\sigma}_{u,ij})e_{jt}\| \leq \|\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i(\sigma_{u,ii} - \widehat{\sigma}_{u,ii})e_{it}\| \\ &\quad + \|\frac{1}{\sqrt{N}}\sum_{(i,j)\in S_L}\xi_i(\Sigma_{u,ij} - \widehat{\sigma}_{u,ij})e_{jt}\| + \|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\xi_i(\Sigma_{u,ij} - \widehat{\sigma}_{u,ij})e_{jt}\| \\ &\equiv b_1 + b_2 + b_3. \end{aligned}$$

We now bound $b_i, i \leq 3$, keeping in mind that

$$\widehat{\sigma}_{u,ii} = \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}^2, \quad \widehat{\sigma}_{u,ij} = s_{ij}(\frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}\widehat{u}_{jt}) \text{ for } i \neq j,$$

where $\widehat{u}_{it}$ is estimated using the regular PC method as in Bai (2003).

First, by the assumption that $\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T}(u_{is}^2 - Eu_{is}^2)\xi_i e_{it} = o_p(1)$, the triangular inequality implies that b_1 is bounded by

$$b_1 = \|\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T}(\widehat{u}_{is}^2 - Eu_{is}^2)\xi_i e_{it}\| \leq \|\frac{2}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{u}_{is} - u_{is})u_{is}\xi_i e_{it}\| + o_p(1)$$

Let H_I denote H_W when $W_T = I_r$ is used as the weight matrix, where the subscript I denotes the "identity weight matrix". Let $\widehat{f}_t^I$ and $\widehat{\lambda}_j^I$ denote the regular PC estimators for the transformed factors and loadings as in Stock and Watson (2002), which correspond to the weighted PC estimators with $W_T = W = I_r$. As shown in Bai (2003)'s Appendix C,

$$u_{it} - \widehat{u}_{it} = (\widehat{f}_t^I - H_I f_t)' H_I^{'-1}\lambda_i + f_t' H_I'(\widehat{\lambda}_i^I - H_I^{'-1}\lambda_i) + (\widehat{f}_t^I - H_I f_t)'(\widehat{\lambda}_i^I - H_I^{'-1}\lambda_i). \quad \text{(B.1)}$$

Thus

$$\|\frac{2}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{u}_{is} - u_{is})u_{is}\xi_i e_{it}\| \leq b_{11} + b_{12} + O_p(\frac{1}{\sqrt{N}} + \frac{\sqrt{N}\log N}{T})$$

where $b_{11} = \|\frac{2}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T}(\widehat{f}_s^I - H_I f_s)' H_I^{'-1}\lambda_i u_{is}\xi_i e_{it}\|$,

$$b_{12} = \|\frac{2}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T} f_s' H_I'(\widehat{\lambda}_i^I - H_I^{'-1}\lambda_i)u_{is}\xi_i e_{it}\| = O_p(\frac{\sqrt{N}\log N}{T} + \sqrt{\frac{\log N}{T}}).$$

It follows from Lemma B.3 that $b_{11} = o_p(1)$. This implies $b_1 = o_p(1)$.

We now bound b_2. Note that $\sum_{(i,j)\in S_L}|\Sigma_{u,ij}| = O(1)$. Since $\max_{i\leq N}\|\xi_i\|$ and $\max_{j\leq N}|e_{jt}|$ are both bounded, there is $C > 0$ so that

$$b_2 \leq \frac{C}{\sqrt{N}}(\sum_{(i,j)\in S_L}|\Sigma_{u,ij}| + \sum_{(i,j)\in S_L}|\widehat{\sigma}_{u,ij}|) = O(\frac{1}{\sqrt{N}}) + \frac{C}{\sqrt{N}}\sum_{(i,j)\in S_L}|\widehat{\sigma}_{u,ij}|.$$

In addition, for any $\epsilon > 0$ and any $M > 0$,

$$P(\frac{1}{N}\sum_{(i,j)\in S_L}|\widehat{\sigma}_{u,ij}| > M\omega_T^2) \leq P(\exists (i,j) \in S_L, \widehat{\sigma}_{u,ij} \neq 0) \leq \epsilon.$$

This implies that $\frac{C}{\sqrt{N}}\sum_{(i,j)\in S_L}|\widehat{\sigma}_{u,ij}| = O_p(\omega_T^2\sqrt{N})$. Hence

$$b_2 = O_p(\frac{\sqrt{N}\log N}{T} + \frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}}) = O_p(\omega_T^2\sqrt{N}).$$

Finally, it follows from the triangular inequality and Lemma B.4 that $b_3 = o_p(1)$. Hence $b = o_p(1)$. Hence $a + b = O_p(\sqrt{N}m_N^2\omega_T^{2-2q}) + o_p(1) = o_p(1)$.

B.2 Proof of Theorems 4.1, 4.2, and 4.3

For Theorem 4.1, let $W = \Sigma_u^{-1}$, define $Q_e = V^{1/2}\Gamma_e'\Sigma_\Lambda^{-1/2}$, and $\Lambda'\Sigma_u^{-1}\Lambda/N \to \Sigma_\Lambda$. Here $\Gamma_e'\Gamma_e = I_r$. By Theorem 3.1, $NVar_1^{-1/2}(\widehat{f}_t^e - H_e f_t) \to^d \mathcal{N}(0,1)$, where $Var_1 = V^{-1}Q_e\Lambda'\Sigma_u^{-1}\Lambda Q_e'V^{-1}$. Let

$$G = \Lambda'\Sigma_u^{-1}\Lambda, Q_1 = V^{1/2}\Gamma_e'(G/N)^{-1/2}$$

and $Var_2 = V^{-1}Q_1\Lambda'\Sigma_u^{-1}\Lambda Q_1'V^{-1} = NV^{-1}$. In addition, $(Var_1 - Var_2)/N = o(1)$. Thus by Slusky's theorem,

$$\sqrt{N}(\widehat{f}_t^e - H_e f_t) \to^d \mathcal{N}(0, V^{-1}).$$

The limiting distribution for $\widehat{\lambda}_j^e$ follows from Theorem 3.1. For $W = \Sigma_u^{-1}$, the limiting distribution of the estimated common component follows from Theorem 3.1 and $\Lambda'\Sigma_u^{-1}\Lambda/N \to \Sigma_{\Lambda.e}$.

Theorem 4.2 is a corollary of Theorem 3.2.

As for Theorem 4.3. note that for any W, $\Xi_W = \Sigma_\Lambda^{-1}\Lambda'W\Sigma_u W\Lambda\Sigma_\Lambda^{-1}/N$. Define $\Sigma_W = (\Lambda'W\Lambda/N)^{-1}\Lambda'W\Sigma_u W\Lambda(\Lambda'W\Lambda/N)^{-1}/N$. Then $\Xi_W - \Sigma_W = o(1)$. It suffices to show $\Sigma_W - \Xi_e$ is semi-positive definite, where $\Xi_e = (\Lambda'\Sigma_u^{-1}\Lambda/N)^{-1}$. Equivalently, we show that

$$f(W) = \Lambda'\Sigma_u^{-1}\Lambda - \Lambda'W\Lambda(\Lambda'W\Sigma_u W\Lambda)^{-1}\Lambda'W\Lambda$$

is semi-positive definite. In fact, let

$$\Delta = I_N - \Sigma_u^{1/2}W\Lambda(\Lambda'W\Sigma_u W\Lambda)^{-1}\Lambda'W\Sigma_u^{1/2},$$

then

$$\begin{aligned} f(W) &= \Lambda'\Sigma_u^{-1/2}(I_N - \Sigma_u^{1/2}W\Lambda(\Lambda'W\Sigma_u W\Lambda)^{-1}\Lambda'W\Sigma_u^{1/2})\Sigma_u^{-1/2}\Lambda \\ &= \Lambda'\Sigma_u^{-1/2}\Delta\Sigma_u^{-1/2}\Lambda. \end{aligned}$$

It is straightforward to show that $\Delta^2 = \Delta$. Hence Δ is semi-positive definite, which implies that $f(W)$ is semi-positive definite.

B.3 Proof of Theorem 4.4

Proof. First of all, since $\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t^e - H_e f_t\|^2 = O_p(m_N^2\omega_T^{2-2q})$, and $\max_{i\leq N}\|H_e^{'-1}\lambda_i - \widehat{\lambda}_i\| = O_p(m_N\omega_T^{1-q})$, we have

$$
\begin{aligned}
&\frac{1}{NT}\|\widehat{F}^e\widehat{\Lambda}^{e'} - F\Lambda\|_F^2 \leq \frac{2}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\|H_e^{'-1}\lambda_i - \widehat{\lambda}_i^e\|^2\|H_e f_t\|^2 \\
&+\frac{2}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\|\widehat{\lambda}_i^e\|^2\|H_e f_t - \widehat{f}_t^e\|^2 \\
\leq\ & 4\max_{i\leq N}\|H_e^{'-1}\lambda_i - \widehat{\lambda}_i\|^2\frac{1}{T}\sum_{t=1}^{T}(\|H_e f_t\|^2 + \|H_e f_t - \widehat{f}_t\|^2) \\
&+4\max_{i\leq N}\|H_e^{'-1}\lambda_i\|^2\frac{1}{T}\sum_{t=1}^{T}\|H_e f_t - \widehat{f}_t\|^2 = O_p(m_N^2\omega_T^{2-2q}).
\end{aligned}
\tag{B.2}
$$

Let $\tilde{V}_e^{-1}$ be the left hand side of (4.6). It then suffices to show

$$
\widehat{V}_e^{-1} - \tilde{V}_e^{-1} = o_p(1).
$$

It follows from (B.2) and $\|F\Lambda'\|_F = O_p(\sqrt{NT})$, $\|\Sigma_u^{-1}\| = O(1)$ that $\widehat{V}_e^{-1} - \tilde{V}_e^{-1} = O_p(m_N\omega_T^{1-q})$. Let $\mathrm{HAC}(f_t u_{jt})$ and $\mathrm{HAC}(\widehat{f}_t^e\widehat{u}_{jt})$ be the HAC covariance estimators of Newey and West (1987), based on $\{f_t u_{jt}\}$ and $\widehat{f}_t^e\widehat{u}_{jt}$ respectively, where

$$
\mathrm{HAC}(\alpha_t) = \frac{1}{T}\sum_{t=1}^{T}\alpha_t\alpha_t' + \sum_{l=1}^{K}(1 - \frac{l}{K+1})\frac{1}{T}\sum_{t=l+1}^{T}(\alpha_t\alpha_{t-l}' + \alpha_{t-l}\alpha_t').
$$

Then $\frac{1}{T}\sum_{t=1}^{T}\|\widehat{f}_t^e - H_e f_t\|^2 = O_p(m_N^2\omega_T^{2-2q})$ and $\max_{i\leq N}\|H_e^{'-1}\lambda_i - \widehat{\lambda}_i\| = O_p(m_N\omega_T^{1-q})$ imply $\max_{j\leq N}\frac{1}{T}\sum_{t=1}^{T}(\widehat{u}_{jt} - u_{jt})^2 = O_p(m_N^2\omega_T^{2-2q})$, and thus

$$
\widehat{\Psi}_j = H_e\mathrm{HAC}(f_t u_{jt})H_e' + O_p(Km_N\omega_T^{1-q})
$$

It is guaranteed by (4.3) that $m_N\omega_T^{1-q} = o(N^{-1/4})$. Hence the assumption $K = o(N^{1/4})$ implies $\widehat{\Psi}_j = H_e\mathrm{HAC}(f_t u_{jt})H_e' + o_p(1)$. It follows from Newey and West (1987) that $\mathrm{HAC}(f_t u_{jt})$ consistently estimates Φ_j. Hence $\mathrm{HAC}(\widehat{f}_t^e\widehat{u}_{jt}) - H_e\Phi_j H_e' = o_p(1)$. By Lemma A.5, $H_e \to^p Q_e^{'-1}$, which gives the consistency of $\mathrm{HAC}(\widehat{f}_t^e\widehat{u}_{jt})$.

In addition, let

$$\tilde{\Theta}_{1T} = \frac{1}{NT^2}\lambda_i' H_e^{-1}\widehat{V}^{-1}\widehat{F}^{e\prime}F\Lambda'\Sigma_u^{-1}\Lambda F'\widehat{F}^{e\prime}\widehat{V}^{-1}H_e'^{-1}\lambda_i.$$

Since $\frac{1}{T}\widehat{F}^{e\prime}\widehat{F}^{e} = I_r$, $\|\widehat{\Sigma}_u^{-1} - \Sigma_u^{-1}\| = o_p(1)$ and $\frac{1}{NT}\|\widehat{F}^e\widehat{\Lambda}^{e\prime} - F\Lambda\|_F^2 = o_p(1)$, we have $\tilde{\Theta}_{1T} - \widehat{\Theta}_{1i} = o_p(1)$. By Lemma A.5, if we replace $\frac{1}{T}\widehat{F}^{e\prime}F$ with Q_e and H_e with $\widehat{V}^{-1}Q_e\Sigma_{\Lambda,e}$, the estimation error introduced by such replacements is negligible. This gives $\tilde{\Theta}_{1T} = \lambda_i'\Xi_e\lambda_i + o_p(1)$. Finally, since $\mathrm{HAC}(\widehat{f}_t^e\widehat{u}_{jt}) - H_e\Phi_j H_e' = o_p(1)$ and $\widehat{f}_t^e - H_e f_t = o_p(1)$, we have

$$\widehat{\Theta}_{2,it} = f_t' H_e' H_e \Phi_j H_e' H_e f_t + o_p(1).$$

By Lemma A.3, $H_e' H_e = \mathrm{cov}(f_t)^{-1} + o_p(1)$. Hence $\widehat{\Theta}_{2,it} \to^p f_t'\Omega_i f_t$.

□

B.4 Proof of Lemma 4.1

We respectively show that for each term of (i)-(iv), its mean and variance are both $o(1)$. Because $\{u_t\}_{t\leq T}$ is serially independent and $\sum_{(i,j)\in S_U} 1 = O(N)$ due to the sparsity, so each of the four terms has mean $O(\frac{\sqrt{N}}{T}) = o(1)$. Let us now study their variances. For notational simplicity, we assume $\dim(\xi_i) = \dim(\lambda_i) = 1$. Recall that ξ_i is the ith column of $\Lambda'\Sigma_u^{-1}$.

(i) Let $w_{is} = u_{is}^2 - Eu_{is}^2$. The variance equals

$$\frac{1}{T^2N}\sum_{i=1}^N \mathrm{var}(\sum_{s=1}^T w_{is}\xi_i e_{it}) + \frac{1}{T^2N}\sum_{i\neq j}\mathrm{cov}(\sum_{s=1}^T w_{is}\xi_i e_{it}, \sum_{s=1}^T w_{js}\xi_j e_{jt})$$

$$\equiv A_1 + A_2.$$

The first term is upper bounded by, due to the Cauchy Schwarz inequality:

$$\frac{1}{T^2N}\sum_{i=1}^N E(\sum_{s=1}^T w_{is}\xi_i e_{it})^2 \leq \frac{1}{TN}\sum_{i=1}^N [E(\frac{1}{\sqrt{T}}\sum_{s=1}^T w_{is})^4]^{1/2}[Ee_{it}^4]^{1/2}\xi_i^2.$$

Note that both $E(\frac{1}{\sqrt{T}}\sum_{s=1}^T w_{is})^4$ and Ee_{it}^4 are bounded uniformly in i. Hence $A_1 = O(\frac{1}{T})$.

For A_2, because of the serial independence and $Ew_{is} = 0$, we have

$$A_2 = \frac{1}{T^2N}\sum_{i\neq j}\sum_{s\neq t}\mathrm{cov}(w_{is}\xi_i e_{it}, w_{js}\xi_j e_{jt}) + \frac{1}{T^2N}\sum_{i\neq j}\mathrm{cov}(w_{it}\xi_i e_{it}, w_{jt}\xi_j e_{jt}).$$

The second term on the right is $O(\frac{N}{T^2}) = o(1)$. The first term is

$\frac{1}{T^2N}\sum_{i\neq j}\sum_{s\neq t}(Ew_{is}w_{js})(Ee_{it}e_{jt})\xi_i\xi_j$. Note that $Ee_{it}e_{jt}=(\Sigma_u^{-1})_{ij}$. This term is bounded by

$$\frac{TN}{T^2N}\max_{ijs}|Ew_{is}w_{js}||\xi_i\xi_j|\|\Sigma_u^{-1}\|_1=O(\frac{1}{T}).$$

Therefore, $A_1+A_2=o(1)$, which implies the desired result.

(ii) Let $w_{ijs}=u_{is}u_{js}-Eu_{is}u_{js}$. The term's variance equals

$$\frac{1}{N^3T^2}\sum_{i=1}^N \mathrm{var}(\sum_{s=1}^T\sum_{j=1}^N w_{ijs}\lambda_j\lambda_i e_{it}\xi_{ik})$$

$$+\frac{1}{N^3T^2}\sum_{i\neq l}\mathrm{cov}(\sum_{s=1}^T\sum_{j=1}^N w_{ijs}\lambda_j\lambda_i e_{it}\xi_{ik},\sum_{s=1}^T\sum_{j=1}^N w_{ljs}\lambda_j\lambda_l e_{lt}\xi_{lk}).$$

Let us call the above two terms B_1 and B_2 respectively. Due to the serial independence and $Ew_{isj}=0$,

$$B_1\leq\frac{1}{N^3T^2}\sum_{i=1}^N E(\sum_{s=1}^T\sum_{j=1}^N w_{ijs}\lambda_j\lambda_i e_{it}\xi_{ik})^2$$

$$=\frac{1}{N^3T^2}\sum_{i,j_1,j_2\leq N}\lambda_{j_1}\lambda_{j_2}\lambda_i^2\xi_{ik}^2\sum_{s=1}^T Ew_{ij_1s}w_{ij_2s}e_{it}^2=O(\frac{1}{T}).$$

On the other hand, because $\|\Sigma_u^{-1}\|=O(1)$,

$$\begin{aligned}
B_2&=\frac{1}{N^3T^2}\sum_{i\neq l}\sum_{j_1=1}^N\sum_{j_2=1}^N\sum_{s\neq t}\mathrm{cov}(w_{ij_1s}\lambda_{j_1}\lambda_i e_{it}\xi_{ik},w_{lj_2s}\lambda_{j_2}\lambda_l e_{lt}\xi_{lk})\\
&+\frac{1}{N^3T^2}\sum_{i\neq l}\sum_{j_1=1}^N\sum_{j_2=1}^N\mathrm{cov}(w_{ij_1t}\lambda_{j_1}\lambda_i e_{it}\xi_{ik},w_{lj_2t}\lambda_{j_2}\lambda_l e_{lt}\xi_{lk})\\
&=\frac{1}{N^3T^2}\sum_{i\neq l}\sum_{j_1=1}^N\sum_{j_2=1}^N\sum_{s\neq t}E(w_{ij_1s}w_{lj_2s})E(e_{it}e_{lt})\xi_{lk}\lambda_{j_1}\lambda_i\xi_{ik},\lambda_{j_2}\lambda_l+O(\frac{N^4}{N^3T^2})\\
&\leq O(\frac{N^2T}{N^3T^2})\sum_{i\neq l}|E(e_{it}e_{lt})|+o(1)\leq O(\frac{N}{TN})\|\Sigma_u^{-1}\|_1+o(1)=o(1).
\end{aligned}$$

Thus $B_1+B_2=o(1)$, which implies the result.

The variances of terms in (iii) and (iv) can be proved to be $o(1)$ in the same way, so we omit the proofs.

B.5 Technical lemmas

Lemma B.1. *For each $t \le T$,*
(i) $\|\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}(u_s'u_l - Eu_s'u_l)\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_i e_{it}\| = o_p(1)$.
(ii) $\|\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}(Eu_s'u_l)\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_i e_{it}\| = o_p(1)$.
(iii) $\|\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}f_l'\Lambda'u_s\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_i e_{it}\| = o_p(1)$.
(iv) $\|\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}f_s'\Lambda'u_l\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_i e_{it}\| = o_p(1)$.

Proof. We can replace $\widehat{f}_l^{I'}$ in each stated term with f_l', because as shown by Fan et al. (2013), $\frac{1}{T}\sum_{l=1}^T\|\widehat{f}_l^I - f_l\|^2 = O_p(\omega_T)$. Thus by Cauchy-Schwarz inequality, such a replacement will introduce an error $O_p(\omega_T)$.
(i) By the Cauchy-Schwarz inequality, the objective is bounded by $O_p(\omega_T)$ plus

$$\frac{1}{\sqrt{T}}\left[\frac{1}{T}\sum_{s=1}^T\|\frac{1}{\sqrt{NT}}\sum_{l=1}^T f_l'(u_s'u_l - Eu_s'u_l)\widehat{V}^{-1}H_I'^{-1}\|^2\right]^{1/2}\left[\frac{1}{T}\sum_{s=1}^T(\frac{1}{N}\sum_{i=1}^N\|\lambda_i u_{is}\xi_i e_{it}\|)^2\right]^{1/2}.$$

The second $(\cdot)^{1/2}$ term is $O_p(1)$. By the assumption that

$$E\|\frac{1}{\sqrt{TN}}\sum_{l=1}^T f_l'(u_s'u_l - Eu_s'u_l)\|^2 = O(1),$$

the first term is $O_p(1/\sqrt{NT})$, which yields the result.
(ii) The objective is bounded by

$$O_p(\frac{1}{N\sqrt{N}T^2})\sum_{i=1}^N\sum_{s,l}^T\|f_l'Eu_s'u_l\|\|\lambda_i u_{is}\xi_i e_{it}\| + o_p(1).$$

Note that $E\sum_{l=1}^T\|f_l'Eu_s'u_l/N\| = O(1)$ by the strong mixing condition. This gives the result.
(iii) The term in $\|.\|$ is an $r\times 1$ vector. Let a_k denote its kth element, $k \le r$. Then $a_k = \mathrm{tr}(a_k) = \frac{1}{NT^2\sqrt{N}}\sum_l\sum_i\sum_s f_l'\Lambda'u_s\widehat{f}_l'\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_{ik}e_{it}$. Using the inequality that $|\mathrm{tr}(AB)| \le \|A\|_F\|B\|_F$, we have

$$\begin{aligned}|a_k| &= |\mathrm{tr}(a_k)| = |\mathrm{tr}(\frac{1}{T}\sum_{l=1}^T f_l\widehat{f}_l^{I'}\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\widehat{V}^{-1}H_I'^{-1}\lambda_i u_{is}\xi_{ik}e_{it}u_s'\Lambda'| \\ &\le \|\frac{1}{NT\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^T\lambda_i u_{is}\xi_{ik}e_{it}u_s'\Lambda'\|_F O_p(1) \qquad \text{(B.3)}\end{aligned}$$

By the assumption that

$$\|\frac{1}{NT\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{j=1}^{N}(u_{is}u_{js}-Eu_{is}u_{js})\xi_{ik}e_{it}\lambda_i\lambda_j'\|_F=o_p(1)$$

and $\max_{i\leq N}\sum_{j=1}^{N}|Eu_{js}u_{is}|=O(1)$, it follows from the triangular inequality that $|a_k|=o_p(1)$. Since there are finitely many a_k $(k\leq r)$, the desired result follows.
(iv) It follows directly from the rate of $\|\frac{1}{T}\sum_{s=1}^{T}f_su_s'\|=O_p(\sqrt{N(\log N)/T})$.

$\square$

Lemma B.2. *For S_U in the partition $\{(i,j):i,j\leq N\}=S_L\cup S_U$, and any $t\leq T$,*
(i) $\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}u_{is}\lambda_j'H_I^{-1}\sum_{l=1}^{T}\widehat{f}_l^Iu_l'u_s\xi_ie_{jt}=o_p(1)$,
(ii) $\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}u_{is}\lambda_j'H_I^{-1}\sum_{l=1}^{T}\widehat{f}_l^If_l'\sum_{v=1}^{N}\lambda_vu_{vs}\xi_ie_{jt}=o_p(1)$,
(iii) $\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}u_{is}\lambda_j'H_I^{-1}\sum_{l=1}^{T}\widehat{f}_l^If_s'\sum_{v=1}^{N}\lambda_vu_{vl}\xi_ie_{jt}=o_p(1)$.

Proof. (i) The term of interest is bounded by $a+b$, where

$$\begin{aligned}a&=\|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U,i\neq j}u_{is}\lambda_j'\sum_{l=1}^{T}f_lu_l'u_s\xi_ie_{jt}\|,\\ b&=\|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U,i\neq j}u_{is}\lambda_j'H_I^{-1}\sum_{l=1}^{T}(\widehat{f}_l^I-H_If_l)u_l'u_s\xi_ie_{jt}\|.\end{aligned}$$

Here a is upper bounded by a_1+a_2, where

$$a_1=\|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U,i\neq j}u_{is}\lambda_j'\sum_{l=1}^{T}f_l(u_l'u_s-Eu_l'u_s)\xi_ie_{jt}\|,$$

and

$$a_2=\|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U,i\neq j}u_{is}\lambda_j'\sum_{l=1}^{T}f_l(Eu_l'u_s)\xi_ie_{jt}\|.$$

Note that a_1 and a_2 can be bounded in the same way as (i)(ii) of Lemma B.1 by the assumption that $\sum_{(i,j)\in S_U,i\neq j}1=O(N)$. We conclude that $a=o_p(1)$.

On the other hand, $b\leq b_1+b_2$ where

$$b_1=\|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U,i\neq j}u_{is}\lambda_j'H_I^{-1}\sum_{l=1}^{T}(\widehat{f}_l^I-H_If_l)(u_l'u_s-Eu_l'u_s)\xi_ie_{jt}\|,$$

and

$$b_2 = \|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{(i,j)\in S_U, i\neq j} u_{is}\lambda_j' H_I^{-1}\sum_{l=1}^{T}(\widehat{f}_l^I - H_I f_l)(Eu_l' u_s)\xi_i e_{jt}\|.$$

Using Cauchy-Schwarz inequality and the strong mixing condition, we conclude that $b = o_p(1)$.

(ii) The $k(\leq r)$th element of the object of interest is bounded by $d_1 + d_2$, where

$$d_1 = |\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}\lambda_j' H_I^{-1}\sum_{l=1}^{T}\widehat{f}_l^I f_l'\sum_{v=1}^{N}\lambda_v(u_{is}u_{vs} - Eu_{is}u_{vs})\xi_{ik}e_{jt}|,$$

$$d_2 = |\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}\lambda_j' H_I^{-1}\sum_{l=1}^{T}\widehat{f}_l^I f_l'\sum_{v=1}^{N}\lambda_v(Eu_{is}u_{vs})\xi_{ik}e_{jt}|.$$

$$\begin{aligned} d_1 &= |\mathrm{tr}(\frac{1}{\sqrt{N}NT^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}\sum_{l=1}^{T}\widehat{f}_l^I f_l'\sum_{v=1}^{N}\lambda_v(u_{is}u_{vs} - Eu_{is}u_{vs})\xi_{ik}e_{jt}\lambda_j' H_I^{-1})| \\ &\leq O_p(\frac{1}{\sqrt{N}NT})\|\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U}\sum_{v=1}^{N}(u_{is}u_{vs} - Eu_{is}u_{vs})\xi_{ik}e_{jt}\lambda_v\lambda_j'\| = o_p(1) \end{aligned}$$

by the assumption that

$$\frac{1}{NT\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\sum_{v=1}^{N}\sum_{s=1}^{T}(u_{is}u_{vs} - Eu_{is}u_{vs})\xi_{ik}e_{jt}\lambda_v\lambda_j' = o_p(1).$$

On the other hand, $d_2 \leq O_p(\frac{1}{N\sqrt{N}})\sum_{i\neq j,(i,j)\in S_U}\sum_{v=1}^{N}|\sigma_{u,iv}|$. Note that $\|\Sigma_{u0}\|_1 = O(1)$, thus $d_2 = O_p(N^{-1/2})$.

(iii) The object of interest is bounded by $e_1 + e_2$, where

$$e_1 = \|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U} u_{is}\lambda_j' H_I^{-1}\sum_{l=1}^{T}(\widehat{f}_l^I - H_I f_l)f_s'\sum_{v=1}^{N}\lambda_v u_{vl}\xi_i e_{jt}\|$$

$$e_2 = \|\frac{1}{N\sqrt{N}T^2}\sum_{s=1}^{T}\sum_{i\neq j,(i,j)\in S_U} u_{is}\lambda_j'\sum_{l=1}^{T} f_l f_s'\sum_{v=1}^{N}\lambda_v u_{vl}\xi_i e_{jt}\|.$$

Since $\max_{i\leq N}\|T^{-1}\sum_{t=1}^{T} f_t u_{it}\| = O_p(\sqrt{\log N/T})$, we conclude that
$e_1 + e_2 = o_p(1)$. □

Lemma B.3. *For each $t \leq T$,*
(i) $\frac{1}{T\sqrt{N}}\sum_{i=1}^{N}\sum_{s=1}^{T}(\widehat{f}_s^I - H_I f_s)' H_I^{'-1}\lambda_i u_{is}\xi_i e_{it} = o_p(1)$,

(ii) $\frac{1}{T\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\sum_{t=1}^T u_{is}\lambda_j' H_I^{-1}(\widehat{f}_s - H_I f_s)\xi_i e_{jt} = o_p(1)$,

Proof. The lemma follows immediately from (A.1) with $H_W = H_I$ and $W_T = I_N$, Lemmas B.1, B.2 and the triangular inequality. □

Let $R_{ij} = \frac{1}{T}\sum_{t=1}^T \widehat{u}_{it}\widehat{u}_{jt}$ where $\widehat{u}_{it}$ is the regular PC estimator of u_{it} as in Bai (2003). Then for the thresholding function $s_{ij}(\cdot)$, $\widehat{\sigma}_{u,ij} = s_{ij}(R_{ij})$. Recall that $\omega_T = \frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}}$.

Lemma B.4. *For each $t \leq T$,*
(i) $\|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\Sigma_{u,ij} - R_{ij})\xi_i e_{jt}\| = o_p(1)$,
(ii) $\|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\sigma}_{u,ij} - R_{ij})\xi_i e_{jt}\| = O_p(\sqrt{N}\omega_T^2) = o_p(1)$.

Proof. (i) Since $R_{ij} = T^{-1}\sum_{t=1}^T \widehat{u}_{it}\widehat{u}_{jt}$, the term of interest equals

$$\frac{2}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T u_{is}(\widehat{u}_{js} - u_{js})\xi_i e_{jt}$$
$$+\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T(\widehat{u}_{is} - u_{is})(\widehat{u}_{js} - u_{js})\xi_i e_{jt}$$
$$+\frac{1}{T\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\sum_{s=1}^T(u_{is}u_{js} - Eu_{is}u_{js})\xi_i e_{jt} \equiv a + b + c.$$

By the assumption $c = o_p(1)$. Also since $\max_{i\leq N}\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it} - u_{it})^2 = O_p(\omega_T^2)$ (e.g., Fan et al. (2013) Lemma C.11), by the assumption $\sum_{i\neq j,(i,j)\in S_U} 1 = O(N)$, and the Cauchy-Schwarz inequality, $b = O_p(\sqrt{N}\omega_T^2)$. We now work out the first term a. Again we use equality (D.1) for $\widehat{u}_{js} - u_{js}$. First,

$$\frac{2}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T u_{is}(\widehat{\lambda}_j^I - H_I^{-1\prime}\lambda_j)' H_I f_s \xi_i e_{jt}$$

$$\leq O_p(\frac{1}{\sqrt{N}})\max_{i\leq N}\|\xi_i\|\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^T u_{jt} f_t\|\max_{j\leq N}\|\widehat{\lambda}_j^I - H_I'^{-1}\lambda_j\|\sum_{i\neq j,(i,j)\in S_U}|e_{jt}|.$$

As $\max_{j\leq N}\|\widehat{\lambda}_j^I - H_I'^{-1}\lambda_j\| = O_p(\omega_T)$, $\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^T u_{jt}f_t\| = O_p(\sqrt{\log N/T})$,
$E\sum_{i\neq j,(i,j)\in S_U}|e_{jt}| = O(N)$, and $\max_i\|\xi_i\| = O(1)$,

$$\frac{2}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T u_{is}(\widehat{\lambda}_j^I - H_I^{-1\prime}\lambda_j)' H_I f_s \xi_i e_{jt} = O_p(\omega_T\sqrt{\frac{N\log N}{T}}) = o_p(1).$$

Also, $\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T u_{is}(\widehat{\lambda}_j^I - H_I^{-1'}\lambda_j)'(\widehat{f}_s^I - H_I f_s)\xi_i e_{jt}$ is bounded by

$$O_p(\frac{1}{\sqrt{N}})\max_{i\leq N}\|\widehat{\lambda}_i^I - H_I^{-1'}\lambda_i\|\left(\frac{1}{T}\sum_{t=1}^T\|\widehat{f}_t^I - H_I f_t\|^2\right)^{1/2}\sum_{i\neq j,(i,j)\in S_U}|e_{jt}| = O_p(\sqrt{N}\omega_T^2).$$

Finally, $\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{t=1}^T u_{is}\lambda_j' H_I^{-1}(\widehat{f}_s - H_I f_s)\xi_i e_{jt} = o_p(1)$, following from Lemma B.3. This implies $a = o_p(1)$. Combining the results above, we obtain the desired result.

(ii) By the definition of the thresholding function, $|s_{ij}(z) - z| \leq a\tau_{ij}^2$ when $|z| > b\tau_{ij}$. Hence $\|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\sigma}_{u,ij} - R_{ij})\xi_i e_{jt}\|$ is upper bounded by (recall $\sum_{(i,j)\in S_U}1 = O(N)$):

$$\begin{aligned}
&\|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U,|R_{ij}|>b\tau_{ij}}(s_{ij}(R_{ij}) - R_{ij})\xi_i e_{jt}\| \\
&+\|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U,|R_{ij}|\leq b\tau_{ij}}(s_{ij}(R_{ij}) - R_{ij})\xi_i e_{jt}\| \\
&\leq O_p(\sqrt{N}\omega_T^2) + \|\frac{1}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U,|R_{ij}|\leq b\tau_{ij}}(s_{ij}(R_{ij}) - R_{ij})\xi_i e_{jt}\| \\
&\equiv O_p(\sqrt{N}\omega_T^2) + v.
\end{aligned}$$

For any $M > 0$, and $\epsilon > 0$,

$$P(v > \sqrt{N}M\omega_T^2) \leq P(\exists(i,j)\in S_U, |R_{ij}| \leq b\tau_{ij}) < \epsilon$$

which yields $v = O_p(\sqrt{N}\omega_T^2)$. This yields the desired result.

□

C Proofs for panel data with interactive effects

Throughout the proof, we denote $w_{ij} = (\Sigma_u^{-1})_{ij}$. We first prove that the estimated covariance matrix is consistent. The following theorem extends the result of Fan et al. (2013) to the panel data model:

Theorem C.1. *Under the Assumptions 3.2, 3.3, 4.1, when $\|\Sigma_u^{-1}\|_1 = O(1)$, for $\omega_T = \sqrt{\frac{\log N}{T}} + \frac{1}{\sqrt{N}}$,*

$$\|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1 = O_p(m_N\omega_T^{1-q}) = \|\widetilde{\Sigma}_u - \Sigma_u\|_1.$$

Proof. Due to the $\sqrt{NT}$-consistency of $\hat{\beta}_0$ achieved by Bai (2009), it is not hard to show that when applying the PC method on $(Y_t - X_t\hat{\beta}_0)$ to estimate $\lambda_i' f_t$, the effect of estimating

β is asymptotically negligible. Hence the same proofs as those of Fan et al. (2013) yield, for $\omega_T = \sqrt{\frac{\log N}{T}} + \frac{1}{\sqrt{N}}$,

$$\max_{i\leq N, j\leq N} |\widetilde{R}_{ij} - \Sigma_{u,ij}| = \max_{i\leq N, j\leq N} |\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} - \Sigma_{u,ij}| = O_p(\omega_T). \tag{C.1}$$

Examining the proof of Theorem A.1 of Fan et al. (2013), we then have $\|\widetilde{\Sigma}_u - \Sigma_u\|_1 = O_p(m_N \omega_T^{1-q})$. We now show the first statement. Note that

$$\|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1 \leq \|(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})(\widetilde{\Sigma}_u - \Sigma_u)\Sigma_u^{-1}\|_1 + \|\Sigma_u^{-1}(\widetilde{\Sigma}_u - \Sigma_u)\Sigma_u^{-1}\|_1 \equiv a + b,$$

where $\|A\|_1 = \max_{j\leq N}\sum_{i=1}^{N} |A_{ij}|$. We have

$$\begin{aligned}
a &\leq \max_{j\leq N} \sum_{i,k,l\leq N} |(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})_{ik}||\Sigma_{u,kl} - \widetilde{\Sigma}_{u,kl}||\Sigma_{u,lj}^{-1}| \\
&\leq \max_{l} \sum_{i,k\leq N} |(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})_{ik}||\Sigma_{u,kl} - \widetilde{\Sigma}_{u,kl}| \max_{j\leq N}\sum_{l} |\Sigma_{u,lj}^{-1}| \\
&\leq \max_{l}\max_{k} \sum_{i\leq N} |(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})_{ik}| \sum_{k\leq N} |\Sigma_{u,kl} - \widetilde{\Sigma}_{u,kl}| \|\Sigma_u^{-1}\|_1 \\
&\leq \|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1 \|\Sigma_u^{-1}\|_1 \max_{l} \sum_{k\leq N} |\Sigma_{u,kl} - \widetilde{\Sigma}_{u,kl}| \\
&= \|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1 \|\Sigma_u^{-1}\|_1 \|\Sigma_u - \widetilde{\Sigma}_u\|_1 = O_p(m_N\omega_T^{1-q}) \|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1.
\end{aligned}$$

In addition,

$$b \leq \max_{j\leq N} \sum_{i,k,l\leq N} |\Sigma_{u,ik}^{-1}||\Sigma_{u,kl} - \widetilde{\Sigma}_{u,kl}||\Sigma_{u,lj}^{-1}| \leq \|\Sigma_u^{-1}\|_1^2 \|\widetilde{\Sigma}_u - \Sigma_u\|_1 = O_p(m_N \omega_T^{1-q}).$$

Hence we have $(1 + o_p(1))\|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\|_1 = O_p(m_N \omega_T^{1-q})$, which implies the result.

□

C.1 Consistency

Lemma C.1. *Under the assumptions of Theorem 5.1,*

$$\|\hat{\beta} - \beta_0\| = o_p(1).$$

Proof. Let $u = (u_1, ..., u_T)'$, and (F_0, Λ_0) denote the true factor and loading matrices. Concentrating out Λ, it can be shown that the estimated $\hat{\beta}$ and $\widehat{F}$ satisfy:

$$
\begin{aligned}
(\hat{\beta}, \widehat{F}) &= \arg\min_{\beta, F'F=TI_r} \frac{1}{NT}\mathrm{tr}(\widetilde{\Sigma}_u^{-1}(Y-X\beta)'M_F(Y-X\beta)) \\
&\quad -\frac{1}{NT}\mathrm{tr}(\widetilde{\Sigma}_u^{-1}u'M_{F_0}u) \\
&= \arg\min_{\beta, F'F=TI_r} S(\beta, F) + R(\beta, F)
\end{aligned}
$$

$X\beta$ is a $T\times N$ matrix with elements of $X_{it}'\beta$, and

$$
\begin{aligned}
S(\beta, F) &= \frac{1}{NT}(\beta-\beta_0)'Z'(\widetilde{\Sigma}_u^{-1}\otimes M_F)Z(\beta-\beta_0) \\
&\quad +\frac{2}{NT}\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}F_0\lambda_{0j}'M_F X_i'(\beta-\beta_0)+\frac{1}{NT}\mathrm{tr}(\widetilde{\Sigma}_u^{-1}\Lambda F_0'M_F F_0\Lambda_0'), \\
R(\beta, F) &= \frac{2}{NT}\mathrm{tr}(\widetilde{\Sigma}_u^{-1}u'M_F F_0\Lambda_0')+\frac{2}{NT}vec(u)'(\widetilde{\Sigma}_u^{-1}\otimes M_F)Z(\beta-\beta_0) \\
&\quad +\frac{1}{NT}\mathrm{tr}(\widetilde{\Sigma}_u^{-1}u'(F_0F_0'-FF')/Tu).
\end{aligned}
$$

It can be further verified that, with $V(F)$ as defined as (5.5),

$$
S(\beta, F) = (\beta-\beta_0)'V(F)(\beta-\beta_0)+(\eta+B^{-1}C(\beta-\beta_0))'B(\eta+B^{-1}C(\beta-\beta_0))\ge 0
$$

where $\eta = vec(M_F F_0)$, $B = (\frac{1}{N}\Lambda_0'\widetilde{\Sigma}_u^{-1}\Lambda_0)\otimes I_T$, and

$$
C = \frac{1}{NT}\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}[\lambda_{0j}\otimes M_F]X_i.
$$

By Lemma C.10, $\sup_{\beta, F'F=TI_r}|R(\beta, F)| = o_p(1)$. Hence

$$
S(\hat{\beta}, \widehat{F}) \le o_p(1)+S(\beta_0, F_0) = o_p(1),
$$

which implies $(\hat{\beta}-\beta_0)'V(\widehat{F})(\hat{\beta}-\beta_0) = o_p(1)$. The consistency of $\hat{\beta}$ follows since $\inf_{F'F=TI_r}\lambda_{\min}(V(F))$ is bounded away from zero in probability.

C.2 Preliminary analysis for the limiting distribution of $\hat{\beta}$

We can write

$$
\hat{\beta} = (\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}X_j)^{-1}(\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}Y_j).
$$

Note that

$$
\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}X_j = Z'(\widetilde{\Sigma}_u^{-1}\otimes M_{\widehat{F}})Z,
$$

hence with $Y_j = X_j\beta_0 + F_0\lambda_{0j} + u_j$,

$$\frac{1}{NT}Z'(\widetilde{\Sigma}_u^{-1} \otimes M_{\widehat{F}})Z(\hat{\beta}-\beta_0) = \frac{1}{NT}\sum_{i,j\leq N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}F_0\lambda_{0j} + \frac{1}{NT}\sum_{i,j\leq N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}u_j$$
$$= I + II. \tag{C.2}$$

We evaluate I and II separately. From now on, we use Λ for Λ_0 to denote the true matrix of loading, without causing any confusion. Let

$$A = (\frac{1}{NT}\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda F_0'\widehat{F})^{-1}$$

and V be a diagonal matrix of the r largest eigenvalues of

$$\frac{1}{NT}(Y - X(\hat{\beta}))'\widetilde{\Sigma}_u^{-1}(Y - X(\hat{\beta})),$$

where $X(\hat{\beta})$ is an $N\times T$ matrix $X(\hat{\beta}) = (X_1\hat{\beta}, ..., X_T\hat{\beta})$. Since $M_{\widehat{F}}\widehat{F} = 0$, we have

$$I = \sum_{i,j\leq N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}(F_0 - \widehat{F}VA)\lambda_{0j}.$$

Next, by the definition of the eigenvalues,

$$\frac{1}{NT}(Y - X(\hat{\beta}))'\widetilde{\Sigma}_u^{-1}(Y - X(\hat{\beta}))\widehat{F} = \widehat{F}V.$$

We thus have

$$\begin{aligned}\widehat{F}VA - F_0 \;\; = \;\; & \frac{1}{NT}\big\{[X(\beta-\hat{\beta})]'\widetilde{\Sigma}_u^{-1}[X(\beta-\hat{\beta})]\widehat{F} \\ & +[X(\beta-\hat{\beta})]'\widetilde{\Sigma}_u^{-1}\Lambda F_0'\widehat{F} + [X(\beta-\hat{\beta})]'\widetilde{\Sigma}_u^{-1}u'\widehat{F} \\ & +F_0\Lambda'\widetilde{\Sigma}_u^{-1}[X(\beta-\hat{\beta})]\widehat{F} + u\widetilde{\Sigma}_u^{-1}[X(\beta-\hat{\beta})]\widehat{F} + F_0\Lambda'\widetilde{\Sigma}_u^{-1}u'\widehat{F} \\ & +u\widetilde{\Sigma}_u^{-1}\Lambda F_0'\widehat{F} + u\widetilde{\Sigma}_u^{-1}u'\widehat{F}\big\}A,\end{aligned} \tag{C.3}$$

where $[X(\beta-\hat{\beta})]$ is a $N\times T$ matrix with elements of $X_{it}'(\beta-\hat{\beta})$.

Substituting into I, we thus have

$$I = \sum_{i=1}^{8} J_i.$$

We define and bound each J_i in the following lemmas (Lemmas C.2 and C.3). Substituting

Lemmas C.2, C.3 to (C.2), we obtain

$$
\begin{aligned}
&\frac{\sqrt{NT}}{NT} Z'(\widetilde{\Sigma}_u^{-1} \otimes M_{\widehat{F}}) Z(\hat{\beta}-\beta_0) \\
&= \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda \left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\widetilde{\Sigma}_u^{-1}) \otimes M_{\widehat{F}}] Z(\hat{\beta}-\beta_0) \\
&- \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda \left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\widetilde{\Sigma}_u^{-1}) \otimes M_{\widehat{F}}] U \\
&+ \frac{\sqrt{NT}}{NT} Z[\widetilde{\Sigma}_u^{-1} \otimes M_{\widehat{F}}] U + O_p(\sqrt{NT} m_N \omega_T^{3-q}) + o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|).
\end{aligned}
$$

It follows from $\|\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}\| = O_p(m_N \omega_T^{1-q})$, $\|\Lambda\| = O(\sqrt{N})$, $\|Z\|^2 = O_p(NT)$ that

$$
\begin{aligned}
&\|\frac{\sqrt{NT}}{NT} Z'((\Sigma_u^{-1} - \widetilde{\Sigma}_u^{-1}) \otimes M_{\widehat{F}}) Z(\hat{\beta}-\beta_0)\| = o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|). \\
&\frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\Sigma_u^{-1}\Lambda \left(\frac{\Lambda'\Sigma_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\Sigma_u^{-1}) \otimes M_{\widehat{F}}] Z(\hat{\beta}-\beta_0) \\
&= \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda \left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\widetilde{\Sigma}_u^{-1}) \otimes M_{\widehat{F}}] Z(\hat{\beta}-\beta_0) \\
&+ o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|).
\end{aligned}
$$

Therefore we have:

$$
\begin{aligned}
&\frac{\sqrt{NT}}{NT} Z'(\Sigma_u^{-1} \otimes M_{\widehat{F}}) Z(\hat{\beta}-\beta_0) \\
&= \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\Sigma_u^{-1}\Lambda \left(\frac{\Lambda'\Sigma_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\Sigma_u^{-1}) \otimes M_{\widehat{F}}] Z(\hat{\beta}-\beta_0) \\
&- \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda \left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\widetilde{\Sigma}_u^{-1}) \otimes M_{\widehat{F}}] U \\
&+ \frac{\sqrt{NT}}{NT} Z'[\widetilde{\Sigma}_u^{-1} \otimes M_{\widehat{F}}] U + O_p(\sqrt{NT} m_N \omega_T^{3-q}) + o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|) \\
&= \frac{\sqrt{NT}}{NT} Z'[\frac{1}{N}\Sigma_u^{-1}\Lambda \left(\frac{\Lambda'\Sigma_u^{-1}\Lambda}{N}\right)^{-1} \Lambda'\Sigma_u^{-1}) \otimes M_{\widehat{F}}] Z(\hat{\beta}-\beta_0) + \\
&\frac{1}{\sqrt{NT}} Z' A(\widetilde{\Sigma}_u^{-1}, \widehat{F}) U + O_p(\sqrt{NT} m_N \omega_T^{3-q}) + o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|)
\end{aligned}
$$

where

$$
A(\widetilde{\Sigma}_u^{-1}, \widehat{F}) = \left[\widetilde{\Sigma}_u^{-1} - \widetilde{\Sigma}_u^{-1}\Lambda \left(\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda\right)^{-1} \Lambda'\widetilde{\Sigma}_u^{-1}\right] \otimes M_{\widehat{F}}.
$$

Hence we have, for $A_{\widehat{F}} = A(\Sigma_u^{-1}, \widehat{F})$,

$$\begin{aligned}\frac{1}{\sqrt{NT}} Z' A_{\widehat{F}} Z(\hat{\beta} - \beta_0) &= \frac{1}{\sqrt{NT}} Z' A(\widetilde{\Sigma}_u^{-1}, \widehat{F}) U \\ &\quad + O_p(\sqrt{NT} m_N \omega_T^{3-q}) + o_p(\sqrt{NT} \|\hat{\beta} - \beta_0\|) \qquad \text{(C.4)}\end{aligned}$$

We need to show that the effect of replacing $\widehat{A} = A(\widetilde{\Sigma}_u^{-1}, \widehat{F})$ with $A_{F_0} \equiv A(\Sigma_u^{-1}, F_0)$ is asymptotically negligible. This is achieved by Proposition 5.1, to be proved below.

Lemma C.2. *We have,*
(i)

$$J_1 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} [X(\beta_0 - \hat{\beta})]' \widetilde{\Sigma}_u^{-1} [X(\beta_0 - \hat{\beta})] \widehat{F} A \lambda_{0j} = O_p(\|\beta_0 - \hat{\beta}\|^2)$$

(ii)

$$J_4 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} F_0 \Lambda' \widetilde{\Sigma}_u^{-1} [X(\beta_0 - \hat{\beta})] \widehat{F} A \lambda_{0j} = o_p(\|\beta_0 - \hat{\beta}\|)$$

(iii)

$$J_5 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} u \widetilde{\Sigma}_u^{-1} [X(\beta_0 - \hat{\beta})] \widehat{F} A \lambda_{0j} = o_p(\|\beta_0 - \hat{\beta}\|)$$

(iv)

$$J_3 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} [X(\beta_0 - \hat{\beta})]' \widetilde{\Sigma}_u^{-1} u' \widehat{F} A \lambda_{0j} = o_p(\|\beta - \hat{\beta}\|)$$

(v)

$$J_8 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} u \widetilde{\Sigma}_u^{-1} u' \widehat{F} A \lambda_{0j} = o_p(\|\hat{\beta} - \beta_0\| + \frac{1}{\sqrt{N}} m_N \omega_T^{2-q}),$$

(vi)

$$J_6 = -\frac{1}{N^2T^2} \sum_{i,j \leq N} \widetilde{\Sigma}_{u,ij}^{-1} X_i' M_{\widehat{F}} F_0 \Lambda' \widetilde{\Sigma}_u^{-1} u' \widehat{F} A \lambda_{0j} = O_p(m_N \omega_T^{1-q} (\|\hat{\beta} - \beta_0\|^2 + \frac{1}{N} + \frac{1}{T})).$$

Proof. (i) It follows immediately from that $\|\widetilde{\Sigma}_u^{-1}\| = O_p(1)$, $\|X\|_F = O_p(\sqrt{NT})$, $\|A\|_F = O_p(1)$, $\|\widehat{F}\|_F = O_p(\sqrt{T})$ and $\|M_{\widehat{F}}\|_F = O_p(1)$.

(ii) Note that $M_{\widehat{F}} F_0 = M_{\widehat{F}}(F_0 - \widehat{F} V A)$ due to $M_{\widehat{F}} \widehat{F} = 0$. Using the same proof that of Proposition A.1 in Bai (2009) to investigate (C.3), we have $\frac{1}{T}\|F_0 - \widehat{F} V A\|_F^2 = O_p(\|\hat{\beta} - \beta_0\|^2 + \frac{1}{N} + \frac{1}{T})$, which implies $\|M_{\widehat{F}} F_0\|_F = o_p(\sqrt{T})$, and the desired result.

(iii) Note that

$$J_5 = -\frac{1}{N^2T^2}(X_1 M_{\widehat{F}} u, ..., X_N M_{\widehat{F}} u)\{\widetilde{\Sigma}_u^{-1} \otimes (\widetilde{\Sigma}_u^{-1}[X(\beta_0 - \hat{\beta})]\widehat{F}A)\}vec(\Lambda).$$

$\|vec(\Lambda)\| = O(\sqrt{N})$, $\|\widetilde{\Sigma}_u^{-1} \otimes (\widetilde{\Sigma}_u^{-1}[X(\beta_0 - \hat{\beta})]\widehat{F}A)\| = O_p(T\sqrt{N}\|\hat{\beta} - \beta_0\|)$. For each $i \leq N$,

$$\frac{1}{T}X_i' M_{\widehat{F}} u = \frac{1}{T}\sum_{t=1}^{T} X_{it}u_t' - \frac{1}{T}X_i'\widehat{F}\frac{1}{T}\sum_{t=1}^{T}(\widehat{F}_t - (VA)^{-1}f_{0t})u_t'$$

$$-\frac{1}{T}X_i'\widehat{F}\frac{1}{T}\sum_{t=1}^{T}(VA)^{-1}f_{0t}u_t'.$$

In addition, $\max_{i,j\leq N}\|\frac{1}{T}\sum_{t=1}^{T} X_{it}u_{jt}\| = O_p(\sqrt{\frac{\log N}{T}}) = \max_{i,j\leq N}\|\frac{1}{T}\sum_{t=1}^{T} f_{0t}u_{jt}\|$.
Hence $\|(X_1 M_{\widehat{F}} u, ..., X_N M_{\widehat{F}} u)\|_F = O_p(TN(\|\hat{\beta} - \beta_0\| + \omega_T))$, and

$$J_5 = O_p(\|\hat{\beta} - \beta_0\|^2 + \omega_T\|\hat{\beta} - \beta_0\|).$$

This then yields the desired result.

(iv) We have

$$\|u'\widehat{F}\|_F \leq \|u'F_0(VA)^{-1}\|_F + \|u'(\widehat{F} - F_0(VA)^{-1}) = O_p(T\sqrt{N}\|\hat{\beta} - \beta_0\| + T + \sqrt{NT})$$

which implies that $J_3 = O_p(\|\hat{\beta} - \beta_0\|^2 + \|\hat{\beta} - \beta_0\|(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{T}}))$.

(v) First, let

$$J_{80} = -\frac{1}{N^2T^2}\sum_{i,j\leq N}\widetilde{\Sigma}_{u,ij}^{-1}X_i' M_{\widehat{F}} u\Sigma_u^{-1}u'\widehat{F}A\lambda_{0j}.$$

Then due to $\frac{1}{T}\|F_0 - \widehat{F}VA\|_F^2 = O_p(\|\hat{\beta} - \beta_0\|^2 + \frac{1}{N} + \frac{1}{T})$, we have

$$J_8 = J_{80} + o_p(\|\hat{\beta} - \beta_0\| + \frac{1}{\sqrt{N}}m_N\omega_T^{2-q}).$$

Also, $Eu\Sigma_u^{-1}u' = NI_T$, due to the serial uncorrelation, so

$$J_{80} = -\frac{1}{N^2T^2}\sum_{i,j\leq N}\widetilde{\Sigma}_{u,ij}^{-1}X_i' M_{\widehat{F}}(u\Sigma_u^{-1}u' - Eu\Sigma_u^{-1}u')\widehat{F}A\lambda_{0j}.$$

The similar proof to that of Lemma A.5 of Bai (2009) yields

$$J_{80} = o_p(\|\beta_0 - \hat{\beta}\|) + O_p(\frac{1}{N\sqrt{N}} + \frac{1}{T\sqrt{N}} + \frac{1}{N\sqrt{T}})$$

which implies the result.

(vi) By $\frac{1}{T}\|F_0 - \widehat{F}VA\|_F^2 = O_p(\|\hat{\beta} - \beta_0\|^2 + \frac{1}{N} + \frac{1}{T})$, we have

$$\|M_{\widehat{F}}F_0\|_F = O_p(\sqrt{T}\|\hat{\beta} - \beta_0\| + \sqrt{\frac{T}{N}} + 1).$$

On the other hand, $\|\Lambda'\Sigma_u^{-1}u'\|_F = O_p(\sqrt{NT})$, $\|u'F_0\|_F = O_p(\sqrt{NT})$, and $\|\Lambda'\Sigma_u^{-1}u'F_0\|_F = O_p(\sqrt{NT})$ because
$\|\frac{1}{\sqrt{NT}}\sum_{i\le N,t\le T}\lambda_i(\Sigma_u^{-1}u_t)_i f_t'\|_F = O_p(1)$. We thus have

$$\begin{aligned}
&\|\Lambda'\widetilde{\Sigma}_u^{-1}u'\widehat{F}\|_F \le \|\Lambda'(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})u'\widehat{F}\|_F + \|\Lambda'\Sigma_u^{-1}u'\widehat{F}\|_F \\
&\le \|\Lambda'(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})u'(\widehat{F} - F_0(VA)^{-1})\|_F + \|\Lambda'(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})u'F_0(VA)^{-1}\|_F \\
&+\|\Lambda'\Sigma_u^{-1}u'(\widehat{F} - F_0(VA)^{-1})\|_F + \|\Lambda'\Sigma_u^{-1}u'F_0(VA)^{-1}\|_F \\
= \;& O_p(N\sqrt{T}m_N\omega_T^{1-q}(\sqrt{T}\|\hat{\beta} - \beta_0\| + \sqrt{\frac{T}{N}} + 1)).
\end{aligned}$$

This implies the desired result.

□

In the lemma below, recall that II was defined in (C.2).

Lemma C.3. *(i)*

$$J_2 = -\frac{1}{N^2T^2}\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}[X(\beta_0 - \hat{\beta})]'\widetilde{\Sigma}_u^{-1}\Lambda F_0'\widehat{F}A\lambda_{0j}$$

$$= \frac{1}{NT}Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda\left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1}\Lambda'\widetilde{\Sigma}_u^{-1})\otimes M_{\widehat{F}}]Z(\hat{\beta} - \beta_0)$$

(ii)

$$J_7 = -\frac{1}{N^2T^2}\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}u\widetilde{\Sigma}_u^{-1}\Lambda F_0'\widehat{F}A\lambda_{0j}$$

$$= \frac{-1}{NT}Z'[\frac{1}{N}\widetilde{\Sigma}_u^{-1}\Lambda\left(\frac{\Lambda'\widetilde{\Sigma}_u^{-1}\Lambda}{N}\right)^{-1}\Lambda'\widetilde{\Sigma}_u^{-1})\otimes M_{\widehat{F}}]U$$

(iii) $II \equiv \frac{1}{NT}\sum_{i,j\le N}\widetilde{\Sigma}_{u,ij}^{-1}X_i'M_{\widehat{F}}u_j = \frac{1}{NT}Z[\widetilde{\Sigma}_u^{-1}\otimes M_{\widehat{F}}]U$.

Proof. The proofs are just straightforward calculations.

□

C.3 Proof of Proposition 5.1

The key step of the proof is the following lemma.

Lemma C.4. *For each $l \leq \dim(\beta)$, the following conditions hold for both $Q_{jt} = \Sigma_{u,j}^{-1\prime} X_{l,t}$ and $Q_{jt} = \Sigma_{u,j}^{-1\prime}(EX_{l,t}f_t')(Ef_tf_t')^{-1}f_t$ (here Q_{jt} is a scalar),*

$$\frac{1}{T\sqrt{NT}}\sum_{i=1}^{N}\sum_{s=1}^{T}(u_{is}^2 - Eu_{is}^2)\sum_{t=1}^{T}Q_{it}e_{it} = o_p(1),$$

$$\frac{1}{T\sqrt{NT}}\sum_{i\neq j,(i,j)\in S_U}\sum_{s=1}^{T}(u_{is}u_{js} - Eu_{is}u_{js})\sum_{t=1}^{T}Q_{jt}e_{it} = o_p(1).$$

Proof. This lemma is proved in Appendix C.5. $\square$

For each $q \leq \dim(\beta)$, let $X_q = (X_{it,q})_{N\times T}$,

$$R = I_T - \frac{1}{T}F_0(Ef_tf_t')^{-1}F_0', \quad G = \frac{1}{T}F^*F^{*\prime}, \quad F^* = F_0(VA)^{-1}.$$

Lemma C.5. *For each $q \leq d = \dim(\beta)$ and $X_{q,i}' = (X_{i1,q}, ..., X_{iT,q})$,*
$\frac{1}{\sqrt{NT}}\mathrm{tr}[\Sigma_u^{-1}(\Sigma_u - \widetilde{\Sigma}_u)\Sigma_u^{-1}X_qRu'] = o_p(1)$.

Proof. To simplify notation, we assume $d = 1$ and write $X = X_q = (X_{it})_{N\times T}$ without loss of generality. Let e_i' and $\Sigma_{u,j}^{-1}$ denote the ith row of $\Sigma_u^{-1}u$ and the jth column of Σ_u^{-1} respectively. Then

$$\begin{aligned}
L &= \frac{1}{\sqrt{NT}}\mathrm{tr}[\Sigma_u^{-1}(\Sigma_u - \widetilde{\Sigma}_u)\Sigma_u^{-1}X_qRu'] \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}(\Sigma_u - \widetilde{\Sigma}_u)_{ij}\Sigma_{u,j}^{-1\prime}XRe_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}(\Sigma_u - \widetilde{\Sigma}_u)_{ij}\Sigma_{u,j}^{-1\prime}Xe_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}(\Sigma_u - \widetilde{\Sigma}_u)_{ij}\Sigma_{u,j}^{-1\prime}(EX_tf_t')(Ef_tf_t')^{-1}F_0'e_i \\
&\quad +\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}(\Sigma_u - \widetilde{\Sigma}_u)_{ij}\Sigma_{u,j}^{-1\prime}(EX_tf_t' - \frac{1}{T}\sum_{t=1}^{T}X_tf_t')(Ef_tf_t')^{-1}\sum_{s=1}^{T}f_se_{is} \\
&= L_1 + L_2 + L_3.
\end{aligned}$$

Let $\widetilde{X}_{jt} = \Sigma_{u,j}^{-1\prime}X_t$, then $\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{X}_{jt}f_t - E\widetilde{X}_{jt}f_t\| = O_p(\sqrt{\frac{\log N}{T}})$ because $\|\Sigma_u^{-1}\|_1 = O(1)$.

$$L_3 \leq O(\frac{N}{\sqrt{NT}})\max_i\|\sum_{s=1}^{T}f_se_{is}\|\max_{j\leq N}\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{X}_{jt}f_t - E\widetilde{X}_{jt}f_t\|\|\Sigma_u - \widetilde{\Sigma}_u\|_1$$

$$= O_p((\log N)\sqrt{\frac{N}{T}}m_N\omega_T^{1-q})$$

which is $o_p(1)$. On the other hand, both L_1 and L_2 are of the form: for some $1\times T$ vector Q_j

$$L_{1,2} = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{j=1}^N(\Sigma_u - \widetilde{\Sigma}_u)_{ij}Q_je_i$$

where $Q_j = \Sigma_{u,j}^{-1'}X$ for L_1 and $Q_j = -\Sigma_{u,j}^{-1'}(EX_tf_t')(Ef_tf_t')^{-1}F_0'$ for L_2. Because $\|\Sigma_u^{-1}\| = O(1)$,

$$\max_{i,j}|Q_je_i| \leq \max_{ij}|\sum_{t=1}^T Q_{jt}e_{it}| = O_p(\sqrt{T\log N}).$$

By definition, when $i\neq j$, $\widetilde{\Sigma}_{u,ij} = 0$ if $|\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}\widehat{u}_{jt}| \leq \tau_{ij}\omega_T$, where τ_{ij} is the threshold constant, bounded away from both zero and infinity with probability approaching one. For any $C>0$, one can pick up a threshold constant in τ_{ij} such that $P(\tau_{ij} > C)\to 1$.

$$\begin{aligned} L_{1,2} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it}^2 - Eu_{it}^2) + \frac{1}{\sqrt{NT}}\sum_{i\neq j,(i,j)\in S_U} Q_je_i(\widetilde{\Sigma}_{u,ij} - \Sigma_{u,ij}) \\ &\quad + \frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L} Q_je_i(\widetilde{\Sigma}_{u,ij} - \Sigma_{u,ij}) \end{aligned}$$

The first and second terms are bounded in Lemmas C.12 and C.13 below, which are $o_p(1)$. We now look at the third term. On one hand,

$$\frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L}|Q_je_i||\Sigma_{u,ij}| = O_p(\sqrt{\frac{T\log N}{NT}})\sum_{(i,j)\in S_L}|\Sigma_{u,ij}| = O_p(\sqrt{\frac{\log N}{N}}).$$

On the other hand, because $\max_{ij}|\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}\widehat{u}_{jt} - Eu_{it}u_{jt}| = O_p(\omega_T)$ (see (C.1)), and $\max_{(i,j)\in S_L}|\Sigma_{u,ij}| = o(\omega_T)$, then for any $\epsilon > 0$, one can pick up large enough $C>0$ so that

$$\begin{aligned} &P(|\frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L}|Q_je_i||\widetilde{\Sigma}_{u,ij}| > T^{-1}) \\ &\leq P(\max_{(i,j)\in S_L}|\widetilde{\Sigma}_{u,ij}| > 0) \leq P(\exists(i,j)\in S_L, |\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}\widehat{u}_{jt}| > \tau_{ij}\omega_T) \\ &\leq P(\max_{ij}|\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}\widehat{u}_{jt}| > \omega_T C) + o(1) \end{aligned}$$

$$\leq P(\max_{ij}|\frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}\widehat{u}_{jt}-\Sigma_{u,ij}|+\max_{(i,j)\in S_L}|\Sigma_{u,ij}|>\omega_T C)+o(1)<\epsilon,$$

which implies $|\frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L}|Q_je_i||\widetilde{\Sigma}_{u,ij}|=O_p(\frac{1}{T})$. Hence

$$|\frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L}Q_je_i(\widetilde{\Sigma}_{u,ij}-\Sigma_{u,ij})|\leq\frac{1}{\sqrt{NT}}\sum_{(i,j)\in S_L}|Q_je_i|(|\widetilde{\Sigma}_{u,ij}|+|\Sigma_{u,ij}|)=o_p(1).$$

Therefore, by Lemmas C.12 and C.13, we have $L_{1,2}=o_p(1)$ when either $Q_j=\Sigma_{u,j}^{-1'}X$ or $Q_j=-\Sigma_{u,j}^{-1'}(EX_tf_t')(Ef_tf_t')^{-1}F_0'$. This proves $L=o_p(1)$. □

Lemma C.6. *For each $q\leq d=\dim(\beta)$ and $X_{q,i}'=(X_{i1,q},...,X_{iT,q})$,*
(i) $\frac{1}{\sqrt{NT}}\mathrm{tr}[(\widetilde{\Sigma}_u^{-1}-\Sigma_u^{-1})(\Sigma_u-\widetilde{\Sigma}_u)\Sigma_u^{-1}X_qM_{\widehat{F}}u']=o_p(1)+o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|)$
(ii) $\frac{1}{\sqrt{NT}}\mathrm{tr}[\Sigma_u^{-1}(\Sigma_u-\widetilde{\Sigma}_u)\Sigma_u^{-1}X_q(G-R)u']=o_p(1)+o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|)$
(iii) $\frac{1}{\sqrt{NT}}\mathrm{tr}[\Sigma_u^{-1}(\Sigma_u-\widetilde{\Sigma}_u)\Sigma_u^{-1}X_q(G-M_{\widehat{F}})u']=o_p(1)+o_p(\sqrt{NT}\|\hat{\beta}-\beta_0\|)$.

Proof. (i) By Theorem C.1,
$\|\widetilde{\Sigma}_u-\Sigma_u\|_1=O_p(m_N\omega_T^{1-q})=\|\widetilde{\Sigma}_u^{-1}-\Sigma_u^{-1}\|_1$. The term of interest is

$$\begin{aligned}&\frac{1}{\sqrt{NT}}\sum_{i,j,k\leq N}(\widetilde{\Sigma}_u-\Sigma_u)_{ik}(\widetilde{\Sigma}_u^{-1}-\Sigma_u^{-1})_{kj}(\Sigma_u^{-1}X_qM_{\widehat{F}}u')_{ji}\\&\leq\|\Sigma_u^{-1}X_qM_{\widehat{F}}u'\|_{\max}\|\widetilde{\Sigma}_u-\Sigma_u\|_1\|\widetilde{\Sigma}_u^{-1}-\Sigma_u^{-1}\|_1\frac{N}{\sqrt{NT}}\\&=O_p((\sqrt{N\log N}+\sqrt{NT}(\hat{\beta}-\beta_0)+\sqrt{T})m_N^2\omega_T^{2-2q}).\end{aligned}\tag{C.5}$$

where we used $\|\Sigma_u^{-1}X_qM_{\widehat{F}}u'\|_{\max}=O_p(\sqrt{T\log N}+T\|\hat{\beta}-\beta_0\|+\frac{T}{\sqrt{N}})$ by Lemma C.11 and $\|\Sigma_u^{-1}\|_1=O(1)$. The desired result follows.

(ii) The objective is bounded by

$$\frac{N}{\sqrt{NT}}\|\Sigma_u^{-1}\|_1^2\|\Sigma_u-\widetilde{\Sigma}_u\|_1\max_{i,j}|X_{q,i}'(R-G)u_j|.$$

The result then follows from Lemma C.11 below.

(iii) Recall the notation $e=\Sigma_u^{-1}u$ and $Q_j=\Sigma_{u,j}^{-1'}X_q$. We have

$$\begin{aligned}&\max_{ij}|Q_j\frac{1}{T}(\widehat{F}-F^*)(\widehat{F}-F^*)'e_i|=O_p(T\|\hat{\beta}-\beta_0\|^2+\frac{T}{N}+1),\\&\max_{ij}|Q_j\frac{1}{T}(\widehat{F}-F^*)(A'V)^{-1}F_0'e_i|=O_p(\sqrt{T}\|\widehat{\beta}-\beta_0\|+\sqrt{\frac{T}{N}}+1).\end{aligned}$$

Substituting

$$G - M_{\widehat{F}} = \frac{1}{T}(F^* - \widehat{F})F^{*\prime} - \frac{1}{T}(F^* - \widehat{F})(F^* - \widehat{F})' + F^* \frac{1}{T}(F^* - \widehat{F}),$$

and noting that $\|\widetilde{\Sigma}_u - \Sigma_u\|_1 = O_p(m_N \omega_T^{1-q})$, we obtain

$$\begin{aligned}
&\frac{1}{\sqrt{NT}} \mathrm{tr}[\Sigma_u^{-1}(\Sigma_u - \widetilde{\Sigma}_u)\Sigma_u^{-1} X_q (G - M_{\widehat{F}}) u'] \\
&= \frac{1}{\sqrt{NT}} \sum_{ij} (\Sigma_{u,ij} - \widetilde{\Sigma}_{u,ij}) Q_j (G - M_{\widehat{F}}) e_i \\
&= \frac{1}{\sqrt{NT}} \sum_{ij} (\Sigma_{u,ij} - \widetilde{\Sigma}_{u,ij}) Q_j \frac{1}{T} F^* (F^* - \widehat{F}) e_i + o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|) \\
&\equiv B + o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|).
\end{aligned}$$

where

$$B = \frac{1}{\sqrt{NT}} \sum_{ij} (\Sigma_{u,ij} - \widetilde{\Sigma}_{u,ij}) Q_j \frac{1}{T} F^* (F^* - \widehat{F}) e_i.$$

We analyze $F^* - \widehat{F}$ in B using (C.3), and study it term by term. It is not difficult to obtain

$$\begin{aligned}
Q_j \frac{1}{T} F^* (F^* - \widehat{F}) e_i &= -\frac{1}{T} \sum_{t=1}^{T} Q_{jt} f_t \frac{1}{NT} V^{-1} [\widehat{F}' F_0 \Lambda' \widetilde{\Sigma}_u^{-1} u' e_i \\
&\quad + \widehat{F}' u \widetilde{\Sigma}_u^{-1} u' e_i] + O_p(T\|\hat{\beta} - \beta_0\| + \log N) \\
&= B_1 + B_2 + O_p(T\|\hat{\beta} - \beta_0\| + \log N),
\end{aligned}$$

where the $O_p(\cdot)$ term is uniform in $j, i \leq N$. Term B_1 equals

$$\begin{aligned}
&-\frac{1}{T} \sum_{t=1}^{T} Q_{jt} f_t \frac{1}{NT} V^{-1} \widehat{F}' F_0 \Lambda' \Big[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1})(u' e_i - E u' e_i) + (\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) E u' e_i \\
&+ \Sigma_u^{-1}(u' e_i - E u' e_i) + \Sigma_u^{-1} E u' e_i \Big] = \sum_{i=1}^{4} B_{1i}.
\end{aligned}$$

For $e' = \Sigma_u^{-1} u'$, the key observation is that $E u' e_i = (0, ..., T, ..., 0)'$, with the ith element being T and others being zero. Hence $\Lambda' \Sigma_u^{-1} E u' e_i = O(T)$, which implies $B_{12} + B_{14} = O_p(\frac{T}{N} + \frac{T}{\sqrt{N}} m_N \omega_T^{1-q})$, and $B_{11} = O_p(m_N \omega_T^{1-q} \sqrt{T \log N})$, where the $O_p(\cdot)$ term is uniform in $j, i \leq N$. Term B_2 can be treated similarly, and is easier. Combining these intermediate results (carrying over B_{13}), we obtain

$$\frac{1}{\sqrt{NT}} \mathrm{tr}[\Sigma_u^{-1}(\Sigma_u - \widetilde{\Sigma}_u)\Sigma_u^{-1} X_q (G - M_{\widehat{F}}) u']$$

$$= \frac{-1}{\sqrt{NT}} \sum_{ij} (\Sigma_{u,ij} - \widetilde{\Sigma}_{u,ij}) \frac{1}{T} \sum_{t=1}^{T} Q_{jt} f_t \frac{1}{NT} V^{-1} \widehat{F}' F_0 \Lambda' \Sigma_u^{-1} (u'e_i - Eu'e_i)$$
$$+ o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|) + O_p(m_N \omega_T^{1-q} \sqrt{\frac{T}{N}} + m_N^2 \omega_T^{2-2q} (\sqrt{T} + \sqrt{N \log N}))$$
$$\leq O_p(\frac{m_N \omega_T^{1-q}}{N}) \sum_i \|\frac{1}{\sqrt{NT}} \Lambda' \Sigma_u^{-1} (u'e_i - Eu'e_i)\| + o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta\|).$$

Because $E\|\frac{1}{\sqrt{NT}}\Lambda'\Sigma_u^{-1}(u'e_i - Eu'e_i)\|^2 = O(1)$, we complete the proof. □

Lemma C.7. *We have*
(i) $\|\frac{1}{\sqrt{NT}} Z'[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) \otimes M_{\widehat{F}}]U\| = o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|)$.
(ii) $\|\frac{1}{\sqrt{NT}} Z' \left\{ [\frac{1}{N} \widetilde{\Sigma}_u^{-1} \Lambda \left(\frac{\Lambda' \widetilde{\Sigma}_u^{-1} \Lambda}{N} \right)^{-1} \Lambda' \widetilde{\Sigma}_u^{-1}) - \frac{1}{N} \Sigma_u^{-1} \Lambda \left(\frac{\Lambda' \Sigma_u^{-1} \Lambda}{N} \right)^{-1} \Lambda' \Sigma_u^{-1})] \otimes M_{\widehat{F}} \right\} U\| = o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|)$.

Proof. Consider part (i). The qth row ($q \leq d$) of $\frac{1}{\sqrt{NT}} Z[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) \otimes M_{\widehat{F}}]U$ can be written as $\frac{1}{\sqrt{NT}} \text{tr}[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) X_q M_{\widehat{F}} u']$ for $u = (u_{it})_{N \times T}$. In addition,

$$\frac{1}{\sqrt{NT}} \text{tr}[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) X_q M_{\widehat{F}} u'] = \frac{1}{\sqrt{NT}} \text{tr}[\Sigma_u^{-1} (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q M_{\widehat{F}} u']$$
$$+ \frac{1}{\sqrt{NT}} \text{tr}[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q M_{\widehat{F}} u']$$
$$= \frac{1}{\sqrt{NT}} \text{tr}[\Sigma_u^{-1} (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q R u']$$
$$+ \frac{1}{\sqrt{NT}} \text{tr}[(\widetilde{\Sigma}_u^{-1} - \Sigma_u^{-1}) (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q M_{\widehat{F}} u']$$
$$+ \frac{1}{\sqrt{NT}} \text{tr}[\Sigma_u^{-1} (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q (M_{\widehat{F}} - G) u']$$
$$+ \frac{1}{\sqrt{NT}} \text{tr}[\Sigma_u^{-1} (\Sigma_u - \widetilde{\Sigma}_u) \Sigma_u^{-1} X_q (G - R) u'].$$

It follows from Lemmas C.5 and C.6 that the four terms on the right hand side are all $o_p(1) + o_p(\sqrt{NT}\|\hat{\beta} - \beta_0\|)$, which concludes the proof for part (i). The proof of part (ii) is very similar to that of part (i).

□

Recall the notation $A(\Sigma_u^{-1}, \widehat{F}) = A_{\widehat{F}}$, $A(\widetilde{\Sigma}_u^{-1}, \widehat{F}) = \widehat{A}$ and $A(\Sigma_u^{-1}, F_0) = A_{F_0}$.

Lemma C.8. $\frac{1}{\sqrt{NT}} Z' A_{\widehat{F}} U = \frac{1}{\sqrt{NT}} Z' A_{F_0} U + O_p(\sqrt{\frac{T}{N}}) + o_p(1)$.

Proof. Recall that $e_t = \Sigma_u^{-1} u_t$ and $f_t = F_{0t}$ denotes the true vector of factors. Let $B =$

$M_{\widehat{F}} - M_{F_0}$. First consider $\frac{1}{\sqrt{NT}} Z'(\Sigma_u^{-1} \otimes B)U$, which equals

$$\begin{aligned}
&\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} B_{st} X_{is} \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} X_{is} \frac{1}{T} (\widehat{F}_s - (VA)'^{-1} f_s)'(VA)'^{-1} f_t \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} X_{is} \frac{1}{T} (\widehat{F}_s - (VA)'^{-1} f_s)'(\widehat{F}_t - (VA)'^{-1} f_t) \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} X_{is} \frac{1}{T} f_s' (VA)^{-1} (\widehat{F}_t - (VA)'^{-1} f_t) \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} X_{is} \frac{1}{T} f_s' ((VA)^{-1} (VA)'^{-1} - (E f_t f_t')^{-1}) f_t.
\end{aligned}$$

These terms can be bounded in the same way as in the proof of Lemma A.8 in Bai (2009), and we reach

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} B_{st} X_{is} = O_p(\sqrt{\frac{T}{N}}) + o_p(1).$$

In addition, if we define $\widetilde{X}_{is}^0 = \sum_{k=1}^{N} \sum_{j=1}^{N} \lambda_i' (\Lambda' \Sigma_u^{-1} \Lambda)^{-1} \lambda_j X_{ks} (\Sigma^{-1})_{kj}$, it then can be shown that

$$\frac{1}{\sqrt{NT}} Z'[(\frac{1}{N} \Sigma_u^{-1} \Lambda \left(\frac{\Lambda' \Sigma_u^{-1} \Lambda}{N} \right)^{-1} \Lambda' \Sigma_u^{-1}) \otimes B]U = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} e_{it} B_{st} \widetilde{X}_{is}^0.$$

This term can also be bounded in the same way as in the proof of Lemma A.8 in Bai (2009). We omit the details.

□

Proof of Proposition 5.1

It follows from Lemmas C.7 and C.8 that (under $N/T \to \infty$)

$$\frac{1}{\sqrt{NT}} Z'(\widehat{A} - A_{\widehat{F}})U = o_p(1) + o_p(\sqrt{NT} \|\hat{\beta} - \beta_0\|),$$

$$\frac{1}{\sqrt{NT}} Z'(A_{\widehat{F}} - A_{F_0})U = o_p(1).$$

Hence

$$\frac{1}{\sqrt{NT}} Z'(\widehat{A} - A_{F_0})U = o_p(1) + o_p(\sqrt{NT} \|\hat{\beta} - \beta_0\|). \tag{C.6}$$

It then follows from (C.4) that, (note that $V(\widehat{F}) = \frac{1}{NT} Z' A_{\widehat{F}} Z$)

$$\sqrt{NT} V(\widehat{F})(\hat{\beta} - \beta_0) = \frac{1}{\sqrt{NT}} Z' A_{F_0} U + o_p(1) + o_p(\sqrt{NT} \|\hat{\beta} - \beta_0\|).$$

which implies, by Assumption 5.1,

$$\sqrt{NT}(\hat{\beta} - \beta_0) = V(\widehat{F})^{-1} \frac{1}{\sqrt{NT}} Z' A_{F_0} U + o_p(1) + o_p(\sqrt{NT} \|\hat{\beta} - \beta_0\|). \tag{C.7}$$

This also implies that, still by Assumption 5.1, there is $C > 0$ so that

$$(1 + o_p(1))\sqrt{NT} \|\hat{\beta} - \beta_0\| \leq \| \frac{C}{\sqrt{NT}} Z' A_{F_0} U \| + o_p(1).$$

Because $\frac{1}{\sqrt{NT}} Z' A_{F_0} U = O_p(1)$ by Assumption 5.2, hence

$$\sqrt{NT}(\hat{\beta} - \beta_0) = O_p(1).$$

It then follows from (C.6) that

$$\frac{1}{\sqrt{NT}} Z'(\widehat{A} - A_{F_0}) U = o_p(1).$$

C.4 Proof of Theorems 5.1, 5.2

By (C.7),

$$\sqrt{NT}(\hat{\beta} - \beta_0) = V(\widehat{F})^{-1} \frac{1}{\sqrt{NT}} Z' A_{F_0} U + o_p(1).$$

In addition, the same proof of Lemma A.9 (i) in Bai (2009) implies that $V(\widehat{F})^{-1} \to^p V(F_0)^{-1}$. We have

$$\sqrt{NT}(\hat{\beta} - \beta_0) = V(F_0)^{-1} \frac{1}{\sqrt{NT}} Z' A_{F_0} U + o_p(1).$$

The limiting distribution then follows immediately from Assumption 5.2. $\square$

Because $\|M_{\widehat{F}} - M_{F_0}\|_F = o_p(1)$ and $\frac{1}{N} \|\widehat{\Lambda} - \Lambda\|^2 = o_p(1)$, which can be proved similarly to Theorem 3.2, and note that the effect of estimating β_0 is negligible due to the $\sqrt{NT}$-consistency of $\widehat{\beta}$. Hence Theorem 5.2 follows.

C.5 Proof of Lemma C.4

Lemma C.9. *When* $u_t \sim \mathcal{N}(0, \Sigma_u)$ *and* $e_t = \Sigma_u^{-1} u_t$*, then*
(i) $Eu_{it}^2 e_{js} = 0$ *for each* $i, j \leq N$*, and (ii)* $Eu_{it}e_{jt} = 0$ *when* $i \neq j$*.*

Proof. (i) For each (i,j), define $a = \frac{\mathrm{cov}(u_{it}, e_{jt})}{\mathrm{var}(u_{it})}$. Let $v = e_{jt} - au_{it}$, then v is Gaussian and $Ev = 0$. Moreover,

$$\mathrm{cov}(v, u_{it}) = \mathrm{cov}(e_{jt}, u_{it}) - a\mathrm{var}(u_{it}) = 0.$$

Hence v and u_{it} are independent, implying $Evu_{it}^2 = 0$. So $Evu_{it}^2 = Ee_{jt}u_{it}^2 - aEu_{it}^3$, which yields $Eu_{it}^2 e_{jt} = 0$.

(ii) The proof is a straightforward calculation of the covariance matrix of (u_t', e_t'). $\square$

We now prove Lemma C.4 by proving the two statements separately:

$$\frac{1}{T\sqrt{NT}} \sum_{i=1}^N \sum_{s=1}^T (u_{is}^2 - Eu_{is}^2) \sum_{t=1}^T Q_{it} e_{it} = o_p(1), \tag{C.8}$$

$$\frac{1}{T\sqrt{NT}} \sum_{i \neq j, (i,j) \in S_U} \sum_{s=1}^T (u_{is}u_{js} - Eu_{is}u_{js}) \sum_{t=1}^T Q_{jt} e_{it} = o_p(1). \tag{C.9}$$

C.5.1 Proof of (C.8)

let

$$G = \frac{1}{T\sqrt{NT}} \sum_{i=1}^N \sum_{s=1}^T (u_{is}^2 - Eu_{is}^2) \sum_{t=1}^T Q_{it} e_{it}.$$

We respectively show that $|EG| = o(1)$ and $\mathrm{var}(G) = o(1)$, which will then imply $G = o_p(1)$.

Expectation Because the data are serially uncorrelated,

$$EG = \frac{1}{T\sqrt{NT}} \sum_{i=1}^N \sum_{s=1}^T E(u_{is}^2 - Eu_{is}^2) Q_{is} e_{is} = \frac{1}{T\sqrt{NT}} \sum_{i=1}^N \sum_{s=1}^T E(u_{is}^2 - Eu_{is}^2) e_{is} EQ_{is} = 0$$

where we used $Eu_{is}^2 e_{is} = 0$ by Lemma C.9 and that Q_{is} and u_s are independent.

Variance

$$\mathrm{var}(G) = \frac{1}{T^3 N} \sum_{i=1}^N \mathrm{var}[\sum_{s=1}^T (u_{is}^2 - Eu_{is}^2) \sum_{t=1}^T Q_{it} e_{it}]$$
$$+ \frac{1}{T^3 N} \sum_{i \neq j} \mathrm{cov}(\sum_{s=1}^T (u_{is}^2 - Eu_{is}^2) \sum_{t=1}^T Q_{it} e_{it}, \sum_{s=1}^T (u_{js}^2 - Eu_{js}^2) \sum_{t=1}^T Q_{jt} e_{jt}) \equiv A_1 + A_2.$$

By the Cauchy-Schwarz inequality,

$$A_1 \leq \frac{1}{T^3 N}\sum_{i=1}^N E[\sum_{s=1}^T (u_{is}^2 - Eu_{is}^2)\sum_{t=1}^T Q_{it}e_{it}]^2$$

$$\leq \frac{1}{TN}\sum_{i=1}^N [E(\frac{1}{\sqrt{T}}\sum_{s=1}^T (u_{is}^2 - Eu_{is}^2))^4]^{1/2}[E(\frac{1}{\sqrt{T}}\sum_{t=1}^T Q_{it}e_{it})^4]^{1/2} = O(\frac{1}{T}).$$

$$A_2 = \frac{1}{T^3 N}\sum_{i\neq j}\sum_{s,t,k,l\leq T} \mathrm{cov}((u_{is}^2 - Eu_{is}^2)Q_{it}e_{it}, (u_{jk}^2 - Eu_{jk}^2)Q_{jl}e_{jl})$$

$$\equiv \frac{1}{T^3 N}\sum_{i\neq j}\sum_{s,t,k,l\leq T} C_{ij,stkl}$$

By Lemma C.9, $Ee_{it}u_{js}^2 = 0$ for any $i, j \leq N, t, s \leq T$. Also, Q_{jt} is independent of (u_t, e_t), and $\{Q_t, u_t\}_{t\leq T}$ is serially independent. Therefore, it is easy to verify that for fixed four integers s, t, k, l, if the set $\{s, t, k, l\}$ contains more than two distinct elements, $C_{ij,stkl} = 0$. Hence if we denote Θ as the set of (s, t, k, l) such that $\{s, t, k, l\}$ contains no more than two distinct elements, then its cardinality satisfies $|\Theta|_0 = O(T^2)$, and

$$\sum_{s,t,k,l\leq T} C_{ij,stkl} = \sum_{(s,t,k,l)\in\Theta} C_{ij,stkl}.$$

Let us partition Θ into $\Theta_1 \cup \Theta_2$ where each element $(s, t, k, l) \in \Theta_1$ contains exactly two distinct integers and each element in Θ_2 contains just one integer (that is, $s = t = k = l$ if $(s, t, k, l) \in \Theta_2$). We know that $\sum_{(s,t,k,l)\in\Theta_2} C_{ij,stkl} = O(T)$. Hence

$$A_2 = \frac{1}{T^3 N}\sum_{i\neq j}\sum_{(s,t,k,l)\in\Theta_1} C_{ij,stkl} + O(\frac{N}{T^2}).$$

Because $Eu_{is}^2 e_{js} = 0$ regardless of (i, j), so

$$A_2 = \frac{1}{T^3 N}\sum_{i\neq j}\sum_{s=1}^T\sum_{t=1}^T [E(u_{is}^2 - Eu_{is}^2)(u_{js}^2 - Eu_{js}^2)]Ee_{it}e_{jt}EQ_{it}Q_{jt} + O(\frac{N}{T^2}).$$

Note that $Ee_{it}e_{jt} = (\Sigma_u^{-1})_{ij}$, and $\|\Sigma_u^{-1}\|_1 = O(1)$. Hence $A_2 = O(\frac{T+N}{T^2}) = o(1)$. This implies $\mathrm{var}(G) = o(1)$, and hence $G = o_p(1)$.

C.5.2 Proof of (C.9)

Let

$$M = \frac{1}{T\sqrt{NT}} \sum_{i\neq j,(i,j)\in S_U} \sum_{s=1}^{T} (u_{is}u_{js} - Eu_{is}u_{js}) \sum_{t=1}^{T} Q_{jt}e_{it}.$$

Expectation For Gaussian errors, $Eu_{is}u_{js}e_{is} = 0$ for all i, j, s. Hence $EM = 0$.

Variance Let $\alpha_{ijs} = u_{is}u_{js} - Eu_{is}u_{js}$. We have,

$$\begin{aligned}
&\mathrm{var}(M) \\
&= \frac{1}{T^3N} \sum_{i\neq j,(i,j)\in S_U} \sum_{m\neq n,(m,n)\in S_U,(m,n)\neq(i,j)} \mathrm{cov}(\sum_{s=1}^{T} \alpha_{ijs} \sum_{t=1}^{T} Q_{jt}e_{it}, \sum_{s=1}^{T} \alpha_{mns} \sum_{t=1}^{T} Q_{mt}e_{nt}) \\
&+ \frac{1}{T^3N} \sum_{i\neq j,(i,j)\in S_U} \mathrm{var}(\sum_{s=1}^{T} \alpha_{ijs} \sum_{t=1}^{T} Q_{jt}e_{it}) \equiv B_2 + B_1.
\end{aligned}$$

Using the Cauchy-Schwarz inequality like in the proof of the first statement. Similarly we can show $B_1 = O_p(\frac{1}{T})$. For B_2, let

$$C_{ijmn,stkl} = \mathrm{cov}(\alpha_{ijs}Q_{jt}e_{it}, \alpha_{mnk}Q_{ml}e_{nl}),$$

$$B_2 = \frac{1}{T^3N} \sum_{i\neq j,(i,j)\in S_U} \sum_{m\neq n,(m,n)\in S_U,(m,n)\neq(i,j)} \sum_{stkl\leq T} C_{ijmn,stkl}.$$

It is straightforward to check that when $\{s, t, k, l\}$ contains more than two distinct elements, $C_{ijmn,stkl} = 0$. In addition, $\sum_{i\neq j,(i,j)\in S_U} 1 = O(N)$. Define Θ_1 as the set of (s, t, k, l) such that $\{s, t, k, l\}$ contains exactly two distinct integers. Then

$$B_2 = \frac{1}{T^3N} \sum_{i\neq j,(i,j)\in S_U} \sum_{m\neq n,(m,n)\in S_U,(m,n)\neq(i,j)} \sum_{(s,t,k,l)\in\Theta_1} C_{ijmn,stkl} + O(\frac{N}{T^2}).$$

Moreover, because $\{u_t, Q_t\}_{t\leq T}$ is serially independent, $EQ_{is}e_{js} = 0$ and $Eu_{is}u_{js}e_{ns} = 0$ for all $i, j, n \leq N, s \leq T$, and $m_N = \max_{i\neq N} \sum_{j=1}^{N} I_{\Sigma_{u,ij}\neq 0} = \max_{i\neq N} \sum_{j:(i,j)\in S_U} 1$, we have

$$\begin{aligned}
&B_2 \\
&= \frac{1}{T^3N} \sum_{i\neq j,(i,j)\in S_U} \sum_{m\neq n,(m,n)\in S_U,(m,n)\neq(i,j)} \sum_{s=1}^{T} \sum_{t=1}^{T} (E\alpha_{ijs}\alpha_{mns})(Ee_{it}e_{nt})(EQ_{mt}Q_{jt}) + o(1) \\
&\leq \frac{1}{TN} \max_{ijmnst} |E\alpha_{ijs}\alpha_{mns}||EQ_{mt}Q_{jt}| \sum_{i=1}^{N} \sum_{n=1}^{N} |(\Sigma_u^{-1})_{in}| \sum_{m:(m,n)\in S_U} \sum_{j:(i,j)\in S_U} 1 + o(1) \\
&\leq O(\frac{m_N^2 N}{TN}) \|\Sigma_u^{-1}\| + o(1) = O(\frac{m_N^2}{T}) + o(1) = o(1).
\end{aligned}$$

Therefore, $\mathrm{var}(M) = B_1 + B_2 = o(1)$. This then implies (with $EM = 0$) that $M = o_p(1)$.

C.6 Further technical lemmas

Lemma C.10. *(i)* $\sup_{F'F/T=I_r} \frac{1}{NT}|\mathrm{tr}[\Sigma_u^{-1}u'(F_0F_0' - FF')u/T]| = o_p(1)$.
$\sup_{F'F/T=I_r} \frac{1}{NT}|\mathrm{tr}[\widetilde{\Sigma}_u^{-1}u'(F_0F_0' - FF')u/T]| = o_p(1)$
(ii) $\sup_{F'F/T=I_r} \frac{1}{NT}\|vec(u)'(\Sigma_u^{-1} \otimes M_F)Z\| = o_p(1)$,
$\sup_{F'F/T=I_r} \frac{1}{NT}\|vec(u)'(\widetilde{\Sigma}_u^{-1} \otimes M_F)Z\| = o_p(1)$.
(iii) $\sup_{F'F/T=I_r} \frac{1}{NT}|\mathrm{tr}(\Sigma_u^{-1}u'M_FF_0\Lambda_0')| = o_p(1)$,
$\sup_{F'F/T=I_r} \frac{1}{NT}|\mathrm{tr}(\widetilde{\Sigma}_u^{-1}u'M_FF_0\Lambda_0')| = o_p(1)$.

Proof. (i)

$$(\frac{1}{NT}\|u\Sigma_u^{-1}u'\|_F)^2 \leq \frac{2}{N^2T^2}\sum_{s,t\leq T}(u_t'\Sigma_u^{-1}u_s - Eu_t'\Sigma_u^{-1}u_s)^2 + \frac{2}{N^2T^2}\sum_{s,t\leq T}(Eu_t'\Sigma_u^{-1}u_s)^2.$$

With $W = \Sigma_u^{-1}$, $\frac{2}{N^2T^2}\sum_{s,t\leq T}(u_t'\Sigma_u^{-1}u_s - Eu_t'\Sigma_u^{-1}u_s)^2 = O_p(\frac{1}{N})$. Also,

$$\begin{aligned}\frac{2}{N^2T^2}\sum_{s,t\leq T}(Eu_t'\Sigma_u^{-1}u_s)^2 &\leq \frac{1}{T^2}\sum_{s,t\leq T}|\frac{1}{N}\sum_{i,j\leq N}w_{ij}Eu_{jt}u_{is}|^2 \\ &\leq \frac{1}{T^2}\|\Sigma_u^{-1}\|_1^2 \max_{i,j,s,t}|Eu_{jt}u_{is}| \max_{i,j\leq N}\sum_{s,t\leq T}|Eu_{jt}u_{is}| = O(\frac{1}{T})\end{aligned}$$

which is due to $\|\Sigma_u^{-1}\|_1 < \infty$ and $\max_{t\leq T, i,j\leq N}\sum_{s=1}^T |Eu_{jt}u_{is}| < \infty$. Therefore, using the inequality $|\mathrm{tr}(AB)| \leq \|A\|_F\|B\|_F$, we have

$$\sup_{F'F/T=I_r}\frac{1}{NT}\mathrm{tr}[\Sigma_u^{-1}u'(F_0F_0' - FF')u/T]$$

$$\leq \frac{1}{NT}\|u\Sigma_u^{-1}u'\|_F \sup_{F'F/T=I_r}\|\frac{1}{T}FF'\|_F = O_p(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{T}}).$$

For the second statement, since $\|\Sigma_u^{-1} - \widetilde{\Sigma}_u^{-1}\| = o_p(1)$, it then follows that

$$\sup_{F'F/T=I_r}\frac{1}{NT}|\mathrm{tr}[(\Sigma_u^{-1} - \widetilde{\Sigma}_u^{-1})u'(F_0F_0' - FF')u/T]| = o_p(1),$$

which yields the result.

(ii) Recall $e_t = \Sigma_u^{-1} u_t$. Let $e_i = (e_{i1}, ..., e_{iT})'$ for $i \leq N$. Then

$$\frac{1}{NT}\|vec(u)'(\Sigma_u^{-1} \otimes M_F)Z\| = \frac{1}{NT}\|\sum_{i=1}^{N} e_i' M_F X_i\|.$$

Under the assumption that $E|\frac{1}{\sqrt{N}}(e_t'e_s - Ee_t'e_s)|^2 < \infty$, the same proof of that of Lemma A.1 in Bai (2009) still goes through, which yields the result. The second state follows immediately from $\|\Sigma_u^{-1} - \widetilde{\Sigma}_u^{-1}\| = o_p(1)$.

(iii) By the definition of M_F, we bound

$$a_1 = \sup_{F'F/T=I_r} \frac{1}{NT}|\text{tr}(\Sigma_u^{-1} u' F_0 \Lambda_0')|$$

and

$$a_2 = \sup_{F'F/T=I_r} \frac{1}{NT}|\text{tr}(\Sigma_u^{-1} u' F F'/T F_0 \Lambda')|.$$

First, $a_1 \leq \sup_{F'F=TI_r} \frac{1}{NT}\|\Lambda_0'\Sigma_u^{-1}\|\|u'F_0\|_F$, which is $o_p(1)$ since $\max_{i\leq N}\|\frac{1}{T}\sum_{t=1}^{T} u_{it} f_t\| = O_p(\sqrt{\frac{\log N}{T}})$. On the other hand, a_2 is bounded by $O_p(\frac{1}{N\sqrt{T}})\|\Lambda_0'\Sigma_u^{-1}u'\|_F$, which is $o_p(1)$. Again, we conclude the proof by noting that $\|\Sigma_u^{-1} - \widetilde{\Sigma}_u^{-1}\| = o_p(1)$.

□

Recall $H = I_T - \frac{1}{T}F_0(Ef_tf_t')^{-1}F_0'$, and $G = \frac{1}{T}F^*F^{*'}$ for $F^* = F_0(VA)^{-1}$.

Lemma C.11. *For each $q \leq d = \dim(\beta)$ and $X_{q,i}' = (X_{i1,q}, ..., X_{iT,q})$,*
(i) $\max_{i,j\leq N}|X_{q,i}' M_{\widehat{F}} u_j| = O_p(\sqrt{T\log N} + T\|\hat{\beta} - \beta_0\| + \frac{T}{\sqrt{N}})$
(ii) $\max_{i,j}|X_{q,i}'(R - G)u_j| = O_p(\sqrt{T\log N}(\|\hat{\beta} - \beta_0\| + \omega))$

Proof. (i) The proof is a straightforward calculation, and very similar to that of Lemma C.2 (iii).
(ii) Because $I_r = \frac{1}{T}\widehat{F}'\widehat{F}$, $Ef_tf_t' = O_p(\frac{1}{\sqrt{T}}) + \frac{1}{T}F_0'F_0$, and

$$\frac{1}{\sqrt{T}}\|\widehat{F} - F_0(VA)^{-1}\| = O_p(\|\hat{\beta} - \beta_0\| + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{T}}),$$

we have

$$H - G = \frac{1}{T}F_0((VA)^{-1}((VA)')^{-1} - (Ef_tf_t')^{-1})F_0' = O_p(\|\hat{\beta} - \beta_0\| + \omega_T)\frac{1}{\sqrt{T}}F_0'$$

which implies the result since $\max_j \frac{1}{\sqrt{T}}\|F_0'u_j\| = O_p(\sqrt{\log N})$.

$\square$

Lemma C.12. *When either $Q_j = \Sigma_{u,j}^{-1'} X$ or $Q_j = -\Sigma_{u,j}^{-1'}(EX_t f_t')(Ef_t f_t')^{-1} F_0'$,*

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it}^2 - Eu_{it}^2) = o_p(1)$$

Proof. First we emphasize that $\widehat{u}_{it} = y_{it} - X_{it}'\hat{\beta}_0 - \widehat{\lambda}_i'\widehat{f}_t$, where $(\hat{\beta}_0, \widehat{\lambda}_i, \widehat{f}_t)$ are obtained in the first-step estimation (that is, by the method of Bai 2009). Throughout Lemmas C.12 and C.13, these notation have the same meanings, without causing confusions. It is easy to show: there is an invertible matrix H so that (for $\omega_T = \frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}}$)

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}(\widehat{f}_t - Hf_t)^2 = O_p(\omega_T^2), \quad \max_{i \le N} \frac{1}{T}\sum_{t=1}^{T}(\widehat{u}_{it} - u_{it})^2 = O_p(\omega_T^2), \\
&\max_{i \le N} |\widehat{\lambda}_i - H'^{-1}\lambda_i| = O_p(\omega_T), \\
&\widehat{f}_s - Hf_s = \frac{1}{TN}\sum_{t=1}^{T}\widehat{f}_t(u_s'u_t + f_t\Lambda'u_s + f_s'\Lambda'u_t) + R_s
\end{aligned} \tag{C.10}$$

where the remaining term R_s depends on $\hat{\beta}_0 - \beta$, which can be negligible because it is $O_p(\frac{1}{\sqrt{NT}})$ uniformly in s. The proof for the above results follows exactly the same lines as those of Fan et al. (2013), noting that the effect of estimating β_0 by $\hat{\beta}_0$ is asymptotically negligible because $\hat{\beta}_0$ is $\sqrt{NT}$-consistent according to Bai (2009). We omit the details to avoid repetitions.

Now $\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it}^2 - Eu_{it}^2)$ is bounded by

$$\left|\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it}^2 - u_{it}^2)\right| + \left|\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{t=1}^{T} (u_{it}^2 - Eu_{it}^2)\right| \equiv B_1 + B_2.$$

Term $B_2 = o_p(1)$ follows from Lemma C.4. Term B_1 is bounded by

$$\left|\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{s=1}^{T} (\widehat{u}_{is} - u_{is})u_{is}\right| + \left|\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it} - u_{it})^2\right| \equiv B_{11} + B_{12}.$$

Note that $\max_i |Q_i e_i| = \max_i |\sum_t Q_{it} e_{it}| = O_p(\sqrt{T \log N})$. So we have $B_{12} = O_p(\sqrt{N \log N}\omega_T^2) = o(1)$ given $N \log N = o(T^2)$.

It then suffices to show $B_{11} = o(1)$. This part is difficult, and we separate it into a number of steps. Note that

$$u_{is}-\widehat{u}_{is}=(\widehat{f}_s-Hf_s)'(\widehat{\lambda}_i-H^{'-1}\lambda_i)+(\widehat{f}_s-Hf_s)'H^{'-1}\lambda_i$$

$$+f_s'H'(\widehat{\lambda}_i-H^{'-1}\lambda_i)+X_{it}'(\hat{\beta}_0-\beta).$$

We consider these terms one by one. By Cauchy Schwarz inequality,

$$|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T(\widehat{f}_s-Hf_s)'(\widehat{\lambda}_i-H^{'-1}\lambda_i)u_{is}|$$

$$\leq \max_i|\widehat{\lambda}_i-H^{'-1}\lambda_i|\max_i|Q_ie_i|(\frac{1}{T}\sum_s(\widehat{f}_s-Hf_s)^2)^{1/2}(\frac{1}{T}\sum_s u_{is}^2)^{1/2}\frac{\sqrt{N}}{\sqrt{T}}=o_p(1).$$

Second, $|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T u_{is}f_s'H'(\widehat{\lambda}_i-H^{'-1}\lambda_i)|=o_p(1)$ because $\max_i\|\frac{1}{T}\sum_{s=1}^T u_{is}f_s\|=O_p(\sqrt{\frac{\log N}{T}})$. The term of $X_{it}'(\hat{\beta}_0-\beta)$ is negligible. We now work on the term of $(\widehat{f}_s-Hf_s)'H^{'-1}\lambda_i$. By the formula $\widehat{f}_s-Hf_s=\frac{1}{TN}\sum_{t=1}^T\widehat{f}_t(u_s'u_t+f_t\Lambda'u_s+f_s'\Lambda'u_t)+R_s$,

$$|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T(\widehat{f}_s-Hf_s)'H^{'-1}\lambda_iu_{is}|\leq\sum_{i=1}^4 C_i.$$

Using (C.10) and by adding and subtracting terms,

$$\begin{aligned}
C_1&=|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T\lambda_i'H^{-1}\frac{1}{TN}\sum_{t=1}^T\widehat{f}_tu_s'u_tu_{is}|\\
&\leq|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\lambda_i'H^{-1}\frac{1}{T^2N}\sum_{s=1}^T(\widehat{f}_s-Hf_s)u_{is}E(u_s'u_s)|\\
&+|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\lambda_i'\frac{1}{TN}\frac{1}{T}\sum_{s=1}^T f_su_{is}E(u_s'u_s)|\\
&+|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T\lambda_i'u_{is}\frac{1}{TN}\sum_{t=1}^T f_t(u_s'u_t-Eu_s'u_t)|\\
&+|\frac{2}{\sqrt{NT}}\sum_{i=1}^N Q_ie_i\frac{1}{T}\sum_{s=1}^T\lambda_i'H^{-1}u_{is}\frac{1}{TN}\sum_{t=1}^T(\widehat{f}_t-Hf_t)(u_s'u_t-Eu_s'u_t)|\\
&=\sum_{i=1}^4 C_{1i}.
\end{aligned}$$

By Cauchy-Schwarz inequality, $C_{11}, C_{12} = o_p(1)$. Also,

$$C_{13} \leq \max_{is} |u_{is}| O_p(\frac{N\sqrt{T \log N}}{NT}) \frac{1}{T} \sum_{s=1}^{T} |\frac{1}{\sqrt{TN}} \sum_{t=1}^{T} f_t(u_s' u_t - Eu_s' u_t)|$$

Note that

$$E\frac{1}{T}\sum_{s=1}^{T} |\frac{1}{\sqrt{TN}} \sum_{t=1}^{T} f_t(u_s' u_t - Eu_s' u_t)| \leq (E|\frac{1}{\sqrt{TN}} \sum_{t=1}^{T} f_t(u_s' u_t - Eu_s' u_t)|^2)^{1/2} = O(1)$$

So $C_{13} = O_p(\sqrt{\frac{\log N}{T}}(\log NT)) = o_p(1)$. Since $E(\frac{1}{\sqrt{N}}(u_s' u_t - Eu_s' u_t)^2) = O(1)$, by Cauchy-Schwarz inequality, $C_{14} = o_p(1)$.

$$\begin{aligned} C_2 &= |\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{s=1}^{T} \lambda_i' H^{-1} \frac{1}{TN} \sum_{t=1}^{T} \widehat{f}_t f_t' \Lambda' u_s u_{is}| \\ &\leq |\frac{2}{TN} \sum_{i=1}^{N} Q_i e_i \lambda_i' H^{-1} \frac{1}{T} \sum_{t=1}^{T} \widehat{f}_t f_t' \frac{1}{\sqrt{NT}} \sum_{s=1}^{T} \sum_{j=1}^{N} \lambda_j (u_{js} u_{is} - Eu_{js} u_{is})| \\ &\quad +|\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \lambda_i' H^{-1} \frac{1}{TN} \sum_{t=1}^{T} \widehat{f}_t f_t' \sum_{j=1}^{N} \lambda_j E u_{js} u_{is}| = o_p(1). \end{aligned}$$

The first term is $o_p(1)$ because

$$E\frac{1}{N}\sum_{i=1}^{N} \|\frac{1}{\sqrt{NT}} \sum_{s=1}^{T} \sum_{j=1}^{N} \lambda_j (u_{js} u_{is} - Eu_{js} u_{is})\|^2 = O(1)$$

and $\max_i |Q_i e_i| = O_p(\sqrt{T \log N})$. The second term is $o_p(1)$ because

$$\max_i \sum_j |Eu_{js} u_{is}| = \|\Sigma_u\|_1 = O(1).$$

$$\begin{aligned} &C_3 = |\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{s=1}^{T} \lambda_i' H^{-1} \frac{1}{TN} \sum_{t=1}^{T} \widehat{f}_t u_t' \Lambda f_s u_{is}| \\ &\leq |\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \lambda_i' \frac{1}{TN} \sum_{t=1}^{T} f_t u_t' \Lambda \frac{1}{T} \sum_{s=1}^{T} f_s u_{is}| \\ &+|\frac{2}{\sqrt{NT}} \sum_{i=1}^{N} Q_i e_i \frac{1}{T} \sum_{s=1}^{T} \lambda_i' H^{-1} \frac{1}{TN} \sum_{t=1}^{T} (\widehat{f}_t - Hf_t) u_t' \Lambda f_s u_{is}| = o_p(1). \end{aligned}$$

The last term involving R_s is negligible. This concludes the proof. $\square$

Lemma C.13. *When either* $Q_j = \Sigma_{u,j}^{-1\prime} X$ *or* $Q_j = -\Sigma_{u,j}^{-1\prime}(EX_t f_t')(Ef_t f_t')^{-1} F_0'$,

$$\frac{1}{\sqrt{NT}} \sum_{i\neq j,(i,j)\in S_U} Q_j e_i(\widetilde{\Sigma}_{u,ij} - \Sigma_{u,ij}) = o_p(1)$$

Proof. The term of interest is bounded by

$$\begin{aligned}
&|\frac{1}{\sqrt{NT}} \sum_{i\neq j,(i,j)\in S_U} Q_j e_i(\frac{1}{T}\sum_{t=1}^{T} u_{it}u_{jt} - \Sigma_{u,ij})| \\
&+|\frac{1}{\sqrt{NT}} \sum_{i\neq j,(i,j)\in S_U} Q_j e_i(\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} - u_{it}u_{jt})| \\
&+|\frac{1}{\sqrt{NT}} \sum_{i\neq j,(i,j)\in S_U} Q_j e_i(\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} - \widetilde{\Sigma}_{u,ij})| \equiv D_1 + D_2 + D_3.
\end{aligned}$$

Term $D_1 = o_p(1)$ follows from Lemma C.4. From now on, we consider the hard-thresholding, that is, for $i \neq j$,

$$\widetilde{\Sigma}_{u,ij} = \frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} I(|\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt}| > \tau_{ij}\omega_T)$$

where τ_{ij} is the threshold constant such that $P(\tau_{ij} < C_1) \to 1$ for some $C_1 > 0$. General thresholding functions can be treated very similarly as in the proof of Lemma B.4. For D_3, we have, for any $\epsilon > 0$,

$$\begin{aligned}
&P(D_3 > T^{-1}) \leq P(\max_{i\neq j,(i,j)\in S_U} |\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} - \widetilde{\Sigma}_{u,ij}| > 0) \\
&\leq P(\exists (i,j) \in S_U, |\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt}| \leq \tau_{ij}\omega_T) \\
&\leq P(\min_{(i,j)\in S_U} |\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt}| \leq C_1\omega_T) + o(1) \\
&\leq P(\min_{(ij)\in S_U} |\Sigma_{u,ij}| - \max_{ij} |\frac{1}{T}\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}_{jt} - \Sigma_{u,ij}| \\
&\leq C_1\omega_T) + o(1) \leq \epsilon + o(1),
\end{aligned}$$

where we use the assumption that $\omega_T = o(\min_{(ij)\in S_U} |\Sigma_{u,ij}|)$. This proves $D_3 = O_p(\frac{1}{T})$. The proof of D_2 follows the same lines of that of term B_1 in Lemma C.12, hence is omitted.

$\square$

D Heteroskedastic WPC

A simple modification to improve the regular PC when cross-sectional heteroskedasticity is present is choosing

$$W = W^h \equiv (\mathrm{diag}(\Sigma_u))^{-1},$$

which can be consistently estimated as follows. First apply the regular PC by taking $W = W_T = I_N$, and obtain consistent estimator $\widehat{C}_{it}$ of the common component $\lambda_i' f_t$ for each $i \leq N, t \leq T$. Define the heteroskedastic weight matrix W_T to be:

$$W_T^h = \mathrm{diag}\{\widehat{\sigma}_{u,11}^{-1}, ..., \widehat{\sigma}_{u,NN}^{-1}\}, \text{ where } \widehat{\sigma}_{u,ii} = \frac{1}{T}\sum_{t=1}^{T}(y_{it} - \widehat{C}_{it})^2.$$

Then in the second step, estimate the factors and loadings with the weight matrix W_T^h.

Let $\widehat{f}_t^h$ and $\widehat{\lambda}_j^h$ denote the WPC estimators for f_t and λ_j with weight $W_T = W_T^h$. Here the superscript h denotes "heteroskedastic PC". To be more specifically, the columns of the $T \times r$ matrix $\widehat{F}^h/\sqrt{T} = (\widehat{f}_1^h, ..., \widehat{f}_T^h)'/\sqrt{T}$ are the eigenvalues corresponding to the largest r eigenvalues of $Y'W_T^hY$, and $\widehat{\Lambda}^h = T^{-1}Y\widehat{F}^h = (\widehat{\lambda}_1^h, ..., \widehat{\lambda}_N^h)'$. We thus term this estimator to be "HPC".

The following assumptions are made, which guarantees the consistency of W_T^h.

Assumption D.1. *(i)* $E\|\frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1}^T(u_{is}^2 - Eu_{is}^2)\sigma_{u,ii}^2\lambda_i u_{it}\| = O(1)$.
(ii) For each $k \leq r$, $E\|\frac{1}{N\sqrt{TN}}\sum_{i=1}^N\sum_{s=1}^T\sum_{j=1}^N(u_{js}u_{is} - Eu_{js}u_{is})\lambda_{ik}\sigma_{u,ii}u_{it}\lambda_j\lambda_i'\|_F = O(1)$.
(iii) $E\|\frac{1}{\sqrt{TN}}\sum_{t=1}^T f_t(u_s'u_t - Eu_s'u_t)\|^2 = O(1)$.

The following result shows that for this choice of W and W_T, the required convergence in Section 2 for the estimated weight matrix (Assumption 3.1) is satisfied.

Lemma D.1. *Let* $W^h = (\mathrm{diag}(\Sigma_u))^{-1}$ *and* $W_T^h = \mathrm{diag}\{\widehat{\sigma}_{u,11}^{-1}, ..., \widehat{\sigma}_{u,NN}^{-1}\}$. *we have*

$$\|W^h - W_T^h\| = O_p(\frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}}),$$

For each $t \leq T$,

$$\|\frac{1}{\sqrt{N}}\Lambda'(W_T^h - W^h)u_t\| = O_p(\frac{1}{\sqrt{N}} + \frac{\sqrt{N}\log N}{T}).$$

Therefore Assumption 3.1 are satisfied when $N(\log N)^2 = o(T^2)$ *and* $T = o(N^2)$.

We now present the limiting distribution for the HPC estimator. Let $\widehat{V}_h$ be the $r \times r$ diagonal matrix of the first r largest eigenvalues of $YW_T^hY'/(TN)$. Let H_h be the $r\times r$ matrix H_W as defined in Section 2 with $W_T = W_T^h$. Specifically, $H_h = \widehat{V}_h^{-1}\frac{1}{T}\sum_{t=1}^T\widehat{f}_t^h f_t'\Lambda'W_T^h\Lambda/N$.

Theorem D.1. *Let Q_h be defined as the same as Q_W with $W = W^h$. For each $t \leq T$ and $j \leq N$,*

$$\sqrt{T}(\widehat{\lambda}_j^h - H_h'^{-1}\lambda_j) \to^d \mathcal{N}(0, Q_h'^{-1}\Phi_j Q_h^{-1}).$$

$$N(V^{-1}Q_h\Lambda' W^h \Sigma_u W^h \Lambda Q_h' V^{-1})^{-1/2}(\widehat{f}_t^h - H_h f_t) \to^d \mathcal{N}(0, I_r).$$

$$\frac{\widehat{\lambda}_i^{h\prime}\widehat{f}_t^h - \lambda_i' f_t}{(\lambda_i'\Xi_h\lambda_i/N + f_t'\Omega_i f_t/T)^{1/2}} \to^d \mathcal{N}(0,1).$$

where $\Xi_h = (\Sigma_\Lambda^h)^{-1}\Lambda' W^h \Sigma_u W^h \Lambda(\Sigma_\Lambda^h)^{-1}/N$, and $\Lambda' W^h \Lambda/N \to \Sigma_\Lambda^h$; Ω_i is defined as in Theorem 3.1.

Numerically, the HPC method improves the finite sample performance from the regular PC method.

Proof of Lemma D.1

First, it was shown by Fan et al. (2013) that $\max_{i\leq N}\widehat{\sigma}_{u,ii}^{-1} = O_p(1)$. Hence $\|W_T^h - W^h\| = \max_{i\leq N}|\widehat{\sigma}_{u,ii}^{-1} - \sigma_{u,ii}^{-1}| = O_p(\max_{i\leq N}|\widehat{\sigma}_{u,ii} - \sigma_{u,ii}|)$. Let $\widehat{u}_{it} = y_{it} - \widehat{C}_{it}$ be the estimated error using the regular PC as in Bai (2003). Then $\widehat{\sigma}_{u,ii} = \frac{1}{T}\sum_{t=1}^T \widehat{u}_{it}^2$. The triangular and Cauchy-Schwarz inequalities imply

$$\max_{i\leq N}|\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it}^2 - u_{it}^2)| \leq \max_{i\leq N}\left[\left(\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}^2\right)^{1/2} + \left(\frac{1}{T}\sum_{t=1}^T u_{it}^2\right)^{1/2}\right] \left(\frac{1}{T}\sum_{t=1}^T(u_{it} - \widehat{u}_{it})^2\right)^{1/2}.$$

On one hand, $\max_{i\leq N}\frac{1}{T}\sum_{t=1}^T\widehat{u}_{it}^2 = O_p(1) = \max_{i\leq N}\frac{1}{T}\sum_{t=1}^T u_{it}^2$. On the other hand, all the conditions in Fan et al. (2013) are satisfied under our assumption, and thus by Lemma C.11 of Fan et al. (2013), $\max_{i\leq N}\frac{1}{T}\sum_{t=1}^T(u_{it} - \widehat{u}_{it})^2 = O_p(1/N + \log N/T)$. Finally, since $\max_{i\leq N}|\frac{1}{T}\sum_{t=1}^T u_{it}^2 - Eu_{it}^2| = O_p(\sqrt{\log N/T})$ (see Lemma A.9), we have $\max_{i\leq N}|\widehat{\sigma}_{u,ii} - \sigma_{u,ii}| = O_p(1/\sqrt{N} + \sqrt{\log N/T})$. This yields the desired rate for $\|W_T^h - W^h\|$.

Let H_I denote H_W when $W_T = I_r$ is used as the weight matrix, where the subscript I denotes the "identity weight matrix". Let $\widehat{f}_t^I$ and $\widehat{\lambda}_j^I$ denote the regular PC estimators for the transformed factors and loadings as in Stock and Watson (2002), which correspond to the weighted PC estimators with $W_T = W = I_r$. As shown in Bai (2003)'s Appendix C,

$$u_{it} - \widehat{u}_{it} = (\widehat{f}_t^I - H_I f_t)' H_I'^{-1}\lambda_i + f_t' H_I'(\widehat{\lambda}_i^I - H_I'^{-1}\lambda_i) + (\widehat{f}_t^I - H_I f_t)'(\widehat{\lambda}_i^I - H_I'^{-1}\lambda_i). \quad \text{(D.1)}$$

Lemma D.2. *For each $t \leq T$,*
(i) $\|\frac{1}{N^2T}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I\prime}(u_s'u_l - Eu_s'u_l)H_I'^{-1}\lambda_i u_{is}\sigma_{u,ii}\lambda_i u_{it}\| = O_p(\frac{\log N}{T} + \frac{1}{N})$.

(ii) $\|\frac{1}{N^2T}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}(Eu_s'u_l)H_I^{'-1}\lambda_iu_{is}\sigma_{u,ii}\lambda_iu_{it}\| = O_p(\frac{\log N}{T}+\frac{1}{N})$.
(iii) $\|\frac{1}{N^2T}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}f_l'\Lambda'u_sH_I^{'-1}\lambda_iu_{is}\sigma_{u,ii}\lambda_iu_{it}\| = O_p(1/\sqrt{NT}+1/N)$.
(iv) $\|\frac{1}{N^2T}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^{I'}f_s'\Lambda'u_lH_I^{'-1}\lambda_iu_{is}\sigma_{u,ii}\lambda_iu_{it}\| = O_p(\log N/T+1/N)$.

Proof. We can replace $\widehat{f}_l^{I'}$ in each stated term with f_l', because as shown by Fan et al. (2013), $\frac{1}{T}\sum_{l=1}^T\|\widehat{f}_l^I-f_l\|^2 = O_p(\log N/T+1/N)$. Thus by Cauchy-Schwarz inequality, such a replacement will introduce an error $O_p(\log N/T+1/N)$.
(i) By the Cauchy-Schwarz inequality, the objective is bounded by
$O_p(\frac{\log N}{T}+\frac{1}{N})$ plus

$$\left[\frac{1}{T}\sum_{s=1}^T\|\frac{1}{T}\sum_{l=1}^Tf_l'(u_s'u_l-Eu_s'u_l)H_I^{'-1}/N\|^2\right]^{1/2}\left[\frac{1}{T}\sum_{s=1}^T(\frac{1}{N}\sum_{i=1}^N\|\lambda_i\|^2\sigma_{u,ii}^2|u_{it}u_{is}|)^2\right]^{1/2}.$$

The second term is $O_p(1)$. By the assumption that
$E\|\frac{1}{\sqrt{TN}}\sum_{l=1}^Tf_l'(u_s'u_l-Eu_s'u_l)\|^2 = O(1)$, the first term is $O_p(1/\sqrt{NT})$, which yields the result.
(ii) The objective is bounded by $\frac{1}{N^2T^2}\sum_{i=1}^N\sum_{s,l}^T\|f_l'Eu_s'u_l\|\|\lambda_i\|^2|u_{is}u_{it}\sigma_{u,ii}|+O_p(\log N/T+1/N)$. Note that $E\sum_{l=1}^T\|f_l'Eu_s'u_l/N\| = O(1)$ by the strong mixing condition. This gives the result.
(iii) The term in $\|.\|$ is an $r\times 1$ vector. Let a_k denote its kth element, $k\le r$. Then $a_k = \mathrm{tr}(a_k) = \frac{1}{NT}\sum_{i=1}^N\sum_{l=1}^T\lambda_{ik}\sigma_{u,ii}\lambda_i'H_I^{'-1}\widehat{f}_l^If_l'\frac{1}{TN}\sum_{s=1}^T\sum_{j=1}^N\lambda_ju_{js}u_{is}$. Using the inequality that $|\mathrm{tr}(AB)|\le\|A\|_F\|B\|_F$, we have

$$\begin{aligned}|a_k| &= |\mathrm{tr}(a_k)| = |\mathrm{tr}(\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^If_l'\frac{1}{N}\sum_{i=1}^N\frac{1}{TN}\sum_{s=1}^T\sum_{j=1}^N\lambda_ju_{js}u_{is}\lambda_{ik}\sigma_{u,ii}\lambda_i'H_I^{'-1})| \\ &\le \|\frac{1}{T}\sum_{l=1}^T\widehat{f}_l^If_l'\|_F\|\frac{1}{N}\sum_{i=1}^N\frac{1}{TN}\sum_{s=1}^T\sum_{j=1}^N\lambda_ju_{js}u_{is}\lambda_{ik}\sigma_{u,ii}\lambda_i'\|_F\|H_I^{'-1}\|_F \\ &= O_p(1)\|\frac{1}{N}\sum_{i=1}^N\frac{1}{TN}\sum_{s=1}^T\sum_{j=1}^Nu_{js}u_{is}\lambda_{ik}\sigma_{u,ii}\lambda_j\lambda_i'\|_F. \end{aligned} \tag{D.2}$$

By the assumption that
$\|\frac{1}{N}\sum_{i=1}^N\frac{1}{\sqrt{TN}}\sum_{s=1}^T\sum_{j=1}^N(u_{js}u_{is}-Eu_{js}u_{is})\lambda_{ik}\sigma_{u,ii}\lambda_j\lambda_i'\|_F = O_p(1)$, and $\max_{i\le N}\sum_{j=1}^N|Eu_{js}u_{is}| = O(1)$, it follows from the triangular inequality that $|a_k| = O_p(1/N+1/\sqrt{NT})$. Since each element a_k is $O_p(1/\sqrt{NT}+1/N)$ and there are finitely many elements ($k\le r$), the desired result follows.
(iv) It follows directly from the rate of convergence $\|\frac{1}{T}\sum_{s=1}^Tf_su_s'\| = O_p(\sqrt{N(\log N)/T})$.
□

Rate for $\|\Lambda'(W_T^h - W^h)u_t/N\|$

Note that

$$\|\Lambda'(W_T^h - W^h)u_t/N\| \leq \|\Lambda' W_T^h((W^h)^{-1} - (W_T^h)^{-1})W^h u_t/N\| \leq a + b,$$

where $a = \|\Lambda' W^h((W^h)^{-1} - (W_T^h)^{-1})W^h u_t/N\|$, and

$$b = \|\Lambda'(W_T^h - W^h)((W^h)^{-1} - (W_T^h)^{-1})W^h u_t/N\|.$$

Since $\lambda_{\min}(W^h)$ is bounded away from zero, thus $\|(W^h)^{-1} - (W_T^h)^{-1}\| = O_p(1/\sqrt{N} + \sqrt{\log N/T})$. This implies $b = O_p(1/N + \log N/T)$. We now bound a.

In fact, $a = \|\frac{1}{N}\sum_{i=1}^N (\widehat{\sigma}_{u,ii} - \sigma_{u,ii})\sigma_{u,ii}^2 \lambda_i u_{it}\|$. By the triangular inequality,

$$a \leq \|\frac{1}{N}\sum_{i=1}^N \frac{1}{T}\sum_{s=1}^T (\widehat{u}_{is}^2 - u_{is}^2)\sigma_{u,ii}^2 \lambda_i u_{it}\| + \|\frac{1}{N}\sum_{i=1}^N \frac{1}{T}\sum_{s=1}^T (u_{is}^2 - Eu_{is}^2)\sigma_{u,ii}^2 \lambda_i u_{it}\|$$

By the assumption that $\|\frac{1}{\sqrt{NT}}\sum_{i=1}^N \sum_{s=1}^T (u_{is}^2 - Eu_{is}^2)\sigma_{u,ii}^2 \lambda_i u_{it}\| = O_p(1)$, the second term is $O_p(1/\sqrt{NT})$. The first term is bounded by

$$\|\frac{1}{NT}\sum_{i=1}^N \sum_{s=1}^T (\widehat{u}_{is} - u_{is})^2 \sigma_{u,ii}^2 \lambda_i u_{it}\| + \|\frac{2}{NT}\sum_{i=1}^N \sum_{s=1}^T (\widehat{u}_{is} - u_{is})u_{is}\sigma_{u,ii}^2 \lambda_i u_{it}\| \equiv a_1 + a_2.$$

We have $a_1 = O_p(\log N/T + 1/N)$. On the other hand, it was shown by Bai (2003) and Fan et al. (2013) that the third term in (D.1) is $O_p(\log N/T + 1/N)$. Hence by (D.1), $a_2 = a_{21} + a_{22} + O_p(\log N/T + 1/N)$, where
$a_{21} = \|\frac{2}{NT}\sum_{i=1}^N \sum_{s=1}^T (\widehat{f}_s^I - H_I f_s)' H_I'^{-1} \lambda_i u_{is}\sigma_{u,ii}\lambda_i u_{it}\|$, and
$a_{22} = \|\frac{2}{NT}\sum_{i=1}^N \sum_{s=1}^T f_s' H_I'(\widehat{\lambda}_i^I - H_I'^{-1}\lambda_i)u_{is}\sigma_{u,ii}\lambda_i u_{it}\| = O_p(\log N/T + \sqrt{\log N/(NT)})$, where we apply the convergence rates for $\frac{1}{T}\sum_{s=1}^T f_s u_{is}$ and $\widehat{\lambda}_i^h - H_I'^{-1}\lambda_i$.

It remains to bound a_{21}. Due to the equality (A.1) of Bai (2003), there is an $r \times r$ matrix V_h with $\|V_h\| = O_p(1)$ such that

$$\widehat{f}_s^I - H_I f_s = V \frac{1}{T}\sum_{l=1}^T \widehat{f}_l^I (u_s' u_l + f_l'\Lambda' u_s + f_s'\Lambda' u_l)/N. \tag{D.3}$$

It then follows from Lemma D.2 that $a_{21} = O_p(\log N/T + 1/N + 1/\sqrt{NT})$. Summarizing the above results, we obtain
$\|\Lambda'(W_T^h - W^h)u_t/N\| = O_p(1/N + (\log N)/T)$.