Supplement for "Augmented Factor Models with Applications to Validating Market Risk Factors and Forecasting Bond Risk Premia"

Jianqing Fan, Yuan Ke, and Yuan Liao

Abstract

This document contains all the technical Lemmas.

Contents

C Technical Results for Section 4

C.1 Bahadur representation of the robust estimator

The main goal is to achieve an expansion for $\widehat{E}(y_{it}|\mathbf{x}_t) - E(y_{it}|\mathbf{x}_t)$ (Proposition C.3). This requires the rates for $\max_{i\leq N}\|\mathbf{b}_{i,\alpha} - \mathbf{b}_i\|$, $\max_{i\leq N}\|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|$, and an expansion of $\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha}$. These are given in the propositions below.

Proposition C.1. *For any* $4 < k < \zeta_2 + 2$,

$$\max_{i\leq N}\|\mathbf{b}_{i,\alpha} - \mathbf{b}_i\| = O(\alpha_T^{-(k-1)}).$$

Proof. Let

$$z_{it} := E(y_{it}|\mathbf{x}_t) - \mathbf{b}_i'\Phi(\mathbf{x}_t).$$

We first prove that for any $0 < k < \zeta_2 + 2$, $\max_{i\leq N}\sup_{\mathbf{x}} E(|e_{it}|^k|\mathbf{x}_t = \mathbf{x}) < \infty$. In fact, uniformly in $\mathbf{x}$ for $\mathbf{x}_t = \mathbf{x}$ and $i \leq N$, as long as $\zeta_2 + 2 > k$

$$\begin{aligned}
E(|e_{it}|^k|\mathbf{x}_t) &= \int_0^\infty P(|e_{it}|^k > x|\mathbf{x}_t)dx \\
&\leq 1 + \int_1^\infty P(|e_{it}|^k > x|\mathbf{x}_t)dx \\
&\leq 1 + \int_1^\infty E(e_{it}^2 1\{|e_{it}| > x^{1/k}\}|\mathbf{x}_t)x^{-2/k}dx \\
&\leq 1 + \int_1^\infty Cx^{-(\zeta_2+2)/k}dx < \infty.
\end{aligned}$$

Since $\zeta_2 > 2$ by assumption, there is $k > 4$ so that $\max_{i\leq N}\sup_{\mathbf{x}} E(|e_{it}|^k|\mathbf{x}_t = \mathbf{x}) < \infty$.

Now recall that $\mathbf{b}_i = \arg\min E(y_{it} - \mathbf{b}_i'\Phi(\mathbf{x}_t))^2$. Hence

$$\begin{aligned}
E[(y_{it} - \mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t))^2 - (y_{it} - \mathbf{b}_i'\Phi(\mathbf{x}_t))^2] &= (\mathbf{b}_{i,\alpha}' - \mathbf{b}_i)'E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'(\mathbf{b}_{i,\alpha} - \mathbf{b}_i) \\
&\geq \underline{c}\|\mathbf{b}_{i,\alpha} - \mathbf{b}_i\|^2
\end{aligned}$$

On the other hand, let $g_\alpha(z) := z^2 - \alpha_T^2\rho(z/\alpha_T)$. Then for $C > 0$ as a generic constant,

$$\begin{aligned}
& E[(y_{it} - \mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t))^2 - (y_{it} - \mathbf{b}_i'\Phi(\mathbf{x}_t))^2] \\
= \; & Eg_\alpha(y_{it} - \mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t)) - Eg_\alpha(y_{it} - \mathbf{b}_i'\Phi(\mathbf{x}_t))
\end{aligned}$$

$$
\begin{aligned}
& +E[\alpha_T^2\rho(\alpha_T^{-1}(y_{it}-\mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t)))-\alpha_T^2\rho(\alpha_T^{-1}(y_{it}-\mathbf{b}_i'\Phi(\mathbf{x}_t)))] \\
\leq_{(1)} \quad & Eg_\alpha(y_{it}-\mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t))-Eg_\alpha(y_{it}-\mathbf{b}_i'\Phi(\mathbf{x}_t)) \\
\leq_{(2)} \quad & E[|2\tilde{z}-\alpha_T\dot{\rho}(\alpha_T^{-1}\tilde{z})||\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b}_{i,\alpha})|], \\
\leq_{(3)} \quad & 2\alpha_T^{-(k-1)}E|\tilde{z}|^k|\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b}_{i,\alpha})| \\
\leq_{(4)} \quad & 2\alpha_T^{-(k-1)}E|z_{it}+e_{it}+(\mathbf{b}_i-\widetilde{\mathbf{b}}_i)'\Phi(\mathbf{x}_t)|^k|\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b}_{i,\alpha})| \\
\leq \quad & C\alpha_T^{-(k-1)}E(C+|(\mathbf{b}_i-\widetilde{\mathbf{b}}_i)'\Phi(\mathbf{x}_t)|^k)|\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b}_{i,\alpha})|
\end{aligned}
$$

where (1) is due to the definition of $\mathbf{b}_{i,\alpha}$; (2) is by the mean value representation: $g_\alpha(z_1)-g_\alpha(z_2)=(2\tilde{z}-\alpha_T\dot{\rho}(\tilde{z}/\alpha_T))(z_1-z_2)$, with $z_1=y_{it}-\mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t)$, $z_2=y_{it}-\mathbf{b}_i'\Phi(\mathbf{x}_t)$, and $\tilde{z}=y_{it}-\widetilde{\mathbf{b}}_i'\Phi(\mathbf{x}_t)$ for some $\widetilde{\mathbf{b}}_i$ lying between $\mathbf{b}_i$ and $\mathbf{b}_{i,\alpha}$; (3) is due to

$$
\begin{aligned}
|2\tilde{z}-\alpha_T\dot{\rho}(\alpha_T^{-1}\tilde{z})| &\leq 2|\tilde{z}|1\{|\tilde{z}|>\alpha_T\} \\
&\leq 2|\tilde{z}|\frac{|\tilde{z}|^{k-1}}{\alpha_T^{k-1}}1\{|\tilde{z}|>\alpha_T\} \\
&\leq 2|\tilde{z}|^k/\alpha_T^{k-1}.
\end{aligned}
$$

(4) follows from $\tilde{z}=y_{it}-E(y_{it}|\mathbf{x}_t)+\mathbf{b}_i'\Phi(\mathbf{x}_t)+z_{it}-\widetilde{\mathbf{b}}_i'\Phi(\mathbf{x}_t)$, and that $e_{it}:=y_{it}-E(y_{it}|\mathbf{x}_t)$.

Next, for ease of presentation, we introduce $M_{it}:=C+|(\mathbf{b}_i-\widetilde{\mathbf{b}}_i)'\Phi(\mathbf{x}_t)|^k$ and $\boldsymbol{\Delta}_i:=\mathbf{b}_i-\mathbf{b}_{i,\alpha}$. Then the above inequality can be further written as:

$$
\begin{aligned}
& E[(y_{it}-\mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t))^2-(y_{it}-\mathbf{b}_i'\Phi(\mathbf{x}_t))^2] \\
\leq \; & C\alpha_T^{-(k-1)}EM_{it}|\Phi(\mathbf{x}_t)'\boldsymbol{\Delta}_i| \\
= \; & C\alpha_T^{-(k-1)}E[M_{it}^2\boldsymbol{\Delta}_i'\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\boldsymbol{\Delta}_i]^{1/2} \\
\leq \; & C\alpha_T^{-(k-1)}[\boldsymbol{\Delta}_i'EM_{it}^2\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\boldsymbol{\Delta}_i]^{1/2} \\
\leq \; & C\alpha_T^{-(k-1)}\|EM_{it}^2\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|^{1/2}\|\boldsymbol{\Delta}_i\|.
\end{aligned}
$$

We now bound $\max_{i\leq N}\|EM_{it}^2\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|=\max_{i\leq N}\sup_{\|\boldsymbol{\nu}\|=1}EM_{it}^2(\Phi(\mathbf{x}_t)'\boldsymbol{\nu})^2$. By the Cauchy-Schwarz inequality, since $\Phi(\mathbf{x}_t)'\boldsymbol{\nu}$ is sub-Gaussian with the universal parameter,

$$
\begin{aligned}
& \sup_{\|\boldsymbol{\nu}\|=1}[EM_{it}^2(\Phi(\mathbf{x}_t)'\boldsymbol{\nu})^2]^2\leq EM_{it}^4\sup_{\|\boldsymbol{\nu}\|=1}E(\Phi(\mathbf{x}_t)'\boldsymbol{\nu})^4\leq CEM_{it}^4 \\
\leq \; & C(C+E|(\mathbf{b}_i-\widetilde{\mathbf{b}}_i)'\Phi(\mathbf{x}_t)|^{4k}) \\
\leq \; & C+CE\|\mathbf{b}_i-\widetilde{\mathbf{b}}_i\|^{4k}\left(\frac{(\mathbf{b}_i'-\widetilde{\mathbf{b}}_i)'}{\|\mathbf{b}_i-\widetilde{\mathbf{b}}_i\|}\Phi(\mathbf{x}_t)\right)^{4k}
\end{aligned}
$$

$$\leq \quad C + C\|\boldsymbol{\Delta}_i\|^{4k} \sup_{\|\boldsymbol{\nu}\|=1} E(\boldsymbol{\nu}'\Phi(\mathbf{x}_t))^{4k} \leq C + C\|\boldsymbol{\Delta}_i\|^{4k}.$$

Therefore, we have proved that uniformly in i,

$$\begin{aligned} E[(y_{it} - \mathbf{b}_{i,\alpha}'\Phi(\mathbf{x}_t))^2 - (y_{it} - \mathbf{b}_i'\Phi(\mathbf{x}_t))^2] &\leq C\alpha_T^{-(k-1)}(C + C\|\boldsymbol{\Delta}_i\|^{4k})^{1/4}\|\boldsymbol{\Delta}_i\| \\ &\leq C\alpha_T^{-(k-1)}(1 + \|\boldsymbol{\Delta}_i\|^k)\|\boldsymbol{\Delta}_i\| \end{aligned}$$

We have also proved that the left hand side is lower bounded by $\underline{c}\|\boldsymbol{\Delta}_i\|^2$. Uniformly in i,

$$\|\boldsymbol{\Delta}_i\| \leq C\alpha_T^{-(k-1)}(1 + \|\boldsymbol{\Delta}_i\|^k).$$

If $\max_i \|\boldsymbol{\Delta}_i\| = O(1)$, then $\|\boldsymbol{\Delta}_i\| \leq C\alpha_T^{-(k-1)}$. Otherwise, $\max_i \|\boldsymbol{\Delta}_i\| \leq C\alpha_T^{-(k-1)} \max_i \|\boldsymbol{\Delta}_i\|^k$, which then implies $1 \leq C(\max_i \|\boldsymbol{\Delta}_i\|/\alpha_T)^{k-1}$. However, note that $\|\boldsymbol{\Delta}_i\| \leq \|\mathbf{b}_i\| + \|\mathbf{b}_{i,\alpha}\| \leq CJ^{1/2}$, and $J = o(\alpha_T^2)$, we have $\max_i \|\boldsymbol{\Delta}_i\|/\alpha_T = o(1)$, which is a contradiction. Therefore, $\max_i \|\boldsymbol{\Delta}_i\| \leq C\alpha_T^{-(k-1)}$. Q.E.D.

The following lemma shows the sieve approximation error is uniformly controlled.

Lemma C.1. *Under Assumption 3.2, there is $\eta \geq 1$, as $J \to \infty$,*

$$\max_{i \leq N} \sup_{\mathbf{x}} |E(y_{it}|\mathbf{x}_t = \mathbf{x}) - \mathbf{b}_i'\Phi(\mathbf{x})| = O(J^{-\eta}).$$

Proof. Recall that for $k \leq K$,

$$\mathbf{v}_k = \arg\min_{\mathbf{v}} E(f_{kt} - \mathbf{v}'\Phi(\mathbf{x}_t))^2 = (E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}E\Phi(\mathbf{x}_t)f_{kt}$$

and that $\mathbf{b}_i = \arg\min_{\mathbf{b} \in \mathbb{R}^J} E[y_{it} - \mathbf{b}'\Phi(\mathbf{x}_t)]^2 = (E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}E\Phi(\mathbf{x}_t)y_{it}$. Also note that $y_{it} = \boldsymbol{\lambda}_i'\mathbf{f}_t + u_{it}$. We have $\mathbf{b}_i = \sum_{k=1}^K \mathbf{v}_k\lambda_{ik}$. Hence

$$\begin{aligned} \max_{i \leq N} \sup_{\mathbf{x}} |E(y_{it}|\mathbf{x}_t = \mathbf{x}) - \mathbf{b}_i'\Phi(\mathbf{x})| &\leq \max_{i \leq N} \sup_{\mathbf{x}} |\sum_{k=1}^K \lambda_{ik}(E(f_{tk}|\mathbf{x}_t = \mathbf{x}) - \mathbf{v}_k'\Phi(\mathbf{x}))| \\ &\leq O(1) \max_k \sup_{\mathbf{x}} |E(f_{tk}|\mathbf{x}_t = \mathbf{x}) - \mathbf{v}_k'\Phi(\mathbf{x})| \\ &= O(J^{-\eta}). \end{aligned}$$

Q.E.D.

We now give the uniform convergence rate of $\widehat{\mathbf{b}}_i$ as well as its Bahadur representation.

Define

$$Q_i(\mathbf{b}) = \frac{1}{T}\sum_{t=1}^T \alpha_T^2 \rho\left(\frac{y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b}}{\alpha_T}\right).$$

Proposition C.2. *When $\alpha_T \leq C\sqrt{T/\log(NJ)}$ for any $C > 0$, and any $4 < k < \zeta_2 + 2$,*

$$\max_{i\leq N} \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\| = O_P(\sqrt{\frac{J\log N}{T}} + \alpha_T^{-(k-1)}).$$

Proof. Let $m_T = \sqrt{\frac{J\log N}{T}}$. We aim to show, for any $\epsilon > 0$, there is $\delta > 0$, when for all large N, T,

$$P(\min_{i\leq N} \inf_{\|\boldsymbol{\nu}\|=\delta} Q_i(\mathbf{b}_{i,\alpha} + m_T\boldsymbol{\nu}) - Q_i(\mathbf{b}_{i,\alpha}) > 0) > 1 - \epsilon.$$

This then implies $\max_i \|\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha}\| = O_P(m_T)$. The result then follows from Proposition C.1.

By the definition of $\mathbf{b}_{i,\alpha}$,

$$E[\Phi(\mathbf{x}_t)\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})] = 0, \qquad e_{it,\alpha} := y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b}_{i,\alpha}.$$

In addition, we have $e_{it} = e_{it,\alpha} + \Delta_{it,\alpha}$, where $\Delta_{it,\alpha} := (\mathbf{b}_{i,\alpha} - \mathbf{b}_i)'\Phi(\mathbf{x}_t) - z_{it}$. Using the formula: $\rho(a+t) - \rho(a) = \dot{\rho}(a)t + \int_0^t (\dot{\rho}(a+x) - \dot{\rho}(a))dx$ for $a = \alpha_T^{-1}e_{it,\alpha}$ and $t = -m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}$,

$$\begin{aligned}
&Q_i(\mathbf{b}_{i,\alpha} + m_T\boldsymbol{\nu}) - Q_i(\mathbf{b}_{i,\alpha}) = -\frac{1}{T}\sum_{t=1}^T m_T\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\Phi(\mathbf{x}_t)'\boldsymbol{\nu} \\
&+\frac{1}{T}\sum_{t=1}^T 1\{\Phi(\mathbf{x}_t)'\boldsymbol{\nu} < 0\}\alpha_T^2 \int_0^{-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}} \dot{\rho}(\alpha_T^{-1}e_{it,\alpha} + x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha})dx \\
&-\frac{1}{T}\sum_{t=1}^T 1\{\Phi(\mathbf{x}_t)'\boldsymbol{\nu} > 0\}\alpha_T^2 \int_{-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}}^0 \dot{\rho}(\alpha_T^{-1}e_{it,\alpha} + x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha})dx.
\end{aligned}$$

By the definition of $\dot{\rho}$, the integrant can be rewritten as:

$$\begin{aligned}
& \dot{\rho}(\alpha_T^{-1}e_{it,\alpha} + x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha}) \\
= & 2x1\{|\alpha_T^{-1}e_{it,\alpha} + x| < 1, |\alpha_T^{-1}e_{it,\alpha}| < 1\} \\
& +(\dot{\rho}(\alpha_T^{-1}e_{it,\alpha} + x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha}))1\{|\alpha_T^{-1}e_{it,\alpha} + x| \geq 1, \text{ or } |\alpha_T^{-1}e_{it,\alpha}| \geq 1\} \\
= & 2x - (\dot{\rho}(\alpha_T^{-1}e_{it,\alpha} + x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha}) - 2x)1\{|\alpha_T^{-1}e_{it,\alpha} + x| \geq 1, \text{ or } |\alpha_T^{-1}e_{it,\alpha}| \geq 1\}.
\end{aligned}$$

In addition, note that

$$|\dot{\rho}(x_1) - \dot{\rho}(x_1)| \leq 2|x_1 - x_2|, \quad \forall x_1, x_2.$$

Thus we can further write:

$$Q_i(\mathbf{b}_{i,\alpha} + m_T\boldsymbol{\nu}) - Q_i(\mathbf{b}_{i,\alpha})$$

$$
\begin{aligned}
&= -\frac{1}{T}\sum_{t=1}^{T} m_T\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\Phi(\mathbf{x}_t)'\boldsymbol{\nu} \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}\alpha_T^2\int_0^{-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}} 2x dx \\
&\quad -\frac{1}{T}\sum_{t=1}^{T} 1\{\Phi(\mathbf{x}_t)'\boldsymbol{\nu}<0\}\alpha_T^2\int_0^{-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}} a(x)b(x)dx \\
&\quad +\frac{1}{T}\sum_{t=1}^{T} 1\{\Phi(\mathbf{x}_t)'\boldsymbol{\nu}>0\}\alpha_T^2\int_{-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu}}^0 a(x)b(x)dx \\
&\geq \inf_{\|\boldsymbol{\nu}\|=\delta}\frac{1}{T}\sum_{t=1}^{T}\alpha_T^2(-m_T\alpha_T^{-1}\Phi(\mathbf{x}_t)'\boldsymbol{\nu})^2 \\
&\quad -\max_i\sup_{\|\boldsymbol{\nu}\|=\delta}|\frac{1}{T}\sum_{t=1}^{T} m_T\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\Phi(\mathbf{x}_t)'\boldsymbol{\nu}| \\
&\quad -\max_i\sup_{\|\boldsymbol{\nu}\|=1}\frac{1}{T}\sum_{t=1}^{T}\alpha_T^2\int_0^{m_T\alpha_T^{-1}|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|} 4xb(x)dx \\
&:= A_1 - A_2 - A_3.
\end{aligned}
$$

In the above,

$$a(x) = \dot{\rho}(\alpha_T^{-1}e_{it,\alpha}+x) - \dot{\rho}(\alpha_T^{-1}e_{it,\alpha}) - 2x$$

and

$$b(x) = 1\{|\alpha_T^{-1}e_{it,\alpha}+x|\geq 1, \text{ or } |\alpha_T^{-1}e_{it,\alpha}|\geq 1\}.$$

We now lower bound A_1 and upper bound A_2, A_3.

First of all, there is $c>0$ independent of δ, with probability approaching one,

$$
\begin{aligned}
A_1 &= \inf_{\|\boldsymbol{\nu}\|=\delta}\boldsymbol{\nu}'\frac{1}{T}\sum_{t=1}^{T} m_T^2\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\boldsymbol{\nu} \\
&\geq \lambda_{\min}(\frac{1}{T}\sum_{t=1}^{T}\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')m_T^2\delta^2 \\
&\geq cm_T^2\delta^2.
\end{aligned}
$$

As for A_2, note that $|\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})|\leq|e_{it,\alpha}|\leq|e_{it}|+|\Delta_{it,\alpha}|$. Uniformly in $i\leq N, j\leq J$, by Holder's inequality, with an arbitrarily small $v>0$, and $p=(1+v)^{-1}$,

$$E(\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t))^2 \leq \alpha_T^{-2}E(\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t))^2$$

$$
\begin{aligned}
&\leq 2\alpha_T^{-2}E(e_{it}^2+\Delta_{it,\alpha}^2)\phi_j(\mathbf{x}_t)^2\\
&\leq 2\alpha_T^{-2}EE\{e_{it}^2|\mathbf{x}_t\}\phi_j(\mathbf{x}_t)^2+2\alpha_T^{-2}E\Delta_{it,\alpha}^2\phi_j(\mathbf{x}_t)^2\\
&\leq C\alpha_T^{-2}((E\{e_{it}^2|\mathbf{x}_t\}^{1+v})^{1/p}+C)\leq C\alpha_T^{-2}.
\end{aligned}
$$

Note that $|\dot\rho|<2$ and $\{\phi_j(\mathbf{x}_t)\}$ is sub-Gaussian, thus by the Bernstein inequality, for $x=2\log(NJ)$,

$$
P(|\frac{1}{T}\sum_{t=1}^T\dot\rho(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t)|>\sqrt{\frac{2E(\dot\rho(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t))^2x}{T}}+\frac{Cx}{T})\leq 2\exp(-x).
$$

Note that when $\alpha_T\leq C\sqrt{T/\log(NJ)}$,

$$
\sqrt{\frac{2E(\dot\rho(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t))^2x}{T}}+\frac{Cx}{T}\leq\sqrt{\frac{C\log(NJ)}{\alpha_T^2T}}+\frac{C\log(NJ)}{T}\leq 2\sqrt{\frac{C\log(NJ)}{\alpha_T^2T}}.
$$

Thus

$$
P(\max_{ij}|\frac{1}{T}\sum_{t=1}^T\dot\rho(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t)|>\sqrt{\frac{C\log(NJ)}{\alpha_T^2T}})\leq CNJ\exp(-2\log(NJ))=\frac{C}{NJ}.
$$

Therefore, with probability approaching one,

$$
\begin{aligned}
A_2 &\leq m_T\alpha_T\delta\max_i\|\frac{1}{T}\sum_{t=1}^T\dot\rho(\alpha_T^{-1}e_{it,\alpha})\Phi(\mathbf{x}_t)'\|\\
&\leq m_T\alpha_T\sqrt{J}\delta\max_{i\leq N,j\leq J}|\frac{1}{T}\sum_{t=1}^T\dot\rho(\alpha_T^{-1}e_{it,\alpha})\phi_j(\mathbf{x}_t)|\\
&\leq \delta m_T\sqrt{\frac{CJ\log(N)}{T}}.
\end{aligned}
$$

As for A_3, note that uniformly for $x\leq m_T\alpha_T^{-1}|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|$, and $e_{it}=e_{it,\alpha}+\Delta_{it,\alpha}$,

$$
\begin{aligned}
&1\{|\alpha_T^{-1}e_{it,\alpha}+x|\geq 1,\ \text{or}\ |\alpha_T^{-1}e_{it,\alpha}|\geq 1\}\\
\leq\ &1\{|\alpha_T^{-1}e_{it,\alpha}+x|\geq 1\}+1\{|\alpha_T^{-1}e_{it,\alpha}|\geq 1\}\\
\leq\ &2\times 1\{|e_{it,\alpha}|>3\alpha_T/4\}+1\{m_T|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|>\alpha_T/4\}\\
\leq\ &2\times 1\{|e_{it}|>\alpha_T/2\}+1\{m_T|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|>\alpha_T/4\}+1\{|\Delta_{it,\alpha}|>\alpha_T/4\}.
\end{aligned}
$$

In addition, with probability at least $1-\epsilon/10$,

$$
\max_i\frac{1}{T}\sum_{t=1}^T 1\{|e_{it}|>\alpha_T/2\}\quad\leq_{(1)}\quad\max_i P(|e_{it}|>\alpha_T/2)
$$

$$
\begin{aligned}
& & & +\sqrt{\frac{\log N}{T}}\max_i P(|e_{it}|>\alpha_T/2)^{1/2}, \\
\frac{1}{T}\sum_t 1\{m_T\|\Phi(\mathbf{x}_t)\|\delta>\alpha_T/4\} & \leq & & 10P(m_T\delta\|\Phi(\mathbf{x}_t)\|>\alpha_T/4)/\epsilon, \\
\max_i \frac{1}{T}\sum_{t=1}^T 1\{|\Delta_{it,\alpha}|>\alpha_T/4\} & \leq & & \max_i \frac{1}{T}\sum_{t=1}^T 1\{\|\Phi(\mathbf{x}_t)\|>C\alpha_T^k\}+1\{|z_{it}|>\alpha_T/4\} \\
& \leq & & 10P(\|\Phi(\mathbf{x}_t)\|>C\alpha_T^k)/\epsilon \\
& & & +CJ^{-2\eta}/\alpha_T^2+\sqrt{\frac{\log N}{T}}CJ^{-\eta}/\alpha_T,
\end{aligned}
$$

where (1) follows from the triangular inequality,

$$
\begin{aligned}
\max_i \frac{1}{T}\sum_{t=1}^T 1\{|e_{it}|>\alpha_T/2\} & \leq \max_i P(|e_{it}|>\alpha_T/2) \\
& \quad +\max_i \left|\frac{1}{T}\sum_{t=1}^T 1\{|e_{it}|>\alpha_T/2\}-P(|e_{it}|>\alpha_T/2)\right|,
\end{aligned}
$$

and we used Bernstein inequality+union bound to bound the second term since the indicator function is bounded. Hence for an arbitrarily small $v>0$, by Holder's inequality, for some generic constant $C>0$, independent of δ,

$$
\begin{aligned}
A_3 & \leq \max_i \sup_{\|\boldsymbol{\nu}\|=\delta} \frac{1}{T}\sum_{t=1}^T 4(m_T|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|)^2[1\{|e_{it}|>\alpha_T/2\} \\
& \quad +1\{m_T|\Phi(\mathbf{x}_t)'\boldsymbol{\nu}|>\alpha_T/4\}+1\{|\Delta_{it,\alpha}|>\alpha_T/4\}] \\
& \leq C\max_i(\frac{1}{T}\sum_{t=1}^T[1\{|e_{it}|>\alpha_T/2\}+1\{m_T\delta\|\Phi(\mathbf{x}_t)\|>\alpha_T/4\}+1\{|\Delta_{it,\alpha}|>\alpha_T/4\}])^{1-v} \\
& \quad \times(\frac{1}{T}\sum_{t=1}^T\|\Phi(\mathbf{x}_t)\|^{2/v})^v(m_T\delta)^2 \\
& \leq (m_T\delta)^2C\Bigg(\max_i P(|e_{it}|>\alpha_T/2)+\sqrt{\frac{\log N}{T}}\max_i P(|e_{it}|>\alpha_T/2)^{1/2} \\
& \quad +10P(m_T\delta\|\Phi(\mathbf{x}_t)\|>\alpha_T/4)/\epsilon \\
& \quad +10P(\|\Phi(\mathbf{x}_t)\|>C\alpha_T^k)/\epsilon+CJ^{-2\eta}/\alpha_T^2+\sqrt{\frac{\log N}{T}}CJ^{-\eta}/\alpha_T\Bigg)^{1-v}(C+E\|\Phi(\mathbf{x}_t)\|^{2/v})^v.
\end{aligned}
$$

We now upper bound $E\|\Phi(\mathbf{x}_t)\|^{2/v}$ and $P(\|\Phi(\mathbf{x}_t)\|>x)$ for any x. Since $\{\phi_j(w_t)\}_{j\leq J}$ is sub-Gaussian, by Lemma 14.12 of Bühlmann and van de Geer (2011),

$$
E\|\Phi(\mathbf{x}_t)\|^{2/v} \leq J^{1/v}E(\max_{j\leq J}\phi_j(\mathbf{x}_t)^{2/v})
$$

$$
\begin{aligned}
&\leq J^{1/v}E(\max_{j\leq J}|\phi_j(\mathbf{x}_t)^{2/v}-E\phi_j(\mathbf{x}_t)^{2/v}|)+J^{1/v}\max_j E\phi_j(\mathbf{x}_t)^{2/v}\\
&\leq J^{1/v}C\log(J).\\
P(\|\Phi(\mathbf{x}_t)\|>x) &\leq P(\max_j|\phi_j(\mathbf{x}_t)|^2J>x^2)\leq J\max_j P(|\phi_j(\mathbf{x}_t)|>x/J^{1/2})\\
&\leq J\exp(-Cx^2/J).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
A_3 &\leq (m_T\delta)^2C\Bigg(\max_i P(|e_{it}|>\alpha_T/2)+\sqrt{\frac{\log N}{T}}\max_i P(|e_{it}|>\alpha_T/2)^{1/2}\\
&\quad +CJ\exp(-C\alpha_T^2/(Jm_T^2\delta^2))/\epsilon\\
&\quad +CJ\exp(-C\alpha_T^{2k}/J)/\epsilon+CJ^{-2\eta}/\alpha_T^2+\sqrt{\frac{\log N}{T}}CJ^{-\eta}/\alpha_T\Bigg)^{1-v}J(\log J)^v\\
&:= (m_T\delta)^2Cl_T.
\end{aligned}
$$

Note that $l_T=o(1)$.

Consequently, for any $\epsilon>0$, there are C,c, and c_ϵ independent of δ (may depend on ϵ), with probability at least $1-\epsilon$, uniformly in $i\leq N$ and $\|\boldsymbol{\nu}\|=\delta$, for $m_T=\sqrt{\frac{J\log N}{T}}$,

$$
\begin{aligned}
Q_i(\mathbf{b}_{i,\alpha}+m_T\boldsymbol{\nu})-Q_i(\mathbf{b}_{i,\alpha}) &\geq m_T^2\delta^2(c-c_\epsilon l_T)-\delta m_TC\sqrt{\frac{J\log N}{T}}\\
&\geq m_T\delta(m_T\delta c/2-Cm_T)>0
\end{aligned}
$$

so long as $\delta c>2C$. Thus $\max_i\|\widehat{\mathbf{b}}_i-\mathbf{b}_{i,\alpha}\|=O_P(m_T)$.

We now prove a simple lemma.

Lemma C.2. *There is $M>0$ for all $x>M$,*

$$
\begin{aligned}
\max_{i\leq N}\sup_{\mathbf{x}}P(|e_{it}|>x|\mathbf{x}_t=\mathbf{x}) &\leq Cx^{-\zeta_2-2}\\
\max_{i\leq N}\sup_{\mathbf{x}}E(|e_{it}|1\{|e_{it}|>x\}|\mathbf{x}_t=\mathbf{x}) &\leq Cx^{-\zeta_2-1}.
\end{aligned}
$$

Proof. Uniformly in $\mathbf{x}=\mathbf{x}_t$ and $i\leq N$,

$$
\begin{aligned}
P(|e_{it}|>x|\mathbf{x}_t) &= E(1\{|e_{it}|>x\}|\mathbf{x}_t)\\
&\leq E(e_{it}^21\{|e_{it}|>x\}|\mathbf{x}_t)x^{-2}\leq Cx^{-\zeta_2-2}\\
E(|e_{it}|1\{|e_{it}|>x\}|\mathbf{x}_t) &\leq E(e_{it}^21\{|e_{it}|>x\}|\mathbf{x}_t)x^{-1}\leq Cx^{-\zeta_2-1}.
\end{aligned}
$$

Lemma C.3. *Uniformly for $i = 1, ..., N$,*

$$\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha} = (2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}\frac{1}{T}\sum_{t=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{it,\alpha})\Phi(\mathbf{x}_t) + \mathbf{R}_{i,b},$$

where $\max_{i\leq N}\|\mathbf{R}_{i,b}\| = O_P(\alpha_T^{-(\zeta_1-1)} + \sqrt{\frac{\log J}{T}})J\sqrt{\frac{J\log N}{T}}$.

Proof. Note that $\nabla Q_i(\mathbf{b}) = -\frac{1}{T}\sum_{t=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}(y_{it}-\Phi(\mathbf{x}_t)'\mathbf{b}))\Phi(\mathbf{x}_t)$. Define $\bar{Q}_i(\mathbf{b}) = EQ_i(\mathbf{b})$,

$$\begin{aligned}\boldsymbol{\mu}_i(\mathbf{b}) &:= \nabla Q_i(\mathbf{b}) - \nabla\bar{Q}_i(\mathbf{b}) \\ &= E\alpha_T\dot{\rho}(\alpha_T^{-1}(y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b}))\Phi(\mathbf{x}_t) - \frac{1}{T}\sum_{t=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}(y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b}))\Phi(\mathbf{x}_t).\end{aligned}$$

The first order condition gives $\nabla Q_i(\widehat{\mathbf{b}}_i) = 0$. By the mean value expansion,

$$\begin{aligned} 0 &= \nabla Q_i(\widehat{\mathbf{b}}_i) - \nabla\bar{Q}_i(\widehat{\mathbf{b}}_i) + \nabla\bar{Q}_i(\widehat{\mathbf{b}}_i) - \nabla\bar{Q}_i(\mathbf{b}_{i,\alpha}) + \nabla\bar{Q}_i(\mathbf{b}_{i,\alpha}) - \nabla Q_i(\mathbf{b}_{i,\alpha}) + \nabla Q_i(\mathbf{b}_{i,\alpha}) \\ &= \boldsymbol{\mu}_i(\widehat{\mathbf{b}}_i) + \nabla\bar{Q}_i(\widehat{\mathbf{b}}_i) - \nabla\bar{Q}_i(\mathbf{b}_{i,\alpha}) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha}) + \nabla Q_i(\mathbf{b}_{i,\alpha}) \\ &= \nabla^2\bar{Q}_i(\widetilde{\mathbf{b}}_i)(\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha}) + \nabla Q_i(\mathbf{b}_{i,\alpha}) + \boldsymbol{\mu}_i(\widehat{\mathbf{b}}_i) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha}).\end{aligned}$$

for some $\widetilde{\mathbf{b}}_i$ in the segment joining $\widehat{\mathbf{b}}_i$ and $\mathbf{b}_{i,\alpha}$. We now proceed by: (i) upper bounding $\max_i\|\boldsymbol{\mu}_i(\widehat{\mathbf{b}}_i) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha})\|$, and (ii) finding the limit of $\nabla^2\bar{Q}_i(\widetilde{\mathbf{b}}_i)$ uniformly in i.

(i) Note that in the proof of Proposition C.2, we have proved that for any $\epsilon > 0$, there is $\delta > 0$, so that the following event holds with probability at least $1 - \epsilon$:

$$\max_i\|\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha}\| \leq \delta m_T, \quad m_T = \sqrt{\frac{J\log N}{T}}.$$

We bound $E\max_i\sup_{\|\mathbf{b}-\mathbf{b}_{i,\alpha}\|\leq\delta m_T}\|\boldsymbol{\mu}_i(\mathbf{b}) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha})\|$. Let $\mu_{ij}(\cdot)$ be the jth element of $\boldsymbol{\mu}_i$, $j \leq J$. Since $\{\mathbf{y}_t, \mathbf{x}_t\}_{t\leq T}$ are serially independent, there exists a Radamacher sequence $\{\varepsilon_t\}_{t\leq T}$ with $P(\varepsilon_t = 1) = P(\varepsilon_t = -1) = 1/2$, that is independent of $\{\mathbf{y}_t, \mathbf{x}_t\}$,

$$\begin{aligned} & E\max_{i\leq N, j\leq J}\sup_{\|\mathbf{b}-\mathbf{b}_{i,\alpha}\|\leq\delta m_T}|\mu_{ij}(\mathbf{b}) - \mu_{ij}(\mathbf{b}_{i,\alpha})| \\ \leq_{(a)} \;& 2E\max_{i\leq N, j\leq J}\sup_{\|\mathbf{b}-\mathbf{b}_{i,\alpha}\|\leq\delta m_T}|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\alpha_T\big(\dot{\rho}(\alpha_T^{-1}(y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b})) \\ & -\dot{\rho}(\alpha_T^{-1}(y_{it} - \Phi(\mathbf{x}_t)'\mathbf{b}_{i,\alpha}))\big)\phi_j(\mathbf{x}_t)|\end{aligned}$$

$$
\begin{aligned}
&\leq_{(b)} && 4E \max_{i\leq N, j\leq J} \sup_{\|\mathbf{b}-\mathbf{b}_{i,\alpha}\|\leq \delta m_T} |\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \Phi(\mathbf{x}_t)'(\mathbf{b}_{i,\alpha}-\mathbf{b})\phi_j(\mathbf{x}_t)| \\
&\leq && 4\delta m_T E \max_{j\leq J} \|\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \phi_j(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\| \leq 4\delta m_T \sqrt{J} E \max_{l,j\leq J} |\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \phi_j(\mathbf{x}_t)\phi_l(\mathbf{x}_t)| \\
&\leq_{(c)} && 4\delta m_T \sqrt{J}\frac{L}{T} \log E \exp\left(L^{-1}\max_{l,j\leq J} |\sum_{t=1}^{T} \varepsilon_t \phi_j(\mathbf{x}_t)\phi_l(\mathbf{x}_t)|\right) \\
&\leq_{(d)} && 4\delta m_T \sqrt{J}\frac{L}{T} \log \sum_{l,j\leq J} E \exp\left(L^{-1} |\sum_{t=1}^{T} \varepsilon_t \phi_j(\mathbf{x}_t)\phi_l(\mathbf{x}_t)|\right) \\
&\leq_{(e)} && 4\delta m_T \sqrt{J}\frac{L}{T} \log \sum_{l,j\leq J} \exp\left(\frac{T}{2(L^2 - LK_0)}\right) \\
&= && 4\delta m_T \sqrt{J}\frac{L}{T} \left(2\log J + \frac{T}{2(L^2 - LK_0)}\right) \\
&= && 4\delta m_T \sqrt{J}\left(\frac{2L\log J}{T} + \sqrt{\frac{c_0 \log J}{4T}}\right) \leq C\delta m_T \sqrt{\frac{J\log J}{T}}.
\end{aligned}
$$

Note that $|\dot{\rho}(\cdot)| \leq 2$ and $\{\phi_j(\cdot)\}$ is sub-Gaussian, hence (a) follows from the symmetrization theorem (see, e.g., Theorem 14.3 of Bühlmann and van de Geer (2011)); since $\dot{\rho}(\cdot)$ is Lipschitz continuous, (b) follows from the contraction theorem (e.g., Theorem 14.4 of Bühlmann and van de Geer (2011)). Let K_0 denote constant parameter of the sub-Gaussianity of $\{\phi_l(\mathbf{x}_t)\phi_j(\mathbf{x}_t)\}_{l,j\leq J}$; for some $c_0 > 0$, let

$$
L = K_0 + \sqrt{\frac{T}{c_0 \log J}}.
$$

Then (c) follows from the Jensen's inequality; (d) follows from the simple inequality that $\exp(\max) \leq \sum \exp$; (e) follows from an inequality of exponential moment of an average for sub-Gaussian random variables (Lemma 14.8 of Bühlmann and van de Geer (2011)).

Therefore,

$$
E \max_i \sup_{\|\mathbf{b}-\mathbf{b}_{i,\alpha}\|\leq \delta m_T} \|\boldsymbol{\mu}_i(\mathbf{b}) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha})\| \leq CJm_T\sqrt{\frac{\log J}{T}} = \frac{CJ^{3/2}(\log N \log J)^{1/2}}{T}.
$$

Hence

$$
\max_i \|\boldsymbol{\mu}_i(\widehat{\mathbf{b}}_i) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha})\| = O_P(J^{3/2}(\log N \log J)^{1/2}/T).
$$

(ii) Note that

$$\nabla\bar{Q}_i(\mathbf{b}) = -E\Phi(\mathbf{x}_t)\alpha_T\dot{\rho}(\alpha_T^{-1}(e_{it}+z_{it})+\alpha_T^{-1}\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b})) = -E\Phi(\mathbf{x}_t)A_{it}(\mathbf{b})$$

where $A_{it}(\mathbf{b}) = E[\alpha_T\dot{\rho}(\alpha_T^{-1}(e_{it}+z_{it})+\alpha_T^{-1}\Phi(\mathbf{x}_t)'(\mathbf{b}_i-\mathbf{b}))|\mathbf{x}_t]$. Let $g_{e,i}$ denote the density of e_{it}, and let P_e denote the conditional probability measure conditioning on $\mathbf{x}_t$. Then careful calculations yield: $\nabla A_{it}(\mathbf{b}) = -2\Phi(\mathbf{x}_t)' + \sum_{j=1}^{8} B_{it,j}(\mathbf{b})\Phi(\mathbf{x}_t)'$, where

$$\begin{aligned}
B_{it,1}(\mathbf{b}) &= -2\alpha_T g_{e,i}(\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,2}(\mathbf{b}) &= -2\alpha_T g_{e,i}(-\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,3}(\mathbf{b}) &= -2P_e((\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)+z_{it}+e_{it}>\alpha_T),\\
B_{it,4}(\mathbf{b}) &= 2((\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)+z_{it})g_{e,i}(\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,5}(\mathbf{b}) &= 2P_e(e_{it}<-\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,6}(\mathbf{b}) &= -2((\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)+z_{it})g_{e,i}(-\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,7}(\mathbf{b}) &= 2[\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}]g_{e,i}(\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}),\\
B_{it,8}(\mathbf{b}) &= -2(-\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it})g_{e,i}(-\alpha_T-(\mathbf{b}_i-\mathbf{b})'\Phi(\mathbf{x}_t)-z_{it}).
\end{aligned}$$

Since $\max_i\|\widehat{\mathbf{b}}_i-\mathbf{b}_i\| = o_P(m_T)$, $\max_{it}|z_{it}| = o_P(\alpha_T)$, $\Phi(\mathbf{x}_t)$ is sub-Gaussian and $J\log N\sqrt{\log T} = o(T)$, we have: with probability approaching one, for any $\epsilon>0$,

$$\max_{i,t}|(\mathbf{b}_i-\widetilde{\mathbf{b}}_i)'\Phi(\mathbf{x}_t)| + \max_{it}|z_{it}| < \epsilon\alpha_T.$$

Hence with probability approaching one,

$$\begin{aligned}
\max_i|\sum_{j\neq 3,5} B_{it,j}(\widetilde{\mathbf{b}}_i)| &\leq C\alpha_T\max_i\sup_{|x|<\epsilon\alpha_T} g_{e,i}(\pm\alpha_T+x) \leq C\alpha_T^{-(\zeta_1-1)},\\
\max_i|B_{it,3}(\widetilde{\mathbf{b}}_i)+B_{it,5}(\widetilde{\mathbf{b}}_i)| &\leq C\max_i P(|e_{it}|>(1-\epsilon)\alpha_T) \leq C\alpha_T^{-(\zeta_2+2)}.
\end{aligned}$$

Hence

$$\|\nabla^2\bar{Q}_i(\widetilde{\mathbf{b}}_i)-2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\| = \|\sum_{j=1}^{8}E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'B_{it,j}(\widetilde{\mathbf{b}}_i)\| = O(J\alpha_T^{-(\zeta_1-1)}+J\alpha_T^{-(\zeta_2+2)}).$$

Consequently, $\widehat{\mathbf{b}}_i-\mathbf{b}_{i,\alpha} = -(2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}\nabla Q_i(\mathbf{b}_{i,\alpha})+\mathbf{R}_{i,b}$, where

$$\max_{i\leq N}\|\mathbf{R}_{i,b}\| \leq \|(2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}\|(\|\nabla^2\bar{Q}_i(\widetilde{\mathbf{b}}_i)-2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|\|\widehat{\mathbf{b}}_i-\mathbf{b}_{i,\alpha}\|$$

$$+\max_i \|\boldsymbol{\mu}_i(\widehat{\mathbf{b}}_i) - \boldsymbol{\mu}_i(\mathbf{b}_{i,\alpha})\|)$$
$$= O_P(\alpha_T^{-(\zeta_1-1)} + \alpha_T^{-(\zeta_2+2)} + \sqrt{\frac{\log J}{T}})Jm_T$$

Proposition C.3. *Let* $\widehat{E}(y_{it}|\mathbf{x}_t) = \widehat{\mathbf{b}}_i'\Phi(\mathbf{x}_t)$. *Then for* $\mathbf{A} = (2E\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')^{-1}$,

$$\widehat{E}(y_{it}|\mathbf{x}_t) = E(y_{it}|\mathbf{x}_t) + \Phi(\mathbf{x}_t)'\mathbf{A}\frac{1}{T}\sum_{s=1}^T \alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})\Phi(\mathbf{x}_s) + R_{1,it} + R_{2,it} + R_{3,it},$$

where (recall that $z_{it} = E(y_{it}|\mathbf{x}_t) - \mathbf{b}_i'\Phi(\mathbf{x}_t)$*)*

$$\begin{aligned} R_{1,it} &:= \Phi(\mathbf{x}_t)'\mathbf{A}\frac{1}{T}\sum_{s=1}^T \alpha_T[\dot{\rho}(\alpha_T^{-1}e_{is,\alpha}) - \dot{\rho}(\alpha_T^{-1}e_{is})]\Phi(\mathbf{x}_s) \\ R_{2,it} &:= \Phi(\mathbf{x}_t)'(\mathbf{R}_{i,b} + \mathbf{b}_{i,\alpha} - \mathbf{b}_i), \qquad R_{3,it} := -z_{it}. \end{aligned}$$

Write $R_{it} := R_{1,it} + R_{2,it} + R_{3,it}$, *then*

$$\begin{aligned} &\max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2 = O_P(J^{1-2\eta} + \alpha_T^{-2(\zeta_1-1)}\frac{J^3\log N}{T} + \frac{J^3\log N\log J}{T^2}), \\ &\max_i \frac{1}{T}\sum_{t=1}^T |\widehat{E}(y_{it}|\mathbf{x}_t) - E(y_{it}|\mathbf{x}_t)|^2 = O_P(\frac{J\log N}{T} + J^{-2\eta}). \end{aligned}$$

Proof. By Lemma C.3 and Proposition C.3,

$$\begin{aligned} &\widehat{E}(y_{it}|\mathbf{x}_t) = E(y_{it}|\mathbf{x}_t) + \Phi(\mathbf{x}_t)'(\widehat{\mathbf{b}}_i - \mathbf{b}_{i,\alpha}) + \Phi(\mathbf{x}_t)'(\mathbf{b}_{i,\alpha} - \mathbf{b}_i) - z_{it} \\ = &\ E(y_{it}|\mathbf{x}_t) + \Phi(\mathbf{x}_t)'\mathbf{A}\frac{1}{T}\sum_{s=1}^T \alpha_T\dot{\rho}(\alpha_T^{-1}e_{is,\alpha})\Phi(\mathbf{x}_s) + \Phi(\mathbf{x}_t)'(\mathbf{R}_{i,b} + \mathbf{b}_{i,\alpha} - \mathbf{b}_i) - z_{it} \\ = &\ E(y_{it}|\mathbf{x}_t) + \Phi(\mathbf{x}_t)'\mathbf{A}\frac{1}{T}\sum_{s=1}^T \alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})\Phi(\mathbf{x}_s) + R_{it}. \end{aligned}$$

On the other hand, uniformly in i, for $a = \lambda_{\max}(\frac{1}{T}\sum_{t=1}^T \Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)')$,

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^T R_{it}^2 &\leq aC\|\mathbf{A}\|^2\|\frac{1}{T}\sum_{s=1}^T \alpha_T[\dot{\rho}(\alpha_T^{-1}e_{is,\alpha}) - \dot{\rho}(\alpha_T^{-1}e_{is})]\Phi(\mathbf{x}_s)\|^2 \\ &\quad + aC\|\mathbf{R}_{i,b} + \mathbf{b}_{i,\alpha} - \mathbf{b}_i\|^2 + C\frac{1}{T}\sum_t z_{it}^2 \\ &\leq C(\frac{1}{T}\sum_s |e_{is,\alpha} - e_{is}|\|\Phi(\mathbf{x}_s)\|)^2 + C\|\mathbf{R}_{i,b}\|^2 + C\|\mathbf{b}_{i,\alpha} - \mathbf{b}\|^2 + C\frac{1}{T}\sum_t z_{it}^2 \end{aligned}$$

$$
\begin{aligned}
&\leq C\frac{1}{T}\sum_s(|z_{it}|^2+\|\mathbf{b}_{i,\alpha}-\mathbf{b}_i\|^2\|\Phi(\mathbf{x}_t)\|^2)\frac{1}{T}\sum_t\|\Phi(\mathbf{x}_t)\|^2+C\|\mathbf{R}_{i,b}\|^2\\
&\quad +C\|\mathbf{b}_{i,\alpha}-\mathbf{b}\|^2+O_P(J^{-2\eta})\\
&= O_P(J)(J^{-2\eta}+J\alpha_T^{-2(k-1)})+O_P(\alpha_T^{-2(\zeta_1-1)}+\alpha_T^{-2(\zeta_2+2)}+\frac{\log J}{T})J^2m_T^2.
\end{aligned}
$$

Also note that $\alpha_T^{-2\zeta_2-4}=O(\log N/T)$. Finally,

$$
\begin{aligned}
&\max_i\frac{1}{T}\sum_{t=1}^T|\widehat{E}(y_{it}|\mathbf{x}_t)-E(y_{it}|\mathbf{x}_t)|^2\\
\leq\ &\max_i\frac{1}{T}\sum_{t=1}^T|\Phi(\mathbf{x}_t)'(\widehat{\mathbf{b}}_i-\mathbf{b}_i)|^2+\max_i\frac{1}{T}\sum_{t=1}^T z_{it}^2\\
\leq\ & a\|\widehat{\mathbf{b}}_i-\mathbf{b}_i\|^2+\max_i\frac{1}{T}\sum_{t=1}^T z_{it}^2\\
=\ & O_P(\frac{J\log N}{T}+\alpha_T^{-2(k-1)}+J^{-2\eta}).
\end{aligned}
$$

The term involving $\alpha_T^{-2(k-1)}$ is negligible since it is smaller than $(\log N/T)^3$.

C.2 Technical lemmas for the loadings

We shall first examine the behavior of $\widetilde{\mathbf{V}}^{-1}$ and $\mathbf{H}$. This is given by the lemma below.

Define

$$\delta_N^2(\mathbf{x})=\|\frac{1}{T}\sum_t E(\mathbf{f}_t|\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|^2+\frac{1}{T}\sum_t\|E(\mathbf{f}_t|\mathbf{x}_t)\|^2.$$

Lemma C.4. *Recall that $\mathbf{V}$ is a $K\times K$ diagonal matrix, whose diagonal elements are the eigenvalues of $\mathbf{\Sigma}_{\Lambda,N}^{1/2}E\{E(\mathbf{f}_t|\mathbf{x}_t)E(\mathbf{f}_t|\mathbf{x}_t)'\}\mathbf{\Sigma}_{\Lambda,N}^{1/2}$. Suppose $J/T+J^{-\eta}+\sqrt{\log N/T}\ll\chi_N$. Then (i)*

$$\|\widetilde{\mathbf{V}}-\mathbf{V}\|=O_P(J^{-\eta}+\sqrt{\frac{\log N}{T}}).$$

(ii) $\|\widetilde{\mathbf{V}}^{-1}\|=O_P(\chi_N^{-1})$. *(iii)* $\delta_N^2(\mathbf{x})=O_P(\chi_N)$. *(iv)* $\|\mathbf{H}\|=O_P(1)$.

Proof. Recall that $\mathbf{\Sigma}_{y|x}=\mathbf{\Lambda}\mathbf{\Sigma}_{f|x}\mathbf{\Lambda}'$. Let $\mathbf{V}$ be a $K\times K$ diagonal matrix, whose diagonal elements are the first K eigenvalues of $\mathbf{\Sigma}_{y|x}/N$, which are also the eigenvalues of $\mathbf{\Sigma}_{f|x}^{1/2}\mathbf{\Sigma}_{\Lambda,N}\mathbf{\Sigma}_{f|x}^{1/2}$. By Assumption 2.1,

$$\lambda_{\min}(\mathbf{V})=\lambda_{\min}(\mathbf{\Sigma}_{f|x}^{1/2}\mathbf{\Sigma}_{\Lambda,N}\mathbf{\Sigma}_{f|x}^{1/2})\geq\underline{c}_\Lambda\chi_N$$

with $\underline{c}_\Lambda > 0$ being a constant. On the other hand, by Proposition C.3,

$$\begin{aligned}
\|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}_{y|x}\|_\infty &\leq \max_{ij} \frac{1}{T}\sum_t |\widehat{E}(y_{it}|\mathbf{x}_t)\widehat{E}(y_{jt}|\mathbf{x}_t) - E(y_{it}|\mathbf{x}_t)E(y_{jt}|\mathbf{x}_t)| \\
&\quad + \max_{ij} |\frac{1}{T}\sum_t E(y_{it}|\mathbf{x}_t)E(y_{jt}|\mathbf{x}_t) - E\{E(y_{it}|\mathbf{x}_t)E(y_{jt}|\mathbf{x}_t)\}| \\
&\leq \max_i \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\| O_P(\|\frac{1}{T}\sum_t \Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|) + O_P(J^{-\eta} + \sqrt{\frac{\log N}{T}}) \\
&= O_P(J^{-\eta} + \sqrt{\frac{\log N}{T}}).
\end{aligned}$$

By Weyl's theorem,

$$\|\widetilde{\mathbf{V}} - \mathbf{V}\| \leq \frac{1}{N}\|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}\| \leq \|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_\infty = O_P(J^{-\eta} + \sqrt{\frac{\log N}{T}}).$$

(ii) Because $J^{-\eta} + \sqrt{\log N/T} \ll \chi_N$, with probability approaching one,

$$\lambda_{\min}(\widetilde{\mathbf{V}}) \geq \lambda_{\min}(\mathbf{V}) - \|\widetilde{\mathbf{V}} - \mathbf{V}\| \geq \underline{c}_\Lambda \chi_N/2.$$

(iii) For notational simplicity, write $\mathbf{g}_t := E(\mathbf{f}_t|\mathbf{x}_t)$. First of all, we show $\mathbf{g}_t$ has a finite fourth moment. In fact, $\mathbf{v}_k := (E\Phi_t\Phi_t')^{-1}E\Phi_t f_{kt}$ has a bounded norm due to Assumption 4.2, thus by Assumption 4.1, $\mathbf{v}_k'\Phi(\mathbf{x}_t)$ has a bounded forth moment. Then by Assumption 4.2, there is $C > 0$,

$$E\|\mathbf{g}_t\|^4 \leq \sup_{\mathbf{x}_t}\max_k C|E(f_{Kt}|\mathbf{x}_t) - \mathbf{v}_k'\Phi(\mathbf{x}_t)|^4 + CE(\mathbf{v}_k'\Phi(\mathbf{x}_t))^4 < O(1). \tag{C.1}$$

Because $E(\mathbf{f}_t|\mathbf{x}_t)$ is independent across t, $\|\frac{1}{T}\sum_{t=1}^T \mathbf{g}_t\mathbf{g}_t' - \mathbf{\Sigma}_{f|x}\| = O_P(\frac{1}{\sqrt{T}})$, implying $\|\frac{1}{T}\sum_{t=1}^T \mathbf{g}_t\mathbf{g}_t'\| \leq O_P(T^{-1/2} + \psi_N) = O_P(\psi_N)$, where the last equality is due to $\psi_N := \lambda_{\max}(\mathbf{\Sigma}_{f|x}) \geq \chi_N \gg \sqrt{1/T}$. Now

$$E\|\mathbf{g}_t\phi_j(\mathbf{x}_t)'\|^2 \leq (E\|\mathbf{g}_t\|^4 E\phi_j(\mathbf{x}_t)^4)^{1/2}$$

So each element of $\mathbf{g}_t\phi_j(\mathbf{x}_t)$ has a bounded second moment uniformly in $j \leq J$. Thus we have $\|\frac{1}{T}\sum_t \mathbf{g}_t\Phi(\mathbf{x}_t)' - E\mathbf{g}_t\Phi(\mathbf{x}_t)'\| = O_P(\sqrt{\frac{J}{T}})$. Similarly, $|\frac{1}{T}\sum_t \|\mathbf{g}_t\|^2 - E\|\mathbf{g}_t\|^2| = O_P(T^{-1/2})$. Hence by Assumption 4.4, recall that $\chi_N := \lambda_{\min}(E\mathbf{g}_t\mathbf{g}_t')$,

$$\delta_N^2(\mathbf{x}) \leq 2\|E\mathbf{g}_t\Phi(\mathbf{x}_t)'\|^2 + 2E\|\mathbf{g}_t\|^2 + O_P(\frac{J}{T})$$

$$\leq \quad 2\|E\Phi(\mathbf{x}_t)\mathbf{g}_t'\mathbf{g}_t\Phi(\mathbf{x}_t)'\| + 2\operatorname{tr} E\mathbf{g}_t\mathbf{g}_t' + O_P(\frac{J}{T})$$
$$\leq \quad C\chi_N + O_P(\frac{J}{T}) = O_P(\chi_N),$$

where the last equality is due to the assumption $J/T \ll \chi_N$.

(iv) By the definition that the columns of $\frac{1}{\sqrt{N}}\widehat{\mathbf{\Lambda}}$ are eigenvectors. We have $\|\frac{1}{N}\mathbf{\Lambda}'\widehat{\mathbf{\Lambda}}\| \leq \|\frac{1}{\sqrt{N}}\mathbf{\Lambda}\| \leq \sqrt{\bar{c}_\Lambda} = O(1)$. So by part (iii)

$$\|\mathbf{H}\| \leq \|\frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t\mathbf{g}_t'\|\|\frac{1}{N}\mathbf{\Lambda}'\widehat{\mathbf{\Lambda}}\|\|\widetilde{\mathbf{V}}^{-1}\| \leq O_P(\chi_N\chi_N^{-1}) = O_P(1).$$

Q.E.D.

Note in Lemma C.5 below that terms $\mathbf{B}_1, \mathbf{B}_4$ and $\mathbf{B}_7$ have two upper bounds, where the second bound uses a simple inequality $\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2 \leq \|\mathbf{M}_\alpha\|^2\|\widehat{\mathbf{\Lambda}}\|^2$. Such a simple inequality is crude, but is sufficient to prove Proposition B.1. On the other hand, given Proposition B.1, a sharper rate for $\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2$ can be found. As a result, the first bounds for $\mathbf{B}_1, \mathbf{B}_4$ and $\mathbf{B}_7$ are used later to achieve sharp rates for $\widehat{\mathbf{g}}(\mathbf{x}_t) - \mathbf{g}(\mathbf{x}_t)$.

Lemma C.5. *(i)* $\|\mathbf{M}_\alpha\|^2 = O_P(NJ/T + NJ^{1-2\eta})$,

(ii) $\|\mathbf{B}_1\|_F^2 = O_P(\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2/(N\chi_N)) = O_P(\|\mathbf{M}_\alpha\|^2/\chi_N)$,
$\|\mathbf{B}_3\|_F^2 = O_P(\|\mathbf{M}_\alpha\|^2/\chi_N)$.

(iii) $\|\mathbf{B}_2\|_F^2 = O_P(N\max_i \frac{1}{T}\sum_{t=1}^{T} R_{it}^2/\chi_N) = \|\mathbf{B}_6\|_F^2$,

(iv) $\|\mathbf{B}_4\|_F^2 = O_P(\|\mathbf{M}_\alpha\|^2\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2/(N^2\chi_N^2)) = O_P(\|\mathbf{M}_\alpha\|_F^4/(N\chi_N^2))$,
$\|\mathbf{B}_8\|_F^2 = O_P(N(\max_i \frac{1}{T}\sum_{t=1}^{T} R_{it}^2)^2/\chi_N^2)$,

(v) $\|\mathbf{B}_5\|_F^2 = O_P(\|\mathbf{M}_\alpha\|^2 J\max_i \frac{1}{T}\sum_{t=1}^{T} R_{it}^2/\chi_N^2)$.

$\|\mathbf{B}_7\|_F^2 = O_P(\max_i \frac{1}{T}\sum_{t=1}^{T} R_{it}^2 J\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2/(N\chi_N^2)) = O_P(\|\mathbf{M}_\alpha\|^2 J\max_i \frac{1}{T}\sum_{t=1}^{T} R_{it}^2/\chi_N^2)$.

Proof. By Lemma C.4, $\delta_N^2(\mathbf{x}) = O_P(\chi_N)$.

(i) Recall that $e_{it} = e_{it,\alpha} + \Delta_{it,\alpha}$, where $\Delta_{it,\alpha} := (\mathbf{b}_{i,\alpha} - \mathbf{b}_i)'\Phi(\mathbf{x}_t) - z_{it}$.

$$E\|\mathbf{M}_\alpha\|_F^2 = E\sum_{i=1}^{N}\|\mathbf{M}_{i,\alpha}\|^2 = \sum_{i=1}^{N}\sum_{j=1}^{J}E(\frac{1}{T}\sum_{s=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})\phi_j(\mathbf{x}_s))^2$$
$$\leq \quad 2\sum_{i=1}^{N}\sum_{j=1}^{J}E(\frac{1}{T}\sum_{s=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is,\alpha})\phi_j(\mathbf{x}_s))^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{J}E(\frac{1}{T}\sum_{s=1}^{T}2|e_{is} - e_{is,\alpha}||\phi_j(\mathbf{x}_s)|)^2$$

$$
\begin{aligned}
&\leq\ 2\sum_{i=1}^N\sum_{j=1}^J\frac{1}{T}\operatorname{var}(\alpha_T\dot\rho(\alpha_T^{-1}e_{is,\alpha})\phi_j(\mathbf{x}_s))\\
&\quad +C\sum_{i=1}^N\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T|(\mathbf{b}_{i,\alpha}-\mathbf{b}_i)'\Phi(\mathbf{x}_s)\phi_j(\mathbf{x}_s)|)^2+C\sum_{i=1}^N\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T|z_{is}\phi_j(\mathbf{x}_s)|)^2\\
&\leq\ O(NJ/T+NJ^2\alpha_T^{-2(k-1)}+NJ^{1-2\eta}),
\end{aligned}
$$

where the first inequality is due to the triangular inequality and $|\dot\rho(t_1)-\dot\rho(t_2)|\leq 2|t_1-t_2|$; the second inequality is due to $E\dot\rho(\alpha_T^{-1}e_{is,\alpha})\Phi(\mathbf{x}_s)=0$ and that $e_{is}-e_{is,\alpha}=(\mathbf{b}_{i,\alpha}-\mathbf{b}_i)'\Phi(\mathbf{x}_s)-z_{is}$.

(ii) The bound for $\|\mathbf{B}_3\|_F^2$ is similar to $\|\mathbf{B}_1\|_F^2$. Since $\|\widetilde{\mathbf{V}}^{-1}\|=O_P(\chi_N^{-1})$,

$$
\begin{aligned}
\|\mathbf{B}_1\|_F^2\ &\leq\ \frac{1}{N^2}\|\mathbf{\Lambda}\|^2\|\chi_N\delta(\mathbf{x})\mathbf{A}\|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2\|\widetilde{\mathbf{V}}^{-1}\|^2=O_P(\frac{\delta_N^2(\mathbf{x})}{\chi_N^2N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2)\\
&=\ O_P(\|\mathbf{M}_\alpha\|^2/\chi_N)
\end{aligned}
$$

(iii) By Proposition C.3, $\frac{1}{T}\sum_t\|\mathbf{R}_t\|^2\leq N\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2$. Hence

$$
\begin{aligned}
\|\mathbf{B}_2\|_F^2\ &\leq\ O_P(1)\|\widetilde{\mathbf{V}}^{-1}\|_F^2\frac{1}{T}\sum_t\|E(\mathbf{f}_t|\mathbf{x}_t)\|^2\frac{1}{T}\sum_t\|\mathbf{R}_t\|^2\\
&=\ O_P(N\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2/\chi_N).
\end{aligned}
$$

The bound for $\|\mathbf{B}_6\|_F^2$ is similar.

(iv) We have

$$
\begin{aligned}
\|\mathbf{B}_4\|_F^2\ &\leq\ \frac{1}{N^2}\|\mathbf{M}_\alpha\mathbf{A}\|^2\|\frac{1}{T}\sum_{t=1}^T\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\mathbf{A}\|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2\|\widetilde{\mathbf{V}}^{-1}\|^2\\
&=\ O_P(\|\mathbf{M}_\alpha\|^2\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2/N^2\chi_N^{-2})\\
&=\ O_P(\|\mathbf{M}_\alpha\|_F^4/(N\chi_N^2)).
\end{aligned}
$$

Also, $\|\mathbf{B}_8\|_F^2\leq\frac{1}{N^2}(\frac{1}{T}\sum_{t=1}^T\|\mathbf{R}_t\|^2)^2\|\widehat{\mathbf{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2=O_P(N(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^2/\chi_N^2)$.

(v) $\mathbf{B}_5$ and $\mathbf{B}_7$ are bounded similarly.

$$
\begin{aligned}
\|\mathbf{B}_7\|_F^2\ &\leq\ \frac{1}{N^2}\|\frac{1}{T}\sum_{t=1}^T\mathbf{R}_t\Phi(\mathbf{x}_t)'\mathbf{A}\|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2\|\widetilde{\mathbf{V}}^{-1}\|^2\\
&=\ O_P(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2 J\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|^2/(N\chi_N^2))
\end{aligned}
$$

$$= O_P(\|\mathbf{M}_\alpha\|^2 J \max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2/\chi_N^2).$$

Q.E.D.

Given Proposition B.1, due to

$$\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F^2 \leq 2\|\mathbf{M}_\alpha\|^2\|\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}\|_F^2 + 2\|\mathbf{M}_\alpha'\mathbf{\Lambda}\|^2\|\mathbf{H}\|_F^2,$$

the rate of convergence for $\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F^2$ can be improved, reaching a sharper bound than $\|\mathbf{M}_\alpha\|^2\|\widehat{\mathbf{\Lambda}}\|_F^2$. This is given in Lemma C.6 below. As a result, rates for $\mathbf{B}_1, \mathbf{B}_4, \mathbf{B}_7$ can be improved as well.

Write

$$a_T^2 := \frac{J}{T} + J^{1-2\eta}, \quad b_{NT}^2 := \frac{J\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|}{T} + \frac{J}{TN} + \frac{J}{T}\alpha_T^{-\zeta_2}.$$

Lemma C.6. *Given Proposition B.1, we have*

$$\begin{aligned}
\frac{1}{N^2}\|\mathbf{M}_\alpha'\mathbf{\Lambda}\|_F^2 &= O_P(b_{NT}^2), \\
\frac{1}{N^2}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F^2 &= O_P(b_{NT}^2) + O_P(\chi_N^{-1}a_T^4).
\end{aligned}$$

Proof. The proof is a straightforward calculation as follows:

$$\begin{aligned}
E\|\mathbf{M}_\alpha'\mathbf{\Lambda}\|_F^2 &= E\|\sum_{i=1}^N \boldsymbol{\lambda}_i \mathbf{M}_{i,\alpha}'\|_F^2 \\
&= E\|\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N \boldsymbol{\lambda}_i \alpha_T \dot{\rho}(\alpha_T^{-1}e_{is})\Phi(\mathbf{x}_s)'\|_F^2 \\
&= \sum_{k=1}^K\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N \lambda_{ik}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})\phi_j(\mathbf{x}_s))^2 \\
&= \sum_{k=1}^K\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N 2\lambda_{ik}e_{is}1\{|e_{is}| < \alpha_T\}\phi_j(\mathbf{x}_s))^2 \\
&\quad + \sum_{k=1}^K\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N \lambda_{ik}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})1\{|e_{is}| \geq \alpha_T\}\phi_j(\mathbf{x}_s))^2 \\
&\leq 8\sum_{k=1}^K\sum_{j=1}^J \operatorname{var}(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N \lambda_{ik}e_{is}\phi_j(\mathbf{x}_s)) \\
&\quad +12\sum_{k=1}^K\sum_{j=1}^J E(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N |\lambda_{ik}e_{is}|1\{|e_{is}| > \alpha_T\}\phi_j(\mathbf{x}_s))^2.
\end{aligned}$$

To bound the first term, Let E_w be the conditional expectation given $\mathbf{x}_s$. We need to bound $\sum_{i,l\leq N}|E_w e_{is}e_{ls}|$. Note that $e_{is}=x_{is}-E(x_{is}|\mathbf{x}_s)=\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_s+u_{is}$. Since $E(\mathbf{u}_s|\mathbf{f}_s,\mathbf{x}_t)=0$, we have

$$
\begin{aligned}
E(\boldsymbol{\gamma}_s\mathbf{u}_s'|\mathbf{x}_s) &= E(\mathbf{f}_s\mathbf{u}_s'|\mathbf{x}_s)-(E\mathbf{f}_s|\mathbf{x}_s)E(\mathbf{u}_s'|\mathbf{x}_s)\\
&= E(\mathbf{f}_s\mathbf{u}_s'|\mathbf{x}_s)=E(\mathbf{f}_sE(\mathbf{u}_s'|\mathbf{x}_s,\mathbf{f}_s)|\mathbf{x}_s)=0.
\end{aligned}
$$

Hence $E_w(e_{is}e_{ls})=E_w(\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_s+u_{is})(\boldsymbol{\lambda}_l'\boldsymbol{\gamma}_s+u_{ls})=\boldsymbol{\lambda}_i'\operatorname{cov}(\boldsymbol{\gamma}_s)\boldsymbol{\lambda}_l+E_w(u_{is}u_{ls})$. Therefore,

$$
\begin{aligned}
& 8\sum_{k=1}^K\sum_{j=1}^J\operatorname{var}(\frac{1}{T}\sum_{s=1}^T\sum_{i=1}^N\lambda_{ik}e_{is}\phi_j(\mathbf{x}_s))=8\sum_{k=1}^K\sum_{j=1}^J\frac{1}{T}\operatorname{var}(\sum_{i=1}^N\lambda_{ik}e_{is}\phi_j(\mathbf{x}_s))\\
=\;& 8\sum_{k=1}^K\sum_{j=1}^J\frac{1}{T}\sum_{i=1}^N\sum_{l=1}^N\lambda_{ik}\lambda_{lk}E\{E_w(e_{is}e_{ls})\phi_j(\mathbf{x}_s)^2\}\\
\leq\;& C\sum_{j=1}^J\frac{1}{T}E\phi_j(\mathbf{x}_s)^2\sup_{\mathbf{x}}\sum_{i=1}^N\sum_{l=1}^N|E_w(e_{is}e_{ls})|\\
\leq\;& \frac{CJ}{T}\sum_{i=1}^N\sum_{l=1}^N|\boldsymbol{\lambda}_i'\operatorname{cov}(\boldsymbol{\gamma}_s)\boldsymbol{\lambda}_l|+\frac{CJ}{T}\sup_{\mathbf{x}}\sum_{i=1}^N\sum_{l=1}^N|E_w(u_{is}u_{ls})|\\
\leq\;& \frac{CJ}{T}N^2\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|+\frac{CJN}{T}\sup_{\mathbf{x}}\max_{i\leq N}\sum_{l=1}^N|E_w(u_{is}u_{ls})|\\
=\;& O(JN^2\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|/T+JN/T).
\end{aligned}
$$

Note that the second term is bounded by

$$
\begin{aligned}
\leq\;& C\sum_{k=1}^K\sum_{j=1}^J\frac{1}{T}\sum_{i=1}^N\sum_{l=1}^N E|e_{is}1\{|e_{is}|>\alpha_T\}|e_{ls}|1\{|e_{ls}|>\alpha_T\}\phi_j(\mathbf{x}_s)^2\\
&+C\sum_{k=1}^K\sum_{j=1}^J\frac{1}{T^2}\sum_{s=1}^T\sum_{i=1}^N\sum_{t\neq s}^T\sum_{l=1}^N E|e_{is}|1\{|e_{is}|>\alpha_T\}|\phi_j(\mathbf{x}_s)|E|e_{lt}|1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)|\\
\leq\;& C\sum_{k=1}^K\sum_{j=1}^J\frac{1}{T}\sum_{i=1}^N\sum_{l=1}^N\sup_{\mathbf{x}}E_w|e_{is}|1\{|e_{is}|>\alpha_T\}|e_{ls}|1\{|e_{ls}|>\alpha_T\}\\
&+C\sum_{k=1}^K\sum_{j=1}^J\sum_{i=1}^N\sum_{l=1}^N(\sup_{\mathbf{x}}E|e_{is}|1\{|e_{is}|>\alpha_T\})^2\\
\leq\;& C\frac{KJ}{T}N^2\max_i\sup_{\mathbf{x}}E_we_{is}^2 1\{|e_{is}|>\alpha_T\}+CKJN^2(\max_i\sup_{\mathbf{x}}E|e_{is}|1\{|e_{is}|>\alpha_T\})^2\\
=\;& O(N^2J\alpha_T^{-\zeta_2}/T+N^2J\alpha_T^{-2(\zeta_2+1)})=O(N^2J\alpha_T^{-\zeta_2}/T).
\end{aligned}
$$

Hence $\frac{1}{N^2}E\|\mathbf{M}_\alpha'\mathbf{\Lambda}\|_F^2 = O(J\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|/T + J/(TN) + J\alpha_T^{-\zeta_2}/T) := O(b_{NT}^2)$.

(ii) Write $a_T^2 := \frac{J}{T} + J^{1-2\eta}$. Proposition (B.1) shows $\frac{1}{N}\|\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}\|_F^2 = O_P(a_T^2\chi_N^{-1})$. In addition, Lemma C.4 implies $\|\mathbf{H}\| = O_P(1)$. Lemma C.5 implies $\frac{1}{N}\|\mathbf{M}_\alpha\|^2 = O_P(a_T^2)$. Thus

$$\begin{aligned}\frac{1}{N^2}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F^2 &\leq \frac{2}{N^2}\|\mathbf{M}_\alpha\|^2\|\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}\|_F^2 + \frac{2}{N^2}\|\mathbf{M}_\alpha'\mathbf{\Lambda}\|^2\|\mathbf{H}\|_F^2 \\ &\leq O_P(a_T^4\chi_N^{-1}) + O_P(b_{NT}^2).\end{aligned}$$

Lemma C.7. *Suppose $J^2/T + J^{-\eta} + \sqrt{\log N/T} \ll \chi_N$.*

$$\begin{aligned}&\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H})\| \leq O_P(\chi_N^{-1/2})(\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F + (\max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2}) \\ \leq\ & O_P(\chi_N^{-1/2})(a_T^4\chi_N^{-1} + b_{NT}^2 + J^{1-2\eta} + \alpha_T^{-2(\zeta_1-1)}\frac{J^3\log N}{T} + \frac{J^3\log N\log J}{T^2})^{1/2}.\end{aligned}$$

In addition, $\|\frac{1}{N}\widehat{\mathbf{\Lambda}}'(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H})\|$ has the same rate of convergence.

Proof. $\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}) = \sum_{i=1}^8 \mathbf{\Lambda}'\mathbf{B}_i$. Keep in mind that $\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|$ and $\|\widehat{\mathbf{\Lambda}}'\mathbf{M}_\alpha\|$ have sharper bounds than $\|\mathbf{\Lambda}\|\|\mathbf{M}_\alpha\|$, $\|\widehat{\mathbf{\Lambda}}\|\|\mathbf{M}_\alpha\|$, given in Lemma C.6.

For $i \neq 3, 4, 5$, we simply use $\|\mathbf{\Lambda}'\mathbf{B}_i\| \leq \|\mathbf{\Lambda}\|\|\mathbf{B}_i\| = O(\sqrt{N})\|\mathbf{B}_i\|$ and Lemma C.5. But note that for $\mathbf{B}_1, \mathbf{B}_7$, the first upper bound in the lemma is used. So

$$\begin{aligned}\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_1\| &= O_P(\chi_N^{-1/2})\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\| \\ \frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_2\| &= O_P(\chi_N^{-1/2}(\max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2}) = \frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_6\| \\ \frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_7\| &= O_P(\chi_N^{-1}(\max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2}J^{1/2}\|\mathbf{M}_\alpha\widehat{\mathbf{\Lambda}}\|/N), \\ \frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_8\| &= O_P(\chi_N^{-1}\max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2).\end{aligned}$$

As for $\mathbf{B}_3, \mathbf{B}_4, \mathbf{B}_5$, we have

$$\begin{aligned}\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_3\| &\leq O_P(\chi_N^{-1})\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|\|\frac{1}{T}\sum_{t=1}^T \Phi(\mathbf{x}_t)E(\mathbf{f}_t|\mathbf{x}_t)'\| \\ &= O_P(\chi_N^{-1/2}\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\| \\ \frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_4\| &\leq O_P(\chi_N^{-1})\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|\|\frac{1}{TN}\sum_{t=1}^T \Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\|\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|\end{aligned}$$

$$= O_P(\chi_N^{-1}\frac{1}{N^2}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|)$$

$$\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{B}_5\| \leq O_P(\chi_N^{-1})\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|\|\frac{1}{T\sqrt{N}}\sum_{t=1}^T\Phi(\mathbf{x}_t)\mathbf{R}_t'\|$$

$$= O_P(\chi_N^{-1}\frac{1}{N}\|\mathbf{\Lambda}'\mathbf{M}_\alpha\|(J\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2}).$$

Hence

$$\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| \leq O(1)\sum_{i=1}^8\|\frac{1}{N}\mathbf{\Lambda}'\mathbf{B}_i\|$$

$$\leq O_P(\chi_N^{-1/2})(\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F+(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2})$$

$$+O_P(\chi_N^{-1})(\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F+(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2})^2$$

$$+O_P(\chi_N^{-1})\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F(J\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2}. \tag{C.2}$$

In addition, by Lemma C.6 with $b_{NT}^2 := \frac{J\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|}{T}+\frac{J}{TN}+\frac{J}{T}\alpha_T^{-\zeta_2}$,

$$\chi_N^{-1}J\frac{1}{N^2}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F^2 = \chi_N^{-1}JO_P(b_{NT}^2) + \chi_N^{-1}JO_P(\chi_N^{-1}a_T^4) = o_P(1). \tag{C.3}$$

The last equality is due to $\operatorname{cov}(\boldsymbol{\gamma}_t)=O(1)$, $\eta\geq 2$, and $J^2/T+J^{-\eta}\ll\chi_N$. By Proposition C.3, with the assumption $J^3\log^2 N=O(T)$ and $\zeta_1>2$,

$$\chi_N^{-1}J\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2 = \chi_N^{-1}JO_P(J^{1-2\eta}+\alpha_T^{-2(\zeta_1-1)}\frac{J^3\log N}{T}+\frac{J^3\log N\log J}{T^2})$$

$$= o_P(1). \tag{C.4}$$

Hence the second and third terms of (C.2) are dominated, so

$$\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| \leq O_P(\chi_N^{-1/2})(\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F+(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2})$$

$$\leq O_P(\chi_N^{-1/2})(a_T^4\chi_N^{-1}+b_{NT}^2+J^{1-2\eta}+\alpha_T^{-2(\zeta_1-1)}\frac{J^3\log N}{T}+\frac{J^3\log N\log J}{T^2})^{1/2}.$$

In addition, $\|\frac{1}{N}\widehat{\mathbf{\Lambda}}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\|\leq\|\mathbf{H}\|\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\|+\frac{1}{N}\|\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H}\|^2$. Note that $\|\mathbf{H}\|=O_P(1)$ and $\frac{1}{N}\|\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H}\|_F^2=O_P(a_T^2\chi_N^{-1})$. Hence

$$\|\frac{1}{N}\widehat{\mathbf{\Lambda}}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| \leq O_P(\chi_N^{-1/2})(\frac{1}{N}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|_F+(\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)^{1/2})$$

$$+O_P(a_T^2\chi_N^{-1}) \tag{C.5}$$

So $\|\frac{1}{N}\widehat{\mathbf{\Lambda}}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})\|$ has the same rate of convergence as $\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})\|$ in the above.

Q.E.D.

Recall Lemma C.4 shows $\|\mathbf{H}\|=O_P(1)$. We now prove $\|\mathbf{H}^{-1}\|=O_P(1)$.

Lemma C.8. *Suppose $J^2/T+J^{-\eta}+\sqrt{\log N/T}\ll\chi_N$.*

$$\|\mathbf{H}'\mathbf{\Sigma}_{\Lambda,N}\mathbf{H}-\mathbf{I}\|=o_P(1)$$

which then implies $\|\mathbf{H}^{-1}\|=O_P(1)$.

Proof. Note that

$$\begin{aligned}
\mathbf{I} &= \frac{1}{N}\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Lambda}} \\
&= \frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})+\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})'\mathbf{\Lambda H} \\
&\quad +\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})+\mathbf{H}'\mathbf{\Sigma}_{\Lambda,N}\mathbf{H}.
\end{aligned}$$

Hence it suffices to show $\frac{1}{N}\|\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H}\|^2=o_P(1)=\|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})'\mathbf{\Lambda H}\|$. By Proposition B.1 with $a_T^2:=\frac{J}{T}+J^{1-2\eta}$ and assumption $J^2/T+J^{-\eta}\ll\chi_N$,

$$\frac{1}{N}\|\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H}\|_F^2 = O_P(a_T^2\chi_N^{-1})=o_P(1).$$

Also by Lemma C.7, $\|\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda H})\|\leq o_P(1)$. Hence $\mathbf{H}'\mathbf{\Sigma}_{\Lambda,N}\mathbf{H}=\mathbf{I}+o_P(1)$. It then follows from the fact that $\mathbf{\Sigma}_{\Lambda,N}=O(1)$, we have $\lambda_{\min}(\mathbf{H}'\mathbf{H})\geq c$ for some $c>0$ with probability approaching one. This then implies $\|\mathbf{H}^{-1}\|=O_P(1)$.

Q.E.D.

Lemma C.9. $\max_{i\leq N}\|\mathbf{M}_{i,\alpha}\|=O_P(J^{-\eta}\sqrt{J}+\sqrt{J(\log N)/T})$

Proof. First, it follows from the proof of Proposition C.2 that

$$\max_i\|\frac{1}{T}\sum_{s=1}^T\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is,\alpha})\Phi(\mathbf{x}_s)\|=O_P(\sqrt{\frac{J\log N}{T}}).$$

Secondly, since $|\dot\rho(t_1)-\dot\rho(t_2)|\leq 2|t_1-t_2|$,

$$
\begin{aligned}
& \max_i \|\frac{1}{T}\sum_{s=1}^T \alpha_T(\dot\rho(\alpha_T^{-1}e_{is})-\dot\rho(\alpha_T^{-1}e_{is,\alpha}))\Phi(\mathbf{x}_s)\| \\
\leq\ & \max_i \|\frac{1}{T}\sum_{s=1}^T 2|e_{is}-e_{is,\alpha}|\Phi(\mathbf{x}_s)\| \\
\leq\ & \max_i \|\frac{1}{T}\sum_{s=1}^T 2|(\mathbf{b}_{i,\alpha}-\mathbf{b}_i)'\Phi(\mathbf{x}_t)-z_{it}|\Phi(\mathbf{x}_s)\| \\
\leq\ & 2\max_i \|\mathbf{b}_{i,\alpha}-\mathbf{b}_i\|O_P(J)+O_P(J^{-\eta}\sqrt{J}) \\
=\ & O_P(J^{-\eta}\sqrt{J}+J\alpha_T^{-(k-1)}).
\end{aligned}
$$

The result then follows from the triangular inequality.

Q.E.D.

C.3 Technical Lemmas for factors

Lemma C.10. $\sum_{t=1}^T\|\mathbf{u}_t'\mathbf{M}_\alpha\|^2=O_P(JN\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|+JN^2/T+J+JN^2\alpha_T^{-\zeta_2})$.

Proof. Note that $E\sum_{t=1}^T\|\mathbf{u}_t'\mathbf{M}_\alpha\|^2 = \frac{1}{T^2}\sum_{t=1}^T\sum_{j=1}^J E(\sum_{i=1}^N\sum_{s=1}^T u_{it}\alpha_T\dot\rho(\alpha_T^{-1}e_{is})\phi_j(\mathbf{x}_s))^2$. We now bound the right hand side. In fact, since $e_{is}=\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_s+u_{is}$,

$$
\begin{aligned}
& E(\sum_{i=1}^N\sum_{s=1}^T u_{it}\alpha_T\dot\rho(\alpha_T^{-1}e_{is})\phi_j(\mathbf{x}_s))^2 \\
\leq\ & 8E(\sum_{i=1}^N\sum_{s=1}^T u_{it}e_{is}1\{|e_{is}|<\alpha_T\}\phi_j(\mathbf{x}_s))^2 \\
& +2E(\sum_{i=1}^N\sum_{s=1}^T u_{it}\alpha_T\dot\rho(\alpha_T^{-1}e_{is})1\{|e_{is}|\geq\alpha_T\}\phi_j(\mathbf{x}_s))^2 \\
\leq\ & CE(\sum_{i=1}^N\sum_{s=1}^T u_{it}e_{is}\phi_j(\mathbf{x}_s))^2+CE(\sum_{i=1}^N\sum_{s=1}^T |u_{it}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|)^2 \\
\leq\ & CE(\sum_{i=1}^N\sum_{s=1}^T u_{it}\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_s\phi_j(\mathbf{x}_s))^2+CE(\sum_{i=1}^N\sum_{s=1}^T (u_{it}u_{is}-E(u_{it}u_{is}))\phi_j(\mathbf{x}_s))^2 \\
& +CE(\sum_{i=1}^N\sum_{s=1}^T (Eu_{it}u_{is})\phi_j(\mathbf{x}_s))^2+CE(\sum_{i=1}^N\sum_{s=1}^T |u_{it}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|)^2.
\end{aligned}
$$

(C.6)

(C.7)

The first term on the right hand side of (C.7) is bounded uniformly in t by

$$\begin{aligned}
& E(\sum_{i=1}^{N}\sum_{s=1}^{T} u_{it}\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_s\phi_j(\mathbf{x}_s))^2 \\
= & \sum_{i=1}^{N}\sum_{l=1}^{N}\boldsymbol{\lambda}_i' E\boldsymbol{\gamma}_t u_{lt}u_{it}\phi_j(\mathbf{x}_t)^2\boldsymbol{\gamma}_t'\boldsymbol{\lambda}_l + \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}\boldsymbol{\lambda}_i' \operatorname{cov}(\boldsymbol{\gamma}_s)E\phi_j(\mathbf{x}_s)^2\boldsymbol{\lambda}_l E u_{lt}u_{it} \\
\leq & \sum_{i=1}^{N}\sum_{l=1}^{N} E[|(Eu_{lt}u_{it}|\mathbf{x}_t,\mathbf{f}_t)|\phi_j(\mathbf{x}_t)^2\|\boldsymbol{\gamma}_t\|^2]\max_i\|\boldsymbol{\lambda}_i\|^2 \\
& +T\sum_{i=1}^{N}\sum_{l=1}^{N}\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|E\phi_j(\mathbf{x}_s)^2|Eu_{lt}u_{it}|\max_i\|\boldsymbol{\lambda}_i\|^2 \\
\leq & NC\sup_{\mathbf{x},\mathbf{f}}\max_i\sum_{l=1}^{N}|(Eu_{lt}u_{it}|\mathbf{x}_t,\mathbf{f}_t)|\sup_{\mathbf{x}}E(\|\boldsymbol{\gamma}_t\|^2|\mathbf{x}_t=\mathbf{x})E\phi_j(\mathbf{x}_t)^2 \\
& +\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|TNC\max_i\sum_{l=1}^{N}|Eu_{lt}u_{it}| \\
\leq & NC\sup_{\mathbf{x},\mathbf{f}}\max_i\sum_{l=1}^{N}|(Eu_{lt}u_{it}|\mathbf{x}_t,\mathbf{f}_t)|\|\operatorname{cov}(\boldsymbol{\gamma}_t)\| + \|\operatorname{cov}(\boldsymbol{\gamma}_s)\|TNC\max_i\sum_{l=1}^{N}|Eu_{lt}u_{it}| \\
= & O(TN\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|).
\end{aligned}$$

The second term of (C.7) : note that for some $v>1$, $E\{Eu_{it}^4|\mathbf{x}_t\}^v<\infty$, uniformly in t,

$$\begin{aligned}
& E(\sum_{i=1}^{N}\sum_{s=1}^{T}(u_{it}u_{is}-E(u_{it}u_{is}))\phi_j(\mathbf{x}_s))^2 \\
= & \sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{l=1}^{N}\sum_{k=1}^{T}E(u_{it}u_{is}-E(u_{it}u_{is}))(u_{lt}u_{lk}-E(u_{lt}u_{lk}))\phi_j(\mathbf{x}_k)\phi_j(\mathbf{x}_s) \\
= & \sum_{i=1}^{N}\sum_{l=1}^{N}E(u_{it}^2-Eu_{it}^2)(u_{lt}^2-Eu_{lt}^2)\phi_j(\mathbf{x}_t)^2 + \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}Eu_{it}u_{lt}Eu_{ls}u_{is}\phi_j(\mathbf{x}_s)^2 \\
\leq & \sum_{i=1}^{N}\sum_{l=1}^{N}E(u_{it}^2-Eu_{it}^2)(u_{lt}^2-Eu_{lt}^2)\phi_j(\mathbf{x}_t)^2 \\
& +CT(\max_i\sum_{l=1}^{N}|Eu_{it}u_{lt}|)(\sup_{\mathbf{x}}\max_l\sum_{i=1}^{N}|Eu_{ls}u_{is}|\mathbf{x})|E\phi_j(\mathbf{x}_s)^2) \\
= & O(N^2+T).
\end{aligned}$$

The third term of (C.7) is bounded as: uniformly in t,

$$E(\sum_{i=1}^{N}\sum_{s=1}^{T}(Eu_{it}u_{is})\phi_j(\mathbf{x}_s))^2 = E(\sum_{i=1}^{N}(Eu_{it}^2)\phi_j(\mathbf{x}_t))^2 = O(N^2).$$

Finally, the fourth term of (C.7) is :

$$
\begin{aligned}
& E(\sum_{i=1}^{N}\sum_{s=1}^{T}|u_{it}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|)^2 \\
= \quad & E\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{l=1}^{N}\sum_{k=1}^{T}|u_{it}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)||u_{lt}e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)| \\
= \quad & \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}E|u_{it}u_{lt}e_{lt}1\{|e_{lt}|>\alpha_T\}\phi_j(\mathbf{x}_t)|E|e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)| \\
& +\sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}e_{it}1\{|e_{it}|>\alpha_T\}u_{lt}e_{lt}1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2 \\
& +E\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{k\neq t}|u_{it}u_{lt}e_{it}1\{|e_{it}|>\alpha_T\}\phi_j(\mathbf{x}_t)|E|e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)| \\
& +\sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}E|u_{it}u_{lt}|E|e_{ls}1\{|e_{ls}|>\alpha_T\}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)^2| \\
& +\sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}\sum_{k\neq s,t}E|u_{it}u_{lt}|E|e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|E|e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)| \\
:= \quad & \sum_{i=1}^{5}a_i.
\end{aligned}
$$

We now study $a_1,\cdots,a_5$ term by term. By Holder's inequality, and the assumption that $E\{E(u_{it}^4|\mathbf{x})\}^v<\infty$, and by repeatedly using Cauchy-Schwarz inequality, uniformly in t,

$$
\begin{aligned}
a_1 &= \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}E|u_{it}u_{lt}e_{lt}1\{|e_{lt}|>\alpha_T\}\phi_j(\mathbf{x}_t)|E|e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)| \\
&\leq \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}(Ee_{lt}^2 1\{|e_{lt}|>\alpha_T\})^{1/2}(Eu_{it}^2u_{lt}^2\phi_j(\mathbf{x}_t)^2)^{1/2}\sup_{\mathbf{x}}E(|e_{is}|1\{|e_{is}|>\alpha_T\}|\mathbf{x})E|\phi_j(\mathbf{x}_s)| \\
&\leq CTN^2\max_i\{E[Eu_{it}^4|\mathbf{x}_t]^v\}^{1/(2v)}\alpha_T^{-(\zeta_2+1)-\zeta_2/2}=O(TN^2\alpha_T^{-(\zeta_2+1)-\zeta_2/2}) \\
a_2 &= \sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}e_{it}1\{|e_{it}|>\alpha_T\}u_{lt}e_{lt}1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2 \\
&\leq \sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_t1\{|e_{it}|>\alpha_T\}u_{lt}\boldsymbol{\lambda}_l'\boldsymbol{\gamma}_t1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2 \\
&\quad +\sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}\boldsymbol{\lambda}_i'\boldsymbol{\gamma}_t1\{|e_{it}|>\alpha_T\}u_{lt}^2 1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}^2 1\{|e_{it}|>\alpha_T\}u_{lt}\boldsymbol{\lambda}_l'\boldsymbol{\gamma}_t 1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2\\
&+\sum_{i=1}^{N}\sum_{l=1}^{N}E|u_{it}^2 1\{|e_{it}|>\alpha_T\}u_{lt}^2 1\{|e_{lt}|>\alpha_T\}|\phi_j(\mathbf{x}_t)^2\\
\leq\;& C\sum_{i=1}^{N}\sum_{l=1}^{N}\max_i(Eu_{it}^4)^{1/2}(E\{E\|\boldsymbol{\gamma}_t\|^4|\mathbf{x}_t\}^v)^{1/(2v)}\\
&+C\sum_{i=1}^{N}\sum_{l=1}^{N}[E(u_{it}u_{lt}^2)^{4/3}]^{3/4}(E\|\boldsymbol{\gamma}_t\|^4\phi_j(\mathbf{x}_t)^8)^{1/4}\\
&+C\sum_{i=1}^{N}\sum_{l=1}^{N}\{E[E(u_{it}^2u_{lt}^2|\mathbf{x}_t)]^v\}^{1/v}=O(N^2)\\
a_3 =\;& E\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{k\neq t}|u_{it}u_{lt}e_{it}1\{|e_{it}|>\alpha_T\}\phi_j(\mathbf{x}_t)|E|e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)|\\
\leq\;& \sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{k\neq t}(E|u_{it}u_{lt}\phi_j(\mathbf{x}_t)|^2)^{1/2}(Ee_{it}^2 1\{|e_{it}|>\alpha_T\})^{1/2}E|e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)|\\
\leq\;& TC\sum_{i=1}^{N}\sum_{l=1}^{N}\{E[E(u_{it}^2u_{lt}^2|\mathbf{x}_t)]^v\}^{1/2v}\alpha_T^{-\zeta_2/2-(\zeta_2+1)}\\
=\;& O(N^2T\alpha_T^{-\zeta_2/2-(\zeta_2+1)})\\
a_4 =\;& \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}E|u_{it}u_{lt}|E|e_{ls}1\{|e_{ls}|>\alpha_T\}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)^2|\\
=\;& O(TN^2\alpha_T^{-\zeta_2})\\
a_5 =\;& \sum_{i=1}^{N}\sum_{s\neq t}\sum_{l=1}^{N}\sum_{k\neq s,t}E|u_{it}u_{lt}|E|e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|E|e_{lk}1\{|e_{lk}|>\alpha_T\}\phi_j(\mathbf{x}_k)|\\
=\;& O(N^2T^2\alpha_T^{-2(\zeta_2+1)}).
\end{aligned}
$$

Therefore, uniformly in $t\leq T$,

$$
E(\sum_{i=1}^{N}\sum_{s=1}^{T}|u_{it}e_{is}1\{|e_{is}|>\alpha_T\}\phi_j(\mathbf{x}_s)|)^2=O(TN^2\alpha_T^{-(\zeta_2+1)-\zeta_2/2}+N^2+TN^2\alpha_T^{-\zeta_2}+N^2T^2\alpha_T^{-2(\zeta_2+1)}).
$$

Consequently, (note that $JN^2\alpha_T^{-\zeta_2}+JN^2T\alpha_T^{-2(\zeta_2+1)}\geq JN^2\sqrt{T}\alpha_T^{-(\zeta_2+1)-\zeta_2/2}$)

$$
E\sum_{t=1}^{T}\|\mathbf{u}_t'\mathbf{M}_\alpha\|^2=O(JN\|\operatorname{cov}(\boldsymbol{\gamma}_s)\|+JN^2/T+J+JN^2\alpha_T^{-\zeta_2}+JN^2T\alpha_T^{-2(\zeta_2+1)}).
$$

Lemma C.11.

$$\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2 = O_P(N^2J^4\log N\log J + J^{2-2\eta}T^2N + T\alpha_T^{-2(\zeta_1-1)}N^2J^4\log N).$$

Proof. Recall that $R_{it} = R_{1,it} + R_{2,it} + R_{3,it}$, where

$$\begin{aligned} R_{1,it} &:= \frac{1}{T}\sum_{k=1}^T \alpha_T[\dot\rho(\alpha_T^{-1}e_{ik,\alpha}) - \dot\rho(\alpha_T^{-1}e_{ik})]\Phi(\mathbf{x}_k)'\mathbf{A}\Phi(\mathbf{x}_t) \\ R_{2,it} &:= \Phi(\mathbf{x}_t)'(\mathbf{R}_{i,b} + \mathbf{b}_{i,\alpha} - \mathbf{b}_i), \qquad R_{3,it} := -z_{it}. \end{aligned}$$

In addition, recall $e_{it} = e_{it,\alpha} + \Delta_{it,\alpha}$, where $\Delta_{it,\alpha} = (\mathbf{b}_{i,\alpha} - \mathbf{b}_i)'\Phi(\mathbf{x}_t) - z_{it}$. For notational simplicity, we also write $H_{kt} := \Phi(\mathbf{x}_k)'\mathbf{A}\Phi(\mathbf{x}_t)$.

$$\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2 \le C\sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}R_{1,it})^2 + C\sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}R_{2,it})^2 + C\sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}R_{3,it})^2.$$

We look at these terms respectively.

bounding the first term

$$\begin{aligned} &\sum_{s=1}^T\sum_{t=1}^T E(\sum_{i=1}^N u_{is}R_{1,it})^2 \\ =\ & \sum_{s=1}^T\sum_{t=1}^T E\{\sum_{i=1}^N u_{is}\frac{1}{T}\sum_{k=1}^T \alpha_T[\dot\rho(\alpha_T^{-1}e_{ik,\alpha}) - \dot\rho(\alpha_T^{-1}e_{ik})]\Phi(\mathbf{x}_k)'\mathbf{A}\Phi(\mathbf{x}_t)\}^2 \\ \le\ & C\sum_{s=1}^T\sum_{t=1}^T E\{\sum_{i=1}^N u_{is}\frac{1}{T}\sum_{k=1}^T \Delta_{ik,\alpha}H_{kt}\}^2 \\ & + C\sum_{s=1}^T\sum_{t=1}^T E\{\sum_{i=1}^N |u_{is}|\frac{1}{T}\sum_{k=1}^T |\Delta_{ik,\alpha}|1\{|e_{ik}| > \alpha_T \text{ or } |e_{it,\alpha}| > \alpha_T\}|H_{kt}|\}^2 \\ :=\ & Ca_1 + Ca_2. \end{aligned}$$

For notational simplicity, let $I_{i,kt} := 1\{|e_{ik}| > \alpha_T \text{ or } |e_{it,\alpha}| > \alpha_T\}$.

$$\begin{aligned} a_1 &= \sum_{s=1}^T\sum_{t=1}^T E\{\sum_{i=1}^N u_{is}\frac{1}{T}\sum_{k=1}^T \Delta_{ik,\alpha}H_{kt}\}^2 \\ &= \frac{1}{T^2}\sum_{s=1}^T\sum_{t=1}^T\sum_{i=1}^N\sum_{j=1}^N\sum_{m=1}^T\sum_{k=1}^T E(Eu_{is}u_{js}|\{\mathbf{x}_l\}_{l\le T})\Delta_{ik,\alpha}H_{kt}\Delta_{jm,\alpha}H_{mt} \end{aligned}$$

$$
\begin{aligned}
&\leq \sup_{\mathbf{x}} \sum_{i=1}^{N} |E(u_{is}u_{js}|\mathbf{x}_s)| \frac{1}{T^2} \sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{m=1}^{T}\sum_{k=1}^{T} E \max_i |\Delta_{ik,\alpha}||H_{kt}\Delta_{jm,\alpha}H_{mt}| \\
&\leq CT^2N(\alpha_T^{-(k-1)}\sqrt{J}+J^{-\eta})^2J^2. \\
a_2 &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T} E\{\sum_{i=1}^{N}\sum_{k=1}^{T}|u_{is}\Delta_{ik,\alpha}I_{i,kt}H_{kt}|\}^2 \\
&\leq \frac{1}{T^2}\sum_{t,k,l\leq T}\sum_{s=t \text{ or } k \text{ or } l}\sum_{i=1}^{N}\sum_{j=1}^{N}(E(u_{is}u_{js}\Delta_{ik,\alpha}H_{kt}\Delta_{jl,\alpha}H_{lt})^2)^{1/2}(EI_{i,kt}I_{j,lt})^{1/2} \\
&\quad +\frac{1}{T^2}\sum_{t,k,l\leq T}\sum_{s\neq t,k,l}\sum_{i=1}^{N}\sum_{j=1}^{N}E|u_{is}u_{js}|(E(\Delta_{ik,\alpha}H_{kt}H_{lt}\Delta_{jl,\alpha})^2)^{1/2}(EI_{i,kt}I_{j,lt})^{1/2} \\
&\leq \frac{CN(N+T)}{T^2}(\alpha_T^{-(k-1)}\sqrt{J}+J^{-\eta})^2J^2\sum_{t,k,l\leq T}(EI_{i,kt}I_{j,lt})^{1/2} \\
&\leq CJ^2NT(N+T)(\alpha_T^{-(k-1)}\sqrt{J}+J^{-\eta})^2\alpha_T^{-(\zeta_2+2)/2}
\end{aligned}
$$

where the last inequality is due to, uniformly in i, j,

$$
\begin{aligned}
P(|e_{it,\alpha}|>\alpha_T) &\leq P(|e_{it}|>3\alpha_T/4)+P(\|\Phi(\mathbf{x}_t)\|>C\alpha_T^k)\leq C\alpha_T^{-(\zeta_2+2)}, \\
\sum_{t,k,l\leq T}(EI_{i,kt}I_{j,lt})^{1/2} &\leq CT^3\alpha_T^{-(\zeta_2+2)/2}.
\end{aligned}
$$

Therefore, $\sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}u_{is}R_{1,it})^2 = O_P((\alpha_T^{-(k-1)}\sqrt{J}+J^{-\eta})^2J^2TN(T+N\alpha_T^{-(\zeta_2+2)/2}))$.

bounding the second term

By Lemma C.3, $\max_{i\leq N}\|\mathbf{R}_{i,b}\|^2 = O_P(\alpha_T^{-2(\zeta_1-1)}+\alpha_T^{-2(\zeta_2+2)}+\frac{\log J}{T})\frac{J^3\log N}{T}$. Hence

$$
\begin{aligned}
\sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}u_{is}R_{2,it})^2 &= \sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}u_{is}\Phi(\mathbf{x}_t)'(\mathbf{R}_{i,b}+\mathbf{b}_{i,\alpha}-\mathbf{b}_i))^2 \\
&\leq 2\sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}u_{is}\Phi(\mathbf{x}_t)'\mathbf{R}_{i,b})^2 \\
&\quad +2\sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}u_{is}\Phi(\mathbf{x}_t)'(\mathbf{b}_{i,\alpha}-\mathbf{b}_i))^2 \\
&:= a_1+a_2, \text{ say}
\end{aligned}
$$

$$
\begin{aligned}
a_1 &\leq 2\sum_{s=1}^{T}\sum_{t=1}^{T}(\sum_{i=1}^{N}\|u_{is}\Phi(\mathbf{x}_t)\|)^2\max_i\|\mathbf{R}_{i,b}\|^2 \\
&= O_P(T^2N^2J)\max_i\|\mathbf{R}_{i,b}\|^2
\end{aligned}
$$

$$
\begin{aligned}
&= O_P((T\alpha_T^{-2(\zeta_1-1)} + T\alpha_T^{-2(\zeta_2+2)} + \log J)N^2J^4\log N). \\
E|a_2| &= 2\sum_{s=1}^T\sum_{i=1}^N\sum_{j=1}^N(\mathbf{b}_{i,\alpha}-\mathbf{b}_i)'Eu_{is}u_{js}\sum_{t=1}^T\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'(\mathbf{b}_{j,\alpha}-\mathbf{b}_j) \\
&\leq 2\sup_{\mathbf{x}}\max_i\sum_{j=1}^N|(Eu_{is}u_{js}|\mathbf{x}_s)|\max_i\|\mathbf{b}_{i,\alpha}-\mathbf{b}_i\|^2\sum_{s=1}^T\sum_{i=1}^N E\|\sum_{t=1}^T\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)\| \\
&\leq O(T^2\max_i\|\mathbf{b}_{i,\alpha}-\mathbf{b}_i\|^2N) = O(T^2N\alpha_T^{-2(k-1)}).
\end{aligned}
$$

Therefore,

$$
\sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}R_{2,it})^2 = O_P((T\alpha_T^{-2(\zeta_1-1)} + T\alpha_T^{-2(\zeta_2+2)} + \log J)N^2J^4\log N + T^2N\alpha_T^{-2(k-1)}).
$$

bounding the third term

$$
E\sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}R_{3,it})^2 = \sum_{s=1}^T\sum_{t=1}^T(\sum_{i=1}^N u_{is}z_{it})^2 = \sum_{s=1}^T\sum_{t=1}^T\sum_{i=1}^N\sum_{j=1}^N Eu_{is}u_{js}z_{it}z_{jt} = O(NT^2J^{-2\eta}).
$$

Hence the result follows.

Lemma C.12.

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^T\|\mathbf{D}_{t2}\|^2 \\
&= O_P(\chi_N^{-1})\left(\frac{1}{N^3}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2 + \frac{1}{N}\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2 + \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2 + \frac{1}{N^2T^2}\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2\right).
\end{aligned}
$$

Proof. First of all, note that $\max_i\sum_j|Eu_{is}u_{js}| < \infty$, hence

$$
E\frac{1}{T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{\Lambda}\|^2 = \sum_{j=1}^K E(\mathbf{u}_s'\boldsymbol{\lambda}_j)^2 = O(N).
$$

In addition, $\frac{1}{T}\sum_{t=1}^T\|\mathbf{D}_{t2}\|^2 = \frac{1}{T}\sum_{t=1}^T\|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'\mathbf{u}_t\|^2 \leq C\sum_{i=1}^8\frac{1}{N^2T}\sum_{t=1}^T\|\mathbf{u}_t'\mathbf{B}_i\|^2$.

$$
\begin{aligned}
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_1\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{\Lambda}\frac{1}{TN}\sum_{t=1}^T E(\mathbf{f}_t|\mathbf{x}_t)\Phi(\mathbf{x}_t)'\mathbf{A}\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2 \\
&= O_P(\chi_N^{-1})\frac{1}{N^4T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{\Lambda}\|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2 = O_P(\chi_N^{-1}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2/N^3),
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_2\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\boldsymbol{\Lambda}\frac{1}{TN}\sum_{t=1}^T E(\mathbf{f}_t|\mathbf{x}_t)\mathbf{R}_t'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&= O_P(\frac{\chi_N^{-1}}{N}\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2),\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_3\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\mathbf{A}\frac{1}{TN}\sum_{t=1}^T\Phi(\mathbf{x}_t)E(\mathbf{f}_t|\mathbf{x}_t)'\boldsymbol{\Lambda}'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2 O_P(\chi_N^{-1}),\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_4\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\mathbf{A}\frac{1}{TN}\sum_{t=1}^T\Phi(\mathbf{x}_t)\Phi(\mathbf{x}_t)'\mathbf{A}\mathbf{M}_\alpha'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&\leq \frac{1}{N^4T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2\|\mathbf{M}_\alpha'\widehat{\boldsymbol{\Lambda}}\|^2 O_P(\chi_N^{-2})\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_5\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\mathbf{A}\frac{1}{TN}\sum_{t=1}^T\Phi(\mathbf{x}_t)\mathbf{R}_t'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&= O_P(\chi_N^{-2}\frac{J}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2)\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_6\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\frac{1}{TN}\sum_{t=1}^T\mathbf{u}_s'\mathbf{R}_t E(\mathbf{f}_t|\mathbf{x}_t)'\boldsymbol{\Lambda}'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&\leq O_P(\frac{\chi_N^{-1}}{N^2T^2}\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2)\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_7\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\frac{1}{TN}\sum_{t=1}^T\mathbf{u}_s'\mathbf{R}_t\Phi(\mathbf{x}_t)'\mathbf{A}\mathbf{M}_\alpha'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&\leq O_P(\chi_N^{-2}\frac{J}{N^4T^2}\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2\|\mathbf{M}_\alpha'\widehat{\boldsymbol{\Lambda}}\|^2)\\
\frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\mathbf{B}_8\|^2 &= \frac{1}{N^2T}\sum_{s=1}^T\|\mathbf{u}_s'\frac{1}{TN}\sum_{t=1}^T\mathbf{R}_t\mathbf{R}_t'\widehat{\boldsymbol{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&\leq O_P(\chi_N^{-2})\frac{1}{N^2T^2}\sum_{s=1}^T\sum_{t=1}^T|\mathbf{u}_s'\mathbf{R}_t|^2\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2.
\end{aligned}
$$

By (C.3) and (C.4), $\chi_N^{-1}J\frac{1}{N^2}\|\mathbf{M}_\alpha'\widehat{\boldsymbol{\Lambda}}\|_F^2 + \chi_N^{-1}J\max_i\frac{1}{T}\sum_{t=1}^T R_{it}^2 = o_P(1)$.

Summarizing, we have

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\|\mathbf{D}_{t2}\|^2 \\
&\quad = O_P(\frac{\chi_N^{-1}}{N^3}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2+\frac{\chi_N^{-1}}{N}\max_i\frac{1}{T}\sum_{t=1}^{T}R_{it}^2 \\
&\qquad +\frac{J\chi_N^{-2}}{N^2T}\sum_{s=1}^{T}\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2\max_i\frac{1}{T}\sum_{t=1}^{T}R_{it}^2 \\
&\qquad +\frac{\chi_N^{-1}}{N^2T}\sum_{s=1}^{T}\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2 \\
&\qquad +\frac{\chi_N^{-2}}{N^4T}\sum_{s=1}^{T}\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2+\frac{\chi_N^{-1}}{N^2T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}|\mathbf{u}_s'\mathbf{R}_t|^2) \\
&\qquad +O_P(\frac{J\chi_N^{-2}}{N^4T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}|\mathbf{u}_s'\mathbf{R}_t|^2\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2) \\
&\qquad +O_P(1)\frac{\chi_N^{-2}}{N^2T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}|\mathbf{u}_s'\mathbf{R}_t|^2\max_i\frac{1}{T}\sum_{t=1}^{T}R_{it}^2 \\
&\quad \leq O_P(\chi_N^{-1})(\frac{1}{N^3}\|\mathbf{M}_\alpha'\widehat{\mathbf{\Lambda}}\|^2+\frac{1}{N}\max_i\frac{1}{T}\sum_{t=1}^{T}R_{it}^2 \\
&\qquad +\frac{1}{N^2T}\sum_{s=1}^{T}\|\mathbf{u}_s'\mathbf{M}_\alpha\|^2+\frac{1}{N^2T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}|\mathbf{u}_s'\mathbf{R}_t|^2).
\end{aligned}
$$

D Proof of Theorem 5.1

The proof of the limiting distribution of S under the null is divided into two major steps.

step 1: Asymptotic expansion: under H_0,

$$S=\frac{1}{TN}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}\mathbf{H}\widehat{\mathbf{W}}\mathbf{H}'\mathbf{\Lambda}'\mathbf{u}_t+o_P(T^{-1/2}).$$

step 2: The effect of estimating $\mathbf{\Sigma}_u$ is first-order negligible:

$$\frac{1}{TN}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}\mathbf{H}\widehat{\mathbf{W}}\mathbf{H}'\mathbf{\Lambda}'\mathbf{u}_t=\frac{1}{TN}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}(\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t+o_P(T^{-1/2}).$$

The result then follows from the asymptotic normality of the first term on the right hand side. We shall prove this using Lindeberg's central limit theorem.

We achieve each step in the following subsections.

D.1 Step 1 asymptotic expansion of S

Proposition D.1. *Under H_0,*

$$S = \frac{1}{TN} \sum_{t=1}^{T} \mathbf{u}_t' \mathbf{\Lambda} \mathbf{H} \widehat{\mathbf{W}} \mathbf{H}' \mathbf{\Lambda}' \mathbf{u}_t + o_P(T^{-1/2})$$

Proof. Since $\|\widehat{\mathbf{W}}\| \leq \max_i \widehat{\sigma}_{ii} = O_P(1)$, it follows from (B.6) that it suffices to prove under H_0, $\frac{N}{T}\sum_{t=1}^T \mathbf{D}_{ti}' \widehat{\mathbf{W}} \frac{1}{N} \mathbf{H}' \mathbf{\Lambda}' \mathbf{u}_t = o_P(T^{-1/2})$, and $\frac{N}{T}\sum_{t=1}^T \|\mathbf{D}_{ti}\|^2 = o_P(T^{-1/2}), i = 2, 3, 4.$

By the proof of Propositions B.2, C.3, Lemmas C.6, C.12 and that $\mathbf{D}_{t3} = \mathbf{C}_{t3}$,$\mathbf{D}_{t4} = \mathbf{C}_{t4}$,

$$\begin{aligned}
\frac{N}{T}\sum_{t=1}^T \|\mathbf{D}_{t4}\|^2 &= O_P(\max_i \frac{N}{T}\sum_{t=1}^T R_{it}^2) \\
&= O_P(NJ^{1-2\eta} + \frac{NJ^3 \log N}{\alpha_T^{2(\zeta_1 - 1)} T} + \frac{NJ^3 \log N \log J}{T^2}) \\
&= o_P(\frac{1}{\sqrt{T}}) \\
\frac{N}{T}\sum_{t=1}^T \|\mathbf{D}_{t3}\|^2 &= O_P(\frac{1}{N}\|\widehat{\mathbf{\Lambda}}' \mathbf{M}_\alpha\|^2) \\
&= O_P(\frac{J}{T} + \frac{NJ\alpha_T^{-\zeta_2}}{T} + J^{2-4\eta} + \alpha_T^{-2(\zeta_1-1)} \frac{J^3 \log N}{TJ^{2\eta-1}}) \\
&= o_P(\frac{1}{\sqrt{T}})
\end{aligned}$$

The last equality holds so long as $N\sqrt{T} = o(J^{2\eta-1})$, $NJ^4 \log N \log J = o(T^{3/2})$, $\zeta_1 > 2$.

By Lemma C.11,

$$\begin{aligned}
\frac{N}{T}\sum_{t=1}^T \|\mathbf{D}_{t2}\|^2 &= O_P(\frac{1}{N^2}\|\mathbf{M}_\alpha' \widehat{\mathbf{\Lambda}}\|^2 + \max_i \frac{1}{T}\sum_{t=1}^T R_{it}^2 + \frac{1}{NT}\sum_{s=1}^T \|\mathbf{u}_s' \mathbf{M}_\alpha\|^2 \\
&\quad + \frac{1}{NT^2}\sum_{s=1}^T \sum_{t=1}^T |\mathbf{u}_s' \mathbf{R}_t|^2) = o_P(\frac{1}{\sqrt{T}}).
\end{aligned}$$

The proof of $\frac{N}{T}\sum_{t=1}^T \mathbf{D}_{ti}' \widehat{\mathbf{W}} \frac{1}{N} \mathbf{H}' \mathbf{\Lambda}' \mathbf{u}_t = o_P(T^{-1/2})$ is given in Lemmas D.1 and D.2. It then leads to the desired result.

D.2 Step 2 Completion of the proof

We now aim to show $\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Sigma}}_u\widehat{\mathbf{\Lambda}}/N = \mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N + o_P(T^{-1/2})$. Once this is done, it then follows from the facts that $\mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N = O_P(1)$ and $(\mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N)^{-1} = O_P(1)$,

$$(\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Sigma}}_u\widehat{\mathbf{\Lambda}}/N)^{-1} = (\mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N)^{-1} + o_P(T^{-1/2}).$$

As a result, by Proposition D.1,

$$\begin{aligned} S &= \frac{1}{TN}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}\mathbf{H}(\mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N)^{-1}\mathbf{H}'\mathbf{\Lambda}'\mathbf{u}_t + o_P(T^{-1/2}) \\ &= \frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t + o_P(T^{-1/2}). \end{aligned}$$

Hence

$$\frac{TS - TK}{\sqrt{2TK}} = \frac{\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t - TK}{\sqrt{2TK}} + o_P(1) \to^d \mathcal{N}(0,1).$$

To finish the proof, we now show two claims:

(1)

$$\frac{\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t - TK}{\sqrt{2TK}} \to^d \mathcal{N}(0,1).$$

(2) $\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Sigma}}_u\widehat{\mathbf{\Lambda}}/N = \mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N + o_P(T^{-1/2})$.

Proof of (1) We define $X_t = \mathbf{u}_t'\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t$ and $s_T^2 = \sum_{t=1}^{T}\mathrm{var}(X_t)$. Then $E(X_t) = \mathrm{tr}\, E((\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{\Lambda}) = K$. Also by Assumption 4.1, $s_T^2/T \to 2K$, hence we have $E\frac{1}{T}\sum_{t=1}^{T}(X_t - K)^2 < \infty$ for all large N, T. For any $\epsilon > 0$, by the dominated convergence theorem, for all large N, T,

$$\frac{1}{T}\sum_{t=1}^{T}E(X_t - K)^2 1\{|X_t - K| > \epsilon s_T\} \leq \frac{1}{T}\sum_{t=1}^{T}E(X_t - K)^2 1\{|X_t - K| > \epsilon\sqrt{KT}\} = o(1).$$

This then implies the Lindeberg condition, $\frac{1}{s_T^2}\sum_{t=1}^{T}E(X_t - K)^2 1\{|X_t - K| > \epsilon s_T\} = o(1)$. Hence by the Lindeberg central limit theorem,

$$\frac{\sum_t X_t - TK}{s_T} \to^d \mathcal{N}(0,1).$$

The result then follows since $s_T^2/T \to 2K$.

Proof of (2) By the triangular inequality,

$$\begin{aligned}\|\frac{1}{N}\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Sigma}}_u\widehat{\mathbf{\Lambda}}-\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}\| &\leq \|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'(\widehat{\mathbf{\Sigma}}_u-\mathbf{\Sigma}_u)\widehat{\mathbf{\Lambda}}\| \\ &+\|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'\mathbf{\Sigma}_u(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| \\ &+\|\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'(\widehat{\mathbf{\Sigma}}_u-\mathbf{\Sigma}_u)(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| \\ &+\|\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'(\widehat{\mathbf{\Sigma}}_u-\mathbf{\Sigma}_u)\mathbf{\Lambda}\mathbf{H}\| \\ &+2\|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}\|.\end{aligned}$$

Using the established bounds for $\|\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H}\|$ in Theorem 3.1, it is straightforward to verify $\|\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'\mathbf{\Sigma}_u(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\| = o_P(T^{-1/2})$. Other terms require sharper bounds yet to be established. These are given in Proposition D.2. It then follows that $\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Sigma}}_u\widehat{\mathbf{\Lambda}}/N = \mathbf{H}'\mathbf{\Lambda}'\mathbf{\Sigma}_u\mathbf{\Lambda}\mathbf{H}/N+o_P(T^{-1/2})$. This completes the proof.

Q.E.D.

D.3 Technical Lemmas for Theorem 5.1

Lemma D.1. *Suppose* $(N+T)J^{1-2\eta}=o(1)$. *Then*

$$\mathrm{tr}(\frac{N}{T}\sum_{t=1}^T\mathbf{D}_{t2}'\widehat{\mathbf{W}}\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'\mathbf{u}_t)=o_P(T^{-1/2})$$

Proof. It suffices to prove $\|\frac{1}{T}\sum_{t=1}^T\mathbf{D}_{t2}\mathbf{u}_t'\mathbf{\Lambda}\|^2=\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\|^2=o_P(\frac{1}{T})$. To this end, we need to decompose $\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H}=\sum_{i=1}^8\mathbf{B}_i$ again as in (B.5). Every term can be bounded using established bounds except for the term involving $\mathbf{B}_3$. More specifically, for $i\neq 3$, we use $\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{B}_i\|^2\leq\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\|_F^2\|\mathbf{B}_i\|^2$. On the other hand,

$$\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\|_F^2\leq 2\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Sigma}_u\|_F^2+2\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'(\mathbf{u}_t\mathbf{u}_t'-\mathbf{\Sigma}_u)\|_F^2.$$

The first term is $O_P(\frac{1}{N})$. As for the second tern,

$$E\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'(\mathbf{u}_t\mathbf{u}_t'-\mathbf{\Sigma}_u)\|_F^2$$

$$
\begin{aligned}
&= \frac{1}{T^2N^2}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathrm{var}(\sum_{j=1}^{N}\lambda_{jk}(u_{jt}u_{it}-Eu_{jt}u_{it}))\\
&= \frac{1}{T^2N^2}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{l=1}^{N}\lambda_{jk}\lambda_{lk}\,\mathrm{cov}(u_{jt}u_{it},u_{lt}u_{it})\\
&= O(\frac{1}{T})+\frac{1}{T^2N^2}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{l\neq i,t}\lambda_{jk}\lambda_{lk}E(u_{jt}u_{it}-\sigma_{ij})u_{it}u_{lt}\\
&= O(\frac{1}{T}).
\end{aligned}
$$

Hence $\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{B}_i\|^2\leq O_P(\frac{1}{T}+\frac{1}{N})\|\mathbf{B}_i\|^2=o(\frac{1}{T})$, for $i\neq 3$, where the last equality holds by straightforward verifying $(\frac{T}{N}+1)\|\mathbf{B}_i\|^2=o(1)$ using Lemma C.5, assuming $(N+T)J^{1-2\eta}=o(1)$.

To allow $N/T\to\infty$, the term involving $\mathbf{B}_3$ requires a different and sharper bound:

$$
\begin{aligned}
\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{B}_3\|^2 &= \|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{M}_\alpha\frac{1}{TN}\sum_{s=1}^{T}\mathbf{A}\Phi(\mathbf{x}_s)E(\mathbf{f}_s|\mathbf{x}_s)'\mathbf{\Lambda}'\widehat{\mathbf{\Lambda}}\widetilde{\mathbf{V}}^{-1}\|^2\\
&\leq \|\frac{1}{TN}\sum_{t=1}^{T}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{M}_\alpha\|^2O_P(1)\\
&= O_P(1)\|\frac{1}{TN}\sum_{t}\mathbf{\Lambda}'\mathbf{u}_t\sum_{i=1}^{N}u_{it}\frac{1}{T}\sum_{s=1}^{T}\alpha_T\dot{\rho}(\alpha_T^{-1}e_{is})\Phi(\mathbf{x}_s)\|^2\\
&\leq O_P(1)\|\frac{1}{T^2N}\sum_{t}\sum_{s=1}^{T}\sum_{i=1}^{N}\mathbf{\Lambda}'\mathbf{u}_tu_{it}u_{is}\Phi(\mathbf{x}_s)\|^2\\
&\quad +O_P(1)\left(\frac{1}{T^2N}\sum_{t}\sum_{s=1}^{T}\sum_{i=1}^{N}\|\mathbf{\Lambda}'\mathbf{u}_t\||u_{it}||u_{is}|1\{|u_{is}|>\alpha_T\}\|\Phi(\mathbf{x}_s)\|\right)^2,
\end{aligned}
\tag{D.1}
$$

where we used the fact that under H_0, $e_{is}=u_{is}$. We respectively bound the two terms on the right hand side.

First term in (D.1) Note that

$$
E\|\frac{1}{T^2N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\mathbf{\Lambda}'\mathbf{u}_tu_{it}u_{is}\Phi(\mathbf{x}_s)\|^2=\frac{1}{T^4N^2}\sum_{l=1}^{J}\sum_{k=1}^{K}E(\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{jk}u_{jt}u_{it}u_{is}\phi_l(\mathbf{x}_s))^2.
$$

We then expand the term on the right hand side, which leads to many additive terms in the expansion. Using the assumption of serial independence to analyze each term, we conclude that

$$E\|\frac{1}{T^2N}\sum_{t=1}^T\sum_{s=1}^T\sum_{i=1}^N\mathbf{\Lambda}'\mathbf{u}_tu_{it}u_{is}\Phi(\mathbf{x}_s)\|^2=O_P(\frac{J}{TN}+\frac{J}{T^2}+\frac{JN}{T^3}).$$

We omit the lengthy details.

Second term in (D.1) As for the second term, first note that under H_0, $u_{it}=e_{it}$. So Lemma C.2 implies $(E|u_{is}|1\{|u_{is}|>\alpha_T\}|\mathbf{x}_t=\mathbf{x})\leq C\alpha_T^{-\zeta_2-1}$. On the other hand, by assumption, for some $C>0$, $\sup_{\mathbf{x}}E(u_{it}^41\{|u_{it}|>\alpha_T\}|\mathbf{x}_t=\mathbf{x})\leq\alpha_T^{-\zeta_5}C$, $E\|\mathbf{\Lambda}'\mathbf{u}_t\|^2=O(N)$. Hence

$$\begin{aligned}
&\frac{1}{T^2N}\sum_t\sum_{s=1}^T\sum_{i=1}^N E\|\mathbf{\Lambda}'\mathbf{u}_t\||u_{it}||u_{is}|1\{|u_{is}|>\alpha_T\}\|\Phi(\mathbf{x}_s)\|\\
=&\frac{1}{T^2N}\sum_t\sum_{i=1}^N E\|\mathbf{\Lambda}'\mathbf{u}_t\|u_{it}^21\{|u_{it}|>\alpha_T\}\|\Phi(\mathbf{x}_t)\|\\
&+\frac{1}{T^2N}\sum_t\sum_{s\neq t}\sum_{i=1}^N E\|\mathbf{\Lambda}'\mathbf{u}_t\||u_{it}|E|u_{is}|1\{|u_{is}|>\alpha_T\}\|\Phi(\mathbf{x}_s)\|\\
\leq&\frac{1}{T^2N}\sum_t\sum_{i=1}^N(E\|\mathbf{\Lambda}'\mathbf{u}_t\|^2)^{1/2}(E\|\Phi(\mathbf{x}_t)\|^2)^{1/2}\sup_{\mathbf{x}}(Eu_{it}^41\{|u_{it}|>\alpha_T\}|\mathbf{x}_t=\mathbf{x})^{1/2}\\
&+\frac{1}{T^2N}\sum_t\sum_{s\neq t}\sum_{i=1}^N(E\|\mathbf{\Lambda}'\mathbf{u}_t\|^2)^{1/2}(Eu_{it}^2)^{1/2}E\|\Phi(\mathbf{x}_s)\|\sup_{\mathbf{x}}(E|u_{is}|1\{|u_{is}|>\alpha_T\}|\mathbf{x}_t=\mathbf{x})\\
=&O_P(\frac{\sqrt{JN}}{T}\alpha_T^{-\zeta_5/2}+\sqrt{NJ}\alpha_T^{-\zeta_2-1}).
\end{aligned}$$

It then implies the second term in (D.1) is $O_P(\frac{JN}{T^2}\alpha_T^{-\zeta_5}+NJ\alpha_T^{-2\zeta_2-2})$.

Thus, when $\zeta_5\geq1$, $T=o(J^{2\eta-1})$

$$\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'\mathbf{B}_3\|^2=O_P(\frac{J}{TN}+\frac{J}{T^2}+\frac{JN}{T^3}+\frac{JN}{T^2}\alpha_T^{-\zeta_5}+NJ\alpha_T^{-2\zeta_2-2})=o_P(\frac{1}{T}).$$

As a result,

$$\|\frac{1}{T}\sum_{t=1}^T\mathbf{D}_{t2}\mathbf{u}_t'\mathbf{\Lambda}\|^2=\|\frac{1}{T}\sum_{t=1}^T\frac{1}{N}\mathbf{\Lambda}'\mathbf{u}_t\mathbf{u}_t'(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})\|^2=o_P(\frac{1}{T}).$$

Lemma D.2. *For $i=3,4$,*

$$\mathrm{tr}(\frac{N}{T}\sum_{t=1}^T\mathbf{D}_{ti}'\widehat{\mathbf{W}}\frac{1}{N}\mathbf{H}'\mathbf{\Lambda}'\mathbf{u}_t)=o_P(T^{-1/2})$$

Proof. Again, it suffices to verify $\|\frac{1}{T}\sum_{t=1}^T \mathbf{D}_{ti}\mathbf{u}_t'\mathbf{\Lambda}\|^2 = o_P(\frac{1}{T})$ for $i = 3, 4$. Note that $\|\frac{1}{T}\sum_{t=1}^T \Phi(\mathbf{x}_t)\mathbf{u}_t'\mathbf{\Lambda}\|^2 = O_P(\frac{NJ}{T})$. Then by definition,

$$\begin{aligned}\|\frac{1}{T}\sum_{t=1}^T \mathbf{D}_{t3}\mathbf{u}_t'\mathbf{\Lambda}\|^2 &= \|\frac{1}{T}\sum_{t=1}^T \frac{1}{N}\widehat{\mathbf{\Lambda}}'\mathbf{M}_\alpha\mathbf{A}\Phi(\mathbf{x}_t)\mathbf{u}_t'\mathbf{\Lambda}\|^2 \\ &\leq O_P(\frac{1}{N^2})\|\widehat{\mathbf{\Lambda}}'\mathbf{M}_\alpha\|^2\|\frac{1}{T}\sum_{t=1}^T \Phi(\mathbf{x}_t)\mathbf{u}_t'\mathbf{\Lambda}\|^2 = o_P(\frac{1}{T}).\end{aligned}$$

On the other hand, recall the definition $R_{it} := R_{1,it} + R_{2,it} + R_{3,it}$, where

$$\begin{aligned}R_{1,it} &:= \Phi(\mathbf{x}_t)'\mathbf{A}\frac{1}{T}\sum_{s=1}^T \alpha_T[\dot{\rho}(\alpha_T^{-1}e_{is,\alpha}) - \dot{\rho}(\alpha_T^{-1}e_{is})]\Phi(\mathbf{x}_s) \\ R_{2,it} &:= \Phi(\mathbf{x}_t)'(\mathbf{R}_{i,b} + \mathbf{b}_{i,\alpha} - \mathbf{b}_i), \qquad R_{3,it} := -z_{it}.\end{aligned}$$

Thus it can be verified similarly that

$$\|\frac{1}{T}\sum_{t=1}^T \mathbf{D}_{t4}\mathbf{u}_t'\mathbf{\Lambda}\|^2 = \|\frac{1}{T}\sum_{t=1}^T \frac{1}{N}\widehat{\mathbf{\Lambda}}'\mathbf{R}_t\mathbf{u}_t'\mathbf{\Lambda}\|^2 = O_P(\frac{1}{NT^2})\sum_{i=1}^N \|\sum_{t=1}^T R_{it}\mathbf{u}_t'\mathbf{\Lambda}\|^2 = o_P(\frac{1}{T}).$$

The verification is very similar as before, and is omitted here.

Proposition D.2. *(i)* $\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Sigma}_u(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}) = o_P(T^{-1/2})$*;*
(ii) $\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Sigma}}_u - \mathbf{\Sigma}_u)\mathbf{\Lambda} = o_P(T^{-1/2})$*;*
(iii) $\|\frac{1}{N}(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H})'(\widehat{\mathbf{\Sigma}}_u - \mathbf{\Sigma}_u)\mathbf{G}\| = o_P(T^{-1/2})$*, for either* $\mathbf{G} = \mathbf{\Lambda}$ *or* $\mathbf{G} = \widehat{\mathbf{\Lambda}}$*.*

Proof. Define $\widetilde{\mathbf{\Lambda}} = \mathbf{\Sigma}_u\mathbf{\Lambda}$. Note that we cannot simply bound these terms by $\frac{1}{N}\|\widetilde{\mathbf{\Lambda}}\|\|\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}\|$ or $\frac{1}{N}\|\mathbf{\Lambda}\|^2\|\widehat{\mathbf{\Sigma}}_u - \mathbf{\Sigma}_u\|$, as these bounds are too crude to achieve the desired rate of convergence when $N/T \to \infty$. More careful analysis is called for.

(i) Proving $\frac{1}{N}\widetilde{\mathbf{\Lambda}}'(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\mathbf{H}) = o_P(T^{-1/2})$ is exactly the same as that of Lemma C.7. Note that replacing $\mathbf{\Lambda}$ with $\widetilde{\mathbf{\Lambda}}$ does not introduce any complications as $\mathbf{\Sigma}_u$ is a diagonal matrix. Hence the proof is omitted here to avoid repetitions.

(ii) For any $k, l \leq K$, the (k, l) element of $\frac{1}{N}\mathbf{\Lambda}'(\widehat{\mathbf{\Sigma}}_u - \mathbf{\Sigma}_u)\mathbf{\Lambda}$ is given by

$$\frac{1}{N}\sum_{i=1}^N \lambda_{ik}\lambda_{il}(\widehat{\sigma}_{ii} - \sigma_{ii}) = \frac{1}{N}\frac{1}{T}\sum_t\sum_{i=1}^N \lambda_{ik}\lambda_{il}(u_{it}^2 - Eu_{it}^2) + \frac{1}{N}\sum_{i=1}^N \lambda_{ik}\lambda_{il}\frac{1}{T}\sum_t(\widehat{u}_{it}^2 - u_{it}^2)$$

As for the first term,

$$
\begin{aligned}
& E\Big|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(u_{it}^2-\sigma_{ii})\Big| \leq \left[E\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(u_{it}^2-\sigma_{ii})\right)^2\right]^{1/2} \\
= & \left[\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{s=1}^{T}\lambda_{ik}\lambda_{il}\lambda_{jk}\lambda_{jl}\operatorname{cov}(u_{it}^2,u_{js}^2)\right]^{1/2} \\
= & \left[\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}^2\lambda_{il}^2\operatorname{var}(u_{it}^2)\right]^{1/2} = o(\frac{1}{\sqrt{T}}).
\end{aligned}
$$

As for the second term, we have

$$
\begin{aligned}
& \Big|\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(\widehat{u}_{it}^2-u_{it}^2)\Big| \leq 2\Big|\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(\widehat{u}_{it}-u_{it})u_{it}\Big| \\
& +\Big|\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(\widehat{u}_{it}-u_{it})^2\Big| \\
\leq & O_P(1)(\frac{1}{T}\sum_{t}\|\widehat{\mathbf{f}}_t-\mathbf{H}^{-1}\mathbf{f}_t\|^2)^{1/2}(\frac{1}{T}\sum_{t}\|\frac{1}{N}\sum_{i=1}^{N}\lambda_{ik}\lambda_{il}u_{it}\boldsymbol{\lambda}_i\|^2)^{1/2} \\
& +O_P(1)(\frac{1}{N}\sum_{i}\|\lambda_{ik}\lambda_{il}(\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}'\boldsymbol{\lambda}_i)\|^2)^{1/2}(\frac{1}{N}\sum_{i}\|\frac{1}{T}\sum_{t}u_{it}\mathbf{f}_t\|^2)^{1/2} \\
& +O(1)\max_i\frac{1}{T}\sum_{t=1}^{T}(\widehat{u}_{it}-u_{it})^2 \\
& +O_P(1)(\frac{1}{T}\sum_{t}\|\widehat{\mathbf{f}}_t-\mathbf{H}^{-1}\mathbf{f}_t\|^2)^{1/2}(\frac{1}{TN}\sum_{it}u_{it}^2)^{1/2}(\frac{1}{N}\sum_{i}\|\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}'\boldsymbol{\lambda}_i\|^2)^{1/2}.
\end{aligned}
$$

Note that $\frac{1}{T}\sum_t\|\widehat{\mathbf{f}}_t-\mathbf{H}^{-1}\mathbf{f}_t\|^2 = O_P\left(\psi_{NT}^2\right)$, $\max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2 = O_P(\psi_{NT}^2+\frac{J\log N}{T})$ by Lemma D.3. Also, $\frac{1}{N}\sum_{i=1}^{N}\|\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}'\boldsymbol{\lambda}_i\|^2 = O_P\left(\frac{J}{T}+\frac{1}{J^{2\eta-1}}+(\frac{\log N}{T})^{\zeta_1}J^3\right)$ by Theorem 4.1. In addition,

$$
\begin{aligned}
E\frac{1}{T}\sum_{t}\|\frac{1}{N}\sum_{i=1}^{N}\lambda_{ik}\lambda_{il}u_{it}\boldsymbol{\lambda}_i\|^2 &= \sum_{m=1}^{K}E(\frac{1}{N}\sum_{i}\lambda_{ik}\lambda_{il}\lambda_{im}u_{it})^2 \\
&= \sum_{m=1}^{K}\frac{1}{N^2}\sum_{i}\lambda_{ik}^2\lambda_{il}^2\lambda_{im}^2Eu_{it}^2 = O(\frac{1}{N}), \\
E\frac{1}{N}\sum_{i}\|\frac{1}{T}\sum_{t}u_{it}\mathbf{f}_t\|^2 &= \frac{1}{N}\sum_{i}\sum_{k}\frac{1}{T^2}\sum_{t}Eu_{it}^2f_{kt}^2 = O(\frac{1}{T}).
\end{aligned}
$$

Hence it is straightforward to verify that $|\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ik}\lambda_{il}(\widehat{u}_{it}^2-u_{it}^2)| = o_P(T^{-1/2})$ so long as $T=o(N^2)$, $T=o(J^{2\eta-1}N)$, $J^4\log N = o(NT)$.

(iii) Let G_{ik} denote the (i,k) element of $\mathbf{G}$, and let δ_{ik} denote the (i,k) element of $\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H}$. Since $\max_i \|\widehat{\boldsymbol{\lambda}}_i-\boldsymbol{\lambda}_i\|=o_P(1)$, we have $\max_{ik}|G_{ik}|=O_P(1)$, regardless of $\mathbf{G}=\mathbf{\Lambda}$ or $\mathbf{G}=\widehat{\mathbf{\Lambda}}$. Then the (l,k) element of the $K\times K$ matrix $\frac{1}{N}(\widehat{\mathbf{\Lambda}}-\mathbf{\Lambda}\mathbf{H})'(\widehat{\mathbf{\Sigma}}_u-\mathbf{\Sigma}_u)\mathbf{G}$ is bounded by

$$|\frac{1}{N}\sum_{i=1}^N \delta_{il}G_{ik}\frac{1}{T}\sum_t(\widehat{u}_{it}^2-\sigma_{ii})|\leq \max_{ilk}|\delta_{il}G_{ik}|\frac{1}{NT}\sum_{i=1}^N|\sum_{t=1}^T(\widehat{u}_{it}^2-u_{it}^2)+(u_{it}^2-\sigma_{ii})|.$$

On one hand, by Lemma D.3,

$$\begin{aligned}\max_{ilk}|\delta_{il}G_{ik}|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T|\widehat{u}_{it}^2-u_{it}^2| &= O_P(\psi_{NT}+\sqrt{\frac{J\log N}{T}})\max_i\|\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}'\boldsymbol{\lambda}_i\| \\ &= o_P(\frac{1}{\sqrt{T}}).\end{aligned}$$

On the other hand,

$$E|\sum_{t=1}^T(u_{it}^2-\sigma_{ii})|\leq \mathrm{var}(\sum_{t=1}^T(u_{it}^2-\sigma_{ii}))^{1/2}=O(T^{1/2}).$$

Hence

$$\max_{ilk}|\delta_{il}G_{ik}|\frac{1}{NT}\sum_{i=1}^N|\sum_{t=1}^T(u_{it}^2-\sigma_{ii})|=O_P(\frac{1}{\sqrt{T}})\max_i\|\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}'\boldsymbol{\lambda}_i\|=o_P(\frac{1}{\sqrt{T}}).$$

Lemma D.3. *Define*

$$\psi_{NT}=\frac{1}{J^{\eta-1/2}}+\frac{1}{\sqrt{N}}+\frac{J^2(\log N\log J)^{1/2}}{T}+(\frac{\log N}{T})^{\zeta_1/2}J^2.$$

Under H_0, when $N=O(T^2)$,

(i) $\frac{1}{T}\sum_{t=1}^T\|\widehat{\mathbf{f}}_t-\mathbf{H}^{-1}\mathbf{f}_t\|^2=O_P\left(\psi_{NT}^2\right)$.

(ii) $\max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2=O_P(\psi_{NT}^2+\frac{J\log N}{T})$.

(iii) $\frac{1}{NT}\sum_i\sum_t|\widehat{u}_{it}^2-u_{it}^2|=O_P(\psi_{NT}+\sqrt{\frac{J\log N}{T}})$.

Proof. (i) By Theorem 3.2, under H_0,

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^T\|\widehat{\mathbf{f}}_t-\mathbf{H}^{-1}\mathbf{f}_t\|^2 &\leq 2\frac{1}{T}\sum_{t=1}^T\|\widehat{\mathbf{g}}(\mathbf{x}_t)-\mathbf{H}^{-1}\mathbf{g}(\mathbf{x}_t)\|^2+2\frac{1}{T}\sum_{t=1}^T\|\widehat{\boldsymbol{\gamma}}_t-\mathbf{H}^{-1}\boldsymbol{\gamma}_t\|^2 \\ &= O_P\left(\frac{1}{N}+\frac{J^4\log N\log J}{T^2}+\frac{1}{J^{2\eta-1}}+(\frac{\log N}{T})^{\zeta_1}J^4\right).\end{aligned}$$

(D.2)

(ii) Uniformly in i, by Theorem 3.1,

$$\begin{aligned}\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2 &\leq C\frac{1}{T}\sum_t\|\widehat{\boldsymbol{\lambda}}_i-\boldsymbol{\lambda}_i\|^2\|\widehat{\mathbf{f}}_t\|^2+C\frac{1}{T}\sum_t\|\boldsymbol{\lambda}_i\|^2\|\widehat{\mathbf{f}}_t-\mathbf{f}_t\|^2\\ &= O_P(\psi_{NT}^2).\end{aligned}$$

(iii) We have, using $|a^2-b^2|\leq|a-b||a+b|$ and the Cauchy-Schwarz inequality,

$$\begin{aligned}(\frac{1}{NT}\sum_i\sum_t|\widehat{u}_{it}^2-u_{it}^2|)^2 &\leq \max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2\frac{1}{NT}\sum_{it}[2(\widehat{u}_{it}-u_{it})^2+4u_{it}^2]\\ &\leq 2(\max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2)^2\\ &\quad+4\max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2\frac{1}{NT}\sum_{it}u_{it}^2\\ &= O_P(\max_i\frac{1}{T}\sum_t(\widehat{u}_{it}-u_{it})^2)\\ &= O_P(\psi_{NT}^2+\frac{J\log N}{T}).\end{aligned}$$

References

Ahn, S. and Horenstein, A. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81** 1203–1227.

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

Bühlmann, P. and van de Geer, S. (2011). *Statistics for high-dimensional data, methods, theory and applications*. The first edition ed. Springer, New York.

Dimatteo, I., Genovese, C. and Kass, R. (2001). Bayesian curve fitting with free-knot splines. *Biometrika* **88** 1055–1071.

Lam, C. and Yao, Q. (2012). Factor modeling for high dimensional time-series: inference for the number of factors. *Annals of Statistics* **40** 694–726.