ONLINE SUPPLEMENT TO "INFERENCE FOR HETEROGENEOUS EFFECTS USING LOW-RANK ESTIMATION OF FACTOR SLOPES"

VICTOR CHERNOZHUKOV, CHRISTIAN HANSEN, YUAN LIAO, AND YINCHU ZHU

Abstract. This document contains the estimation algorithm for the multivariate case, additional simulations and all technical proofs.

Contents

Appendix A. Estimation Algorithm for the Multivariate Case 2
Appendix B. Additional Simulations 4
B.1. Policy relevant tests 4
B.2. Simulations for Dynamic Model 5
Appendix C. Proof of Proposition 2.1: the algorithm convergence 7
Appendix D. Nuclear-Norm Matrix Estimators 11
D.1. Level of the Score 12
D.2. Useful Claims 12
D.3. The RSC condition 14
D.4. Proof of Proposition D.1: convergence of $\widetilde{\Theta}_S, \widetilde{M}_S$ 18
D.5. Proof of Proposition D.1: convergence of $\widetilde{\Lambda}_S, \widetilde{A}_S$ 20
Appendix E. Proof of Theorem 3.1 22
E.1. Behavior of $\widehat{f}_t$ 22
E.2. Behavior of $\widehat{\lambda}_i$ 25
E.3. Proof of normality of $\widehat{\theta}_{it}$ 27
E.4. Two special cases 28
E.5. Covariance estimations 29
Appendix F. Policy Relevant Tests 31
F.1. Group effect dynamics of two periods 31
F.2. Test of group homogeneity 33
F.3. Test of joint significance in a given time period 35
Appendix G. Technical lemmas 36
G.1. The effect of $\widehat{e}_{it} - e_{it}$ in the factor model 36
G.2. Behavior of the preliminary 46
G.3. Technical lemmas for $\widehat{f}_t$ 61
G.4. Technical lemmas for $\widehat{\lambda}_i$ 73
G.5. Technical lemmas for covariance estimations 90

References 94

Appendix A. Estimation Algorithm for the Multivariate Case

Consider

$$\begin{aligned} y_{it} &= \sum_{r=1}^{R} x_{it,r}\theta_{it,r} + \alpha_i' g_t + u_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T. \\ \theta_{it,r} &= \lambda_{i,r}' f_{t,r}. \end{aligned}$$

where $(x_{it,1}, ..., x_{it,R})'$ is an R-dimensional vector of covariate. Each coefficient $\theta_{it,r}$ admits a factor structure with $\lambda_{i,r}$ and $f_{t,r}$ as the "loadings" and "factors"; the factors and loadings are $\theta_{it,r}$ specific, but are allowed to have overlap. Here $(R, \dim(\lambda_{i,1}), ..., \dim(\lambda_{i,R}))$ are all assumed fixed.

Suppose $x_{it,r} = \mu_{it,r} + e_{it,r}$. For instance, $\mu_{it,r} = l_{i,r}' w_{t,r}$ also admits a factor structure. Then after partialing out $\mu_{it,r}$, the model can also be written as:

$$\dot{y}_{it} = \sum_{r=1}^{R} e_{it,r}\lambda_{i,r}' f_{t,r} + \alpha_i' g_t + u_{it}, \quad \dot{y}_{it} = y_{it} - \sum_{r=1}^{R} \mu_{it,r}\theta_{it,r}$$

As such, it is straightforward to extend the estimation algorithm to the multivariate case. Let X_r be the $N \times T$ matrix of $x_{it,r}$, Let Θ_r be the $N \times T$ matrix of $\theta_{it,r}$. We first estimate these low rank matrices using penalized nuclear-norm regression. We then apply sample splitting, and employ steps 2-4 to iteratively estimate $(f_{t,r}, \lambda_{i,r})$. The formal algorithm is stated as follows.

Algorithm A.1. Estimate $\theta_{it,r}$ as follows.

Step 1. Estimate the number of factors. Run nuclear-norm penalized regression:

$$(\widetilde{M},\widetilde{\Theta}_r) := \arg\min_{M,\Theta_r} \|Y - M - \sum_{r=1}^{R} X_r \odot \Theta_r\|_F^2 + \nu_0\|M\|_n + \sum_{r=1}^{R}\nu_r\|\Theta\|_n.$$

Estimate $K_r = \dim(\lambda_{i,r}), K_0 = \dim(\alpha_i)$ by

$$\widehat{K}_r = \sum_i 1\{\psi_i(\widetilde{\Theta}_r) \geq (\nu_r\|\widetilde{\Theta}_r\|)^{1/2}\}, \quad \widehat{K}_0 = \sum_i 1\{\psi_i(\widetilde{M}) \geq (\nu_0\|\widetilde{M}\|)^{1/2}\}.$$

Step 2. Estimate the structure $x_{it,r} = \mu_{it,r} + e_{it,r}$.

In the factor model, use the PC estimator to obtain $(\widehat{l'_{i,r}w_{t,r}}, \widehat{e}_{it,r})$ for all $i = 1,...,N, t = 1,..,T$ and $r = 1,...,R$.

Step 3: Sample splitting. Randomly split the sample into $\{1,...,T\}/\{t\} = I \cup I^c$, so that $|I|_0 = [(T-1)/2]$. Denote by Y_I, $X_{I,r}$ as the $N \times |I|_0$ matrices of $(y_{is}, x_{is,r})$ for observations at $s \in I$. Estimate the low-rank matrices Θ and M as in step 1, with (Y, X_r) replaced with $(Y_I, X_{I,r})$, and obtain $(\widetilde{M}_I, \widetilde{\Theta}_{I,r})$.

Let $\widetilde{\Lambda}_{I,r} = (\widetilde{\lambda}_{1,r}, ..., \widetilde{\lambda}_{N,r})'$ be the $N \times \widehat{K}_r$ matrix, whose columns are defined as $\sqrt{N}$ times the first $\widehat{K}_r$ eigenvectors of $\widetilde{\Theta}_{I,r}\widetilde{\Theta}'_{I,r}$. Let $\widetilde{A}_I = (\widetilde{\alpha}_1, ..., \widetilde{\alpha}_N)'$ be the $N \times \widehat{K}_0$ matrix, whose columns are defined as $\sqrt{N}$ times the first $\widehat{K}_0$ eigenvectors of $\widetilde{M}_I\widetilde{M}'_I$.

Step 4. Estimate the "partial-out" components.

Substitute in $(\widetilde{\alpha}_i, \widetilde{\lambda}_{i,r})$, and define

$$(\widetilde{f}_{s,r}, \widetilde{g}_s) := \arg\min_{f_{s,r},g_s} \sum_{i=1}^{N}(y_{is} - \widetilde{\alpha}'_i g_s - \sum_{r=1}^{R} x_{is,r}\widetilde{\lambda}'_{i,r}f_{s,r})^2, \quad s \in I^c \cup \{t\}.$$

and

$$(\dot{\lambda}_{i,r}, \dot{\alpha}_i) = \arg\min_{\lambda_{i,r},\alpha_i} \sum_{s \in I^c\cup\{t\}} (y_{is} - \alpha'_i\widetilde{g}_s - \sum_{r=1}^{R} x_{is,r}\lambda'_{i,r}\widetilde{f}_{s,r})^2, \quad i = 1,...,N.$$

Step 5. Estimate $(f_{t,r}, \lambda_{i,r})$ for inferences.

For all $s \in I^c \cup \{t\}$, let

$$(\widehat{f}_{I,s,r}, \widehat{g}_{I,s}) := \arg\min_{f_{s,r},g_s} \sum_{i=1}^{N}(\widehat{y}_{is} - \widetilde{\alpha}'_i g_s - \sum_{r=1}^{R} \widehat{e}_{is,r}\widetilde{\lambda}'_{i,r}f_{s,r})^2.$$

Fix $i \leq N$,

$$(\widehat{\lambda}_{I,i,r}, \widehat{\alpha}_{I,i}) = \arg\min_{\lambda_{i,r},\alpha_i} \sum_{s\in I^c\cup\{t\}} (\widehat{y}_{is} - \alpha_i'\widehat{g}_{I,s} - \sum_{r=1}^{R} \widehat{e}_{is,r}\lambda_{i,r}'\widehat{f}_{I,s,r})^2.$$

where $\widehat{y}_{is} = y_{is} - \widehat{l_{i,r}'w_{s,r}}\dot{\lambda}_{i,r}'\widetilde{f}_{s,r}$, $\widehat{e}_{is,r} = x_{is,r} - \widehat{l_{i,r}'w_{s,r}}$ in the factor model.

Step 6. Estimate $\theta_{it,r}$. Repeat steps 3-5 with I and I^c exchanged, and obtain $(\widehat{\lambda}_{I^c,i,r}, \widehat{f}_{I^c,s,r} : s \in I \cup \{t\}, i \leq N, r \leq R)$. Define

$$\widehat{\theta}_{it,r} := \frac{1}{2}[\widehat{\lambda}_{I,i,r}'\widehat{f}_{I,t,r} + \widehat{\lambda}_{I^c,i,r}'\widehat{f}_{I^c,t,r}].$$

The asymptotic variance can be estimated by $\widehat{v}_{\lambda,r} + \widehat{v}_{f,r}$, where

$$\begin{aligned}
\widehat{v}_{\lambda,r} &= \frac{1}{2N}(\widehat{\lambda}_{I,r,\mathcal{G}}'\widehat{V}_{\lambda,1}^{I-1}\widehat{V}_{\lambda,2}^{I}\widehat{V}_{\lambda,1}^{I-1}\widehat{\lambda}_{I,r,\mathcal{G}} + \widehat{\lambda}_{I^c,r,\mathcal{G}}'\widehat{V}_{\lambda,1}^{I^c-1}\widehat{V}_{\lambda,2}^{I^c}\widehat{V}_{\lambda,1}^{I^c-1}\widehat{\lambda}_{I^c,r,\mathcal{G}}) \\
\widehat{v}_{f} &= \frac{1}{2T|\mathcal{G}|_0}(\widehat{f}_{I,t,r}'\widehat{V}_{I,f}\widehat{f}_{I,t,r} + \widehat{f}_{I^c,t,r}'\widehat{V}_{I^c,f}\widehat{f}_{I^c,t,r}) \\
\widehat{V}_{S,f} &= \frac{1}{|\mathcal{G}|_0|S|_0}\sum_{s\notin S}\sum_{i\in\mathcal{G}}\widehat{\Omega}_{S,i,r}\widehat{f}_{S,s,r}\widehat{f}_{S,s,r}'\widehat{\Omega}_{S,i,r}\widehat{e}_{is,r}^2\widehat{u}_{is}^2 \\
\widehat{V}_{\lambda,1}^{S} &= \frac{1}{N}\sum_j \widehat{\lambda}_{S,r,j}\widehat{\lambda}_{S,r,j}'\widehat{e}_{jt,r}^2 \\
\widehat{V}_{\lambda,2}^{S} &= \frac{1}{N}\sum_j \widehat{\lambda}_{S,r,j}\widehat{\lambda}_{S,r,j}'\widehat{e}_{jt,r}^2\widehat{u}_{jt}^2 \\
\widehat{\lambda}_{S,r,\mathcal{G}} &= \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\widehat{\lambda}_{S,r,i}, \quad \widehat{\Omega}_{S,i,r} = (\frac{1}{|S|_0}\sum_{s\in S}\widehat{f}_{S,s,r}\widehat{f}_{S,s,r}')^{-1}(\frac{1}{T}\sum_{s=1}^{T}\widehat{e}_{is,r}^2)^{-1}
\end{aligned}$$

It is also straightforward to extend the univariate asymptotic analysis to the multivariate case, and establish the asymptotic normality for $\widehat{\theta}_{it,r}$. The proof techniques are the same, subjected to more complicated notation. Therefore our proofs below focus on the univariate case.

Appendix B. Additional Simulations

B.1. **Policy relevant tests.** We run simulations for testing the group homogeneous effect:

$$H_0^1 : \bar{\theta}_{\mathcal{G}_1,t} = ... = \bar{\theta}_{\mathcal{G}_J,t}$$

for a given time t. We generate outcomes as

$$y_{it} = \alpha_i'g_t + x_{it,1}\theta_{it} + x_{it,2}\beta_{it} + u_{it}$$

where $\theta_{it} = \lambda_{i,1}' f_{t,1}$ and $\beta_{it} = \lambda_{i,2}' f_{t,2}$, and $\bar{\theta}_{\mathcal{G}_j,t}$ denotes the average of θ_{it} in group $\mathcal{G}_j$. The DGP is generated similar to that of the simulation section of the main text, except that we impose the within-group homogeneity assumption. That is, let $\bar{\lambda}_{\mathcal{G}_j}$, $j = 1, ..., J$ be predetermined values, we set

$$\lambda_{i,1} = \bar{\lambda}_{\mathcal{G}_j}, \text{ for all } i \in \mathcal{G}_j$$

These predetermined values are set as

$$\bar{\lambda}_{\mathcal{G}_j} = \begin{cases} 1 & \text{under the null} \\ 1 + Z_j N^{-a} & \text{under the alternative} \end{cases}, \quad j = 1, ..., J$$

where $Z_j \in \mathcal{N}(0,1)$. The parameter $a > 0$ measures the deviation from the null under local alternatives. We set $J = 10$ and each group contains N/J elements.

The second test we conducted is the test of joint significance:

$$H_0^2 : \theta_{i,t_0} = 0, \forall i.$$

As discussed, the test is equivalent to testing $f_{t_0,1} = 0$ at a given t_0, and in this simulation, $t_0 = 1$. The DGP is still similar as before, except that we impose null and local alternatives on $f_{t_0,1}$:

$$f_{t_0,1} = \begin{cases} 0 & \text{under the null} \\ N^{-a} & \text{under the alternative} \end{cases}.$$

We use the chi-square tests as presented in Theorem 3.2. Table B.1 reports the frequencies of rejection out of 1000 simulations with nominal level 0.05. The results show desired size and reasonable power for both tests.

B.2. **Simulations for Dynamic Model.** We now consider an example with a lagged dependent variable to allow for dynamics. Specifically, we consider the model

$$y_{it} = \alpha_i' g_t + x_{it}\lambda_{i,1}' f_{t,1} + y_{i,t-1}\lambda_{i,2}' f_{t,2} + u_{it}, \tag{B.1}$$

where $x_{it} = l_i' w_t + e_{it}$. Our asymptotic theory does not allow the lagged variable $y_{i,t-1}$ because the presence of this variable is incompatible with an exact representation within a static factor model structure. Nevertheless, as this dynamic model is interesting, we investigate the finite sample performance of our method in this case with the understanding that it is not covered by our formal theory.

TABLE B.1. Frequencies of Rejection in Policy Tests

N	T	Null and local alternatives			
			Test of Group Homogeneity		
		null	$N^{-3/4}$	$N^{-1/2}$	$N^{-1/4}$
100	100	0.047	0.230	0.990	1.000
200	200	0.039	0.790	1.000	1.000
			Test of Joint Significance		
		null	$N^{-1/2}$	$N^{-1/4}$	1
100	100	0.054	0.144	0.874	1.000
200	200	0.057	0.214	0.940	1.000

Note: This table reports the rejection frequencies out of 1000 simulations for the two policy tests. For the group homogeneity test, the local alternatives are $\bar{\lambda}_{\mathcal{G}_j} = 1 + \mathcal{N}(0, N^{-2a})$ with $a \in \{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}\}$. For the join significance test, the local alternatives are $f_t = N^{-a}$ with $a \in \{\frac{1}{2}, \frac{1}{4}, 0\}$.

We calibrate the parameters of the data generating process for this simulation by estimating (B.1) using the data from Acemoglu et al. (2008). Here y_{it} is the democracy score of country i in period t, and x_{it} is the GDP per capita over the same period. From the data, we estimate that all low-rank matrices have two factors with the exception of $\alpha_i' g_t$, which is estimated to have one factor. We then generate all factors and loadings from normal distributions with means and covariances matching with the sample means and covariances of the estimated factors and loadings obtained from the actual democracy-income data. We generate initial conditions y_{i1} independently from $0.3\mathcal{N}(0,1) + 0.497$, whose parameters are calibrated from the real data at time $t = 1$. u_{it} is generated independently from a $\mathcal{N}(0, \sigma^2)$ with $\sigma = 0.1287$ calibrated from the real data. Finally, y_{it} is generated iteratively according to the model using the draws of initial conditions.

Let $z_{it} := y_{i,t-1}$. Figure B.2 plots the first twenty eigenvalues of the $N \times N$ sample covariance of z_{it} when $N = T = 100$, averaged over 100 repetitions. The plot reveals one very spiked eigenvalue. Therefore, we estimate a one-factor model for z_{it} in our estimation procedure.

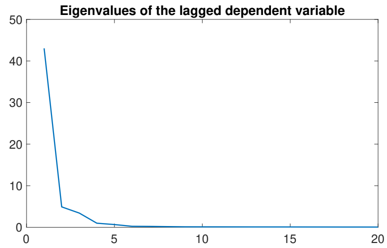

FIGURE B.1. Sorted eigenvalues of the sample covariance of the lagged dependent variable, averaged over 100 repetitions.

In this design, we only report results from applying the full multi-step procedure including sample-splitting and partialing-out. We report results for standardized estimates using the estimated standard error for $t = i = 1$ and $r \in \{1, 2\}$ for $\theta_{r,it} = \lambda_{i,r}' f_{t,r}$ and $\widehat{v}_{r,\lambda} + \widehat{v}_{r,f}$ a corresponding estimate of the asymptotic variance. All results are based on 1000 simulation replications.

Table B.2 reports the sample means and standard deviations of the t-statistics as well as coverage probabilities of feasible 95% confidence intervals, and we report the histograms of the standardized estimates in Figure B.2. Looking at the histograms, we can see that the bias in the estimated effects of x_{it} is relatively small while there is a noticeable downward bias in the estimated effect of $y_{i,t-1}$. We also see from the numeric results that the standard deviation of the standardized effect of x_{it} appears to be systematically too large, i.e. the estimated standard errors used for standardization are too small, which translates into some distortions in coverage probability. Coverage of the effect of lagged y seems reasonably accurate despite the large estimated bias. Importantly, we see that performance improves on all dimensions as $N = T$ becomes larger.

APPENDIX C. PROOF OF PROPOSITION 2.1: THE ALGORITHM CONVERGENCE

Proof. Recall that $\Theta_{k+1} = S_{\tau\nu_1/2}(\Theta_k - \tau A_k)$, where

$$A_k = X \odot (X \odot \Theta_k - Y + M_k).$$

By Lemma C.2, set $\Theta = \widetilde{\Theta}$ and $M = \widetilde{M}$, and replace k with subscript m,

$$F(\widehat{\Theta}, \widehat{M}) - F(\Theta_{m+1}, M_{m+1}) \geq \frac{1}{\tau}\left(\|\Theta_{m+1} - \widehat{\Theta}\|_F^2 - \|\Theta_m - \widehat{\Theta}\|_F^2\right).$$

TABLE B.2. Estimation and inference results in dynamic model simulation

$N = T$	mean		standard deviation		coverage probability	
	x_{it}	$y_{i,t-1}$	x_{it}	$y_{i,t-1}$	x_{it}	$y_{i,t-1}$
50	0.019	-0.345	1.342	1.140	0.856	0.919
100	-0.096	-0.267	1.178	1.015	0.920	0.932
150	-0.017	-0.276	1.117	0.942	0.926	0.944

Note: This table reports the simulation mean, standard deviation, and coverage probability of 95% confidence intervals for the estimated effect of x_{it} and $y_{i,t-1}$ for $i = t = 1$. Simulation mean and standard deviation are for estimates centered around the true parameter values and normalized in each replication by the estimated standard error. Thus, ideally, the entries in the mean panel would be 0 and the entries in the standard deviation panel would be 1. Results are based on 1000 simulation replications.

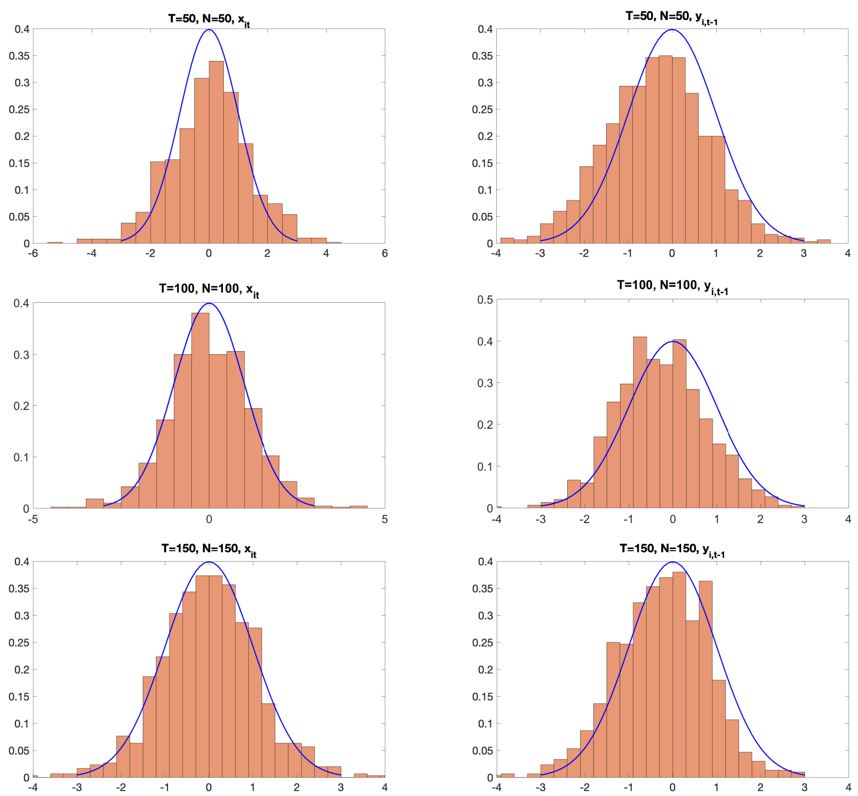

FIGURE B.2. Histograms of standardized estimates in dynamic models ($\widehat{\theta}_{11} - \theta_{11}$ divided by the estimated asymptotic standard deviation). The left three plots are the effect of $x_{i,t}$; the right three plots are for the effect of $y_{i,t-1}$. The standard normal density function is superimposed on the histograms.

Let $m = 1, ..., k$, and sum these inequalities up, since $F(\Theta_{m+1}, M_{m+1}) \geq F(\Theta_{k+1}, M_{k+1})$ by Lemma C.1,

$$\begin{aligned} & kF(\widehat{\Theta}, \widehat{M}) - kF(\Theta_{k+1}, M_{k+1}) \geq kF(\widehat{\Theta}, \widehat{M}) - \sum_{m=1}^{k} F(\Theta_{m+1}, M_{m+1}) \\ \geq\ & \frac{1}{\tau}\left(\|\Theta_{k+1} - \widehat{\Theta}\|_F^2 - \|\Theta_1 - \widehat{\Theta}\|_F^2\right) \geq -\frac{1}{\tau}\|\Theta_1 - \widehat{\Theta}\|_F^2. \end{aligned}$$

□

The above proof depends on the following lemmas.

Lemma C.1. *We have: (i)*

$$\begin{aligned} \Theta_{k+1} &= \arg\min_{\Theta} p(\Theta, \Theta_k, M_k) + \nu_1\|\Theta\|_n, \\ p(\Theta, \Theta_k, M_k) &:= \tau^{-1}\|\Theta_k - \Theta\|_F^2 - 2\mathsf{tr}\left((\Theta_k - \Theta)' A_k\right). \end{aligned}$$

(ii) For any $\tau \in (0, 1/\max x_{it}^2)$,

$$F(\Theta, M_k) \leq p(\Theta, \Theta_k, M_k) + \nu_1\|\Theta\|_n + \nu_0\|M_k\|_n + \|Y - M_k - X \odot \Theta_k\|_F^2.$$

(iii) $F(\Theta_{k+1}, M_{k+1}) \leq F(\Theta_{k+1}, M_k) \leq F(\Theta_k, M_k)$.

Proof. (i) We have $\|\Theta_k - \tau A_k - \Theta\|_F^2 = \|\Theta_k - \Theta\|_F^2 + \tau^2\|A_k\|_F^2 - 2\mathsf{tr}[(\Theta_k - \Theta)' A_k]\tau$. So

$$\begin{aligned} & \arg\min_{\Theta} \|\Theta_k - \tau A_k - \Theta\|_F^2 + \tau\nu_1\|\Theta\|_n \\ =\ & \arg\min_{\Theta} \|\Theta_k - \Theta\|_F^2 - 2\mathsf{tr}[(\Theta_k - \Theta)' A_k]\tau + \tau\nu_1\|\Theta\|_n \\ =\ & \arg\min_{\Theta} \tau \cdot p(\Theta, \Theta_k, M_k) + \tau\nu_1\|\Theta\|_n. \end{aligned}$$

On the other hand, it is well known that $\Theta_{k+1} = S_{\tau\nu_1/2}(\Theta_k - \tau A_k)$ is the solution to the first problem in the above equalities (Ma et al., 2011). This proves (i).

(ii) Note that for $\Theta_k = (\theta_{k,it})$ and $\Theta = (\theta_{it})$, and any $\tau^{-1} > \max_{it} x_{it}^2$, we have

$$\|X \odot (\Theta_k - \Theta)\|_F^2 = \sum_{it} x_{it}^2(\theta_{k,it} - \theta_{it})^2 < \tau^{-1}\|\Theta_k - \Theta\|_F^2.$$

So

$$\begin{aligned} F(\Theta, M_k) &= \|Y - M_k - X \odot \Theta\|_F^2 + \nu_0\|M_k\|_n + \nu_1\|\Theta\|_n \\ &= \|Y - M_k - X \odot \Theta_k\|_F^2 + \|X \odot (\Theta_k - \Theta)\|_F^2 - 2\mathsf{tr}[A_k'(\Theta_k - \Theta)] \\ &\quad + \nu_0\|M_k\|_n + \nu_1\|\Theta\|_n \\ &\leq \|Y - M_k - X \odot \Theta_k\|_F^2 + \tau^{-1}\|\Theta_k - \Theta\|_F^2 - 2\mathsf{tr}[A_k'(\Theta_k - \Theta)] \end{aligned}$$

$$
\begin{aligned}
&+\nu_0\|M_k\|_n + \nu_1\|\Theta\|_n \\
= \;& p(\Theta, \Theta_k, M_k) + \|Y - M_k - X \odot \Theta_k\|_F^2 + \nu_0\|M_k\|_n + \nu_1\|\Theta\|_n.
\end{aligned}
$$

(iii) By definition, $p(\Theta_k, \Theta_k, M_k) = 0$. So

$$
\begin{aligned}
F(\Theta_{k+1}, M_{k+1}) &= \|Y - X \odot \Theta_{k+1} - M_{k+1}\|_F^2 + \nu_1\|\Theta_{k+1}\|_n + \nu_0\|M_{k+1}\|_n \\
&\leq^{(a)} \|Y - X \odot \Theta_{k+1} - M_k\|_F^2 + \nu_1\|\Theta_{k+1}\|_n + \nu_0\|M_k\|_n \\
&= F(\Theta_{k+1}, M_k) \\
&\leq^{(b)} p(\Theta_{k+1}, \Theta_k, M_k) + \nu_1\|\Theta_{k+1}\|_n + \nu_0\|M_k\|_n + \|Y - M_k - X \odot \Theta_k\|_F^2 \\
&\leq^{(c)} p(\Theta_k, \Theta_k, M_k) + \nu_1\|\Theta_k\|_n + \nu_0\|M_k\|_n + \|Y - M_k - X \odot \Theta_k\|_F^2 \\
&= F(\Theta_k, M_k).
\end{aligned}
$$

(a) is due to the definition of M_{k+1}; (b) is due to (ii); (c) is due to (i).

$\square$

Lemma C.2. *For any $\tau \in (0, 1/\max x_{it}^2)$, any (Θ, M) and any $k \geq 1$,*

$$
F(\Theta, M) - F(\Theta_{k+1}, M_{k+1}) \geq \frac{1}{\tau}\left(\|\Theta_{k+1} - \Theta\|_F^2 - \|\Theta_k - \Theta\|_F^2\right).
$$

Proof. The proof is similar to that of Lemma 2.3 of Beck and Teboulle (2009), with the extension that M_{k+1} is updated after Θ_{k+1}. The key difference here is that, while an update to M_{k+1} is added to the iteration, we show that the lower bound does not depend on M_{k+1} or M_k. Therefore, the convergence property of the algorithm depends mainly on the step of updating Θ.

Let $\partial\|A\|_n$ be an element that belongs to the subgradient of $\|A\|_n$. Note that $\partial\|A\|_n$ is convex in A. Also, $\|Y - X \odot \Theta - M\|_F^2$ is convex in (Θ, M), so for any Θ, M, we have the following three inequalities:

$$
\begin{aligned}
\|Y - X \odot \Theta - M\|_F^2 &\geq \|Y - X \odot \Theta_k - M_k\|_F^2 \\
&\quad -2\mathsf{tr}[(\Theta - \Theta_k)'(X \odot (Y - X \odot \Theta_k - M_k))] \\
&\quad -2\mathsf{tr}[(M - M_k)'(Y - X \odot \Theta_k - M_k)] \\
\nu_1\|\Theta\|_n &\geq \nu_1\|\Theta_{k+1}\|_n + \nu_1\mathsf{tr}[(\Theta - \Theta_{k+1})'\partial\|\Theta_{k+1}\|_n] \\
\nu_0\|M\|_n &\geq \nu_0\|M_k\|_n + \nu_0\mathsf{tr}[(M - M_k)'\partial\|M_k\|_n].
\end{aligned}
$$

In addition,

$$
\begin{aligned}
-F(\Theta_{k+1}, M_{k+1}) &\geq -F(\Theta_{k+1}, M_k) \\
&\geq -p(\Theta_{k+1}, \Theta_k, M_k) - \nu_1\|\Theta_{k+1}\|_n - \nu_0\|M_k\|_n - \|Y - M_k - X \odot \Theta_k\|_F^2.
\end{aligned}
$$

where the two inequalities are due to Lemma C.1. Sum up the above inequalities,

$$
\begin{array}{rcl}
 & & F(\Theta, M) - F(\Theta_{k+1}, M_{k+1}) \geq (A) \\
(A) & := & -2\mathsf{tr}[(\Theta - \Theta_k)'(X \odot (Y - X \odot \Theta_k - M_k))] - 2\mathsf{tr}[(M - M_k)'(Y - X \odot \Theta_k - M_k)] \\
 & & +\nu_1 \mathsf{tr}[(\Theta - \Theta_{k+1})' \partial \|\Theta_{k+1}\|_n] + \nu_0 \mathsf{tr}[(M - M_k)' \partial \|M_k\|_n] - p(\Theta_{k+1}, \Theta_k, M_k).
\end{array}
$$

We now simplify (A). Since $k \geq 1$, both M_k and Θ_{k+1} should satisfy the KKT condition. By Lemma C.1, they are:

$$
\begin{array}{rcl}
0 & = & \nu_1 \partial \|\Theta_{k+1}\|_n - \tau^{-1} 2(\Theta_k - \Theta_{k+1}) + 2A_k \\
0 & = & \nu_0 \partial \|M_k\|_n - 2(Y - X \odot \Theta_k - M_k).
\end{array}
$$

Plug in, we have

$$
\begin{array}{rcl}
(A) & = & \tau^{-1} 2\mathsf{tr}[(\Theta - \Theta_{k+1})'(\Theta_k - \Theta_{k+1})] - \tau^{-1} \|\Theta_k - \Theta_{k+1}\|_F^2 \\
 & = & \frac{1}{\tau} \left(\|\Theta_{k+1} - \Theta\|_F^2 - \|\Theta_k - \Theta\|_F^2 \right).
\end{array}
$$

□

Appendix D. Nuclear-Norm Matrix Estimators

Recall that $\widetilde{M}_S$ and $\widetilde{\Theta}_S$ respectively are the estimated low-rank matrices obtained by the nuclear-norm penalized estimations on sample $S \in \{I, I^c, \{1, ..., T\}\}$. Given the above assumptions, we have consistency in the Frobenius norm for the estimated low-rank matrices.

Proposition D.1. *Suppose* $2\|X \odot U\| < (1-c)\nu_1$, $2\|U\| < (1-c)\nu_0$ *and* $\nu_0 \asymp \nu_1$. *Then under Assumption 3.1, for* $S \in \{I, I^c, \{1, ..., T\}\}$ *(i)*

$$
\frac{1}{NT} \|\widetilde{M}_S - M_S\|_F^2 = O_P\left(\frac{\nu_0^2 + \nu_1^2}{NT}\right) = \frac{1}{NT} \|\widetilde{\Theta}_S - \Theta_S\|_F^2.
$$

(ii) Additionally with Assumption 3.3, there are square matrices H_{S1}, H_{S2}, *so that*

$$
\frac{1}{N} \|\widetilde{A}_S - AH_{S1}\|_F^2 = O_P\left(\frac{\nu_0^2 + \nu_1^2}{NT}\right), \quad \frac{1}{N} \|\widetilde{\Lambda}_S - \Lambda H_{S2}\|_F^2 = O_P\left(\frac{\nu_0^2 + \nu_1^2}{NT}\right).
$$

(iii) Furthermore,

$$
P(\widehat{K}_1 = K_1, \quad \widehat{K}_2 = K_2) \to 1.
$$

Proposition D.1 extends the usual low-rank results to the multi-dimensional case. Convergence using the entire sample is sufficient for consistently selecting the rank, as shown in result (iii), while convergence using the subsamples serves for post-SVT inference after sample-splitting. The proof is divided in the following subsections.

D.1. Level of the Score.

Lemma D.1. *In the presence of serial correlations in x_{it} (Assumption 3.4), $\|X \odot U\| = O_P(\sqrt{N+T}) = \|U\|$. In addition, the chosen ν_0, ν_1 in Section 2.4 are $O_P(\sqrt{N+T})$.*

Proof. The assumption that Ω_{NT} contains independent sub-Gaussian columns ensures that, by the eigenvalue-concentration inequality for sub-Gaussian random vectors (Theorem 5.39 of Vershynin (2010)):

$$\|\Omega_{NT}\Omega_{NT}' - \mathsf{E}\Omega_{NT}\Omega_{NT}'\| = O_P(\sqrt{NT} + N).$$

In addition, let w_i be the $T \times 1$ vector of $\{x_{it}u_{it} : t \leq T\}$. We have, for each (i, j, t, s),

$$\mathsf{E}(w_i w_j')_{s,t} = \mathsf{E}x_{it}x_{js}u_{it}u_{js} = \begin{cases} \mathsf{E}x_{it}^2 u_{it}^2, & i = j, t = s \\ 0, & \text{otherwise} \end{cases}$$

due to the conditional cross-sectional and serial independence in u_{it}. Then for the (i, j)'th entry of $\mathsf{E}\Omega_{NT}\Omega_{NT}'$,

$$\begin{aligned} (\mathsf{E}\Omega_{NT}\Omega_{NT}')_{i,j} &= (\mathsf{E}(X \odot U)\Sigma_T^{-1}(X \odot U)')_{i,j} \\ &= \mathsf{E}w_i'\Sigma_T^{-1}w_j = tr(\Sigma_T^{-1}\mathsf{E}w_j w_i') = \begin{cases} \sum_{t=1}^T (\Sigma_T^{-1})_{tt}\mathsf{E}x_{it}^2 u_{it}^2, & i = j \\ 0, & i \neq j. \end{cases} \end{aligned}$$

Hence $\|\mathsf{E}\Omega_{NT}\Omega_{NT}'\| \leq O(T)$. This implies $\|\Omega_{NT}\Omega_{NT}'\| \leq O(T + N)$. Hence $\|X \odot U\| \leq \|\Omega_{NT}\|\|\Sigma_T^{1/2}\| \leq O_P(\max\{\sqrt{N}, \sqrt{T}\})$. The rate for $\|U\|$ follows from the same argument. The second claim that ν_0, ν_1 satisfy the same rate constraint follows from the same argument, by replacing U with Z, and Assumption 3.4 is still satisfied by Z and $X \odot Z$.

□

D.2. Useful Claims.

The proof of Proposition D.1 uses some claims that are proved in the following lemma. Let us first recall the notations. Define $U_2 D_2 V_2' = \Theta$ and $U_1 D_1 V_1' = M$ as the singular value decompositions of the true values Θ and M. Further decompose, for $j = 1, 2$,

$$U_j = (U_{j,r}, U_{j,c}), \quad V_j = (V_{j,r}, V_{j,c})$$

Here $(U_{j,r}, V_{j,r})$ corresponds to the nonzero singular values, while $(U_{j,c}, V_{j,c})$ corresponds to the zero singular values. In addition, for any $N \times T/2$ matrix Δ, let

$$\mathcal{P}_j(\Delta) = U_{j,c}U_{j,c}'\Delta V_{j,c}V_{j,c}', \quad \mathcal{M}_j(\Delta) = \Delta - \mathcal{P}_j(\Delta).$$

Here $U_{j,c}U_{j,c}'$ and $V_{j,c}V_{j,c}'$ respectively are the projection matrices onto the columns of $U_{j,c}$ and $V_{j,c}$. Therefore, $\mathcal{M}_1(\cdot)$ and $\mathcal{M}_2(\cdot)$ can be considered as the projection matrices onto the "low-rank" spaces of Θ and M respectively, and $\mathcal{P}_1(\cdot)$ and $\mathcal{P}_2(\cdot)$ ar projections onto their orthogonal spaces.

Lemma D.2 (claims)**.** *Same results below also apply to $\mathcal{P}_2$ Θ, and samples on I^c. For any matrix Δ,*

(i) $\|\mathcal{P}_1(\Delta) + M\|_n = \|\mathcal{P}_1(\Delta)\|_n + \|M\|_n$.

(ii) $\|\Delta\|_F^2 = \|\mathcal{M}_1(\Delta)\|_F^2 + \|\mathcal{P}_1(\Delta)\|_F^2$

(iii) $\mathsf{rank}(\mathcal{M}_1(\Delta)) \leq 2K_1$*, where* $K_1 = \mathsf{rank}(M)$.

(iv) $\|\Delta\|_F^2 = \sum_j \sigma_j^2$ *and* $\|\Delta\|_n^2 \leq \|\Delta\|_F^2 \mathsf{rank}(\Delta)$*, with* σ_j *as the singular values of* Δ*.*

(v) For any Δ_1, Δ_2*,* $|\mathsf{tr}(\Delta_1\Delta_2)| \leq \|\Delta_1\|_n \|\Delta_2\|$*, Here* $\|.\|$ *denotes the operator norm.*

Proof. (i) Note that $M = U_{1,r}D_{1,r}V_{1,r}'$ where $D_{1,r}$ are the subdiagonal matrix of nonzero singular values. The claim follows from Lemma 2.3 of Recht et al. (2010).

(ii) Write

$$U_1'\Delta V_1 = \begin{pmatrix} A & B \\ C & U_{1,c}'\Delta V_{1,c} \end{pmatrix} = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & U_{1,c}'\Delta V_{1,c} \end{pmatrix} := H_2 + H_1.$$

Then $\mathcal{P}_1(\Delta) = U_1 H_1 V_1'$ and $\mathcal{M}_1(\Delta) = U_1 H_2 V_1'$. So

$$\|\mathcal{P}_1(\Delta)\|_F^2 = \mathsf{tr}(U_1 H_1 V_1' V_1 H_1' U_1') = \mathsf{tr}(H_1 H_1') = \|H_1\|_F^2.$$

Similarly, $\|\mathcal{M}_1(\Delta)\|_F^2 = \|H_2\|_F^2$. So

$$\|H_2\|_F^2 + \|H_1\|_F^2 = \|U_1'\Delta V_1\|_F^2 = \|\Delta\|_F^2.$$

(iii) This is Lemma 1 of Negahban and Wainwright (2011).

(iv) The first is a basic equality, and the second follows from the Cauchy-Schwarz inequality.

(v) Let $UDV' = \Delta_1$ be the singular value decomposition of Δ_1, then

$$|\mathsf{tr}(\Delta_1\Delta_2)| = |\sum_i D_{ii}(V'\Delta_2 U)_{ii}| \leq \max_i |(V'\Delta_2 U)_{ii}| \sum_i D_{ii} \leq \|\Delta_1\|_n \|\Delta_2\|.$$

D.3. **The RSC condition.** We now provide primitive conditions for the RSC condition: Assumption 3.1. Recall

$$\mathcal{C}(c_1, c_2) = \Bigg\{(\Delta_1, \Delta_2) : \|\mathcal{P}_1(\Delta_1)\|_n + \|\mathcal{P}_2(\Delta_2)\|_n \leq c\|\mathcal{M}_1(\Delta_1)\|_n + c\|\mathcal{M}_2(\Delta_2)\|_n,$$
$$\frac{1}{NT}\|\Delta_1\|_F^2 + \frac{1}{NT}\|\Delta_2\|_F^2 \geq \frac{c_2}{\sqrt{NT}}\Bigg\}$$

We shall consider the general structure of x_{it}:

$$x_{it} = \mu_{it} + e_{it}$$

where μ_{it} is the mean part of x_{it}, and it can be either deterministic or random. We shall denote the conditional expectation $\mathsf{E}_\mu(\cdot) = \mathsf{E}(\cdot|\mu_{it} : t \leq T, i \leq N)$ when μ_{it} is random. Correspondingly, denote by $P_\mu(\cdot)$ as the conditional probability.

Lemma D.3. *For any $c_1 > 0$, there are $\kappa_c, c_2, B > 0$, uniformly for $(\Delta_1, \Delta_2) \in \mathcal{C}(c_1, c_2)$,*

$$\|\Delta_1 + \Delta_2 \odot X\|_F^2 \geq \kappa_c\|\Delta_1\|_F^2 + \kappa_c\|\Delta_2\|_F^2 - (N+T)B$$

provided that the following two conditions hold:

(i) variations in X: for all $i \leq N, t \leq T$, all eigenvalues of Σ_{it} are bounded from below by a nonrandom constant $c_0 > 0$ almost surely, where

$$\Sigma_{it} = \begin{pmatrix} 1 & \mathsf{E}_\mu x_{it} \\ \mathsf{E}_\mu x_{it} & \mathsf{E}_\mu x_{it}^2 \end{pmatrix} = \begin{pmatrix} 1 & \mu_{it} \\ \mu_{it} & \mu_{it}^2 + \mathsf{E}_\mu e_{it}^2 \end{pmatrix}$$

(ii) concentration inequality: define

$$\mathcal{B}(x) := \{(\Delta_1, \Delta_2) \in \mathcal{C}(c_1, c_2), \frac{1}{NT}\|\Delta_1\|_F^2 + \frac{1}{NT}\|\Delta_2\|_F^2 \leq x\}.$$

For any $C_0 > 0$, $x > 0$ and sufficiently large constant $B > 0$, there are constants $C', c' > 0$ that only depend on C_0, almost surely for $\{\mu_{it}\}$,

$$P_\mu\left(\sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T (Z_\mu(\Delta)_{it} - \mathsf{E}_\mu Z(\Delta)_{it})\right| > \frac{N+T}{NT}B + C_0 x\right) \leq C'\exp(-c'NTx^2) \tag{D.1}$$

where $Z_\mu(\Delta)_{it} := (x_{it}\Delta_{2,it} + \Delta_{1,it})^2$.

Condition (ii) of Lemma D.3 requires a concentration inequality over $\mathcal{B}(x)$, which is an intersection of the low-rank set with

$$\left\{(\Delta_1,\Delta_2): \frac{c_2}{\sqrt{NT}} \leq \frac{1}{NT}\|\Delta_1\|_F^2 + \frac{1}{NT}\|\Delta_2\|_F^2 \leq x\right\}. \tag{D.2}$$

This condition is often verifiable using more primitive conditions on restricted strong convexity sets, e.g., Lemma 14 of Klopp (2014) and result (32) of Negahban and Wainwright (2012). Below, we provide one set of primitive conditions to verify it. Apparently when $x < \frac{c_2}{\sqrt{NT}}$, $\mathcal{B}(x)$ is an empty set so (D.1) is trivially satisfied. Thus we focus on $x > \frac{c_2}{\sqrt{NT}}$.

Lemma D.4. *Suppose:*

(i') for all (i,t) $|\Delta_{1,it}| + |\Delta_{2,it}| < C$ when $(\Delta_1,\Delta_2) \in \mathcal{B}(x)$ and $|x_{it}| < C$;

(ii') x_{it} is independent across (i,t), conditioning on $\{\mu_{it} : i \leq N, t \leq T\}$.

Then the concentration inequality (D.1) and thus condition (ii) of Lemma D.3 holds.

Remark D.1. The primitive condition (i') requires the estimation errors Δ_1 and Δ_2 have bounded entrywise $\|.\|_\infty$-norm. We can always restrict the parameter space so that the regularization problem is optimized on sufficiently large compact sets to satisfy this condition. Additionally, for the iid Bernoulli regressors x_{it} as in the matrix completion problem, it is shown that the estimation error of the nuclear-norm penalized regression satisfies this condition (Theorem 1 of Chen et al. (2019)).

One may also add an additional $\|.\|_\infty$-norm constraint to the nuclear-norm penalized regression, such as $\|M\|_\infty < C$ and $\|\Theta\|_\infty < C$. Note that C may be arbitrarily large and independent of (N,T), so *neither* C is an additional tuning parameter *nor* the $\|.\|_\infty$- constraint is an additional regularization in the usual sense. The $\|.\|_\infty$-constraint is simply a formalization to require that elements the estimator do not diverge.

Proof. **Proof of Lemma D.3**

We use the standard peeling argument. For notational simplicity, write

$$\Delta = (\Delta_1,\Delta_2), \quad X(\Delta) := \|\Delta_1 + \Delta_2 \odot X\|_F^2.$$

Let $c_0 = \min_{it} \psi_{\min}(\Sigma_{it})$. Then

$$\mathsf{E}_\mu X(\Delta) = \sum_{it} (\Delta_{1,it}, \Delta_{2,it}) \Sigma_{it} (\Delta_{1,it}, \Delta_{2,it})' \geq c_0 \|\Delta_1\|_F^2 + c_0 \|\Delta_2\|_F^2.$$

Define event, for sufficiently large $B > 0$,

$$\mathcal{E}_{NT}(\Delta) := \{|X(\Delta) - \mathsf{E}_\mu X(\Delta)| > \frac{1}{2} \mathsf{E}_\mu X(\Delta) + (N+T)B\}.$$

We aim to show $P(\exists \Delta \in \mathcal{C}(c_1, c_2), \mathcal{E}_{NT}(\Delta)) \to 0$. Once this is achieved, then $P(\forall \Delta \in \mathcal{C}(c_1, c_2), \mathcal{E}_{NT}(\Delta)^c) \to 1$. On the event $\mathcal{E}_{NT}(\Delta)^c$, the RSC condition holds for $\kappa_c = c_0/2$, because:

$$X(\Delta) \geq \frac{1}{2} \mathsf{E}_\mu X(\Delta) - (N+T)B \geq \kappa_c \|\Delta_1\|_F^2 + \kappa_c \|\Delta_2\|_F^2 - (N+T)B.$$

To prove $P(\exists \Delta \in \mathcal{C}(c_1, c_2), \mathcal{E}_{NT}(\Delta)) \to 0$, let

$$\Gamma_l = \{\Delta \in \mathcal{C}(c_1, c_2) : 2^l v_{NT} \leq \frac{1}{NT} \mathsf{E}_\mu X(\Delta) \leq 2^{l+1} v_{NT}\}$$

where $v_{NT} = \frac{B}{\sqrt{NT}}$, and $l \in \mathbb{N}$. We let $c_2 = 2c_0^{-1} B$ in the definition of $\mathcal{C}(c_1, c_2)$. Then for $\Delta \in \mathcal{C}(c_1, c_2)$ we have

$$\frac{1}{NT} \|\Delta_1\|_F^2 + \frac{1}{NT} \|\Delta_2\|_F^2 \geq 2 c_0^{-1} v_{NT}.$$

Then $\frac{1}{NT} \mathsf{E}_\mu X(\Delta) \geq \frac{1}{NT} c_0 \|\Delta_1\|_F^2 + \frac{1}{NT} c_0 \|\Delta_2\|_F^2 \geq 2 v_{NT}$. Hence there is $l \in \mathbb{N}$ so that $\Delta \in \Gamma_l$ as long as $\Delta \in \mathcal{C}(c_1, c_2)$. Thus $\mathcal{C}(c_1, c_2) \subset \cup_{l=1}^\infty \Gamma_l$.

In addition, if $\Delta \in \Gamma_l$, then $\mathcal{E}_{NT}(\Delta)$ implies

$$|X(\Delta) - \mathsf{E}_\mu X(\Delta)| - (N+T)B > \frac{1}{2} \mathsf{E}_\mu X(\Delta) \geq \frac{1}{2} NT 2^l v_{NT} = \frac{1}{4} NT 2^{l+1} v_{NT}$$

and

$$\frac{1}{NT} \|\Delta_1\|_F^2 + \frac{1}{NT} \|\Delta_2\|_F^2 \leq \frac{c_0^{-1}}{NT} \mathsf{E}_\mu X(\Delta) \leq c_0^{-1} 2^{l+1} v_{NT}.$$

This also implies $\Gamma_l \subset \mathcal{B}(x_l)$ with $x_l := c_0^{-1} 2^{l+1} v_{NT}$. Hence for c' that only depends on c_0 (but not on B), (D.1) yields

$$\begin{aligned}
& P(\exists \Delta \in \mathcal{C}(c_1, c_2), \mathcal{E}_{NT}(\Delta)) = \mathsf{E} P_\mu(\exists \Delta \in \mathcal{C}(c_1, c_2), \mathcal{E}_{NT}(\Delta)) \\
\leq \ & \mathsf{E} \sum_{l=1}^\infty P_\mu(\sup_{\Delta \in \mathcal{B}(x_l)} |X(\Delta) - \mathsf{E}_\mu X(\Delta)| > \frac{1}{4} NT 2^{l+1} v_{NT} + (N+T)B) \\
= \ & \mathsf{E} \sum_{l=1}^\infty P_\mu \left(\sup_{\Delta \in \mathcal{B}(x_l)} \left| \frac{1}{NT} \sum_{it} (Z_\mu(\Delta)_{it} - \mathsf{E}_\mu Z(\Delta)_{it}) \right| > \frac{c_0}{4} x_l + \frac{(N+T)B}{NT} \right)
\end{aligned}$$

$$\leq \quad \mathsf{E}\sum_{l=1}^{\infty} C' \exp(-c' 4^{l+1} B^2) \leq \frac{\exp(-16c'B^2)}{1-\exp(-16c'B^2)} < \epsilon$$

for any arbitrarily small $\epsilon > 0$, where the last inequality holds for sufficiently large B.

Proof of Lemma D.4

Let

$$Z := \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(Z_\mu(\Delta)_{it} - \mathsf{E}_\mu Z(\Delta)_{it}) \right|$$

where we recall $Z_\mu(\Delta)_{it} := (x_{it}\Delta_{2,it} + \Delta_{1,it})^2$.

We first bound $\mathsf{E}_\mu Z$, then bound the tail probability of $Z > \mathsf{E}_\mu Z$.

As for $\mathsf{E}_\mu Z$, for any $\Delta \in \mathcal{B}(x) \subset \mathcal{C}(c_1, c_2)$,

$$\begin{aligned}
\|\Delta_1\|_n + \|\Delta_2\|_n &= \|\mathcal{P}_1(\Delta_1) + \mathcal{M}_1(\Delta_1)\|_n + \|\mathcal{P}_2(\Delta_2) + \mathcal{M}_2(\Delta_2)\|_n \\
&\leq (1+c_1)\|\mathcal{M}_1(\Delta_1)\|_n + (1+c_1)\|\mathcal{M}_2(\Delta_2)\|_n \\
&\leq (1+c_1)\sqrt{\mathrm{rank}(M)}\|\mathcal{M}_1(\Delta_1)\|_F + (1+c_1)\sqrt{\mathrm{rank}(\Theta)}\|\mathcal{M}_2(\Delta_2)\|_F \\
&\leq C(\|\Delta_1\|_F + \|\Delta_2\|_F) \leq C\sqrt{2}\sqrt{NTx},
\end{aligned}$$

where the first and third inequalities follow from the definition of $\mathcal{B}(x)$ and $\mathcal{C}(c_1, c_2)$ while the second inequality follows from Lemma D.2. Thus

$$\|\Delta_1\|_n + \|\Delta_2\|_n \leq C^*\sqrt{NTx}. \tag{D.3}$$

where C^* only depends c_1 and the ranks of the low rank matrices M and Θ taking their true values. Next, let ϵ_{it} be an iid Rademacher sequence. Let $G_1 = (\epsilon_{it})_{N\times T}$ and $G_2 = (\epsilon_{it}x_{it})_{N\times T}$, both are $N \times T$ matrices. Then by the expected operator norm for subGaussian matrices, (e.g., Exercise 4.4.6 of Vershynin (2018)), $\mathsf{E}_\mu\|G_1\| + \mathsf{E}_\mu\|G_2\| \leq C(\sqrt{N+T})$ where C is nonrandom. Then

$$\begin{aligned}
\mathsf{E}_\mu Z &\leq^{(i)} 2\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_\mu(\Delta)_{it}\epsilon_{it} \right| \\
&\leq^{(ii)} C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \epsilon_{it}x_{it}\Delta_{2,it} \right| + C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \epsilon_{it}\Delta_{1,it} \right| \\
&= C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\mathrm{tr}(G_1\Delta_1') \right| + C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \left| \frac{1}{NT}\mathrm{tr}(G_2\Delta_2') \right| \\
&\leq C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \frac{1}{NT}\|G_1\|\|\Delta_1\|_n + C\mathsf{E}_\mu \sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)} \frac{1}{NT}\|G_2\|\|\Delta_2\|_n
\end{aligned}$$

$$\begin{aligned}
&\leq \quad C\frac{\sqrt{N+T}}{NT}\sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)}\|\Delta_1\|_n + C\frac{\sqrt{N+T}}{NT}\sup_{(\Delta_1,\Delta_2)\in\mathcal{B}(x)}\|\Delta_2\|_n \\
&\leq^{(iii)} \quad C\frac{\sqrt{N+T}}{NT}\sqrt{NTx} = 2C\sqrt{\frac{N+T}{NTC_0}}\sqrt{\frac{C_0}{4}x} \leq \frac{C_0 x}{8} + B\frac{N+T}{NT}.
\end{aligned} \tag{D.4}$$

Inequality (i) follows from the standard symmetrization argument; Inequality (ii) uses the contraction inequality (e.g., (2.3) of Koltchinskii (2011)), which requires $|\Delta_{it}| < C$ and $|x_{it}| < C$ for all (i,t); inequality (iii) follows from (D.3) for $\Delta \in \mathcal{B}(x)$; the last inequality holds for $B > \frac{2C^2}{C_0}$;

We now bound the tail probability of $Z > \mathsf{E}_\mu Z$. The condition $|\Delta_{it}| < C$ and $|x_{it}| < C$ implies there is $C_x > 0$

$$(x_{it}\Delta_{2,it} + \Delta_{1,it})^2 < C_x.$$

Also, conditionally on $\{\mu_{it}\}$, x_{it} are independent across (i,t). So we can apply Massart inequality (e.g., Theorem 14.2 of Bülhmann and van der Geer (2011)), to reach $P_\mu(Z \geq \mathsf{E}_\mu Z + t) \leq \exp(-\frac{NTt^2}{8C_x^2})$ for any $t > 0$. Now set $t = \frac{7C_0 x}{8}$.

$$\begin{aligned}
&P_\mu\left(Z > \frac{N+T}{NT}B + C_0 x\right) \leq P_\mu\left(Z > \mathsf{E}_\mu Z - \frac{C_0 x}{8} + C_0 x\right) = P_\mu\left(Z > \mathsf{E}_\mu Z + \frac{7C_0 x}{8}\right) \\
&\leq \quad \exp(-\frac{NTt^2}{8C_x^2}) = \exp(-\frac{49C_0^2}{512C_x^2}NTx^2).
\end{aligned}$$

The first inequality follows from (D.4). Note that C_x does not depend on B.

□

D.4. **Proof of Proposition D.1: convergence of $\widetilde{\Theta}_S, \widetilde{M}_S$.** In the proof below, we set $S = I$. That is, we consider estimation using only the data with $t \in I$. We set $T_0 = |I|_0$. The proof carries over to $S = I^c$ or $S = I \cup I^c$. We suppress the subscript S for notational simplicity. Let $\Delta_1 = \widetilde{M} - M$ and $\Delta_2 = \widetilde{\Theta} - \Theta$. Then

$$\|Y - \widetilde{M} - X \odot \widetilde{\Theta}\|_F^2 = \|\Delta_1 + X \odot \Delta_2\|_F^2 + \|U\|_F^2 - 2\mathsf{tr}[U'(\Delta_1 + X \odot \Delta_2)].$$

Note that $\mathsf{tr}(U'(X \odot \Delta_2)) = \mathsf{tr}(\Delta_2'(X \odot U))$. Thus by claim (v),

$$\begin{aligned}
|2\mathsf{tr}[U'(\Delta_1 + X \odot \Delta_2)]| &\leq 2\|U\|\|\Delta_1\|_n + 2\|X \odot U\|\|\Delta_2\|_n \\
&\leq (1-c)\nu_0\|\Delta_1\|_n + (1-c)\nu_1\|\Delta_2\|_n.
\end{aligned}$$

Thus $\|Y-\widetilde{M}-X\odot\widetilde{\Theta}\|_F^2\leq\|Y-M-X\odot\Theta\|_F^2$ (evalued at the true parameters) implies

$$\begin{aligned}\|\Delta_1+X\odot\Delta_2\|_F^2+\nu_0\|\widetilde{M}\|_n+\nu_1\|\widetilde{\Theta}\|_n &\leq (1-c)\nu_0\|\Delta_1\|_n+(1-c)\nu_1\|\Delta_2\|_n\\ &\quad +\nu_0\|M\|_n+\nu_1\|\Theta\|_n.\end{aligned}$$

Now

$$\begin{aligned}\|\widetilde{M}\|_n &= \|\Delta_1+M\|_n=\|M+\mathcal{P}_1(\Delta_1)+\mathcal{M}_1(\Delta_1)\|_n\\ &\geq \|M+\mathcal{P}_1(\Delta_1)\|_n-\|\mathcal{M}_1(\Delta_1)\|_n\\ &= \|M\|_n+\|\mathcal{P}_1(\Delta_1)\|_n-\|\mathcal{M}_1(\Delta_1)\|_n,\end{aligned}$$

where the last equality follows from claim (i). Similar lower bound applies to $\|\widetilde{\Theta}\|_n$. Therefore,

$$\begin{aligned}&\|\Delta_1+X\odot\Delta_2\|_F^2+c\nu_0\|\mathcal{P}_1(\Delta_1)\|_n+c\nu_1\|\mathcal{P}_2(\Delta_2)\|_n\\ \leq\ & (2-c)\nu_0\|\mathcal{M}_1(\Delta_1)\|_n+(2-c)\nu_1\|\mathcal{M}_2(\Delta_2)\|_n. \qquad \text{(D.5)}\end{aligned}$$

In the case U is Gaussian, $\|U\|$ and $\|X\odot U\|\asymp\max\{\sqrt{N},\sqrt{T}\}$, while in the more general case, set $\nu_0\asymp\nu_1\asymp\max\{\sqrt{N},\sqrt{T}\}$. Inequality (D.5) implies there is $c_1>0$,

$$\|\mathcal{P}_1(\Delta_1)\|_n+\|\mathcal{P}_2(\Delta_2)\|_n\leq c_1\|\mathcal{M}_1(\Delta_1)\|_n+c_1\|\mathcal{M}_2(\Delta_2)\|_n.$$

We now apply Assumption 3.1: Recall the set

$$\mathcal{C}(c_1,c_2)=\Big\{(\Delta_1,\Delta_2):\|\mathcal{P}_1(\Delta_1)\|_n+\|\mathcal{P}_2(\Delta_2)\|_n\leq c_1\|\mathcal{M}_1(\Delta_1)\|_n+c_1\|\mathcal{M}_2(\Delta_2)\|_n,$$
$$\|\Delta_1\|_F^2+\|\Delta_2\|_F^2\geq c_2\sqrt{NT}\Big\}.$$

Consider two cases. Case 1: $\|\Delta_1\|_F^2+\|\Delta_2\|_F^2\leq c_2\sqrt{NT}$. Then $\|\Delta_1\|_F^2+\|\Delta_2\|_F^2\leq O_P(\nu_0^2+\nu_1^2)$.

Case 2: $\|\Delta_1\|_F^2+\|\Delta_2\|_F^2\geq c_2\sqrt{NT}$. Then $(\Delta_1,\Delta_2)\in\mathcal{C}(c_1,c_2)$, so by Assumption 3.1, for some constants $\kappa,\eta>0$, with large probability

$$\|\Delta_1+\Delta_2\odot X\|_F^2\geq\kappa\|\Delta_1\|_F^2+\kappa\|\Delta_2\|_F^2-(N+T)\eta.$$

Then (D.5) and the proved claims show that for a generic $C>0$,

$$\begin{aligned}\|\Delta_1\|_F^2+\|\Delta_2\|_F^2 &\leq C\nu_0\|\mathcal{M}_1(\Delta_1)\|_n+C\nu_1\|\mathcal{M}_2(\Delta_2)\|_n+O_P(N+T)\\ &\leq^{\text{claim (iv)}} C\nu_0\|\mathcal{M}_1(\Delta_1)\|_F\sqrt{\mathsf{rank}(\mathcal{M}_2(\Delta_1))}\end{aligned}$$

$$
\begin{aligned}
& +C\nu_1\|\mathcal{M}_2(\Delta_2)\|_F\sqrt{\mathsf{rank}(\mathcal{M}_2(\Delta_2))}+O_P(N+T) \\
\leq^{\text{claim (iii)}} \quad & C\nu_0\|\mathcal{M}_1(\Delta_1)\|_F\sqrt{2K_1}+C\nu_1\|\mathcal{M}_2(\Delta_2)\|_F\sqrt{2K_2}+O_P(N+T) \\
\leq^{\text{claim (ii)}} \quad & C\nu_0\|\Delta_1\|_F+C\nu_1\|\Delta_2\|_F+O_P(N+T) \\
\leq \quad & C\max\{\nu_0,\nu_1\}\sqrt{\|\Delta_1\|_F^2+\|\Delta_2\|_F^2}+O_P(N+T).
\end{aligned}
$$

Thus $\|\Delta_1\|_F^2+\|\Delta_2\|_F^2\leq O_P(\nu_0^2+\nu_1^2)$.

D.5. **Proof of Proposition D.1: convergence of $\widetilde{\Lambda}_S,\widetilde{A}_S$.** We proceed the proof in the following steps.

step 1: bound the eigenvalues

Replace $\nu_0^2+\nu_1^2$ with $O_P(N+T)$, then

$$\|\widetilde{\Theta}_S-\Theta_S\|_F^2=O_P(N+T).$$

Let $S_f=\frac{1}{T_0}\sum_{t\in I}f_tf_t'$, $\Sigma_f=\frac{1}{T}\sum_{t=1}^T f_tf_t'$ and $S_\Lambda=\frac{1}{N}\Lambda'\Lambda$. Let $\psi_{I,1}^2\geq...\geq\psi_{I,K_1}^2$ be the K_1 nonzero eigenvalues of $\frac{1}{NT_0}\Theta_I\Theta_I'=\frac{1}{N}\Lambda S_f\Lambda'$. Let $\widetilde{\psi}_1^2\geq...\geq\widetilde{\psi}_{K_2}^2$ be the first K_2 nonzero singular values of $\frac{1}{NT_0}\widetilde{\Theta}_I\widetilde{\Theta}_I'$. Also, let ψ_j^2 be the j th largest eigenvalue of $\frac{1}{N}\Lambda\Sigma_f\Lambda'$. Note that $\psi_1^2...\psi_{K_1}^2$ are the same as the eigenvalues of $\Sigma_f^{1/2}S_\Lambda\Sigma_f^{1/2}$. Hence by Assumption 3.3, there are constants $c_1,...,c_{K_1}>0$, so that

$$\psi_j^2=c_j,\quad j=1,...,K_1.$$

Then by Weyl's theorem, for $j=1,...,\min\{T_0,N\}$, with the assumption that $\|S_f-\Sigma_f\|=O_P(\frac{1}{\sqrt{T}})$, $|\psi_{I,j}^2-\psi_j^2|\leq\frac{1}{N}\|\Lambda(S_f-\Sigma_f)\Lambda'\|\leq O(1)\|S_f-\Sigma_f\|=O_P(\frac{1}{\sqrt{T}})$. This also implies $\|\Theta_I\|=\psi_{I,1}\sqrt{NT_0}=\sqrt{(c_1+o_P(1))T_0N}$.

Still by Weyl's theorem, for $j=1,...,\min\{T_0,N\}$,

$$
\begin{aligned}
|\widetilde{\psi}_j^2-\psi_{I,j}^2| &\leq \frac{1}{NT_0}\|\widetilde{\Theta}_I\widetilde{\Theta}_I'-\Theta_I\Theta_I'\| \\
&\leq \frac{2}{NT_0}\|\Theta_I\|\|\widetilde{\Theta}_I-\Theta_I\|+\frac{1}{NT_0}\|\widetilde{\Theta}_I-\Theta_I\|^2=O_P(\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{T}}).
\end{aligned}
$$

implying

$$|\widetilde{\psi}_j^2-\psi_j^2| = O_P(\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{T}}).$$

Then for all $j\leq K_1$, with probability approaching one,

$$
\begin{aligned}
|\psi_{j-1}^2-\widetilde{\psi}_j^2| &\geq |\psi_{j-1}^2-\psi_j^2|-|\psi_j^2-\widetilde{\psi}_j^2|\geq(c_{j-1}-c_j)/2 \\
|\widetilde{\psi}_j^2-\psi_{j+1}^2| &\geq |\psi_j^2-\psi_{j+1}^2|-|\widetilde{\psi}_j^2-\psi_j^2|\geq(c_j-c_{j+1})/2
\end{aligned}
\tag{D.6}
$$

with $\psi_{K_1+1}^2=c_{K_1+1}=0$ because $\Theta_I\Theta_I'$ has at most K_1 nonzero eigenvalues.

step 2: characterize the eigenvectors

Next, we show that there is a $K_1 \times K_1$ matrix H_1, so that the columns of $\frac{1}{\sqrt{N}}\Lambda H_1$ are the first K_1 eigenvectors of $\Lambda\Sigma_f\Lambda'$. Let $L = S_\Lambda^{1/2}\Sigma_f S_\Lambda^{1/2}$. Let R be a $K_1 \times K_1$ matrix whose columns are the eigenvectors of L. Then $D = R'LR$ is a diagonal matrix of the eigenvalues of L that are distinct nonzeros according to Assumption 3.3. Let $H_1 = S_\Lambda^{-1/2}R$. Then

$$\begin{aligned}\frac{1}{N}\Lambda\Sigma_f\Lambda'\Lambda H_1 &= \Lambda S_\Lambda^{-1/2}S_\Lambda^{1/2}\Sigma_f S_\Lambda^{1/2}S_\Lambda^{1/2}H_1 = \Lambda S_\Lambda^{-1/2}RR'LR \\ &= \Lambda H_1 D.\end{aligned}$$

Now $\frac{1}{N}(\Lambda H_1)'\Lambda H_1 = H_1' S_\Lambda H_1 = R'R = I$. So the columns of $\Lambda H_1/\sqrt{N}$ are the eigenvectors of $\Lambda\Sigma_f\Lambda'$, corresponding to the eigenvalues in D.

Importantly, the rotation matrix H_1, by definition, depends only on S_Λ, Σ_f, which is time-invariant, and does not depend on the splitted sample.

step 3: prove the convergence

We first assume $\widehat{K}_1 = K_1$. The proof of the consistency is given in step 4 below. Once this is true, then the following argument can be carried out conditional on the event $\widehat{K}_1 = K_1$. Apply Davis-Kahan sin-theta inequality, and by (D.6),

$$\begin{aligned}&\|\frac{1}{\sqrt{N}}\widetilde{\Lambda}_I - \frac{1}{\sqrt{N}}\Lambda H_1\|_F \le \frac{\frac{1}{N}\|\Lambda\Sigma_f\Lambda' - \frac{1}{T_0}\widetilde{\Theta}_I\widetilde{\Theta}_I'\|}{\min_{j\le K_2}\min\{|\psi_{j-1}^2 - \widetilde{\psi}_j^2|, |\widetilde{\psi}_j^2 - \psi_{j+1}^2|\}} \\ \le\ & O_P(1)\frac{1}{N}\|\Lambda\Sigma_f\Lambda' - \frac{1}{T_0}\widetilde{\Theta}_I\widetilde{\Theta}_I'\| \\ \le\ & O_P(1)\frac{1}{N}\|\Lambda(\Sigma_f - S_f)\Lambda'\| + \frac{1}{NT_0}\|\Theta_I\Theta_I' - \widetilde{\Theta}_I\widetilde{\Theta}_I'\| = O_P(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}).\end{aligned}$$

step 4: prove $P(\widehat{K}_1 = K_1) = 1$.

Note that $\psi_j(\widetilde{\Theta}) = \widetilde{\psi}_j\sqrt{NT}$. By step 1, for all $j \le K_1$, $\widetilde{\psi}_j^2 \ge c_j - o_P(1) \ge c_j/2$ with probability approaching one. Also, $\widetilde{\psi}_{K_1+1}^2 \le O_P(T^{-1/2} + N^{-1/2})$, implying that

$$\min_{j\le K_1}\psi_j(\widetilde{\Theta}) \ge c_{K_1}\sqrt{NT}/2, \quad \max_{j>K_1}\psi_j(\widetilde{\Theta}) \le O_P(T^{-1/4} + N^{-1/4})\sqrt{NT}.$$

In addition, $\nu_0^{1/2-\epsilon}\|\widetilde{\Theta}\|^{1/2+\epsilon} \asymp (\sqrt{N+T})^{1/2-\epsilon}(\sqrt{NT})^{1/2+\epsilon}$ for some small $\epsilon \in (0,1)$,

$$\min_{j\le K_1}\psi_j(\widetilde{\Theta}) \ge \nu_0^{1/2-\epsilon}\|\widetilde{\Theta}\|^{1/2+\epsilon}, \quad \max_{j>K_1}\psi_j(\widetilde{\Theta}) \le o_P(1)\nu_0^{1/2-\epsilon}\|\widetilde{\Theta}\|^{1/2+\epsilon}.$$

This proves the consistency of $\widehat{K}_1$.

Finally, the proof of the convergence for $\widetilde{A}_I$ and the consistency of $\widehat{K}_2$ follows from the same argument. Q.E.D.

Appendix E. Proof of Theorem 3.1

In the factor model

$$x_{it} = l_i' w_t + e_{it},$$

we write $\widehat{e}_{it} = x_{it} - \widehat{l_i' w_t}$, $\widehat{\mu}_{it} = \widehat{l_i' w_t}$ and $\mu_{it} = l_i' w_t$. The proof in this section works for both models.

Let

$$C_{NT} = \min\{\sqrt{N}, \sqrt{T}\}.$$

First recall that $(\widetilde{f}_s, \dot{\lambda}_i)$ are computed as the preliminary estimators in step 3. The main technical requirement of these estimators is that their estimation effects are negligible, specifically, there is a rotation matrix H_1 that is independent of the sample splitting, for each fixed $t \notin I$,

$$\begin{aligned} \|\frac{1}{N}\sum_j (\dot{\lambda}_j - H_1'\lambda_j)e_{jt}\|^2 &= O_P(C_{NT}^{-4}), \\ \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}} \|\frac{1}{T}\sum_{s\notin I} f_s(\widetilde{f}_s - H_1^{-1}f_s)'\mu_{is}e_{is}\|^2 &= O_P(C_{NT}^{-4}). \end{aligned}$$

These are given in Lemmas G.5 and G.6 for the factor model.

E.1. Behavior of $\widehat{f}_t$.

Recall that for each $t \notin I$,

$$(\widehat{f}_{I,t}, \widehat{g}_{I,t}) := \arg\min_{f_t, g_t} \sum_{i=1}^N (\widehat{y}_{it} - \widetilde{\alpha}_i' g_t - \widehat{e}_{it}\widetilde{\lambda}_i' f_t)^2.$$

For notational simplicity, we simply write $\widehat{f}_t = \widehat{f}_{I,t}$ and $\widehat{g}_t = \widehat{g}_{I,t}$, but keep in mind that $\widetilde{\alpha}$ and $\widetilde{\lambda}$ are estimated through the low rank estimations on data I. Note that $\widetilde{\lambda}_i$ consistently estimates λ_i up to a rotation matrix H_1', so $\widehat{f}_t$ is consistent for $H_1^{-1}f_t$. However, as we shall explain below, it is difficult to establish the asymptotic normality for $\widehat{f}_t$ centered at $H_1^{-1}f_t$. Instead, we obtain a new centering quantity, and obtain an expansion for

$$\sqrt{N}(\widehat{f}_t - H_f f_t)$$

with a new rotation matrix H_f that is also independent of t. For the purpose of inference for θ_{it}, this is sufficient.

Let $\widehat{w}_{it} = (\widetilde{\lambda}_i'\widehat{e}_{it}, \widetilde{\alpha}_i')'$, and $\widehat{B}_t = \frac{1}{N}\sum_i \widehat{w}_{it}\widehat{w}_{it}'$. Define $w_{it} = (\lambda_i' e_{it}, \alpha_i')'$, and

$$\widehat{Q}_t = \frac{1}{N}\sum_i \widehat{w}_{it}(\mu_{it}\lambda_i' f_t - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_t + u_{it}).$$

We have

$$\begin{aligned} \begin{pmatrix} \widehat{f}_t \\ \widehat{g}_t \end{pmatrix} &= \widehat{B}_t^{-1}\frac{1}{N}\sum_i \widehat{w}_{it}(y_{it} - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_t) \\ &= \begin{pmatrix} H_1^{-1} f_t \\ H_2^{-1} g_t \end{pmatrix} + \widehat{B}_t^{-1}\widehat{S}_t \begin{pmatrix} H_1^{-1} f_t \\ H_2^{-1} g_t \end{pmatrix} + \widehat{B}_t^{-1}\widehat{Q}_t \end{aligned} \tag{E.1}$$

where

$$\widehat{S}_t = \frac{1}{N}\sum_i \begin{pmatrix} \widetilde{\lambda}_i\widehat{e}_{it}(\lambda_i' H_1 e_{it} - \widetilde{\lambda}_i'\widehat{e}_{it}) & \widetilde{\lambda}_i\widehat{e}_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i') \\ \widetilde{\alpha}_i(\lambda_i' H_1 e_{it} - \widetilde{\lambda}_i'\widehat{e}_{it}) & \widetilde{\alpha}_i(\alpha_i' H_2 - \widetilde{\alpha}_i') \end{pmatrix}.$$

Note that the "upper block" of $\widehat{B}_t^{-1}\widehat{S}_t$ is not first-order negligible. Essentially this is due to the fact that the moment condition

$$\frac{\partial}{\partial\lambda_i}\mathsf{E}\lambda_i e_{it}(\dot{y}_{it} - \alpha_i' g_t - e_{it}\lambda_i' f_t) \neq 0,$$

so is not "Neyman orthogonal" with respect to λ_i. On the other hand, we can get around such difficulty. In Lemma G.7 below, we show that $\widehat{B}_t$ and $\widehat{S}_t$ both converge in probability to block diagonal matrices that are independent of t. So g_t and f_t are "orthogonal", and

$$\widehat{B}_t^{-1}\widehat{S}_t = \begin{pmatrix} \bar{H}_3 & 0 \\ 0 & \bar{H}_4 \end{pmatrix} + o_P(N^{-1/2}).$$

Define $H_f := H_1^{-1} + \bar{H}_3 H_1^{-1}$. Then (E.1) implies that

$$\widehat{f}_t = H_f f_t + \text{upper block of } \widehat{B}_t^{-1}\widehat{Q}_t.$$

Therefore $\widehat{f}_t$ converges to f_t up to a new rotation matrix H_f, which equals H_1^{-1} up to an $o_P(1)$ term $\bar{H}_3 H_1^{-1}$. While the effect of $\bar{H}_3$ is not negligible, it is "absorbed" into the new rotation matrix. As such, we are able to establish the asymptotic normality for $\sqrt{N}(\widehat{f}_t - H_f f_t)$.

Proposition E.1. *For each fixed $t \notin I$, we have*

$$\widehat{f}_t - H_f f_t = H_f(\frac{1}{N}\sum_i \lambda_i\lambda_i'\mathsf{E}e_{it}^2)^{-1}\frac{1}{N}\sum_i \lambda_i e_{it} u_{it} + O_P(C_{NT}^{-2}).$$

Proof. Define

$$B = \frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i\lambda_i' H_1 \mathsf{E} e_{it}^2 & 0 \\ 0 & H_2'\alpha_i\alpha_i' H_2 \end{pmatrix}, \quad S = \frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')\mathsf{E} e_{it}^2 & 0 \\ 0 & \widetilde{\alpha}_i(\alpha_i' H_2 - \widetilde{\alpha}_i') \end{pmatrix}.$$

Both B, S are independent of t due to the stationarity of e_{it}^2. But S depends on the sample splitting through $(\widetilde{\lambda}_i, \widetilde{\alpha}_i)$.

From (E.1),

$$\begin{aligned}
\begin{pmatrix} \widehat{f}_t \\ \widehat{g}_t \end{pmatrix} &= (B^{-1}S + \mathsf{I})\begin{pmatrix} H_1^{-1} f_t \\ H_2^{-1} g_t \end{pmatrix} + B^{-1}\frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i e_{it} \\ H_2'\alpha_i \end{pmatrix} u_{it} \\
&\quad + \sum_{d=1}^{6} A_{dt}, \text{ where} \\
A_{1t} &= (\widehat{B}_t^{-1}\widehat{S}_t - B^{-1}S)\begin{pmatrix} H_1^{-1} f_t \\ H_2^{-1} g_t \end{pmatrix} \\
A_{2t} &= (\widehat{B}_t^{-1} - B^{-1})\frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i e_{it} \\ H_2'\alpha_i \end{pmatrix} u_{it} \\
A_{3t} &= \widehat{B}_t^{-1}\frac{1}{N}\sum_i \begin{pmatrix} \widetilde{\lambda}_i\widehat{e}_{it} - H_1'\lambda_i e_{it} \\ \widetilde{\alpha}_i - H_2'\alpha_i \end{pmatrix} u_{it} \\
A_{4t} &= \widehat{B}_t^{-1}\frac{1}{N}\sum_i \begin{pmatrix} \widetilde{\lambda}_i\widehat{e}_{it} - H_1'\lambda_i e_{it} \\ \widetilde{\alpha}_i - H_2'\alpha_i \end{pmatrix} (\mu_{it}\lambda_i' f_t - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_t) \\
A_{5t} &= (\widehat{B}_t^{-1} - B^{-1})\frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i e_{it} \\ H_2'\alpha_i \end{pmatrix} (\mu_{it}\lambda_i' f_t - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_t) \\
A_{6t} &= B^{-1}\frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i e_{it} \\ H_2'\alpha_i \end{pmatrix} (\mu_{it}\lambda_i' f_t - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_t).
\end{aligned} \tag{E.2}$$

Note that $B^{-1}S$ is a block-diagonal matrix, with the upper block being

$$\bar{H}_3 := H_1^{-1}(\frac{1}{N}\sum_i \lambda_i\lambda_i'\mathsf{E} e_{it}^2)^{-1}\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')\mathsf{E} e_{it}^2.$$

Define

$$H_f := (\bar{H}_3 + \mathsf{I})H_1^{-1}.$$

Fixed $t \notin I$, in the factor model, in Lemma G.9 we show that $\sum_{d=2}^5 A_{dt} = O_P(C_{NT}^{-2})$, the "upper block" of A_{6t}, $\frac{1}{N}\sum_i \lambda_i e_{it}(l_i' w_t \lambda_i' f_t - \widehat{l_i' w_t}\dot{\lambda}_i'\widetilde{f}_t) = O_P(C_{NT}^{-2})$

and the upper block of A_{1t} is $O_P(C_{NT}^{-2})$. Therefore,

$$\widehat{f}_t = H_f f_t + H_1^{-1}(\frac{1}{N}\sum_i \lambda_i\lambda_i' \mathsf{E} e_{it}^2)^{-1}\frac{1}{N}\sum_i \lambda_i e_{it} u_{it} + O_P(C_{NT}^{-2}).$$

Given that $\bar{H}_3 = O_P(C_{NT}^{-1})$ we have $H_1^{-1} = H_f + O_P(C_{NT}^{-1})$. By $\frac{1}{N}\sum_i \lambda_i e_{it} u_{it} = O_P(N^{-1/2})$,

$$\widehat{f}_t = H_f f_t + H_f(\frac{1}{N}\sum_i \lambda_i\lambda_i' \mathsf{E} e_{it}^2)^{-1}\frac{1}{N}\sum_i \lambda_i e_{it} u_{it} + O_P(C_{NT}^{-2}). \tag{E.3}$$

□

E.2. **Behavior of $\widehat{\lambda}_i$.** Recall that fix $i \leq N$,

$$(\widehat{\lambda}_{I,i}, \widehat{\alpha}_{I,i}) = \arg\min_{\lambda_i,\alpha_i} \sum_{s\notin I} (\widehat{y}_{is} - \alpha_i' \widehat{g}_{I,s} - \widehat{e}_{is}\lambda_i' \widehat{f}_{I,s})^2.$$

For notational simplicity, we simply write $\widehat{\lambda}_i = \widehat{\lambda}_{I,i}$ and $\widehat{\alpha}_i = \widehat{\alpha}_{I,i}$, but keep in mind that $\widetilde{\alpha}$ and $\widetilde{\lambda}$ are estimated through the low rank estimations on data I. Write

$$T_0 = |I^c|_0.$$

$$(B^{-1}S + \mathsf{I})\begin{pmatrix} H_1^{-1} & 0 \\ 0 & H_2^{-1} \end{pmatrix} := \begin{pmatrix} H_f & 0 \\ 0 & H_g \end{pmatrix}$$

$$\widehat{D}_i = \frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{f}_s \widehat{f}_s' \widehat{e}_{is}^2 & \widehat{f}_s \widehat{g}_s' \widehat{e}_{is} \\ \widehat{g}_s \widehat{f}_s' \widehat{e}_{is} & \widehat{g}_s \widehat{g}_s' \end{pmatrix}, \quad D_i = \frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} H_f f_s f_s' H_f' \mathsf{E} e_{is}^2 & 0 \\ 0 & H_g g_s g_s' H_g' \end{pmatrix} \tag{E.4}$$

Proposition E.2. *Fixed a group $\mathcal{G} \subset \{1, ..., N\}$. Write $\Omega_i := (\frac{1}{T}\sum_{s=1}^T f_s f_s' \mathsf{E} e_{is}^2)^{-1}$.*

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\widehat{\lambda}_{I,i} - H_f'^{-1}\lambda_i) = H_f'^{-1}\frac{1}{|\mathcal{G}|_0 T_0}\sum_{i\in\mathcal{G}}\sum_{s\notin I}\Omega_i f_s e_{is} u_{is} + O_P(C_{NT}^{-2}).$$

Proof. By definition and (E.2),

$$\begin{aligned}
\begin{pmatrix} \widehat{\lambda}_i \\ \widehat{\alpha}_i \end{pmatrix} &= \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{f}_s \widehat{e}_{is} \\ \widehat{g}_s \end{pmatrix}(y_{is} - \widehat{\mu}_{is}\dot{\lambda}_i' \widetilde{f}_s) \\
&= \begin{pmatrix} H_f'^{-1}\lambda_i \\ H_g'^{-1}\alpha_i \end{pmatrix} + D_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} H_f f_s e_{is} \\ H_g g_s \end{pmatrix} u_{is} + \sum_{d=1}^{6} R_{di}, \qquad \text{where,} \\
R_{1i} &= (\widehat{D}_i^{-1} - D_i^{-1})\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} H_f f_s e_{is} \\ H_g g_s \end{pmatrix} u_{is}
\end{aligned}$$

$$\begin{aligned}
R_{2i} &= \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} H_f f_s(\widehat{e}_{is}-e_{is})u_{is} \\ 0 \end{pmatrix} \\
R_{3i} &= (\widehat{D}_i^{-1}-D_i^{-1})\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{f}_s\widehat{e}_{is} \\ \widehat{g}_s \end{pmatrix}(\mu_{is}\lambda_i' f_s-\widehat{\mu}_{is}\dot{\lambda}_i'\widetilde{f}_s) \\
R_{4i} &= D_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{f}_s\widehat{e}_{is} \\ \widehat{g}_s \end{pmatrix}(\mu_{is}\lambda_i' f_s-\widehat{\mu}_{is}\dot{\lambda}_i'\widetilde{f}_s) \\
R_{5i} &= \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{e}_{is}\mathsf{I} & 0 \\ 0 & \mathsf{I} \end{pmatrix}\begin{pmatrix} \widehat{f}_s-H_f f_s \\ \widehat{g}_s-H_g g_s \end{pmatrix}u_{is} \\
R_{6i} &= \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix} \widehat{f}_s\widehat{e}_{is} \\ \widehat{g}_s \end{pmatrix}(\lambda_i' H_f^{-1},\alpha_i' H_g^{-1})\begin{pmatrix} \widehat{e}_{is}\mathsf{I} & 0 \\ 0 & \mathsf{I} \end{pmatrix}\begin{pmatrix} \widehat{f}_s-H_f f_s \\ \widehat{g}_s-H_g g_s \end{pmatrix} \quad \text{(E.5)}
\end{aligned}$$

Let r_{di} be the upper blocks of R_{di}. Lemma G.10 shows $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i^{-1}-D_i^{-1}\|^2 = O_P(C_{NT}^{-2})$. So $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{1i}\| = O_P(C_{NT}^{-2})$. Lemma G.2 shows $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I} f_s(\widehat{e}_{is}-e_{is})u_{is}\|^2 = O_P(C_{NT}^{-4})$, so $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{2i}\| = O_P(C_{NT}^{-2})$. By Lemma G.12,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}\widehat{f}_s\widehat{e}_{is}(\mu_{it}\lambda_i' f_s-\widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\|^2 = O_P(C_{NT}^{-4})$$

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}\widehat{g}_s(\mu_{it}\lambda_i' f_s-\widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\|^2 = O_P(C_{NT}^{-2}).$$

So $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i}\| = O_P(C_{NT}^{-2}) = \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{4i}\|$. Next, by Lemma G.13, $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{5i}\| = O_P(C_{NT}^{-2})$. By Lemma G.14, $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{6i}\| = O_P(C_{NT}^{-2})$. Therefore,

$$\sum_{d=1}^{6}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{di}\| = O_P(C_{NT}^{-2}).$$

Now write $S_{f,I} = \frac{1}{T_0}\sum_{s\notin I} f_s f_s'$ and $S_f = \frac{1}{T}\sum_{s=1}^{T} f_s f_s'$,

$$\begin{aligned}
\Omega_{I,i} &:= (\frac{1}{T_0}\sum_{s\notin I} f_s f_s'\mathsf{E}e_{is}^2)^{-1} = S_{f,I}^{-1}(\mathsf{E}e_{is}^2)^{-1} \\
\Omega_i &:= (\frac{1}{T_0}\sum_{s=1}^{T} f_s f_s'\mathsf{E}e_{is}^2)^{-1} = S_f^{-1}(\mathsf{E}e_{is}^2)^{-1}.
\end{aligned}$$

Then by the assumption $S_{f,I}=S_f+O_P(T^{-1/2})$. It remains to bound, for $\omega_{is}=e_{is}u_{is}$,

$$
\begin{aligned}
&\|\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\Omega_{I,i}-\Omega_i]\frac{1}{T_0}\sum_{s\notin I}f_s\omega_{is}\|^2=\|\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\Omega_{I,i}(S_{f,I}-S_f)\Omega_i\frac{1}{T_0}\sum_{s\notin I}f_s\omega_{is}\mathsf{E}e_{is}^2\|^2\\
\leq\ & \|S_{f,I}-S_f\|^2\|S_f^{-1}\|\|S_{f,I}^{-1}\|\|\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T_0}\sum_{s\notin I}f_s\omega_{is}(\mathsf{E}e_{is}^2)^{-1}\|_F^2\leq O_P(T^{-2}|\mathcal{G}|_0^{-1}).
\end{aligned}
$$

This finishes the proof.

□

E.3. **Proof of normality of $\widehat{\theta}_{it}$.** Fix $t\notin I$. Suppose T is odd and $|I|_0=|I^c|_0=(T-1)/2$. Let

$$
\begin{aligned}
Q_I &:= \frac{1}{|\mathcal{G}|_0|I|_0}\sum_{i\in\mathcal{G}}\sum_{s\notin I}\Omega_i f_s e_{is}u_{is}=O_P(C_{NT}^{-1}|\mathcal{G}|_0^{-1/2}),\\
Q_{I^c} &:= \frac{1}{|\mathcal{G}|_0|I|_0}\sum_{i\in\mathcal{G}}\sum_{s\notin I^c}\Omega_i f_s e_{is}u_{is}=O_P(C_{NT}^{-1}|\mathcal{G}|_0^{-1/2}),\\
Q &:= \frac{1}{|\mathcal{G}|_0 T}\sum_{i\in\mathcal{G}}\sum_{s=1}^{T}\Omega_i f_s e_{is}u_{is},\\
J &:= (\frac{1}{N}\sum_i\lambda_i\lambda_i'\mathsf{E}e_{it}^2)^{-1}\frac{1}{N}\sum_i\lambda_i e_{it}u_{it}=O_P(C_{NT}^{-1})\\
\bar\lambda_{\mathcal{G}} &:= \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\lambda_i.
\end{aligned}
$$

Also $Q_I=O_P(\frac{1}{\sqrt{T|\mathcal{G}|_0}})=Q_{I^c}$ follows from $\max_i\sum_j|\mathsf{Cov}(u_{is}e_{is},u_{js}e_{js}|F)|<C$, and $\max_i\|\Omega_i\|\leq\lambda_{\min}^{-1}(\frac{1}{T}\sum_s f_sf_s')(\min_i\mathsf{E}e_{is}^2)^{-1}<C$. By Propositions E.1 E.2, $\widehat{f}_t-H_ff_t=H_fJ+O_P(C_{NT}^{-2})$ and $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\widehat{\lambda}_{I,i}-H_f'^{-1}\lambda_i)=H_f'^{-1}Q_I+O_P(C_{NT}^{-2})$. Therefore,

$$
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\widehat{\lambda}_{I,i}'\widehat{f}_{I,t}-\lambda_i'f_t)=f_t'Q_I+\bar\lambda_{\mathcal{G}}'J+O_P(C_{NT}^{-2}).
$$

Exchanging I with I^c, we have, for $t\notin I^c$,

$$
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\widehat{\lambda}_{I^c,i}'\widehat{f}_{I^c,t}-\lambda_i'f_t) = f_t'Q_{I^c}+\bar\lambda_{\mathcal{G}}'J+O_P(C_{NT}^{-2}).
$$

Note that the fixed $t\notin I\cup I^c$, so take the average:

$$
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\widehat{\theta}_{it}-\theta_{it}) = f_t'\frac{Q_I+Q_{I^c}}{2}+\bar\lambda_{\mathcal{G}}'J+O_P(C_{NT}^{-2})
$$

$$\begin{aligned}
&= f_t' \frac{1}{|\mathcal{G}|_0 T} \sum_{i\in\mathcal{G}} \left[\sum_{s\in I^c\cup\{t\}} \Omega_i f_s e_{is} u_{is} + \sum_{s\in I\cup\{t\}} \Omega_i f_s e_{is} u_{is} \right] + \bar{\lambda}_{\mathcal{G}}' J \\
&\quad + O_P(C_{NT}^{-2}) \\
&= f_t' Q + \bar{\lambda}_{\mathcal{G}}' J + O_P(C_{NT}^{-2}) \\
&= \frac{1}{\sqrt{T|\mathcal{G}|_0}} \zeta_{NT} + \frac{1}{\sqrt{N}} \xi_{NT} + O_P(C_{NT}^{-2}). \qquad \text{(E.6)}
\end{aligned}$$

where

$$\begin{aligned}
\zeta_{NT} &= f_t' \frac{1}{\sqrt{|\mathcal{G}|_0 T}} \sum_{i\in\mathcal{G}} \sum_{s=1}^{T} \Omega_i f_s e_{is} u_{is}, \\
\xi_{NT} &= \bar{\lambda}_{\mathcal{G}}' V_{\lambda 1}^{-1} \frac{1}{\sqrt{N}} \sum_i \lambda_i e_{it} u_{it} \\
V_{\lambda 1} &= \frac{1}{N} \sum_i \lambda_i \lambda_i' \mathsf{E} e_{it}^2, \quad V_{\lambda 2} = \mathsf{Var}(\frac{1}{\sqrt{N}} \sum_i \lambda_i e_{it} u_{it} | F) \\
V_\lambda &= V_{\lambda 1}^{-1} V_{\lambda 2} V_{\lambda 1}^{-1}, \quad V_f = \frac{1}{T} \sum_{s=1}^{T} \mathsf{Var}(\frac{1}{\sqrt{|\mathcal{G}|_0}} \sum_{i\in\mathcal{G}} \Omega_i f_s e_{is} u_{is} | F), \\
\Sigma_{NT} &= \frac{1}{T|\mathcal{G}|_0} f_t' V_f f_t + \frac{1}{N} \bar{\lambda}_{\mathcal{G}}' V_\lambda \bar{\lambda}_{\mathcal{G}}.
\end{aligned}$$

Next, $\mathsf{Cov}(\xi_{NT}, \zeta_{NT}|F) = \frac{1}{\sqrt{|\mathcal{G}|_0 TN}} \sum_{i\in\mathcal{G}} \sum_j \bar{\lambda}_{\mathcal{G}}' V_{\lambda 1}^{-1} \lambda_j f_t' \Omega_i f_t \mathsf{E}(\omega_{it}\omega_{jt}|F) = o_P(1)$. We can then use the same argument as the proof of Theorem 3 in Bai (2003) to claim that

$$\frac{\frac{1}{\sqrt{T|\mathcal{G}|_0}} \zeta_{NT} + \frac{1}{\sqrt{N}} \xi_{NT}}{\Sigma_{NT}^{1/2}} \to^d \mathcal{N}(0,1).$$

Also, when either $|\mathcal{G}|_0 = o(N)$ or $N = o(T^2)$ holds, $\Sigma_{NT}^{-1/2} O_P(C_{NT}^{-2}) = o_P(1)$. Therefore, $\Sigma_{NT}^{-1/2} \frac{1}{|\mathcal{G}|_0} \sum_{i\in\mathcal{G}} (\widehat{\theta}_{it} - \theta_{it}) \to^d \mathcal{N}(0,1)$.

E.4. **Two special cases.** We now consider two special cases.

Case I. fix an $i \le N$ and set $\mathcal{G} = \{i\}$. In this case we are interested in inference about individual θ_{it}. We have

$$V_{f,i} = \Omega_i \mathsf{E}(\frac{1}{T} \sum_{s=1}^{T} f_s f_s' e_{is}^2 u_{is}^2 | F) \Omega_i, \quad \Sigma_{NT,i} := \frac{1}{T} f_t' V_{f,i} f_t + \frac{1}{N} \lambda_i' V_\lambda \lambda_i.$$

So $\Sigma_{NT,i}^{-1/2} (\widehat{\theta}_{it} - \theta_{it}) \to^d \mathcal{N}(0,1)$.

Case II. $\mathcal{G}=\{1,...,N\}$. Then $|\mathcal{G}|_0=N$. Set $\bar\lambda=\frac{1}{N}\sum_i\lambda_i$. We have

$$\sqrt{N}(\bar\lambda' V_\lambda\bar\lambda)^{-1/2}\frac{1}{N}\sum_{i=1}^N(\widehat\theta_{it}-\theta_{it})\to^d\mathcal{N}(0,1).$$

E.5. **Covariance estimations.** Suppose $C_{NT}^{-1}\max_{is}e_{is}^2u_{is}^2=o_P(1)$. We now consider estimating Σ_{NT} in the general case, with the assumption that $e_{it}u_{it}$ are cross sectionally independent given F. Then

$$V_{\lambda 2}=\frac{1}{N}\sum_i\lambda_i\lambda_i'\mathsf{E}(e_{it}^2u_{it}^2|F),\quad V_f=\frac{1}{T|\mathcal{G}|_0}\sum_{s=1}^T\sum_{i\in\mathcal{G}}\Omega_i f_s f_s'\Omega_i\mathsf{E}(e_{is}^2u_{is}^2|F)$$

The main goal is to consistently estimate the above two quantities. For notational simplicity, we assume $\dim(f_t)=\dim(\alpha_i)=1$.

E.5.1. *Consistency for V_λ.* We have $\widehat V_{\lambda 2,I}:=\frac{1}{N}\sum_i\widehat\lambda_{I,i}\widehat\lambda_{I,i}'\widehat e_{it}^2\widehat u_{it,I}^2$, where $\widehat u_{it,I}=y_{it}-x_{it}\widehat\lambda_{I,i}'\widehat f_{I,t}-\widehat\alpha_{I,i}'\widehat g_{I,t}$.

We remove the subscript "I" for notational simplicity.

Step 1. Show $\|\widehat V_{\lambda 2,I}-\widetilde V_{\lambda 2}\|=o_P(1)$ where $\widetilde V_{\lambda 2}=H_f'^{-1}\frac{1}{N}\sum_i\lambda_i\lambda_i'e_{it}^2u_{it}^2H_f^{-1}$.

By Lemma G.14 $\max_i[\|\widehat\lambda_i-H_f'^{-1}\lambda_i\|+\|\widehat\alpha_i-H_g'^{-1}\alpha_i\|]=o_P(1)$. By Lemma G.17,

$$\frac{1}{N}\sum_i|\widehat u_{it}-u_{it}|^2(1+e_{it}^2)\ \le\ \max_{it}|\widehat u_{it}-u_{it}|^2O_P(1)=o_P(1).$$

So we have

$$\begin{aligned}\|\widehat V_{\lambda 2}-\widetilde V_{\lambda 2}\| &\le \frac{1}{N}\sum_i(\widehat\lambda_i-\lambda_i)^2\widehat e_{it}^2\widehat u_{it}^2+\frac{4}{N}\sum_i|(\widehat\lambda_i-\lambda_i)\lambda_i|(\widehat u_{it}-u_{it})^2\max_i\widehat e_{it}^2\\ &+\frac{4}{N}\sum_i(\widehat\lambda_i-\lambda_i)\lambda_i\widehat e_{it}^2u_{it}^2+\frac{2}{N}\sum_i\lambda_i^2(\widehat e_{it}^2-e_{it}^2)(\widehat u_{it}-u_{it})^2\\ &+\frac{2}{N}\sum_i\lambda_i^2(\widehat e_{it}^2-e_{it}^2)u_{it}^2+\frac{2}{N}\sum_i\lambda_i^2e_{it}^2(\widehat u_{it}-u_{it})u_{it}+\frac{1}{N}\sum_i\lambda_i^2e_{it}^2(\widehat u_{it}-u_{it})^2\\ &\le o_P(1).\end{aligned}$$

Step 2. Show $\|\widetilde V_{\lambda 2}-H_f'^{-1}V_{\lambda 2}H_f^{-1}\|=o_P(1)$.

It suffices to show $\mathsf{Var}(\frac{1}{N}\sum_i\lambda_i^2e_{it}^2u_{it}^2|F)\to 0$ almost surely. Almost surely in F,

$$\mathsf{Var}(\frac{1}{N}\sum_i\lambda_i^2e_{it}^2u_{it}^2|F)\le\frac{\max_i\|\lambda_i\|^2}{N}\max_{i\le N}\mathsf{E}(e_{it}^4u_{it}^4|F)\to 0$$

given that $\max_{i\le N}\mathsf{E}(e_{it}^4u_{it}^4|F)<C$ almost surely.

The consistency for $V_{\lambda 1}$ follows from the same argument.

Therefore

$$\widehat{\lambda}_{I,\mathcal{G}}'\widehat{V}_{\lambda 1,I}^{-1}\widehat{V}_{\lambda 2,I}\widehat{V}_{\lambda 1,I}^{-1}\widehat{\lambda}_{I,\mathcal{G}}-\bar{\lambda}_{\mathcal{G}}'V_\lambda\bar{\lambda}_{\mathcal{G}}=o_P(1)$$

E.5.2. *Consistency of V_f.* Write $\widehat{\Omega}_{I,i}=(\frac{1}{T_0}\sum_{s\notin I}\widehat{f}_{I,s}\widehat{f}_{I,s}')^{-1}\widehat{\sigma}_i^{-2}$ for $\widehat{\sigma}_i^2=\frac{1}{T}\sum_{s=1}^T\widehat{e}_{is}^2$. $\widehat{V}_{f,I}=\frac{1}{T_0|\mathcal{G}|_0}\sum_{s\notin I}\sum_{i\in\mathcal{G}}\widehat{\Omega}_{I,i}\widehat{f}_{I,s}\widehat{f}_{I,s}'\widehat{\Omega}_{I,i}\widehat{e}_{is}^2\widehat{u}_{is,I}^2$. We aim to show the consistency of $\widehat{V}_{f,I}$. We remove the subscript "I" for notational simplicity. Then simply write

$$\begin{aligned}\widehat{V}_f &= \widehat{S}_f^{-1}\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}\widehat{f}_s\widehat{f}_s'\widehat{e}_{is}^2\widehat{u}_{is}^2\widehat{\sigma}_i^{-4}\widehat{S}_f^{-1}\\ \widetilde{V}_f &= H_f'^{-1}S_f^{-1}\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}f_sf_s'e_{is}^2u_{is}^2\sigma_i^{-4}S_f^{-1}H_f^{-1}\end{aligned}$$

where $\sigma_i^2=\mathsf{E}e_{it}^2$.

Step 1. Show $\|\widehat{V}_f-\widetilde{V}_f\|=o_P(1)$. First note that

$$\widehat{S}_f-H_fS_fH_f'=\frac{1}{T}\sum_t\widehat{f}_t\widehat{f}_t'-H_ff_tf_t'H_f'=O_P(T^{-1}+N^{-1/2}).$$

We show $\|\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}\widehat{f}_s\widehat{f}_s'\widehat{e}_{is}^2\widehat{u}_{is}^2\widehat{\sigma}_i^{-4}-H_f\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}f_sf_s'e_{is}^2u_{is}^2\sigma_i^{-4}H_f'\|^2=o_P(1)$. Ignoring H_f, it is bounded by

$$\begin{aligned}&O_P(C_{NT}^{-2})\max_s\|\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\widehat{e}_{is}^2\widehat{u}_{is}^2\widehat{\sigma}_i^{-4}\|^2+O_P(C_{NT}^{-2})\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}\|f_sf_s'\widehat{u}_{is}^2\widehat{\sigma}_i^{-4}\|^2\\&+\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}(\widehat{u}_{is}-u_{is})^2\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}\|f_sf_s'e_{is}^2(\widehat{u}_{is}+u_{is})\widehat{\sigma}_i^{-4}\|^2+O_P(1)\max_i|\widehat{\sigma}_i^4-\sigma_i^4|^2\end{aligned}$$

By Lemma G.16, $\max_{is}|\widehat{u}_{is}-u_{is}|=o_P(1)$, $\max_i|\widehat{\sigma}_i^2-\sigma_i^2|=o_P(1)$, $C_{NT}^{-1}\max_{is}e_{is}^2u_{is}^2=o_P(1)$. So the above is $o_P(1)$.

Step 2. Show $\|\widetilde{V}_f-H_f'^{-1}V_fH_f^{-1}\|=O_P(\frac{1}{\sqrt{T|\mathcal{G}|_0}})$.

It suffices to prove $\mathsf{Var}(\frac{1}{T|\mathcal{G}|_0}\sum_s\sum_{i\in\mathcal{G}}f_s^2e_{is}^2u_{is}^2\sigma_i^{-4}|F)\to 0$ almost surely. It is in fact bounded by

$$\frac{1}{T^2|\mathcal{G}|_0^2}\sum_s\sum_{i\in\mathcal{G}}f_s^4\mathsf{E}(e_{is}^4u_{is}^4|F)\max_i\sigma_i^{-8}=O_P(\frac{1}{T|\mathcal{G}|_0})$$

given that $\frac{1}{T}\sum_t\|f_t\|^4<C$ almost surely. So

$$\widehat{f}_{I,t}'\widehat{V}_{I,f}\widehat{f}_{I,t}-f_t'V_ff_t=O_P(\frac{1}{T}+\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{T|\mathcal{G}|_0}})+o_P(1)=o_P(1).$$

Appendix F. Policy Relevant Tests

F.1. **Group effect dynamics of two periods.** Consider testing

$$H_0 : \bar{\theta}_{\mathcal{G},t_1} = \bar{\theta}_{\mathcal{G},t_2}.$$

This test involves two time periods t_1, t_2, but we do not have to run the algorithm twice to respectively estimate them. Instead, a simple modification of Algorithm 2.2 yields unbiased estimated θ_{it} at both t_1, t_2. The main modification is to replace t by the two periods when doing sample splitting. The modified algorithm is as follows,

Algorithm F.1. Estimate $\theta_{i,t_1}, \theta_{i,t_2}$ as follows.
Step 1: Same as Algorithm 2.2

Step 2: Sample splitting. Randomly split the sample into $\{1, ..., T\}/\{t_1, t_2\} = I \cup I^c$, so that $|I|_0 = [(T-2)/2]$. Denote respectively by Y_I, X_I as the $N \times |I|_0$ matrices of (y_{is}, x_{is}) for observations $s \in I$. Obtain $(\widetilde{\Lambda}, \widetilde{A})$ on the two split sample as in Algorithm 2.2.

Step 3-6. Same as Algorithm 2.2, with $\{t_1, t_2\}$ in place of $\{t\}$. Obtain $(\widehat{\lambda}_{I,i}, \widehat{f}_{I,t})$, $(\widehat{\lambda}_{I^c,i}, \widehat{f}_{I^c,t})$ for $t \in \{t_1, t_2\}$, and hence

$$\begin{aligned} \widehat{\theta}_{it} &:= \frac{1}{2}[\widehat{\lambda}_{I,i}'\widehat{f}_{I,t} + \widehat{\lambda}_{I^c,i}'\widehat{f}_{I^c,t}], \quad t \in \{t_1, t_2\}, \\ \widehat{\theta}_{\mathcal{G},t} &:= \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\widehat{\theta}_{it}. \end{aligned}$$

We shall focus on the case $f_{t_1} \neq f_{t_2}$ for the two periods. If $f_{t_1} = f_{t_2}$, then the problem is not just testing the equality for a particular group, but would be for all individuals. In that case, the test statistic would be simply based on $\widehat{f}_{t_1} - \widehat{f}_{t_2}$, whose asymptotic null distribution is also easy to derive.

Theorem F.1. *Further suppose $N = o(T^2)$, $N = o(T^{3/2}|\mathcal{G}|_0)$ and $\|f_{t_1} - f_{t_2}\| > c > 0$. Then under the null $H_0 : \bar{\theta}_{\mathcal{G},t_1} = \bar{\theta}_{\mathcal{G},t_2}$*

$$\widehat{\sigma}_{NT}^{-1}(\widehat{\theta}_{\mathcal{G},t_1} - \widehat{\theta}_{\mathcal{G},t_2}) \to^d \mathcal{N}(0,1)$$

where

$$\widehat{v}_\lambda = \frac{1}{2}\widehat{\lambda}_{I,\mathcal{G}}'\widehat{V}_{\lambda 1,I}^{-1}\widehat{V}_{\lambda 2,I}\widehat{V}_{\lambda 1,I}^{-1}\widehat{\lambda}_{I,\mathcal{G}} + \frac{1}{2}\widehat{\lambda}_{I^c,\mathcal{G}}'\widehat{V}_{\lambda 1,I^c}^{-1}\widehat{V}_{\lambda 2,I^c}\widehat{V}_{\lambda 1,I^c}^{-1}\widehat{\lambda}_{I^c,\mathcal{G}}$$

$$\begin{aligned}
\widehat{v}_{t_1,t_2} &= \frac{1}{2}(\widehat{f}_{I,t_1}-\widehat{f}_{I,t_2})'\widehat{V}_{I,f}(\widehat{f}_{I,t_1}-\widehat{f}_{I,t_2})+\frac{1}{2}(\widehat{f}_{I^c,t_1}-\widehat{f}_{I^c,t_2})'\widehat{V}_{I^c,f}(\widehat{f}_{I^c,t_1}-\widehat{f}_{I^c,t_2})\\
\widehat{\sigma}^2_{NT,S} &= \frac{1}{T|\mathcal{G}|_0}\widehat{v}_{t_1,t_2}+\frac{2}{N}\widehat{v}_\lambda.
\end{aligned}$$

Proof. The asymptotic analysis of $\widehat{\theta}_{i,t_1}$ and $\widehat{\theta}_{i,t_2}$ is the same as that of $\widehat{\theta}_{i,t}$. Thus the same expansion (F.1) holds for $\{t_1,t_2\}$ in place of $\{t\}$. Therefore, for $\omega_{it}=e_{it}u_{it}$, and $X:=\frac{1}{\sqrt{T|\mathcal{G}|_0}}\sum_{i\in\mathcal{G}}\sum_{s=1}^T\Omega_i f_s\omega_{is}$,

$$\begin{aligned}
\widehat{\theta}_{\mathcal{G},t_1}-\widehat{\theta}_{\mathcal{G},t_2} &= \frac{1}{\sqrt{T|\mathcal{G}|_0}}(f_{t_1}-f_{t_2})'X+\frac{1}{\sqrt{N}}\bar{\lambda}_{\mathcal{G}}'V_{\lambda 1}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N\lambda_i(\omega_{i,t_1}-\omega_{i,t_2})+O_P(C_{NT}^{-2})\\
&:= \zeta_1+\zeta_2+O_P(C_{NT}^{-2}).
\end{aligned}$$

Let

$$\begin{aligned}
\widetilde{\sigma}^2_{NT,1} &= \frac{1}{T|\mathcal{G}|_0}(f_{t_1}-f_{t_2})'V_f(f_{t_1}-f_{t_2}),\quad \widetilde{\sigma}^2_{NT,2}=\frac{2}{N}\bar{\lambda}_{\mathcal{G}}'V_\lambda\bar{\lambda}_{\mathcal{G}}.\\
\widetilde{\sigma}^2_{NT} &= \widetilde{\sigma}^2_{NT,1}+\widetilde{\sigma}^2_{NT,2}.
\end{aligned}$$

We now apply the same argument as the proof of Theorem 3 in Bai (2003) to claim the asymptotic normality of $(\zeta_1+\zeta_2)$. We have $(\widetilde{\sigma}^{-1}_{NT,1}\zeta_1,\widetilde{\sigma}^{-1}_{NT,2}\zeta_2)\to^d\mathcal{N}(0,I_2)$. By the almost sure representation there exist random vectors $(\xi_1,\xi_2)=^d(\widetilde{\sigma}^{-1}_{NT,1}\zeta_1,\widetilde{\sigma}^{-1}_{NT,2}\zeta_2)$, and $(Z_1,Z_2)=^d\mathcal{N}(0,I_2)$ such that $(\xi_1,\xi_2)\to(Z_1,Z_2)$ almost surely. Now

$$\begin{aligned}
\frac{\zeta_1+\zeta_2}{\widetilde{\sigma}_{NT}} &=^d \frac{\widetilde{\sigma}_{NT,1}\xi_1+\widetilde{\sigma}_{NT,2}\xi_2}{\widetilde{\sigma}_{NT}}\\
&= \frac{\widetilde{\sigma}_{NT,1}Z_1+\widetilde{\sigma}_{NT,2}Z_2}{\widetilde{\sigma}_{NT}}+\frac{\widetilde{\sigma}_{NT,1}(\xi_1-Z_1)+\widetilde{\sigma}_{NT,2}(\xi_2-Z_2)}{\widetilde{\sigma}_{NT}}\\
&=^d \mathcal{N}(0,1)+o_P(1)\to^d\mathcal{N}(0,1).
\end{aligned}$$

Next, $|\widetilde{\sigma}^{-1}_{NT}|\leq|(\frac{2}{N}\bar{\lambda}_{\mathcal{G}}'V_\lambda\bar{\lambda}_{\mathcal{G}})^{-1/2}|=O(\sqrt{N})$. So $\widetilde{\sigma}^{-1}_{NT}O_P(C_{NT}^{-2})=o_P(1)$ as long as $N=o(T^2)$. This implies

$$\widetilde{\sigma}^{-1}_{NT}(\widehat{\theta}_{\mathcal{G},t_1}-\widehat{\theta}_{\mathcal{G},t_2})\to^d\mathcal{N}(0,1).$$

To prove $(\widehat{\sigma}^2_{NT,S}-\widetilde{\sigma}^2_{NT,S})/\widetilde{\sigma}^2_{NT,S}=o_P(1)$, we first show $\frac{1}{T|\mathcal{G}|_0}\Delta/\widetilde{\sigma}^2_{NT,S}=o_P(1)$ where

$$\Delta:=\widehat{v}_{t_1,t_2}-(f_{t_1}-f_{t_2})'V_f(f_{t_1}-f_{t_2}),$$

by following the proof for the consistency of V_f in Section E.5.2. But note that here the consistency of $\widehat{v}_{t_1,t_2}$ is no longer sufficient, because under the null, it is likely $f_{t_1} = f_{t_2}$, in which case $\widetilde{\sigma}^2_{NT} = O(N^{-1})$. Then we need to show

$$\frac{N}{T|\mathcal{G}|_0}\Delta = o_P(1).$$

To obtain a sharper bound for Δ, we refine the proof of Section E.5.2. We assume serial homoskedasticity and redefine, for $\widehat{\sigma}^2_i = \frac{1}{T}\sum_{s=1}^T \widehat{e}^2_{is}$, $\widehat{\sigma}^2_{ui} = \frac{1}{T}\sum_{s=1}^T \widehat{u}^2_{is}$

$$\widehat{V}_f = \widehat{S}_f^{-1}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\widehat{\sigma}^2_{ui}\widehat{\sigma}_i^{-2}, \quad \widetilde{V}_f = H_f'^{-1}S_f^{-1}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\mathsf{E}u^2_{is}\sigma_i^{-2}H_f^{-1}.$$

Then $\|\widehat{V}_f - \widetilde{V}_f\| = O_P(C_{NT}^{-1})$ and $\|\widetilde{V}_f - H_f'^{-1}V_fH_f^{-1}\| = O_P(\frac{1}{\sqrt{T|\mathcal{G}|_0}})$. Hence $\widehat{f}_{I,t_1} - \widehat{f}_{I,t_2})'\widehat{V}_{I,f}(\widehat{f}_{I,t_1} - \widehat{f}_{I,t_2}) - (f_{t_1} - f_{t_2})'V_f(f_{t_1} - f_{t_2}) = O_P(C_{NT}^{-1})$, implying $\Delta = O_P(C_{NT}^{-1})$. Hence $\frac{N}{T|\mathcal{G}|_0}\Delta = o_P(1)$ holds so long as $N = o(T^{3/2}|\mathcal{G}|_0)$.

Next, $\frac{1}{N}(\widehat{v}_\lambda - \bar{\lambda}_{\mathcal{G}}'V_\lambda\bar{\lambda}_{\mathcal{G}})/\widetilde{\sigma}^2_{NT,S} = o_P(1)$ is straightforward, where the consistency of $\widehat{v}_\lambda$ is sufficient. Hence $(\widehat{\sigma}^2_{NT,S} - \widetilde{\sigma}^2_{NT,S})/\widetilde{\sigma}^2_{NT,S} = o_P(1)$. Therefore under the null

$$\widehat{\sigma}_{NT}^{-1}(\widehat{\theta}_{\mathcal{G},t_1} - \widehat{\theta}_{\mathcal{G},t_2}) \to^d \mathcal{N}(0,1).$$

□

F.2. **Test of group homogeneity.** We now study the test of

$$H_0 : \bar{\theta}_{\mathcal{G}_1,t} = ... = \bar{\theta}_{\mathcal{G}_J,t}.$$

Under the null, it is assumed that $\bar{\lambda}_{\mathcal{G}_1} = ... = \bar{\lambda}_{\mathcal{G}_J}$. Let $\widehat{\theta}_{\mathcal{G}_j,t}$ be the estimated $\bar{\theta}_{\mathcal{G}_j,t}$. Because J is assumed to be fixed, the expansion (F.1) also holds jointly for $\{\widehat{\theta}_{\mathcal{G}_j,t} : j \leq J\}$. Therefore,

$$\widehat{\theta}_{\mathcal{G}_j,t} - \bar{\theta}_{\mathcal{G}_j,t} = \frac{1}{|\mathcal{G}_j|_0\sqrt{T}}f_t'\sum_{i\in\mathcal{G}_j}\Omega_i m_i + \frac{1}{\sqrt{N}}\bar{\lambda}_{\mathcal{G}_j}'A + O_P(C_{NT}^{-2}).$$

Putting together, jointly we have

$$\Delta := \begin{pmatrix}\widehat{\theta}_{\mathcal{G}_1,t} - \bar{\theta}_{\mathcal{G}_1,t}\\ \vdots \\ \widehat{\theta}_{\mathcal{G}_J,t} - \bar{\theta}_{\mathcal{G}_J,t}\end{pmatrix} = \frac{1}{\sqrt{T}}[I_J \otimes f_t']M + \frac{1}{\sqrt{N}}\bar{\Lambda}A_t + O_P(C_{NT}^{-2}) \qquad \text{(F.1)}$$

where

$$M = \begin{pmatrix}\frac{1}{|\mathcal{G}_1|_0}\sum_{i\in\mathcal{G}_1}\Omega_i m_i\\ \vdots \\ \frac{1}{|\mathcal{G}_J|_0}\sum_{i\in\mathcal{G}_J}\Omega_i m_i\end{pmatrix}, \quad \bar{\Lambda} = \begin{pmatrix}\bar{\lambda}_{\mathcal{G}_1}'\\ \vdots \\ \bar{\lambda}_{\mathcal{G}_J}'\end{pmatrix}$$

$$A_t = V_{\lambda 1}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{it}u_{it}, \quad m_i = \frac{1}{\sqrt{T}}\sum_{s=1}^{T} f_s e_{is}u_{is}. \tag{F.2}$$

We further assume that if i, j belong two groups, then $e_{it}u_{it}$ are independent given F and groups are mutually disjoint. Then $\mathsf{Var}(\frac{1}{\sqrt{T}}[I_J \otimes f_t']M|F)$ is a block diagonal matrix:

$$\mathsf{Var}(\frac{1}{\sqrt{T}}[I_J \otimes f_t']M|F) = \frac{1}{T}\begin{pmatrix} \frac{1}{|\mathcal{G}_1|_0} f_t' V_{f,\mathcal{G}_1} f_t & 0 & \cdots \\ & \ddots & \\ \cdots & 0 & \frac{1}{|\mathcal{G}_J|_0} f_t' V_{f,\mathcal{G}_J} f_t \end{pmatrix} := \frac{1}{T}D$$

where $V_{f,\mathcal{G}_j} = \frac{1}{T}\sum_{s=1}^{T}\mathsf{Var}\left(\frac{1}{\sqrt{|\mathcal{G}_j|_0}}\sum_{i\in\mathcal{G}_j}\Omega_i f_s e_{is}u_{is}\middle| F\right)$. Additionally, let the $(J-1)\times J$ matrix selection matrix be:

$$\Xi = \begin{pmatrix} 1 & -1 & 0 & \cdots & & \\ 0 & 1 & -1 & 0 & \cdots & \\ \vdots & & & & & \\ 0 & \cdots & & 1 & -1 & \end{pmatrix}, \quad \widehat{D}_S = \begin{pmatrix} \frac{1}{|\mathcal{G}_1|_0}\widehat{f}_{S,t}'\widehat{V}_{fS,\mathcal{G}_1}\widehat{f}_{S,t} & 0 & \cdots \\ & \ddots & \\ \cdots & 0 & \frac{1}{|\mathcal{G}_J|_0}\widehat{f}_{S,t}'\widehat{V}_{fS,\mathcal{G}_J}\widehat{f}_{S,t} \end{pmatrix}$$

for $S = I$ and I^c. Let $\widehat{D} = \frac{1}{2}(\widehat{D}_I + \widehat{D}_{I^c})$. We note that $\Xi\bar{\Lambda} = 0$ and $\Xi(\bar{\theta}_{\mathcal{G}_1,t}, ..., \bar{\theta}_{\mathcal{G}_J,t})' = 0$ under the null. So

$$\begin{aligned}\sqrt{T}(\Xi D\Xi')^{-1/2}\Xi\Delta &= \sqrt{T}(\Xi D\Xi')^{-1/2}\Xi\frac{1}{\sqrt{T}}[I_J \otimes f_t']M \\ &\quad + \sqrt{T}(\Xi D\Xi')^{-1/2}\Xi O_P(C_{NT}^{-2}).\end{aligned} \tag{F.3}$$

It remains to show three statements:
(i) $\sqrt{T}(\Xi D\Xi')^{-1/2}\Xi O_P(C_{NT}^{-2}) = o_P(1)$.
(ii) $\sqrt{T}(\Xi D\Xi')^{-1/2}\Xi\frac{1}{\sqrt{T}}[I_J \otimes f_t']M|F \to^d \mathcal{N}(0, I)$.
(iii) $[(\Xi\widehat{D}\Xi')^{-1/2} - (\Xi D\Xi')^{-1/2}](\Xi D\Xi')^{1/2} = o_P(1)$.

Recalling $C_{NT}^{-2} = O(1/N + 1/T)$, $\lambda_{\min}(\Xi\Xi') \geq c$, and $\min_{j\leq J}\lambda_{\min}(V_{f,\mathcal{G}_j}) > c$. Hence

$$\|(\Xi D\Xi')^{-1/2}\| \leq \sqrt{\frac{1}{\lambda_{\min}(\Xi\Xi')\min_{j\leq J}(\frac{1}{|\mathcal{G}_j|_0}f_t'V_{f,\mathcal{G}_j}f_t)}} \leq O_P(\|f_t\|^{-1})\sqrt{\max_{j\leq J}|\mathcal{G}_j|_0}$$

This implies (i) when both $\max_{j\leq J}|\mathcal{G}_j|_0 = o(T)$ and $T\max_{j\leq J}|\mathcal{G}_j|_0 = o(N^2)$. In addition, it is straightforward to check the multivariate Lindeberg-Feller condition, so that (ii) holds. As for (iii), let $\widehat{B} = \Xi\widehat{D}\Xi'$ and $B = \Xi D\Xi'$. Both matrices are positive definite then

$$\|\widehat{B}-B\| = O_P(\frac{1}{\min_{j\leq J}|\mathcal{G}_j|_0})\max_j \|\widehat{f}_t'\widehat{V}_{f,\mathcal{G}_J}\widehat{f}_t - f_t'V_{f,\mathcal{G}_J}f_t\| = o_P(\frac{1}{\min_{j\leq J}|\mathcal{G}_j|_0}).$$

Hence with the assumption that $\max_j|\mathcal{G}_j|_0/\min_j|\mathcal{G}_j|_0 = O(1)$,

$$\begin{aligned}\|\widehat{B}^{-1/2}-B^{-1/2}\| &\leq \frac{1}{\sqrt{\lambda_{\max}^{-1}(\widehat{B})}+\sqrt{\lambda_{\max}^{-1}(B)}}\|\widehat{B}^{-1}-B^{-1}\| \\ &\leq \sqrt{\lambda_{\min}(B)}\|\widehat{B}-B\|\|\widehat{B}^{-1}\|\|B^{-1}\|,\end{aligned}$$

so $[\widehat{B}^{-1/2}-B^{-1/2}]B^{1/2} = o_P(\max_j|\mathcal{G}_j|_0/\min_j|\mathcal{G}_j|_0)^{3/2} = o_P(1)$. This gives (iii). (F.3) and (i)(ii) together imply $\sqrt{T}(\Xi D\Xi')^{-1/2}\Xi\Delta \to^d \mathcal{N}(0,I)$. This, combined with (iii), implies under the null,

$$\sqrt{T}(\Xi\widehat{D}\Xi')^{-1/2}\Xi\begin{pmatrix}\widehat{\theta}_{\mathcal{G}_1,t}\\ \vdots \\ \widehat{\theta}_{\mathcal{G}_J,t}\end{pmatrix} \to^d \mathcal{N}(0,I_{J-1}).$$

This leads to the desired asymptotic null distribution:

$$T(\widehat{\theta}_{\mathcal{G}_1,t},...,\widehat{\theta}_{\mathcal{G}_J,t})'\Xi'(\Xi\widehat{D}\Xi')^{-1}\Xi\begin{pmatrix}\widehat{\theta}_{\mathcal{G}_1,t}\\ \vdots \\ \widehat{\theta}_{\mathcal{G}_J,t}\end{pmatrix} \to^d \chi^2(J-1).$$

F.3. **Test of joint significance in a given time period.** Consider testing:

$$H_0 : \theta_{it} = 0 \text{ for all } i.$$

With the factor slope structure, we assume that H_0 is equivalent to

$$H_0' : f_t = 0.$$

This becomes a more tractable low dimensional testing problem. By Proposition E.1, for both $S = I$ and $S = I^c$,

$$\widehat{f}_{S,t} - H_{S,f}f_t = H_{S,f}V_{\lambda 1}^{-1}\frac{1}{N}\sum_i \lambda_i e_{it}u_{it} + O_P(C_{NT}^{-2}).$$

Then the above expansion still holds with $\widehat{f}_{S,t}$ and $H_{S,f}$ replaced by $\frac{1}{2}(\widehat{f}_{I,t}+\widehat{f}_{I^c,t})$ and $\bar{H}_f = \frac{1}{2}H_{I,f}+\frac{1}{2}H_{I^c,f}$.

We first derive the probability limit of $\bar{H}_f$. In the proof of Proposition E.1, $H_{S,f} := (\bar{H}_3+\mathsf{I})H_1^{-1} = H_1^{-1}+o_P(1)$ because $\bar{H}_3 = o_P(1)$. In addition, the definition of H_1 is given in step 2 of the proof of Proposition D.1, which only depends on (Λ, F). Thus $\|\bar{H}_f - H_1^{-1}\| = o_P(1)$, and H_1^{-1} is nonrandom conditionally on F. We have, under H_0', conditionally on F, when $N = o(T^2)$,

$$\sqrt{N}(H_1^{-1}V_\lambda H_1'^{-1})^{-1/2}\frac{1}{2}(\widehat{f}_{I,t} + \widehat{f}_{I^c,t}) \to^d \mathcal{N}(0, I) \tag{F.4}$$

which then also holds unconditionally using the dominated convergence theorem. Additionally, the consistency proof in Section E.5.1 shows $\|\widehat{V}_{\lambda 2,S} - H_f'^{-1}V_{\lambda 2}H_f^{-1}\| = o_P(1)$ and $\|\widehat{V}_{\lambda 1,S} - H_f'^{-1}V_{\lambda 1}H_f^{-1}\| = o_P(1)$ for both $S = I$ and $S = I^c$. Therefore for

$$\widehat{V}_\lambda := \frac{1}{2}\widehat{V}_{\lambda 1,I}^{-1}\widehat{V}_{\lambda 2,I}\widehat{V}_{\lambda 1,I}^{-1} + \frac{1}{2}\widehat{V}_{\lambda 1,I^c}^{-1}\widehat{V}_{\lambda 2,I^c}\widehat{V}_{\lambda 1,I^c}^{-1}$$

we have $\|\widehat{V}_\lambda - H_1^{-1}V_\lambda H_1'^{-1}\| = o_P(1)$. Next, the eigenvalues of $H_1^{-1}V_\lambda H_1'^{-1}$ are bounded away from zero, thus (F.4) still holds with $\widehat{V}_\lambda$ in place of $H_1^{-1}V_\lambda H_1'^{-1}$. This yields

$$N\frac{1}{2}(\widehat{f}_{I,t} + \widehat{f}_{I^c,t})'\widehat{V}_\lambda^{-1}\frac{1}{2}(\widehat{f}_{I,t} + \widehat{f}_{I^c,t}) \to^d \chi^2(\dim(f_t)).$$

Appendix G. Technical lemmas

G.1. **The effect of $\widehat{e}_{it} - e_{it}$ in the factor model.** Here we present the intermediate results when x_{it} admits a factor structure:

$$x_{it} = l_i' w_t + e_{it}.$$

Let $\widehat{w}_t$ be the PC estimator of w_t. Then $\widehat{l}_i' = \frac{1}{T}\sum_s x_{is}\widehat{w}_s'$ and $\widehat{e}_{it} - e_{it} = l_i' H_x(\widehat{w}_t - H_x^{-1}w_t) + (\widehat{l}_i' - l_i' H_x)\widehat{w}_t$.

Lemma G.1. *(i)* $\max_{it}|\widehat{e}_{it} - e_{it}| = O_P(\phi_{NT})$, *where*

$$\phi_{NT} := (C_{NT}^{-2}(\max_i \frac{1}{T}\sum_s e_{is}^2)^{1/2} + b_{NT,4} + b_{NT,5})(1 + \max_{t\leq T}\|w_t\|) + (b_{NT,1} + C_{NT}^{-1}b_{NT,2} + b_{NT,3}).$$

So $\max_{it}|\widehat{e}_{it} - e_{it}|\max_{it}|e_{it}| = O_P(1)$.

(ii) All terms below are $O_P(C_{NT}^{-2})$, *for a fixed* t: $\frac{1}{N}\sum_i(\widehat{e}_{it} - e_{it})^4$, $\frac{1}{N}\sum_i(\widehat{e}_{it} - e_{it})^2$, *and* $\frac{1}{N}\sum_i(\widehat{e}_{it} - e_{it})^2 e_{it}^2$, $\frac{1}{N}\sum_i \lambda_i\lambda_i'(\widehat{e}_{it} - e_{it})e_{it}$, $\frac{1}{N}\sum_i \lambda_i(\widehat{e}_{it} - e_{it})u_{it}$.

(iii) $\frac{1}{N}\sum_i \lambda_i\alpha_i'(\widehat{e}_{it} - e_{it}) = O_P(C_{NT}^{-1})$ *for a fixed* t.

Proof. Below we first simplify the expansion of $\widehat{e}_{it} - e_{it}$. Let $K_3 = \dim(l_i)$. Let $\mathcal{Q}$ be a diagonal matrix consisting of the reciprocal of the first K_3 eigenvalues of

$XX'/(NT)$. Let

$$\begin{aligned}\zeta_{st} &= \frac{1}{N}\sum_i (e_{is}e_{it} - \mathsf{E}e_{is}e_{it}), \quad \eta_t = \frac{1}{N}\sum_i l_i e_{it},\\ \sigma^2 &= \frac{1}{N}\sum_i \mathsf{E}e_{it}^2.\end{aligned}$$

For the PC estimator, there is a rotation matrix $\bar{H}_x$, by (A.1) of Bai (2003), (which can be simplified due to the serial independence in e_{it})

$$\begin{aligned}\widehat{w}_t - \bar{H}_x w_t &= \mathcal{Q}\frac{\sigma^2}{T}(\widehat{w}_t - \bar{H}_x w_t) + \mathcal{Q}[\frac{\sigma^2}{T}\bar{H}_x + \frac{1}{TN}\sum_{is}\widehat{w}_s l_i' e_{is}]w_t\\ &\quad + \mathcal{Q}\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st} + \mathcal{Q}\frac{1}{T}\sum_s \widehat{w}_s w_s'\eta_t.\end{aligned}$$

Move $\mathcal{Q}\frac{\sigma^2}{T}(\widehat{w}_t - \bar{H}_x w_t)$ to the left hand side (LHS); then LHS becomes $(\mathsf{I} - \mathcal{Q}\frac{\sigma^2}{T})(\widehat{w}_t - \bar{H}_x w_t)$. Note that $\|\mathcal{Q}\| = O_P(1)$ so $\mathcal{Q}_1 := (\mathsf{I} - \mathcal{Q}\frac{\sigma^2}{T})^{-1}$ exists whose eigenvalues all converge to one. Then multiply $\mathcal{Q}_1$ on both sides, we reach

$$\begin{aligned}\widehat{w}_t - \bar{H}_x w_t &= \mathcal{Q}_1\mathcal{Q}[\frac{\sigma^2}{T}\bar{H}_x + \frac{1}{TN}\sum_{is}\widehat{w}_s l_i' e_{is}]w_t\\ &\quad + \mathcal{Q}_1\mathcal{Q}\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st} + \mathcal{Q}_1\mathcal{Q}\frac{1}{T}\sum_s \widehat{w}_s w_s'\eta_t.\end{aligned}$$

Next, move $\mathcal{Q}_1\mathcal{Q}[\frac{\sigma^2}{T}\bar{H}_x + \frac{1}{TN}\sum_{is}\widehat{w}_s l_i' e_{is}]w_t$ to LHS, combined with $-\bar{H}_x w_t$, then LHS becomes $\widehat{w}_t - H_x^{-1}w_t$, where $H_x^{-1} = (\mathsf{I} + \mathcal{Q}_1\mathcal{Q}\frac{\sigma^2}{T})\bar{H}_x + \mathcal{Q}_1\mathcal{Q}\frac{1}{TN}\sum_{is}\widehat{w}_s l_i' e_{is}$, and $\mathcal{Q}_1\mathcal{Q}\frac{1}{TN}\sum_{is}\widehat{w}_s l_i' e_{is} = o_P(1)$. So the eigenvalues of H_x^{-1} converge to those of $\bar{H}_x$, which are well known to be bounded away from both zero and infinity (Bai, 2003). Finally, let $R_1 = \mathcal{Q}_1\mathcal{Q}$ and $R_2 = \mathcal{Q}_1\mathcal{Q}\frac{1}{T}\sum_s \widehat{w}_s w_s'$, we reach

$$\widehat{w}_t - H_x^{-1}w_t = R_1\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st} + R_2\eta_t. \tag{G.1}$$

with $\|R_1\| + \|R_2\| = O_P(1)$.

Also, $\widehat{l}_i' = \frac{1}{T}\sum_s x_{is}\widehat{w}_s'$, for $\mathcal{Q}_3 = -H_x(R_1\frac{1}{T^2}\sum_{m,s\le T}\zeta_{ms}\widehat{w}_m\widehat{w}_s' + R_2\frac{1}{T}\sum_s \eta_s\widehat{w}_s') = O_P(C_{NT}^{-2})$,

$$\begin{aligned}\widehat{e}_{it} - e_{it} &= l_i' H_x(\widehat{w}_t - H_x^{-1}w_t) + (\widehat{l}_i' - l_i' H_x)\widehat{w}_t\\ &= l_i' H_x(\widehat{w}_t - H_x^{-1}w_t) + \frac{1}{T}\sum_s e_{is}(\widehat{w}_s' - w_s' H_x^{-1\prime})\widehat{w}_t\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{T}\sum_s e_{is}w_s' H_x^{-1'}\widehat{w}_t + l_i' H_x \frac{1}{T}\sum_s (H_x^{-1}w_s - \widehat{w}_s)\widehat{w}_s'\widehat{w}_t \\
= \;& \frac{1}{T}\sum_s e_{is}w_s' H_x^{-1'}\widehat{w}_t + \frac{1}{T^2}\sum_{s,m\le T} e_{is}\zeta_{ms}\widehat{w}_m' R_1'\widehat{w}_t + \frac{1}{T}\sum_s e_{is}\eta_s' R_2'\widehat{w}_t \\
&+ l_i'\mathcal{Q}_3\widehat{w}_t + l_i' H_x R_1 \frac{1}{T}\sum_s \widehat{w}_s\zeta_{st} + l_i' H_x R_2 \eta_t. \qquad \text{(G.2)}
\end{aligned}
$$

(i) We first show that $\max_{t\le T}\|\widehat{w}_t - H_x^{-1}w_t\| = O_P(1)$. Define

$$
\begin{aligned}
b_{NT,1} &= \max_{t\le T}\|\frac{1}{NT}\sum_{is} w_s(e_{is}e_{it} - \mathsf{E}e_{is}e_{it})\| \\
b_{NT,2} &= (\max_{t\le T}\frac{1}{T}\sum_s(\frac{1}{N}\sum_i e_{is}e_{it} - \mathsf{E}e_{is}e_{it})^2)^{1/2} \\
b_{NT,3} &= \max_{t\le T}\|\frac{1}{N}\sum_i l_i e_{it}\| \\
b_{NT,4} &= \max_i \|\frac{1}{T}\sum_s e_{is}w_s\| \\
b_{NT,5} &= \max_i \|\frac{1}{NT}\sum_{js} l_j(e_{js}e_{is} - \mathsf{E}e_{js}e_{is})\|
\end{aligned}
$$

Then

$$
\begin{aligned}
\max_{t\le T}\|\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st}\| &\le O_P(1)\max_{t\le T}\|\frac{1}{T}\sum_s w_s\zeta_{st}\| + O_P(C_{NT}^{-1})(\max_{t\le T}\frac{1}{T}\sum_s \zeta_{st}^2)^{1/2} \\
&= O_P(b_{NT,1} + C_{NT}^{-1}b_{NT,2}).
\end{aligned}
$$

Then by assumption, $\max_{t\le T}\|\widehat{w}_t - H_x^{-1}w_t\| = O_P(b_{NT,1} + C_{NT}^{-1}b_{NT,2} + b_{NT,3}) = O_P(1)$. As such $\max_{t\le T}\|\widehat{w}_t\| = O_P(1) + \max_{t\le T}\|w_t\|$.

In addition,

$$
\begin{aligned}
&\max_i \|\frac{1}{T^2}\sum_{m\le T}\sum_{s\le T} e_{is}\zeta_{ms}\widehat{w}_m'\| \le O_P(C_{NT}^{-1})(\max_i \frac{1}{T}\sum_{s\le T} e_{is}^2)^{1/2}(\frac{1}{T^2}\sum_{s,t\le T}\zeta_{st}^2)^{1/2} \\
&+O_P(1)(\max_i \frac{1}{T}\sum_s e_{is}^2)^{1/2}(\frac{1}{T}\sum_s \|\frac{1}{TN}\sum_{jt}(e_{jt}e_{js} - \mathsf{E}e_{jt}e_{js})w_t\|^2)^{1/2} \\
= \;& O_P(C_{NT}^{-2})(\max_i \frac{1}{T}\sum_s e_{is}^2)^{1/2}.
\end{aligned}
$$

So $\max_{it}|\widehat{e}_{it} - e_{it}| = O_P(\phi_{NT})$, where

$$
\phi_{NT} := (C_{NT}^{-2}(\max_i \frac{1}{T}\sum_s e_{is}^2)^{1/2} + b_{NT,4} + b_{NT,5})(1 + \max_{t\le T}\|w_t\|) + (b_{NT,1} + C_{NT}^{-1}b_{NT,2} + b_{NT,3}).
$$

(ii) Let $a \in \{1,2,4\}$, and $b \in \{0,1,2\}$, and a bounded constant sequence c_i, consider, up to a $O_P(1)$ multiplier that is independent of (t,i),

$$\begin{aligned}
\frac{1}{N}\sum_i c_i e_{it}^b(\widehat{e}_{it}-e_{it})^a &= \frac{1}{N}\sum_i c_i e_{it}^b(\frac{1}{T}\sum_s e_{is}w_s)^a\widehat{w}_t^a \\
&+\frac{1}{N}\sum_i c_i e_{it}^b(\frac{1}{T^2}\sum_{s,m\le T} e_{is}\zeta_{ms}\widehat{w}_m)^a\widehat{w}_t^a \\
&+\frac{1}{N}\sum_i c_i e_{it}^b(\frac{1}{T}\sum_s e_{is}\eta_s')^a\widehat{w}_t^a+\frac{1}{N}\sum_i c_i e_{it}^b l_i^a\mathcal{Q}_3^a\widehat{w}_t^a \\
&+\frac{1}{N}\sum_i c_i e_{it}^b l_i^a(\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st})^a+\frac{1}{N}\sum_i c_i e_{it}^b l_i^a\eta_t^a. \quad \text{(G.3)}
\end{aligned}$$

Note that

$$\begin{aligned}
\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st} &\le O_P(1)\frac{1}{TN}\sum_{is} w_s(e_{is}e_{it}-\mathsf{E}e_{is}e_{it}) \\
&\quad +(\frac{1}{T}\sum_s(\frac{1}{N}\sum_i(e_{is}e_{it}-\mathsf{E}e_{is}e_{it}))^2)^{1/2}O_P(C_{NT}^{-1})=O_P(C_{NT}^{-2}) \\
\frac{1}{N}\sum_i e_{it}^b(\frac{1}{T}\sum_s e_{is}w_s)^a &= O(T^{-a/2}),\quad a=2,4,\quad b=0,2 \\
\frac{1}{N}\sum_i(\frac{1}{T^2}\sum_{s,m\le T} e_{is}\zeta_{ms}\widehat{w}_m)^4 &\le (\frac{1}{T^2}\sum_{m,s\le T}\zeta_{ms}^2)^2(\frac{1}{T}\sum_m w_m^2)^2\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}^2)^2 \\
&\quad +\frac{1}{N}\sum_i(\frac{1}{T}\sum_{m\le T}(\frac{1}{T}\sum_{s\le T} e_{is}\zeta_{ms})^2)^2 O_P(C_{NT}^{-4})=O_P(C_{NT}^{-4}) \\
\frac{1}{N}\sum_i(\frac{1}{T^2}\sum_{s,m\le T} e_{is}\zeta_{ms}\widehat{w}_m)^2 &\le \frac{1}{N}\sum_i\frac{1}{T}\sum_m(\frac{1}{T}\sum_s e_{is}\zeta_{ms})^2\frac{1}{T}\sum_m w_m^2 \\
&\quad +\frac{1}{N}\sum_i\frac{1}{T}\sum_{m\le T}(\frac{1}{T}\sum_{s\le T} e_{is}\zeta_{ms})^2 O_P(C_{NT}^{-2})=O_P(C_{NT}^{-2}) \\
\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}\eta_s')^4 &\le \frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}^2)^2(\frac{1}{T}\sum_t \eta_t^2)^2=O_P(C_{NT}^{-4}) \\
\frac{1}{N}\sum_i e_{it}^2(\frac{1}{T}\sum_s e_{is}\eta_s')^2 &= O_P(C_{NT}^{-2}) \\
\frac{1}{N}\sum_i e_{it}^2 l_i^2(\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st})^2 &\le \frac{1}{N}\sum_i e_{it}^2 l_i^2(\frac{1}{T}\sum_s w_s\zeta_{st})^2+O_P(C_{NT}^{-2})=O_P(C_{NT}^{-2}) \\
\frac{1}{T}\sum_t(\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st})^2 &\le O_P(C_{NT}^{-4}) \\
\frac{1}{N}\sum_i c_i e_{it}(\frac{1}{T}\sum_s e_{is}w_s) &= O_P(1)(\frac{1}{N^2T^2}\sum_{ij\le N}\sum_{sl\le T} c_ic_j\mathsf{E}w_sw_l\mathsf{E}(e_{it}e_{jt}e_{is}e_{jl}|W))^{1/2}
\end{aligned}$$

$$
\begin{aligned}
& = O_P(C_{NT}^{-2}) \\
\frac{1}{N}\sum_i c_i e_{it}(\frac{1}{T^2}\sum_{s,m\leq T} e_{is}\zeta_{ms}\widehat{w}_m) & \leq O_P(C_{NT}^{-2}) + \frac{1}{N}\sum_i c_i e_{it}(\frac{1}{T}\sum_{s\leq T}\frac{1}{T}\sum_{m\leq T} e_{is}\zeta_{ms}w_m) \\
& \leq O_P(C_{NT}^{-2}) + O_P(1)(\frac{1}{T}\sum_{s\leq T}\mathsf{E}(\frac{1}{T}\sum_{m\leq T}\zeta_{ms}w_m)^2)^{1/2} \\
& = O_P(C_{NT}^{-2}).
\end{aligned}
$$

where the last equality follows from the following:

$$
\begin{aligned}
& \frac{1}{T}\sum_{s\leq T}\mathsf{E}(\frac{1}{T}\sum_{m\leq T}\zeta_{ms}w_m)^2 = O(T^{-2}) + \frac{1}{T}\sum_{s\leq T}\frac{1}{T^2}\sum_{t\neq s}\mathsf{E}w_s w_t \mathsf{Cov}(\zeta_{ss},\zeta_{ts}|W) \\
& +\frac{1}{T}\sum_{s\leq T}\frac{1}{T}\sum_{m\neq s}\frac{1}{T}\sum_{t\leq T}\mathsf{E}w_m w_t\mathsf{Cov}(\zeta_{ms},\zeta_{ts}|W) \\
= \;& O(T^{-2}) + \frac{1}{T}\sum_{s\leq T}\frac{1}{T}\sum_{t\neq s}\frac{1}{T}\mathsf{E}w_t^2\frac{1}{N^2}\sum_{ij}\mathsf{E}(e_{is}e_{js}|W)\mathsf{E}(e_{it}e_{jt}|W) = O(C_{NT}^{-4}).
\end{aligned}
$$

With the above results ready, we can proceed proving (ii)(iii) as follows.

Now for $a=4, b=0$, $c_i=1$, up to a $O_P(1)$ multiplier

$$
\begin{aligned}
\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4 & = \frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}w_s)^4 + \frac{1}{N}\sum_i(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^4 \\
& +\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}\eta_s')^4 + \mathcal{Q}_3^4 + (\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^4 + \eta_t^4 \\
& \leq O_P(C_{NT}^{-4}).
\end{aligned}
$$

For $a=2$, $b=0$, $c_i=1$,

$$
\begin{aligned}
\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2 & \leq \frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}w_s)^2 + \frac{1}{N}\sum_i(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^2 \\
& +\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}\eta_s')^2 + \mathcal{Q}_3^2 + (\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2 + \eta_t^2 \\
& \leq O_P(C_{NT}^{-2}).
\end{aligned}
$$

For $a=2$, $b=2$, $c_i=1$,

$$
\begin{aligned}
\frac{1}{N}\sum_i e_{it}^2(\widehat{e}_{it}-e_{it})^2 & = \frac{1}{N}\sum_i e_{it}^2(\frac{1}{T}\sum_s e_{is}w_s)^2 + \frac{1}{N}\sum_i e_{it}^2(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^2 \\
& +\frac{1}{N}\sum_i e_{it}^2(\frac{1}{T}\sum_s e_{is}\eta_s')^2 + \mathcal{Q}_3^2 + (\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2 + \eta_t^2 \\
& \leq O_P(C_{NT}^{-2}).
\end{aligned}
$$

Next, let $a=b=1$ and c_i be any element of $\lambda_i\lambda_i'$,

$$\begin{aligned}\frac{1}{N}\sum_i c_ie_{it}(\widehat{e}_{it}-e_{it}) &= \frac{1}{N}\sum_i c_ie_{it}(\frac{1}{T}\sum_s e_{is}w_s)+\frac{1}{N}\sum_i c_ie_{it}(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)\\&+\frac{1}{N}\sum_i c_ie_{it}(\frac{1}{T}\sum_s e_{is}\eta_s')+\frac{1}{N}\sum_i c_ie_{it}l_i\mathcal{Q}_3\\&+\frac{1}{N}\sum_i c_ie_{it}l_i(\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st})+\frac{1}{N}\sum_i c_ie_{it}l_i\eta_t\\&\leq O_P(C_{NT}^{-2}).\end{aligned}$$

Next, ignoring an $O_P(1)$ multiplier,

$$\begin{aligned}\frac{1}{N}\sum_i \lambda_i(\widehat{e}_{it}-e_{it})u_{it} &\leq \frac{1}{N}\sum_i\lambda_i\frac{1}{T}\sum_s e_{is}w_su_{it}\\&+\frac{1}{N}\sum_i\lambda_i\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_mu_{it}+\frac{1}{N}\sum_i\lambda_i\frac{1}{T}\sum_s e_{is}\eta_su_{it}\\&+\frac{1}{N}\sum_i\lambda_il_i'\mathcal{Q}_3u_{it}+\frac{1}{N}\sum_i\lambda_il_iu_{it}\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st}+\frac{1}{N}\sum_i\lambda_il_i\eta_tu_{it}\\&\leq O_P(C_{NT}^{-2}).\end{aligned}$$

(iii) Let $a=1,b=0$ and c_i be any element of $\lambda_i\alpha_i'$,

$$\begin{aligned}\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it}) &= \frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}w_s)+\frac{1}{N}\sum_i c_i(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)\\&+\frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}\eta_s')+\mathcal{Q}_3+(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})+\eta_t\\&\leq O_P(C_{NT}^{-1}),\end{aligned}$$

where the dominating term is $\eta_t=O_P(C_{NT}^{-1})$.

$\square$

Lemma G.2. *Assume* $\max_{it}|e_{it}|C_{NT}^{-1}=O_P(1)$ *and* $\mathsf{E}e_{it}^8<C$.

Let c_i *be a non-random bounded sequence.*

(i) $\max_{t\leq T}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4=O_P(1+\max_{t\leq T}\|w_t\|^4+b_{NT,2}^4)C_{NT}^{-4}+O_P(b_{NT,1}^4+b_{NT,3}^4)$.

$\max_{t\leq T}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2e_{it}^2\leq O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)\max_{it}|e_{it}|^2C_{NT}^{-2}$
$+O_P(b_{NT,1}^2+b_{NT,3}^2)\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2$.

$\max_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})e_{it}|\leq O_P(1+\max_{t\leq T}\|w_t\|+b_{NT,2})\max_{it}|e_{it}|C_{NT}^{-1}$
$+\max_{t\leq T}\|\frac{1}{N}\sum_i c_ie_{it}\|_FO_P(b_{NT,1}+b_{NT,3})$.

$\max_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})^2|\leq O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)C_{NT}^{-2}+O_P(b_{NT,1}^2+b_{NT,3}^2)$.

$\max_{t\leq T}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2\leq O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)C_{NT}^{-2}+O_P(b_{NT,1}^2+b_{NT,3}^2)$
$\max_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})|\leq O_P(1+\max_{t\leq T}\|w_t\|)C_{NT}^{-2}+O_P(b_{NT,1}+b_{NT,3}+C_{NT}^{-1}b_{NT,2})$.
$\max_i\frac{1}{T}\sum_t(\widehat{e}_{it}-e_{it})^2\leq O_P(b_{NT,4}^2+b_{NT,5}^2+C_{NT}^{-2})$.
(ii) All terms below are $O_P(C_{NT}^{-2})$: $\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^4$
$\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^2$, $\frac{1}{NT}\sum_{it}e_{it}^2(\widehat{e}_{it}-e_{it})^2$, $\frac{1}{T}\sum_t|\frac{1}{N}\sum_i c_i(e_{it}-\widehat{e}_{it})|^2m_t^2$, *where* $m_t^2=1+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|e_{jt}f_t\|^2(\|f_t\|+\|g_t\|)^2$.
(iii) All terms below are $O_P(C_{NT}^{-4})$: $\frac{1}{T}\sum_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})e_{it}|^2$,
$\frac{1}{T}\sum_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})^2|^2$, $\frac{1}{T}\sum_t|\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})u_{it}|^2$,
$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}f_s(\widehat{e}_{is}-e_{is})u_{is}\|^2$, $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}f_se_{is}(\widehat{e}_{is}-e_{is})\lambda_i'f_s\|^2$,
$\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\frac{1}{N}\sum_i c_if_t^2e_{jt}(e_{it}-\widehat{e}_{it})\|^2$

Proof. (i) First, in the proof of Lemma G.1(i), we showed $\max_{t\leq T}\|\widehat{w}_t\|\leq O_P(1+\max_{t\leq T}\|w_t\|)$. By the proof of Lemma G.1(ii),

$$
\begin{aligned}
&\max_{t\leq T}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4\leq\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}w_s)^4\max_{t\leq T}\widehat{w}_t^4+\frac{1}{N}\sum_i(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^4\max_{t\leq T}\widehat{w}_t^4\\
&+\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}\eta_s')^4\max_{t\leq T}\widehat{w}_t^4+\mathcal{Q}_3^4\max_{t\leq T}\widehat{w}_t^4+\max_{t\leq T}(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^4+\max_{t\leq T}\eta_t^4\\
\leq\ & O_P(1+\max_{t\leq T}\|w_t\|^4+b_{NT,2}^4)(C_{NT}^{-4})+O_P(b_{NT,1}^4+b_{NT,3}^4).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2(\widehat{e}_{it}-e_{it})^2 &= \max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2(\frac{1}{T}\sum_s e_{is}w_s)^2\max_{t\leq T}\widehat{w}_t^2+\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2l_i^4\mathcal{Q}_3^2\max_{t\leq T}\widehat{w}_t^2\\
&\quad+\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^2\max_{t\leq T}\widehat{w}_t^2\\
&\quad+\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2(\frac{1}{T}\sum_s e_{is}\eta_s')^2\max_{t\leq T}\widehat{w}_t^2\\
&\quad+\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2l_i^2(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2+\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2l_i^2\eta_t^2\\
&\leq O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)\max_{it}|e_{it}|^2C_{NT}^{-2}\\
&\quad+O_P(b_{NT,1}^2+b_{NT,3}^2)\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2.\\
\frac{1}{N}\sum_i c_ie_{it}(\widehat{e}_{it}-e_{it}) &\leq \max_{t\leq T}\frac{1}{N}\sum_i c_ie_{it}(\frac{1}{T}\sum_s e_{is}w_s)\widehat{w}_t\\
&\quad+\max_{t\leq T}\frac{1}{N}\sum_i c_ie_{it}(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)\widehat{w}_t
\end{aligned}
$$

$$
\begin{aligned}
& +\max_{t\leq T}\frac{1}{N}\sum_i c_i e_{it}(\frac{1}{T}\sum_s e_{is}\eta_s')\widehat{w}_t+\max_{t\leq T}\frac{1}{N}\sum_i c_i e_{it} l_i \mathcal{Q}_3 \max_{t\leq T}\widehat{w}_t \\
& +\max_{t\leq T}\frac{1}{N}\sum_i c_i e_{it} l_i(\frac{1}{T}\sum_s \widehat{w}_s\zeta_{st})+\max_{t\leq T}\frac{1}{N}\sum_i c_i e_{it} l_i \max_{t\leq T}\eta_t \\
\leq\ & O_P(1+\max_{t\leq T}\|w_t\|+b_{NT,2})\max_{it}|e_{it}|C_{NT}^{-1} \\
& +\max_{t\leq T}\|\frac{1}{N}\sum_i c_i e_{it} l_i\|_F O_P(b_{NT,1}+b_{NT,3}).
\end{aligned}
$$

$$
\begin{aligned}
\max_{t\leq T}\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})^2 \leq\ & \frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}w_s)^2\max_{t\leq T}\widehat{w}_t^2+\max_{t\leq T}(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2+\max_{t\leq T}\eta_t^2 \\
& +\frac{1}{N}\sum_i c_i(\frac{1}{T^2}\sum_{s,m\leq T} e_{is}\zeta_{ms}\widehat{w}_m)^2\max_{t\leq T}\widehat{w}_t^2 \\
& +\frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}\eta_s')^2\max_{t\leq T}\widehat{w}_t^2+\max_{t\leq T}\mathcal{Q}_3^2\widehat{w}_t^2 \\
\leq\ & O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)C_{NT}^{-2}+O_P(b_{NT,1}^2+b_{NT,3}^2),
\end{aligned}
$$

$$
\begin{aligned}
\max_{t\leq T}\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it}) \leq\ & \frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}w_s)\max_{t\leq T}\widehat{w}_t \\
& +\frac{1}{N}\sum_i c_i(\frac{1}{T^2}\sum_{s,m\leq T} e_{is}\zeta_{ms}\widehat{w}_m)\max_{t\leq T}\widehat{w}_t \\
& +\frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}\eta_s')\max_{t\leq T}\widehat{w}_t+\mathcal{Q}_3\max_{t\leq T}\widehat{w}_t \\
& +\max_{t\leq T}(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})+\max_{t\leq T}\eta_t \\
\leq\ & O_P(1+\max_{t\leq T}\|w_t\|)C_{NT}^{-2}+O_P(b_{NT,1}+b_{NT,3}+C_{NT}^{-1}b_{NT,2}).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\max_i\frac{1}{T}\sum_t(\widehat{e}_{it}-e_{it})^2 \leq\ & \max_i\frac{1}{T}\sum_t\widehat{w}_t^2(\frac{1}{T}\sum_s e_{is}w_s'H_x^{-1'})^2+\max_i\frac{1}{T}\sum_t\widehat{w}_t^2(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m'R_1')^2 \\
& +\max_i\frac{1}{T}\sum_t\widehat{w}_t^2(\frac{1}{T}\sum_s e_{is}\eta_s'R_2')^2+\max_i\frac{1}{T}\sum_t\widehat{w}_t^2(l_i'\mathcal{Q}_3)^2 \\
& +\max_i\frac{1}{T}\sum_t(l_i'H_xR_1\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2+\max_i\frac{1}{T}\sum_t\eta_t^2(l_i'H_xR_2)^2 \\
\leq\ & O_P(b_{NT,4}^2+b_{NT,5}^2+C_{NT}^{-2}).
\end{aligned}
$$

(ii) Note that $\max_{t\leq T}\|\widehat{w}_t\|^2 = O_P(1)+O_P(1)\max_{t\leq T}\|w_t\|^2 \leq O_P(1)+o_P(C_{NT})$, where the last inequality follows from the assumption that $\max_{t\leq T}\|w_t\|^2 = o_P(C_{NT})$.

$$
\begin{aligned}
\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^4 &\leq \frac{1}{T}\sum_t\|\widehat{w}_t\|^2\max_{t\leq T}\|\widehat{w}_t\|^2[\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}w_s)^4 + O_P(1)(\frac{1}{T^2}\sum_{s,m\leq T}\zeta_{ms}^2)^2] \\
&\quad +\frac{1}{T}\sum_t\|\widehat{w}_t\|^2\max_{t\leq T}\|\widehat{w}_t\|^2[\frac{1}{N}\sum_i(\frac{1}{T}\sum_s e_{is}\eta_s)^4 + \mathcal{Q}_3^4] \\
&\quad +\frac{1}{T}\sum_t(\frac{1}{T}\sum_s\zeta_{st}^2)^2 O_P(C_{NT}^{-4}) + \frac{1}{T}\sum_t\eta_t^4 + \frac{1}{T}\sum_t(\frac{1}{T}\sum_s w_s\zeta_{st})^4 \\
&\leq (O_P(1)+o_P(C_{NT}))C_{NT}^{-4} + O_P(C_{NT}^{-2}) \\
&\quad +O_P(1)\frac{1}{T}\sum_t\frac{1}{T^4}\sum_{s,k,l,m\leq T}\mathsf{E}w_l w_m w_k w_s \mathsf{E}(\zeta_{st}\zeta_{kt}\zeta_{lt}\zeta_{mt}|W) \\
&= O_P(C_{NT}^{-2}).
\end{aligned}
$$

Similarly, $\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^2 = O_P(C_{NT}^{-2})$.

$$
\begin{aligned}
\frac{1}{NT}\sum_{it}e_{it}^2(\widehat{e}_{it}-e_{it})^2 &\leq \frac{1}{NT}\sum_{it}e_{it}^2\widehat{w}_t^2(\frac{1}{T}\sum_s e_{is}\eta_s)^2 + \frac{1}{NT}\sum_{it}e_{it}^2\widehat{w}_t^2(\frac{1}{T}\sum_s e_{is}w_s)^2 \\
&\quad +\frac{1}{NT}\sum_{it}e_{it}^2\widehat{w}_t^2(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^2 + \frac{1}{NT}\sum_{it}e_{it}^2\widehat{w}_t^2(l_i'\mathcal{Q}_3)^2 \\
&\quad +\frac{1}{NT}\sum_{it}e_{it}^2(l_i'\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2 + \frac{1}{NT}\sum_{it}e_{it}^2(l_i'\eta_t)^2 = O_P(C_{NT}^{-2}).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\frac{1}{T}\sum_t(\frac{1}{N}\sum_i c_i(e_{it}-\widehat{e}_{it}))^2 m_t^2 &\leq \frac{1}{T}\sum_t m_t^2\widehat{w}_t^2(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_s e_{is}w_s)^2 \\
&\quad +\frac{1}{T}\sum_t m_t^2\widehat{w}_t^2(\frac{1}{N}\sum_i c_i\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m)^2 \\
&\quad +\frac{1}{T}\sum_t m_t^2\widehat{w}_t^2(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_s e_{is}\eta_s)^2 \\
&\quad +\frac{1}{T}\sum_t m_t^2\widehat{w}_t^2\mathcal{Q}_3^2 + \frac{1}{T}\sum_t m_t^2(\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st})^2 + \frac{1}{T}\sum_t m_t^2\eta_t^2 \\
&= O_P(C_{NT}^{-2}).
\end{aligned}
$$

(iii)

$$
\frac{1}{T}\sum_t(\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})e_{it})^2
$$

$$
\begin{aligned}
\leq\ & \max_{t\leq T}\|\widehat{w}_t-H_xw_t\|^2\frac{1}{T}\sum_t[(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_s e_{is}w_se_{it})^2+\frac{1}{T}\sum_m(\frac{1}{N}\sum_i c_ie_{it}\frac{1}{T}\sum_{s\leq T}e_{is}\zeta_{ms})^2] \\
&+\frac{1}{T}\sum_t w_t^2[(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_s e_{is}w_se_{it})^2+\frac{1}{T}\sum_m(\frac{1}{N}\sum_i c_ie_{it}\frac{1}{T}\sum_{s\leq T}e_{is}\zeta_{ms})^2] \\
&+\frac{1}{T}\sum_t\widehat{w}_t^2(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_s e_{is}\eta_se_{it})^2+\frac{1}{T}\sum_t\widehat{w}_t^2(\frac{1}{N}\sum_i c_il_i'e_{it})^2\mathcal{Q}_3^2 \\
&+\frac{1}{T}\sum_t(\frac{1}{N}\sum_i c_il_ie_{it})^2\frac{1}{T}\sum_s\zeta_{st}^2+\frac{1}{T}\sum_t\eta_t^2(\frac{1}{N}\sum_i c_il_ie_{it})^2=O_P(C_{NT}^{-4}).
\end{aligned}
$$

Next, to bound $\frac{1}{T}\sum_t(\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})^2)^2$, we first bound $\frac{1}{T}\sum_t\widehat{w}_t^4$. By (G.1),

$$
\begin{aligned}
\frac{1}{T}\sum_t\widehat{w}_t^4 &\leq O_P(1)+\frac{1}{T}\sum_t\|\widehat{w}_t-H_x^{-1}w_t\|^4 \\
&\leq O_P(1)+\frac{1}{T}\sum_t\|\eta_t\|^4+\frac{1}{T}\sum_t\|\frac{1}{T}\sum_s w_s\zeta_{st}\|^4+\frac{1}{T}\sum_t(\frac{1}{T}\sum_s\zeta_{st}^2)^2O_P(C_{NT}^{-4}) \\
&= O_P(1).
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}(\frac{1}{N}\sum_i c_i(\widehat{e}_{it}-e_{it})^2)^2 \\
\leq\ & \frac{1}{T}\sum_{t\notin I}\widehat{w}_t^4(\frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}w_s)^2)^2+\frac{1}{T}\sum_{t\notin I}\widehat{w}_t^4(\frac{1}{N}\sum_i c_i(\frac{1}{T}\sum_s e_{is}\eta_s)^2)^2 \\
&+\frac{1}{T}\sum_{t\notin I}\widehat{w}_t^4\mathcal{Q}_3^4+\frac{1}{T}\sum_{t\notin I}\eta_t^4+\frac{1}{T}\sum_{t\notin I}(\frac{1}{T}\sum_s w_s\zeta_{st})^4+\frac{1}{T}\sum_{t\notin I}(\frac{1}{T}\sum_s\zeta_{st}^2)^2O_P(C_{NT}^{-4}) \\
&+\frac{1}{T}\sum_{t\notin I}\widehat{w}_t^4(\frac{1}{N}\sum_i c_i\frac{1}{T}\sum_{m\leq T}(\frac{1}{T}\sum_{s\leq T}e_{is}\zeta_{ms})^2)^2=O_P(C_{NT}^{-4}).
\end{aligned}
$$

And up to an $O_P(1)$ as a product term,

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_t(\widehat{e}_{it}-e_{it})u_{it}\|^2 \\
\leq\ & \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_tu_{it}\widehat{w}_t\|^2(\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m\|^2+\|\frac{1}{T}\sum_s e_{is}w_s\|^2+\|\frac{1}{T}\sum_s e_{is}\eta_s\|^2+\|\mathcal{Q}_3\|^2) \\
&+\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_tu_{it}l_i\eta_t\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_tu_{it}\frac{1}{T}\sum_s\widehat{w}_s\zeta_{st}\|^2.
\end{aligned}
$$

It is easy to see all terms except for the last one is $O_P(C_{NT}^{-4})$. To see the last one, note that is is bounded by ,

$$
\begin{aligned}
&O_P(1)\frac{1}{T}\sum_s(\widehat{w}_s-H_xw_s)^2\frac{1}{T}\sum_s\frac{1}{T}\sum_{t\notin I}\zeta_{st}^2+O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_tu_{it}\frac{1}{T}\sum_s w_s\zeta_{st}\|^2\\
\leq\ & O_P(C_{NT}^{-4})+O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}f_tu_{jt}\frac{1}{T}\sum_s w_s\frac{1}{N}\sum_i(e_{is}e_{it}-\mathsf{E}e_{is}e_{it})\|^2\\
=\ & O_P(C_{NT}^{-4})
\end{aligned}
$$

since e_{it} is conditionally serially independent given (U,W,F) and u_{it} is conditionally serially independent given (E,W,F).

The conclusion that $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}f_se_{is}(\widehat{e}_{is}-e_{is})\lambda_i'f_s\|^2=O_P(C_{NT}^{-4})$ follows similarly, due to $\max_j\sum_{i\leq N}|(\mathsf{E}e_{it}e_{jt}|F,W)|<C$.

Next, for $a:=\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{N}\sum_i\|\frac{1}{T_0}\sum_{t\in I^c}c_if_t^2e_{jt}\widehat{w}_t\|^2=O_P(C_{NT}^{-2})$,

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\frac{1}{N}\sum_i c_if_t^2e_{jt}(e_{it}-\widehat{e}_{it})\|^2\\
\leq\ & a\frac{1}{N}\sum_i[\|\frac{1}{T}\sum_s e_{is}w_s\|^2+\|\frac{1}{T^2}\sum_{s,m\leq T}e_{is}\zeta_{ms}\widehat{w}_m'\|^2+\|\frac{1}{T}\sum_s e_{is}\eta_s\|^2+\|\mathcal{Q}_3\|^2]\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}f_t^2e_{jt}\frac{1}{T}\sum_s\widehat{w}_s\frac{1}{N}\sum_i(e_{is}e_{it}-\mathsf{E}e_{is}e_{it})\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i l_i\frac{1}{T_0}\sum_{t\in I^c}f_t^2(e_{jt}e_{it}-\mathsf{E}e_{jt}e_{it})\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i l_i\frac{1}{T_0}\sum_{t\in I^c}f_t^2\mathsf{E}e_{jt}e_{it}\|^2\\
=\ & O_P(C_{NT}^{-4}).
\end{aligned}
$$

□

G.2. **Behavior of the preliminary.** Recall that

$$(\widetilde{f}_s,\widetilde{g}_s):=\arg\min_{f_s,g_s}\sum_{i=1}^N(y_{is}-\widetilde{\alpha}_i'g_s-x_{is}\widetilde{\lambda}_i'f_s)^2,\quad s\notin I.$$

and

$$(\dot{\lambda}_i,\dot{\alpha}_i)=\arg\min_{\lambda_i,\alpha_i}\sum_{s\notin I}(y_{is}-\alpha_i'\widetilde{g}_s-x_{is}\lambda_i'\widetilde{f}_s)^2,\quad i=1,...,N.$$

The goal of this section is to show that the effect of the preliminary estimation is negligible. Specifically, we aim to show, for each fixed $t\notin I$,

$$\|\frac{1}{N}\sum_j(\dot{\lambda}_j-H_1'\lambda_j)e_{jt}\|^2\ =\ O_P(C_{NT}^{-4}),$$

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I} f_s(\widetilde{f}_s - H_1^{-1}f_s)'\mu_{is}e_{is}\|^2 \quad = \quad O_P(C_{NT}^{-4}).$$

Throughout the proof below, we treat $|I^c| = T$ instead of $T/2$ to avoid keeping the constant "2". In addition, for notational simplicity, we write $\widetilde{\Lambda} = \widetilde{\Lambda}_I$ and $\widetilde{A} = \widetilde{A}_I$ by suppressing the subscripts, but we should keep in mind that $\widetilde{\Lambda}$ and $\widetilde{A}$ are estimated on data D_I as defined in step 2. In addition, let E_I and Var_I be the conditional expectation and variance, given D_I. Recall that X_s be the vector of x_{is} fixing $s \leq T$, and $M_{\widetilde{\alpha}} = I_N - \widetilde{A}(\widetilde{A}'\widetilde{A})^{-1}\widetilde{A}'$; X_i be the vector of x_{is} fixing $i \leq N$, and $M_{\widetilde{g}} = I - \widetilde{G}(\widetilde{G}'\widetilde{G})^{-1}\widetilde{G}'$, for $\widetilde{G}$ as the $|I^c|_0 \times K_1$ matrix of $\widetilde{g}_s$. Define $\widetilde{F}$ similarly. Let L denote $N \times K_3$ matrix of l_i, so $X_s = Lw_s + e_s$. Also let W be $T \times K_3$ matrix of w_t.

Define

$$\begin{aligned}
\widetilde{D}_{fs} &= \frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)\widetilde{\Lambda} \\
D_{fs} &= \frac{1}{N}\Lambda'(\mathsf{diag}(X_s)M_{\alpha}\mathsf{diag}(X_s)\Lambda \\
\bar{D}_{fs} &= \frac{1}{N}\Lambda'\mathsf{E}((\mathsf{diag}(e_s)M_{\alpha}\mathsf{diag}(e_s))\Lambda + \frac{1}{N}\Lambda'(\mathsf{diag}(Lw_s)M_{\alpha}\mathsf{diag}(Lw_s)\Lambda \\
\widetilde{D}_{\lambda i} &= \frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)\widetilde{F} \\
D_{\lambda i} &= \frac{1}{T}F'(\mathsf{diag}(X_i)M_g\mathsf{diag}(X_i))F \\
\bar{D}_{\lambda i} &= \frac{1}{T}F'\mathsf{E}(\mathsf{diag}(E_i)M_g\mathsf{diag}(E_i))F + \frac{1}{T}F'(\mathsf{diag}(Wl_i)M_g\mathsf{diag}(Wl_i))F
\end{aligned}$$

By the stationarity, D_f does not depend on s.

Lemma G.3. *Suppose* $\max_{it} e_{it}^2 + \max_{t\leq T}\|w_t\|^2 = o_P(C_{NT})$, $\max_{t\leq T}\mathsf{E}\|w_t\|^4 = O(1)$. $\frac{1}{N}\sum_{ij}|\mathsf{Cov}(e_{is}, e_{js}|w_s)| < \infty$ *and* $\|\mathsf{E}e_se_s'\| < \infty$. $\frac{1}{N^3}\sum_{ijkl}\mathsf{Cov}(e_{is}e_{js}, e_{ks}e_{ls}) < C$. *Also, there is* $c > 0$, *so that* $\min_s \min_j \psi_j(D_{fs}) > c$. *Then*

(i) $\max_s \|\widetilde{D}_{fs}^{-1}\| = O_P(1)$.

(ii) $\frac{1}{T}\sum_{s\notin I}\|\widetilde{D}_{fs}^{-1} - (H_1'\bar{D}_{fs}H_1)^{-1}\|^2 = O_P(C_{NT}^{-2})$.

Proof. (i) The eigenvalues of D_{fs} are bounded from zero uniformly in $s \leq T$. Also,

$$\begin{aligned}
\widetilde{D}_{fs} - H_1'D_{fs}H_1 &= \sum_l \delta_l, \qquad \text{where} \\
\delta_1 &= \frac{1}{N}(\widetilde{\Lambda} - \Lambda H_1)'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)\widetilde{\Lambda},
\end{aligned}$$

$$\begin{aligned} \delta_2 &= \frac{1}{N} H_1' \Lambda' \mathsf{diag}(X_s) M_{\widetilde{\alpha}} \mathsf{diag}(X_s)(\widetilde{\Lambda} - \Lambda H_1) \\ \delta_3 &= \frac{1}{N} H_1' \Lambda' \mathsf{diag}(X_s)(M_{\widetilde{\alpha}} - M_\alpha) \mathsf{diag}(X_s) \Lambda H_1. \end{aligned} \tag{G.4}$$

We now bound each term uniformly in $s \leq T$. The first term is

$$\frac{1}{N}(\widetilde{\Lambda} - \Lambda H_1)' \mathsf{diag}(X_s) M_{\widetilde{\alpha}} \mathsf{diag}(X_s) \widetilde{\Lambda} \leq O_P(1) \frac{1}{\sqrt{N}} \|\widetilde{\Lambda} - \Lambda H_1\|_F \max_{is} x_{is}^2 = o_P(1)$$

provided that $\max_{is} x_{is}^2 = o_P(C_{NT})$. The second term is bounded similarly. The third term is bounded by

$$O_P(1) \max_{is} x_{is}^2 \|M_{\widetilde{\alpha}} - M_\alpha\| = O_P(1) \frac{1}{\sqrt{N}} \|\widetilde{A} - AH_2\|_F \max_{is} x_{is}^2 = o_P(1).$$

This implies $\max_s \|\widetilde{D}_{fs} - H_1' D_{fs} H_1\| = o_P(1)$. In addition, because of the convergence of $\|\frac{1}{\sqrt{N}}(\widetilde{\Lambda} - \Lambda H_1)\|_F$, we have $\min_j \psi_j(H_1' H_1) \geq C \min_j \psi_j(\frac{1}{N} H_1' \Lambda' \Lambda H_1)$, bounded away from zero. Thus $\min_s \min_j \psi_j(H_1' D_{fs} H_1) \geq \min_s \min_j \psi_j(D_{fs}) C$, bounded away from zero. This together with $\max_s \|\widetilde{D}_{fs} - H_1' D_{fs} H_1\| = o_P(1)$ imply $\min_s \|\widetilde{D}_{fs}^{-1}\| = O_P(1)$.

(ii) By (G.4),

$$\begin{aligned} & \frac{1}{T} \sum_{s \notin I} \|\widetilde{D}_{fs}^{-1} - (H_1' D_{fs} H_1)^{-1}\|^2 \\ \leq & \frac{1}{T} \sum_s \|\widetilde{D}_{fs} - (H_1' D_{fs} H_1)\|^2 \|(H_1' D_{fs} H_1)^{-2}\| \max_s \|\widetilde{D}_{fs}^{-2}\| \\ \leq & O_P(1) \frac{1}{T} \sum_s \|\widetilde{D}_{fs} - (H_1' D_{fs} H_1)\|^2 \\ = & O_P(1) \sum_{l=1}^{3} \frac{1}{T} \sum_s \|\delta_l\|^2. \end{aligned}$$

We bound each term below. Up to a $O_P(1)$ product,

$$\begin{aligned} \frac{1}{T} \sum_{s \notin I} \|\delta_1\|^2 &\leq \frac{1}{N^2} \sum_{ij} \|\widetilde{\lambda}_i - H_1' \lambda_i\|^2 \|\widetilde{\lambda}_j\|^2 \frac{1}{T} \sum_{s \notin I} x_{is}^2 x_{js}^2 \\ &\leq \frac{1}{N^2} \sum_{ij} \|\widetilde{\lambda}_i - H_1' \lambda_i\|^2 \|\widetilde{\lambda}_j\|^2 \frac{1}{T} \sum_{s \notin I} (1 + e_{is}^2 e_{js}^2) \\ &\quad + \frac{1}{N^2} \sum_{ij} \|\widetilde{\lambda}_i - H_1' \lambda_i\|^2 \|\widetilde{\lambda}_j\|^2 (\frac{1}{T} \sum_{s \notin I} e_{is}^4)^{1/2} \end{aligned}$$

$\{e_{is}\}$ is serially independent, so e_{is}, e_{js} are independent of $\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2\|\widetilde{\lambda}_j - H_1'\lambda_j\|^2$ for $s \notin I$. Take the conditional expectation E_I.

$$\begin{aligned} \frac{1}{T}\sum_{s\notin I}\|\delta_1\|^2 &\leq \frac{1}{N^2}\sum_{ij}\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2\|\widetilde{\lambda}_j\|^2\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I(1+e_{is}^2e_{js}^2) \\ &+\frac{1}{N^2}\sum_{ij}\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2\|\widetilde{\lambda}_j\|^2\mathsf{E}_I(\frac{1}{T}\sum_{s\notin I}e_{is}^4)^{1/2} = O_P(C_{NT}^{-2}). \end{aligned}$$

Term of $\frac{1}{T}\sum_{s\notin I}\|\delta_2\|^2$ is bounded similarly.

$$\frac{1}{T}\sum_{s\notin I}\|\delta_3\|^2 \leq O_P(1)\|M_{\widetilde{\alpha}} - M_\alpha\| = O_P(C_{NT}^{-2}).$$

Next,

$$\begin{aligned} \frac{1}{T}\sum_s\|D_{fs} - \bar{D}_{fs}\|^2 &\leq \frac{1}{T}\sum_s\|\frac{1}{N}\sum_{ij}\lambda_i\lambda_j'M_{\alpha,ij}(x_{is}x_{js} - \mathsf{E}e_{is}e_{js} - l_i'w_sl_j'w_s)\|_F^2 \\ \leq & \frac{1}{T}\sum_s\|\frac{1}{N}\sum_{ij}\lambda_i\lambda_j'M_{\alpha,ij}(e_{is}e_{js} - \mathsf{E}e_{is}e_{js})\|_F^2 \\ &+\frac{2}{T}\sum_s\|\frac{1}{N}\sum_{ij}\lambda_i\lambda_j'M_{\alpha,ij}l_i'w_se_{js}\|_F^2 \end{aligned}$$

We assume $\dim(\lambda_i) = \dim(p_i) = 1$. As for the first term, it is less than

$$\begin{aligned} &\frac{1}{T}\sum_s\mathsf{Var}(\frac{1}{N}\sum_{ij}\lambda_i\lambda_j'M_{\alpha,ij}e_{is}e_{js}) \\ \leq\ &\frac{1}{T}\sum_s\frac{1}{N^4}\sum_{ijkl}|\mathsf{Cov}(e_{is}e_{js}, e_{ks}e_{ls})| = O(N^{-1}) \end{aligned}$$

provided that $\frac{1}{N^3}\sum_{ijkl}\mathsf{Cov}(e_{is}e_{js}, e_{ks}e_{ls}) < C$. As for the second term, it is less than

$$O_P(1)\frac{1}{T}\sum_s\mathsf{E}w_s^2\frac{1}{N^2}\sum_{ij}|\mathsf{Cov}(e_{js}, e_{ls}|w_s)| = O(N^{-1})$$

provided that $\frac{1}{N}\sum_{ij}|\mathsf{Cov}(e_{is}, e_{js}|w_s)| < \infty$ and $\|\mathsf{E}e_se_s'\| < \infty$. So $\frac{1}{T}\sum_s\|D_{fs} - \bar{D}_{fs}\|^2 = O_P(N^{-1})$.

□

Lemma G.4. *Suppose* $\max_{it} e_{it}^4 = O_P(\min\{N,T\})$. *(i)* $\frac{1}{T}\|\widetilde{F} - FH_1^{-1'}\|_F^2 = O_P(C_{NT}^{-2}) = \frac{1}{T}\|\widetilde{G} - GH_2^{-1'}\|_F^2$, *and* $\frac{1}{T}\sum_{t\notin I}\|\widetilde{f}_t - H_1^{-1}f_t\|^2e_{it}^2u_{it}^2 = O_P(C_{NT}^{-2})$.

(ii) $\max_i \|\widetilde{D}_{\lambda i}^{-1}\| = O_P(1)$.
(iii) $\frac{1}{N}\sum_i \|\widetilde{D}_{\lambda i}^{-1} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1}\|^2 = O_P(C_{NT}^{-2})$.

Proof. (i) The proof is straightforward given the expansion of (G.7) and $\max_s \|\widetilde{D}_{fs}^{-1}\| = O_P(1)$. So we omit the details for brevity. (ii) Note that

$$
\begin{aligned}
\widetilde{D}_{\lambda i} - H_1^{-1}D_{\lambda i}H_1^{-1'} &= \sum_l \delta_l, \quad \text{where} \\
\delta_1 &= \frac{1}{T}(\widetilde{F} - FH_1^{-1'})'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)\widetilde{F}, \\
\delta_2 &= \frac{1}{T}H_1^{-1}F'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(\widetilde{F} - FH_1^{-1'}) \\
\delta_3 &= \frac{1}{T}H_1^{-1}F'\mathsf{diag}(X_i)(M_{\widetilde{g}} - M_g)\mathsf{diag}(X_i)FH_1^{-1'}. \quad \text{(G.5)}
\end{aligned}
$$

The proof is very similar to that of Lemma G.3.

(iii)

$$
\begin{aligned}
&\frac{1}{N}\sum_i \|\widetilde{D}_{\lambda i}^{-1} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1}\|^2 \\
\leq\ & \frac{1}{N}\sum_i \|\widetilde{D}_{\lambda i} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})\|^2 \max_i \|(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-2}\|\|\widetilde{D}_{\lambda i}^{-2}\| \\
\leq\ & O_P(1)\frac{1}{N}\sum_i \|\widetilde{D}_{\lambda i} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})\|^2 \\
=\ & O_P(1)\sum_{l=1}^{3}\frac{1}{N}\sum_i \|\delta_l\|^2 + O_P(1)\frac{1}{N}\sum_i \|D_{\lambda i} - \bar{D}_{\lambda i}\|^2.
\end{aligned}
$$

We now bound each term. With the assumption that $\max_{it} x_{it}^4 = O_P(\min\{N, T\})$,

$$
\begin{aligned}
\frac{1}{N}\sum_i \|\delta_1\|^2 &\leq \frac{1}{T}\|\widetilde{F} - FH_1^{-1'}\|_F^2 \|M_{\widetilde{g}} - M_g\|^2 \frac{1}{T}\|\widetilde{F}\|_F^2 \frac{1}{N}\sum_i \|\mathsf{diag}(X_i)\|^4 \\
&\quad + \frac{1}{T^2}\|\widetilde{F} - FH_1^{-1'}\|_F^4 \frac{1}{N}\sum_i \|\mathsf{diag}(X_i)\|^4 \\
&\quad + \frac{1}{N}\sum_i \|\frac{1}{T}(\widetilde{F} - FH_1^{-1'})'\mathsf{diag}(X_i)M_g\mathsf{diag}(X_i)FH_1^{-1'}\|^2 \\
&\leq O_P(C_{NT}^{-4})\max_{it} x_{it}^4 + \frac{1}{NT}\sum_i\sum_t x_{it}^2(\sum_s M_{g,ts}x_{is}f_s)^2 O_P(C_{NT}^{-2}) \\
&\leq O_P(C_{NT}^{-2}).
\end{aligned}
$$

$\frac{1}{N}\sum_i \|\delta_2\|^2$ is bounded similarly. $\frac{1}{N}\sum_i \|\delta_3\|^2 \leq O_P(1)\|M_{\widetilde{g}} - M_g\|^2 = O_P(C_{NT}^{-2})$. Finally,

$$\begin{aligned}\frac{1}{N}\sum_i \|D_{\lambda i} - \bar{D}_{\lambda i}\|^2 &\leq \frac{1}{N}\sum_i \|\frac{1}{T}\sum_{st} f_s f_t' M_{g,st}(x_{is}x_{it} - \mathsf{E}e_{is}e_{it} - l_i'w_s l_i'w_t)\|_F^2 \\ &\leq O_P(1)\frac{1}{N}\sum_i \|\frac{1}{T}\sum_{st} f_s f_t' M_{g,st}(e_{is}e_{it} - \mathsf{E}e_{is}e_{it})\|_F^2 \\ &\quad + O_P(1)\frac{1}{N}\sum_i \|\frac{1}{T}\sum_{st} f_s f_t' M_{g,st} w_s e_{it}\|_F^2 \\ &\leq O_P(1)\frac{1}{N}\sum_i \|\frac{1}{T}\sum_t f_t^2 w_t e_{it}\|_F^2 + O_P(1)\frac{1}{N}\sum_i \|\frac{1}{T}\sum_t f_t g_t e_{it}\|_F^2 \\ &= O_P(C_{NT}^{-2}).\end{aligned}$$

Lemma G.5. *Suppose* $\mathsf{Var}(u_s|e_t, e_s, w_s) < C$ *and* $C_{NT}^{-1}\max_{is}|x_{is}|^2 = O_P(1)$.

(i) For each fixed $t \notin I$, $\frac{1}{N}\sum_i (H_1'\lambda_i - \dot{\lambda}_i)e_{it} = O_P(C_{NT}^{-2})$.

(ii) $\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i c_i e_{it}(\dot{\lambda}_i - H_1'\lambda_i)\|^2 = O_P(C_{NT}^{-4})$ *for any deterministic and bounded sequence* c_i.

Proof. Given that the $|I^c| \times 1$ vector $y_i = G\alpha_i + \mathsf{diag}(X_i)F\lambda_i + u_i$ where u_i is a $|I^c| \times 1$ vector and G is an $|I^c| \times K_1$ matrix, we have

$$\begin{aligned}\dot{\lambda}_i &= \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}y_i \\ &= H_1'\lambda_i + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'} - \widetilde{G})H_2'\alpha_i \\ &\quad + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'} - \widetilde{F})H_1'\lambda_i + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i.\end{aligned}$$

(i)

Step 1: given $\mathsf{E}e_{it}^4 f_t^2 + e_{it}^2 f_t^2 w_t^2 < C$,

$$\begin{aligned}&\frac{1}{\sqrt{N}}\sum_i e_{it}(\widetilde{D}_{\lambda i}^{-1} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1})\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'} - \widetilde{G})H_2'\alpha_i \\ \leq\ & O_P(\sqrt{N}C_{NT}^{-2}) + O_P(\sqrt{N}C_{NT}^{-2})(\frac{1}{T}\sum_s(\widetilde{f}_s - H_1^{-1}f_s)^2)^{1/2}\max_{is}|x_{is}| \\ \leq\ & O_P(\sqrt{N}C_{NT}^{-2}) + O_P(\sqrt{N}C_{NT}^{-3})\max_{is}|x_{is}| = O_P(\sqrt{N}C_{NT}^{-2}).\end{aligned}$$

Step 2: $\bar{D}_{\lambda i}$ is nonrandom given W, G, F,

$$\begin{aligned}&\frac{1}{\sqrt{N}}\sum_i e_{it}(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'} - \widetilde{G})H_2'\alpha_i \\ \leq\ & O_P(\sqrt{N}C_{NT}^{-1})(a^{1/2} + b^{1/2}) \quad \text{where}\end{aligned}$$

$$
\begin{aligned}
a &= \frac{1}{T}\sum_s(\widetilde{f}_s - H_1^{-1}f_s)^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
b &= \frac{1}{T}\sum_s f_s^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2.
\end{aligned}
$$

We now bound each term. As for b, note that for each fixed t,

$$
\begin{aligned}
& \mathsf{E}((\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2|F,G,W,u_s) \\
\leq\ & \mathsf{E}(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}e_{is}|F,G,W,u_s)^2 + \mathsf{E}(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}l_i'w_s|F,G,W,u_s)^2 \\
\leq\ & \frac{C}{N} + \frac{C\|w_s\|^2}{N} + (\mathsf{E}\frac{1}{N}\sum_i \alpha_i e_{it}e_{is}\bar{D}_{\lambda i}^{-1}|F,G,W,u_s)^2 \qquad \text{(G.6)}
\end{aligned}
$$

with the assumption $\frac{1}{N}\sum_{ij}|\mathsf{Cov}(e_{jt}e_{js}, e_{it}e_{is}|F,G,W,u_s)| < C$. So

$$
\begin{aligned}
\mathsf{E}b &\leq \frac{1}{T}\sum_s \mathsf{E}f_s^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \leq O(C_{NT}^{-2}). \\
a &= \frac{1}{T}\sum_s f_s^2(\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1 - \widetilde{\Lambda}))^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\quad +\frac{1}{T}\sum_s(\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1 - \widetilde{\Lambda})H_1^{-1}f_s)^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\quad +\frac{1}{T}\sum_s(\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s)^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\leq O_P(C_{NT}^{-2})\max_{is} x_{is}^4\frac{1}{T}\sum_s \mathsf{E}f_s^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\quad +O_P(1)\frac{1}{T}\sum_s(\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s)^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\leq O_P(C_{NT}^{-2}) + O_P(1)\frac{1}{T}\sum_s \mathsf{E}(\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_{\alpha}u_s)^2(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \\
&\quad +O_P(C_{NT}^{-2})\max_{is} x_{is}^2\frac{1}{TN}\sum_s \mathsf{E}\|u_s\|^2\mathsf{E}((\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2|u_s) \\
&:= O_P(C_{NT}^{-2}) + O_P(1)(a.1 + a.2).
\end{aligned}
$$

We now respectively bound $a.1$ and $a.2$. As for $a.1$, note that $\mathsf{Var}(u_s|e_t,e_s) < C$ almost surely, thus

$$
\begin{aligned}
& \mathsf{E}(\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_\alpha u_s|e_t,e_s)^2 = \frac{1}{N}\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_\alpha\mathsf{Var}(u_s|e_t,e_s)M_\alpha\mathsf{diag}(X_s)\Lambda \\
\leq\ & C\frac{1}{N}\frac{1}{N}\Lambda'\mathsf{diag}(X_s)^2\Lambda.
\end{aligned}
$$

As for $a.2$, we use (G.6). Thus,

$$
\begin{aligned}
a.1 &\leq \frac{1}{T}\sum_s \mathsf{E}(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2\mathsf{E}(\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_\alpha u_s|e_t,e_s)^2 \\
&\leq \frac{C}{N}\frac{1}{T}\sum_s \mathsf{E}(\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2 \ \frac{1}{N}\Lambda'\mathsf{diag}(X_s)^2\Lambda = O(N^{-1}). \\
a.2 &\leq C_{NT}^{-2}\max_{is} x_{is}^2\frac{1}{TN}\sum_s \mathsf{E}\|u_s\|^2\mathsf{E}((\frac{1}{N}\sum_i \alpha_i e_{it}\bar{D}_{\lambda i}^{-1}x_{is})^2|u_s) \\
&\leq C_{NT}^{-2}\max_{is} x_{is}^2\frac{1}{TN}\sum_s \mathsf{E}\|u_s\|^2\frac{C}{N} \\
&\quad +C_{NT}^{-2}\max_{is} x_{is}^2\frac{1}{TN}\sum_s \mathsf{E}\|u_s\|^2(\mathsf{E}\frac{1}{N}\sum_i \alpha_i e_{it}e_{is}\bar{D}_{\lambda i}^{-1}|F,u_s)^2 \\
&\leq O(C_{NT}^{-2}).
\end{aligned}
$$

Put together, $a^{1/2}+b^{1/2} = O(C_{NT}^{-1})$. So the first term in the expansion of $\frac{1}{\sqrt{N}}\sum_i(H_1'\lambda_i-\dot{\lambda}_i)e_{it}$ is

$$
\frac{1}{\sqrt{N}}\sum_i e_{it}\widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'}-\widetilde{G})H_2'\alpha_i = O_P(\sqrt{N}C_{NT}^{-2}).
$$

Step 3:

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}(\widetilde{D}_{\lambda i}^{-1}-(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1})\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'}-\widetilde{F})H_1'\lambda_i \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2})(\frac{1}{NT}\sum_i e_{it}^2\|F'\mathsf{diag}(X_i)M_g\mathsf{diag}(X_i)\|^2)^{1/2} \\
&+O_P(\sqrt{N}C_{NT}^{-2})(\frac{1}{N}\sum_i e_{it}^2\|\mathsf{diag}(X_i)\|^4)^{1/2}[\|M_g-M_{\widetilde{g}}\|\frac{1}{\sqrt{T}}\|\widetilde{F}\|+\frac{1}{\sqrt{T}}\|\widetilde{F}-FH_1^{-1}\|] \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2})(\frac{1}{NT}\sum_i\sum_s e_{it}^2 f_s^2 x_{is}^4)^{1/2} \\
&+O_P(\sqrt{N}C_{NT}^{-2})(\frac{1}{NT}\sum_i\sum_s x_{is}^2 e_{it}^2 g_s^2(\frac{1}{T}\sum_k f_k g_k x_{ik})^2)^{1/2} \\
&+O_P(\sqrt{N}C_{NT}^{-3})\max_{it} x_{it}^2 = O_P(\sqrt{N}C_{NT}^{-2}).
\end{aligned}
$$

Step 4: note $\frac{1}{T}\sum_s(\frac{1}{N}\sum_i \lambda_i e_{it}\bar{D}_{\lambda i}^{-1}e_{is}^2)^2 = O_P(C_{NT}^{-2})$,

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1}\frac{1}{T}(\widetilde{F}-FH_1^{-1'})'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'}-\widetilde{F})H_1'\lambda_i \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-3})\max_{it} x_{it}^2
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{N}}\sum_i e_{it}(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1\prime})^{-1}\frac{1}{T}(\widetilde{F}-FH_1^{-1\prime})'\mathsf{diag}(X_i)M_g\mathsf{diag}(X_i)(FH_1^{-1\prime}-\widetilde{F})H_1'\lambda_i \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2})+\sqrt{N}\frac{1}{T}\sum_s(\widetilde{f}_s-H_1^{-1}f_s)^2\frac{1}{N}\sum_i\lambda_ie_{it}\bar{D}_{\lambda i}^{-1}x_{is}^2 \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2})+\max_{is}x_{is}^2O_P(\sqrt{N}C_{NT}^{-2})\frac{1}{NT}\sum_s g_s^2w_s^2\sum_i\lambda_ie_{it}\bar{D}_{\lambda i}^{-1}l_i^2 \\
&+\max_{is}x_{is}^2O_P(\sqrt{N}C_{NT}^{-2})\frac{1}{T}\sum_s w_s^2\frac{1}{N}\sum_j\lambda_j^2x_{js}^2f_s^2\frac{1}{N}\sum_i\lambda_ie_{it}\bar{D}_{\lambda i}^{-1}l_i^2 \\
&+\sqrt{N}\frac{1}{T}\sum_s w_s^2(\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_\alpha u_s)^2\frac{1}{N}\sum_i\lambda_ie_{it}\bar{D}_{\lambda i}^{-1}l_i^2 \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2}).
\end{aligned}
$$

Step 5: First we bound $\frac{1}{NT}\sum_i\sum_s(\widetilde{f}_s-H_1^{-1}f_s)^2e_{is}^2$. We have

$$
\begin{aligned}
\widetilde{f}_s &= \widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}y_s \\
&= H_1^{-1}f_s+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s \\
&\quad+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s \\
&\quad+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s.
\end{aligned}
$$

So

$$
\begin{aligned}
&\frac{1}{NT}\sum_i\sum_s(\widetilde{f}_s-H_1^{-1}f_s)^2e_{is}^2 \\
=\ & \frac{1}{NT}\sum_i\sum_{s\notin I}e_{is}^2\frac{1}{N}\|\Lambda'\mathsf{diag}(X_s)\|^2[g_s^2+\frac{1}{N}\|u_s\|^2]O_P(C_{NT}^{-2}) \\
&+O_P(1)\frac{1}{N}\sum_j(H_1\lambda_j-\widetilde{\lambda}_j)^2\max_j\frac{1}{NT}\sum_i\sum_{s\notin I}\mathsf{E}_Ie_{is}^2f_s^2\frac{1}{N}\|\Lambda'\mathsf{diag}(X_s)\|^2x_{js}^2 \\
&+\frac{1}{NT}\sum_i\sum_s e_{is}^2(\frac{1}{N}\Lambda'\mathsf{diag}(X_s)M_\alpha u_s)^2 \\
&+\frac{1}{NT}\sum_i\sum_s e_{is}^2\frac{1}{N}\|\mathsf{diag}(X_s)M_\alpha u_s\|^2O_P(C_{NT}^{-2}) \\
=\ & O_P(C_{NT}^{-2}).
\end{aligned}
$$

Therefore

$$
\frac{1}{\sqrt{N}}\sum_i e_{it}(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1\prime})^{-1}\frac{1}{T}F'\mathsf{diag}(X_i)(M_{\widetilde{g}}-M_g)\mathsf{diag}(X_i)(FH_1^{-1\prime}-\widetilde{F})H_1'\lambda_i
$$

$$
\begin{aligned}
&\leq O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{N}\sum_i \lambda_i^2 e_{it}^2 \frac{1}{T}\sum_k f_k^2 x_{ik}^2)^{1/2}(\frac{1}{NT}\sum_i\sum_s(\widetilde{f}_s - H_1^{-1}f_s)^2 e_{is}^2)^{1/2}\\
&\leq O_P(\sqrt{N}C_{NT}^{-2}),
\end{aligned}
$$

where the last equality is due to $\frac{1}{NT}\sum_i\sum_{s\notin I}(\widetilde{f}_s - H_1^{-1}f_s)^2 e_{is}^2 = O_P(C_{NT}^{-2})$.

Step 6:

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}(H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1}\frac{1}{T}F'\mathsf{diag}(X_i)M_g\mathsf{diag}(X_i)(FH_1^{-1'} - \widetilde{F})H_1'\lambda_i\\
\leq\ & O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{T}\sum_s f_s^2(\frac{1}{N}\sum_i \lambda_i x_{is}^2 e_{it}\bar{D}_{\lambda i}^{-1})^2)^{1/2}\\
&+O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{T}\sum_s g_s^2(\frac{1}{N}\sum_i x_{is}\lambda_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}\sum_k f_k x_{ik} g_k)^2)^{1/2}\\
=\ & O_P(\sqrt{N}C_{NT}^{-1})(a^{1/2}+b^{1/2}).
\end{aligned}
$$

We aim to show $a = O_P(C_{NT}^{-2}) = b$.

$$
\begin{aligned}
\mathsf{E}a \ :=\ & \frac{1}{T}\sum_s \mathsf{E}f_s^2(\frac{1}{N}\sum_i \lambda_i x_{is}^2 e_{it}\bar{D}_{\lambda i}^{-1})^2\\
\leq\ & \frac{1}{T}\sum_s \frac{1}{N^2}\sum_{ij}\mathsf{E}f_s^2\lambda_i\bar{D}_{\lambda i}^{-1}\lambda_j\bar{D}_{\lambda j}^{-1}\mathsf{E}(e_{is}^2 e_{js}^2|F)\mathsf{Cov}(e_{it},e_{jt}|F,G,W)\\
&+\frac{1}{T}\sum_s \mathsf{E}f_s^2\frac{1}{N^2}\sum_{ij}\lambda_i l_i^2 w_s^2\bar{D}_{\lambda i}^{-1}\lambda_j\mu_j^2\bar{D}_{\lambda j}^{-1}\mathsf{Cov}(e_{it},e_{jt}|F,G,W)\\
&+\frac{2}{T}\sum_s \mathsf{E}f_s^2\frac{1}{N^2}\sum_{ij}\lambda_i l_i w_s\bar{D}_{\lambda i}^{-1}\bar{D}_{\lambda j}^{-1}\lambda_i l_j w_s\mathsf{E}(e_{js}e_{is}|F)\mathsf{Cov}(e_{jt},e_{it}|F,G,W)\\
=\ & O_P(N^{-1}).\\
\mathsf{E}b \ :=\ & \frac{1}{T}\sum_s \mathsf{E}g_s^2(\frac{1}{N}\sum_i x_{is}\lambda_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}\sum_k f_k x_{ik}g_k)^2\\
=\ & O_P(C_{NT}^{-2}).
\end{aligned}
$$

Therefore the second term is

$$
\frac{1}{\sqrt{N}}\sum_i e_{it}\widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'} - \widetilde{F})H_1'\lambda_i = O_P(\sqrt{N}C_{NT}^{-2}).
$$

Step 7: $\frac{1}{\sqrt{N}}\sum_i e_{it}\widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i$.

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}(\widetilde{D}_{\lambda i}^{-1} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1})\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i\\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2}) + O_P(1)\frac{1}{\sqrt{N}}\sum_i e_{it}(\widetilde{D}_{\lambda i}^{-1} - (H_1^{-1}\bar{D}_{\lambda i}H_1^{-1'})^{-1})\frac{1}{T}F'\mathsf{diag}(X_i)M_g u_i
\end{aligned}
$$

$$
\begin{aligned}
&\leq O_P(\sqrt{N}C_{NT}^{-2}) + O_P(\sqrt{N}C_{NT}^{-1})\frac{1}{N}\sum_i |e_{it}||\frac{1}{T}\sum_s f_s x_{is}u_{is}| \\
&\quad + O_P(\sqrt{N}C_{NT}^{-1})\frac{1}{N}\sum_i |e_{it}||\frac{1}{T}\sum_s f_s x_{is} g_s||\frac{1}{T}\sum_k g_k u_{ik}| \\
&= O_P(\sqrt{N}C_{NT}^{-2}).
\end{aligned}
$$

Also,

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}(\widetilde{F} - FH_1^{-1'})'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i \\
\leq\ & \frac{1}{\sqrt{N}}\sum_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}(\widetilde{F} - FH_1^{-1'})'\mathsf{diag}(X_i)M_g u_i \\
&+ O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{NT}\sum_i e_{it}^2\|u_i\|^2)^{1/2}(\frac{1}{NT}\sum_i \|(\widetilde{F} - FH_1^{-1'})'\mathsf{diag}(X_i)\|^2)^{1/2} \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2}) + O_P(\sqrt{N}C_{NT}^{-1})\|\frac{1}{N}\sum_i \frac{1}{\sqrt{T}}\mathsf{diag}(X_i)M_g u_i e_{it}\bar{D}_{\lambda i}^{-1}\| \\
\leq\ & O_P(\sqrt{N}C_{NT}^{-2}) + O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{T}\sum_s(\frac{1}{N}\sum_i x_{is}M_{g,s}' u_i e_{it}\bar{D}_{\lambda i}^{-1})^2)^{1/2} \\
=\ & O_P(\sqrt{N}C_{NT}^{-2}),
\end{aligned}
$$

where the last equality is due to $\frac{1}{T}\sum_s(\frac{1}{N}\sum_i x_{is}M_{g,s}'u_i e_{it}\bar{D}_{\lambda i}^{-1})^2 = O_P(C_{NT}^{-2})$, proved as follows:

$$
\begin{aligned}
&\frac{1}{T}\sum_s \mathsf{E}(\frac{1}{N}\sum_i x_{is}M_{g,s}'u_i e_{it}\bar{D}_{\lambda i}^{-1})^2 \\
\leq\ & \frac{1}{T}\sum_s \mathsf{E}(\frac{1}{N}\sum_i x_{is}u_{is}e_{it}\bar{D}_{\lambda i}^{-1})^2 + \frac{1}{T}\sum_s \mathsf{E}(\frac{1}{NT}\sum_i\sum_k x_{is}g_k g_s u_{ik}e_{it}\bar{D}_{\lambda i}^{-1})^2 \\
\leq\ & O(T^{-1}) + \frac{1}{T}\sum_{s\neq t}\frac{1}{N^2}\sum_{ij}\mathsf{E}|\mathsf{E}(e_{jt}e_{it}|F)\mathsf{E}(u_{is}u_{js}x_{js}x_{is}|F)| \\
&+ \frac{1}{T}\sum_s \mathsf{E}\frac{1}{NT}\sum_i\sum_k \bar{D}_{\lambda i}^{-1}\frac{1}{NT}\sum_j\sum_l \bar{D}_{\lambda j}^{-1}|x_{js}g_l g_s u_{jl}x_{is}g_k g_s u_{ik}||\mathsf{Cov}(e_{it}, e_{jt}|F)| \\
\leq\ & O(T^{-1}) + O(N^{-1})
\end{aligned}
$$

where the last equality is due to $\max_j \sum_i |\mathsf{Cov}(e_{it}, e_{jt}|F)| < C$.

Next,

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}F'\mathsf{diag}(X_i)(M_{\widetilde{g}} - M_g)u_i \\
=\ & \mathsf{tr}\frac{1}{\sqrt{N}}\sum_i u_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}F'\mathsf{diag}(X_i)(M_{\widetilde{g}} - M_g)
\end{aligned}
$$

$$\begin{aligned}
&\leq O_P(\sqrt{N}C_{NT}^{-1})\frac{1}{T}\|\frac{1}{N}\sum_i u_i e_{it}\bar{D}_{\lambda i}^{-1}F'\mathsf{diag}(X_i)\|_F \\
&\leq O_P(\sqrt{N}C_{NT}^{-1})(\frac{1}{T^2}\sum_{sk}\mathsf{E}(\frac{1}{N}\sum_i u_{is}e_{it}\bar{D}_{\lambda i}^{-1}f_k x_{ik})^2)^{1/2} = O_P(\sqrt{N}C_{NT}^{-2})
\end{aligned}$$

where the last equality is due to $\frac{1}{T^2}\sum_{sk}\mathsf{E}(\frac{1}{N}\sum_i u_{is}e_{it}\bar{D}_{\lambda i}^{-1}f_k x_{ik})^2 = O_P(C_{NT}^{-2})$.

$$\begin{aligned}
&\frac{1}{T^2}\sum_{sk}\mathsf{E}(\frac{1}{N}\sum_i u_{is}e_{it}\bar{D}_{\lambda i}^{-1}f_k x_{ik})^2 \\
=\ &\frac{1}{T^2}\sum_{sk}\mathsf{E}\frac{1}{N}\sum_i \bar{D}_{\lambda i}^{-1}\frac{1}{N}\sum_j \bar{D}_{\lambda j}^{-1}f_k^2 x_{jk}u_{is}x_{ik}u_{js}e_{jt}e_{it} \\
\leq\ &O(T^{-1}) + \frac{1}{T^2}\sum_{sk}\frac{1}{N^2}\sum_{ij}\mathsf{E}\bar{D}_{\lambda i}^{-1}\bar{D}_{\lambda j}^{-1}f_k^2\mathsf{E}(x_{jk}u_{is}x_{ik}u_{js}|F)\mathsf{Cov}(e_{jt},e_{it}|F) \\
\leq\ &O_P(C_{NT}^{-2}).
\end{aligned}$$

Step 8: since u_{it} is conditionally serially independent given E, F,

$$\begin{aligned}
&\mathsf{E}(\frac{1}{\sqrt{N}}\sum_i e_{it}\bar{D}_{\lambda i}^{-1}\frac{1}{T}F'\mathsf{diag}(X_i)M_g u_i)^2 \\
=\ &\mathsf{E}\frac{1}{N}\sum_{ij}\sum_s \bar{D}_{\lambda j}^{-1}\bar{D}_{\lambda i}^{-1}\frac{1}{T}(F'\mathsf{diag}(X_i)M_{g,s})^2 e_{jt}e_{it}\frac{1}{T}u_{js}u_{is} \\
\leq\ &\frac{C}{T^2 N}\sum_{ij}\sum_s \mathsf{E}u_{js}u_{is}\mathsf{E}((F'\mathsf{diag}(X_i)M_{g,s})^2 e_{jt}e_{it}|U) \\
\leq\ &\frac{C}{T^2 N}\sum_{ij}\sum_s |\mathsf{Cov}(u_{js},u_{is})| = O(T^{-1}) = O_P(C_{NT}^{-2}).
\end{aligned}$$

Together, $\frac{1}{\sqrt{N}}\sum_i (H_1'\lambda_i - \dot{\lambda}_i)e_{it} = O_P(\sqrt{T}C_{NT}^{-2})$.

(ii) The proof is the same as that of part (i), by substituting in the expansion of $H_1'\lambda_i - \dot{\lambda}_i$, hence we ignore it for brevity.

□

Lemma G.6. *For any bounded deterministic sequence* c_i, $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I} f_s w_s' e_{is}c_i'(\widetilde{f}_s - H_1^{-1}f_s)\|^2 = O_P(C_{NT}^{-4})$.

Proof. For $y_s = Ag_s + \mathsf{diag}(X_s)\Lambda f_s + u_s$,

$$\begin{aligned}
\widetilde{f}_s &= \widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}y_s \\
&= H_1^{-1}f_s + \widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2 - \widetilde{A})H_2^{-1}g_s \\
&\quad + \widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1 - \widetilde{\Lambda})H_1^{-1}f_s
\end{aligned}$$

$$+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s. \tag{G.7}$$

Without loss of generality, we assume $\dim(g_s)=\dim(f_s)=1$, as we can always work with their elements given that the number of factors is fixed. It suffices to prove, for $h_s:=f_sw_s$,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}(\widetilde{f}_s-H_1^{-1}f_s)\|^2=O_P(C_{NT}^{-4}).$$

We plug in each term in the expansion of $\widetilde{f}_s$:

$$\begin{aligned}&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s\|^2\\&+\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\|^2\\&+\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s\|^2.\end{aligned}$$

First term: $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s\|^2$. By Lemma G.3,

$$\begin{aligned}&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}(\widetilde{D}_{fs}^{-1}-(H_1'\bar{D}_{fs}H_1)^{-1})\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s\|^2\\ \leq\ & O_P(C_{NT}^{-4})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_s\|\frac{1}{\sqrt{N}}\widetilde{\Lambda}'\mathsf{diag}(X_s)\|^2\|g_s\|^2\|h_se_{is}\|^2\\ \leq\ & O_P(C_{NT}^{-4})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_s\|\frac{1}{\sqrt{N}}\widetilde{\Lambda}'\mathsf{diag}(X_s)\|^2\|g_s\|^2\|h_se_{is}\|^2=O_P(C_{NT}^{-4}).\\&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s\|^2\\ \leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T\sqrt{N}}\sum_{s\notin I}h_sg_se_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\widetilde{\Lambda}'\mathsf{diag}(X_s)\|^2\\ \leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\widetilde{\lambda}_j^2[(\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_sg_se_{is}e_{js})^2+(\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_sg_se_{is}w_s)^2].\end{aligned}$$

Because e_{is} is serially independent conditionally on (W,F,G), s the above is

$$\begin{aligned}\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\mathsf{E}_I(\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_sg_se_{is}e_{js})^2+\mathsf{E}_I(\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_sg_se_{is}w_s)^2]\\ =\ & O_P(C_{NT}^{-4}).\end{aligned}$$

Put together the first term is $O_P(C_{NT}^{-4})$.

Second term: $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\|^2$.

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(\widetilde{D}_{fs}^{-1}-(H_1'\bar{D}_{fs}H_1)^{-1})\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\|^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_s\|\frac{1}{\sqrt{N}}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)\|^2\|h_s\|^2\|f_s e_{is}\|^2=O_P(C_{NT}^{-2}).
\end{aligned}
$$

The same proof leads to

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)(M_{\widetilde{\alpha}}-M_\alpha)\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\|^2\\
\leq\ & O_P(C_{NT}^{-2}).
\end{aligned}
$$

Now let $M_{\widetilde{\alpha},ij}$ be the (i,j) th component of $M_{\widetilde{\alpha}}$. Let $z_{js}=\sum_k\widetilde{\lambda}_k M_{\alpha,kj}x_{ks}$,

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_\alpha\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\|^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_s f_s h_s e_{is}\bar{D}_{fs}^{-1}\frac{1}{\sqrt{N}}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_\alpha\mathsf{diag}(X_s)\|^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_s f_s h_s\bar{D}_{fs}^{-1}e_{is}z_{js}x_{js})^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_s\mathsf{E}_I(f_s h_s\bar{D}_{fs}^{-1}e_{is}x_{js}^2)^2(\widetilde{\lambda}_j M_{\alpha,jj})^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{N}\sum_{k\neq j}(\frac{1}{T}\sum_s\mathsf{E}_I f_s h_s\bar{D}_{fs}^{-1}x_{ks}e_{is}x_{js})^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\mathsf{Var}_I(\frac{1}{T}\sum_s f_s h_s\bar{D}_{fs}^{-1}e_{is}z_{js}x_{js})\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_s\mathsf{E}_I(f_s h_s\bar{D}_{fs}^{-1}e_{is}x_{js}^2)^2(\widetilde{\lambda}_j-H_1'\lambda_j)^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_s\mathsf{E}_I f_s h_s\bar{D}_{fs}^{-1}e_{is}e_{js}^2)^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_s\mathsf{E}_I f_s h_s\bar{D}_{fs}^{-1}e_{is}l_j w_s e_{js})^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{N}\sum_{k\neq j}(\frac{1}{T}\sum_s\mathsf{E}_I f_s h_s\bar{D}_{fs}^{-1}e_{ks}e_{is}e_{js})^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{N}\sum_{k\neq j}(\frac{1}{T}\sum_s\mathsf{E}_I f_s h_s\bar{D}_{fs}^{-1}e_{ks}e_{is}l_j w_s)^2
\end{aligned}
$$

$$+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{N}\sum_{k\neq j}(\frac{1}{T}\sum_s \mathsf{E}_I f_s h_s \bar{D}_{fs}^{-1} l_k w_s e_{is}e_{js})^2+O_P(C_{NT}^{-4})=O_P(C_{NT}^{-4})$$

given that $\sum_j |\mathsf{E}_I(e_{is}e_{js}|f_s,w_s)|+\frac{1}{N}\sum_{k\neq j}|\mathsf{E}_I(e_{ks}e_{is}e_{js}|f_s,w_s)|<\infty$. Put together, the second term is $O_P(C_{NT}^{-4})$.

Third term: $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s\|^2$.

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(\widetilde{D}_{fs}^{-1}-(H_1'\bar{D}_{fs}H_1)^{-1})\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s\|^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{s\notin I}|h_s e_{is}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s|^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I|h_s e_{is}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s|^2\\
=\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I h_s^2 e_{is}^2\frac{1}{N^2}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{E}_I(u_s u_s'|X_s,f_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)\widetilde{\Lambda}\\
\leq\ & O_P(C_{NT}^{-2}N^{-1})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I h_s^2 e_{is}^2\frac{1}{N}\|\widetilde{\Lambda}'\mathsf{diag}(X_s)\|^2\\
=\ & O_P(C_{NT}^{-2}N^{-1})=O_P(C_{NT}^{-4}).
\end{aligned}$$

Next, due to $\mathsf{E}(u_s|f_s,D_I,e_s)=0$, and e_s is conditionally serially independent given (f_s,u_s),

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)(M_{\widetilde{\alpha}}-M_\alpha)u_s\|^2\\
=\ & \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\mathsf{tr}\frac{1}{T}\sum_{s\notin I}u_s h_s e_{is}(H_1'D_xH_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)(M_{\widetilde{\alpha}}-M_\alpha)]^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}u_s h_s e_{is}(H_1'\bar{D}_{fs}^{-1}H_1)^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)\|_F^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\mathsf{E}_I\|\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}u_s h_s e_{is}\frac{1}{\sqrt{N}}x_{js}\|_F^2\widetilde{\lambda}_j^2\\
\leq\ & O_P(C_{NT}^{-2})\max_{ij}\mathsf{E}_I\mathsf{E}_I(\|\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}u_s h_s e_{is}\frac{1}{\sqrt{N}}x_{js}\|_F^2|X,F,W)\\
\leq\ & O_P(C_{NT}^{-2})\max_{ijk}\frac{1}{T}\mathsf{Var}_I(\bar{D}_{fs}^{-1}u_{ks}h_s e_{is}x_{js})\leq O_P(C_{NT}^{-2}T^{-1})=O_P(C_{NT}^{-4}).
\end{aligned}$$

Next,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_s e_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}(\widetilde{\Lambda}-\Lambda H_1)'\mathsf{diag}(X_s)M_\alpha u_s\|^2$$

$$
\begin{aligned}
&\leq O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{N}\sum_j(\widetilde{\lambda}_j-H_1'\lambda_j)\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_se_{is}M_{\alpha,j}'u_sx_{js}\|^2\\
&\leq O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j(\frac{1}{T}\sum_{s\notin I}\bar{D}_{fs}^{-1}h_se_{is}M_{\alpha,j}'u_sx_{js})^2\\
&\leq O_P(C_{NT}^{-2}T^{-1})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I\bar{D}_{fs}^{-2}h_s^2x_{js}^2e_{is}^2M_{\alpha,j}'\mathsf{Var}_I(u_s|e_s,w_s,f_s)M_{\alpha,j}\\
&\leq O_P(C_{NT}^{-2}T^{-1})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{N}\sum_j\frac{1}{T}\sum_{s\notin I}\mathsf{E}_I\bar{D}_{fs}^{-2}h_s^2x_{js}^2e_{is}^2\|M_{\alpha,j}\|^2\\
&= O_P(C_{NT}^{-2}T^{-1})=O_P(C_{NT}^{-4}).
\end{aligned}
$$

Finally, due to the conditional serial independence,

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}h_se_{is}(H_1'\bar{D}_{fs}H_1)^{-1}\frac{1}{N}H_1'\Lambda'\mathsf{diag}(X_s)M_\alpha u_s\|^2\\
&\leq O_P(N^{-2}T^{-2})\sum_{s\notin I}\mathsf{E}_I\bar{D}_{fs}^{-2}h_s^2e_{is}^2\Lambda'\mathsf{diag}(X_s)M_\alpha\mathsf{E}_I(u_su_s'|E,G,F)M_\alpha\mathsf{diag}(X_s)\Lambda\\
&\leq O_P(N^{-2}T^{-1})\mathsf{E}_I\bar{D}_{fs}^{-2}h_s^2e_{is}^2\|\Lambda'\mathsf{diag}(X_s)\|^2=O_P(N^{-1}T^{-1})=O_P(C_{NT}^{-4}).
\end{aligned}
$$

Put together, the third term is $O_P(C_{NT}^{-4})$. This finishes the proof. □

G.3. **Technical lemmas for $\widehat{f}_t$.**

Lemma G.7. *Assume $\frac{1}{N}\sum_{ij}|\mathsf{Cov}(e_{it}^2,e_{jt}^2)|<C$. For each fixed t,*
(i) $\widehat{B}_t-B=O_P(C_{NT}^{-1})$.
(ii) The upper two blocks of $\widehat{B}_t^{-1}\widehat{S}_t-B^{-1}S$ are both $O_P(C_{NT}^{-2})$.

Proof. Throughout the proof, we assume $\dim(\alpha_i)=\dim(\lambda_i)=1$ without loss of generality.

(i) $\widehat{B}_t-B=b_1+b_2$, where

$$
\begin{aligned}
b_1 &= \frac{1}{N}\sum_i\begin{pmatrix}\widetilde{\lambda}_i\widetilde{\lambda}_i'\widehat{e}_{it}^2-H_1'\lambda_i\lambda_i'H_1e_{it}^2 & \widetilde{\lambda}_i\widetilde{\alpha}_i'\widehat{e}_{it}-H_1'\lambda_i\alpha_iH_2'e_{it}\\ \widetilde{\alpha}_i\widetilde{\lambda}_i'\widehat{e}_{it}-H_2'\alpha_i\lambda_iH_1'e_{it} & \widetilde{\alpha}_i\widetilde{\alpha}_i'-H_2'\alpha_i\alpha_i'H_2\end{pmatrix}\\
b_2 &= \frac{1}{N}\sum_i\begin{pmatrix}H_1'\lambda_i\lambda_i'H_1(e_{it}^2-\mathsf{E}e_{it}^2) & H_1'\lambda_i\alpha_iH_2'e_{it}\\ H_2'\alpha_i\lambda_iH_1'e_{it} & 0\end{pmatrix}.
\end{aligned}
$$

To prove the convergence of b_1, first note that

$$
\begin{aligned}
&\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2-e_{it}^2)=\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}-e_{it})^2+\frac{2}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}-e_{it})e_{it}\\
&\leq \frac{1}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)(\widetilde{\lambda}_i-H_1'\lambda_i)'(\widehat{e}_{it}-e_{it})^2+\frac{1}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)\lambda_i'H_1'(\widehat{e}_{it}-e_{it})^2
\end{aligned}
$$

$$
\begin{aligned}
&+O_P(1)\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2+\frac{2}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)(\widetilde{\lambda}_i-H_1'\lambda_i)'(\widehat{e}_{it}-e_{it})e_{it} \\
&+\frac{2}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)\lambda_i'H_1'(\widehat{e}_{it}-e_{it})e_{it}+O_P(1)\frac{1}{N}\sum_i\lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})e_{it} \\
\leq\;& O_P(C_{NT}^{-2})\max_{it}|\widehat{e}_{it}-e_{it}|\max_{it}|e_{it}|+O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4)^{1/2} \\
&+O_P(1)\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2+O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2e_{it}^2)^{1/2} \\
&+O_P(1)\frac{1}{N}\sum_i\lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})e_{it}=O_P(C_{NT}^{-2}),
\end{aligned}
\tag{G.8}
$$

by Lemma G.1. In addition, still by Lemma G.1,

$$
\begin{aligned}
&\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\alpha}_i'(\widehat{e}_{it}-e_{it})=\frac{1}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)(\widetilde{\alpha}_i-H_2'\alpha_i)'(\widehat{e}_{it}-e_{it}) \\
&+\frac{1}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)\alpha_i'H_2(\widehat{e}_{it}-e_{it})+O_P(1)\frac{1}{N}\sum_i\lambda_i\alpha_i'(\widehat{e}_{it}-e_{it}) \\
\leq\;& O_P(C_{NT}^{-2})\max_{it}|\widehat{e}_{it}-e_{it}|+O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2)^{1/2} \\
&+O_P(1)\frac{1}{N}\sum_i\lambda_i\alpha_i'(\widehat{e}_{it}-e_{it})=O_P(C_{NT}^{-2}).
\end{aligned}
\tag{G.9}
$$

So the first term of b_1 is, due to the serial independence of e_{it}^2,

$$
\begin{aligned}
&\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'\widehat{e}_{it}^2-H_1'\lambda_i\lambda_i'H_1e_{it}^2 \\
\leq\;& \frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2-e_{it}^2)+\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1)e_{it}^2 \\
\leq\;& O_P(C_{NT}^{-2})+O_P(1)\frac{1}{N}\sum_i\|\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1\|_F\mathsf{E}_Ie_{it}^2 \\
\leq\;& O_P(C_{NT}^{-2})+O_P(1)\frac{1}{N}\sum_i\|\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1\|_F=O_P(C_{NT}^{-1}).
\end{aligned}
$$

The second term is,

$$
\begin{aligned}
&\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\alpha}_i'\widehat{e}_{it}-H_1'\lambda_i\alpha_iH_2'e_{it} \\
\leq\;& O_P(C_{NT}^{-1})+O_P(1)\frac{1}{N}\sum_i\|\widetilde{\lambda}_i\widetilde{\alpha}_i'-H_1'\lambda_i\alpha_iH_2'\|_F\mathsf{E}_I|e_{it}|\leq O_P(C_{NT}^{-1}).
\end{aligned}
$$

The third term of b_1 is bounded similarly. The last term of b_1 is easy to show to be $O_P(C_{NT}^{-1})$. As for b_2, by the assumption that $\frac{1}{N}\sum_{ij}|\mathsf{Cov}(e_{it}^2, e_{jt}^2)| < C$, thus $b_2 = O_P(N^{-1/2})$. Hence $\widehat{B}_t - B = O_P(C_{NT}^{-1})$.

(ii) We have

$$B_t^{-1}\widehat{S}_t - B^{-1}S = (B_t^{-1} - B^{-1})(\widehat{S}_t - S) + (B_t^{-1} - B^{-1})S + B^{-1}(\widehat{S}_t - S).$$

We first bound the four blocks of $\widehat{S}_t - S$. We have $\widehat{S}_t - S = c_t + d_t$,

$$\begin{aligned} c_t &= \frac{1}{N}\sum_i \begin{pmatrix} \widetilde{\lambda}_i\lambda_i' H_1 e_{it}(\widehat{e}_{it} - e_{it}) + \widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2 - e_{it}^2) & \widetilde{\lambda}_i(\widehat{e}_{it} - e_{it})(\alpha_i' H_2 - \widetilde{\alpha}_i') \\ \widetilde{\alpha}_i\lambda_i' H_1(e_{it} - \widehat{e}_{it}) + \widetilde{\alpha}_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(\widehat{e}_{it} - e_{it}) & 0 \end{pmatrix} \\ d_t &= \frac{1}{N}\sum_i \begin{pmatrix} (\widetilde{\lambda}_i\lambda_i' H_1 - \widetilde{\lambda}_i\widetilde{\lambda}_i')e_{it}^2 - H_1'\lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')\mathsf{E}e_{it}^2 & \widetilde{\lambda}_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i') \\ \widetilde{\alpha}_i(\lambda_i' H_1 - \widetilde{\lambda}_i')e_{it} & 0 \end{pmatrix}. \end{aligned}$$

Call each block of c_t to be $c_{t,1}...c_{t,4}$ in the clockwise order. Note that $c_{t,4} = 0$.

As for $c_{t,1}$, it follows from Lemma G.1 that

$$\begin{aligned} &\frac{1}{N}\sum_i \widetilde{\lambda}_i\lambda_i' H_1 e_{it}(\widehat{e}_{it} - e_{it}) \leq (\frac{1}{N}\sum_i e_{it}^2(\widehat{e}_{it} - e_{it})^2)^{1/2} O_P(C_{NT}^{-1}) \\ &+ O_P(1)\frac{1}{N}\sum_i \lambda_i\lambda_i e_{it}(\widehat{e}_{it} - e_{it}) \leq O_P(C_{NT}^{-2}). \end{aligned}$$

We have also shown $\frac{1}{N}\sum_i \widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2 - e_{it}^2) = O_P(C_{NT}^{-2})$. Thus $c_{t,1} = O_P(C_{NT}^{-2})$.

For $c_{t,2}$, from Lemma G.1,

$$\begin{aligned} &\frac{1}{N}\sum_i \widetilde{\lambda}_i(\widehat{e}_{it} - e_{it})(\alpha_i' H_2 - \widetilde{\alpha}_i') \\ \leq\ & O_P(C_{NT}^{-2})\max_i |e_{it} - \widehat{e}_{it}| + O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i (\widehat{e}_{it} - e_{it})^2)^{1/2} = O_P(C_{NT}^{-2}). \end{aligned}$$

For the third term of c_t, similarly, $\frac{1}{N}\sum_i \widetilde{\alpha}_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(\widehat{e}_{it} - e_{it}) = O_P(C_{NT}^{-2})$. Also, by Lemma G.1,

$$\begin{aligned} &\frac{1}{N}\sum_i \widetilde{\alpha}_i\lambda_i' H_1(e_{it} - \widehat{e}_{it}) \\ \leq\ & O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i (\widehat{e}_{it} - e_{it})^2)^{1/2} + O_P(1)\frac{1}{N}\sum_i \lambda_i\alpha_i'(\widehat{e}_{it} - e_{it}) \\ \leq\ & O_P(C_{NT}^{-1}). \end{aligned}$$

So $c_{t,3} = O_P(C_{NT}^{-1})$.

As for d_t, we first prove that $\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(e_{it}^2 - \mathsf{E}e_{it}^2) = O_P(C_{NT}^{-1}N^{-1/2})$. Note that $\mathsf{E}_I\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(e_{it}^2 - \mathsf{E}e_{it}^2) = 0$. Let Υ_t be an $N\times 1$ vector of e_{it}^2, and $\mathsf{diag}(\Lambda)$ be diagonal matrix consisting of elements of Λ. Then

$$\|\mathsf{Var}(\Upsilon_t)\| \le \max_i \sum_j |\mathsf{Cov}(e_{it}^2, e_{jt}^2)| < C,$$

and $\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')e_{it}^2 = \frac{1}{N}(H_1\Lambda - \widetilde{\Lambda})'\mathsf{diag}(\Lambda)\Upsilon_t$. So

$$\begin{aligned} &\mathsf{Var}_I(\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')e_{it}^2) = \frac{1}{N^2}(H_1\Lambda - \widetilde{\Lambda})'\mathsf{diag}(\Lambda)\mathsf{Var}(\Upsilon_t)\mathsf{diag}(\Lambda)(H_1\Lambda - \widetilde{\Lambda}) \\ \le\ & C\frac{1}{N^2}\|H_1\Lambda - \widetilde{\Lambda}\|_F^2 = O_P(C_{NT}^{-2}N^{-1}). \end{aligned}$$

This implies $\frac{1}{N}\sum_i \lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(e_{it}^2 - \mathsf{E}e_{it}^2) = O_P(C_{NT}^{-1}N^{-1/2})$.

Thus the first term of d_t is

$$\begin{aligned} &\frac{1}{N}\sum_i (\widetilde{\lambda}_i\lambda_i' H_1 - \widetilde{\lambda}_i\widetilde{\lambda}_i')e_{it}^2 - H_1'\lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')\mathsf{E}e_{it}^2 \\ \le\ & \frac{1}{N}\sum_i (\widetilde{\lambda}_i - H_1'\lambda_i)(\lambda_i' H_1 - \widetilde{\lambda}_i')e_{it}^2 + \frac{1}{N}\sum_i H_1'\lambda_i(\lambda_i' H_1 - \widetilde{\lambda}_i')(e_{it}^2 - \mathsf{E}e_{it}^2) \\ =\ & O_P(1)\frac{1}{N}\sum_i (\widetilde{\lambda}_i - H_1'\lambda_i)(\lambda_i' H_1 - \widetilde{\lambda}_i')\mathsf{E}_I e_{it}^2 + O_P(C_{NT}^{-1}N^{-1/2}) = O_P(C_{NT}^{-2}). \end{aligned}$$

As for the second term of d_t, $\frac{1}{N}\sum_i \widetilde{\lambda}_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i')$, note that

$$\begin{aligned} &\frac{1}{N}\sum_i (\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i') = O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i e_{it}^2(\alpha_i' H_2 - \widetilde{\alpha}_i')^2)^{1/2} \\ \le\ & O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i (\alpha_i' H_2 - \widetilde{\alpha}_i')^2\mathsf{E}_I e_{it}^2)^{1/2} = O_P(C_{NT}^{-2}). \end{aligned}$$

And, $\mathsf{E}_I\frac{1}{N}\sum_i \widetilde{\lambda}_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i') = 0$.

$$\begin{aligned} &\mathsf{Var}_I(\frac{1}{N}\sum_i \lambda_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i')) = \frac{1}{N^2}\mathsf{Var}_I((AH_2 - \widetilde{A})'\mathsf{diag}(\Lambda)e_t) \\ \le\ & \frac{1}{N^2}(AH_2 - \widetilde{A})'\mathsf{diag}(\Lambda)\mathsf{Var}(e_t)\mathsf{diag}(\Lambda)(AH_2 - \widetilde{A}) = O_P(C_{NT}^{-2}N^{-1}), \end{aligned}$$

implying $\frac{1}{N}\sum_i \widetilde{\lambda}_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i') = O_P(C_{NT}^{-2})$.

Finally the third term of d_t, $\frac{1}{N}\sum_i \widetilde{\lambda}_i e_{it}(\alpha_i' H_2 - \widetilde{\alpha}_i')$, is bounded similarly. So $d_t = O_P(C_{NT}^{-2})$. Put together, we have: $\widehat{S}_t - S = c_t + d_t = O_P(C_{NT}^{-1})$. On the other hand, the upper two blocks of $\widehat{S}_t - S$ are $O_P(C_{NT}^{-2})$, determined by $c_{t,1}, c_{t,2}$ and the upper blocks of d_t. In addition, note that both B, S are block diagonal matrices, and the diagonal blocks of S are $O_P(C_{NT}^{-1})$. Due to

$$\widehat{B}_t^{-1}\widehat{S}_t - B^{-1}S = (\widehat{B}_t^{-1} - B^{-1})(\widehat{S}_t - S) + (\widehat{B}_t^{-1} - B^{-1})S + B^{-1}(\widehat{S}_t - S),$$

hence the upper blocks of $\widehat{B}_t^{-1}\widehat{S}_t - B^{-1}S$ are both $O_P(C_{NT}^{-2})$.

□

Lemma G.8. *Suppose* $\max_{t\leq T} |\frac{1}{N}\sum_i \alpha_i\lambda_i e_{it}| = O_P(1)$, $C_{NT}^{-1}\max_{it}|e_{it}|^2 + \max_t \|\frac{1}{N}\sum_i \lambda_i\lambda_i'\bar{e}_i e_{it}\|_F = o_P(1)$. *Then*

(i) $\max_{t\leq T}\|\widehat{B}_t^{-1}\| = O_P(1)$, $\max_{t\leq T}|\widehat{B}_t - B_t| = o_P(1)$.

(ii) $\frac{1}{T}\sum_{s\notin I}\|\widehat{B}_s^{-1} - B^{-1}\|^2 = O_P(C_{NT}^{-2})$.

(iii) Write

$$\widehat{S}_t - S = \begin{pmatrix}\Delta_{t1}\\ \Delta_{t2}\end{pmatrix},$$

whose partition matches with that of $(f_t', g_t')'$. *Then* $\frac{1}{T}\sum_{t\notin I}\|\Delta_{t1}\|^2 = O_P(C_{NT}^{-4})$ *and* $\frac{1}{T}\sum_{t\notin I}\|\Delta_{t2}\|^2 = O_P(C_{NT}^{-2})$.

(iv) $\max_{t\leq T}\|\widehat{S}_t\| = O_P(1)$ *and* $\max_{t\leq T}\|\widehat{S}_t - S\| = o_P(1)$.

Proof. Define

$$B_t = \frac{1}{N}\sum_i \begin{pmatrix} H_1'\lambda_i\lambda_i'H_1 e_{it}^2 & 0\\ 0 & H_2'\alpha_i\alpha_i'H_2\end{pmatrix}.$$

Then $\widehat{B}_t - B_t = b_{1t} + b_{2t}$,

$$\begin{aligned} b_{1t} &= \frac{1}{N}\sum_i \begin{pmatrix}\widetilde{\lambda}_i\widetilde{\lambda}_i'\widehat{e}_{it}^2 - H_1'\lambda_i\lambda_i'H_1e_{it}^2 & \widetilde{\lambda}_i\widetilde{\alpha}_i'\widehat{e}_{it} - H_1'\lambda_i\alpha_iH_2'e_{it}\\ \widetilde{\alpha}_i\widetilde{\lambda}_i'\widehat{e}_{it} - H_2'\alpha_i\lambda_iH_1'e_{it} & \widetilde{\alpha}_i\widetilde{\alpha}_i' - H_2'\alpha_i\alpha_i'H_2\end{pmatrix}\\ b_{2t} &= \frac{1}{N}\sum_i \begin{pmatrix} 0 & H_1'\lambda_i\alpha_iH_2'e_{it}\\ H_2'\alpha_i\lambda_iH_1'e_{it} & 0\end{pmatrix}.\end{aligned}$$

(i) We now show $\max_{t\leq T}|b_{1t}| = o_P(1)$. In addition, by assumption $\max_{t\leq T}|b_{2t}| = o_P(1)$. Thus $\max_{t\leq T}|\widehat{B}_t - B_t| = o_P(1)$, and thus $\max_{t\leq T}\|\widehat{B}_t^{-1}\| = o_P(1) + \max_{t\leq T}\|B_t^{-1}\| = O_P(1)$, by the assumption that $\|B_t^{-1}\| = O_P(1)$ uniformly in t. To show $\max_{t\leq T}|b_{1t}| = o_P(1)$, note that:

First term:

$$\max_t \|\frac{1}{N}\sum_i \widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2 - e_{it}^2)\|$$

$$
\begin{aligned}
&\leq O_P(1)(\frac{1}{N}\sum_i\|\widetilde{\lambda}_i-H_1'\lambda_i\|^2)^{1/2}\max_t[\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4+(\widehat{e}_{it}-e_{it})^2e_{it}^2]^{1/2}\\
&\quad+\frac{1}{N}\sum_i\|\widetilde{\lambda}_i-H_1'\lambda_i\|^2[2\max_t|(\widehat{e}_{it}-e_{it})e_{it}|+\max_{t\leq T}(\widehat{e}_{it}-e_{it})^2]\\
&\quad+O_P(1)\max_t\|\frac{1}{N}\sum_i\lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})e_{it}\|_F\\
&\quad+O_P(1)\max_t\|\frac{1}{N}\sum_i\lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})^2\|_F\\
&\leq O_P(1+\max_{t\leq T}\|w_t\|^2+b_{NT,2}^2)C_{NT}^{-2}+O_P(b_{NT,1}^2+b_{NT,3}^2)\\
&\quad+O_P(1+\max_{t\leq T}\|w_t\|+b_{NT,2})\max_{it}|e_{it}|C_{NT}^{-1}+O_P(b_{NT,1}+b_{NT,3})(\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2)^{1/2}C_{NT}^{-1}\\
&\quad+O_P(\phi_{NT}^2C_{NT}^{-2})+\phi_{NT}\max_{it}|e_{it}|C_{NT}^{-2}+\max_{t\leq T}\|\frac{1}{N}\sum_i c_ie_{it}l_i\|_FO_P(b_{NT,1}+b_{NT,3})\\
&= o_P(1),
\end{aligned}
$$

given assumptions $C_{NT}^{-1}(b_{NT,2}+\max_{t\leq T}\|w_t\|)\max_{it}|e_{it}|=o_P(1)$, $\phi_{NT}\max_{it}|e_{it}|=O_P(1)$, $\max_{t\leq T}\|\frac{1}{N}\sum_i e_{it}\alpha_i\lambda_i'\|_F=o_P(1)$, and $(b_{NT,1}+b_{NT,3})[(\max_{t\leq T}\frac{1}{N}\sum_i e_{it}^2)^{1/2}C_{NT}^{-1}+1]=o_P(1)$.

In addition, $\max_{t\leq T}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1)e_{it}^2\|\leq O_P(C_{NT}^{-1})\max_{it}e_{it}^2=o_P(1)$. So the first term of $\max_{t\leq T}|b_{1t}|$ is $o_P(1)$.

Second term,

$$
\begin{aligned}
&\max_t\|\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\alpha}_i'(\widehat{e}_{it}-e_{it})\|_F\\
&\leq \max_{it}|\widehat{e}_{it}-e_{it}|\frac{1}{N}\sum_i\|(\widetilde{\lambda}_i-H_1'\lambda_i)(\widetilde{\alpha}_i-H_2'\alpha_i)'\|_F\\
&\quad+O_P(1)(\frac{1}{N}\sum_i\|\widetilde{\lambda}_i-H_1'\lambda_i\|^2+(\frac{1}{N}\sum_i\|\widetilde{\alpha}_i-H_2'\alpha_i\|^2)^{1/2}\max_t(\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2)^{1/2}\\
&\quad+O_P(1)\max_t\|\frac{1}{N}\sum_i\lambda_i\alpha_i'(\widehat{e}_{it}-e_{it})\|_F\\
&\leq O_P(\phi_{NT}+1+\max_{t\leq T}\|w_t\|)C_{NT}^{-2}+O_P(b_{NT,1}+b_{NT,3}+C_{NT}^{-1}b_{NT,2})=o_P(1).
\end{aligned}
$$

Next,

$$
\max_{t\leq T}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widetilde{\alpha}_i'-H_1'\lambda_i\alpha_iH_2')e_{it}\|_F=O_P(C_{NT}^{-1})\max_{it}|e_{it}|=o_P(1).
$$

So the second term of $\max_{t\leq T}|b_{1t}|$ is $o_P(1)$. Similarly, the third term is also $o_P(1)$. Finally, the last term $\|\frac{1}{N}\sum_i\widetilde{\alpha}_i\widetilde{\alpha}_i'-H_2'\alpha_i\alpha_i'H_2\|_F=o_P(1)$.

(ii) Because we have proved $\max_{t\leq T}\|\widehat{B}_t^{-1}\|=O_P(1)$, it suffices to prove $\frac{1}{T}\sum_{s\notin I}\|\widehat{B}_s-B\|_F^2=O_P(C_{NT}^{-2})$, or $\frac{1}{T}\sum_{s\notin I}\|b_{1t}\|_F^2=O_P(C_{NT}^{-2})=\frac{1}{T}\sum_{s\notin I}\|b_{2t}\|_F^2$.

First term of b_{1t}: by (G.8), the first term is bounded by, by Lemma G.2,

$$
\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'\widehat{e}_{it}^2-H_1'\lambda_i\lambda_i'H_1e_{it}^2\|_F^2\\
\leq\;&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2-e_{it}^2)\|_F^2+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1)e_{it}^2\|_F^2\\
\leq\;&O_P(C_{NT}^{-4})\max_{it}|\widehat{e}_{it}-e_{it}|^2\max_{it}|e_{it}|^2+O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^4\\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}(\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2)^2+O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}\frac{1}{N}\sum_i(\widehat{e}_{it}-e_{it})^2e_{it}^2\\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}(\frac{1}{N}\sum_i\lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})e_{it})^2\\
&+O_P(1)\mathsf{tr}(\frac{1}{N^2}\sum_{ij}(\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1)(\widetilde{\lambda}_j\widetilde{\lambda}_j'-H_1'\lambda_j\lambda_i'H_1)\frac{1}{T}\sum_{t\notin I}\mathsf{E}_Ie_{jt}^2e_{it}^2)\\
\leq\;&O_P(C_{NT}^{-2})+O_P(1)(\frac{1}{N}\sum_i\|(\widetilde{\lambda}_i\widetilde{\lambda}_i'-H_1'\lambda_i\lambda_i'H_1\|_F)^2\\
\leq\;&O_P(C_{NT}^{-4}\phi_{NT}^2)\max_{it}|e_{it}|^2+O_P(C_{NT}^{-2})=O_P(C_{NT}^{-2}).
\end{aligned}
\tag{G.10}
$$

Second term of b_{1t}:

$$
\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\alpha}_i'\widehat{e}_{it}-H_1'\lambda_i\alpha_iH_2'e_{it}\|_F^2\\
\leq\;&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\lambda}_i\widetilde{\alpha}_i'(\widehat{e}_{it}-e_{it})\|_F^2+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widetilde{\alpha}_i'-H_1'\lambda_i\alpha_iH_2')e_{it}\|_F^2\\
\leq\;&O_P(C_{NT}^{-4})\max_i\frac{1}{T}\sum_t(\widehat{e}_{it}-e_{it})^2+O_P(C_{NT}^{-2})\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^2\\
&+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\lambda_i\alpha_i'(\widehat{e}_{it}-e_{it})\|_F^2\\
&+O_P(1)\mathsf{tr}\frac{1}{N^2}\sum_{ij}(\widetilde{\lambda}_i\widetilde{\alpha}_i'-H_1'\lambda_i\alpha_iH_2')(\widetilde{\lambda}_j\widetilde{\alpha}_j'-H_1'\lambda_j\alpha_jH_2')\frac{1}{T}\sum_{t\notin I}\mathsf{E}_Ie_{it}e_{jt}\\
\leq\;&O_P(b_{NT,4}^2+b_{NT,5}^2+C_{NT}^{-2})C_{NT}^{-4}+O_P(C_{NT}^{-2})=O_P(C_{NT}^{-2}).
\end{aligned}
\tag{G.11}
$$

The third term of b_{1t} is bounded similarly. Finally, it is straightforward to see that fourth term of b_{1t} is $\|\frac{1}{N}\sum_i\widetilde{\alpha}_i\widetilde{\alpha}_i'-H_2'\alpha_i\alpha_i'H_2\|_F^2=O_P(C_{NT}^{-2})$. Thus $\frac{1}{T}\sum_{s\notin I}\|b_{1t}\|_F^2=O_P(C_{NT}^{-2})$.

As for b_{2t}, it suffices to prove

$$\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \alpha_i\lambda_i' e_{it}\|_F^2 = O_P(1)\frac{1}{T}\sum_{t\notin I}\mathsf{E}\|\frac{1}{N}\sum_i \alpha_i\lambda_i' e_{it}\|_F^2 = O_P(N^{-1}).$$

Thus $\frac{1}{T}\sum_{s\notin I}\|b_{2t}\|_F^2 = O_P(C_{NT}^{-2})$.

(iii) Note that $\widehat{S}_t - S = c_t + d_t$, where

$$\begin{aligned}
c_t &= \frac{1}{N}\sum_i \begin{pmatrix} \widetilde\lambda_i\lambda_i' H_1 e_{it}(\widehat e_{it}-e_{it}) + \widetilde\lambda_i\widetilde\lambda_i'(\widehat e_{it}^2 - e_{it}^2) & \widetilde\lambda_i(\widehat e_{it}-e_{it})(\alpha_i' H_2 - \widetilde\alpha_i') \\ \widetilde\alpha_i\lambda_i' H_1(e_{it}-\widehat e_{it}) + \widetilde\alpha_i(\lambda_i' H_1 - \widetilde\lambda_i')(\widehat e_{it}-e_{it}) & 0 \end{pmatrix} \\
d_t &= \frac{1}{N}\sum_i \begin{pmatrix} (\widetilde\lambda_i\lambda_i' H_1 - \widetilde\lambda_i\widetilde\lambda_i')e_{it}^2 - H_1'\lambda_i(\lambda_i' H_1 - \widetilde\lambda_i')\mathsf{E}e_{it}^2 & \widetilde\lambda_i e_{it}(\alpha_i' H_2 - \widetilde\alpha_i') \\ \widetilde\alpha_i(\lambda_i' H_1 - \widetilde\lambda_i')e_{it} & 0 \end{pmatrix}.
\end{aligned}$$

As for the upper blocks of c_t, first note that

$$\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}(\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2|e_{it}|)^2 \\
=\ & O_P(1)\frac{1}{N^2}\sum_{ij}\|\widetilde\lambda_i - H_1'\lambda_i\|^2\|\widetilde\lambda_j - H_1'\lambda_j\|^2\frac{1}{T}\sum_{t\notin I}\mathsf{E}_I|e_{jt}e_{it}| \\
=\ & O_P(1)(\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2)^2 = O_P(C_{NT}^{-4}).
\end{aligned}$$

Then by Lemma G.2,

$$\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde\lambda_i\lambda_i' H_1 e_{it}(\widehat e_{it}-e_{it})\|_F^2 \\
\leq\ & O_P(1)\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2\frac{1}{NT}\sum_i\sum_{t\notin I}e_{it}^2(\widehat e_{it}-e_{it})^2 \\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\lambda_i\lambda_i e_{it}(\widehat e_{it}-e_{it})\|_F^2 \\
=\ & O_P(C_{NT}^{-4}) \\
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde\lambda_i\widetilde\lambda_i'(\widehat e_{it}^2 - e_{it}^2)\|_F^2 \\
\leq\ & (\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2)^2\max_{it}[(\widehat e_{it}-e_{it})^4] \\
&+\frac{1}{T}\sum_{t\notin I}(\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2|e_{it}|)^2\max_{it}|\widehat e_{it}-e_{it}|^2 \\
&+\frac{1}{N}\sum_i\|\widetilde\lambda_i - H_1'\lambda_i\|^2\frac{1}{NT}\sum_{t\notin I}\sum_i[(\widehat e_{it}-e_{it})^4 + (\widehat e_{it}-e_{it})^2 e_{it}^2]
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})e_{it}\|_F^2+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \lambda_i\lambda_i'(\widehat{e}_{it}-e_{it})^2\|_F^2 \\
=\ & O_P(C_{NT}^{-4}).
\end{aligned}
$$

Also,

$$
\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \widetilde{\lambda}_i(\widehat{e}_{it}-e_{it})(\alpha_i'H_2-\widetilde{\alpha}_i')\|_F^2 \\
\leq\ & \frac{1}{N}\sum_i\|\alpha_i'H_2-\widetilde{\alpha}_i'\|^2\frac{1}{N}\sum_i\|\widetilde{\lambda}_i-H_1'\lambda_i\|^2\max_{it}(\widehat{e}_{it}-e_{it})^2 \\
&+O_P(1)\frac{1}{N}\sum_i\|\alpha_i'H_2-\widetilde{\alpha}_i'\|^2\frac{1}{N}\sum_i\frac{1}{T}\sum_{t\notin I}(\widehat{e}_{it}-e_{it})^2=O_P(C_{NT}^{-4}).
\end{aligned}
$$

Similarly, $\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\alpha}_i(\lambda_i'H_1-\widetilde{\lambda}_i')(\widehat{e}_{it}-e_{it})\|_F^2=O_P(C_{NT}^{-4})$. Finally,

$$
\begin{aligned}
\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\widetilde{\alpha}_i\lambda_i'H_1(e_{it}-\widehat{e}_{it})\|_F^2 &= O_P(C_{NT}^{-4})+O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\alpha_i\lambda_i'(e_{it}-\widehat{e}_{it})\|_F^2 \\
&= O_P(C_{NT}^{-4}).
\end{aligned}
$$

So if we let the two upper blocks of c_t be $c_{t,1},c_{t,2}$, and the third block of c_t be $c_{t,3}$, then $\frac{1}{T}\sum_{t\notin I}\|c_{t,1}\|_F^2+\frac{1}{T}\sum_{t\notin I}\|c_{t,2}\|_F^2=O_P(C_{NT}^{-4})$ while $\frac{1}{T}\sum_{t\notin I}\|c_{t,3}\|_F^2=O_P(C_{NT}^{-2})$.

As for d_t, let $d_{t,1},...,d_{t,3}$ denote the nonzero blocks. Then $d_{t,1}$ depends on $\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i\lambda_i(\lambda_i'H_1-\widetilde{\lambda}_i')(e_{it}^2-\mathsf{E}e_{it}^2)\|_F^2$. Let Υ_t be an $N\times 1$ vector of e_{it}^2, and $\mathsf{diag}(\Lambda)$ be diagonal matrix consisting of elements of Λ. Suppose $\dim(\lambda_i)=1$ (focus on each element), then $\frac{1}{T}\sum_{t\notin I}\|d_{t,1}\|_F^2$ is bounded by

$$
\begin{aligned}
&\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i\lambda_i'H_1-\widetilde{\lambda}_i\widetilde{\lambda}_i')e_{it}^2-H_1'\lambda_i(\lambda_i'H_1-\widetilde{\lambda}_i')\mathsf{E}e_{it}^2\|_F^2 \\
\leq\ & \frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i-H_1'\lambda_i)(\lambda_i'H_1-\widetilde{\lambda}_i')e_{it}^2\|_F^2 \\
&+\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i H_1'\lambda_i(\lambda_i'H_1-\widetilde{\lambda}_i')(e_{it}^2-\mathsf{E}e_{it}^2)\|_F^2 \\
\leq\ & O_P(1)\frac{1}{N^2}\sum_{ij}\|\widetilde{\lambda}_i-H_1'\lambda_i\|^2\|\widetilde{\lambda}_j-H_1'\lambda_j\|^2\frac{1}{T}\sum_{t\notin I}\mathsf{E}_I e_{jt}^2e_{it}^2 \\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}\frac{1}{N^2}(\widetilde{\Lambda}-\Lambda H_1)'\mathsf{diag}(\Lambda)\mathsf{Var}_I(\Upsilon_t)\mathsf{diag}(\Lambda)(\widetilde{\Lambda}-\Lambda H_1) \\
=\ & O_P(C_{NT}^{-4}).
\end{aligned}
$$

In addition, using the same technique, it is easy to show

$$\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \widetilde{\lambda}_i e_{it}(\alpha_i' H_2-\widetilde{\alpha}_i')\|_F^2 = O_P(C_{NT}^{-4}) = \frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i \widetilde{\alpha}_i(\lambda_i' H_1-\widetilde{\lambda}_i')e_{it}\|_F^2.$$

Hence $\frac{1}{T}\sum_{t\notin I}\|d_{t,k}\|_F^2 = O_P(C_{NT}^{-4})$ for $k=1,2,3$. Together, we have

$$\frac{1}{T}\sum_{t\notin I}\|\Delta_{t1}\|^2 \leq \frac{1}{T}\sum_{t\notin I}\|c_{t,1}+d_{t,1}\|^2 + \frac{1}{T}\sum_{t\notin I}\|c_{t,2}+d_{t,2}\|^2 = O_P(C_{NT}^{-4})$$

and $\frac{1}{T}\sum_{t\notin I}\|\Delta_{t2}\|^2 = \frac{1}{T}\sum_{t\notin I}\|c_{t,3}+d_{t,3}\|^2 = O_P(C_{NT}^{-2})$.

(iv) Recall that $\widehat{S}_t - S = c_t + d_t$. It suffices to show $\max_{t\leq T}\|c_t\| = o_P(1) = \max_{t\leq T}\|d_t\|$.

$$\begin{aligned}\max_{t\leq T}\|c_t\| &\leq O_P(1)\max_{it}|\widehat{e}_{it}-e_{it}|(|e_{it}|+1) = o_P(1).\\ \max_{t\leq T}\|d_t\| &\leq (\max_{it}|e_{it}| + \max_{it}|e_{it}|^2+1)O_P(C_{NT}^{-1}) = o_P(1).\end{aligned}$$

□

Lemma G.9. *For terms defined in (E.2), and for each fixed $t\notin I$,*

(i) $\sum_{d=2}^5 A_{dt} = O_P(C_{NT}^{-2})$

(ii) For the "upper block" of A_{6t}, $\frac{1}{N}\sum_i \lambda_i e_{it}(l_i' w_t\lambda_i' f_t - \widehat{l_i' w_t}\dot{\lambda}_i'\widetilde{f}_t) = O_P(C_{NT}^{-2})$.

(iii) The upper block of A_{1t} is $O_P(C_{NT}^{-2})$.

Proof. (i) Term A_{2t}. Given $B_t^{-1}-B^{-1} = O_P(C_{NT}^{-1})$ and the cross-sectional weak correlations in u_{it}, it is easy to see $A_{2t} = O_P(C_{NT}^{-2})$.

Term A_{3t}. It suffices to prove:

$$\begin{aligned}&\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widehat{e}_{it} - H_1'\lambda_i e_{it})u_{it} = O_P(C_{NT}^{-2})\\ &\frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)u_{it} = O_P(C_{NT}^{-2}).\end{aligned}\tag{G.12}$$

First, let Υ_t be an $N\times 1$ vector of $e_{it}u_{it}$. Due to the serial independence of (u_{it},e_{it}), we have

$$\begin{aligned}&\mathsf{E}_I\frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it} = 0,\\ &\mathsf{Var}_I(\frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it}) = \frac{1}{N^2}(\widetilde{\Lambda}-\Lambda H_1)'\mathsf{Var}_I(\Upsilon_t)(\widetilde{\Lambda}-\Lambda H_1)\\ \leq\ & O_P(C_{NT}^{-2}N^{-1})\max_i\sum_j|\mathsf{Cov}(e_{it}u_{it}, e_{jt}u_{jt})| = O_P(C_{NT}^{-2}N^{-1}).\end{aligned}$$

Similarly, $\mathsf{E}_I \frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)u_{it} = 0$ and $\mathsf{Var}_I(\frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)u_{it}) = O_P(C_{NT}^{-2}N^{-1})$.

$$\frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it} = O_P(C_{NT}^{-1}N^{-1/2})$$
$$\frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)u_{it} = O_P(C_{NT}^{-1}N^{-1/2})$$
$$\frac{1}{N}\sum_i\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2u_{it}^2 = O_P(1)\frac{1}{N}\sum_i\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2\mathsf{E}_I u_{it}^2 = O_P(C_{NT}^{-2}).$$
(G.13)

Thus, the first term of (G.12) is

$$\begin{aligned}
&\frac{1}{N}\sum_i(\widetilde{\lambda}_i\widehat{e}_{it} - H_1'\lambda_i e_{it})u_{it} \le (\frac{1}{N}\sum_i\|\widetilde{\lambda}_i - H_1'\lambda_i\|^2u_{it}^2)^{1/2}(\frac{1}{N}\sum_i(\widehat{e}_{it} - e_{it})^2)^{1/2} \\
&+H_1'\frac{1}{N}\sum_i\lambda_i(\widehat{e}_{it} - e_{it})u_{it} + \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it} \\
\le\ & O_P(C_{NT}^{-1})(\frac{1}{N}\sum_i(\widehat{e}_{it} - e_{it})^2)^{1/2} + O_P(C_{NT}^{-2}) + O_P(1)\frac{1}{N}\sum_i\lambda_i(\widehat{e}_{it} - e_{it})u_{it} \\
=\ & O_P(C_{NT}^{-2}).
\end{aligned}$$

where the last equality follows from Lemma G.1.

Term A_{4t}. Recall that $\mu_{it} = l_i'w_t$. It suffices to prove each of the following terms is $O_P(C_{NT}^{-2})$:

$$\begin{aligned}
C_{1t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}(\widehat{e}_{it} - e_{it})(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{2t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)(\widehat{e}_{it} - e_{it})^2(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{3t} &= \frac{1}{N}\sum_i H_1'\lambda_i(\widehat{e}_{it} - e_{it})^2(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{4t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}(\widehat{e}_{it} - e_{it})\lambda_i'H_1\widetilde{f}_t \\
C_{5t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)(\widehat{e}_{it} - e_{it})^2\lambda_i'H_1\widetilde{f}_t \\
C_{6t} &= \frac{1}{N}\sum_i H_1'\lambda_i(\widehat{e}_{it} - e_{it})^2\lambda_i'H_1\widetilde{f}_t \\
C_{7t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}\mu_{it}(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t
\end{aligned}$$

$$
\begin{aligned}
C_{8t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)(\widehat{e}_{it}-e_{it})\mu_{it}(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{9t} &= \frac{1}{N}\sum_i H_1'\lambda_i(\widehat{e}_{it}-e_{it})\mu_{it}(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{10t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}\mu_{it}\lambda_i'H_1(\widetilde{f}_t - H_1^{-1}f_t) \\
C_{11t} &= \frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)(\widehat{e}_{it}-e_{it})\mu_{it}\lambda_i'H_1(\widetilde{f}_t - H_1^{-1}f_t) \\
C_{12t} &= \frac{1}{N}\sum_i H_1'\lambda_i(\widehat{e}_{it}-e_{it})\mu_{it}\lambda_i'H_1(\widetilde{f}_t - H_1^{-1}f_t) \\
C_{13t} &= \frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)(\widehat{e}_{it}-e_{it})(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{14t} &= \frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)(\widehat{e}_{it}-e_{it})\lambda_i'H_1\widetilde{f}_t \\
C_{15t} &= \frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)\mu_{it}(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
C_{16t} &= \frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)\mu_{it}\lambda_i'H_1(\widetilde{f}_t - H_1^{-1}f_t). \qquad \text{(G.14)}
\end{aligned}
$$

The proof follows from repeatedly applying the Cauchy-Schwarz inequality and is straightforward. In addition, it is also straightforward to apply Cauchy-Schwarz to prove that

$$
\frac{1}{T}\sum_{t\notin I}\|C_{dt}\|^2 = O_P(C_{NT}^{-4}), \quad d = 1, ..., 16. \qquad \text{(G.15)}
$$

Term A_{5t}. Note that

$$
A_{5t} = (\widehat{B}_t^{-1} - B^{-1})\begin{pmatrix} H_1'\sum_{d=1}^4 B_{dt} \\ H_2'\sum_{d=5}^8 B_{dt}\end{pmatrix}
$$

where B_{dt} are defined below. Given $B_t^{-1} - B^{-1} = O_P(C_{NT}^{-1})$, it suffices to prove the following terms are $O_P(C_{NT}^{-2})$.

$$
\begin{aligned}
B_{1t} &= \frac{1}{N}\sum_i \lambda_i e_{it}(\widehat{e}_{it}-e_{it})(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\
B_{2t} &= \frac{1}{N}\sum_i \lambda_i e_{it}(\widehat{e}_{it}-e_{it})\lambda_i'H_1\widetilde{f}_t \\
B_{3t} &= \frac{1}{N}\sum_i \lambda_i e_{it}l_i'w_t(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t
\end{aligned}
$$

$$\begin{aligned} B_{4t} &= \frac{1}{N}\sum_i \lambda_i e_{it} l_i' w_t \lambda_i' H_1(\widetilde{f}_t - H_1^{-1} f_t) \\ B_{5t} &= \frac{1}{N}\sum_i \alpha_i(\widehat{e}_{it} - e_{it})(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \end{aligned} \tag{G.16}$$

and that the following terms are $O_P(C_{NT}^{-1})$:

$$\begin{aligned} B_{6t} &= \frac{1}{N}\sum_i \alpha_i(\widehat{e}_{it} - e_{it})\lambda_i' H_1 \widetilde{f}_t \\ B_{7t} &= \frac{1}{N}\sum_i \alpha_i l_i' w_t(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t \\ B_{8t} &= \frac{1}{N}\sum_i \alpha_i l_i' w_t \lambda_i' H_1(\widetilde{f}_t - H_1^{-1} f_t). \end{aligned} \tag{G.17}$$

In fact, $B_{1t}, B_{5,t} \sim B_{8t}$ follow immediately from the Cauchy-Schwarz inequality. $B_{2t} = O_P(C_{NT}^{-2})$ due to Lemma G.1. By Lemma G.5 that $\frac{1}{\sqrt{N}}\sum_i (H_1'\lambda_i - \dot{\lambda}_i)e_{it} = O_P(\sqrt{N}C_{NT}^{-2})$. So $t \notin I$, $B_{3t} = O_P(1)\frac{1}{N}\sum_i \lambda_i e_{it}\mu_i(\dot{\lambda}_i - H_1'\lambda_i)' = O_P(C_{NT}^{-2})$. In addition, for each fixed t, $\widetilde{f}_t - H_1^{-1}f_t = O_P(C_{NT}^{-1})$. Thus

$$B_{4t} = O_P(C_{NT}^{-1})\frac{1}{N}\sum_i \lambda_i e_{it}\mu_i\lambda_i' = O_P(C_{NT}^{-2}).$$

So $A_{5t} = O_P(C_{NT}^{-2})$.

(ii) "upper block" of A_{6t}. Note that $l_i'w_t - \widehat{l_i'w_t} = \widehat{e}_{it} - e_{it}$.

$$A_{6t} = B^{-1}\begin{pmatrix} H_1'\sum_{d=1}^4 B_{dt} \\ H_2'\sum_{d=5}^8 B_{dt} \end{pmatrix}$$

So we only need to look at $\sum_{d=1}^4 B_{dt}$. From the proof of (i), we have $B_{dt} = O_P(C_{NT}^{-2})$, for $d = 1, ..., 4$. It follows immediately that $\frac{1}{N}\sum_i \lambda_i e_{it}(l_i'w_t\lambda_i' f_t - \widehat{l_i'w_t}\dot{\lambda}_i'\widetilde{f}_t) = O_P(C_{NT}^{-2})$.

(iii) Lastly, note that the upper block of A_{1t} is determined by the upper blocks of $\widehat{B}_t^{-1}\widehat{S}_t - B^{-1}S$, and are both $O_P(C_{NT}^{-2})$ by Lemma G.7.

□

G.4. **Technical lemmas for $\widehat{\lambda}_i$.**

Lemma G.10. *Suppose* $\max_i \|\frac{1}{T}\sum_s f_s f_s'(e_{is}^2 - \mathsf{E}e_{is}^2)\| = o_P(1) = \max_i \|\frac{1}{T}\sum_s f_s g_s' e_{is}\|$.
(i) $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}} \|\widehat{D}_i - D_i\|^2 = O_P(C_{NT}^{-2})$.
(ii) $\max_{i\le N} \|\widehat{D}_i^{-1} - D_i^{-1}\| = o_P(1)$.

(iii) $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i^{-1}-D_i^{-1}\|^2=O_P(C_{NT}^{-2})$.

Proof. $\widehat{D}_i - D_i$ is a two-by-two block matrix. The first block is

$$
\begin{aligned}
d_{1i} &:= \frac{1}{T_0}\sum_{s\notin I}\widehat{f}_s\widehat{f}_s'\widehat{e}_{is}^2 - H_f f_s f_s' H_f' \mathsf{E}e_{is}^2 \\
&= \frac{1}{T_0}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)(\widehat{f}_s - H_f f_s)'[(\widehat{e}_{is}-e_{is})^2+(\widehat{e}_{is}-e_{is})e_{is}+e_{is}^2] \\
&\quad +\frac{1}{T_0}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)f_s'H_f'[(\widehat{e}_{is}-e_{is})^2+(\widehat{e}_{is}-e_{is})e_{is}+e_{is}^2] \\
&\quad +\frac{1}{T_0}\sum_{s\notin I}H_f f_s(\widehat{f}_s - H_f f_s)'[(\widehat{e}_{is}-e_{is})^2+(\widehat{e}_{is}-e_{is})e_{is}+e_{is}^2] \\
&\quad +\frac{1}{T_0}\sum_{s\notin I}H_f f_s f_s' H_f'[(\widehat{e}_{is}-e_{is})^2+(\widehat{e}_{is}-e_{is})e_{is}]+\frac{1}{T_0}\sum_{s\notin I}H_f f_s f_s' H_f'(e_{is}^2-\mathsf{E}e_{is}^2).
\end{aligned}
$$

It follows from Lemma G.2 that $\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^2e_{it}^2=O_P(C_{NT}^{-2})$ $\frac{1}{NT}\sum_{it}(\widehat{e}_{it}-e_{it})^4=O_P(C_{NT}^{-4})$ and Lemma G.1 $\max_{it}|\widehat{e}_{it}-e_{it}|e_{it}=O_P(1)$.

$$
\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_{1i}\|^2 &\le O_P(1)\frac{1}{T_0}\sum_{s\notin I}\|\widehat{f}_s - H_f f_s\|^2(1+\|f_t\|^2+\|f_t\|^2\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}e_{it}^4) \\
&\quad +O_P(C_{NT}^{-2}) = O_P(C_{NT}^{-2}).
\end{aligned}
$$

The second block is

$$
\begin{aligned}
d_{2i} &:= \frac{1}{T_0}\sum_{s\notin I}\widehat{f}_s\widehat{g}_s'\widehat{e}_{is} = \frac{1}{T_0}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)(\widehat{g}_s - H_g g_s)'\widehat{e}_{is} \\
&\quad +\frac{1}{T_0}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)g_s'H_g'\widehat{e}_{is}+\frac{1}{T_0}\sum_{s\notin I}H_f f_s(\widehat{g}_s - H_g g_s)'\widehat{e}_{is} \\
&\quad +\frac{1}{T_0}\sum_{s\notin I}H_f f_s g_s'H_g'(\widehat{e}_{is}-e_{is})+\frac{1}{T_0}\sum_{s\notin I}H_f f_s g_s' H_g' e_{is}
\end{aligned}
$$

So $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_{2i}\|^2 = O_P(C_{NT}^{-2})$. The third block is similar. The fourth block is $d_{4i}=\frac{1}{T_0}\sum_{s\notin I}\widehat{g}_s\widehat{g}_s' - H_g g_s g_s' H_g'$. So $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_{4i}\|^2=O_P(C_{NT}^{-2})$.

Hence $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i-D_i\|^2=O_P(C_{NT}^{-2})$.

(ii) Note that

$$
\max_i\|d_{1i}\| = o_P(1)+O_P(1)\max_i\|\frac{1}{T}\sum_s f_s f_s'(e_{is}^2-\mathsf{E}e_{is}^2)\| = o_P(1)
$$

$$\max_i \|d_{2i}\| = o_P(1) + O_P(1)\max_i \|\frac{1}{T}\sum_s f_s g_s' e_{is}\| = o_P(1).$$

So $\max_i \|\widehat{D}_i - D_i\| = o_P(1)$. Now because $\min_i \lambda_{\min}(D_i) \geq c$, with probability approaching one $\min_i \lambda_{\min}(\widehat{D}_i) \geq \min_i \lambda_{\min}(D_i) - \max_i \|\widehat{D}_i - D_i\| > c/2$. So $\max_i \|D_i^{-1}\| + \max_i \|\widehat{D}_i^{-1}\| = O_P(1)$. So

$$\max_i \|\widehat{D}_i^{-1} - D_i^{-1}\| \leq \max_i \|\widehat{D}_i^{-1}\| \max_i \|D_i^{-1}\| \max_i \|\widehat{D}_i - D_i\| = o_P(1).$$

(iii) Because $\min_i \lambda_{\min}(D_i) \geq c$,

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i^{-1} - D_i^{-1}\|^2 &\leq \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i^{-1}\|^2\|D_i^{-1}\|^2\|\widehat{D}_i - D_i\|^2 \\ &\leq \max_i[\|\widehat{D}_i^{-1} - D_i^{-1}\| + \|D_i^{-1}\|]^2 \frac{C}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i - D_i\|^2 \\ &= O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i - D_i\|^2 = O_P(C_{NT}^{-2}).\end{aligned}$$

□

Lemma G.11. $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)e_{is}f_s'\|^2 = O_P(C_{NT}^{-4})$.

This lemma is needed to prove the performance for $\widehat{\lambda}_i$, which controls the effect of $\widehat{f}_s - H_f f_s$ on the estimation of λ_i.

Proof. Use (E.2), $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)e_{is}f_s'\|^2$ is bounded by, up to a multiplier of order $O_P(1)$,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\frac{1}{N}\sum_{j=1}^N \lambda_j e_{jt}u_{jt}e_{it}f_t'\|^2 + \sum_{d=1}^6 \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{dt}e_{it}f_t'\|^2.$$

where a_{dt} is the upper block of A_{dt}. We shall assume $\dim(f_t) = 1$ without loss of generality. The proof for $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{1t}e_{it}f_t'\|^2$ is the hardest, and we shall present it at last.

Step 1. Show the first term.

Let $\nu_t = (\lambda_1 e_{1t}, ..., \lambda_N e_{Nt})'$. Then

$$\begin{aligned}&\mathsf{E}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\frac{1}{N}\sum_{j=1}^N \lambda_j e_{jt}u_{jt}e_{it}f_t'\|^2 = \frac{1}{N^2T^2}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\sum_{t\notin I}\mathsf{E}(\nu_t' u_t e_{it}f_t)^2 \\ \leq\ & \frac{1}{N^2T^2}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\sum_{t\notin I}\mathsf{E}f_t^2\nu_t'\mathsf{E}(u_tu_t'|E,F)\nu_t e_{it}^2 \leq \frac{1}{N^2T^2}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\sum_{t\notin I}\sum_{j=1}^N \mathsf{E}f_t^2\lambda_j^2 e_{jt}^2 e_{it}^2\end{aligned}$$

$$= \quad O(N^{-1}T^{-1}) = O(C_{NT}^{-4}).$$

Step 2. Show $\sum_{d=2}^{6}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{dt}e_{it}f_t'\|^2$.

Up to a multiplier of order $O_P(1)$, by Lemma G.8,

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{2t}e_{it}f_t'\|^2\\
\leq\;& \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}|\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|\|\widehat{B}_t^{-1}-B^{-1}\|[\|\frac{1}{N}\sum_j\lambda_je_{jt}u_{jt}\|+\|\frac{1}{N}\sum_j\alpha_ju_{jt}\|]|^2\\
\leq\;& \frac{1}{T}\sum_{t\notin I}\|\widehat{B}_t-B\|^2\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|^2[\|\frac{1}{N}\sum_j\lambda_je_{jt}u_{jt}\|+\|\frac{1}{N}\sum_j\alpha_ju_{jt}\|]^2\\
=\;& O_P(C_{NT}^{-4})\\
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{3t}e_{it}f_t'\|^2\\
\leq\;& \max_{t\leq T}\|\widehat{B}_t^{-1}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|\|\frac{1}{N}\sum_j(\widetilde{\lambda}_j\widehat{e}_{jt}-H_1'\lambda_je_{jt})u_{jt}\|)^2\\
&+\max_{t\leq T}\|\widehat{B}_t^{-1}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|\|\frac{1}{N}\sum_j(\widetilde{\alpha}_j-H_2'\alpha_j)u_{jt}\|)^2\\
\leq\;& O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j(\widetilde{\lambda}_j-H_1'\lambda_j)e_{jt}u_{jt}\|^2+O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j(\widetilde{\alpha}_j-H_2'\alpha_j)u_{jt}\|^2\\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j(\widetilde{\lambda}_j-H_1'\lambda_j)(\widehat{e}_{jt}-e_{jt})u_{jt}\|^2\\
&+O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j\lambda_j(\widehat{e}_{jt}-e_{jt})u_{jt}\|^2=O_P(C_{NT}^{-4}).
\end{aligned}$$

The first term is bounded by, for $\omega_{it}=e_{it}u_{it}$,

$$O_P(1)\frac{1}{T}\sum_{t\notin I}\frac{1}{N^2}(\widetilde{\Lambda}-\Lambda H_1)'\mathsf{Var}_I(\omega_t)(\widetilde{\Lambda}-\Lambda H_1)\leq O_P(1)\frac{1}{N^2}\|\widetilde{\Lambda}-\Lambda H_1\|^2=O_P(C_{NT}^{-4}).$$

The second term $\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j(\widetilde{\alpha}_j-H_2'\alpha_j)u_{jt}\|^2$ is bounded similarly. The third term follows from Cauchy-Schwarz. The last term follows from Lemma G.2. Next,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{4t}e_{it}f_t'\|^2\leq\max_{t\leq T}\|\widehat{B}_t^{-1}\|^2\sum_{d=1}^{16}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|\|C_{dt}\|)^2=O_P(C_{NT}^{-4}),$$

where C_{dt}'s are defined in the proof of Lemma G.9. Applying the simple Cauchy-Schwarz proves $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\frac{1}{T}\sum_{t\notin I}\|e_{it}f_t'\|\|C_{dt}\|)^2=O_P(C_{NT}^{-4})$ for all $d\leq 16$.

Next, for B_{dt} defined in the proof of Lemma G.9, repeatedly use Cauchy-Schwarz,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{5t}e_{it}f_t'\|^2 \leq O_P(C_{NT}^{-2})\sum_{d=1}^{8}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{dt}e_{it}f_t'\|^2 = O_P(C_{NT}^{-4}).$$

Also,

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{6t}e_{it}f_t'\|^2 &\leq \sum_{d=1}^{4}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{dt}e_{it}f_t'\|^2 = O_P(C_{NT}^{-4}), \quad \text{where}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{1t}e_{it}f_t'\|^2 &= O_P(C_{NT}^{-4}) \text{ Cauchy-Schwarz}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{2t}e_{it}f_t'\|^2 &\leq O_P(1)\frac{1}{T}\sum_{t\notin I}|\frac{1}{N}\sum_j\lambda_j^2e_{jt}(\widehat{e}_{jt}-e_{jt})|^2 = O_P(C_{NT}^{-4})\\
&\quad \text{(Lemma G.2)}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{3t}e_{it}f_t'\|^2 &\leq O_P(1)\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j l_j\lambda_je_{jt}(\dot{\lambda}_j-H_1'\lambda_j)\|^2 = O_P(C_{NT}^{-4})\\
&\quad \text{(Lemma G.5)}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T}\sum_{t\notin I}\|B_{4t}e_{it}f_t'\|^2 &\leq O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_j\lambda_j^2l_je_{jt}\|^2 = O_P(C_{NT}^{-4}).
\end{aligned}$$

Step 3. Finally, we bound $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{1t}e_{it}f_t'\|^2$.

Note that

$$A_{1t} = (\widehat{B}_t^{-1}\widehat{S}_t - B^{-1}S)\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix} := (A_{1t,a}+A_{1t,b}+A_{1t,c})\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix},$$

where

$$\begin{aligned}
A_{1t,a} &= (\widehat{B}_t^{-1}-B^{-1})(\widehat{S}_t-S)\\
A_{1t,b} &= (\widehat{B}_t^{-1}-B^{-1})S\\
A_{1t,c} &= B^{-1}(\widehat{S}_t-S).
\end{aligned}$$

Let $(a_{1t,a}, a_{1t,b}, a_{1t,c})$ respectively be the upper blocks of $(A_{1t,a}, A_{1t,b}, A_{1t,c})$. Then

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{1t}e_{it}f_t'\|^2 \leq \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}(\frac{1}{T}\sum_{t\notin I}(\|a_{1t,a}\|+\|a_{1t,b}\|+\|a_{1t,c}\|)\|e_{it}f_t\|(\|f_t\|+\|g_t\|))^2.$$

By the Cauchy-Schwarz and Lemma G.8, and B is a block diagonal matrix,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\frac{1}{T}\sum_{t\notin I}\|a_{1t,b}\|\|e_{it}f_t\|(\|f_t\|+\|g_t\|)]^2 \leq O_P(1)\frac{1}{T}\sum_{t\notin I}\|A_{1t,b}\|^2$$

$$\leq O_P(\|S\|^2)\frac{1}{T}\sum_{t\notin I}\|\widehat{B}_t - B\|^2 = O_P(C_{NT}^{-4}).$$

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\frac{1}{T}\sum_{t\notin I}\|a_{1t,c}\|\|e_{it}f_t\|(\|f_t\|+\|g_t\|)]^2 \leq O_P(1)\frac{1}{T}\sum_{t\notin I}\|a_{1t,c}\|^2$$

$$\leq O_P(1)\frac{1}{T}\sum_{t\notin I}\|\Delta_{t1}\|^2 = O_P(C_{NT}^{-4}),$$

where Δ_{t1} is defined in Lemma G.8, the upper block of $\widehat{S}_t - S$.

The treatment of $\frac{1}{T}\sum_{t\notin I}\|a_{1t,a}\|\|e_{it}f_t\|(\|f_t\|+\|g_t\|)$ is slightly different. Note that $\max_{t\leq T}\|\widehat{B}_t^{-1}\| + \|B^{-1}\| = O_P(1)$, shown in Lemma G.8. Partition

$$\widehat{S}_t - S = \begin{pmatrix} \Delta_{t1} \\ (\Delta_{t2,1}+\Delta_{t2,2}, 0) \end{pmatrix}$$

where the notation Δ_{t1} is defined in the proof of Lemma G.8. The proof of Lemma G.8 also gives

$$\begin{aligned}
\Delta_{t2,1} &= \frac{1}{N}\sum_i \widetilde{\alpha}_i(\lambda_i'H_1 - \widetilde{\lambda}_i')e_{it} + \frac{1}{N}\sum_i \widetilde{\alpha}_i(\lambda_i'H_1 - \widetilde{\lambda}_i')(\widehat{e}_{it} - e_{it}) \\
&\quad + \frac{1}{N}\sum_i (\widetilde{\alpha}_i - H_2'\alpha_i)\lambda_i'H_1(e_{it} - \widehat{e}_{it}) \\
\Delta_{t2,2} &= H_2'\frac{1}{N}\sum_i \alpha_i\lambda_i'(e_{it} - \widehat{e}_{it})H_1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|a_{1t,a}\| &\leq \|(\widehat{B}_t^{-1} - B^{-1})(\widehat{S}_t - S)\| \\
&\leq (\max_t \|\widehat{B}_t^{-1}\| + \|B\|^{-1})(\|\Delta_{t1}\| + \|\Delta_{t2,1}\|) + \|\widehat{B}_t^{-1} - B^{-1}\|\|\Delta_{t2,2}\|. \qquad \text{(G.18)}
\end{aligned}$$

Note that the above bound treats Δ_{t1} and Δ_{t2} differently because by the proof of Lemma G.8, $\frac{1}{T}\sum_{t\notin I}\|\Delta_{t1}\|^2 = O_P(C_{NT}^{-4}) = \frac{1}{T}\sum_{t\notin I}\|\Delta_{t2,1}\|^2$ but the rate of convergence for $\frac{1}{T}\sum_{t\notin I}\|\Delta_{t2,2}\|^2$ is slower $(= O_P(C_{NT}^{-2}))$.

Hence

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}[\frac{1}{T}\sum_{t\notin I}\|a_{1t,a}\|\|e_{it}f_t\|(\|f_t\|+\|g_t\|)]^2 \leq O_P(1)\frac{1}{T}\sum_{t\notin I}\|\Delta_{t1}\|^2 + \|\Delta_{t2,1}\|^2$$

$$+\frac{1}{T}\sum_{t\notin I}\|\widehat{B}_t^{-1} - B^{-1}\|^2\frac{1}{T}\sum_{t\notin I}\|\Delta_{t2,2}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|e_{it}f_t\|^2(\|f_t\|+\|g_t\|)^2$$

$$\leq^{(a)} \quad O_P(C_{NT}^{-4}) + O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}\|\Delta_{t2,2}\|^2\|\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|e_{it}f_t\|^2(\|f_t\|+\|g_t\|)^2$$
$$\leq^{(b)} \quad O_P(C_{NT}^{-4}).$$

where (a) follows from the proof of Lemma G.8, while (b) follows from Lemma G.1. Thus

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}a_{1t}e_{it}f_t'\|^2 = O_P(C_{NT}^{-4}).$$

Also note that in the above proof, we have also proved (to be used later)

$$\frac{1}{T}\sum_{t\notin I}[\|a_{1t,b}\|^2+\|a_{1t,c}\|^2+\|\Delta_{t1}\|^2+\|\Delta_{t2,1}\|^2] \leq O_P(C_{NT}^{-4}) \tag{G.19}$$

□

Lemma G.12. *for* $\mu_{it} = l_i'w_t$, $\widehat{\mu}_{it} = \widehat{l_i'w_t}$,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}\widehat{f}_s\widehat{e}_{is}(\mu_{it}\lambda_i'f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\|^2 = O_P(C_{NT}^{-4})$$
$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T}\sum_{s\notin I}\widehat{g}_s(\mu_{it}\lambda_i'f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\|^2 = O_P(C_{NT}^{-2}).$$

Proof. It suffices to prove that the following statements:

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,d}\|^2 = O_P(C_{NT}^{-4}), d = 1\sim 6, \quad \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,7}\|^2 = O_P(C_{NT}^{-2}),$$

where

$$R_{3i,1} := \frac{1}{T}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)(\widehat{e}_{is}-e_{is})\mu_{it}\lambda_i'f_s + \frac{1}{T}\sum_{s\notin I}(\widehat{f}_s\widehat{e}_{is} - H_f f_s e_{is})(\mu_{it}\lambda_i'f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)$$
$$R_{3i,2} := \frac{1}{T}\sum_{s\notin I}f_s e_{is}(\widehat{e}_{is}-e_{is})(\dot{\lambda}_i'\widetilde{f}_s - \lambda_i'f_s) + \frac{1}{T}\sum_{s\notin I}f_s e_{is}\mu_{is}(\dot{\lambda}_i - H_1'\lambda_i)'(\widetilde{f}_s - H_1^{-1}f_s)$$
$$R_{3i,3} := \frac{1}{T}\sum_{s\notin I}f_s e_{is}\mu_{is}f_s'H_1^{-1'}(\dot{\lambda}_i - H_1'\lambda_i)$$
$$R_{3i,4} := \frac{1}{T}\sum_{s\notin I}f_s e_{is}(\widehat{e}_{is}-e_{is})\lambda_i'f_s$$
$$R_{3i,5} := \frac{1}{T}\sum_{s\notin I}f_s e_{is}\mu_{is}\lambda_i'(\widetilde{f}_s - H_1^{-1}f_s)$$
$$R_{3i,6} := \frac{1}{T}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)e_{is}\mu_{is}\lambda_i'f_s$$

$$R_{3i,7} := \frac{1}{T}\sum_{s\notin I}\widehat{g}_s(\mu_{is}\lambda_i' f_s - \widehat{\mu}_{is}\dot{\lambda}_i'\widetilde{f}_s). \tag{G.20}$$

So

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,d}\|^2 &= O_P(C_{NT}^{-4}), \quad d=1,2,3, \text{by Cauchy-Schwarz}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,4}\|^2 &= O_P(C_{NT}^{-4}), \text{ Lemma G.2}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,5}\|^2 &= O_P(C_{NT}^{-4}), \text{ Lemma G.6}\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,6}\|^2 &= O_P(C_{NT}^{-4}), \text{ Lemma G.11 },\\
\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|R_{3i,7}\|^2 &= O_P(C_{NT}^{-2}), \text{ by Cauchy-Schwarz.}
\end{aligned}$$

□

Lemma G.13. *Let r_{5i} be the upper block of R_{5i}, where*

$$R_{5i} = \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix}\widehat{e}_{is}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix}\begin{pmatrix}\widehat{f}_s - H_f f_s\\ \widehat{g}_s - H_g g_s\end{pmatrix}u_{is}.$$

Then $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{5i}\| = O_P(C_{NT}^{-2})$.

Proof. For any matrix A of the same size R_{5i}, we write $\mathcal{U}(A)$ to denote its upper block (the first $\dim(f_t)$-submatrix of A). Then in this notation $r_{5i} = \mathcal{U}(R_{5i})$.

First, by Cauchy-Schwarz

$$\begin{aligned}
&\left[\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\widehat{D}_i^{-1} - D_i^{-1}\|\left(\|\frac{1}{T}\sum_{s\notin I}\widehat{e}_{is}(\widehat{f}_s - H_f f_s)u_{is}\| + \|\frac{1}{T}\sum_{s\notin I}(\widehat{g}_s - H_g g_s)u_{is}\|\right)\right]^2\\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\left(\frac{1}{T}\sum_{s\notin I}(\widehat{f}_s - H_f f_s)^2\frac{1}{T}\sum_{s\notin I}u_{is}^2\widehat{e}_{is}^2 + \frac{1}{T}\sum_{s\notin I}(\widehat{g}_s - H_g g_s)^2\frac{1}{T}\sum_{s\notin I}u_{is}^2\right)\\
\leq\ & O_P(C_{NT}^{-4}).
\end{aligned}$$

Now let $\bar{R}_{5i}$ be defined as R_{5i} but with $\widehat{D}_i^{-1}$ replaced with D_i^{-1}. Substitute in the expansion (E.2), then

$$\bar{R}_{5i} = D_i^{-1}\frac{1}{T_0}\sum_{t\notin I}\begin{pmatrix}\widehat{e}_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix}B^{-1}\frac{1}{N}\sum_j\begin{pmatrix}H_1'\lambda_j e_{jt}\\ H_2'\alpha_j\end{pmatrix}u_{jt}u_{it} + \sum_{d=1}^{6}D_i^{-1}\frac{1}{T_0}\sum_{t\notin I}\begin{pmatrix}\widehat{e}_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix}A_{dt}u_{it}.$$

Both B^{-1} and D_i^{-1} are block diagonal; let b, d_i respectively denote their first diagonal block. Then

$$\mathcal{U}(\bar{R}_{5i}) = \frac{1}{T_0}\sum_{t\neq I} d_i\widehat{e}_{it}bH_1'\frac{1}{N}\sum_j \lambda_j e_{jt}u_{jt}u_{it} + \sum_{d=1}^{6} d_i\frac{1}{T_0}\sum_{t\neq I}\widehat{e}_{it}u_{it}\mathcal{U}(A_{dt}).$$

The goal is now to prove $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\mathcal{U}(\bar{R}_{5i})\| = O_P(C_{NT}^{-2})$.

First, let $\nu_t = (\lambda_j u_{jt} : j \leq N)$, a column vector.

$$\begin{aligned}
&\left[\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\neq I} d_i(\widehat{e}_{it}-e_{it})bH_1'\frac{1}{N}\sum_j \lambda_j e_{jt}u_{jt}u_{it}\|\right]^2 \\
\leq\ & O_P(C_{NT}^{-2})\frac{1}{T_0}\sum_{t\neq I}\frac{1}{N^2}\mathsf{E}\mathsf{Var}(\nu_t' e_t u_{it}|U) \leq O_P(C_{NT}^{-2})\frac{1}{T_0}\sum_{t\neq I}\frac{1}{N^2}\mathsf{E}u_{it}^2\nu_t'\mathsf{Var}(e_t|U)\nu_t \\
\leq\ & O_P(C_{NT}^{-2}N^{-1})\frac{1}{T_0}\sum_{t\neq I}\mathsf{E}u_{it}^2\frac{1}{N}\sum_{j=1}^{N}\lambda_j^2 u_{jt}^2 = O_P(C_{NT}^{-4}).
\end{aligned}$$

Next, for $\omega_{it} = e_{it}u_{it}$,

$$\begin{aligned}
&\left[\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\neq I} d_i e_{it}bH_1'\frac{1}{N}\sum_j \lambda_j e_{jt}u_{jt}u_{it}\|\right]^2 \\
\leq\ & O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\mathsf{E}|\frac{1}{NT}\sum_t\sum_j e_{it}\lambda_j e_{jt}u_{jt}u_{it}|^2 \\
\leq\ & O_P(1)\frac{1}{N^2}(\max_i\sum_j|\mathsf{E}\omega_{jt}\omega_{it}|)^2 + O_P(1)\frac{1}{N^2T^2}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\sum_t\mathsf{Var}(\sum_j\lambda_j\omega_{jt}\omega_{it}) \\
\leq\ & O_P(N^{-2}) + O_P(1)\frac{1}{N^2T^2}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\sum_t\sum_{jk}|\mathsf{Cov}(\omega_{jt}\omega_{it},\omega_{kt}\omega_{it})| \\
=\ & O_P(N^{-2}+N^{-1}T^{-1}) = O_P(C_{NT}^{-4}).
\end{aligned}$$

We now show $\sum_{d=1}^{6}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\widehat{e}_{it}u_{it}\mathcal{U}(A_{dt})\| = O_P(C_{NT}^{-2})$. For each $d \leq 6$,

$$\begin{aligned}
&[\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}(\widehat{e}_{it}-e_{it})u_{it}\mathcal{U}(A_{dt})\|]^2 \\
\leq\ & \frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\frac{1}{T_0}\sum_{t\neq I}(\widehat{e}_{it}-e_{it})^2u_{it}^2d_i^2\frac{1}{T_0}\sum_{t\neq I}\|A_{dt}\|^2 = O_P(C_{NT}^{-4}),
\end{aligned}$$

where it follows from applying Cauchz-Schwarz that $\frac{1}{T_0}\sum_{t\neq I}\|A_{dt}\|^2 = O_P(C_{NT}^{-2})$. It remains to prove $\sum_{d=1}^{6}\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}\mathcal{U}(A_{dt})\| = O_P(C_{NT}^{-2})$.

term $\mathcal{U}(A_{1t})$. Using the notation of the proof of Lemma G.11, we have

$$\mathcal{U}(A_{1t}) := (a_{1t,a} + a_{1t,b} + a_{1t,c}) \begin{pmatrix} H_1^{-1} f_t \\ H_2^{-1} g_t \end{pmatrix},$$

Also, (G.18) and (G.19) yield, for $m_{it} := \|d_i \omega_{it}\|(\|f_t\| + \|g_t\|)$,

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \|d_i \frac{1}{T_0} \sum_{t \neq I} \omega_{it} \mathcal{U}(A_{dt})\| \\
\leq\ & [\frac{1}{T_0} \sum_{t \neq I} \|a_{1t,b} + a_{1t,c}\|^2]^{1/2} [\frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \frac{1}{T_0} \sum_{t \neq I} m_{it}^2]^{1/2} + \frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \frac{1}{T_0} \sum_{t \neq I} \|m_{it} a_{1t,a}\| \\
\leq\ & O_P(C_{NT}^{-2}) + [\frac{1}{T_0} \sum_{t \neq I} \|\Delta_{t1} + \Delta_{t2,1}\|^2]^{1/2} [\frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \frac{1}{T_0} \sum_{t \neq I} m_{it}^2]^{1/2} \\
& + \frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \frac{1}{T_0} \sum_{t \neq I} m_{it} \|\widehat{B}_t^{-1} - B^{-1}\| \|\Delta_{t2,2}\| \\
\leq\ & O_P(C_{NT}^{-2}) + O_P(C_{NT}^{-1}) [\frac{1}{T} \sum_t \|\Delta_{t2,2}\|^4]^{1/4} \\
\leq\ & O_P(C_{NT}^{-2}) + O_P(C_{NT}^{-1}) [\frac{1}{T} \sum_t \frac{1}{N} \sum_i (e_{it} - \widehat{e}_{it})^4]^{1/4} = O_P(C_{NT}^{-2}).
\end{aligned}$$

terms $\mathcal{U}(A_{2t}), \mathcal{U}(A_{4t}), \mathcal{U}(A_{5t})$. $\frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \|d_i \frac{1}{T_0} \sum_{t \neq I} \omega_{it} \mathcal{U}(A_{dt})\| = O_P(C_{NT}^{-2})$, $d = 2, 4, 5$ follow from the simple Cauchy-Schwarz.

term $\mathcal{U}(A_{3t})$. Because $\max_{t \leq T} \|\widehat{B}_t^{-1}\| = O_P(1)$,

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \|d_i \frac{1}{T_0} \sum_{t \neq I} \omega_{it} \mathcal{U}(A_{3t})\| \leq \frac{1}{|\mathcal{G}|_0} \sum_{i \in \mathcal{G}} \|d_i \frac{1}{T_0} \sum_{t \neq I} \omega_{it} A_{3t}\| \\
\leq\ & [\frac{1}{T_0} \sum_{t \neq I} \|\frac{1}{N} \sum_j \lambda_i u_{jt} (\widehat{e}_{jt} - e_{jt})\|^2]^{1/2} + [\frac{1}{T_0} \sum_{t \neq I} \|\frac{1}{N} \sum_j u_{jt} (\widetilde{\lambda}_j - H_1' \lambda_j)(\widehat{e}_{jt} - e_{jt})\|^2]^{1/2} \\
& + [\frac{1}{T_0} \sum_{t \neq I} \|\frac{1}{N} \sum_j \omega_{jt} (\widetilde{\lambda}_j - H_1' \lambda_j)\|^2]^{1/2} + [\frac{1}{T_0} \sum_{t \neq I} \|\frac{1}{N} \sum_j u_{jt} (\widetilde{\alpha}_j - H_2' \alpha_j)\|^2]^{1/2} \\
=\ & O_P(C_{NT}^{-2}).
\end{aligned}$$

The first term follows from Lemma G.2. The second term follows from Cauchy-Schwarz. The third and fourth are due to, for instance,

$$\begin{aligned}
& \frac{1}{T_0} \sum_{t \neq I} \mathsf{E}_I \|\frac{1}{N} \sum_j \omega_{jt} (\widetilde{\lambda}_j - H_1' \lambda_j)\|^2 \\
\leq\ & \frac{1}{T_0} \sum_{t \neq I} \frac{1}{N^2} \mathsf{Var}_I(\omega_t' (\widetilde{\Lambda} - \Lambda H_1)) \leq \frac{1}{T_0} \sum_{t \neq I} \frac{1}{N^2} \|\widetilde{\Lambda} - \Lambda H_1\|^2 \|\mathsf{Var}_I(\omega_t)\|
\end{aligned}$$

$$\leq \quad O(C_{NT}^{-4}).$$

term $\mathcal{U}(A_{6t})$. Note that B^{-1} is block diagonal. Let b be the first block of B^{-1}. Recall that Lemma G.9 shows

$$\mathcal{U}(A_{6t}) = bH_1'\sum_{d=1}^4 B_{dt}.$$

So $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}\mathcal{U}(A_{6t})\| \leq O_P(1)\sum_{d=1}^4\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}B_{dt}\|$.

For $d = 1, 4$, we apply Cauchy-Schwarz,

$$\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}B_{dt}\| \leq O_P(1)[\frac{1}{T}\sum_t\|B_{dt}\|^2]^{1/2} = O_P(C_{NT}^{-2}).$$

For $d = 2$, we still apply Cauchy-Schwarz and Lemma G.2 that

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}B_{2t}\| \leq O_P(C_{NT}^{-2}) + O_P(1)[\frac{1}{T}\sum_t(\frac{1}{N}\sum_i\lambda_i^2 e_{it}(\widehat{e}_{it}-e_{it}))^2]^{1/2}\\
=\;& O_P(C_{NT}^{-2}).
\end{aligned}$$

Finally, we bound for $d = 3$.

$$\begin{aligned}
&[\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|d_i\frac{1}{T_0}\sum_{t\neq I}\omega_{it}B_{3t}\|]^2 \leq O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\neq I}\omega_{it}B_{3t}\|^2\\
\leq\;& O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\neq I}\omega_{jt}\frac{1}{N}\sum_i l_i\lambda_i e_{it}w_t(\dot{\lambda}_i - H_1'\lambda_i)'\widetilde{f}_t\|^2\\
\leq\;& O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{N}\sum_i\|\frac{1}{T_0}\sum_{t\neq I}\omega_{jt}e_{it}w_t(\widetilde{f}_t - H_1^{-1}f_t)\|^2\\
&+O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{N}\sum_i\|\frac{1}{T_0}\sum_{t\neq I}\omega_{jt}e_{it}w_t f_t\|^2\\
=\;& O_P(C_{NT}^{-4}) + O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{N}\sum_i\|\frac{1}{T_0}\sum_{t\neq I}e_{jt}u_{jt}e_{it}w_t f_t\|^2\\
=\;& O_P(C_{NT}^{-4}).
\end{aligned}$$

The last equality is due to the conditional serial independence of (e_t, u_t) and that $\mathsf{E}(u_t|E, F, W) = 0$.

$\square$

Lemma G.14. *Let r_{6i} be the upper block of R_{6i}, where*

$$R_{6i} = \widehat{D}_i^{-1}\frac{1}{T_0}\sum_{s\notin I}\begin{pmatrix}\widehat{f}_s\widehat{e}_{is}\\ \widehat{g}_s\end{pmatrix}(\lambda_i'H_f^{-1}, \alpha_i'H_g^{-1})\begin{pmatrix}\widehat{e}_{is}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix}\begin{pmatrix}\widehat{f}_s - H_f f_s\\ \widehat{g}_s - H_g g_s\end{pmatrix}.$$

Then $\frac{1}{|\mathcal{G}|_0}\sum_{i\in\mathcal{G}}\|r_{6i}\| = O_P(C_{NT}^{-2})$.

Proof. Write $B^{-1} = \mathsf{diag}(b_1, b_2)$,

$$\begin{aligned}
\Gamma^j &:= \Gamma^j_0 + \Gamma^j_1 + ... + \Gamma^j_6,\\
\Gamma^j_0 &:= \frac{1}{T_0}\sum_{t\notin I}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j' H_f^{-1}\widehat{e}_{jt}b_1, \alpha_j' H_g^{-1}b_2)\frac{1}{N}\sum_i\begin{pmatrix}H_1'\lambda_i e_{it}\\ H_2'\alpha_i\end{pmatrix}u_{it}\\
\Gamma^j_d &:= \frac{1}{T_0}\sum_{t\notin I}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j' H_f^{-1}\widehat{e}_{jt}, \alpha_j' H_g^{-1})A_{dt}, \quad d = 1, ..., 6.
\end{aligned}$$

Then $R_{6j} = \widehat{D}_j^{-1}\Gamma^j$. Let d_j be the first diagonal block of D_j^{-1} (which is a block diagonal matrix), We have

$$\begin{aligned}
[\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|r_{6j}\|]^2 &\leq \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\widehat{D}_j^{-1} - D_j^{-1}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j\|^2 + [\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|d_j\mathcal{U}(\Gamma^j)\|]^2\\
&\leq O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j\|^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma^j)\|^2\\
&\leq O_P(1)\sum_{d=2}^{5}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_d\|^2 + O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_0\|^2\\
&\quad + O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_1\|^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma^j_1)\|^2\\
&\quad + O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_6\|^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma^j_6)\|^2.
\end{aligned}$$

In the above inequalities, we treat individual Γ^j_d differently. This is because we aim to show the following:

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_d\|^2 &= O_P(C_{NT}^{-4}), \quad d = 0, 2\sim 5,\\
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_d\|^2 &= O_P(C_{NT}^{-2}), \quad d = 1, 6,\\
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma^j_d)\|^2 &= O_P(C_{NT}^{-4}), \quad d = 1, 6.
\end{aligned}$$

That is, while it is not likely to prove $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma^j_d\|^2$ are also $O_P(C_{NT}^{-4})$ for $d = 1, 6$, it can be proved that their upper bounds are.

Step 1: $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_d^j\|^2$ for $d=0,2\sim 5$. Let

$$m_t=\max_{jt}|\widehat{e}_{jt}^4-e_{jt}^4|+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}e_{jt}^4+f_t^2+g_t^2+\max_{jt}|\widehat{e}_{jt}^2-e_{jt}^2|+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}e_{jt}^2+1.$$

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_0^j\|^2 \leq\;& \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\widehat{f}_t\frac{1}{N}\sum_i\widehat{e}_{jt}^2u_{it}\lambda_j\lambda_i e_{it}\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\widehat{f}_t\frac{1}{N}\sum_i\widehat{e}_{jt}u_{it}\alpha_j\alpha_i\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\widehat{g}_t\frac{1}{N}\sum_i u_{it}\lambda_j\lambda_i\widehat{e}_{jt}e_{it}\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}\widehat{g}_t\frac{1}{N}\sum_i u_{it}\alpha_j\alpha_i\|^2\\
\leq\;& O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}m_t\|\frac{1}{N}\sum_i u_{it}\lambda_i e_{it}\|^2+O_P(C_{NT}^{-2})\frac{1}{T}\sum_{t\notin I}m_t\|\frac{1}{N}\sum_i u_{it}\alpha_i\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}g_t\frac{1}{N}\sum_i u_{it}\lambda_j\lambda_i e_{jt}e_{it}\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}g_t\frac{1}{N}\sum_i u_{it}\alpha_j\alpha_i\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}e_{jt}^2f_t\frac{1}{N}\sum_i u_{it}\lambda_j\lambda_i e_{it}\|^2+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T}\sum_{t\notin I}e_{jt}f_t\frac{1}{N}\sum_i u_{it}\alpha_j\alpha_i\|^2\\
\leq\;& O_P(C_{NT}^{-4}).
\end{aligned}$$

Now let

$$\mathcal{D}=\frac{1}{T_0}\sum_{t\in I^c}\|\widehat{f}_t-H_f f_t\|^2+\|\widehat{g}_t-H_g g_t\|^2=O_P(C_{NT}^{-2}).$$

$$\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_2^j\|^2\\
=\;&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})(\widehat{B}_t^{-1}-B^{-1})\frac{1}{N}\sum_i\begin{pmatrix}H_1'\lambda_i e_{it}\\ H_2'\alpha_i\end{pmatrix}u_{it}\|^2\\
\leq\;&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\mathcal{D}\frac{1}{T_0}\sum_{t\in I^c}(|\widehat{e}_{jt}|^2+1)^2\left\|\frac{1}{N}\sum_i\begin{pmatrix}\lambda_i e_{it}\\ \alpha_i\end{pmatrix}u_{it}\right\|^2\max_{t\leq T}\|\widehat{B}_t^{-1}-B^{-1}\|\\
&+\frac{1}{T_0}\sum_{t\in I^c}\|\widehat{B}_t^{-1}-B^{-1}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}\left\|\begin{pmatrix}f_t\widehat{e}_{jt}\\ g_t\end{pmatrix}\right\|^2(|\widehat{e}_{jt}|+1)^2\left\|\frac{1}{N}\sum_i\begin{pmatrix}\lambda_i e_{it}\\ \alpha_i\end{pmatrix}u_{it}\right\|^2\\
\leq\;&O_P(C_{NT}^{-2})\frac{1}{T_0}\sum_{t\in I^c}n_t\left\|\frac{1}{N}\sum_i\begin{pmatrix}\lambda_i e_{it}\\ \alpha_i\end{pmatrix}u_{it}\right\|^2=O_P(C_{NT}^{-4}),
\end{aligned}$$

where

$$n_t=\max_{t\leq T}\|\widehat{B}_t^{-1}-B^{-1}\|+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}(|e_{jt}|+1+\max_{jt}|e_{jt}-\widehat{e}_{jt}|)^2[\|f_t\|^2+\|g_t\|^2+\|f_te_{jt}\|^2].$$

Next, let

$$\mathcal{A}_t = \frac{1}{N}\sum_i \begin{pmatrix} \widetilde{\lambda}_i \widehat{e}_{it} - H_1'\lambda_i e_{it} \\ \widetilde{\alpha}_i - H_2'\alpha_i \end{pmatrix} u_{it}.$$

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_3^j\|^2 &= \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})\widehat{B}_t^{-1}\mathcal{A}_t\|^2\\
&\leq \frac{1}{T_0}\sum_{t\in I^c}\mathcal{D}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}(e_{jt}^2+1)^2\|\mathcal{A}_t\|^2 + O_P(1)\frac{1}{T_0}\sum_{t\in I^c}\|\mathcal{A}_t\|^2\\
&\leq O_P(C_{NT}^{-4}) + \frac{1}{T_0}\sum_{t\in I^c}\|\frac{1}{N}\sum_i(\widetilde{\alpha}_i - H_2'\alpha_i)u_{it}\|^2\\
&\quad+\frac{1}{T_0}\sum_{t\in I^c}\|\frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it}\|^2 + \frac{1}{T_0}\sum_{t\in I^c}\|\frac{1}{N}\sum_i\lambda_i(\widehat{e}_{it}-e_{it})u_{it}\|^2\\
&\leq O_P(C_{NT}^{-4}),
\end{aligned}$$

where the last equality is due to Lemma G.2 and :

$$\begin{aligned}
&\frac{1}{T_0}\sum_{t\in I^c}\mathsf{E}[|\frac{1}{N}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)e_{it}u_{it}|^2|E,I]\\
=\;&\frac{1}{T_0}\sum_{t\in I^c}\frac{1}{N^2}(\widetilde{\Lambda}-\Lambda H_1)'\mathsf{diag}(e_t)\mathsf{Var}(u_t|E,I)\mathsf{diag}(e_t)(\widetilde{\Lambda}-\Lambda H_1)\\
\leq\;& O_P(1)\frac{1}{N^2}\sum_i(\widetilde{\lambda}_i - H_1'\lambda_i)^2\frac{1}{T_0}\sum_{t\in I^c}\mathsf{E}_I e_{it}^2 = O_P(C_{NT}^{-4}).
\end{aligned}$$

Next, for C_{dt} defined in the proof of Lemma G.9,

$$\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_4^j\|^2 &= \sum_{d=1}^{16}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})\widehat{B}_t^{-1}C_{dt}\|^2\\
&\leq \sum_{d=1}^{16}\mathcal{D}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}|e_{jt}|^4\|C_{dt}\|^2 + O_P(1)\sum_{d=1}^{16}\frac{1}{T_0}\sum_{t\in I^c}\|C_{dt}\|^2\\
&= O_P(C_{NT}^{-4}).
\end{aligned}$$

For B_{dt} defined in the proof of Lemma G.9, and

$$n_t = \|f_t\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}(e_{jt}^2+|e_{jt}|)^2 + \|g_t\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}(1+|e_{jt}|)^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}(|e_{jt}|^2+1)^2,$$

$$\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_5^j\|^2 = \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\sum_{d=1}^{8}\frac{1}{T_0}\sum_{t\in I^c}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})(\widehat{B}_t^{-1}-B^{-1})B_{dt}\|^2$$

$$
\begin{aligned}
&\leq [\max_{t\leq T}\|\widehat{B}_t^{-1}-B^{-1}\|\mathcal{D}+O_P(C_{NT}^{-2})]\sum_{d=1}^{8}\frac{1}{T_0}\sum_{t\in I^c}n_t\|B_{dt}\|^2\\
&\leq O_P(C_{NT}^{-2})\sum_{d=1}^{8}\frac{1}{T_0}\sum_{t\in I^c}n_t\|B_{dt}\|^2=O_P(C_{NT}^{-4}),
\end{aligned}
$$

where the last equality follows from simply applying Cauchy-Schwarz to show $\frac{1}{T_0}\sum_{t\in I^c}n_t\|B_{dt}\|^2=O_P(C_{NT}^{-2})$ for $d=1\sim 8$.

Step 2: $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_d^j\|^2$ for $d=1,6$.

By Lemma G.8, $\max_{t\leq T}[\|\widehat{B}_t^{-1}+\|\widehat{S}_t\|]=O_P(1)$, $\frac{1}{T}\sum_t\|\widehat{B}_t^{-1}\widehat{S}_t-B^{-1}S\|^2=O_P(C_{NT}^{-2})$.

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_1^j\|^2\\
&=\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})(\widehat{B}_t^{-1}\widehat{S}_t-B^{-1}S)\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix}\|^2\\
&\leq O_P(C_{NT}^{-2})\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}|\widehat{e}_{jt}|^2(|\widehat{e}_{jt}|^2+1)\|\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix}\|^2\\
&\quad+\frac{1}{T}\sum_t\|(\widehat{B}_t^{-1}\widehat{S}_t-B^{-1}S)\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}\|\begin{pmatrix}f_t\widehat{e}_{jt}\\ g_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})\|^2\|\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix}\|^2\\
&\leq O_P(C_{NT}^{-2}).\\
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\Gamma_6^j\|^2\\
&=\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\sum_{d=1}^{6}\frac{1}{T_0}\sum_{t\notin I}\begin{pmatrix}\widehat{f}_t\widehat{e}_{jt}\\ \widehat{g}_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})B^{-1}\begin{pmatrix}H_1'\sum_{d=1}^4 B_{dt}\\ H_2'\sum_{d=5}^8 B_{dt}\end{pmatrix}\|^2\\
&\leq O_P(C_{NT}^{-2})\sum_{d=1}^{8}\frac{1}{T_0}\sum_{t\in I^c}n_t\|B_{dt}\|^2\\
&\quad+\sum_{d=1}^{8}\frac{1}{T}\sum_t\|B_{dt}\|^2\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}\|\begin{pmatrix}f_t\widehat{e}_{jt}\\ g_t\end{pmatrix}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})\|^2=O_P(C_{NT}^{-2}).
\end{aligned}
$$

Step 3: $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma_d^j)\|^2$ for $d=1,6$. Let

$$
q_t=\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\sqrt{e_{jt}^4+e_{jt}^2(\|f_t\|^2+\|g_t\|^2)(\|f_t\|^2+1)}
$$

$$
\begin{aligned}
\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma_1^j)\|^2 &= \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt}(\lambda_j'H_f^{-1}\widehat{e}_{jt},\alpha_j'H_g^{-1})(A_{1t,a}+A_{1t,b}+A_{1t,c})\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix}\|^2\\
&\leq O_P(C_{NT}^{-4})+O_P(C_{NT}^{-2})\frac{1}{T_0}\sum_{t\in I^c}q_t\|\widehat{S}_t-S\|^2\\
&\quad +O_P(1)\frac{1}{T_0}\sum_{t\in I^c}[\|\Delta_{t1,1}\|^2+\|\Delta_{t1,2}\|^2]+O_P(1)\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}f_t^2e_{jt}\Delta_{t2,1}\|^2
\end{aligned}
$$

where $A_{1t,a}=(\widehat{B}_t^{-1}-B^{-1})(\widehat{S}_t-S)$, $A_{1t,b}=(\widehat{B}_t^{-1}-B^{-1})S$, $A_{1t,c}=B^{-1}(\widehat{S}_t-S)$,

$$
\widehat{S}_t-S = \begin{pmatrix}\Delta_{t1,1} & \Delta_{t1,2}\\ \Delta_{t2,1} & 0\end{pmatrix}
$$

$$
\begin{aligned}
\Delta_{t1,1} &= \frac{1}{N}\sum_i\widetilde{\lambda}_i\lambda_i'H_1e_{it}(\widehat{e}_{it}-e_{it})+\widetilde{\lambda}_i\widetilde{\lambda}_i'(\widehat{e}_{it}^2-e_{it}^2)+(\widetilde{\lambda}_i\lambda_i'H_1-\widetilde{\lambda}_i\widetilde{\lambda}_i')e_{it}^2-H_1'\lambda_i(\lambda_i'H_1-\widetilde{\lambda}_i')\mathsf{E}e_{it}^2\\
\Delta_{t1,2} &= \frac{1}{N}\sum_i\widetilde{\lambda}_i(\widehat{e}_{it}-e_{it})(\alpha_i'H_2-\widetilde{\alpha}_i')+\widetilde{\lambda}_ie_{it}(\alpha_i'H_2-\widetilde{\alpha}_i')\\
\Delta_{t2,1} &= \frac{1}{N}\sum_i\widetilde{\alpha}_i\lambda_i'H_1(e_{it}-\widehat{e}_{it})+\widetilde{\alpha}_i(\lambda_i'H_1-\widetilde{\lambda}_i')(\widehat{e}_{it}-e_{it})+\widetilde{\alpha}_i(\lambda_i'H_1-\widetilde{\lambda}_i')e_{it}
\end{aligned}
$$

By Lemma G.8, $\frac{1}{T_0}\sum_{t\in I^c}[\|\Delta_{t1,1}\|^2+\|\Delta_{t1,2}\|^2]=O_P(C_{NT}^{-4})$. It is straightforward to use Cauchy-Schwarz to show $\frac{1}{T_0}\sum_{t\in I^c}q_t\|\widehat{S}_t-S\|^2=O_P(C_{NT}^{-2})$. In addition,

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}f_t^2e_{jt}\Delta_{t2,1}\|^2\leq\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\frac{1}{N}\sum_i\alpha_i\lambda_if_t^2e_{jt}(e_{it}-\widehat{e}_{it})\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i\frac{1}{T_0}\sum_{t\in I^c}f_t^2e_{jt}\alpha_i(\lambda_i'H_1-\widetilde{\lambda}_i')(\widehat{e}_{it}-e_{it})\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i\frac{1}{T_0}\sum_{t\in I^c}f_t^2\alpha_i(\lambda_i'H_1-\widetilde{\lambda}_i')(e_{it}e_{jt}-\mathsf{E}e_{it}e_{jt})\|^2\\
&+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i\frac{1}{T_0}\sum_{t\in I^c}f_t^2\alpha_i(\lambda_i'H_1-\widetilde{\lambda}_i')\mathsf{E}e_{it}e_{jt}\|^2+O_P(C_{NT}^{-4}).
\end{aligned}
$$

By Lemma G.2, $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\frac{1}{N}\sum_i\alpha_i\lambda_if_t^2e_{jt}(e_{it}-\widehat{e}_{it})\|^2=O_P(C_{NT}^{-4})$. The middle two terms follow from Cauchy-Schwarz. Also, by Cauchy-Schwarz

$$
\begin{aligned}
&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{N}\sum_i\alpha_i(\lambda_i'H_1-\widetilde{\lambda}_i')\frac{1}{T_0}\sum_{t\in I^c}f_t^2\mathsf{E}e_{it}e_{jt}\|^2\\
&= O_P(C_{NT}^{-4})\max_j\sum_{i=1}^N|\mathsf{E}e_{it}e_{jt}|=O_P(C_{NT}^{-4}).
\end{aligned}
$$

Hence $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t^2 e_{jt}\Delta_{t2,1}\|^2 = O_P(C_{NT}^{-4})$. Thus $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma_1^j)\|^2 = O_P(C_{NT}^{-4})$.

Finally,

$$\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\mathcal{U}(\Gamma_6^j)\|^2 = \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt}(\widehat{e}_{jt}\lambda_j' H_f^{-1}, \alpha_j' H_g^{-1})B^{-1}\begin{pmatrix} H_1' & 0 \\ 0 & H_2\end{pmatrix}\begin{pmatrix}\sum_{d=1}^4 B_{dt} \\ \sum_{d=5}^8 B_{dt}\end{pmatrix}\|^2$$

It suffices to prove

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt}^2 B_{dt}\|^2 &= O_P(C_{NT}^{-4}), \quad d = 1\sim 4\\ \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt} B_{dt}\|^2 &= O_P(C_{NT}^{-4}), \quad d = 5\sim 8.\end{aligned}$$

This is proved by Lemma G.15. □

Lemma G.15. *For B_{dt} defined in (G.16) and (G.17),*

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt}^2 B_{dt}\|^2 &= O_P(C_{NT}^{-4}), \quad d = 1\sim 4\\ \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt} B_{dt}\|^2 &= O_P(C_{NT}^{-4}), \quad d = 5\sim 8.\end{aligned}$$

Proof. For $r = 1, 2$ and all $d = 1, .., 8$,

$$\begin{aligned}&\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c}\widehat{f}_t\widehat{e}_{jt}^r B_{dt}\|^2 \le \max_{jt}|\widehat{e}_{jt}^r - e_{jt}^r|\frac{1}{T_0}\sum_{t\in I^c}\|\widehat{f}_t - H_f f_t\|^2\frac{1}{T_0}\sum_{t\in I^c}\|B_{dt}\|^2\\ &+\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\frac{1}{T_0}\sum_{t\in I^c}\|\widehat{f}_t - H_f f_t\|^2 e_{jt}^{2r}\frac{1}{T_0}\sum_{t\in I^c}\|B_{dt}\|^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t\widehat{e}_{jt}^r B_{dt}\|^2\\ \le\ & O_P(C_{NT}^{-4}) + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt}^r B_{dt}\|^2 + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t(\widehat{e}_{jt}^r - e_{jt}^r)B_{dt}\|^2.\end{aligned}$$

It remains to show the following three steps.

Step 1. Show $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt}^2 B_{dt}\|^2 = O_P(C_{NT}^{-4})$ for $d = 1\sim 4$.

When $d = 1, 4$, it follows from Cauchy-Schwarz.

$$\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt}^2 B_{2t}\|^2 = O_P(C_{NT}^{-4}) + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t^2 e_{jt}^2\frac{1}{N}\sum_i \lambda_i^2 e_{it}(\widehat{e}_{it} - e_{it})\|^2,$$

which is $O_P(C_{NT}^{-4})$ due to Lemma G.2.

Next, by Lemma G.5, for any fixed sequence c_i, $\frac{1}{T}\sum_{t\notin I}\|\frac{1}{N}\sum_i c_i e_{it}(\dot\lambda_i - H_1'\lambda_i)\|^2 = O_P(C_{NT}^{-4})$. So

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt}^2 B_{3t}\|^2 &= O_P(C_{NT}^{-4}) + \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t^2 e_{jt}^2 \frac{1}{N}\sum_i \lambda_i e_{it} l_i' w_t(\dot\lambda_i - H_1'\lambda_i)'\|^2 \\ &\leq O_P(C_{NT}^{-4}).\end{aligned}$$

Step 2. Show $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt} B_{dt}\|^2 = O_P(C_{NT}^{-4})$ for $d = 5 \sim 8$.

When $d = 5$, it follows from Cauchy-Schwarz.

Next by Lemma G.2,

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt} B_{6t}\|^2 &= \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t^2 e_{jt}\frac{1}{N}\sum_i \alpha_i(\widehat{e}_{it} - e_{it})\lambda_i\|^2 + O_P(C_{NT}^{-4}) \\ &\leq O_P(C_{NT}^{-4}). \\ \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt} B_{7t}\|^2 &= \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t^2 e_{jt} w_t\|^2 \|\frac{1}{N}\sum_i \alpha_i l_i(\dot\lambda_i - H_1'\lambda_i)\|^2 + O_P(C_{NT}^{-4}) \\ &\leq O_P(C_{NT}^{-4}).\end{aligned}$$

Next, by Lemma G.6,

$$\begin{aligned}\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt} B_{8t}\|^2 &= \frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t e_{jt}(\widetilde{f}_t - H_1^{-1}f_t)w_t\|^2 \|\frac{1}{N}\sum_i \alpha_i l_i \lambda_i\|^2 \\ &\leq O_P(C_{NT}^{-4}).\end{aligned}$$

Step 3. Show $\frac{1}{|\mathcal{G}|_0}\sum_{j\in\mathcal{G}}\|\frac{1}{T_0}\sum_{t\in I^c} f_t(\widehat{e}_{jt}^r - e_{jt}^r)B_{dt}\|^2 = O_P(C_{NT}^{-4})$ for $d = 1 \sim 8$.

It is bounded by Cauchy-Schwarz.

$\square$

G.5. **Technical lemmas for covariance estimations.**

Lemma G.16. *Suppose uniformly over $i \leq N$, the following terms are $o_P(1)$: $q\|\frac{1}{T}\sum_t f_t e_{it} u_{it}\|, q\|\frac{1}{T}\sum_t g_t u_{it}\|, q\|\frac{1}{T}\sum_s f_s w_s' e_{is}\|$, $q|\frac{1}{T}\sum_{tj} c_j(m_{jt}m_{it} - \mathsf{E}m_{jt}m_{it})|$ for $m_{it} \in \{e_{it}u_{it}, u_{it}\}$, $|\frac{1}{T}\sum_t e_{it}^2 - \mathsf{E}e_{it}^2|$. Suppose $q\max_{t\leq T}\|\frac{1}{N}\sum_i \lambda_i\omega_{it} + \alpha_i u_{it}\| = o_P(1)$ and $qC_{NT}^{-1}(\max_{it}|e_{it}u_{it}|) = o_P(1)$ (which are satisfied given our assumptions), where*

$$q := \max_{it}(\|f_t x_{it}\| + 1 + |x_{it}|).$$

Then (i) $\max_i[\|\widehat{\lambda}_i - H_f'^{-1}\lambda_i\| + \|\widehat{\alpha}_i - H_g'^{-1}\alpha_i\|]q = o_P(1)$.

(ii) $\max_{t\leq T}[\|\widehat{f}_t - H_f f_t\| + \|\widehat{g}_t - H_g g_t\|]q = o_P(1)$.

(iii) $\max_{is}|\widehat{u}_{is} - u_{is}| = o_P(1)$,

(iv) $\max_i |\widehat{\sigma}_i^2 - \sigma_i^2| = o_P(1)$, *where* $\sigma_i^2 = \mathsf{E}e_{it}^2$ *and* $\widehat{\sigma}_i^2 = \frac{1}{T}\sum_t \widehat{e}_{it}^2$.

Proof. (i) In view of (E.5), it suffices to show $\max_i \|R_{di}\|q = o_P(1)$ for all $d = 1 \sim 6$. First, by our assumption, $q\max_i \|R_{1i}\| = o_P(1)$. Also by Lemma G.10, $\max_i \|\widehat{D}_i^{-1} - D_i^{-1}\| = o_P(1)$. Also, $q\max_i \|R_{2i}\| = o_P(1)$ follows from $q\max_{it} |\widehat{e}_{it} - e_{it}||e_{it} + u_{it}| = o_P(1)$.

Term $\max_i[\|R_{3i}\| + \|R_{4i}\|]$**.** It suffices to prove $q\max_i \|R_{3i,d}\| = o_P(1)$ for $R_{3i,d}$ defined in Lemma G.12.

$$\begin{aligned}\max_i \|R_{3i,1}\|q &\leq o_P(1) + q\max_i \|\frac{1}{T}\sum_s(\widehat{f}_s - H_f f_s)(\widehat{e}_{is} - e_{is})(\mu_{it}\lambda_i' f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\| \\ &\quad + q\max_i \|\frac{1}{T}\sum_s(\widehat{f}_s - H_f f_s)e_{is}(\mu_{it}\lambda_i' f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\| \\ &\quad + qO_P(1)\max_i \|\frac{1}{T}\sum_s f_s(\widehat{e}_{is} - e_{is})(\mu_{it}\lambda_i' f_s - \widehat{\mu}_{it}\dot{\lambda}_i'\widetilde{f}_s)\| \\ &= o_P(1)\end{aligned}$$

following from the Cauchy-Schwarz, $q\max_{it} |\widehat{e}_{it} - e_{it}||e_{it}| = o_P(1)$, $qC_{NT}^{-1}\max_{it} e_{it}^2 = o_P(1)$ and Lemma G.17. Similarly, $\max_i \|R_{3i,d}\|q = o_P(1)$ for all $d = 2 \sim 7$.

Term $\max_i \|R_{5i}\|$**.** By the assumption:

$$q\max_i \|\frac{1}{T}\sum_{tj}\lambda_j(\omega_{jt}\omega_{it} - \mathsf{E}\omega_{jt}\omega_{it})\| + q\max_i \|\frac{1}{T}\sum_{tj}\alpha_j(u_{jt}u_{it} - \mathsf{E}u_{jt}u_{it})\| = o_P(1).$$

It suffices to prove for A_{dt} defined in (E.2),

$$q\sum_{d=1}^{6}\max_i \|\frac{1}{T}\sum_t \begin{pmatrix}\widehat{e}_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix} A_{dt}u_{it}\| = o_P(1).$$

First, for $A_{1t,a} = (\widehat{B}_t^{-1} - B^{-1})(\widehat{S}_t - S)$, $A_{1t,b} = (\widehat{B}_t^{-1} - B^{-1})S$, $A_{1t,c} = B^{-1}(\widehat{S}_t - S)$,

$$\begin{aligned}& q\max_i \|\frac{1}{T}\sum_t \begin{pmatrix}\widehat{e}_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix} A_{1t}u_{it}\| \leq q\max_i \|\frac{1}{T}\sum_t \begin{pmatrix}e_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix} A_{1t,a}\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix} u_{it}\| \\ & + q\max_i \|\frac{1}{T}\sum_t \begin{pmatrix}e_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix} A_{1t,b}\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix} u_{it}\| \\ & + q\max_i \|\frac{1}{T}\sum_t \begin{pmatrix}e_{it}\mathsf{I} & 0\\ 0 & \mathsf{I}\end{pmatrix} A_{1t,c}\begin{pmatrix}H_1^{-1}f_t\\ H_2^{-1}g_t\end{pmatrix} u_{it}\| + o_P(1) \\ = \; & qO_P(C_{NT}^{-1})(\max_{it}|e_{it}u_{it}|) + o_P(1) = o_P(1).\end{aligned}$$

Similarly, $d = 2 \sim 6$, Cauchy-Schwarz and $qO_P(C_{NT}^{-1})(\max_{it} |e_{it}u_{it}|) = o_P(1)$ imply

$$q \max_i \| \frac{1}{T} \sum_t \begin{pmatrix} \widehat{e}_{it}\mathsf{I} & 0 \\ 0 & \mathsf{I} \end{pmatrix} A_{dt} u_{it} \| \leq o_P(1).$$

Term $\max_i \|R_{6i}\|$**.** Lemma G.10 shows $\max_i \|\widehat{D}_i^{-1}\| = O_P(1)$. It suffices to prove $q \max_i \|\Gamma_d^i\| = o_P(1)$, where Γ_d^i is as defined in the proof of Lemma G.14.

$$\begin{aligned} q \max_j \|\Gamma_0^j\| &= q \max_j \| \frac{1}{T_0} \sum_{t \notin I} \begin{pmatrix} \widehat{f}_t \widehat{e}_{jt} \\ \widehat{g}_t \end{pmatrix} (\lambda_j' H_f^{-1} \widehat{e}_{jt} b_1, \alpha_j' H_g^{-1} b_2) \frac{1}{N} \sum_i \begin{pmatrix} H_1' \lambda_i e_{it} \\ H_2' \alpha_i \end{pmatrix} u_{it} \| \\ &= qO_P(C_{NT}^{-1}) \max_j |\widehat{e}_{jt}^2 + \widehat{e}_{jt}| = o_P(1) \\ \max_j \|\Gamma_d^j\| &= q \max_j \| \frac{1}{T_0} \sum_{t \notin I} \begin{pmatrix} \widehat{f}_t \widehat{e}_{jt} \\ \widehat{g}_t \end{pmatrix} (\lambda_j' H_f^{-1} \widehat{e}_{jt}, \alpha_j' H_g^{-1}) A_{dt} \| = o_P(1) \end{aligned}$$

due to $O_P(C_{NT}^{-1}) \max_j |\widehat{e}_{jt}^2 + \widehat{e}_{jt}| = o_P(1)$ and $\max_i \|\dot{\lambda}_i - H_1'\lambda_i\| = o_P(1)$.

(ii) In view of (E.2), we have $\max_{t \leq T} \| \frac{1}{N} \sum_i \lambda_i \omega_{it} + \alpha_i u_{it} \| = o_P(1)$. So it suffices to prove $\max_{t \leq T} \|A_{dt}\| q = o_P(1)$ for $d = 1 \sim 8$.

The conclusion of Lemma G.8 can be strengthened to ensure that

$q \max_{t \leq T}(\|f_t\| + \|g_t\|) \max_{t \leq T}[\|\widehat{S}_t - S\| + \|\widehat{B}_t^{-1} - B^{-1}\|] = o_P(1)$. So this is true for $d = 1, 2$. Next, $C_{NT}^{-1} q = o_P(1)$ and Lemma G.17 shows $q \max_t \|\widetilde{f}_t - H_1^{-1} f_t\| = o_P(1)$. So it is easy to verify using the Cauchy-Schwarz that it also holds true for $d = 3 \sim 8$.

(iii) Uniformly in i, s, by part (ii),

$$\begin{aligned} |\widehat{u}_{is} - u_{is}| &\leq |x_{it}| \|\widehat{\lambda}_i - \lambda_i\| \|\widehat{f}_t\| + C|x_{it}| \|\widehat{f}_t - f_t\| \\ &\quad + \|\widehat{\alpha}_i - \alpha_i\| \|\widehat{g}_t\| + C\|\widehat{g}_t - g_t\| = o_P(1). \end{aligned}$$

(iv) Uniformly in $i \leq N$,

$$\frac{1}{T} \sum_t \widehat{e}_{it}^2 - \mathsf{E} e_{it}^2 \leq |\frac{1}{T} \sum_t \widehat{e}_{it}^2 - e_{it}^2| + \frac{1}{T} \sum_t e_{it}^2 - \mathsf{E} e_{it}^2 = o_P(1).$$

$\square$

Lemma G.17. *Suppose the following are* $o_P(1)$ *(which are satisfied given our assumptions):* $\max_{it} |\widehat{e}_{it} - e_{it}| |e_{it} + u_{it}|$, $C_{NT}^{-1} \max_{it} |x_{it}|^3 \|f_t\|$, $\max_i |\frac{1}{T} \sum_t u_{it}^2 - \mathsf{E} u_{it}^2|$, $q \max_i \|\frac{1}{T} \sum_t f_t e_{it} u_{it}\|$, $q \max_i \|\frac{1}{T} \sum_t f_t w_t u_{it}\|$, $q \max_{it} |x_{it}| \max_i \|\frac{1}{T} \sum_t u_{it} g_t\|$, $\max_i \|\frac{1}{T} \sum_t \omega_{it}^2 - \mathsf{E}\omega_{it}^2\| + \max_i \|\frac{1}{T} \sum_t u_{it}^2 w_t^2 - \mathsf{E} u_{it}^2 w_t^2\|$,

$\max_{t\leq T}|\frac{1}{N}\sum_i u_{it}^2 - \mathsf{E}u_{it}^2|$, $q\max_{t\leq T}[\|\frac{1}{N}\sum_i \lambda_i e_{it}u_{it}\| + \|\frac{1}{N}\sum_i \lambda_i w_t u_{it}\|]$, $q\max_{it}\max_{t\leq T}\|\frac{1}{N}\sum_i u_{it}\alpha_i\|$, $\max_{t\leq T}|\frac{1}{N}\sum_i \omega_{it}^2 - \mathsf{E}\omega_{it}^2|$, $\max_i\|\frac{1}{N}\sum_i u_{it}^2 w_t w_t' - \mathsf{E}u_{it}^2 w_t w_t'\|$,

where

$$q := \max_{it}(\|f_t x_{it}\| + 1 + |x_{it}|).$$

Then (i) $q\max_{it}\|\dot{\lambda}_i - H_1'\lambda_i\| = o_P(1)$. *(ii)* $q\max_t\|\widetilde{f}_t - H_1^{-1}f_t\| = o_P(1)$.

Proof. Recall

$$\begin{aligned}
\dot{\lambda}_i &= \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}y_i \\
&= H_1'\lambda_i + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'} - \widetilde{G})H_2'\alpha_i \\
&\quad + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'} - \widetilde{F})H_1'\lambda_i + \widetilde{D}_{\lambda i}^{-1}\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i.
\end{aligned}$$

(i) By Lemma G.4, $\max_i\|\widetilde{D}_{\lambda i}^{-1}\| = O_P(1)$.

First we show $q\max_i\|\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}(GH_2^{-1'} - \widetilde{G})\| = o_P(1)$. It is in fact bounded by

$$O_P(C_{NT}^{-1})\max_{it}|x_{it}|q = o_P(1).$$

Next, $q\max_i\|\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}\mathsf{diag}(X_i)(FH_1^{-1'} - \widetilde{F})H_1'\lambda_i\|$ is bounded by

$$O_P(C_{NT}^{-1})q\max_{it}|x_{it}|^2 = o_P(1)$$

Finally, note that $\max_i\frac{1}{T}\|u_i\|^2 = \max_i|\frac{1}{T}\sum_t u_{it}^2 - \mathsf{E}u_{it}^2| + O_P(1) = O_P(1)$. So

$$\begin{aligned}
& q\max_i\|\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)M_{\widetilde{g}}u_i\| \\
\leq\ & qO_P(C_{NT}^{-1})\max_{it}|x_{it}|\max_i\frac{1}{\sqrt{T}}\|u_i\| + q\max_i\|\frac{1}{T}\sum_t\widetilde{f}_t x_{it}u_{it}\| \\
& + O_P(1)q\max_i\|\frac{1}{T}\widetilde{F}'\mathsf{diag}(X_i)G\|\frac{1}{T}\max_i\|G'u_i\| \\
\leq\ & o_P(1) + O_P(C_{NT}^{-1})q\max_i(\frac{1}{T}\sum_t u_{it}^2x_{it}^2)^{1/2} \\
& + O_P(1)q\max_i[\|\frac{1}{T}\sum_t f_t e_{it}u_{it}\| + \|\frac{1}{T}\sum_t f_t w_t u_{it}\|] \\
& + O_P(1)q\max_{it}|x_{it}|\max_i\|\frac{1}{T}\sum_t u_{it}g_t\| \\
\leq\ & o_P(1) + O_P(C_{NT}^{-1})q\max_i(\frac{1}{T}\sum_t\omega_{it}^2 - \mathsf{E}\omega_{it}^2)^{1/2}
\end{aligned}$$

$$+O_P(C_{NT}^{-1})q\max_i(\frac{1}{T}\sum_t u_{it}^2w_t^2-\mathsf{E}u_{it}^2w_t^2)^{1/2}=o_P(1).$$

(ii) Recall that

$$\begin{aligned}\widetilde{f}_s &= \widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}y_s\\ &= H_1^{-1}f_s+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}(AH_2-\widetilde{A})H_2^{-1}g_s\\ &\quad+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}\mathsf{diag}(X_s)(\Lambda H_1-\widetilde{\Lambda})H_1^{-1}f_s\\ &\quad+\widetilde{D}_{fs}^{-1}\frac{1}{N}\widetilde{\Lambda}'\mathsf{diag}(X_s)M_{\widetilde{\alpha}}u_s.\end{aligned}$$

By Lemma G.3, $\max_s\|\widetilde{D}_{fs}^{-1}\|=O_P(1)$. The rest of the proof is very similar to part (i), so we omit details to avoid repetitions. The only extra assumption, parallel to those of part (i), is that the following terms are $o_P(1)$: $\max_{t\leq T}|\frac{1}{N}\sum_i u_{it}^2-\mathsf{E}u_{it}^2|$, $q\max_{t\leq T}[\|\frac{1}{N}\sum_i\lambda_ie_{it}u_{it}\|+\|\frac{1}{N}\sum_i\lambda_iw_tu_{it}\|]$, $q\max_{it}\max_{t\leq T}\|\frac{1}{N}\sum_i u_{it}\alpha_i\|$, $\max_{t\leq T}|\frac{1}{N}\sum_i\omega_{it}^2-\mathsf{E}\omega_{it}^2|$, $\max_i\|\frac{1}{N}\sum_i u_{it}^2w_tw_t'-\mathsf{E}u_{it}^2w_tw_t'\|$.

□

References

Acemoglu, D., Johnson, S., Robinson, J. A. and Yared, P. (2008). Income and democracy. *American Economic Review* **98** 808–42.

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

Beck, A. and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences* **2** 183–202.

Bülhmann, P. and van der Geer, S. (2011). Statistics for high dimensional data. *Statistics (New York). Springer-Verlag, Berlin* .

Chen, Y., Chi, Y., Fan, J., Ma, C. and Yan, Y. (2019). Noisy matrix completion: Understanding statistical guarantees for convex relaxation via nonconvex optimization. *arXiv preprint arXiv:1902.07698* .

Klopp, O. (2014). Noisy low-rank matrix completion with general sampling distribution. *Bernoulli* **20** 282–303.

Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems: Ecole d'Eté de Probabilités de Saint-Flour XXXVIII-2008*, vol. 2033. Springer Science & Business Media.

Ma, S., Goldfarb, D. and Chen, L. (2011). Fixed point and bregman iterative methods for matrix rank minimization. *Mathematical Programming* **128** 321–353.

Negahban, S. and Wainwright, M. J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *The Annals of Statistics* **39** 1069–1097.

Negahban, S. and Wainwright, M. J. (2012). Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *Journal of Machine Learning Research* **13** 1665–1697.

Recht, B., Fazel, M. and Parrilo, P. A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review* **52** 471–501.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027* .

Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, vol. 47. Cambridge university press.

Department of Economics, MIT, Cambridge, MA 02139
Email address: vchern@mit.edu

Booth School of Business, University of Chicago, Chicago, IL 60637
Email address: Christian.Hansen@chicagobooth.edu

Department of Economics, Rutgers University, New Brunswick, NJ 08901
Email address: yuan.liao@rutgers.edu

Department of Economics, Brandeis University, 415 South St, Waltham, MA 02453
Email address: yinchuzhu@brandeis.edu