Internet Appendix to Target PCA: Transfer Learning Large Dimensional Panel Data

Junting Duan* Markus Pelger† Ruoxuan Xiong‡

August 25, 2023

Abstract

This Internet Appendix collects the detailed proofs for all the theoretical statements in the main text, the data description and additional simulation results.

Keywords: Factor Analysis, Principal Components, Transfer Learning, Multiple Data Sets, Large-Dimensional Panel Data, Large N and T, Missing Data, Weak Factors, Causal Inference
JEL classification: C14, C38, C55, G12

*Stanford University, Department of Management Science & Engineering, Email: duanjt@stanford.edu.
†Stanford University, Department of Management Science & Engineering, Email: mpelger@stanford.edu.
‡Emory University, Department of Quantitative Theory and Methods, Email: ruoxuan.xiong@emory.edu.

Contents

IA.A Data for Empirical Study

Table IA.1: Selected target time series from the interest and exchange rates category

abbreviation	description
FEDFUNDS	Effective Federal Funds Rate
TB3MS	3-Month Treasury Bill
TB6MS	6-Month Treasury Bill
GS1	1-Year Treasury Rate
GS5	5-Year Treasury Rate
GS10	10-Year Treasury Rate
AAA	Moody's Seasoned Aaa Corporate Bond Yield
BAA	Moody's Seasoned Baa Corporate Bond Yield
TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
T1YFFM	1-Year Treasury C Minus FEDFUNDS
T5YFFM	5-Year Treasury C Minus FEDFUNDS
T10YFFM	10-Year Treasury C Minus FEDFUNDS
AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
EXJPUSx	Japan / U.S. Foreign Exchange Rate
EXUSUKx	U.S. / U.K. Foreign Exchange Rate
EXCAUSx	Canada / U.S. Foreign Exchange Rate

This table lists the selected target time series from the interest and exchange rates category of the FRED-MD data. Further details are in the appendix of McCracken and Ng (2016). The time series selected here are the ones without missing values from 01/1960 to 12/2020.

Table IA.2: Quarterly-observed target time series from the national income & product accounts and the flow of funds categories

Description
Real Gross National Product
Gross National Product
Real Gross Domestic Product
Gross Domestic Product
Gross Value Added: GDP: Business
Gross Domestic Product: Services
Gross Domestic Product: Goods
Gross Domestic Product: Durable Goods
Gross Domestic Product: Nondurable Goods
Gross Domestic Purchases
Gross Domestic Product: Terms of Trade Index
Gross Domestic Product: Research and Development
All Sectors; Corporate and Foreign Bonds; Liability
All Sectors; Total Debt Securities; Liability
All Sectors; U.S. Government Agency Securities; Liability
Banks in U.S.-Affiliated Areas; Debt Securities; Asset
Domestic Nonfinancial Sectors; Debt Securities; Asset
Domestic Financial Sectors; Corporate Equities; Liability
Domestic Financial Sectors; U.S. Direct Investment Abroad; Asset
Domestic Financial Sectors; Debt Securities; Liability
Domestic Financial Sectors; Debt Securities; Asset
Domestic Financial Sectors; Total Liabilities and Equity
Domestic Financial Sectors; Short-Term Loans Including Security Repurchase Agreements; Liability
Domestic Financial Sectors; Total Assets (Does Not Include Land)
Domestic Financial Sectors; Total Currency and Deposits; Asset
Domestic Financial Sectors; Total Financial Assets
Federal Government; Checkable Deposits and Currency; Asset
Federal Government; Debt Securities; Liability
Federal Government Retirement Funds; Debt Securities; Asset
Federal Government; Total Financial Assets
Federal Government; Total Liabilities
Federal Government; Total Assets (Does Not Include Land)
Federal Government; Treasury Securities; Liability
Federal Government; Total Mortgages; Asset
Finance Companies; Debt Securities; Liability
GSEs and Agency- and GSE-Backed Mortgage Pools; U.S. Government Agency Securities; Liability
Households and Nonprofit Organizations; Corporate Equities; Asset, Market Value Levels
Mutual Funds; Debt Securities; Asset (Market Value)
Nonfinancial Corporate Business; Corporate Equities; Liability, Market Value Levels
Nonfinancial Noncorporate Business; Real Estate at Market Value, Market Value Levels
Rest of the World; Corporate Bonds; Asset, Transactions
Rest of the World; U.S. Corporate Equities; Asset, Transactions
Rest of the World; U.S. Corporate Equities; Asset
Rest of the World; Foreign Direct Investment in U.S.; Asset (Current Cost)
Rest of the World; Total Financial Assets
Rest of the World; Total Liabilities and Equity
Rest of the World; Treasury Securities; Asset
Security Brokers and Dealers; Debt Securities; Asset
State and Local Governments; Debt Securities; Asset
State and Local Governments; U.S. Government Loans; Liability
State and Local Governments; Total Liabilities
State and Local Governments; Municipal Securities; Liability
State and Local Governments; Total Currency and Deposits; Asset
State and Local Governments; Total Financial Assets
State and Local Governments; Total Mortgages; Asset
State and Local Governments; Net Worth (IMA)
State and Local Governments; Trade Payables; Liability, Transactions
State and Local Governments; Treasury Securities, Including SLGS; Asset

This table lists the selected quarterly target time series from the national income & product accounts and the flow of funds categories from the FRED database.

IA.B Additional Simulation Results

Table IA.3: Relative MSE for missing-at-random pattern

Observation Pattern of Y	$\mathcal{M}$	T-PCA	XP_Y	$\text{XP}_{Z^{(1)}}$	SE-PCA
	obs	0.720	0.760	1.088	1.015
$p = 0.3$, $\sigma_{e_x} = 16$	miss	0.727	0.787	1.139	1.132
	all	0.725	0.778	1.124	1.096
	obs	0.396	0.408	0.932	0.559
$p = 0.5$, $\sigma_{e_x} = 16$	miss	0.398	0.414	0.964	0.596
	all	0.397	0.411	0.948	0.578
	obs	0.268	0.272	0.867	0.383
$p = 0.7$, $\sigma_{e_x} = 16$	miss	0.269	0.275	0.891	0.401
	all	0.268	0.273	0.875	0.388
	obs	0.440	0.760	0.440	0.946
$p = 0.3$, $\sigma_{e_x} = 4$	miss	0.430	0.787	0.430	1.049
	all	0.433	0.778	0.433	1.018
	obs	0.256	0.408	0.256	0.538
$p = 0.5$, $\sigma_{e_x} = 4$	miss	0.253	0.414	0.253	0.573
	all	0.255	0.411	0.255	0.556
	obs	0.183	0.272	0.183	0.373
$p = 0.7$, $\sigma_{e_x} = 4$	miss	0.182	0.275	0.182	0.391
	all	0.183	0.273	0.183	0.379
	obs	0.325	0.760	0.397	0.927
$p = 0.3$, $\sigma_{e_x} = 1$	miss	0.314	0.787	0.386	1.028
	all	0.317	0.778	0.390	0.998
	obs	0.184	0.408	0.224	0.530
$p = 0.5$, $\sigma_{e_x} = 1$	miss	0.182	0.414	0.220	0.564
	all	0.183	0.411	0.222	0.547
	obs	0.127	0.272	0.158	0.368
$p = 0.7$, $\sigma_{e_x} = 1$	miss	0.127	0.275	0.157	0.385
	all	0.127	0.273	0.158	0.373

This table reports the relative MSE of different estimation methods for missing-at-random and different parameter choices. We compare T-PCA (our benchmark method), XP_Y (PCA on Y), $\text{XP}_{Z^{(1)}}$ (PCA on concatenated panel) and SE-PCA (separate PCA). We generate a two-factor model as follows: Factors $F_t \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, loadings $(\Lambda_x)_i \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2), (\Lambda_y)_i \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, and errors $(e_x)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_{e_x}), (e_y)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_{e_y})$ with different σ_{e_x} and fixed $\sigma_{e_y} = 4$. The entries of Y are missing independently at random with observation probability p. We set $N_x = N_y = T = 200$ and run 200 simulations for each setup.

Table IA.4: Relative MSE for low-frequency observation pattern

Observation Pattern of Y	$\mathcal{M}$	T-PCA	XP$_Y$	XP$_{Z^{(1)}}$	SE-PCA
	obs	0.450	-	1.096	1.160
$p = 0.3$, $\sigma_{e_x} = 16$	miss	1.058	-	1.346	1.234
	all	0.872	-	1.264	1.210
	obs	0.291	-	0.846	1.059
$p = 0.5$, $\sigma_{e_x} = 16$	miss	1.029	-	1.119	1.104
	all	0.656	-	0.979	1.080
	obs	0.227	-	0.816	1.016
$p = 0.7$, $\sigma_{e_x} = 16$	miss	1.011	-	1.033	1.051
	all	0.460	-	0.879	1.026
	obs	0.431	-	0.537	0.591
$p = 0.3$, $\sigma_{e_x} = 8$	miss	0.565	-	0.611	0.622
	all	0.522	-	0.586	0.611
	obs	0.273	-	0.329	0.480
$p = 0.5$, $\sigma_{e_x} = 8$	miss	0.476	-	0.491	0.502
	all	0.373	-	0.408	0.490
	obs	0.210	-	0.258	0.436
$p = 0.7$, $\sigma_{e_x} = 8$	miss	0.434	-	0.436	0.450
	all	0.276	-	0.310	0.440
	obs	0.374	-	0.372	0.355
$p = 0.3$, $\sigma_{e_x} = 4$	miss	0.363	-	0.366	0.369
	all	0.365	-	0.367	0.364
	obs	0.230	-	0.230	0.243
$p = 0.5$, $\sigma_{e_x} = 4$	miss	0.256	-	0.256	0.250
	all	0.242	-	0.242	0.246
	obs	0.173	-	0.173	0.196
$p = 0.7$, $\sigma_{e_x} = 4$	miss	0.208	-	0.208	0.202
	all	0.183	-	0.183	0.198

This table reports the relative MSE of different estimation methods for a low-frequency observation pattern and different parameter choices. We compare T-PCA (our benchmark method), XP$_Y$ (PCA on Y), XP$_{Z^{(1)}}$ (PCA on concatenated panel) and SE-PCA (separate PCA). We generate a two-factor model as follows: Factors $F_t \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, loadings $(\Lambda_x)_i \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, $(\Lambda_y)_i \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, and errors $(e_x)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{e_x}^2)$ and $(e_y)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{e_y}^2)$ with different σ_{e_x} and fixed $\sigma_{e_y} = 4$. The rows of Y are missing-at-random with observation probability p. We set $N_x = N_y = T = 200$ and run 200 simulations for each setup. Note that in this setting, XP$_Y$ is not applicable and separate-PCA degenerates to PCA using only the panel X.

Table IA.5: Relative MSE for missingness depends on loadings pattern

Observation Pattern of Y	$\mathcal{M}$	T-PCA	XP$_Y$	XP$_{Z^{(1)}}$	SE-PCA
	obs	1.014	1.104	1.161	1.384
$p_1 = 0.2$, $\sigma_{e_x} = 8$	miss	1.097	1.286	1.213	1.738
	all	1.077	1.243	1.200	1.655
	obs	0.466	0.503	0.629	0.675
$p_1 = 0.4$, $\sigma_{e_x} = 8$	miss	0.473	0.522	0.636	0.740
	all	0.470	0.513	0.633	0.712
	obs	0.305	0.322	0.451	0.444
$p_1 = 0.6$, $\sigma_{e_x} = 8$	miss	0.303	0.324	0.452	0.465
	all	0.304	0.323	0.451	0.452
	obs	0.499	0.545	0.614	0.688
$p_1 = 0.2$, $\sigma_{e_x} = 6$	miss	0.546	0.641	0.643	0.866
	all	0.535	0.619	0.636	0.824
	obs	0.245	0.264	0.319	0.354
$p_1 = 0.4$, $\sigma_{e_x} = 6$	miss	0.251	0.277	0.325	0.389
	all	0.249	0.272	0.323	0.374
	obs	0.162	0.172	0.216	0.238
$p_1 = 0.6$, $\sigma_{e_x} = 6$	miss	0.162	0.174	0.217	0.249
	all	0.162	0.173	0.216	0.242
	obs	0.219	0.238	0.262	0.280
$p_1 = 0.2$, $\sigma_{e_x} = 4$	miss	0.252	0.293	0.287	0.356
	all	0.244	0.280	0.281	0.338
	obs	0.111	0.119	0.132	0.150
$p_1 = 0.4$, $\sigma_{e_x} = 4$	miss	0.117	0.129	0.138	0.165
	all	0.114	0.125	0.135	0.158
	obs	0.072	0.077	0.087	0.102
$p_1 = 0.6$, $\sigma_{e_x} = 4$	miss	0.074	0.080	0.089	0.107
	all	0.073	0.078	0.087	0.104

This table reports the relative MSE of different estimation methods when missingness depends on the loadings and for different parameter choices. We compare T-PCA (our benchmark method), XP$_Y$ (PCA on Y), XP$_{Z^{(1)}}$ (PCA on concatenated panel) and SE-PCA (separate PCA). We generate a two-factor model as follows: Factors $F_t \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, loadings $(\Lambda_y)_i \overset{i.i.d.}{\sim} \mathcal{N}(0, I_2)$, $(\Lambda_x)_{i2} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and errors $(e_x)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_{e_x})$ and $(e_y)_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_{e_y})$. We assume $(\Lambda_x)_{i1} = 0$ and $\sigma_{e_y} = 0.5 \cdot \sigma_{e_x}$. Furthermore, we define a unit-specific characteristic $S_i = \mathbb{1}(|(\Lambda_y)_{i,2}| > 0.1)$. The entries of Y are missing independently with observation probability p_1 if $S_i = 1$ and $p_2 = 1$ if $S_i = 0$. We set $N_x = N_y = T = 200$ and run 200 simulations for each setup.

IA.C Proofs

IA.C.1 Proof of Theorem 1

According to Assumptions G3.2, Σ_{Λ_x} and $\Sigma_{\Lambda_y,t}$ are positive semi-definite, and $\Sigma_{\Lambda_x} + \Sigma_{\Lambda_y,t}$ is positive definite. Therefore, when $\gamma = r \cdot N_x/N_y$ with some positive constant r, the weighted second moment matrix

$$\Sigma_{\Lambda,t}^{(\gamma)} = \lim \frac{N_x}{N_x+N_y}\left(\Sigma_{\Lambda_x} + \gamma\frac{N_y}{N_x}\cdot\Sigma_{\Lambda_y,t}\right) = \lim \frac{N_x}{N_x+N_y}\left(\Sigma_{\Lambda_x} + r\cdot\Sigma_{\Lambda_y,t}\right)$$

is positive definite. Similarly, we can show that the matrix $\Sigma_{\Lambda}^{(\gamma)} = \lim \frac{N_x}{N_x+N_y}\left(\Sigma_{\Lambda_x} + r\cdot\Sigma_{\Lambda_y}\right)$ is positive definite. By Assumption G1, Σ_{Λ_x} is not positive definite. When $\gamma \neq r\cdot N_x/N_y$ for any constant r and $N_y/N_x \to 0$, the second moment matrix $\Sigma_{\Lambda,t}^{(\gamma)}$ is not positive definite as well.

In the following, we prove that when $\gamma = r\cdot N_x/N_y$, we can consistently estimate the loadings, factors, and thus the common components of Y. Based on the definition of target-PCA in Section 2.4, we can plug the expression of $\tilde{\Sigma}^{Z^{(\gamma)}}$ into $\frac{1}{N_x+N_y}\tilde{\Sigma}^{Z^{(\gamma)}}\tilde{\Lambda}^{(\gamma)} = \tilde{\Lambda}^{(\gamma)}\tilde{D}^{(\gamma)}$ and obtain the decomposition for the estimated combined loadings $\tilde{\Lambda}^{(\gamma)}$ as

$$\begin{aligned}
\tilde{\Lambda}_i^{(\gamma)} &= \frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\tilde{\Sigma}_{ji}^{Z^{(\gamma)}} \\
&= H_i^{(\gamma)}\Lambda_i^{(\gamma)} + \underbrace{\frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t e_{tj}^{(\gamma)}}_{(a)} \\
&\quad + \underbrace{\frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t e_{ti}^{(\gamma)}}_{(b)} \\
&\quad + \underbrace{\frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}e_{tj}^{(\gamma)}e_{ti}^{(\gamma)}}_{(c)},
\end{aligned}$$

where $H_i^{(\gamma)} = (\tilde{D}^{(\gamma)})^{-1}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_tF_t^\top$. As is shown, the estimated combined loadings are related to the true combined loadings through $\tilde{\Lambda}_i^{(\gamma)} = H_i^{(\gamma)}\Lambda_i^{(\gamma)} + (a) + (b) + (c)$ up to the rotation matrix $H_i^{(\gamma)}$ for index i. Since the estimation of factors requires the same rotation matrix for all $\Lambda_i^{(\gamma)}$, we consider the unified rotation matrix $H^{(\gamma)}$ defined as

$$H^{(\gamma)} = \frac{1}{T(N_x+N_y)}(\tilde{D}^{(\gamma)})^{-1}\tilde{\Lambda}^{(\gamma)\top}\Lambda^{(\gamma)}F^\top F.$$

This yields the decomposition $\tilde{\Lambda}_i^{(\gamma)} = H^{(\gamma)}\Lambda_i^{(\gamma)} + (H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)} + (a) + (b) + (c)$, based on which we derive the consistency result of the estimated loadings.

To simplify notation, we define the following four terms

$$\eta_{ij} = \frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} \Lambda_i^{(\gamma)\top} F_t e_{tj}^{(\gamma)}, \qquad \xi_{ij} = \frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} \Lambda_j^{(\gamma)\top} F_t e_{ti}^{(\gamma)},$$

$$\gamma(i,j) = \frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} \mathbb{E}[e_{ti}^{(\gamma)} e_{tj}^{(\gamma)}], \qquad \zeta_{ij} = \frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} e_{ti}^{(\gamma)} e_{tj}^{(\gamma)} - \gamma(i,j).$$

We omit the superscript (γ) in these four terms to save space. It holds that

$$\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)} = \frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\left(\tilde{\Lambda}_j^{(\gamma)}\eta_{ij} + \tilde{\Lambda}_j^{(\gamma)}\xi_{ij} + \tilde{\Lambda}_j^{(\gamma)}\zeta_{ij} + \tilde{\Lambda}_j^{(\gamma)}\gamma(i,j)\right).$$

We provide bounds for $\eta_{ij}, \xi_{ij}, \zeta_{ij}$ and $\gamma(i,j)$ in the following.

Lemma 1. *Under Assumption G2 and Assumption G3, suppose $N_y/N_x \to c \in [0,\infty)$ and $\gamma = r \cdot N_x/N_y$, then as $T, N_x, N_y \to \infty$, we have*

1. $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} \eta_{ij}^2 = O_p\left(\frac{1}{T}\right)$;
2. $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} \xi_{ij}^2 = O_p\left(\frac{1}{T}\right)$;
3. $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} \zeta_{ij}^2 = O_p\left(\frac{1}{T}\right)$;
4. $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} \gamma^2(i,j) \le \frac{C}{N_y}$.

Proof. 1. By Assumptions G2, G3.2 and G3.4, there is

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\eta_{ij}^2\right] &= \frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left[\Lambda_i^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} F_t e_{tj}^{(\gamma)}\right]^2 \\
&\le \frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left\|\Lambda_i^{(\gamma)}\right\|^2 \cdot \mathbb{E}\left\|\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} F_t e_{tj}^{(\gamma)}\right\|^2 \\
&\le \frac{C}{T}.
\end{aligned}$$

As a result, it holds that $\frac{1}{(N_x+N_y)^2}\sum_{i,j}\eta_{ij}^2 = O_p\left(\frac{1}{T}\right)$.
2. By the same arguments, we can show that $\frac{1}{(N_x+N_y)^2}\sum_{i,j}\xi_{ij}^2 = O_p\left(\frac{1}{T}\right)$.

3. Following from Assumption G3.3(e), it holds that

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\zeta_{ij}^2\right] &= \frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left[\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}\left(e_{ti}^{(\gamma)}e_{tj}^{(\gamma)}-\mathbb{E}(e_{ti}^{(\gamma)}e_{tj}^{(\gamma)})\right)\right]^2 \\
&\le \frac{C}{(N_x+N_y)^2}\left(\sum_{i,j=N_x+1}^{N_x+N_y}\gamma^2\frac{1}{|Q_{ij}^Z|}+\sum_{i=1}^{N_x}\sum_{j=N_x+1}^{N_x+N_y}\gamma\frac{1}{|Q_{ij}^Z|}+\sum_{i,j=1}^{N_x}\frac{1}{|Q_{ij}^Z|}\right) \\
&\le \frac{C}{T}.
\end{aligned}
$$

Therefore, $\frac{1}{(N_x+N_y)^2}\sum_{i,j}\zeta_{ij}^2 = O_p\left(\frac{1}{T}\right)$ as claimed.

4. By definition, we have

$$
\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\gamma^2(i,j) = \frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\frac{1}{|Q_{ij}^Z|^2}\left(\sum_{t\in Q_{ij}^Z}\mathbb{E}\left[e_{ti}^{(\gamma)}e_{tj}^{(\gamma)}\right]\right)^2.
$$

According to Assumption G3.3(c), the RHS of the above equation can be bounded by

$$
\begin{aligned}
\text{RHS} \le&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_y}\gamma^2(\tau_{ij}^{(e_y)})^2+\frac{2}{(N_x+N_y)^2}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\gamma(\tau_{ij}^{(e_y,e_x)})^2+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x}(\tau_{ij}^{(e_x)})^2 \\
\le&\frac{C}{(N_x+N_y)^2}\left(N_y\frac{N_x^2}{N_y^2}+N_x\frac{N_x}{N_y}+N_x\right) \\
\le&\frac{C}{N_y}.
\end{aligned}
$$

□

Lemma 2. *Under Assumption G2 and Assumption G3, suppose $N_y/N_x \to c \in [0,\infty)$ and $\gamma = r\cdot N_x/N_y$. Then as $T, N_x, N_y \to \infty$, we have*

1. $\frac{1}{(N_x+N_y)^2}\tilde{\Lambda}^{(\gamma)\top}\left((\tilde{Z}^{(\gamma)\top}\tilde{Z}^{(\gamma)})\odot\left[\frac{1}{|Q_{ij}^Z|}\right]\right)\tilde{\Lambda}^{(\gamma)} = \tilde{D}^{(\gamma)} \xrightarrow{p} D^{(\gamma)}$;
2. $\frac{1}{(N_x+N_y)^2}\tilde{\Lambda}^{(\gamma)\top}\left(((F\Lambda^{(\gamma)\top})\odot W^Z)^\top((F\Lambda^{(\gamma)\top})\odot W^Z)\odot\left[\frac{1}{|Q_{ij}^Z|}\right]\right)\tilde{\Lambda}^{(\gamma)} \xrightarrow{p} D^{(\gamma)}$;
3. $\frac{1}{(N_x+N_y)^2}\tilde{\Lambda}^{(\gamma)\top}\left(\Lambda^{(\gamma)}\frac{F^\top F}{T}\Lambda^{(\gamma)\top}\right)\tilde{\Lambda}^{(\gamma)} \xrightarrow{p} D^{(\gamma)}$;

where $\tilde{Z}^{(\gamma)} = Z^{(\gamma)}\odot W^Z$ and $D^{(\gamma)} = diag\left(d_1^{(\gamma)},\cdots,d_k^{(\gamma)}\right)$ are the eigenvalues of $\Sigma_F\Sigma_\Lambda^{(\gamma)}$.

Proof. Let $\lambda \in \mathbb{R}^{(N_x+N_y)\times 1}$. Define $\Gamma = \{\lambda | \lambda^\top \lambda = N_x + N_y\}$, and let

$$R(\lambda) = \frac{1}{(N_x+N_y)^2}\lambda^\top \left((\tilde{Z}^{(\gamma)\top}\tilde{Z}^{(\gamma)}) \odot \left[\frac{1}{|Q^Z_{ij}|}\right]\right)\lambda,$$

$$\tilde{R}(\lambda) = \frac{1}{(N_x+N_y)^2}\lambda^{(\gamma)\top} \left(((F\Lambda^{(\gamma)\top}) \odot W^Z)^\top ((F\Lambda^{(\gamma)\top}) \odot W^Z) \odot \left[\frac{1}{|Q^Z_{ij}|}\right]\right)\lambda,$$

$$R^*(\lambda) = \frac{1}{(N_x+N_y)^2}\lambda^\top \left(\Lambda^{(\gamma)}\frac{F^\top F}{T}\Lambda^{(\gamma)\top}\right)\lambda.$$

First of all, we prove that as $T, N_x, N_y \to \infty$,

$$(1) = \sup_{\lambda\in\Gamma} \frac{1}{(N_x+N_y)^2}\left|\lambda^\top \left(((e^{(\gamma)})^\top \odot (W^Z)^\top)(e^{(\gamma)} \odot W^Z) \odot \left[\frac{1}{|Q^Z_{ij}|}\right]\right)\lambda\right| \xrightarrow{p} 0,$$

$$(2) = \sup_{\lambda\in\Gamma} \frac{1}{(N_x+N_y)^2}\left|\lambda^\top \left(((e^{(\gamma)})^\top \odot (W^Z)^\top))(F\Lambda^{(\gamma)\top} \odot W^Z) \odot \left[\frac{1}{|Q^Z_{ij}|}\right]\right)\lambda\right| \xrightarrow{p} 0.$$

Observe that

$$\begin{aligned}
(1) &= \sup_{\lambda\in\Gamma} \frac{1}{(N_x+N_y)^2}\left|\sum_{i,j=1}^{N_x+N_y}\lambda_i\lambda_j\frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}} e^{(\gamma)}_{ti}e^{(\gamma)}_{tj}\right| \\
&\le \sup_{\lambda\in\Gamma} \underbrace{\left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\lambda_i^2\lambda_j^2\right)^{1/2}}_{=1} \cdot \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\left(\frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}} e^{(\gamma)}_{ti}e^{(\gamma)}_{tj}\right)^2\right)^{1/2} \\
&= \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}(\zeta_{ij}+\gamma(i,j))^2\right)^{1/2}.
\end{aligned}$$

According to Lemma 1, $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\zeta_{ij}^2 = o_p(1)$ and $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\gamma^2(i,j) = o(1)$. As a result, $(1) \xrightarrow{p} 0$. Consider (2), we have

$$\begin{aligned}
(2) &= \sup_{\lambda\in\Gamma} \frac{1}{(N_x+N_y)^2}\left|\sum_{i,j=1}^{N_x+N_y}\lambda_i\lambda_j\frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}} e^{(\gamma)}_{ti}F_t^\top\Lambda^{(\gamma)}_j\right| \\
&\le \sup_{\lambda\in\Gamma} \underbrace{\left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\lambda_i^2\lambda_j^2\right)^{1/2}}_{=1} \cdot \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\xi_{ij}^2\right)^{1/2} = o_p(1)
\end{aligned}$$

following from Lemma 1. Combining these two terms, it holds that

$$
\begin{aligned}
\sup_{\lambda\in\Gamma} |R(\lambda) - \tilde{R}(\lambda)| \leq & \sup_{\lambda\in\Gamma} \frac{1}{(N_x+N_y)^2} \left| \lambda^\top \left(((e^{(\gamma)})^\top \odot (W^Z)^\top)(e^{(\gamma)} \odot W^Z) \odot \left[\frac{1}{|Q_{ij}^Z|} \right] \right) \lambda \right| \\
& + \sup_{\lambda\in\Gamma} \frac{2}{(N_x+N_y)^2} \left| \lambda^\top \left(((e^{(\gamma)})^\top \odot (W^Z)^\top))(F\Lambda^\top \odot W^Z) \odot \left[\frac{1}{|Q_{ij}^Z|} \right] \right) \lambda \right| \\
& \xrightarrow{p} 0.
\end{aligned}
$$

Furthermore, for any $\lambda \in \Gamma$,

$$
\begin{aligned}
\tilde{R}(\lambda) - R^*(\lambda) &= \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \lambda_i \lambda_j \Lambda_i^{(\gamma)\top} \underbrace{\left(\frac{1}{|Q_{ij}^Z|} \sum_{t\in Q_{ij}^Z} F_t F_t^\top - \frac{1}{T} F^\top F \right)}_{\Delta_{F,ij}} \Lambda_j^{(\gamma)} \\
&\leq \underbrace{\left(\frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \lambda_i^2 \lambda_j^2 \right)^{1/2}}_{=1} \cdot \left(\frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \left(\Lambda_i^{(\gamma)\top} \Delta_{F,ij} \Lambda_j^{(\gamma)} \right)^2 \right)^{1/2}.
\end{aligned}
$$

By Assumptions G3.1 and G3.2, we have

$$
\begin{aligned}
& \mathbb{E}\left[\frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \left(\Lambda_i^{(\gamma)\top} \Delta_{F,ij} \Lambda_j^{(\gamma)} \right)^2 \right] \\
\leq & \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \mathbb{E}\left[\left\| \Lambda_i^{(\gamma)} \right\|^2 \left\| \Lambda_j^{(\gamma)} \right\|^2 \right] \cdot \mathbb{E}\|\Delta_{F,ij}\|^2 \ \leq \frac{C}{T}.
\end{aligned}
$$

Therefore, $\sup_{\lambda\in\Gamma} |\tilde{R}(\lambda) - R^*(\lambda)| \xrightarrow{p} 0$. We can also derive $\sup_{\lambda\in\Gamma} |R(\lambda) - R^*(\lambda)| \xrightarrow{p} 0$ by the decomposition $R(\lambda) - R^*(\lambda) = R(\lambda) - \tilde{R}(\lambda) + \tilde{R}(\lambda) - R^*(\lambda)$. As a result, $|\sup_{\lambda\in\Gamma} R(\lambda) - \sup_{\lambda\in\Gamma} R^*(\lambda)| \leq \sup_{\lambda\in\Gamma} |R(\lambda) - R^*(\lambda)| \xrightarrow{p} 0$. Since $\sup_{\lambda\in\Gamma} R^*(\lambda) \xrightarrow{p} d_1^{(\gamma)}$ where $d_1^{(\gamma)}$ is the largest eigenvalue of $\Sigma_F \Sigma_\Lambda^{(\gamma)}$, we have $\sup_{\lambda\in\Gamma} R(\lambda) \xrightarrow{p} d_1^{(\gamma)}$. By definition, $\tilde{\Lambda}_1^{(\gamma)} = \arg\sup_{\lambda\in\Gamma} R(\lambda)$, thus $R(\tilde{\Lambda}_1^{(\gamma)}) = \tilde{d}_1^{(\gamma)} \xrightarrow{p} d_1^{(\gamma)}$ and $\tilde{R}(\tilde{\Lambda}_1^{(\gamma)}), R^*(\tilde{\Lambda}_1^{(\gamma)}) \xrightarrow{p} d_1^{(\gamma)}$. We repeat this procedure sequentially using the orthonormal subspace of Γ and finally complete our proof. $\square$

Lemma 3. *Under Assumption G2 and Assumption G3, suppose $N_y/N_x \to c \in [0, \infty)$ and $\gamma = r \cdot N_x/N_y$. For $T, N_x, N_y \to \infty$ it holds that*

1. $\frac{1}{N_x+N_y} \Lambda^{(\gamma)\top} \tilde{\Lambda}^{(\gamma)} \xrightarrow{p} Q^{(\gamma)} = \Sigma_F^{-1/2} \Upsilon^{(\gamma)} (D^{(\gamma)})^{1/2}$, *where the diagonal entries of $D^{(\gamma)}$ are eigenvalues of $\Sigma_F^{1/2} \Sigma_\Lambda^{(\gamma)} \Sigma_F^{1/2}$, and $\Upsilon^{(\gamma)}$ is the corresponding eigenvector matrix such that $\Upsilon^{(\gamma)\top} \Upsilon^{(\gamma)} = I$;*
2. $H^{(\gamma)} \xrightarrow{p} (Q^{(\gamma)})^{-1}$, *where $H^{(\gamma)} = T^{-1}(N_x+N_y)^{-1}(\tilde{D}^{(\gamma)})^{-1} \tilde{\Lambda}^{(\gamma)\top} \Lambda^{(\gamma)} F^\top F$.*

Proof. Left-multiplying $\frac{1}{N_x+N_y}\tilde{\Sigma}^{Z^{(\gamma)}}\tilde{\Lambda}^{(\gamma)} = \tilde{\Lambda}^{(\gamma)}\tilde{D}^{(\gamma)}$ on both sides by $\frac{1}{N_x+N_y}(\frac{F^\top F}{T})^{1/2}\Lambda^{(\gamma)\top}$, we have

$$\left(\frac{F^\top F}{T}\right)^{1/2}\frac{\Lambda^{(\gamma)\top}}{N_x+N_y}\tilde{\Sigma}^{Z^{(\gamma)}}\frac{\tilde{\Lambda}^{(\gamma)}}{N_x+N_y} = \left(\frac{F^\top F}{T}\right)^{1/2}\left(\frac{\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}}{N_x+N_y}\right)\tilde{D}^{(\gamma)}.$$

This can be rewritten as

$$\left(\frac{F^\top F}{T}\right)^{1/2}\frac{\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}}{N_x+N_y}\left(\frac{F^\top F}{T}\right)\frac{\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}}{N_x+N_y} + d^{(\gamma)}_{T,N_yN_x} = \left(\frac{F^\top F}{T}\right)^{1/2}\left(\frac{\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}}{N_x+N_y}\right)\tilde{D}^{(\gamma)},$$

where

$$d^{(\gamma)}_{T,N_xN_y} = \left(\frac{F^\top F}{T}\right)^{1/2}\frac{\Lambda^{(\gamma)\top}}{N_x+N_y}\left(\tilde{\Sigma}^{Z^{(\gamma)}} - \Lambda^{(\gamma)}\frac{F^\top F}{T}\Lambda^{(\gamma)\top}\right)\frac{\tilde{\Lambda}^{(\gamma)}}{N_x+N_y}.$$

Plugging the expansion of $\tilde{\Sigma}^{Z^{(\gamma)}}_{ij}$ into $d^{(\gamma)}_{T,N_xN_y}$, we obtain for any i and j,

$$\begin{aligned}\left(\tilde{\Sigma}^{Z^{(\gamma)}} - \Lambda^{(\gamma)}\frac{F^\top F}{T}\Lambda^{(\gamma)\top}\right)_{ij} =& \Lambda_i^{(\gamma)\top}\underbrace{\left[\frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}}F_tF_t^\top - \frac{1}{T}F^\top F\right]}_{\Delta_{F,ij}}\Lambda_j^{(\gamma)} + \frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}}F_t^\top\Lambda_i^{(\gamma)}e_{tj}^{(\gamma)} \\ &+ \frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}}F_t^\top\Lambda_j^{(\gamma)}e_{ti}^{(\gamma)} + \frac{1}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}}e_{ti}^{(\gamma)}e_{tj}^{(\gamma)} \\ =& \Lambda_i^{(\gamma)\top}\Delta_{F,ij}\Lambda_j^{(\gamma)} + \eta_{ij} + \xi_{ij} + \zeta_{ij} + \gamma(i,j).\end{aligned}$$

Each (r,s)-th entry of the component $\frac{\Lambda^{(\gamma)\top}}{N_x+N_y}\left(\tilde{\Sigma}^{Z^{(\gamma)}} - \Lambda^{(\gamma)}\frac{F^\top F}{T}\Lambda^{(\gamma)\top}\right)\frac{\tilde{\Lambda}^{(\gamma)}}{N_x+N_y}$ of $d^{(\gamma)}_{T,N_xN_y}$ can be bounded by

$$\begin{aligned}&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_{ir}^{(\gamma)}\tilde{\Lambda}_{js}^{(\gamma)}\left(\Lambda_i^{(\gamma)\top}\Delta_{F,ij}\Lambda_j^{(\gamma)} + \eta_{ij} + \xi_{ij} + \zeta_{ij} + \gamma(i,j)\right) \\ \le& \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}(\Lambda_{ir}^{(\gamma)})^2(\tilde{\Lambda}_{js}^{(\gamma)})^2\right)^{1/2}\cdot\left[\left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\left(\Lambda_i^{(\gamma)\top}\Delta_{F,ij}\Lambda_j^{(\gamma)}\right)^2\right)^{1/2}\right. \\ &+\left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\eta_{ij}^2\right)^{1/2} + \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\xi_{ij}^2\right)^{1/2} \\ &\left.+\left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\zeta_{ij}^2\right)^{1/2} + \left(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\gamma^2(i,j)\right)^{1/2}\right].\end{aligned}$$

Based on Lemma 1 and the proof of Lemma 2, we can deduce that $d^{(\gamma)}_{T,N_xN_y} = o_p(1)$. Let

$$B^{(\gamma)}_{T,N_xN_y} = \left(\frac{F^\top F}{T}\right)^{1/2}\frac{\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}}{N_x+N_y}\left(\frac{F^\top F}{T}\right)^{1/2},$$

and

$$R^{(\gamma)}_{T,N_xN_y} = \left(\frac{F^\top F}{T}\right)^{1/2} \frac{\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}}{N_x+N_y}.$$

It holds that

$$(B^{(\gamma)}_{T,N_xN_y} + d^{(\gamma)}_{T,N_xN_y}(R^{(\gamma)}_{T,N_xN_y})^{-1})R^{(\gamma)}_{T,N_xN_y} = R^{(\gamma)}_{T,N_xN_y}\tilde{D}^{(\gamma)}.$$

Note that $B^{(\gamma)}_{T,N_xN_y} + d^{(\gamma)}_{T,N_xN_y}(R^{(\gamma)}_{T,N_xN_y})^{-1} \xrightarrow{p} B^{(\gamma)} = \Sigma_F^{1/2}\Sigma_\Lambda^{(\gamma)}\Sigma_F^{1/2}$ by Assumption G3 and the fact that $d^{(\gamma)}_{T,N_xN_y} = o_p(1)$. Because the eigenvalues of $B^{(\gamma)}$ are distinct, the eigenvalues of $B^{(\gamma)}_{T,N_xN_y} + d^{(\gamma)}_{T,N_xN_y}(R^{(\gamma)}_{T,N_xN_y})^{-1}$ will also be distinct for large T, N_x and N_y by the continuity of eigenvalues. With similar arguments as the proof of Proposition 1 in Bai (2003), it holds that

$$\frac{\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}}{N_x+N_y} \xrightarrow{p} \Sigma_F^{-1/2}\Upsilon^{(\gamma)}(D^{(\gamma)})^{1/2} =: Q^{(\gamma)},$$

where $D^{(\gamma)}$ is the diagonal matrix consisting of eigenvalues of $\Sigma_F^{1/2}\Sigma_\Lambda^{(\gamma)}\Sigma_F^{1/2}$, and $\Upsilon^{(\gamma)}$ is the corresponding eigenvector matrix such that $\Upsilon^{(\gamma)\top}\Upsilon^{(\gamma)} = I$. Note that the eigenvalues of $\Sigma_F^{1/2}\Sigma_\Lambda^{(\gamma)}\Sigma_F^{1/2}$ are the same as the eigenvalues of $\Sigma_F\Sigma_\Lambda^{(\gamma)}$. Furthermore, it holds that

$$\begin{aligned} H^{(\gamma)} &= T^{-1}(N_x+N_y)^{-1}(\tilde{D}^{(\gamma)})^{-1}\tilde{\Lambda}^{(\gamma)\top}\Lambda^{(\gamma)}F^\top F \\ &\xrightarrow{p} (D^{(\gamma)})^{-1}(D^{(\gamma)})^{1/2}\Upsilon^{(\gamma)\top}\Sigma_F^{-1/2}\Sigma_F = (D^{(\gamma)})^{-1/2}\Upsilon^{(\gamma)\top}\Sigma_F^{1/2} = (Q^{(\gamma)})^{-1}. \end{aligned}$$

□

Proof. Proof of Theorem 1.1:

(1) Proof of $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 = O_p(\frac{1}{\delta_{N_y,T}})$.

Observe that

$$\begin{aligned} \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \le{}& \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \\ &+ \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}\right\|^2. \end{aligned}$$

We bound the two terms on the RHS respectively in the following. For the first term, we have the decomposition

$$\begin{aligned} &\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)} \\ &= \frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\left(\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\eta_{ij} + \sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\xi_{ij} + \sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij} + \sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\gamma(i,j)\right), \end{aligned}$$

which follows that

$$
\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2\\
\leq& 4\left\|(\tilde{D}^{(\gamma)})^{-1}\right\|^2\cdot\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\frac{1}{(N_x+N_y)^2}\left(\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\eta_{ij}\right\|^2+\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\xi_{ij}\right\|^2\right.\\
&\left.+\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij}\right\|^2+\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\gamma(i,j)\right\|^2\right).
\end{aligned}
$$

Each term $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\frac{1}{(N_x+N_y)^2}\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}\right\|^2$ with $\phi_{ij}=\eta_{ij},\xi_{ij},\zeta_{ij},\gamma(i,j)$ on the RHS can be bounded by

$$
\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\frac{1}{(N_x+N_y)^2}\left\|\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}\right\|^2\leq\underbrace{\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\|\tilde{\Lambda}_j^{(\gamma)}\|^2}_{O_p(1)}\cdot\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\phi_{ij}^2.
$$

By Lemma 1 and $\left\|(\tilde{D}^{(\gamma)})^{-1}\right\|=O_p(1)$ proved in Lemma 2, we conclude that

$$
\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2=O_p\big(\frac{1}{T}\big)+O_p\big(\frac{1}{N_y}\big)=O_p\big(\frac{1}{\delta_{N_y,T}}\big),
$$

where $\delta_{N_y,T}=\min(N_y,T)$.

The second part $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|(H_i^{(\gamma)}-H^{(\gamma)})\Lambda_i^{(\gamma)}\right\|^2$ can be bounded by

$$
\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|(H_i^{(\gamma)}-H^{(\gamma)})\Lambda_i^{(\gamma)}\right\|^2\\
=&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\left(\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_tF_t^\top-\frac{1}{T}\sum_{t=1}^{T}F_tF_t^\top\right)\Lambda_i^{(\gamma)}\right\|^2\\
\leq&\underbrace{\left\|(\tilde{D}^{(\gamma)})^{-1}\right\|^2}_{O_p(1)}\cdot\underbrace{\left(\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\|\tilde{\Lambda}_j^{(\gamma)}\|^2\right)}_{O_p(1)}\cdot\Delta^{(\gamma)},
\end{aligned}
$$

where $\Delta^{(\gamma)}:=\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\|\Lambda_i^{(\gamma)}\|^2\|\Lambda_j^{(\gamma)}\|^2\cdot\left\|\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_tF_t^\top-\frac{1}{T}\sum_{t=1}^{T}F_tF_t^\top\right\|^2$. Since there is $\mathbb{E}\left[\Delta^{(\gamma)}\right]\leq\frac{C}{T}$, we conclude that $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|(H_i^{(\gamma)}-H^{(\gamma)})\Lambda_i^{(\gamma)}\right\|^2=O_p\big(\frac{1}{T}\big)$.

Combining these two parts, we have

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2=O_p(\frac{1}{T})+O_p(\frac{1}{\delta_{N_y,T}})=O_p(\frac{1}{\delta_{N_y,T}})$$

as claimed.

(2) Proof of $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-(H^{(\gamma)\top})^{-1}F_t\right\|^2=O_p(\frac{1}{\delta_{N_y,T}})$.

We derive the estimated factors $\tilde{F}_t$ by regressing the observed $Z_{ti}^{(\gamma)}$ on $\tilde{\Lambda}_i^{(\gamma)}$ as

$$\tilde{F}_t=\left(\sum_{i=1}^{N_x+N_y}W_{ti}^Z\tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top}\right)^{-1}\left(\sum_{i=1}^{N_x+N_y}W_{ti}^Z Z_{ti}^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)}\right),\quad t=1,\cdots,T.$$

For analysis, we define an auxiliary $\tilde{F}_t^*$ as

$$\tilde{F}_t^*:=\left(\sum_{i=1}^{N_x+N_y}W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right)^{-1}\left(\sum_{i=1}^{N_x+N_y}W_{ti}^Z Z_{ti}^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)}\right).$$

We have the decomposition

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-(H^{(\gamma)\top})^{-1}F_t\right\|^2\leq\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^*-(H^{(\gamma)\top})^{-1}F_t\right\|^2+\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^*-\tilde{F}_t\right\|^2.$$

We bound the two terms on the RHS in the following. For the first term, $\tilde{F}_t^*$ can be decomposed as

$$\begin{aligned}\tilde{F}_t^*=&(H^{(\gamma)\top})^{-1}F_t+(H^{(\gamma)\top})^{-1}(\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right)\\&+(H^{(\gamma)\top})^{-1}(\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}(H^{(\gamma)})^{-1}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})Z_{ti}^{(\gamma)}\right),\end{aligned}$$

where $\tilde{\Sigma}_{\Lambda,t}^{(\gamma)}=\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\xrightarrow{p}\Sigma_{\Lambda,t}^{(\gamma)}\succ 0$. Therefore, it holds that

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^*-(H^{(\gamma)\top})^{-1}F_t\right\|^2\leq&\frac{C}{T}\sum_{t=1}^{T}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^2\\&+\frac{C}{T}\sum_{t=1}^{T}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})Z_{ti}^{(\gamma)}\right\|^2.\end{aligned}$$

By Assumption G3.3, there is

$$\mathbb{E}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)}\right\|^2 = \sum_{r=1}^{k}\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left[W_{ti}^Z W_{tj}^Z \Lambda_{ir}^{(\gamma)}\Lambda_{jr}^{(\gamma)}\right]\mathbb{E}\left[e_{ti}^{(\gamma)} e_{tj}^{(\gamma)}\right] \leq \frac{C}{N_y},$$

and based on the first step,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})Z_{ti}^{(\gamma)}\right\|^2 \\
\leq &\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left(Z_{ti}^{(\gamma)}\right)^2 \cdot \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 = O_p(\frac{1}{\delta_{N_y,T}}).
\end{aligned}$$

Thus, $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^* - (H^{(\gamma)\top})^{-1}F_t\right\|^2 = O_p(\frac{1}{\delta_{N_y,T}})$.

Consider the second term $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^* - \tilde{F}_t\right\|^2$. The difference $\tilde{F}_t^* - \tilde{F}_t$ can be expanded as

$$\begin{aligned}
\tilde{F}_t^* - \tilde{F}_t = &\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top}\right)^{-1} \cdot \\
&\left[\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right]\tilde{F}_t^*.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right\| \\
\leq &\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left(\left\|\tilde{\Lambda}_i^{(\gamma)}\right\| + \left\|H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|\right)\left\|\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right\| \\
\leq &\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left(\left\|\tilde{\Lambda}_i^{(\gamma)}\right\| + \left\|H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|\right)^2\right)^{1/2} \cdot \left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2\right)^{1/2}.
\end{aligned}$$

By $\frac{1}{N_x+N_y}\tilde{\Lambda}^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)} = I_k$ and the consistency results of loadings, we have

$$\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right\| = O_p(\frac{1}{\sqrt{\delta_{N_y,T}}}).$$

Therefore, $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} \xrightarrow{p} (Q^{(\gamma)})^{-1}\Sigma_{\Lambda,t}^{(\gamma)}((Q^{(\gamma)})^{-1})^\top \succ 0$. To simplify notation, we let $\Delta_{\Lambda,t} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}$. Similar to the first term, the second term can be bounded by

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^* - \tilde{F}_t\right\|^2 \le \frac{C}{T}\sum_{t=1}^{T}\|\Delta_{\Lambda,t}\|^2\left\|\tilde{F}_t^*\right\|^2.$$

For $\Delta_{\Lambda,t}$, we have

$$\begin{aligned}
\Delta_{\Lambda,t} =& \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top} \\
=& \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left[H^{(\gamma)}W_{ti}^Z\Lambda_i^{(\gamma)}(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})^\top + W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right] \\
&+ \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left[W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}(H_i^{(\gamma)}\Lambda_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})^\top + W_{ti}^Z(H_i^{(\gamma)}\Lambda_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})(H^{(\gamma)}\Lambda_i^{(\gamma)})^\top\right] \\
&+ \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})^\top,
\end{aligned}$$

which implies that

$$\begin{aligned}
\|\Delta_{\Lambda,t}\|^2 \le& 10\cdot\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} H^{(\gamma)}W_{ti}^Z\Lambda_i^{(\gamma)}(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})^\top\right\|^2 \\
&+ 10\cdot\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}(H_i^{(\gamma)} - H^{(\gamma)})^\top\right\|^2 \\
&+ 5\cdot\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})^\top\right\|^2.
\end{aligned}$$

According to the proof of the first term, we have

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^*\right\|^2 \le \frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^* - (H^{(\gamma)\top})^{-1}F_t\right\|^2 + \frac{1}{T}\sum_{t=1}^{T}\left\|(H^{(\gamma)\top})^{-1}F_t\right\|^2 = O_p(1).$$

Based on the proof of the consistency results of loadings,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^{*}-\tilde{F}_t\right\|^2 \leq \frac{C}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t^{*}\right\|^2 \Bigg[& \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \cdot \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \\
& + \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \cdot \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\left(H_i^{(\gamma)}-H^{(\gamma)}\right)\Lambda_i^{(\gamma)}\right\|^2 \\
& + \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \cdot \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)}\right\|^2 \Bigg] \\
= O_p(\frac{1}{\delta_{N_y,T}}).
\end{aligned}$$

Combining this with the first term, we complete our proof. □

IA.C.2 Proof of Theorem 2

IA.C.2.1 Proof of Theorem 2.1

Lemma 4. *Under Assumptions G2 and G3, suppose that $N_y/N_x \to c \in [0,\infty)$ and $\gamma = r\cdot N_x/N_y$, it holds that*

$$(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)} = \left(\frac{1}{T}F^\top F\right)^{-1}\left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1} + o_p(1).$$

Proof. Based on the definition of $H^{(\gamma)}$ and the fact that $H^{(\gamma)} = (Q^{(\gamma)})^{-1}+o_p(1) = \left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}\right)^{-1} + o_p(1)$, we obtain

$$\begin{aligned}
(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)} &= \left(\frac{1}{T}F^\top F\right)^{-1}\left(\frac{1}{N_x+N_y}\tilde{\Lambda}^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1}\tilde{D}^{(\gamma)}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)} \\
&= \left(\frac{1}{T}F^\top F\right)^{-1}H^{(\gamma)\top}H^{(\gamma)} + o_p(1).
\end{aligned}$$

Additionally, by Theorem 1,

$$\begin{aligned}
H^{(\gamma)\top}H^{(\gamma)} &= \left(\frac{1}{N_x+N_y}\tilde{\Lambda}^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1}\left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\tilde{\Lambda}^{(\gamma)}\right)^{-1} + o_p(1) \\
&= \left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1}(H^{(\gamma)\top}H^{(\gamma)})^{-1}\left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1} + o_p(1).
\end{aligned}$$

Left-multiplying both sides by $\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}$, we have

$$H^{(\gamma)\top}H^{(\gamma)} = \left(\frac{1}{N_x+N_y}\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}\right)^{-1} + o_p(1).$$

Plugging this into $(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}$, we complete our proof. $\square$

Lemma 5. *Suppose $N_y/N_x \to c \in [0,\infty)$ and let $\gamma = r \cdot N_x/N_y$. Under Assumptions G2, G3 and G4, we have for any $i = N_x+1,\cdots,N_x+N_y$,*

1. $\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\eta_{ij} = O_p\big(\frac{1}{\sqrt{T\delta_{N_y,T}}}\big)$;
2. $\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\xi_{ij} = O_p\big(\frac{1}{\sqrt{T}}\big)$;
3. $\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\gamma(i,j) = O_p\big(\frac{1}{\sqrt{N_y\delta_{N_y,T}}}\big)$;
4. $\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij} = O_p\big(\frac{1}{\sqrt{T\delta_{N_y,T}}}\big)$.

Proof. Observe that we can decompose each term $\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}$ with $\phi_{ij} = \eta_{ij}, \xi_{ij}, \gamma(i,j)$ and ζ_{ij} as

$$\begin{aligned}\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij} =& \frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}(\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)})\phi_{ij} \\ &+ \frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}H^{(\gamma)}\Lambda_j^{(\gamma)}\phi_{ij},\end{aligned}$$

where the first term $\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}(\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)})\phi_{ij}$ can be bounded by

$$\begin{aligned}&\left\|\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}(\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)})\phi_{ij}\right\| \\ \leq& \underbrace{\left(\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\|\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)}\|^2\right)^{1/2}}_{O_p\left(\frac{1}{\sqrt{\delta_{N_y,T}}}\right) \text{ by Theorem 1}} \cdot \left(\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\phi_{ij}^2\right)^{1/2}.\end{aligned}$$

We analyze $\frac{1}{\sqrt{\gamma}}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}$ with $\phi_{ij} = \eta_{ij}, \xi_{ij}, \gamma(i,j), \zeta_{ij}$ respectively in the following.

1. $\phi_{ij} = \eta_{ij}$: For $i = N_x+1,\cdots,N_x+N_y$, it holds that

$$\mathbb{E}\left[\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\eta_{ij}^2\right] \leq \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\mathbb{E}\left\|\Lambda_i^{(\gamma)}\right\|^2 \cdot \mathbb{E}\left\|\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t e_{tj}^{(\gamma)}\right\|^2 \leq \frac{C\gamma}{T},$$

where the last inequality follows from Assumption G3.2 and Assumption G3.4. Additionally, by Assumption G4.1, there is

$$\begin{aligned}\left\| \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\eta_{ij}\right\| &= \left\| \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t e_{tj}^{(\gamma)}\right\| \\ &\leq \left\| \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t^\top e_{tj}^{(\gamma)}\right\| \cdot \left\|\Lambda_i^{(\gamma)}\right\| = O_p\big(\frac{\sqrt{\gamma}}{\sqrt{N_y T}}\big).\end{aligned}$$

Combining these parts, we have $\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\phi_{ij} = O_p\big(\frac{1}{\sqrt{T\delta_{N_y,T}}}\big)$.

2. $\phi_{ij}=\xi_{ij}$: Similar to the previous part, we can show that $\mathbb{E}\left[\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\xi_{ij}^2\right] \leq \frac{C\gamma}{T}$. According to Assumption G4.6, we have

$$\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\xi_{ij} = \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t e_{ti}^{(\gamma)} = O_p\big(\frac{\sqrt{\gamma}}{\sqrt{T}}\big).$$

Therefore, $\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\xi_{ij} = O_p\big(\frac{1}{\sqrt{T}}\big)$ as claimed.

3. $\phi_{ij}=\gamma(i,j)$: Based on Assumption G3.3(c), for any $i=N_x+1,\cdots,N_x+N_y$,

$$\begin{aligned}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\gamma^2(i,j) &= \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left(\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}\mathbb{E}\left[e_{ti}^{(\gamma)}e_{tj}^{(\gamma)}\right]\right)^2 \\ &\leq \frac{1}{N_x+N_y}\sum_{j=1}^{N_y}\gamma^2(\tau_{(i-N_x),j}^{(e_y)})^2 + \frac{1}{N_x+N_y}\sum_{j=1}^{N_x}\gamma(\tau_{(i-N_x),j}^{(e_y,e_x)})^2 \leq C\frac{\gamma}{N_y}.\end{aligned}$$

Moreover, there is

$$\begin{aligned}\mathbb{E}\left\|\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\gamma(i,j)\right\| &\leq \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\mathbb{E}\left\|\Lambda_j^{(\gamma)}\right\|\cdot|\gamma(i,j)| \\ &\leq \frac{C}{N_x+N_y}\sum_{j=1}^{N_y}\gamma^{3/2}\tau_{(i-N_x),j}^{(e_y)} + \frac{C}{N_x+N_y}\sum_{j=1}^{N_x}\gamma^{1/2}\tau_{(i-N_x),j}^{(e_y,e_x)} \\ &\leq C\frac{\sqrt{\gamma}}{N_y}.\end{aligned}$$

As a result, we conclude that $\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\gamma(i,j) = O_p\big(\frac{1}{\sqrt{N_y\delta_{N_y,T}}}\big)$.

4. $\phi_{ij} = \zeta_{ij}$: According to Assumption G3.3(e), for any $i = N_x + 1, \cdots, N_x + N_y$ it holds that

$$\mathbb{E}\left[\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\zeta_{ij}^2\right] = \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\mathbb{E}\left[\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}[e_{ti}^{(\gamma)}e_{tj}^{(\gamma)} - \mathbb{E}(e_{ti}^{(\gamma)}e_{tj}^{(\gamma)})]\right]^2 \leq \frac{C\gamma}{T},$$

and by Assumption G4.3,

$$\begin{aligned}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\zeta_{ij} &= \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\left(\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}[e_{ti}^{(\gamma)}e_{tj}^{(\gamma)} - \mathbb{E}(e_{ti}^{(\gamma)}e_{tj}^{(\gamma)})]\right)\\ &= O_p(\frac{\sqrt{\gamma}}{\sqrt{N_yT}}).\end{aligned}$$

Combining these terms, we have $\frac{1}{\sqrt{\gamma}} \cdot \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij} = O_p(\frac{1}{\sqrt{T\delta_{N_y,T}}})$. We complete our proof. □

Proof. Proof of Theorem 2.1:

For any $i = 1, \cdots, N_y$, we have the decomposition

$$\sqrt{T}\left((\tilde{\Lambda}_y)_i - H^{(\gamma)}(\Lambda_y)_i\right) = \sqrt{T}\left((\tilde{\Lambda}_y)_i - H_{i+N_x}^{(\gamma)}(\Lambda_y)_i\right) + \sqrt{T}\left(H_{i+N_x}^{(\gamma)} - H^{(\gamma)}\right)(\Lambda_y)_i.$$

For simplicity, we denote $i' = i + N_x$.

Step 1 – For the first term $\sqrt{T}\left((\tilde{\Lambda}_y)_i - H_{i+N_x}^{(\gamma)}(\Lambda_y)_i\right)$, observe that

$$\begin{aligned}(\tilde{\Lambda}_y)_i - H_{i+N_x}^{(\gamma)}(\Lambda_y)_i &= \frac{1}{\sqrt{\gamma}}\left(\tilde{\Lambda}_{i'}^{(\gamma)} - H_{i'}^{(\gamma)}\Lambda_{i'}^{(\gamma)}\right)\\ &= (\tilde{D}^{(\gamma)})^{-1}\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left(\tilde{\Lambda}_j^{(\gamma)}\eta_{i'j} + \tilde{\Lambda}_j^{(\gamma)}\xi_{i'j} + \tilde{\Lambda}_j^{(\gamma)}\zeta_{i'j} + \tilde{\Lambda}_j^{(\gamma)}\gamma(i',j)\right).\end{aligned}$$

According to Lemma 5, when $\sqrt{T}/N_y \to 0$, there is

$$\begin{aligned}\sqrt{T}\left((\tilde{\Lambda}_y)_i - H_{i+N_x}^{(\gamma)}(\Lambda_y)_i\right) &= \sqrt{T}(\tilde{D}^{(\gamma)})^{-1}\frac{1}{\sqrt{\gamma}}\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\xi_{i'j} + o_p(1)\\ &= \underbrace{\sqrt{T}\cdot(\tilde{D}^{(\gamma)})^{-1}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}H^{(\gamma)}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\frac{1}{|Q_{i'j}^Z|}\sum_{t\in Q_{i'j}^Z}F_t(e_y)_{ti}}_{\omega_{\Lambda_i,1}} + o_p(1).\end{aligned}$$

By Assumption G4.6, for any $i = 1, \cdots, N_y$ and $i' = i + N_x$,

$$\frac{\sqrt{T}}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \frac{1}{|Q_{i'j}^Z|} \sum_{t \in Q_{i'j}^Z} F_t (e_y)_{ti} \xrightarrow{d} \mathcal{N}(0, \Gamma_{\Lambda_y, i}^{(\gamma),\text{obs}}).$$

From Lemma 2 and Lemma 3, $(\tilde{D}^{(\gamma)})^{-1} \xrightarrow{p} (D^{(\gamma)})^{-1}$ and $H^{(\gamma)} \xrightarrow{p} (Q^{(\gamma)})^{-1}$. Based on Slutsky's Theorem, it holds that

$$\begin{aligned} \sqrt{T} \left((\tilde{\Lambda}_y)_i - H_{i+N_x}^{(\gamma)} (\Lambda_y)_i \right) &= \sqrt{T} \cdot \omega_{\Lambda_i, 1} + o_p(1) \\ &\xrightarrow{d} \mathcal{N} \left(0, (D^{(\gamma)})^{-1} (Q^{(\gamma)})^{-1} \Gamma_{\Lambda_y, i}^{(\gamma),\text{obs}} ((Q^{(\gamma)})^{-1})^\top (D^{(\gamma)})^{-1} \right). \end{aligned}$$

Step 2 – For the second term $\sqrt{T} \left(H_{i+N_x}^{(\gamma)} - H^{(\gamma)} \right) (\Lambda_y)_i$, we have

$$\begin{aligned} \left(H_{i+N_x}^{(\gamma)} - H^{(\gamma)} \right) (\Lambda_y)_i &= (\tilde{D}^{(\gamma)})^{-1} \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \tilde{\Lambda}_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \underbrace{\left(\frac{1}{|Q_{i'j}^Z|} \sum_{t \in Q_{i'j}^Z} F_t F_t^\top - \frac{1}{T} \sum_{t=1}^{T} F_t F_t^\top \right)}_{\Delta_{F,i'j}} (\Lambda_y)_i \\ &= (\tilde{D}^{(\gamma)})^{-1} \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} (\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)} \Lambda_j^{(\gamma)}) \Lambda_j^{(\gamma)\top} \Delta_{F,i'j} (\Lambda_y)_i \\ &\quad + \underbrace{(\tilde{D}^{(\gamma)})^{-1} \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} H^{(\gamma)} \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \Delta_{F,i'j} (\Lambda_y)_i}_{\omega_{\Lambda_i, 2}}. \end{aligned}$$

The first part on the RHS can be bounded by

$$\begin{aligned} &\left\| \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} (\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)} \Lambda_j^{(\gamma)}) \Lambda_j^{(\gamma)\top} \Delta_{F,i'j} (\Lambda_y)_i \right\|^2 \\ \leq &\underbrace{\frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \left\| \tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)} \Lambda_j^{(\gamma)} \right\|^2}_{O_p(\frac{1}{\delta_{N_y,T}}) \text{ by Theorem 1}} \cdot \|(\Lambda_y)_i\|^2 \cdot \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \left\| \Lambda_j^{(\gamma)} \right\|^2 \|\Delta_{F,i'j}\|^2, \end{aligned}$$

where

$$\mathbb{E} \left[\frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \|\Lambda_j^{(\gamma)}\|^2 \|\Delta_{F,i'j}\|^2 \right] = \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \mathbb{E} \left\| \Lambda_j^{(\gamma)} \right\|^2 \mathbb{E} \left\| \Delta_{F,i'j} \right\|^2 \leq \frac{C}{T}.$$

As a result, $\frac{\sqrt{T}}{N_x+N_y} \sum_{j=1}^{N_x+N_y} (\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)} \Lambda_j^{(\gamma)}) \Lambda_j^{(\gamma)\top} \Delta_{F,i'j} (\Lambda_y)_i = o_p(1)$. Consider the second part

$\omega_{\Lambda_i,2}$. By Assumption G4.8 and Slutsky's theorem, we have

$$
\begin{aligned}
\sqrt{T}\omega_{\Lambda_i,2} =& \sqrt{T}(\tilde{D}^{(\gamma)})^{-1}\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y} H^{(\gamma)}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,i'j}(\Lambda_y)_i \\
=& (\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\frac{\sqrt{T}}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\left(\frac{1}{|Q^Z_{i'j}|}\sum_{t\in Q^Z_{i'j}} F_tF_t^\top - \frac{1}{T}\sum_{t=1}^T F_tF_t^\top\right)(\Lambda_y)_i \\
\xrightarrow{d}& \mathcal{N}\left(0, (D^{(\gamma)})^{-1}(Q^{(\gamma)})^{-1}\Gamma^{(\gamma),\text{miss}}_{\Lambda_y,i}((Q^{(\gamma)})^{-1})^\top (D^{(\gamma)})^{-1}\right) \qquad \mathcal{G}^t-\text{stably},
\end{aligned}
$$

where $\Gamma^{(\gamma),\text{miss}}_{\Lambda_y,i} = h^{(\gamma)}_{i+N_x}((\Lambda_y)_i)$.

Step 3 – Observe that $\omega_{\Lambda_i,1}$ and $\omega_{\Lambda_i,2}$ are asymptotically independent. Combining the results from the first two steps, we have

$$
\begin{aligned}
\sqrt{T}\left((\tilde{\Lambda}_y)_i - H^{(\gamma)}(\Lambda_y)_i\right) &= \sqrt{T}\left(\omega_{\Lambda_i,1}+\omega_{\Lambda_i,2}\right) + o_p(1) \\
&\xrightarrow{d} \mathcal{N}\left(0, (D^{(\gamma)})^{-1}(Q^{(\gamma)})^{-1}\left(\Gamma^{(\gamma),\text{obs}}_{\Lambda_y,i}+\Gamma^{(\gamma),\text{miss}}_{\Lambda_y,i}\right)((Q^{(\gamma)})^{-1})^\top (D^{(\gamma)})^{-1}\right)
\end{aligned}
$$

$\mathcal{G}^t$ - stably. If we left-multiply $(\tilde{\Lambda}_y)_i - H^{(\gamma)}(\Lambda_y)_i$ by $(H^{(\gamma)})^{-1}$, according to Lemma 3 and Lemma 4, it holds that

$$
\begin{aligned}
&\sqrt{T}\left((H^{(\gamma)})^{-1}(\tilde{\Lambda}_y)_i - (\Lambda_y)_i\right) \\
\xrightarrow{d}& \mathcal{N}\left(0, \Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\left(\Gamma^{(\gamma),\text{obs}}_{\Lambda_y,i}+\Gamma^{(\gamma),\text{miss}}_{\Lambda_y,i}\right)(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}\right) \quad \mathcal{G}^t-\text{stably},
\end{aligned}
$$

or equivalently,

$$
\sqrt{T}(\Sigma^{(\gamma)}_{\Lambda_y,i})^{-1/2}\left((H^{(\gamma)})^{-1}(\tilde{\Lambda}_y)_i - (\Lambda_y)_i\right) \xrightarrow{d} \mathcal{N}(0, I_k),
$$

where $\Sigma^{(\gamma)}_{\Lambda_y,i} = \Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\left(\Gamma^{(\gamma),\text{obs}}_{\Lambda_y,i}+\Gamma^{(\gamma),\text{miss}}_{\Lambda_y,i}\right)(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}$. □

IA.C.2.2 Proof of Theorem 2.2

Lemma 6. *Suppose that $N_y/N_x \to c \in [0,\infty)$ and $\gamma = r\cdot N_x/N_y$ for some constant r. Under Assumptions G2, G3 and Assumption G4, we have for any t,*

1. $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W^Z_{ti}(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}})$;
2. $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W^Z_{ti}(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}})$;
3. $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} (\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})\Lambda_i^{(\gamma)\top} = O_p(\frac{1}{\delta_{N_y,T}})$;
4. $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W^Z_{ti}(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})\Lambda_i^{(\gamma)\top} = O_p(\frac{1}{\delta_{N_y,T}})$.

Proof. 1. It holds that

$$
\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right)e_{ti}^{(\gamma)}\\
=&(\tilde{D}^{(\gamma)})^{-1}\Bigg(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\eta_{ij}e_{ti}^{(\gamma)}+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\xi_{ij}e_{ti}^{(\gamma)}\\
&+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij}e_{ti}^{(\gamma)}+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\gamma(i,j)e_{ti}^{(\gamma)}\Bigg).
\end{aligned}
$$

Each $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}e_{ti}^{(\gamma)}$ with $\phi_{ij}=\eta_{ij},\xi_{ij},\zeta_{ij}$ and $\gamma(i,j)$ can be decomposed as

$$
\begin{aligned}
&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\phi_{ij}e_{ti}^{(\gamma)}\\
=&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)}\right)\phi_{ij}e_{ti}^{(\gamma)}+\frac{1}{(N_x+N_y)^2}H^{(\gamma)}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\phi_{ij}e_{ti}^{(\gamma)},
\end{aligned}
$$

where the first part on the RHS can be bounded by

$$
\begin{aligned}
&\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)}\right)\phi_{ij}e_{ti}^{(\gamma)}\right\|^2\\
\le&\underbrace{\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)}\right\|^2}_{O_p(\frac{1}{\delta_{N_y,T}})\text{ by Theorem 1}}\cdot\underbrace{\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}(e_{ti}^{(\gamma)})^2}_{O_p(1)}\cdot\underbrace{\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\phi_{ij}^2}_{O_p(\frac{1}{\delta_{N_y,T}})\text{ by Lemma 1}}=O_p(\frac{1}{\delta_{N_y,T}^2}).
\end{aligned}
$$

We analyze the second part $\frac{1}{(N_x+N_y)^2}H^{(\gamma)}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\phi_{ij}e_{ti}^{(\gamma)}$ for $\phi_{ij}=\eta_{ij},\xi_{ij},\zeta_{ij}$ and $\gamma(i,j)$ in the following.

For $\phi_{ij}=\eta_{ij}$, we have

$$
\begin{aligned}
&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\eta_{ij}e_{ti}^{(\gamma)}\\
=&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\Lambda_i^{(\gamma)\top}F_s\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\\
&+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\Lambda_i^{(\gamma)\top}F_s\left(e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}-\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\right).
\end{aligned}
$$

Based on Assumptions G3.1, G3.2 and G3.3(d), it holds that

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\Lambda_i^{(\gamma)\top}F_s\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\right\| \\
\leq&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left[\|\Lambda_j^{(\gamma)}\|\|\Lambda_i^{(\gamma)}\|\right]\cdot\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\mathbb{E}\|F_s\|\cdot\left|\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\right| \\
\leq&\frac{C\gamma}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\frac{1}{T}\sum_{s=1}^{T}\left|\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\right|\leq\frac{C}{N_yT}.
\end{aligned}$$

Additionally, by Assumption G4.4, we have

$$\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\Lambda_i^{(\gamma)\top}F_s\left(e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}-\mathbb{E}\left[e_{sj}^{(\gamma)}e_{ti}^{(\gamma)}\right]\right)=O_p(\frac{1}{\delta_{N_y,T}}).$$

As a result, $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\eta_{ij}e_{ti}^{(\gamma)}=O_p(\frac{1}{\delta_{N_y,T}})$. Combining the first part and the fact $\|H^{(\gamma)}\|=O_p(1)$, we have $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\eta_{ij}e_{ti}^{(\gamma)}=O_p(\frac{1}{\delta_{N_y,T}})$. By similar arguments, we can show that $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\xi_{ij}e_{ti}^{(\gamma)}=O_p(\frac{1}{\delta_{N_y,T}})$.

For $\phi_{ij}=\zeta_{ij}$, we have

$$\begin{aligned}
&\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\zeta_{ij}e_{ti}^{(\gamma)}\right\| \\
=&\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\left(e_{si}^{(\gamma)}e_{sj}^{(\gamma)}-\mathbb{E}\left[e_{si}^{(\gamma)}e_{sj}^{(\gamma)}\right]\right)e_{ti}^{(\gamma)}\right\|=O_p(\frac{1}{\delta_{N_y,T}})
\end{aligned}$$

following from Assumption G4.4. Therefore, $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\tilde{\Lambda}_j^{(\gamma)}\zeta_{ij}e_{ti}^{(\gamma)}=O_p(\frac{1}{\delta_{N_y,T}})$.

Finally for $\phi_{ij}=\gamma(i,j)$, by Assumption G3.3(c), it holds that

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\gamma(i,j)e_{ti}^{(\gamma)}\right\| \\
=&\mathbb{E}\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z\Lambda_j^{(\gamma)}\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\mathbb{E}\left[e_{si}^{(\gamma)}e_{sj}^{(\gamma)}\right]e_{ti}^{(\gamma)}\right\| \\
\leq&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\mathbb{E}\left[\left\|\Lambda_j^{(\gamma)}\right\|\left|e_{ti}^{(\gamma)}\right|\right]\cdot\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}\left|\mathbb{E}\left[e_{si}^{(\gamma)}e_{sj}^{(\gamma)}\right]\right|\leq\frac{C}{N_y}.
\end{aligned}$$

So, $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_j^{(\gamma)}\gamma(i,j)e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}})$.

Combining the four terms and the fact that $\|(\tilde{D}^{(\gamma)})^{-1}\| = O_p(1)$, we derive our result.

2. We have the decomposition

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right)e_{ti}^{(\gamma)} = \underbrace{\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})e_{ti}^{(\gamma)}}_{O_p(\frac{1}{\delta_{N_y,T}})\text{ by Lemma 6.1}}$$
$$+\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}.$$

The leading term of the second part can be bounded

$$\left\|\frac{1}{(N_x+N_y)^2}(\tilde{D}^{(\gamma)})^{-1}\sum_{i,j=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,ij}\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^2$$
$$\le \left\|(\tilde{D}^{(\gamma)})^{-1}\right\|^2\left\|H^{(\gamma)}\right\|^2\cdot\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\cdot$$
$$\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^2,$$

where $\Delta_{F,ij} = \frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z} F_sF_s^\top - \frac{1}{T}\sum_{s=1}^T F_sF_s^\top$. According to Assumption G4.5, it holds that $\mathbb{E}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^4 \le C/T^2N_y^2$. As a result, we have

$$\mathbb{E}\left[\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^2\right]$$
$$\le\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left(\mathbb{E}\left\|\Lambda_j^{(\gamma)}\right\|^4\cdot\mathbb{E}\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right\|^4\right)^{1/2} \le \frac{1}{TN_y},$$

which implies that

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}}).$$

Thus, $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right)e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}})$ as claimed.

3. $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left(\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right)\Lambda_i^{(\gamma)\top}$ can be decomposed as

$$\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left(\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right)\Lambda_i^{(\gamma)\top}\\
&=(\tilde{D}^{(\gamma)})^{-1}\Bigg(\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\eta_{ij}+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\xi_{ij}\\
&\quad+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\zeta_{ij}+\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\gamma(i,j)\Bigg).
\end{aligned}$$

For each $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\phi_{ij}$ with $\phi_{ij}=\eta_{ij},\xi_{ij},\zeta_{ij}$ and $\gamma(i,j)$, we have

$$\begin{aligned}
&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\phi_{ij}\\
=&\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\left(\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)}\right)\Lambda_i^{(\gamma)\top}\phi_{ij}+\frac{1}{(N_x+N_y)^2}H^{(\gamma)}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\phi_{ij}.
\end{aligned}$$

The first term on the RHS can be bounded by

$$\begin{aligned}
&\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}(\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)})\Lambda_i^{(\gamma)\top}\phi_{ij}\right\|^2\\
\leq&\underbrace{\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\tilde{\Lambda}_j^{(\gamma)}-H^{(\gamma)}\Lambda_j^{(\gamma)}\right\|^2}_{O_p(\frac{1}{\delta_{N_y,T}})\text{ by Theorem 1}}\cdot\underbrace{\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\Lambda_i^{(\gamma)}\right\|^2}_{O_p(1)}\cdot\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\phi_{ij}^2=O_p(\frac{1}{\delta_{N_y,T}^2})
\end{aligned}$$

following from Lemma 1. We analyze the second term $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\phi_{ij}$ in the following.

For $\phi_{ij}=\eta_{ij}$, by Assumption G4.2,

$$\begin{aligned}
\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\eta_{ij}\right\|&=\left\|\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_t\Lambda_j^{(\gamma)\top}e_{tj}^{(\gamma)}\right\|\\
&=O_p(\frac{1}{\sqrt{N_yT}}).
\end{aligned}$$

Therefore, $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\eta_{ij}=O_p(\frac{1}{\delta_{N_y,T}})$. By the same arguments, we can show that $\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\tilde{\Lambda}_j^{(\gamma)}\Lambda_i^{(\gamma)\top}\xi_{ij}=O_p(\frac{1}{\delta_{N_y,T}})$.

For $\phi_{ij} = \zeta_{ij}$, by Assumption G4.4 it holds that

$$
\begin{aligned}
\left\| \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \Lambda_j^{(\gamma)} \Lambda_i^{(\gamma)\top} \zeta_{ij} \right\| = & \left\| \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \Lambda_j^{(\gamma)} \Lambda_i^{(\gamma)\top} \frac{1}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} \left(e_{ti}^{(\gamma)} e_{tj}^{(\gamma)} - \mathbb{E}\left[e_{ti}^{(\gamma)} e_{tj}^{(\gamma)} \right] \right) \right\| \\
= & O_p\left(\frac{1}{\delta_{N_y,T}}\right).
\end{aligned}
$$

Thus, $\frac{1}{(N_x+N_y)^2} \sum_{i,j}^{N_x+N_y} \tilde{\Lambda}_j^{(\gamma)} \Lambda_i^{(\gamma)\top} \zeta_{ij} = O_p(\frac{1}{\delta_{N_y,T}})$.

Finally, for $\phi_{ij} = \gamma(i,j)$, we have

$$
\begin{aligned}
\mathbb{E} \left\| \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \Lambda_j^{(\gamma)} \Lambda_i^{(\gamma)\top} \gamma(i,j) \right\| &= \mathbb{E} \left\| \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \Lambda_j^{(\gamma)} \Lambda_i^{(\gamma)\top} \frac{1}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} \mathbb{E}\left[e_{ti}^{(\gamma)} e_{tj}^{(\gamma)} \right] \right\| \\
&\leq \frac{1}{(N_x+N_y)^2} \sum_{i,j=1}^{N_x+N_y} \mathbb{E}\left[\|\Lambda_j^{(\gamma)}\| \|\Lambda_i^{(\gamma)}\| \right] \cdot \frac{1}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} \left| \mathbb{E}\left[e_{ti}^{(\gamma)} e_{tj}^{(\gamma)} \right] \right| \\
&\leq \frac{C}{N_y},
\end{aligned}
$$

where the last equality follows from Assumption G3.3(c).

Combining the four terms and the fact that $\|(\tilde{D}^{(\gamma)})^{-1}\| = O_p(1)$, we complete our proof.

4. We omit the proof of $\frac{1}{N_x+N_y} \sum_{i=1}^{N_x+N_y} W_{ti}^Z \left(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)} \Lambda_i^{(\gamma)} \right) \Lambda_i^{(\gamma)\top} = O_p(\frac{1}{\delta_{N_y,T}})$, which is similar to the proof of Lemma 6.3. □

Proof. Proof of Theorem 2.2:

We derive the estimated factors $\tilde{F}_t$ by regressing the observed $Z_{ti}^{(\gamma)}$ on $\tilde{\Lambda}_i^{(\gamma)}$, i.e.

$$
\tilde{F}_t = \left(\sum_{i=1}^{N_x+N_y} W_{ti}^Z \tilde{\Lambda}_i^{(\gamma)} \tilde{\Lambda}_i^{(\gamma)\top} \right)^{-1} \left(\sum_{i=1}^{N_x+N_y} W_{ti}^Z Z_{ti}^{(\gamma)} \tilde{\Lambda}_i^{(\gamma)} \right), \quad t = 1, \cdots, T.
$$

Similar with the proof of Theorem 1.1, we resort to the auxiliary $\tilde{F}_t^*$, which is defined as

$$
\tilde{F}_t^* = \left(\sum_{i=1}^{N_x+N_y} W_{ti}^Z H^{(\gamma)} \Lambda_i^{(\gamma)} \Lambda_i^{(\gamma)\top} H^{(\gamma)\top} \right)^{-1} \left(\sum_{i=1}^{N_x+N_y} W_{ti}^Z Z_{ti}^{(\gamma)} \tilde{\Lambda}_i^{(\gamma)} \right).
$$

Step 1 – In the first step, we analyze $\tilde{F}_t^*$. We have the decomposition

$$H^{(\gamma)\top}\tilde{F}_t^* = F_t + (\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)}\right)$$
$$+ (\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}(H^{(\gamma)})^{-1}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \left(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right)\left(\Lambda_i^{(\gamma)\top}F_t + e_{ti}^{(\gamma)}\right)\right),$$

where $\tilde{\Sigma}_{\Lambda,t}^{(\gamma)} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top} \xrightarrow{p} \Sigma_{\Lambda,t}^{(\gamma)}$ is positive definite by Theorem 1.

Consider the first part $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)}$, we have

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)} = \frac{\sqrt{N_x}}{N_x+N_y}\frac{1}{\sqrt{N_x}}\sum_{i=1}^{N_x}(\Lambda_x)_i(e_x)_{ti} + \frac{\gamma\sqrt{N_y}}{N_x+N_y}\frac{1}{\sqrt{N_y}}\sum_{i=1}^{N_y} W_{ti}^Y(\Lambda_y)_i(e_y)_{ti}.$$

If all the factors in F_y are strong factors in Y, then each entry of $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)}$ will converge at the same rate of $\sqrt{N_y}$, which is determined by the second term. By Assumption G4.7,

$$\frac{\sqrt{N_y}}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)} \xrightarrow{d} \mathcal{N}\left(0, \Gamma_{F,t}^{(\gamma),\text{obs}}\right),$$

where $\Gamma_{F,t}^{(\gamma),\text{obs}}$ is a positive definite matrix. We let $\omega_{F,1} \coloneqq (\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_i^{(\gamma)} e_{ti}^{(\gamma)})$. Then $\sqrt{N_y}\cdot\omega_{F,1} \xrightarrow{d} \mathcal{N}(0, (\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\Gamma_{F,t}^{(\gamma),\text{obs}}(\Sigma_{\Lambda,t}^{(\gamma)})^{-1})$. If some factor $F_{t,w}$ is weak in Y whose loading $\sum_{i=1}^{N_y}(\Lambda_y)_{i,w}^2$ grows at the rate $g(N_y)$ which is sub-linear or constant in N_y, then for $F_{t,w}$ there is

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_{i,w}^{(\gamma)} e_{ti}^{(\gamma)} = \frac{\sqrt{N_x}}{N_x+N_y}\frac{1}{\sqrt{N_x}}\sum_{i=1}^{N_x}(\Lambda_x)_{i,w}(e_x)_{ti}$$
$$+ \frac{\gamma\sqrt{g(N_y)}}{N_x+N_y}\frac{1}{\sqrt{g(N_y)}}\sum_{i=1}^{N_y} W_{ti}^Y(\Lambda_y)_{i,w}(e_y)_{ti}.$$

If $g(N_y)N_x/N_y^2 \to \infty$, the convergence rate of $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_{i,w}^{(\gamma)} e_{ti}^{(\gamma)}$ is $O(N_y/\sqrt{g(N_y)})$, which is determined by the second term; otherwise, the convergence rate is $O(\sqrt{N_x})$. Combining these two cases, Assumption G4.7 assumes that

$$\frac{\sqrt{N_w}}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z \Lambda_{i,w}^{(\gamma)} e_{ti}^{(\gamma)} \xrightarrow{d} \mathcal{N}\left(0, \Gamma_{F_w,t}^{(\gamma),\text{obs}}\right),$$

where $N_w = \min\left(N_y^2/g(N_y), N_x\right)$ and $\Gamma_{F_w,t}^{(\gamma),\text{obs}}$ is positive definite. If we can assume that the loadings of $F_{t,w}$ are orthogonal to the loadings of other factors, then we have

$$\sqrt{N_w}\,(\omega_{F,1})_w \xrightarrow{d} \mathcal{N}\left(0, (\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\Gamma_{F_w,t}^{(\gamma),\text{obs}}(\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\right),$$

where $\Sigma_{\Lambda,t,w}^{(\gamma)}$ is the diagonal entry of $\Sigma_{\Lambda,t}^{(\gamma)}$ corresponding to the weak factor.

Now, we consider the second part $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})(\Lambda_i^{(\gamma)\top}F_t + e_{ti}^{(\gamma)})$. Observe that

$$\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)}\right)\Lambda_i^{(\gamma)\top}F_t \\
=&\underbrace{\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z\left(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)}\right)\Lambda_i^{(\gamma)\top}F_t}_{\Delta_{t,1}} + \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t.
\end{aligned}$$

Let $\Delta_{F,ij} = \frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z} F_sF_s^\top - \frac{1}{T}\sum_{s=1}^T F_sF_s^\top$. The second part $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t$ can be further decomposed as

$$\begin{aligned}
&\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y} W_{ti}^Z(H_i^{(\gamma)} - H^{(\gamma)})\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t \\
=\;&\underbrace{(\tilde{D}^{(\gamma)})^{-1}\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\left(\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)}\right)\Lambda_j^{(\gamma)\top}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t}_{\Delta_{t,2}} \\
&+\underbrace{(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t}_{\Delta_{t,3}}.
\end{aligned}$$

For $\Delta_{t,3}$, we have

$$\Delta_{t,3} = (\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t = (\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\mathbf{X}_t^{(\gamma)}F_t,$$

where $\mathbf{X}_t^{(\gamma)} = \frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}$ is asymptotically normal with convergence rate $\sqrt{T}$ from Assumption G4.8. We denote $\omega_{F,2} := (\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\mathbf{X}_t^{(\gamma)}F_t$.

For $\Delta_{t,2}$, since $\mathbb{E}\left\|\Delta_{F,ij}\right\|^2 \leq \frac{C}{T}$ and $\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)}\right\|^2 = O_p(\frac{1}{\delta_{N_y,T}})$ by Theorem 1,

$$\begin{aligned}
&\left\|(\tilde{D}^{(\gamma)})^{-1}\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\left(\tilde{\Lambda}_j^{(\gamma)} - H^{(\gamma)}\Lambda_j^{(\gamma)}\right)\Lambda_j^{(\gamma)\top}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}F_t\right\|^2 \\
&\leq O_p(\frac{1}{\delta_{N_y,T}})\cdot\left(\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\left\|\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Delta_{F,ij}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\right\|^2\right) \\
&\leq O_p(\frac{1}{\delta_{N_y,T}})\cdot\left(\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\cdot\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\Lambda_i^{(\gamma)}\right\|^2\left\|\Delta_{F,ij}\right\|\right)^2\right),
\end{aligned}$$

where

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\left\|\Lambda_j^{(\gamma)}\right\|^2\cdot\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\left\|\Lambda_i^{(\gamma)}\right\|^2\left\|\Delta_{F,ij}\right\|\right)^2\right] \\
&= \frac{1}{(N_x+N_y)^3}\sum_{i,j,l=1}^{N_x+N_y}\mathbb{E}\left[\|\Lambda_i^{(\gamma)}\|^2\|\Lambda_j^{(\gamma)}\|^2\|\Lambda_l^{(\gamma)}\|^2\right]\cdot\mathbb{E}\left[\|\Delta_{F,ij}\|\|\Delta_{F,lj}\|\right] \\
&\leq \frac{1}{(N_x+N_y)^3}\sum_{i,j,l=1}^{N_x+N_y}\mathbb{E}\left[\|\Lambda_i^{(\gamma)}\|^2\|\Lambda_j^{(\gamma)}\|^2\|\Lambda_l^{(\gamma)}\|^2\right]\cdot\left(\mathbb{E}\|\Delta_{F,ij}\|^2\cdot\mathbb{E}\|\Delta_{F,lj}\|^2\right)^{1/2} \leq \frac{C}{T}.
\end{aligned}$$

As a result, $\Delta_{t,2} = O_p(\frac{1}{\delta_{N_y,T}})$. By Lemma 6, $\Delta_{t,1} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H_i^{(\gamma)}\Lambda_i^{(\gamma)})\Lambda_i^{(\gamma)\top}F_t = O_p(\frac{1}{\delta_{N_y,T}})$ and $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z(\tilde{\Lambda}_i^{(\gamma)} - H^{(\gamma)}\Lambda_i^{(\gamma)})e_{ti}^{(\gamma)} = O_p(\frac{1}{\delta_{N_y,T}})$. For $\sqrt{T}/N_y \to 0$, they are all small order terms compared to $\Delta_{t,3}$. Therefore, we have $H^{(\gamma)\top}\tilde{F}_t^* = F_t + \omega_{F,1} + \omega_{F,2} + O_p(\frac{1}{\delta_{N_y,T}})$.

Step 2 – Next, we analyze the difference between $\tilde{F}_t$ and $\tilde{F}_t^*$. We have the decomposition

$$\begin{aligned}
\tilde{F}_t^* - \tilde{F}_t = &\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top}\right)^{-1}\cdot \\
&\left[\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\right]\tilde{F}_t^*.
\end{aligned}$$

We let $\Delta_{\Lambda,t} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} - \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}$. According to the proof of Theorem 1, we have

$$\begin{aligned}\Delta_{\Lambda,t} =& \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Big[H^{(\gamma)}W_{ti}^{Z}\Lambda_i^{(\gamma)}(\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)})^\top + W_{ti}^{Z}(\tilde{\Lambda}_i^{(\gamma)}-H_i^{(\gamma)}\Lambda_i^{(\gamma)})\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}\Big]\\ &+\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Big[W_{ti}^{Z}H^{(\gamma)}\Lambda_i^{(\gamma)}(H_i^{(\gamma)}\Lambda_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})^\top + W_{ti}^{Z}(H_i^{(\gamma)}\Lambda_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})(H^{(\gamma)}\Lambda_i^{(\gamma)})^\top\Big]\\ &+\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^{Z}(\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})(\tilde{\Lambda}_i^{(\gamma)}-H^{(\gamma)}\Lambda_i^{(\gamma)})^\top.\end{aligned}$$

By Lemma 6 and Theorem 1, the first and third parts on the RHS are at the order $O_p(\frac{1}{\delta_{N_y,T}})$. Thus, $\Delta_{\Lambda,t} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}\Big[W_{ti}^{Z}H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}(H_i^{(\gamma)}-H^{(\gamma)})^\top + W_{ti}^{Z}(H_i^{(\gamma)}-H^{(\gamma)})(H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top})^\top\Big]+$ $O_p\big(\frac{1}{\delta_{N_y,T}}\big)$. As analyzed in the first step, the leading term of $H_i^{(\gamma)}-H^{(\gamma)}$ is

$$\frac{1}{N_x+N_y}(\tilde{D}^{(\gamma)})^{-1}\sum_{j=1}^{N_x+N_y}H^{(\gamma)}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Delta_{F,ij},$$

and thus, the leading term of $\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^{Z}(H_i^{(\gamma)}-H^{(\gamma)})\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top}$ is $(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\mathbf{X}_t^{(\gamma)}H^{(\gamma)\top}$. As a result,

$$\Delta_{\Lambda,t} = (\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\mathbf{X}_t^{(\gamma)}H^{(\gamma)\top} + H^{(\gamma)}\mathbf{X}_t^{(\gamma)}H^{(\gamma)\top}((\tilde{D}^{(\gamma)})^{-1})^\top + O_p(\frac{1}{\delta_{N_y,T}}).$$

Note that $\mathbf{X}_t^{(\gamma)}$ is asymptotically normal with convergence rate $\sqrt{T}$, we have

$$\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^{Z}\tilde{\Lambda}_i^{(\gamma)}\tilde{\Lambda}_i^{(\gamma)\top} = \frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^{Z}H^{(\gamma)}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}H^{(\gamma)\top} + O_p(\frac{1}{\sqrt{T}}).$$

According to the first step, $H^{(\gamma)\top}\tilde{F}_t^{*} = F_t + O_p(\frac{1}{\sqrt{\delta_{N_y,T}}})$. Combining this with $\Delta_{\Lambda,t}$, we derive

$$\begin{aligned}H^{(\gamma)\top}(\tilde{F}_t^{*}-\tilde{F}_t) =& \underbrace{(\tilde{\Sigma}_{\Lambda,t}^{(\gamma)})^{-1}(H^{(\gamma)})^{-1}\Big[(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\mathbf{X}_tH^{(\gamma)\top} + H^{(\gamma)}\mathbf{X}_tH^{(\gamma)\top}((\tilde{D}^{(\gamma)})^{-1})^\top\Big](H^{(\gamma)\top})^{-1}F_t}_{\omega_{F,3}}\\ &+O_p(\frac{1}{\delta_{N_y,T}}).\end{aligned}$$

Step 3 – In the final step, we analyze the asymptotic distribution of $\tilde{F}_t$, which is determined by

$\omega_{F,1}$, $\omega_{F,2}$ and $\omega_{F,3}$ from the first two steps, i.e.,

$$H^{(\gamma)\top}\tilde{F}_t - F_t = \omega_{F,1} + \omega_{F,2} - \omega_{F,3} + O_p(\frac{1}{\delta_{N_y,T}}).$$

We have $\omega_{F,2} - \omega_{F,3} = -\tilde{\Sigma}_{\Lambda,t}^{(\gamma)} \cdot \mathbf{X}_t^{(\gamma)} H^{(\gamma)\top}((\tilde{D}^{(\gamma)})^{-1})^\top (H^{(\gamma)\top})^{-1} F_t$. According to Lemma 4, there is $(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)} \xrightarrow{p} \left(\frac{F^\top F}{T}\right)^{-1}\left(\frac{\Lambda^{(\gamma)\top}\Lambda^{(\gamma)}}{N_x+N_y}\right)^{-1}$. By Assumption G4.8 and Slutsky's Theorem, it holds that

$$\sqrt{T}\left(\omega_{F,2} - \omega_{F,3}\right) \xrightarrow{d} \mathcal{N}\left(0, (\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\Gamma_{F,t}^{(\gamma),\text{miss}}(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\right) \quad \mathcal{G}^t - \text{stably},$$

where $\Gamma_{F,t}^{(\gamma),\text{miss}} = g_t^{(\gamma)}((\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}F_t)$ with function $g_t^{(\gamma)}(\cdot)$ defined in Assumption G4.8. Additionally, $\omega_{F,1}$ and $\omega_{F,2} - \omega_{F,3}$ are asymptotically independent. If all the factors in F_y are strong factors in Y, we can deduce that

$$\begin{aligned}&\sqrt{\delta_{N_y,T}}\left(H^{(\gamma)\top}\tilde{F}_t - F_t\right)\\ &\xrightarrow{d} \mathcal{N}\left(0, (\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\left[\text{plim}\left(\frac{\delta_{N_y,T}}{N_y}\Gamma_{F,t}^{(\gamma),\text{obs}} + \frac{\delta_{N_y,T}}{T}\Gamma_{F,t}^{(\gamma),\text{miss}}\right)\right](\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\right) \quad \mathcal{G}^t - \text{stably}.\end{aligned}$$

If some factor $F_{t,w}$ is weak in Y and its loadings are orthogonal to the loadings of the other factors, then

$$\sqrt{T}\left(\omega_{F,2} - \omega_{F,3}\right)_w \xrightarrow{d} \mathcal{N}\left(0, (\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\Gamma_{F_w,t}^{(\gamma),\text{miss}}(\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\right) \quad \mathcal{G}^t - \text{stably},$$

where $\Gamma_{F_w,t}^{(\gamma),\text{miss}}$ corresponds to the weak factor in $\Gamma_{F,t}^{(\gamma),\text{miss}}$, and thus,

$$\begin{aligned}&\sqrt{\delta_{N_w,T}}\left((H^{(\gamma)\top}\tilde{F}_t)_w - F_{t,w}\right)\\ &\xrightarrow{d} \mathcal{N}\left(0, (\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\left[\text{plim}\left(\frac{\delta_{N_w,T}}{N_w}\Gamma_{F_w,t}^{(\gamma),\text{obs}} + \frac{\delta_{N_w,T}}{T}\Gamma_{F_w,t}^{(\gamma),\text{miss}}\right)\right](\Sigma_{\Lambda,t,w}^{(\gamma)})^{-1}\right) \quad \mathcal{G}^t - \text{stably},\end{aligned}$$

where $\delta_{N_w,T} = \min(N_w, T)$ and $N_w = \min(N_y^2/g(N_y), N_x)$. □

IA.C.2.3 Proof of Theorem 2.3

Proof. For any $t = 1, \cdots, T$ and $i = 1, \cdots, N_y$, we have the decomposition

$$\begin{aligned}\tilde{C}_{ti} - C_{ti} &= \tilde{F}_t^\top(\tilde{\Lambda}_y)_i - F_t^\top(\Lambda_y)_i\\ &= \tilde{F}_t^\top\left((\tilde{\Lambda}_y)_i - H^{(\gamma)}(\Lambda_y)_i\right) + \left(\tilde{F}_t^\top H^{(\gamma)} - F_t^\top\right)(\Lambda_y)_i.\end{aligned}$$

From Theorem 2.1 and Theorem 2.2, it holds that

$$\begin{aligned}\sqrt{\delta_{N_y,T}}(\tilde{C}_{ti}-C_{ti}) =&\sqrt{\delta_{N_y,T}}F_t^\top(H^{(\gamma)})^{-1}\left(\omega_{\Lambda,1}+\omega_{\Lambda,2}\right)\\&+\sqrt{\delta_{N_y,T}}(\Lambda_y)_i^\top\left(\omega_{F,1}+\omega_{F,2}-\omega_{F,3}\right)+o_p(1).\end{aligned}$$

Plugging the expression of $\omega_{\Lambda,1},\omega_{\Lambda,2},\omega_{F,1},\omega_{F,2}$ and $\omega_{F,3}$ into the RHS, we obtain

$$\begin{aligned}&\sqrt{\delta_{N_y,T}}(\tilde{C}_{ti}-C_{ti})\\=&F_t^\top(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\frac{\sqrt{\delta_{N_y,T}}}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\frac{1}{|Q_{i'j}^Z|}\sum_{t\in Q_{i'j}^Z}F_t(e_y)_{ti}\\&+F_t^\top(H^{(\gamma)})^{-1}(\tilde{D}^{(\gamma)})^{-1}H^{(\gamma)}\sqrt{\delta_{N_y,T}}X_{i+N_x}(\Lambda_y)_i\\&+(\Lambda_y)_i^\top\sqrt{\delta_{N_y,T}}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\right)^{-1}\left(\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}W_{ti}^Z\Lambda_i^{(\gamma)}e_{ti}^{(\gamma)}\right)\\&-(\Lambda_y)_i^\top\sqrt{\delta_{N_y,T}}\tilde{\Sigma}_{\Lambda,t}^{(\gamma)}\mathbf{X}_t^{(\gamma)}H^{(\gamma)\top}((\tilde{D}^{(\gamma)})^{-1})^\top(H^{(\gamma)\top})^{-1}F_t+o_p(1),\end{aligned}$$

where $X_i,\mathbf{X}_t$ are defined as $X_i=\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\left(\frac{1}{|Q_{ij}^Z|}\sum_{s\in Q_{ij}^Z}F_sF_s^\top-\frac{1}{T}\sum_{s=1}^T F_sF_s^\top\right)$, $\mathbf{X}_t=\frac{1}{N_x+N_y}\sum_{i=1}^{N_x+N_y}X_iW_{ti}^Z\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}$, and $\tilde{\Sigma}_{\Lambda,t}^{(\gamma)}=\frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}W_{tj}^Z\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}$. Note that X_i and $\mathbf{X}_t$ are correlated and are asymptotically independent of other terms in $\sqrt{\delta_{N_y,T}}(\tilde{C}_{ti}-C_{ti})$. Therefore, by Assumption G4, Lemma 4 and proof of Theorems 2.1 and 2.2, we can conclude that

$$\begin{aligned}&\sqrt{\delta_{N_y,T}}(\tilde{C}_{ti}-C_{ti})\xrightarrow{d}\\&\mathcal{N}\Bigg(0,\operatorname{plim}\Bigg(\frac{\delta_{N_y,T}}{T}F_t^\top\Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\Gamma_{\Lambda_y,i}^{(\gamma),\text{obs}}(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}F_t+\frac{\delta_{N_y,T}}{N_y}(\Lambda_y)_i^\top(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\Gamma_{F,t}^{(\gamma),\text{obs}}(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}(\Lambda_y)_i\\&+\frac{\delta_{N_y,T}}{T}(\Lambda_y)_i^\top(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\Gamma_{F,t}^{(\gamma),\text{miss}}(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}(\Lambda_y)_i+\frac{\delta_{N_y,T}}{T}F_t^\top\Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\Gamma_{\Lambda_y,i}^{(\gamma),\text{miss}}(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}F_t\\&-2\cdot\frac{\delta_{N_y,T}}{T}(\Lambda_y)_i^\top(\Sigma_{\Lambda,t}^{(\gamma)})^{-1}\Gamma_{\Lambda_y,F,i,t}^{(\gamma),\text{miss,cov}}(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}F_t\Bigg)\Bigg)\quad\mathcal{G}^t-\text{stably},\end{aligned}$$

where $\Gamma_{\Lambda_y,F,i,t}^{(\gamma),\text{miss,cov}}=g_{i,t}^{(\gamma),\text{cov}}((\Lambda_y)_i,(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1}F_t)$ with function $g_{i,t}^{(\gamma),\text{cov}}(\cdot)$ defined in Assumption G4.8. Equivalently, we have

$$\sqrt{\delta_{N_y,T}}(\Sigma_{C,ti}^{(\gamma)})^{-1/2}(\tilde{C}_{ti}-C_{ti})\xrightarrow{d}\mathcal{N}(0,1),$$

where

$$
\begin{aligned}
\Sigma_{C,ti}^{(\gamma)} =& \frac{\delta_{N_y,T}}{T} F_t^\top \Sigma_F^{-1} (\Sigma_\Lambda^{(\gamma)})^{-1} \Gamma_{\Lambda_y,i}^{(\gamma),\text{obs}} (\Sigma_\Lambda^{(\gamma)})^{-1} \Sigma_F^{-1} F_t + \frac{\delta_{N_y,T}}{N_y} (\Lambda_y)_i^\top (\Sigma_{\Lambda,t}^{(\gamma)})^{-1} \Gamma_{F,t}^{(\gamma),\text{obs}} (\Sigma_{\Lambda,t}^{(\gamma)})^{-1} (\Lambda_y)_i \\
&+ \frac{\delta_{N_y,T}}{T} (\Lambda_y)_i^\top (\Sigma_{\Lambda,t}^{(\gamma)})^{-1} \Gamma_{F,t}^{(\gamma),\text{miss}} (\Sigma_{\Lambda,t}^{(\gamma)})^{-1} (\Lambda_y)_i + \frac{\delta_{N_y,T}}{T} F_t^\top \Sigma_F^{-1} (\Sigma_\Lambda^{(\gamma)})^{-1} \Gamma_{\Lambda_y,i}^{(\gamma),\text{miss}} (\Sigma_\Lambda^{(\gamma)})^{-1} \Sigma_F^{-1} F_t \\
&- 2 \cdot \frac{\delta_{N_y,T}}{T} (\Lambda_y)_i^\top (\Sigma_{\Lambda,t}^{(\gamma)})^{-1} \Gamma_{\Lambda_y,F,i,t}^{(\gamma),\text{miss,cov}} (\Sigma_\Lambda^{(\gamma)})^{-1} \Sigma_F^{-1} F_t.
\end{aligned}
$$

We complete our proof. □

IA.C.3 Proof of Proposition 5

We prove that Assumptions G2 and S1 imply Assumption G3, and Assumptions G2, S1, and S2 imply Assumption G4 in the following.

IA.C.3.1 Assumptions G2 and S1 imply Assumption G3

1. Assumptions G3.1 and G3.2 hold under Assumptions G2, S1.1 and S1.2

Proof. Since $F_t \overset{i.i.d.}{\sim} (0, \Sigma_F)$, by LLN we have $\frac{1}{T}\sum_{t=1}^T F_t F_t^\top \xrightarrow{p} \Sigma_F$. Additionally, since $\mathbb{E}\|F_t\|^4$ is bounded, there is

$$
\begin{aligned}
\mathbb{E}\left\|\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^T F_t F_t^\top - \Sigma_F\right)\right\|^2 &= \sum_{p,q=1}^k T \cdot \mathbb{E}\left[\left(\frac{1}{T}\sum_{t=1}^T F_{t,p}F_{t,q} - (\Sigma_F)_{pq}\right)^2\right] \\
&= \sum_{p,q=1}^k \mathbb{E}\left[F_{t,p}^2 F_{t,q}^2\right] - (\Sigma_F)_{pq}^2 \le C,
\end{aligned}
$$

where $F_{t,p}$ denotes the p-th factor of F_t and $(\Sigma_F)_{pq}$ denotes the (p,q)-th entry of Σ_F. Since the observation matrix W^Y is independent of factors F, by similar arguments we can show that $\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} F_t F_t^\top \xrightarrow{p} \Sigma_F$ and $\mathbb{E}\left\|\sqrt{|Q_{ij}^Z|}\left(\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z} F_t F_t^\top - \Sigma_F\right)\right\|^2 \le C$. For factor loadings, since $(\Lambda_x)_i \overset{i.i.d.}{\sim} (0, \Sigma_{\Lambda_x})$, by LLN we have $\frac{1}{N_x}\sum_{i=1}^{N_x} (\Lambda_x)_i (\Lambda_x)_i^\top \xrightarrow{p} \Sigma_{\Lambda_x}$. Other assumptions in Assumption G3.2 automatically hold under Assumption S1.2. □

2. Assumption G3.3 holds under Assumption S1.3

Proof. When Assumption S1.3 holds, Assumption G3.3(a) automatically holds. Let $I(\cdot)$ be an indicator function where $I(A) = 1$ if event A happens and $I(A) = 0$ otherwise. Under Assumption S1.3, there are $\gamma_{st,i}^{(e_x)} = \mathbb{E}[(e_x)_{ti}(e_x)_{si}] = \sigma_{e_x}^2 \cdot I(t=s)$ and $\gamma_{st,i}^{(e_y)} = \mathbb{E}[(e_y)_{ti}(e_y)_{si}] = \sigma_{e_y}^2 \cdot I(t=s)$. Let $\gamma_{st} = (\sigma_{e_x}^2 + \sigma_{e_y}^2)\cdot I(t=s)$. It satisfies $|\gamma_{st,i}^{(e_x)}| \le \gamma_{st}$, $|\gamma_{st,i}^{(e_y)}| \le \gamma_{st}$ and $\sum_{s=1}^T \gamma_{st} \le C$ for all t. By the same arguments, we can prove Assumption G3.3(c). For Assumption G3.3(d), there is

$\tau^{(e_x)}_{ij,ts} = \sigma^2_{e_x} \cdot I(i = j, t = s)$. So $\sum_{j=1}^{N_x} \sum_{s=1}^{T} |\tau^{(e_x)}_{ij,ts}|$ is bounded for all i, t. Similar arguments hold for $\tau^{(e_y)}_{ij,ts}$ and $\tau^{(e_x,e_y)}_{ij,ts}$. We denote $v^{(y)}_{t,ij} = (e_y)_{ti}(e_y)_{tj} - \mathbb{E}[(e_y)_{ti}(e_y)_{tj}]$. We have $\mathbb{E}[v^{(y)}_{t,ij}] = 0$, and since $\mathbb{E}(e_y)^8_{ti}$ is bounded,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{|Q^Y_{ij}|^{1/2}} \sum_{t\in Q^Y_{ij}} v^{(y)}_{t,ij}\right]^4 &= \frac{1}{|Q^Y_{ij}|^2} \sum_{t,s,u,w\in Q^Y_{ij}} \mathbb{E}\left[v^{(y)}_{t,ij} v^{(y)}_{s,ij} v^{(y)}_{u,ij} v^{(y)}_{w,ij}\right] \\
&= \frac{1}{|Q^Y_{ij}|^2} \left[3 \sum_{t,s\in Q^Y_{ij}} \mathbb{E}\left[(v^{(y)}_{t,ij})^2 (v^{(y)}_{s,ij})^2\right] + \sum_{t\in Q^Y_{ij}} \mathbb{E}\left[(v^{(y)}_{t})^4\right]\right] \leq C.
\end{aligned}$$

By similar arguments, we can prove that $\mathbb{E}\Big[\frac{1}{T^{1/2}} \sum_{t=1}^{T} \big((e_x)_{ti}(e_x)_{tj} - \mathbb{E}[(e_x)_{ti} \cdot (e_x)_{tj}]\big)\Big]^4$, and additionally, $\mathbb{E}\Big[\frac{1}{|Q^Y_{jj}|^{1/2}} \sum_{t\in Q^Y_{jj}} \big((e_x)_{ti}(e_y)_{tj} - \mathbb{E}[(e_x)_{ti}(e_y)_{tj}]\big)\Big]^4$, are bounded. □

3. Assumption G3.4 holds under Assumption S1.4

Proof. Since F, e_y, and W^Y are independent, it is easy to see that for any $i, j = 1, \cdots, N_y$,

$$\mathbb{E}\left\|\frac{1}{\sqrt{|Q^Y_{ij}|}} \sum_{t\in Q^Y_{ij}} F_t (e_y)_{tj}\right\|^2 = \sum_{p=1}^{k} \frac{1}{|Q^Y_{ij}|} \sum_{t\in Q^Y_{ij}} \mathbb{E}[(e_y)^2_{tj}] \cdot \mathbb{E}[F^2_{t,p}] \leq C.$$

Similarly, $\mathbb{E}\left\|\frac{1}{\sqrt{|Q^Y_{ii}|}} \sum_{t\in Q^Y_{ii}} F_t (e_x)_{tj'}\right\|^2 \leq C$ and $\mathbb{E}\left\|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t (e_x)_{tj'}\right\|^2 \leq C$ for any $j' = 1, \cdots, N_x$. □

IA.C.3.2 Assumptions G2, S1 and S2 imply Assumption G4

1. Assumption G4.1 holds under Assumptions G2 and S1

Proof. Since factors F, loadings Λ_x, Λ_y, and errors e_x, e_y are all i.i.d. with zero means, and $|Q^Y_{ij}|/T$

is bounded away from 0 for all i, j, we have

$$\mathbb{E}\left\|\sqrt{\frac{T}{N_y}}\sum_{j=1}^{N_y}(\Lambda_y)_j\frac{1}{|Q_{ij}^Y|}\sum_{t\in Q_{ij}^Y}F_t^\top(e_y)_{tj}\right\|^2 = \sum_{p,q=1}^{k}\mathbb{E}\left[\left(\sqrt{\frac{T}{N_y}}\sum_{j=1}^{N_y}(\Lambda_y)_{j,p}\frac{1}{|Q_{ij}^Y|}\sum_{t\in Q_{ij}^Y}F_{t,q}(e_y)_{tj}\right)^2\right]$$
$$=\sum_{p,q=1}^{k}\frac{T}{N_y}\sum_{j,l=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\frac{1}{|Q_{il}^Y|}\sum_{t\in Q_{ij}^Y}\sum_{s\in Q_{il}^Y}\mathbb{E}\Big[(\Lambda_y)_{j,p}(\Lambda_y)_{l,p}F_{t,q}F_{s,q}(e_y)_{tj}(e_y)_{sl}\Big]$$
$$=\sum_{p,q=1}^{k}\frac{T}{N_y}\sum_{j=1}^{N_y}\frac{1}{|Q_{ij}^Y|^2}\sum_{t\in Q_{ij}^Y}\mathbb{E}\Big[(\Lambda_y)_{j,p}^2F_{t,q}^2(e_y)_{tj}^2\Big]\le C.$$

Similarly we can prove that $\mathbb{E}\left\|\sqrt{\frac{T}{N_x}}\sum_{j=1}^{N_x}(\Lambda_x)_j\frac{1}{|Q_{ii}^Y|}\sum_{t\in Q_{ii}^Y}F_t^\top(e_x)_{tj}\right\|^2\le C.$ □

2. Assumption G4.2 holds under Assumptions G2 and S1

Proof. It holds that

$$\mathbb{E}\left\|\frac{\sqrt{N_yT}}{N_yN_x}\sum_{i=1}^{N_y}\sum_{j=1}^{N_x}(\Lambda_x)_j(\Lambda_x)_j^\top\frac{1}{|Q_{ii}^Y|}\sum_{t\in Q_{ii}^Y}F_t(\Lambda_y)_i^\top(e_y)_{ti}\right\|^2$$
$$=\sum_{p,q=1}^{k}\frac{T}{N_yN_x^2}\mathbb{E}\left[\left(\sum_{i=1}^{N_y}\sum_{j=1}^{N_x}\sum_{r=1}^{k}\frac{1}{|Q_{ii}^Y|}\sum_{t\in Q_{ii}^Y}(\Lambda_x)_{j,p}(\Lambda_x)_{j,r}F_{t,r}(\Lambda_y)_{i,q}(e_y)_{ti}\right)^2\right]$$
$$=\sum_{p,q,r,m=1}^{k}\frac{T}{N_yN_x^2}\sum_{i=1}^{N_y}\sum_{j,l=1}^{N_x}\frac{1}{|Q_{ii}^Y|^2}\sum_{t\in Q_{ii}^Y}\mathbb{E}\big[(\Lambda_x)_{j,p}(\Lambda_x)_{j,r}(\Lambda_x)_{l,p}(\Lambda_x)_{l,m}\big]\cdot\mathbb{E}[F_{t,r}F_{t,m}]\cdot\mathbb{E}[(\Lambda_y)_{i,q}^2]\cdot\mathbb{E}[(e_y)_{ti}^2]$$
$$\le\sum_{p,q,r,m=1}^{k}\sum_{r,m=1}^{k}\frac{T}{|Q_{ii}^Y|}C\ \le\ C,$$

where the first inequality holds since Λ_x has bounded fourth moments. By similar arguments, we can prove the other three bounds in Assumption G4.2. □

3. Assumption G4.3 holds under Assumptions G2 and S1

Proof. For simplicity, we let $v_{t,ij}^{(y)}=(e_y)_{ti}(e_y)_{tj}-\mathbb{E}\left[(e_y)_{ti}(e_y)_{tj}\right]$. We have $\mathbb{E}[v_{t,ij}^{(y)}]=0$. According to Assumption S1, $v_{t,ij}^{(y)}$ is independent of $v_{s,hl}^{(y)}$ for any $s\neq t$ and is independent of loadings Λ_y. As a result,

$$\begin{aligned}
\mathbb{E}\left\|\sqrt{\frac{T}{N_y}}\sum_{j=1}^{N_y}(\Lambda_y)_j\frac{1}{|Q_{ij}^Y|}\sum_{t\in Q_{ij}^Y}v_{t,ij}^{(y)}\right\|^2 &= \sum_{p=1}^{k}\frac{T}{N_y}\mathbb{E}\left[\left(\sum_{j=1}^{N_y}(\Lambda_y)_{j,p}\frac{1}{|Q_{ij}^Y|}\sum_{t\in Q_{ij}^Y}v_{t,ij}^{(y)}\right)^2\right]\\
&= \sum_{p=1}^{k}\frac{T}{N_y}\sum_{j,l=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\frac{1}{|Q_{il}^Y|}\sum_{t\in Q_{ij}^Y}\sum_{s\in Q_{il}^Y}\mathbb{E}\Big[(\Lambda_y)_{j,p}(\Lambda_y)_{l,p}v_{t,ij}^{(y)}v_{s,il}^{(y)}\Big]\\
&= \sum_{p=1}^{k}\frac{T}{N_y}\sum_{j=1}^{N_y}\frac{1}{|Q_{ij}^Y|^2}\sum_{t\in Q_{ij}^Y}\mathbb{E}\left[(\Lambda_y)_{j,p}^2\right]\cdot\mathbb{E}\left[(v_{t,ij}^{(y)})^2\right]\\
&\le C.
\end{aligned}$$

Similarly, we can prove that $\mathbb{E}\left\|\sqrt{\frac{T}{N_x}}\sum_{j=1}^{N_x}(\Lambda_x)_j\frac{1}{|Q_{ii}^Y|}\sum_{t\in Q_{ii}^Y}\left((e_x)_{tj}(e_y)_{ti}-\mathbb{E}[(e_x)_{tj}(e_y)_{ti}]\right)\right\|^2\le C.$ □

4. Assumption G4.4 holds under Assumptions G2 and S1

Proof. In the following, we only prove that $\mathbb{E}\left\|\sqrt{\frac{T}{N_y^3}}\sum_{i,j=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\sum_{s\in Q_{ij}^Y}W_{ti}^Y(\Lambda_y)_j(e_y)_{ti}((e_y)_{si}(e_y)_{sj}-\mathbb{E}[(e_y)_{si}(e_y)_{sj}])\right\|^2\le C$ and other bounds can be proved similarly. As before, we define $v_{t,ij}^{(y)}=(e_y)_{ti}(e_y)_{tj}-\mathbb{E}[(e_y)_{ti}(e_y)_{tj}]$. It holds that

$$\begin{aligned}
&\mathbb{E}\left\|\sqrt{\frac{T}{N_y^3}}\sum_{i,j=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\sum_{s\in Q_{ij}^Y}W_{ti}^Y(\Lambda_y)_j(e_y)_{ti}v_{s,ij}^{(y)}\right\|^2\\
=&\sum_{p=1}^{k}\frac{T}{N_y^3}\sum_{i,j,h,l=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\frac{1}{|Q_{hl}^Y|}\sum_{s\in Q_{ij}^Y}\sum_{u\in Q_{hl}^Y}\mathbb{E}\left[W_{ti}^YW_{th}^Y(\Lambda_y)_{j,p}(\Lambda_y)_{l,p}(e_y)_{ti}(e_y)_{th}v_{s,ij}^{(y)}v_{u,hl}^{(y)}\right]\\
=&\sum_{p=1}^{k}\frac{T}{N_y^3}\sum_{i,j,h,l=1}^{N_y}\frac{1}{|Q_{ij}^Y|}\frac{1}{|Q_{hl}^Y|}\sum_{s\in Q_{ij}^Y\cap Q_{hl}^Y}\mathbb{E}\left[W_{ti}^YW_{th}^Y(\Lambda_y)_{j,p}(\Lambda_y)_{l,p}\right]\cdot\mathbb{E}\left[(e_y)_{ti}(e_y)_{th}v_{s,ij}^{(y)}v_{s,hl}^{(y)}\right].
\end{aligned}$$

Note that if the indices i,h,j,l take four different values, the RHS of the above equation will equal zero. Thus, the RHS of the above equation can be bounded by C. □

5. Assumption G4.5 holds under Assumptions G2 and S1

Proof. For any $i, j = 1, \cdots, N_y$, we let $\Delta_{F,ij} = \frac{1}{|Q_{ij}^Y|} \sum_{s \in Q_{ij}^Y} F_s F_s^\top - \frac{1}{T} \sum_{s=1}^T F_s F_s^\top$. It holds that

$$\begin{aligned}
\mathbb{E} \left\| \sqrt{\frac{T}{N_y}} \sum_{i=1}^{N_y} \Delta_{F,ij} W_{ti}^Y (\Lambda_y)_i (e_y)_{ti} \right\|^2 &= \frac{T}{N_y} \sum_{i,l=1}^{N_y} \mathbb{E} \left[W_{ti}^Y W_{tl}^Y (\Lambda_y)_i^\top (\Lambda_y)_l \Delta_{F,ij} \Delta_{F,lj} (e_y)_{ti} (e_y)_{tl} \right] \\
&= \frac{T}{N_y} \sum_{i=1}^{N_y} \mathbb{E} \Big[W_{ti}^Y (\Lambda_y)_i^\top (\Lambda_y)_i \Delta_{F,ij} \Delta_{F,lj} \Big] \cdot \mathbb{E} \left[(e_y)_{ti}^2 \right].
\end{aligned}$$

We will prove in part 8 that $\mathbb{E}\|\Delta_{F,ij}\|^2 \leq C/T$. Once this holds, we can bound the RHS of the above equation by C. We can prove other bounds following similar arguments. ☐

6. Assumption G4.6 holds under Assumptions G2, S1 and S2

Proof. Since factors $F_t \overset{i.i.d.}{\sim} (0, \Sigma_F)$, idiosyncratic errors $(e_y)_{ti} \overset{i.i.d.}{\sim} (0, \sigma_{e_y}^2)$, and they are independent of the observation pattern, by CLT we have $\frac{1}{|Q_{ij}^Y|^{1/2}} \sum_{t \in Q_{ij}^Y} F_t (e_y)_{ti} \xrightarrow{d} \mathcal{N}(0, \sigma_{e_y}^2 \Sigma_F)$. Therefore, by Assumption G2.2,

$$\frac{\sqrt{T}}{|Q_{ij}^Y|} \sum_{t \in Q_{ij}^Y} F_t (e_y)_{ti} \xrightarrow{d} \mathcal{N} \left(0, \frac{1}{q_{ij}} \sigma_{e_y}^2 \Sigma_F \right), \qquad \forall\, i, j = 1, \cdots, N_y.$$

Based on Slutsky's Theorem,

$$\frac{1}{N_x} \sum_{j=1}^{N_x} (\Lambda_x)_j (\Lambda_x)_j^\top \frac{\sqrt{T}}{|Q_{ii}^Y|} \sum_{t \in Q_{ii}^Y} F_t (e_y)_{ti} \xrightarrow{d} \mathcal{N} \left(0, \frac{1}{q_{ii}} \Sigma_{\Lambda_x} \sigma_{e_y}^2 \Sigma_F \Sigma_{\Lambda_x} \right).$$

For any i, j, l, we have the asymptotic covariance matrix

$$\begin{aligned}
\mathrm{ACov} \left(\frac{\sqrt{T}}{|Q_{ij}^Y|} \sum_{t \in Q_{ij}^Y} F_t (e_y)_{ti},\ \frac{\sqrt{T}}{|Q_{il}^Y|} \sum_{t \in Q_{il}^Y} F_t (e_y)_{ti} \right) &= \lim_{T \to \infty} \frac{T}{|Q_{ij}^Y| \cdot |Q_{il}^Y|} \sum_{t \in Q_{ij}^Y} \sum_{s \in Q_{il}^Y} \mathbb{E} \left[F_t F_s^\top (e_y)_{ti} (e_y)_{si} \right] \\
&= \frac{q_{ij,il}}{q_{ij} q_{il}} \sigma_{e_y}^2 \Sigma_F.
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{1}{N_y} \sum_{j=1}^{N_y} (\Lambda_y)_j (\Lambda_y)_j^\top \frac{\sqrt{T}}{|Q_{ij}^Y|} \sum_{t \in Q_{ij}^Y} F_t (e_y)_{ti} &= \mathbb{E} \left[(\Lambda_y)_j (\Lambda_y)_j^\top \right] \cdot \frac{1}{N_y} \sum_{j=1}^{N_y} \frac{\sqrt{T}}{|Q_{ij}^Y|} \sum_{t \in Q_{ij}^Y} F_t (e_y)_{ti} + o_p(1) \\
&\xrightarrow{d} \mathcal{N} \left(0, \lim_{N_y \to \infty} \frac{1}{N_y^2} \sum_{j,l=1}^{N_y} \frac{q_{ij,il}}{q_{ij} q_{il}} \Sigma_{\Lambda_y} \sigma_{e_y}^2 \Sigma_F \Sigma_{\Lambda_y} \right) \\
&= \mathcal{N} \left(0, \omega_i^{(2,3)} \Sigma_{\Lambda_y} \sigma_{e_y}^2 \Sigma_F \Sigma_{\Lambda_y} \right),
\end{aligned}$$

where $\omega_i^{(2,3)}$ is defined in Assumption S2.2. Furthermore, the two parts in Assumption G4.6 are jointly asymptotically normal with covariance matrix $\frac{1}{q_{ii}}\Sigma_{\Lambda_x}\sigma_{e_y}^2\Sigma_F\Sigma_{\Lambda_y}$. Combining these terms, $\Gamma_{\Lambda_y,i}^{(\gamma),\text{obs}}$ equals to

$$\Gamma_{\Lambda_y,i}^{(\gamma),\text{obs}} = \lim \frac{N_x^2}{(N_x+N_y)^2}\sigma_{e_y}^2\left[\frac{1}{q_{ii}}\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)\Sigma_F\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)+\left(\omega_i^{(2,3)}-\frac{1}{q_{ii}}\right)r^2\Sigma_{\Lambda_y}\Sigma_F\Sigma_{\Lambda_y}\right].$$

□

7. Assumption G4.7 holds under Assumptions G2, S1 and S2

Proof. By Assumptions S1, S2, and CLT, we have

$$\frac{1}{\sqrt{N_x}}\sum_{i=1}^{N_x}(\Lambda_x)_i(e_x)_{ti} \xrightarrow{d} \mathcal{N}(0,\sigma_{e_x}^2\Sigma_{\Lambda_x}),$$

and

$$\frac{1}{\sqrt{N_y}}\sum_{i=1}^{N_y}W_{ti}^Y(\Lambda_y)_i(e_y)_{ti} \xrightarrow{d} \mathcal{N}(0,\sigma_{e_y}^2\Sigma_{\Lambda_y,t}).$$

Furthermore, $\frac{1}{\sqrt{N_x}}\sum_{i=1}^{N_x}(\Lambda_x)_i(e_x)_{ti}$ and $\frac{1}{\sqrt{N_y}}\sum_{i=1}^{N_y}W_{ti}^Y(\Lambda_y)_i(e_y)_{ti}$ are asymptotically independent. So $\Gamma_{F,t}^{(\gamma),\text{obs}}$ defined in Assumption G4.7 simplifies to

$$\Gamma_{F,t}^{(\gamma),\text{obs}} = \lim \frac{N_x^2}{(N_x+N_y)^2}\left(\frac{N_y}{N_x}\sigma_{e_x}^2\Sigma_{\Lambda_x}+r^2\sigma_{e_y}^2\Sigma_{\Lambda_y,t}\right).$$

Suppose there is some weak factor F_w in Y whose loading $\sum_{i=1}^{N_y}(\Lambda_y)_{i,w}^2$ grows at the rate $g(N_y) = p_wN_y$, where p_w is defined in Assumption S1.2. For this weak factor F_w, p_w decays to 0 but is nonzero as N_y grows. We have $\frac{1}{\sqrt{g(N_y)}}\sum_{i=1}^{N_y}W_{ti}^Y(\Lambda_y)_{i,w}(e_y)_{ti} \xrightarrow{d} \mathcal{N}(0,\sigma_{e_y}^2\Sigma_{\Lambda_y,t,w})$. Then, there is $\Gamma_{F_w,t}^{(\gamma),\text{obs}} = \lim\frac{N_x^2}{(N_x+N_y)^2}\left(\frac{N_w}{N_x}\sigma_{e_x}^2\Sigma_{\Lambda_x}+r^2\frac{p_wN_w}{N_y}\sigma_{e_y}^2\Sigma_{\Lambda_y,t,w}\right)$, where $N_w = \min(N_y^2/g(N_y),N_x)$. □

8. Assumption G4.8 holds under Assumptions G2, S1 and S2

Proof. It suffices to show that $\left(\sqrt{T}\cdot\text{vec}(X_i^{(\gamma)}),\sqrt{T}\cdot\text{vec}(\mathbf{X}_t^{(\gamma)})\right)$ is asymptotically normal.

Step 1 – $\sqrt{T}\cdot vec(X_i^{(\gamma)})$ is asymptotically normal

Observe that $X_i^{(\gamma)} = \frac{1}{N_x+N_y}\sum_{j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}(\Lambda_j^{(\gamma)})^\top\left(\frac{1}{|Q_{ij}^Z|}\sum_{t\in Q_{ij}^Z}F_tF_t^\top-\frac{1}{T}\sum_{t=1}^TF_tF_t^\top\right)$, where $\Lambda^{(\gamma)} = [\Lambda_x;\sqrt{\gamma}\Lambda_y]\in\mathbb{R}^{(N_x+N_y)\times k}$ is the combined loadings. For any $k\times k$ matrices A and B, there is $\mathrm{vec}(AB) = (I_k \otimes A) \cdot \mathrm{vec}(B)$. Therefore, $\sqrt{T} \cdot \mathrm{vec}(X_i^{(\gamma)})$ can be written as

$$\sqrt{T} \cdot \mathrm{vec}(X_i^{(\gamma)}) = \frac{1}{N_x + N_y} \sum_{j=1}^{N_x+N_y} \left(I_k \otimes \Lambda_j^{(\gamma)} (\Lambda_j^{(\gamma)})^\top \right) \mathrm{vec} \left(\frac{\sqrt{T}}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} F_t F_t^\top - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t F_t^\top \right).$$

To simplify notation, we let $v_{ij} = \mathrm{vec}\Big(\frac{\sqrt{T}}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} F_t F_t^\top - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t F_t^\top \Big)$. Furthermore, we define $q_{ij}^Z = \lim_{T\to\infty} |Q_{ij}^Z|/T$ and $q_{ij,hl}^Z = \lim_{T\to\infty} |Q_{ij}^Z \cap Q_{hl}^Z|/T$ for any i, j, h, l. According to CLT,

$$v_{ij} = \sqrt{T} \left(\frac{1}{|Q_{ij}^Z|} - \frac{1}{T} \right) \sum_{t \in Q_{ij}^Z} \mathrm{vec}(F_t F_t^\top) - \frac{1}{\sqrt{T}} \sum_{t \notin Q_{ij}^Z} \mathrm{vec}(F_t F_t^\top) \xrightarrow{d} \mathcal{N} \left(0, \left(\frac{1}{q_{ij}^Z} - 1 \right) \Xi_F \right),$$

where $\Xi_F = \mathrm{Var}(\mathrm{vec}(F_t F_t^\top))$. Additionally, the asymptotic covariance of v_{ij} and v_{hl} for any i, j, h, l can be calculated as

$$\begin{aligned}
&\mathrm{ACov}(v_{ij}, v_{hl}) \\
&= \mathrm{ACov}\left(\frac{\sqrt{T}}{|Q_{ij}^Z|} \sum_{t \in Q_{ij}^Z} \mathrm{vec}(F_t F_t^\top) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathrm{vec}(F_t F_t^\top), \frac{\sqrt{T}}{|Q_{hl}^Z|} \sum_{t \in Q_{hl}^Z} \mathrm{vec}(F_t F_t^\top) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathrm{vec}(F_t F_t^\top) \right) \\
&= \lim_{T\to\infty} \frac{T}{|Q_{ij}^Z| \cdot |Q_{hl}^Z|} \left| Q_{ij}^Z \cap Q_{hl}^Z \right| \cdot \mathrm{Var}(\mathrm{vec}(F_t F_t^\top)) - 2\mathrm{Var}(\mathrm{vec}(F_t F_t^\top)) + \mathrm{Var}(\mathrm{vec}(F_t F_t^\top)) \\
&= \left(\frac{q_{ij,hl}^Z}{q_{ij}^Z q_{hl}^Z} - 1 \right) \Xi_F.
\end{aligned}$$

Particularly, we have $\mathrm{ACov}(v_{ij}, v_{il}) = (q_{ij,il}^Z / (q_{ij}^Z q_{il}^Z) - 1)\Xi_F$. When $i = N_x + 1, \cdots, N_x + N_y$, if j or l is chosen from $1, \cdots, N_x$, then $q_{ij,il}^Z / (q_{ij}^Z q_{il}^Z) = 1/q_{i'i'}$ with $i' = i - N_x$; otherwise, $q_{ij,il}^Z / (q_{ij}^Z q_{il}^Z) = q_{i'j',i'l'}/(q_{i'j'} q_{i'l'})$, where $j' = j - N_x$ and $l' = l - N_x$. Let $u_{jl} = (I_k \otimes \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top}) \Xi_F (I_k \otimes \Lambda_l^{(\gamma)} \Lambda_l^{(\gamma)\top})$. Observe that u_{jl} is independent with u_{mn} for distinct j, l, m, n. Thus, for any $p, q = 1, \cdots, k^2$, we have

$$\mathbb{E} \left[\frac{1}{(N_x + N_y)^2} \sum_{j,l=1}^{N_x+N_y} \left(\frac{q_{ij,hl}^Z}{q_{ij}^Z q_{hl}^Z} - 1 \right) (u_{jl,pq} - \mathbb{E}[u_{jl,pq}]) \right]^2 = O\Big(\frac{1}{N_y} \Big).$$

When q_{ij} and $q_{ij,hl}$ are independent of $(\Lambda_x)_m(\Lambda_x)_m^\top$ and $(\Lambda_y)_m(\Lambda_y)_m^\top$ for any i,j,h,l,m,

$$\begin{aligned}
&\mathrm{ACov}\left(\sqrt{T}\cdot\mathrm{vec}(X_i^{(\gamma)}),\sqrt{T}\cdot\mathrm{vec}(X_i^{(\gamma)})\right)\\
=&\lim\frac{1}{(N_x+N_y)^2}\sum_{j,l=1}^{N_x+N_y}\left(I_k\otimes\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\right)\cdot\mathrm{Cov}(v_{ij},v_{il})\cdot\left(I_k\otimes\Lambda_l^{(\gamma)}\Lambda_l^{(\gamma)\top}\right)\\
=&\lim\frac{1}{(N_x+N_y)^2}\sum_{j,l=1}^{N_x+N_y}\left(\frac{q^Z_{ij,hl}}{q^Z_{ij}q^Z_{hl}}-1\right)\left(I_k\otimes\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\right)\Xi_F\left(I_k\otimes\Lambda_l^{(\gamma)}\Lambda_l^{(\gamma)\top}\right)\\
=&\lim\frac{N_x^2}{(N_x+N_y)^2}\Bigg[\left(\frac{1}{q_{i'i'}}-1\right)\left(I_k\otimes\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)\right)\Xi_F\left(I_k\otimes\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)\right)\\
&+\left(\omega_{i'}^{(2,3)}-\frac{1}{q_{i'i'}}\right)\left(I_k\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left(I_k\otimes r\Sigma_{\Lambda_y}\right)\Bigg],
\end{aligned}$$

where $r=\gamma\cdot N_y/N_x$ and $i'=i-N_x$. Therefore, for any $i=N_x+1,\cdots,N_x+N_y$,

$$\begin{aligned}
&\sqrt{T}\cdot\mathrm{vec}(X_i^{(\gamma)})\\
\xrightarrow{d}\ &\mathcal{N}\Bigg(0,\lim\frac{N_x^2}{(N_x+N_y)^2}\Bigg[\left(\frac{1}{q_{i'i'}}-1\right)\left(I_k\otimes\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)\right)\Xi_F\left(I_k\otimes\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)\right)\\
&\qquad+\left(\omega_{i'}^{(2,3)}-\frac{1}{q_{i'i'}}\right)\left(I_k\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left(I_k\otimes r\Sigma_{\Lambda_y}\right)\Bigg]\Bigg).
\end{aligned}$$

Step 2 – $\sqrt{T}\cdot vec(\mathbf{X}_t^{(\gamma)})$ is asymptotically normal

In Assumption G4.8, we have $\sqrt{T}\cdot\mathbf{X}_t^{(\gamma)}=\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\Big(\frac{\sqrt{T}}{|Q^Z_{ij}|}\sum_{t\in Q^Z_{ij}}F_tF_t^\top-\frac{1}{\sqrt{T}}\sum_{t=1}^T F_tF_t^\top\Big)W^Z_{ti}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}$, and its vectorized form can be written as

$$\sqrt{T}\cdot\mathrm{vec}(\mathbf{X}_t^{(\gamma)})=\frac{1}{(N_x+N_y)^2}\sum_{i,j=1}^{N_x+N_y}W^Z_{ti}\big(\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\otimes I_k\big)\big(I_k\otimes\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\big)v_{ij},$$

where v_{ij} is defined in the first step. With similar arguments as in Step 1, we can show the existence of the following limit

$$\begin{aligned}
\lim\frac{1}{(N_x+N_y)^4}\sum_{i,j,h,l=1}^{N_x+N_y}&\left(\frac{q^Z_{ij,hl}}{q^Z_{ij}q^Z_{hl}}-1\right)W^Z_{ti}W^Z_{tl}\left(\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\otimes I_k\right)\left(I_k\otimes\Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\right)\\
&\Xi_F\left(I_k\otimes\Lambda_h^{(\gamma)}\Lambda_h^{(\gamma)\top}\right)\left(\Lambda_l^{(\gamma)}\Lambda_l^{(\gamma)\top}\otimes I_k\right).
\end{aligned}$$

Observe that when $i,j=1,\cdots,N_x$ or $h,l=1,\cdots,N_x$, $q^Z_{ij,hl}/(q^Z_{ij}q^Z_{hl})=1$; When $i,h=1,\cdots,N_x$ and $j,l=N_x+1,\cdots,N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij}q^Z_{hl})=q_{j'l'}/(q_{j'j'}q_{l'l'})$ with $j'=j-N_x$ and $l'=l-N_x$; When $h=1,\cdots,N_x$ and $i,j,l=N_x+1,\cdots,N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij}q^Z_{hl})=q_{i'j',l'l'}/(q_{i'j'}q_{l'l'})$ with

$i' = i - N_x$; When $i,j,h,l = N_x+1,\cdots,N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij}q^Z_{hl}) = q_{i'j',h'l'}/(q_{i'j'}q_{h'l'})$. By symmetry, other cases of i,j,h,l can be considered similarly. Additionally, observe that for any $k\times k$ matrices A and B, there are $(A\otimes I_k)(I_k\otimes B) = A\otimes B$ and $(I_k\otimes A)(B\otimes I_k) = B\otimes A$. As a result, we have

$$
\begin{aligned}
&\mathrm{ACov}\left(\sqrt{T}\cdot \mathrm{vec}(\mathbf{X}_t^{(\gamma)}), \sqrt{T}\cdot \mathrm{vec}(\mathbf{X}_t^{(\gamma)})\right)\\
=&\lim \frac{1}{(N_x+N_y)^4}\sum_{i,j,l,h=1}^{N_x+N_y} W^Z_{ti}W^Z_{tl}\left(\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\otimes \Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\right)\mathrm{Cov}(v_{ij},v_{lh})\left(\Lambda_l^{(\gamma)}\Lambda_l^{(\gamma)\top}\otimes \Lambda_h^{(\gamma)}\Lambda_h^{(\gamma)\top}\right)\\
=&\lim \frac{1}{(N_x+N_y)^4}\sum_{i,j,h,l=1}^{N_x+N_y} \left(\frac{q^Z_{ij,hl}}{q^Z_{ij}q^Z_{hl}}-1\right)\left(W^Z_{ti}\Lambda_i^{(\gamma)}\Lambda_i^{(\gamma)\top}\otimes \Lambda_j^{(\gamma)}\Lambda_j^{(\gamma)\top}\right)\Xi_F\left(W^Z_{tl}\Lambda_l^{(\gamma)}\Lambda_l^{(\gamma)\top}\otimes \Lambda_h^{(\gamma)}\Lambda_h^{(\gamma)\top}\right)\\
=&\lim \frac{N_x^4}{(N_x+N_y)^4}\frac{1}{N_y^2}\sum_{i,j=1}^{N_y}\left(\frac{q_{ij}}{q_{ii}q_{jj}}-1\right)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+\lim \frac{N_x^4}{(N_x+N_y)^4}\frac{1}{N_y^3}\sum_{i,j,l=1}^{N_y}\left(\frac{q_{jj,il}}{q_{jj}q_{il}}-1\right)\Big[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)+\\
&\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Big]\\
&+\lim \frac{N_x^4}{(N_x+N_y)^4}\frac{1}{N_y^4}\sum_{i,j,l,h=1}^{N_y}\left(\frac{q_{il,jh}}{q_{jh}q_{il}}-1\right)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\\
=&\lim \frac{N_x^4}{(N_x+N_y)^4}\Bigg[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\Big((\omega^{(1)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(2)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)+\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\Big((\omega^{(2)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(3)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)\Bigg].
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\sqrt{T}\cdot\mathrm{vec}(\mathbf{X}_t^{(\gamma)})\xrightarrow{d}\\
&\mathcal{N}\Bigg(0,\lim \frac{N_x^4}{(N_x+N_y)^4}\Bigg[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\Big((\omega^{(1)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(2)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)+\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\Big((\omega^{(2)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(3)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)\Bigg]\Bigg).
\end{aligned}
$$

Step 3 – $\sqrt{T}\cdot vec(X_i^{(\gamma)})$ and $\sqrt{T}\cdot vec(\mathbf{X}_t^{(\gamma)})$ are jointly asymptotically normal

We only consider the case where $i = N_x+1,\cdots,N_x+N_y$. It is easy to see that $\sqrt{T}\cdot\mathrm{vec}(X_i^{(\gamma)})$ and $\sqrt{T}\cdot\mathrm{vec}(\mathbf{X}_t^{(\gamma)})$ are jointly asymptotically normal since both of their randomnesses come from

v_{ij}. By similar arguments, we can show the existence of the following limit

$$\lim \frac{1}{(N_x+N_y)^3} \sum_{j,l,h=1}^{N_x+N_y} \left(\frac{q^Z_{ij,hl}}{q^Z_{ij} q^Z_{hl}} - 1 \right) W^Z_{tl} \left(\Lambda_l^{(\gamma)} \Lambda_l^{(\gamma)\top} \otimes I_k \right) \left(I_k \otimes \Lambda_h^{(\gamma)} \Lambda_h^{(\gamma)\top} \right) \Xi_F \left(I_k \otimes \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \right).$$

Observe that when $h, l = 1, \cdots, N_x$, $q^Z_{ij,hl}/(q^Z_{ij} q^Z_{hl}) = 1$; When $j, h = 1, \cdots, N_x, l = N_x+1, \cdots, N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij} q^Z_{hl}) = q_{i'l'}/(q_{i'i'} q_{l'l'})$ with $i' = i - N_x$ and $l' = l - N_x$; When $h = 1, \cdots, N_x, j, l = N_x+1, \cdots, N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij} q^Z_{hl}) = q_{i'j',l'l'}/(q_{i'j'} q_{l'l'})$ with $j' = j - N_x$; When $j, h, l = N_x+1, \cdots, N_x+N_y$, $q^Z_{ij,hl}/(q^Z_{ij} q^Z_{hl}) = q_{i'j',h'l'}/(q_{i'j'} q_{h'l'})$, where $h' = h - N_x$. Other cases can be similarly considered. As a result, we have

$$\begin{aligned}
&\mathrm{ACov}\left(\sqrt{T} \cdot \mathrm{vec}(\mathbf{X}_t^{(\gamma)}), \sqrt{T} \cdot \mathrm{vec}(X_i^{(\gamma)})\right) \\
=& \lim \frac{1}{(N_x+N_y)^3} \sum_{j,h,l=1}^{N_x+N_y} W^Z_{tl} \left(\Lambda_l^{(\gamma)} \Lambda_l^{(\gamma)\top} \otimes I_k \right) \left(I_k \otimes \Lambda_h^{(\gamma)} \Lambda_h^{(\gamma)\top} \right) \mathrm{Cov}(v_{lh}, v_{ij}) \left(I_k \otimes \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \right) \\
=& \lim \frac{1}{(N_x+N_y)^3} \sum_{j,l,h=1}^{N_x+N_y} \left(\frac{q^Z_{ij,hl}}{q^Z_{ij} q^Z_{hl}} - 1 \right) \left(W^Z_{tl} \Lambda_l^{(\gamma)} \Lambda_l^{(\gamma)\top} \otimes I_k \right) \left(I_k \otimes \Lambda_h^{(\gamma)} \Lambda_h^{(\gamma)\top} \right) \Xi_F \left(I_k \otimes \Lambda_j^{(\gamma)} \Lambda_j^{(\gamma)\top} \right) \\
=& \lim \frac{N_x^3}{(N_x+N_y)^3} \Big[\left(\Sigma_{\Lambda_x} \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x} \right) \Xi_F \left((\omega_{i'}^{(1)} - 1)(I_k \otimes \Sigma_{\Lambda_x}) + (\omega_{i'}^{(2,2)} - 1)\big(I_k \otimes r\Sigma_{\Lambda_y}\big) \right) \\
&\quad + \left(r\Sigma_{\Lambda_y,t} \otimes r\Sigma_{\Lambda_y} \right) \Xi_F \left((\omega_{i'}^{(2,1)} - 1)(I_k \otimes \Sigma_{\Lambda_x}) + (\omega_{i'}^{(3)} - 1)\big(I_k \otimes r\Sigma_{\Lambda_y}\big) \right) \Big],
\end{aligned}$$

where $\omega_i^{(1)}, \omega_i^{(2,1)}, \omega_i^{(2,2)}$ and $\omega_i^{(3)}$ are defined in Assumption S2. For the special case where $\omega_i^{(1)} = \omega_i^{(2,1)} = \omega_i^{(2,2)} = \omega_i^{(3)} = \omega_i$, the asymptotic covariance matrix is simplified to

$$\begin{aligned}
&\mathrm{ACov}\left(\sqrt{T} \cdot \mathrm{vec}(\mathbf{X}_t^{(\gamma)}), \sqrt{T} \cdot \mathrm{vec}(X_i^{(\gamma)})\right) \\
=& (\omega_{i'} - 1) \cdot \lim \frac{N_x^3}{(N_x+N_y)^3} \left((\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t}) \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x} \right) \Xi_F \left(I_k \otimes \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y} \right) \right).
\end{aligned}$$

□

IA.C.4 Proof of Corollary 1

Proof. According to Proposition 5, Theorem 2 holds under the simplified assumptions in Corollary 1. As a result, we just need to calculate the asymptotic variances in Theorem 2 under the simplified model.

1. The asymptotic variance of loadings:

We have $\Sigma_\Lambda^{(\gamma)} = \lim \frac{N_x}{N_x+N_y} \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y} \right)$ and $\Sigma_{\Lambda,t}^{(\gamma)} = \lim \frac{N_x}{N_x+N_y} \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t} \right)$, where $r =$

$\gamma \cdot N_y/N_x$. According to the proof of Proposition 5, $\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{obs}}$ defined in Assumption G4.6 equals to

$$\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{obs}} = \lim \frac{N_x^2}{(N_x+N_y)^2} \cdot \sigma_{e_y}^2 \left[\frac{1}{q_{ii}} \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right) \Sigma_F \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right) + \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right) r^2 \Sigma_{\Lambda_y} \Sigma_F \Sigma_{\Lambda_y} \right].$$

As a result, the first part of $\Sigma_{\Lambda_y,i}^{(\gamma)}$ is

$$\begin{aligned}
&\Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{obs}}(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1} \\
=&\frac{1}{q_{ii}}\sigma_{e_y}^2\Sigma_F^{-1} + \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right)\sigma_{e_y}^2 r^2 \Sigma_F^{-1}\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\Sigma_{\Lambda_y}\Sigma_F\Sigma_{\Lambda_y}\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\Sigma_F^{-1}.
\end{aligned}$$

Next, consider $\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{miss}} = h_{i+N_x}^{(\gamma)}((\Lambda_y)_i)$ defined in Theorem 2.1. From the proof of Proposition 5, for any $i = 1, \cdots, N_y$,

$$\begin{aligned}
&\sqrt{T} \cdot \mathrm{vec}(X_{i+N_x}^{(\gamma)}) \\
\xrightarrow{d}\ & \mathcal{N}\Bigg(0, \lim \frac{N_x^2}{(N_x+N_y)^2}\Bigg[\left(\frac{1}{q_{ii}} - 1\right)\left(I_k \otimes \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)\right)\Xi_F\left(I_k \otimes \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)\right) \\
&\qquad + \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right)\left(I_k \otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left(I_k \otimes r\Sigma_{\Lambda_y}\right)\Bigg]\Bigg),
\end{aligned}$$

where $\Xi_F = \mathrm{Var}(\mathrm{vec}(F_tF_t^\top))$. Therefore, $\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{miss}}$ can be calculated as

$$\begin{aligned}
\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{miss}} =& \lim((\Lambda_y)_i^\top \otimes I_k)\mathbb{E}\left[T \cdot \mathrm{vec}(X_{i+N_x})\mathrm{vec}(X_{i+N_x})^\top\right]((\Lambda_y)_i \otimes I_k) \\
=& \lim \frac{N_x^2}{(N_x+N_y)^2}\Bigg[\left(\frac{1}{q_{ii}} - 1\right)\left((\Lambda_y)_i^\top \otimes \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)\right)\Xi_F\left((\Lambda_y)_i \otimes \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)\right) \\
&\qquad + \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right)\left((\Lambda_y)_i^\top \otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left((\Lambda_y)_i \otimes r\Sigma_{\Lambda_y}\right)\Bigg].
\end{aligned}$$

Thus, the second part of $\Sigma_{\Lambda_y,i}^{(\gamma)}$ is

$$\begin{aligned}
&\Sigma_F^{-1}(\Sigma_\Lambda^{(\gamma)})^{-1}\Gamma_{\Lambda_y,i}^{(\gamma),\mathrm{miss}}(\Sigma_\Lambda^{(\gamma)})^{-1}\Sigma_F^{-1} \\
=&\left(\frac{1}{q_{ii}} - 1\right)\Sigma_F^{-1}((\Lambda_y)_i^\top \otimes I_k)\Xi_F\left((\Lambda_y)_i \otimes I_k\right)\Sigma_F^{-1} \\
&+ \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right)\Sigma_F^{-1}\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\left((\Lambda_y)_i^\top \otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left((\Lambda_y)_i \otimes r\Sigma_{\Lambda_y}\right)\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\Sigma_F^{-1}.
\end{aligned}$$

Combining the two parts, we can deduce that

$$\sqrt{T}(\Sigma_{\Lambda_y,i}^{(\gamma)})^{-1/2}\left((H^{(\gamma)})^{-1}(\tilde{\Lambda}_y)_i - (\Lambda_y)_i\right) \xrightarrow{d} \mathcal{N}\left(0, I_k\right),$$

where

$$\Sigma^{(\gamma)}_{\Lambda_y,i} = \frac{1}{q_{ii}}\sigma^2_{e_y}\Sigma_F^{-1} + \left(\frac{1}{q_{ii}} - 1\right)\Sigma_F^{-1}((\Lambda_y)_i^\top \otimes I_k)\Xi_F((\Lambda_y)_i \otimes I_k)\Sigma_F^{-1} + \left(\omega_i^{(2,3)} - \frac{1}{q_{ii}}\right)\Sigma_F^{-1}$$
$$\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\left[\sigma^2_{e_y} r^2 \Sigma_{\Lambda_y}\Sigma_F\Sigma_{\Lambda_y} + \left((\Lambda_y)_i^\top \otimes r\Sigma_{\Lambda_y}\right)\Xi_F\left((\Lambda_y)_i \otimes r\Sigma_{\Lambda_y}\right)\right]\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1}\Sigma_F^{-1}.$$

2. The asymptotic variance of factors:
   Based on the proof of Proposition 5, we have

$$\Gamma^{(\gamma),\mathrm{obs}}_{F,t} = \lim \frac{N_x^2}{(N_x+N_y)^2}\left(\frac{N_y}{N_x}\sigma^2_{e_x}\Sigma_{\Lambda_x} + r^2\sigma^2_{e_y}\Sigma_{\Lambda_y,t}\right).$$

As a result, the first part of $\Sigma^{(\gamma)}_{F,t}$ can be calculated as

$$(\Sigma^{(\gamma)}_{\Lambda,t})^{-1}\Gamma^{(\gamma),\mathrm{obs}}_{F,t}(\Sigma^{(\gamma)}_{\Lambda,t})^{-1} = \lim\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t}\right)^{-1}\left(\frac{N_y}{N_x}\sigma^2_{e_x}\Sigma_{\Lambda_x} + r^2\sigma^2_{e_y}\Sigma_{\Lambda_y,t}\right)\left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t}\right)^{-1}.$$

Consider the second part where $\Gamma^{(\gamma),\mathrm{miss}}_{F,t} = g_t^{(\gamma)}\left((\Sigma^{(\gamma)}_\Lambda)^{-1}\Sigma_F^{-1}F_t\right)$. We have proved in Proposition 5

$$\begin{aligned}
&\sqrt{T}\cdot\mathrm{vec}(\mathbf{X}^{(\gamma)}_t) \xrightarrow{d} \\
&\mathcal{N}\Bigg(0, \lim\frac{N_x^4}{(N_x+N_y)^4}\Bigg[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\Big((\omega^{(1)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(2)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big) + \left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\Big((\omega^{(2)}-1)\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\\
&+(\omega^{(3)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)\Bigg]\Bigg).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\Gamma^{(\gamma),\mathrm{miss}}_{F,t}\\
&=\lim\left(I_k\otimes F_t^\top\Sigma_F^{-1}(\Sigma^{(\gamma)}_\Lambda)^{-1}\right)\mathbb{E}\left[T\cdot\mathrm{vec}(\mathbf{X}^{(\gamma)}_t)\mathrm{vec}(\mathbf{X}^{(\gamma)}_t)^\top\right]\left(I_k\otimes(\Sigma^{(\gamma)}_\Lambda)^{-1}\Sigma_F^{-1}F_t\right)\\
&=\lim\frac{N_x^2}{(N_x+N_y)^2}\left(I_k\otimes F_t^\top\Sigma_F^{-1}(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y})^{-1}\right)\Bigg[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\Big((\omega^{(1)}-1)\\
&\quad\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right) + (\omega^{(2)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big) + \left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Xi_F\Big((\omega^{(2)}-1)\\
&\quad\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right) + (\omega^{(3)}-1)\left(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y}\right)\Big)\Bigg]\left(I_k\otimes(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y})^{-1}\Sigma_F^{-1}F_t\right).
\end{aligned}$$

If all the factors in F_y are strong factors in Y, combining the two parts, we have

$$\sqrt{\delta_{N_y,T}}(\Sigma^{(\gamma)}_{F,t})^{-1/2}\left(H^{(\gamma)\top}\tilde{F}_t - F_t\right) \xrightarrow{d} \mathcal{N}(0, I_k),$$

where

$$\Sigma_{F,t}^{(\gamma)} = \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t}\right)^{-1} \left[\frac{\delta_{N_yT}}{N_y} \left(\frac{N_y}{N_x} \sigma_{e_x}^2 \Sigma_{\Lambda_x} + r^2 \sigma_{e_y}^2 \Sigma_{\Lambda_y,t} \right) + \frac{\delta_{N_yT}}{T} \left(I_k \otimes F_t^\top \Sigma_F^{-1} \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y}\right)^{-1} \right) \right.$$
$$\left[\left(\Sigma_{\Lambda_x} \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x}\right) \Xi_F \Big((\omega^{(1)} - 1) \left(\Sigma_{\Lambda_x} \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x}\right) + (\omega^{(2)} - 1) \right.$$
$$\left(r\Sigma_{\Lambda_y,t} \otimes r\Sigma_{\Lambda_y}\right) \Big) + \left(r\Sigma_{\Lambda_y,t} \otimes r\Sigma_{\Lambda_y}\right) \Xi_F \Big((\omega^{(2)} - 1) \left(\Sigma_{\Lambda_x} \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x}\right) + (\omega^{(3)} - 1)$$
$$\left. \left. \left(r\Sigma_{\Lambda_y,t} \otimes r\Sigma_{\Lambda_y}\right) \Big) \right] \left(I_k \otimes (\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y})^{-1} \Sigma_F^{-1} F_t \right) \right] \left(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t}\right)^{-1}.$$

If some factor F_w is weak in Y, then the asymptotic distribution of the estimation of this weak factor is

$$\sqrt{\delta_{N_w,T}} (\Sigma_{F_w,t}^{(\gamma)})^{-1/2} \left((H^{(\gamma)\top} \tilde{F}_t)_w - F_{t,w} \right) \xrightarrow{d} \mathcal{N}(0, I_k).$$

We have

$$\begin{aligned} \Sigma_{F_w,t}^{(\gamma)} =& (\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t})_w^{-1} \left[\frac{\delta_{N_w,T}}{N_w} \left(\frac{N_w}{N_x} \sigma_{e_x}^2 \Sigma_{\Lambda_x,w} + r^2 \frac{p_w N_w}{N_y} \sigma_{e_y}^2 \Sigma_{\Lambda_y,t,w} \right) \right] (\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t})_w^{-1} \\ &+ \frac{\delta_{N_w,T}}{T} \cdot \Sigma_{F_w,t}^{(\gamma),miss}, \end{aligned}$$

where $N_w = \min(N_y/p_w, N_x)$, $\frac{1}{N_y p_w} \sum_{i=1}^{N_y} W_{ti}^Y (\Lambda_y)_{i,w} (\Lambda_y)_{i,w}^\top \xrightarrow{p} \mathcal{N}(0, \Sigma_{\Lambda_y,t,w})$, $(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t})_w^{-1}$, $\Sigma_{\Lambda_x,w}$ and $\Sigma_{F_w,t}^{(\gamma),\text{miss}}$ are respectively the diagonal block in $(\Sigma_{\Lambda_x} + r\Sigma_{\Lambda_y,t})^{-1}$, Σ_{Λ_x} and $\Sigma_{F,t}^{(\gamma),\text{miss}}$ corresponding to the weak factors.

3. The asymptotic variance of common components:

For $\Gamma_{\Lambda_y,F,i,t}^{(\gamma),\text{miss,cov}} = g_{i,t}^{(\gamma),\text{cov}}\left((\Lambda_y)_i, (\Sigma_\Lambda^{(\gamma)})^{-1} \Sigma_F^{-1} F_t\right)$, define $\Psi_{i,t}^{(\gamma),\text{cov}} = \lim \mathbb{E}\left[T \cdot \text{vec}(X_{i+N_x}^{(\gamma)}) \text{vec}(\mathbf{X}_t^{(\gamma)})^\top \right]$ with $i = 1, \cdots, N_y$. It holds that $\Gamma_{\Lambda_y,F,i,t}^{(\gamma),\text{miss,cov}} = \left(I_k \otimes F_t^\top \Sigma_F^{-1} \left(\Sigma_\Lambda^{(\gamma)}\right)^{-1} \right) \Psi_{i,t}^{(\gamma),\text{cov}} ((\Lambda_y)_i \otimes I_k)$. For general $\omega_i^{(1)}, \omega_i^{(2,1)}, \omega_i^{(2,2)}$ and $\omega_i^{(3)}$, we have

$$\begin{aligned} \Psi_{i,t}^{(\gamma),\text{cov}} =& \lim \frac{N_x^3}{(N_x + N_y)^3} \Big[\left(\Sigma_{\Lambda_x} \otimes r\Sigma_{\Lambda_y} + r\Sigma_{\Lambda_y,t} \otimes \Sigma_{\Lambda_x}\right) \Xi_F \Big((\omega_i^{(1)} - 1)(I_k \otimes \Sigma_{\Lambda_x}) + (\omega_i^{(2,2)} - 1) \\ & (I_k \otimes r\Sigma_{\Lambda_y}) \Big) + (r\Sigma_{\Lambda_y,t} \otimes r\Sigma_{\Lambda_y}) \Xi_F \Big((\omega_i^{(2,1)} - 1)(I_k \otimes \Sigma_{\Lambda_x}) + (\omega_i^{(3)} - 1)(I_k \otimes r\Sigma_{\Lambda_y}) \Big) \Big]. \end{aligned}$$

As a result, there is

$$
\begin{aligned}
&\Sigma^{(\gamma),\text{miss,cov}}_{\Lambda_y,F,i,t}\\
=&(\Sigma^{(\gamma)}_{\Lambda,t})^{-1}\Gamma^{(\gamma),\text{miss,cov}}_{\Lambda_y,F,i,t}(\Sigma^{(\gamma)}_{\Lambda})^{-1}\Sigma_F^{-1}\\
=&\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y,t}\right)^{-1}\left(I_k\otimes F_t^\top\Sigma_F^{-1}\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)^{-1}\right)\Big[\left(\Sigma_{\Lambda_x}\otimes r\Sigma_{\Lambda_y}+r\Sigma_{\Lambda_y,t}\otimes\Sigma_{\Lambda_x}\right)\Xi_F\\
&\left((\omega_i^{(1)}-1)((\Lambda_y)_i\otimes\Sigma_{\Lambda_x})+(\omega_i^{(2,2)}-1)((\Lambda_y)_i\otimes r\Sigma_{\Lambda_y})\right)+(r\Sigma_{\Lambda_y,t}\otimes r\Sigma_{\Lambda_y})\Xi_F\\
&\left((\omega_i^{(2,1)}-1)((\Lambda_y)_i\otimes\Sigma_{\Lambda_x})+(\omega_i^{(3)}-1)\left((\Lambda_y)_i\otimes r\Sigma_{\Lambda_y}\right)\right)\Big]\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)^{-1}\Sigma_F^{-1}.
\end{aligned}
$$

In the special case where $\omega_i^{(1)}=\omega_i^{(2,1)}=\omega_i^{(2,2)}=\omega_i^{(3)}=\omega_i$, $\Sigma^{(\gamma),\text{miss,cov}}_{\Lambda_y,F,i,t}$ can be simplified as

$$
\begin{aligned}
\Sigma^{(\gamma),\text{miss,cov}}_{\Lambda_y,F,i,t}=&(\omega_i-1)r\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y,t}\right)^{-1}\Big[\left(I_k\otimes F_t^\top\Sigma_F^{-1}\left(\Sigma_{\Lambda_x}+r\Sigma_{\Lambda_y}\right)^{-1}\right)(\Sigma_{\Lambda_x}\otimes\Sigma_{\Lambda_y})+\\
&\Sigma_{\Lambda_y,t}\otimes(F_t^\top\Sigma_F^{-1})\Big]\Xi_F\left((\Lambda_y)_i\otimes I_k\right)\Sigma_F^{-1}.
\end{aligned}
$$

This completes the proof. □

IA.C.5 Proof of Proposition 1

Proof. We prove Proposition 1 as a special case of the general Theorem 1. In the following, we prove that the assumptions in Proposition 1 imply the general model specified by Assumptions G1, G2, G3 and G4.

Under the data generating process described in Section 3.1, the first factor is a strong factor in target Y and is not contained in auxiliary panel X, so Assumption G1 holds. Since there is no missing observation in Y, Assumption G2 automatically holds. The factors $F_t\overset{i.i.d.}{\sim}(0,\Sigma_F)$, where Σ_F is a diagonal matrix with diagonal elements equal to σ_F^2. The loadings in X follow

$$
\frac{1}{N_x}\sum_{i=1}^{N_y}(\Lambda_x)_i(\Lambda_x)_i^\top\xrightarrow{p}\Sigma_{\Lambda_x}=\begin{bmatrix}0&0\\0&\sigma^2_{\Lambda_x}\end{bmatrix}.
$$

And for loadings in Y, we have

$$
\frac{1}{N_y}\sum_{i=1}^{N_y}(\Lambda_y)_i(\Lambda_y)_i^\top\xrightarrow{p}\Sigma_{\Lambda_y}=\begin{bmatrix}\sigma^2_{\Lambda_y}&0\\0&0\end{bmatrix}.
$$

Thus, $\Sigma_{\Lambda_x}+\Sigma_{\Lambda_y}$ is a positive definite matrix. Additionally, the factors and loadings have bounded fourth moments. These conditions imply Assumptions G3.1 and G3.2. In the data generating process in Section 3.1, the idiosyncratic errors have bounded eighth moments and are drawn i.i.d. from $(e_x)_{ti}\overset{i.i.d.}{\sim}(0,\sigma^2_{e_x})$ and $(e_y)_{ti}\overset{i.i.d.}{\sim}(0,\sigma^2_{e_y})$. As a result, Assumptions G3.3 and G3.4 hold. It is straightforward to show that the moment conditions in Assumption G4 hold under this model. □

IA.C.6 Proof of Proposition 2

Proof. The data generating process and observation pattern in Section 3.2 are a special case of the simplified factor model in Assumptions S1 and S2.

Since factors, loadings, and idiosyncratic errors are i.i.d. distributed, we automatically obtain Assumption S1. Under the missing-at-random observation pattern with observed probability p, we can show that the quantities in Assumption S2 satisfy $\omega_i^{(2,3)} = 1/p$, $\omega_i^{(1)} = \omega_i^{(2,1)} = \omega_i^{(2,2)} = \omega_i^{(3)} = 1$, and $\omega^{(1)} = \omega^{(2)} = \omega^{(3)} = 1$. Plugging these parameters into the asymptotic variance of the estimated common components of Y in Corollary 1.3, we have

$$\begin{aligned}\Sigma_{C,ti}^{(\gamma)} =& \frac{\delta_{N_y T}}{T}\frac{\sigma_{e_y}^2}{p\sigma_F^2}F_t^2 + \frac{\delta_{N_y T}}{T}\left(\frac{1}{p}-1\right)\sigma_F^{-4}\mathrm{Var}(F_t^2)(\Lambda_y)_i^2 F_t^2 \\ &+ \frac{\delta_{N_y T}}{N_y}(\Lambda_y)_i^2\left(\sigma_{\Lambda_x}^2 + \gamma\frac{N_y}{N_x}p\sigma_{\Lambda_y}^2\right)^{-2}\left(\frac{N_y}{N_x}\sigma_{\Lambda_x}^2\sigma_{e_x}^2 + \gamma^2\frac{N_y^2}{N_x^2}p\sigma_{\Lambda_y}^2\sigma_{e_y}^2\right).\end{aligned}$$

The partial derivative of $\Sigma_{C,ti}^{(\gamma)}$ with respect to γ equals to

$$\begin{aligned}\frac{\partial\Sigma_{C,ti}^{(\gamma)}}{\partial\gamma} =& \frac{\delta_{N_y T}}{N_y}(\Lambda_y)_i^2\left(\sigma_{\Lambda_x}^2 + \gamma\frac{N_y}{N_x}p\sigma_{\Lambda_y}^2\right)^{-2}\Bigg[-2\frac{N_y}{N_x}p\sigma_{\Lambda_y}^2\left(\sigma_{\Lambda_x}^2 + \gamma\frac{N_y}{N_x}p\sigma_{\Lambda_y}^2\right)^{-1} \\ &\cdot\left(\frac{N_y}{N_x}\sigma_{\Lambda_x}^2\sigma_{e_x}^2 + \gamma^2\frac{N_y^2}{N_x^2}p\sigma_{\Lambda_y}^2\sigma_{e_y}^2\right) + 2\gamma\frac{N_y^2}{N_x^2}p\sigma_{\Lambda_y}^2\sigma_{e_y}^2\Bigg].\end{aligned}$$

We let $\partial\Sigma_{C,ti}^{(\gamma)}/\partial\gamma = 0$ and obtain $\gamma^* = \sigma_{e_x}^2/\sigma_{e_y}^2$. Furthermore, $\partial^2\Sigma_{C,ti}^{(\gamma^*)}/\partial\gamma^2 > 0$. As a result, the optimized γ that minimizes $\Sigma_{C,ti}^{(\gamma)}$ is $\gamma^* = \sigma_{e_x}^2/\sigma_{e_y}^2$ for any i and t. This completes the proof. □

IA.C.7 Proof of Proposition 3

Proof. If we select $\gamma = r$ for some constant r, then target-PCA degenerates to the PCA estimator in Xiong and Pelger (2023) applied only to auxiliary data X. Therefore, when all the factors can be identified in X, Theorem 1 holds with convergence rate $\delta_{N_x,T} = \min(N_x, T)$, Theorem 2.2 holds with convergence rate $\sqrt{\delta_{N_x,T}}$, and the asymptotic variance is independent of Y. □

IA.C.8 Proof of Proposition 4

Proof. The proof of Proposition 4 is analogous to the proof of Theorems 1 and 2 and follows the same arguments. Let $\delta_{N_x,T,M} = \min(\delta_{N_x,T}, M)$. When $M \to \infty$ and $\gamma = r \cdot N_x$ for some constant r, the upper bound for $(N_x+N_y)^{-2} \sum_{i,j=1}^{N_x+N_y} \gamma^2(i,j)$ in Lemma 1 changes to $C/\delta_{N_x,T,M}$, while Lemmas 2 and 3 remain the same. As a result, Theorem 1 holds with convergence rate $\delta_{N_x,T,M}$. For Theorem 2, we replace the convergence rate $\delta_{N_y,T}$ in Lemma 6 by $\delta_{N_x,T,M}$. With modified assumptions, we can prove the statements of Theorem 2.2 with convergence rate $\sqrt{\delta_{N_x,T,M}}$. □

IA.C.9 Proof of Proposition 6

Proof. The proof of Proposition 6 is analogous to the proof of Theorem 2. Specifically, when the number of time periods, for which any two and four units are observed, is proportional to T^α, the rates based on T in Lemmas 1, 5 and 6 will simply be replaced by T^α, while Lemmas 2, 3 and 4 continue to hold. This will change the convergence rate in Theorem 1 to $\delta_{N_y,T^\alpha} = \min(N_y, T^\alpha)$ and the convergence rate in Theorem 2 to $\sqrt{T^\alpha}$ for the loadings of Y, and $\sqrt{\delta_{N_y,T^\alpha}}$ for the factors and common components of Y. □

References

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

McCracken, M. W., and S. Ng (2016): "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, 34(4), 574–589.

Xiong, R., and M. Pelger (2023): "Large dimensional latent factor modeling with missing observations and applications to causal inference," *Journal of Econometrics*, 233(1), 271–301.