MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Decomposition of Bilateral Trade Flows Using a Three-Dimensional Panel Data Model

Yufeng Mao, Bin Peng, Mervyn Silvapulle, Param Silvapulle and Yanrong Yang

May 2021

Working Paper 07/21

Decomposition of Bilateral Trade Flows Using a Three-Dimensional Panel Data Model[1]

Yufeng Mao♯, Bin Peng♯, Mervyn Silvapulle♯, Param Silvapulle♯
and Yanrong Yang*

♯Monash University and *Australian National University

Abstract

This study decomposes the bilateral trade flows using a three-dimensional panel data model. Under the scenario that all three dimensions diverge to infinity, we propose an estimation approach to identify the number of global shocks and country-specific shocks sequentially, and establish the asymptotic theories accordingly. From the practical point of view, being able to separate the pervasive and nonpervasive shocks in a multi-dimensional panel data is crucial for a range of applications, such as, international financial linkages, migration flows, etc. In the numerical studies, we first conduct intensive simulations to examine the theoretical findings, and then use the proposed approach to investigate the international trade flows from two major trading groups (APEC and EU) over 1982-2019, and quantify the network of bilateral trade.

Keywords: Three-Dimensional Panel Data, Bilateral Trade, Asymptotic Theory

JEL classification: C23, P45

[1]Peng acknowledges the Australian Research Council Discovery Grants Program for its financial support under Grant Number DP210100476.

1 Introduction

All countries of the world are nowadays connected with each other more or less through varieties of bilateral trade. Getting reliable and up-to-date statistics on exports and imports of different countries is thus crucial in order to provide a detailed insight into the most recent trading patterns. Given an increasing interest in understanding such a complex network, we see the rising popularity of multi-dimensional models over the past decade, e.g., Moench et al. (2013), Beck et al. (2016), Andreou et al. (2019), Choi et al. (2020), Kapetanios et al. (2020), just to name a few. Excellent reviews on the applications and theoretical developments of multi-dimensional panel data models can be respectively seen in Baltagi et al. (2015) and Breitung and Eickmeier (2016) for instance.

Despite a vast amount of research on two dimensional factor models (see Bai and Ng, 2008 for an excellent review), it seems that the literature of multi-dimensional models has not even settled on how to effectively distinguish pervasive and nonpervasive economic shocks, where pervasive and nonpervasive shocks refer to those affecting the entire network and those affecting only a part of the network respectively (e.g., Wang, 2010; Ergemen and Rodríguez-Caballero, 2017). In this regard, the presentation (2.4) of Section 2 provides a clear visualization using matrix form.

To solve the aforementioned issue, different algorithms have been proposed (e.g., Breitung and Eickmeier, 2016 and references therein), but from the theoretical point of view the progress has not been pushed forward much since Wang (2010). We now comment on the relevant literature. Ergemen and Rodríguez-Caballero (2017) extend the study of Wang (2010) to allow for long run dependence, and both papers numerically rely on some initial estimates on the pervasive and nonpervasive factors. In our view, the requirement on initial estimates is due to the fact that both studies aim to estimate the pervasive and nonpervasive factors in one objective function, which as a consequence leads to a complex minimization problem. Thus, the numerical implementation often becomes complex, and is hard to be justified. In another two works, both Choi et al. (2018) and Andreou et al. (2019) propose sequential procedures to identify and estimate pervasive and nonpervasive shocks, in which canonical correlation analysis (CCA) are adopted. However, only two of the three dimensions are allowed to diverge in both studies. Han (2019) considers a shrinkage estimation approach to explore the group effects of the factor structure, which can be computationally expensive, as the choice of tuning parameter often plays an important role in practice.

From the practical point of view, being able to separate the pervasive and nonpervasive shocks in a multi-dimensional panel data is crucial for a range of applications. First, as mentioned in the beginning of the paper, accounting for pervasive and nonpervasive shocks reveals a detailed network structure of the international trade. We will come back to it in the empirical study section. A second example is better understanding business-cycle fluctuations across countries and regions (Kose et al., 2003). Along this line of research, identifying the common fluctuations across macroeconomic aggregates worldwide has always been one of the priorities (e.g., Gregory et al., 1997). The emergence of multi-dimensional panel data models provides an excellent framework to facilitate the investigation. Another field which urgently calls for development on multi-dimensional panel data models is associated with migration flows. As well understood, the rate of migration between two countries does not depend solely on their relative attractiveness, but also on the one of alternative destinations (Bertoli and Fernández-Huertas Moraga, 2013). Given the increasing mobility of the entire population, how to better capture the bilateral flows therefore becomes vital now more than ever. Other examples requiring multi-dimensional panel data models can also be found in Choi et al. (2018), Kapetanios et al. (2020), etc.

Having presented the above challenges and necessities, in this study, we specifically consider a three-dimensional panel data model with unobserved global (pervasive) and country-specific (nonpervasive) factors, which has been exposed in the literature but has not been fully solved to the best of the authors' knowledge. On theory, our contributions are the following three-fold: (1). under the scenario that all three dimensions can diverge to infinity, we propose an estimation approach to identify the number of global shocks and country-specific shocks sequentially; (2). the newly proposed approach is easy to implement, and the asymptotic theories are established accordingly; (3). we further conduct intensive numerical studies to examine the finite sample performance of the newly proposed approach using both simulated and real datasets. In the empirical study, we then apply the approach to decompose the network of bilateral trade using country level data from two major trading groups (APEC and EU) over the period 1982-2019. We find that the country-specific shocks become more volatile in recent years, which may indicate the increasing instability of the inward and outward bilateral trade costs over the past couple of decades. In addition, we show that the trade flows involving China mainland, Germany and the United States show relatively strong sensitivity to global shocks, which reflects the fact that, in general, they are leading export and import countries worldwide. We note that the relationship among Canada, Mexico, and the United States is also highly sensitive to different shocks, which somewhat reflects the fact that all three of them are highly economically related through North American Free Trade Agreement (NAFTA) that eliminates some trade barriers and promotes the trading activities.

The structure of this paper is as follows. Section 2 presents the model with the estimation approach, and establishes the asymptotic properties accordingly. In Section 3, we conduct intensive simulations to examine the finite sample performance of the newly proposed approach. Section 4 provides an empirical study using country level bilateral trade data. Section 5 concludes. Due to the limit of space, the preliminary lemmas and the proofs are given in the online supplementary appendices.

Before proceeding further, it is convenient to introduce some notation: $\|\cdot\|_F$ denotes the Euclidean norm of a vector or the Frobenius norm of a matrix; for a matrix $\mathbf{A}$, its spectral norm is defined as $\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}\{\mathbf{A}'\mathbf{A}\}}$, where $\lambda_{\max}\{\cdot\}$ denotes the maximum eigenvalue; $\mathbf{M_A} = \mathbf{I} - \mathbf{P_A}$ denotes the orthogonal projection matrix generated by matrix $\mathbf{A}$, where $\mathbf{P_A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ and $\mathbf{A}$ is a matrix with full column rank; let $\to_P$ and $\to_D$ denote convergence in probability and in distribution, respectively; we write $a \asymp b$ if $a = O_P(b)$ and $b = O_P(a)$; let $\text{diag}(\mathbf{A}, \mathbf{B})$ denotes the block-diagonal matrix that takes $\mathbf{A}$ and $\mathbf{B}$ as the upper left and lower right blocks; $\text{vec}(\mathbf{A})$ stands for the vectorization operation; $\mathbb{I}(\cdot)$ stands for the indicator function.

2 Model & Methodology

In this section, we first present the model, then provide the estimation approach, and finally establish the asymptotic theories accordingly.

2.1 The Setup

Having presented our motivations in Section 1, we specifically consider the next model in this study.

$$y_{ijt} = \boldsymbol{\gamma}_{ij}'\boldsymbol{g}_t + \boldsymbol{\lambda}_{E,ij}'\boldsymbol{f}_{E,it} + \boldsymbol{\lambda}_{I,ij}'\boldsymbol{f}_{I,jt} + u_{ijt}, \tag{2.1}$$

where $i = 1, \ldots, M$ index the exporters, $j = 1, \ldots, N$ index the importers, and $t = 1, \ldots, T$ index the time periods. We observe y_{ijt}'s only, and u_{ijt}'s are the idiosyncratic error terms. $\boldsymbol{g}_t$ is an $r_g \times 1$ unobservable global factor, which is regarded as global shocks and may capture the globalisation trends. Some detailed explanation on the globalisation trends can be found in Kapetanios et al. (2020), and we shall be more specific on this so-called "trend" in the empirical study of Section 4. $\boldsymbol{f}_{E,it}$ and $\boldsymbol{f}_{I,jt}$ represent the unobservable $r_{E,i} \times 1$ and $r_{I,j} \times 1$ country-specific factors. Specifically, $\boldsymbol{f}_{E,it}$ is referred to as an exporter factor which affects all import partners associated with export country i and $\boldsymbol{f}_{I,jt}$ is referred to as an importer factor which affects all export partners associated with import country j. The country-specific factors may capture the unobservable multilateral trade resistances (MTRs) that are different for exporters and importers. Loosely speaking, MTRs refer to the barriers which each of exporter and importer face in their trade with all their trading partners. We refer interested readers to Anderson and Van Wincoop (2003) for a comprehensive discussion on MTR. $\boldsymbol{\gamma}_{ij}$, $\boldsymbol{\lambda}_{E,ij}$ and $\boldsymbol{\lambda}_{I,ij}$ are the corresponding factor loadings. Throughout this paper, we always use the subscript $_g$ to denote the variables associated with the global factors, and use the subscripts $_E$ and $_I$ to denote the variables associated with the exporters and importers respectively.

The model (2.1) is in fact not new, and has been mentioned in Breitung and Eickmeier (2016, eq. 18), Choi et al. (2018, eq. 1), and Kapetanios, Serlenga and Shin (2020, eq. 2) among others for different purposes. In what follows, we propose an easily implemented methodology to recover the structure of the right hand side of (2.1), when all three dimensions are allowed to diverge to infinity. Precisely, we first estimate the numbers of global and country-specific factors (i.e., the values of r_g, $r_{E,i}$'s and $r_{I,j}$'s), and then establish inferences for global and country-specific shocks.

Remark 2.1. *Before proceeding further, we comment on an important identification issue. For simplicity, we suppose that* $r_g = 1$ *and* $g_t \equiv 1$, *and suppose further that*

$$\boldsymbol{f}_{E,it} = \boldsymbol{f}_E + \boldsymbol{\eta}_{E,it} \quad \textit{with} \quad E[\boldsymbol{\eta}_{E,it}] = 0,$$
$$\boldsymbol{f}_{I,it} = \boldsymbol{f}_I + \boldsymbol{\eta}_{I,it} \quad \textit{with} \quad E[\boldsymbol{\eta}_{I,it}] = 0.$$

Then, the model (2.1) *becomes*

$$y_{ijt} = \boldsymbol{\gamma}_{ij}^* + \boldsymbol{\lambda}_{E,ij}'\boldsymbol{\eta}_{E,it} + \boldsymbol{\lambda}_{I,ij}'\boldsymbol{\eta}_{I,jt} + u_{ijt}, \tag{2.2}$$

where $\boldsymbol{\gamma}_{ij}^* = \boldsymbol{\gamma}_{ij} + \boldsymbol{\lambda}_{E,ij}'\boldsymbol{f}_E + \boldsymbol{\lambda}_{I,ij}'\boldsymbol{f}_I$. *It then infers that for a model having a multi-layer factor structure, only one layer can have non-zero mean factors.*

Having said Remark 2.1, without loss of generality, we assume that

$$E[\boldsymbol{f}_{E,it}] = \mathbf{0} \quad \text{and} \quad E[\boldsymbol{f}_{I,jt}] = \mathbf{0} \tag{2.3}$$

for country-specific factors throughout this study.

As repeatedly pointed out in the literature (e.g., Wang, 2010; Breitung and Eickmeier, 2016; Choi et al., 2018), investigating (2.1) relies on how to utilize the sparse structure of the next presentation.

$$\begin{pmatrix} y_{11t} \\ \vdots \\ y_{M1t} \\ \vdots \\ y_{1Nt} \\ \vdots \\ y_{MNt} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\gamma}_{11}' & \boldsymbol{\lambda}_{E,11}' & \cdots & \mathbf{0} & \boldsymbol{\lambda}_{I,11}' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\gamma}_{M1}' & \mathbf{0} & \cdots & \boldsymbol{\lambda}_{E,M1}' & \boldsymbol{\lambda}_{I,M1}' & \cdots & \mathbf{0} \\ \vdots & & \vdots & & & \vdots & \\ \boldsymbol{\gamma}_{1N}' & \boldsymbol{\lambda}_{E,1N}' & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\lambda}_{I,1N}' \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\gamma}_{MN}' & \mathbf{0} & \cdots & \boldsymbol{\lambda}_{E,MN}' & \mathbf{0} & \cdots & \boldsymbol{\lambda}_{I,MN}' \end{pmatrix} \begin{pmatrix} \boldsymbol{g}_t \\ \boldsymbol{f}_{E,1t} \\ \vdots \\ \boldsymbol{f}_{E,Mt} \\ \boldsymbol{f}_{I,1t} \\ \vdots \\ \boldsymbol{f}_{I,Nt} \end{pmatrix} + \begin{pmatrix} u_{11t} \\ \vdots \\ u_{M1t} \\ \vdots \\ u_{1Nt} \\ \vdots \\ u_{MNt} \end{pmatrix} \tag{2.4}$$

In view of (2.4), a few facts emerge:

1. In order to estimate (2.4), one needs to identify the number of factors for each $\boldsymbol{g}_t$, $\boldsymbol{f}_{E,it}$ and $\boldsymbol{f}_{I,jt}$. Traditional PCA usually requires a low rank setting. However, having $\boldsymbol{g}_t$, $\boldsymbol{f}_{E,it}$'s and $\boldsymbol{f}_{I,jt}$'s in one column as in (2.4) yields a factor with a diverging dimension, which suggests that recovering all factors and loadings in one goal seems to be challenging. Thus, it motivates us to consider a multiple steps approach below.

2. The country-specific factors associated with exporters and importers are interchangeable, as the sparse structure associated with the corresponding factor loadings depends on how we rank y_{ijt} with respect to i and j only. Thus, we would expect to recover the exporter and importer factors in a parallel manner.

3. As clearly seen in (2.4), $\boldsymbol{g}_t$ has an impact on every single y_{ijt}, although the magnitude depends on the value of $\boldsymbol{\gamma}_{ij}$. However, $\boldsymbol{f}_{E,it}$ or $\boldsymbol{f}_{I,jt}$ affects only an asymptotically negligible subset of y_{ijt}'s due to the sparse structure. From the signal-to-noise ratio point of view, we expect that the global factors are easier to be identified. Intuitively speaking, they can be estimated first if principal component analysis (PCA) is employed. As the country-specific factors contain the second tier of signal, they should be recovered after removing the dominating ones.

In Section 2.1 below, we propose a multi-step estimation approach based on the aforementioned points.

2.2 The Estimation Approach

We are now ready to present the estimation approach, which is a procedure involving multiple steps. The outline is as follows.

Step 1 Conduct PCA to identify the number of global factors r_g, and estimate the global factor, which contains the strongest "signal" as explained under (2.4).

Step 2 Remove the estimated global factor, then simultaneously conduct multiple PCA to estimate the number of country-specific factors $r_{E,i}$' and $r_{I,j}$'s, and recover the country-specific factors, which contain "signals" weaker than the global factor but stronger than the error terms.

First, we write (2.1) in matrix form to facilitate the development. Throughout, the subscript $\bullet$ always stands for including all available sample in the corresponding dimension for notational simplicity.

$$\mathbf{Y} = \mathbf{\Gamma}\mathbf{G}' + \mathbf{\Lambda}_E\mathbf{F}_E' + \mathbf{\Lambda}_I\mathbf{F}_I' + \mathbf{U}, \tag{2.5}$$

where the response variables and error terms are defined by

$$\begin{aligned}
\mathbf{Y} &= (\mathbf{Y}_{\bullet\bullet 1}, \dots, \mathbf{Y}_{\bullet\bullet T}), \quad \mathbf{Y}_{\bullet\bullet t} = (y_{11t}, \dots, y_{M1t}, \dots, y_{1Nt}, \dots, y_{MNt})', \\
\mathbf{U} &= (\mathbf{U}_{\bullet\bullet 1}, \dots, \mathbf{U}_{\bullet\bullet T}), \quad \mathbf{U}_{\bullet\bullet t} = (u_{11t}, \dots, u_{M1t}, \dots, u_{1Nt}, \dots, u_{MNt})';
\end{aligned} \tag{2.6}$$

the global factors and loadings are defined by

$$\mathbf{G} = (\boldsymbol{g}_1, \dots, \boldsymbol{g}_T)', \quad \mathbf{\Gamma} = (\boldsymbol{\gamma}_{11}, \dots, \boldsymbol{\gamma}_{M1}, \dots, \boldsymbol{\gamma}_{1N}, \dots, \boldsymbol{\gamma}_{MN})'; \tag{2.7}$$

and the country-specific factors and loadings are defined by

$$\begin{aligned}
\mathbf{F}_E &= (\mathbf{F}_{E,1}, \dots, \mathbf{F}_{E,M}), \quad \mathbf{F}_{E,i} = (\boldsymbol{f}_{E,i1}, \dots, \boldsymbol{f}_{E,iT})', \\
\mathbf{F}_I &= (\mathbf{F}_{I,1}, \dots, \mathbf{F}_{I,N}), \quad \mathbf{F}_{I,j} = (\boldsymbol{f}_{I,j1}, \dots, \boldsymbol{f}_{I,jT})', \\
\mathbf{\Lambda}_E &= (\mathrm{diag}\{\mathbf{\Lambda}_{E,\bullet 1}'\}, \dots, \mathrm{diag}\{\mathbf{\Lambda}_{E,\bullet N}'\})', \quad \mathbf{\Lambda}_{E,\bullet j} = (\boldsymbol{\lambda}_{E,1j}, \dots, \boldsymbol{\lambda}_{E,Mj})', \\
\mathbf{\Lambda}_I &= \mathrm{diag}\{\mathbf{\Lambda}_{I,\bullet 1}, \dots, \mathbf{\Lambda}_{I,\bullet N}\}, \quad \mathbf{\Lambda}_{I,\bullet j} = (\boldsymbol{\lambda}_{I,1j}, \dots, \boldsymbol{\lambda}_{I,Mj})'.
\end{aligned} \tag{2.8}$$

With the above notations in hand, we are ready to present the details of each step with necessary discussions.

Step 1 — Conduct PCA on $\frac{1}{MNT}\mathbf{Y}'\mathbf{Y}$ as follows.

$$\widehat{\mathbf{G}}\mathbf{V}_g = \frac{1}{MNT}\mathbf{Y}'\mathbf{Y}\widehat{\mathbf{G}}, \tag{2.9}$$

in which $\frac{1}{T}\widehat{\mathbf{G}}'\widehat{\mathbf{G}} = \mathbf{I}_{k_{\max}}$, $\mathbf{V}_g = \text{diag}\{\widehat{\rho}_{g,1}, \ldots, \widehat{\rho}_{g,k_{\max}}\}$ with $\widehat{\rho}_{g,1} \geq \cdots \geq \widehat{\rho}_{g,k_{\max}}$ being the largest $k_{\max}$ eigenvalues, $k_{\max}$ $(> r_g)$ is a user-specified fixed large integer. By (2.9), we implement the following two sub-steps.

Step 1.1 Estimate the number of global factors r_g by

$$\widehat{r}_g = \underset{0\leq k\leq k_{\max}}{\text{argmin}} \left\{ \frac{\widehat{\rho}_{g,k+1}}{\widehat{\rho}_{g,k}} \cdot \mathbb{I}\left(\widehat{\rho}_{g,k} \geq \omega_{MNT}\right) + \mathbb{I}\left(\widehat{\rho}_{g,k} < \omega_{MNT}\right) \right\}, \tag{2.10}$$

where $\omega_{MNT} = 1/\ln(\max\{M, N, T\})$, and $\widehat{\rho}_{g,0} = 1$ is a mock eigenvalue.

Step 1.2 Estimate $\mathbf{G}$ by letting $\widehat{\mathbf{G}}$ include the first $\widehat{r}_g$ columns only, where we have slightly abused the notation $\widehat{\mathbf{G}}$. The loading matrix is estimated by $\widehat{\mathbf{\Gamma}} = \frac{1}{T}\mathbf{Y}\widehat{\mathbf{G}}$.

Step 2 includes two parallel sections: **Part 1** and **Part 2**.

Part 1 — For each $j = 1, \ldots, N$, conduct PCA:

$$\widehat{\mathbf{F}}_{I,j}\mathbf{V}_{I,j} = \frac{1}{MT}(\mathbf{Y}_{I,j} - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{Y}_{I,j} - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')\widehat{\mathbf{F}}_{I,j}, \tag{2.11}$$

where $\mathbf{Y}_{I,j} = (\mathbf{Y}_{\bullet j1}, \ldots, \mathbf{Y}_{\bullet jT})$ with $\mathbf{Y}_{\bullet jt} = (y_{1jt}, \ldots, y_{Mjt})'$, $\widehat{\mathbf{\Gamma}}_{I,j}$ includes the M rows of $\widehat{\mathbf{\Gamma}}$ corresponding the j^{th} importer, $\frac{1}{T}\widehat{\mathbf{F}}_{I,j}'\widehat{\mathbf{F}}_{I,j} = \mathbf{I}_{k_{\max}}$, and $\mathbf{V}_{I,j} = \text{diag}\{\widehat{\rho}_{Ij,1}, \ldots, \widehat{\rho}_{Ij,k_{\max}}\}$ with $\widehat{\rho}_{Ij,1} \geq \cdots \geq \widehat{\rho}_{Ij,k_{\max}}$ being the largest $k_{\max}$ $(> r_{I,j})$ eigenvalues. By (2.11), implement the followings.

Part 1.1 Estimate $r_{I,j}$ by

$$\widehat{r}_{I,j} = \underset{0\leq k\leq k_{\max}}{\text{argmin}} \left\{ \frac{\widehat{\rho}_{Ij,k+1}}{\widehat{\rho}_{Ij,k}} \cdot \mathbb{I}\left(\widehat{\rho}_{Ij,k} \geq \omega_{MNT}\right) + \mathbb{I}\left(\widehat{\rho}_{Ij,k} < \omega_{MNT}\right) \right\}, \tag{2.12}$$

where $\widehat{\rho}_{Ij,0} = 1$ is a mock eigenvalue.

Part 1.2 Estimate $\mathbf{F}_{I,j} = (\boldsymbol{f}_{I,j1}, \ldots, \boldsymbol{f}_{I,jT})'$ by letting $\widehat{\mathbf{F}}_{I,j}$ include the first $\widehat{r}_{I,j}$ columns only. The loading matrix $\mathbf{\Lambda}_{I,\bullet j}$ defined in (2.8) is estimated by $\widehat{\mathbf{\Lambda}}_{I,\bullet j} = \frac{1}{T}(\mathbf{Y}_{I,j} - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')\widehat{\mathbf{F}}_{I,j}$.

Part 2 — For each $i = 1, \ldots, M$, conduct PCA:

$$\widehat{\mathbf{F}}_{E,i}\mathbf{V}_{E,i} = \frac{1}{NT}(\mathbf{Y}_{E,i} - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}')'(\mathbf{Y}_{E,i} - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}')\widehat{\mathbf{F}}_{E,i}, \tag{2.13}$$

where $\mathbf{Y}_{E,i} = (\mathbf{Y}_{i\bullet 1}, \ldots, \mathbf{Y}_{i\bullet T})$ with $\mathbf{Y}_{i\bullet t} = (y_{i1t}, \ldots, y_{iNt})'$, $\widehat{\mathbf{\Gamma}}_{E,i}$ includes the N rows of $\widehat{\mathbf{\Gamma}}$ corresponding to the i^{th} exporter, $\frac{1}{T}\widehat{\mathbf{F}}'_{E,i}\widehat{\mathbf{F}}_{E,i} = \mathbf{I}_{k_{\max}}$, and $\mathbf{V}_{E,i} = \mathrm{diag}\{\widehat{\rho}_{Ei,1}, \ldots, \widehat{\rho}_{Ei,k_{\max}}\}$ with $\widehat{\rho}_{Ei,1} \geq \cdots \geq \widehat{\rho}_{Ei,k_{\max}}$ being the largest $k_{\max}$ ($> r_{E,i}$) eigenvalues. By (2.13), we conduct the followings.

Part 2.1 Estimate $r_{E,i}$ by

$$\widehat{r}_{E,i} = \underset{0 \leq k \leq k_{\max}}{\mathrm{argmin}} \left\{ \frac{\widehat{\rho}_{Ei,k+1}}{\widehat{\rho}_{Ei,k}} \cdot \mathbb{I}\left(\widehat{\rho}_{Ei,k} \geq \omega_{MNT}\right) + \mathbb{I}\left(\widehat{\rho}_{Ei,k} < \omega_{MNT}\right) \right\}, \tag{2.14}$$

where $\widehat{\rho}_{Ei,0} = 1$ is a mock eigenvalue.

Part 2.2 Estimate $\mathbf{F}_{E,i} = (\boldsymbol{f}_{E,i1}, \ldots, \boldsymbol{f}_{E,iT})'$ by letting $\widehat{\mathbf{F}}_{E,i}$ include the first $\widehat{r}_{E,i}$ columns only. The loading matrix $\mathbf{\Lambda}_{E,i\bullet} = (\boldsymbol{\lambda}_{E,i1}, \ldots, \boldsymbol{\lambda}_{E,iN})'$ is estimated by $\widehat{\mathbf{\Lambda}}_{E,i\bullet} = \frac{1}{T}(\mathbf{Y}_{E,i} - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}')\widehat{\mathbf{F}}_{E,i}$.

Remark 2.2. *We make a few comments on the estimation approach. (1). The use of eigenvalue ratio in* (2.10), (2.12) *and* (2.14) *is in the same spirit of Lam and Yao (2012) and Ahn and Horenstein (2013). (2). The threshold ω_{MNT} is to bypass a technical challenge raised in Lam and Yao (2012, eq. 3.3), and the mock eigenvalues $\widehat{\rho}_{g,0}$, $\widehat{\rho}_{Ij,0}$'s and $\widehat{\rho}_{Ei,0}$'s are designed to capture the cases where there are no global factors, or some of the country-specific factors do not exist. From the dimension reduction point of view, it is crucial to have a procedure which accounts for zero factors under the three dimensional panel data framework. (3). $k_{\max}$ is a user-defined fixed integer. Practically, one can adopt any reasonable large value which suits the empirical study (e.g., Fan et al., 2013; Pelger and Xiong, 2019).*

2.3 Consistency

In this subsection, we show that the number of factors can be identified consistently in each step with necessary conditions. The asymptotic distributions are established in the next subsection.

To facilitate the development, we impose the following conditions.

Assumption 1.

1. *As $T \to \infty$, $\frac{1}{T}\mathbf{G}'\mathbf{G} \to_P \mathbf{\Sigma}_{\mathbf{G}}$, where $\mathbf{\Sigma}_{\mathbf{G}}$ is a deterministic positive definite matrix. Also,* $\max_{t \geq 1} E\|\boldsymbol{g}_t\|_F^4 < \infty$.

2. *Suppose that* (2.3) *holds. Moreover,* $\max_{i\geq 1,t\geq 1} E\|\boldsymbol{f}_{E,it}\|_F^4 < \infty$ *and* $\|\mathbf{F}_E\|_2 = O_P(\sqrt{T} \vee \sqrt{M})$. *Also,* $\max_{j\geq 1,t\geq 1} E\|\boldsymbol{f}_{I,jt}\|_F^4 < \infty$ *and* $\|\mathbf{F}_I\|_2 = O_P(\sqrt{T} \vee \sqrt{N})$.

Assumption 2.

1. *As* $(M,N) \to (\infty,\infty)$, $\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \to_P \mathbf{\Sigma}_{\mathbf{\Gamma}}$, *where* $\mathbf{\Sigma}_{\mathbf{\Gamma}}$ *is a deterministic positive definite matrix. Also,* $\max_{i\geq 1,j\geq 1} E\|\boldsymbol{\gamma}_{ij}\|_F^4 < \infty$.

2. *Suppose that* $\max_{i\geq 1,j\geq 1} E\|\boldsymbol{\lambda}_{E,ij}\|_F^4 < \infty$ *and* $\max_{i\geq 1,j\geq 1} \|\boldsymbol{\lambda}_{E,ij}\|_F = O_P(\sqrt{\ln(MN)})$. *Also,* $\max_{i\geq 1,j\geq 1} E\|\boldsymbol{\lambda}_{I,ij}\|_F^4 < \infty$ *and* $\max_{i\geq 1,j\geq 1} \|\boldsymbol{\lambda}_{I,ij}\|_F = O_P(\sqrt{\ln(MN)})$.

Assumption 3.

1. *Let* $\{u_{ijt} \mid i \geq 1, j \geq 1, t \geq 1\}$ *be independent of the other variables. Let* $\mathcal{F}_{-\infty}^0$ *and* $\mathcal{F}_\tau^\infty$ *denote the* σ*-algebras generated by* $\{\mathbf{U}_{\bullet\bullet t} \mid t \leq 0\}$ *and* $\{\mathbf{U}_{\bullet\bullet t} \mid t \geq \tau\}$ *respectively, where* $\mathbf{U}_{\bullet\bullet t} = (u_{11t}, \ldots, u_{M1t}, \ldots, u_{1Nt}, \ldots, u_{MNt})'$. *Define the mixing coefficient* $\alpha(\tau) = \sup_{A\in\mathcal{F}_{-\infty}^0, B\in\mathcal{F}_\tau^\infty} |\Pr(A)\Pr(B) - \Pr(AB)|$.

   (a) *Let* $\{\mathbf{U}_{\bullet\bullet t} \mid t \geq 1\}$ *be strictly stationary and* α*-mixing such that for some* $\nu > 0$, $\max_{i\geq 1,j\geq 1} E|u_{ijt}|^{4+\nu} < \infty$, *and the mixing coefficient satisfies* $\sum_{t=1}^\infty [\alpha(t)]^{\nu/(2+\nu)} < \infty$.

   (b) $E[u_{ijt}] = 0$, $\max_{i\geq 1,j\geq 1} \sigma_{ij}^2 < \infty$ *and* $\sum_{(i,j)\neq(m,n)} |\sigma_{ijmn}| = O(MN)$, *where* $\sigma_{ij}^2 = E[u_{ijt}^2]$ *and* $\sigma_{ijmn} = E[u_{ijt}u_{mnt}]$ *for* $t \geq 1$. *In addition, suppose that* $\sum_{i,m=1}^M \sum_{j,n=1}^N \sum_{t,s=1}^T |E[u_{ijt}u_{mns}]| = O(MNT)$.

2. *Suppose that* $r_g < \infty$, $\max_{i\geq 1} r_{E,i} < \infty$, *and* $\max_{j\geq 1} r_{I,j} < \infty$.

Assumption 1 imposes restrictions on the global and country-specific factors, which are not more restrictive than Assumption 1.i of Choi et al. (2018). The conditions on the spectral norm of $\mathbf{F}_E$ and $\mathbf{F}_I$ are widely adopted in the literature (e.g., Li, Qian and Su, 2016, Assumption A.1.iii and Lu and Su, 2016, Assumption A.1.v). Extensive discussions with examples on this type of assumption can be found in Moon and Weidner (2015).

Assumption 2 puts restrictions on the loadings associated with the global and country-specific factors. The bounds on $\max_{i\geq 1,j\geq 1} \|\boldsymbol{\lambda}_{E,ij}\|_F$ and $\max_{i\geq 1,j\geq 1} \|\boldsymbol{\lambda}_{I,ij}\|_F$ are fairly standard. See Assumption A7 of Connor et al. (2012) for example.

Assumption 3.1 assumes that the error terms u_{ijt}'s follow stationary time series process over t, and simultaneously allow for weak cross-sectional dependence over i and j. Assumption 3.2 requires r_g, $r_{E,i}$'s and $r_{I,j}$'s to be bounded, which nests $r_g = 0$, $r_{E,i} = 0$ and $r_{I,j} = 0$ as special cases.

Under these conditions, we present the first theorem of this paper below.

Theorem 2.1. *Under Assumptions 1-3, as* $(M, N, T) \to (\infty, \infty, \infty)$*,*

1. *in* **Step 1.1**, $\Pr(\widehat{r}_g = r_g) \to 1$*;*

2. *in* **Step 1.2**, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$*, where* $\mathbf{H} = \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}} \cdot (\mathbf{V}_g^{\dagger})^{-1}$*, and* $\mathbf{V}_g^{\dagger}$ *is the* $r_g \times r_g$ *leading principal submatrix of* $\mathbf{V}_g$*.*

Theorem 2.1.1 shows that r_g can be estimated consistently, while Theorem 2.1.2 indicates that we can only recover $\mathbf{G}$ up to a rotation matrix. From the signal-to-noise ratio point of view, only the space spanned by the global factors can be recovered in **Step 1**.

Remark 2.3. *It is noteworthy that when establishing Theorem 2.1, no harsh conditions are imposed between the global factor structure and the country-specific ones. In this sense, although the rate of Theorem 2.1 is slow, we show that the global factors can be identified from the data first with minimum cost. In the traditional literature, the fact has barely been mentioned. To the best of the authors' knowledge, the only exception is Remark 4 of Han (2019). In Appendix A of the online supplementary file, we provide a sharper rate for the estimation of the global factor when more structures are adopted. The details are summarized in Lemma A.4.*

Having presented the results associated with the global factors, we investigate the country-specific ones, and further impose the following conditions.

Assumption 4.

1. *For* $i = 1, \ldots, M$ *and* $j = 1, \ldots, N$*, suppose that the following conditions hold:*

   (a) $\frac{1}{T}\|\mathbf{G}'\mathbf{F}_{E,i}\|_F = O_P(T^{a_{E,i}})$ *and* $\frac{1}{T}\|\mathbf{G}'\mathbf{F}_{I,j}\|_F = O_P(T^{a_{I,j}})$*, where* $\mathbf{F}_{E,i}$ *and* $\mathbf{F}_{I,j}$ *are defined in* (2.8)*,* $\max_{i\geq 1} a_{E,i} < 0$*, and* $\max_{j\geq 1} a_{I,j} < 0$*;*

   (b) $\max_{i\geq 1}\|\frac{1}{T}\mathbf{F}'_{E,i}\mathbf{F}_{E,i} - \mathbf{\Sigma}_{\mathbf{F}_{E,i}}\|_F = o_P(1)$ *and* $\max_{j\geq 1}\|\frac{1}{T}\mathbf{F}'_{I,j}\mathbf{F}_{I,j} - \mathbf{\Sigma}_{\mathbf{F}_{I,j}}\|_F = o_P(1)$*, where* $\mathbf{\Sigma}_{\boldsymbol{F}_{E,i}}$ *and* $\mathbf{\Sigma}_{\boldsymbol{F}_{I,j}}$ *are deterministic positive definite matrices;*

   (c) $\max_{i\geq 1}\|\frac{1}{N}\mathbf{\Lambda}'_{E,i\bullet}\mathbf{\Lambda}_{E,i\bullet} - \mathbf{\Sigma}_{\mathbf{\Lambda}_{E,i\bullet}}\|_F = o_P(1)$ *and* $\max_{j\geq 1}\|\frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j} - \mathbf{\Sigma}_{\mathbf{\Lambda}_{I,\bullet j}}\|_F = o_P(1)$*, where* $\mathbf{\Lambda}_{E,i\bullet} = (\boldsymbol{\lambda}_{E,i1}, \ldots, \boldsymbol{\lambda}_{E,iN})'$*,* $\mathbf{\Lambda}_{I,\bullet j}$ *is defined under* (2.8)*, and* $\mathbf{\Sigma}_{\mathbf{\Lambda}_{E,i\bullet}}$ *and* $\mathbf{\Sigma}_{\mathbf{\Lambda}_{I,\bullet j}}$ *are deterministic positive definite matrices.*

2. *Suppose that* $\max_{j\geq 1}\sum_{i\neq m}\sigma_{ijmj} = O(M)$*, and* $\max_{i\geq 1}\sum_{j\neq n}\sigma_{ijin} = O(N)$*, where* σ_{ijmn} *is defined in Assumption 3.*

Assumption 4.1.(a) requires certain orthogonality between the global factors and country-specific factors. Specifically, the values of $a_{E,i}$ and $a_{I,j}$ measure the degree of orthogonality between the global and country-specific factors. If $a_{E,i} = a_{I,j} = -\infty$, this condition essentially reduces to Assumption A of Ando and Bai (2017), where they show the necessity of orthogonality in order to identify the common and group-specific factors under a two-dimensional panel data framework. Similar discussions on orthogonality can also be seen in Andreou et al. (2019). Assumptions 4.1.(b) and 4.1.(c) impose more conditions on the blocks of factors and loadings associated with exporters and importers, which are fairly standard. Assumption 4.2 further regulates the weak cross-sectional dependence of the error terms.

With Assumption 4 in hand, the country-specific factor structures can be successfully recovered in **Step 2**. The details are summarized in the next theorem.

Theorem 2.2. *Under Assumptions 1-4, as* $(M, N, T) \to (\infty, \infty, \infty)$*,*

1. *For* $j = 1, \dots, N$*,*
   (a) *in* **Part 1.1** of **Step 2**, $\Pr(\widehat{r}_{I,j} = r_{I,j}) \to 1$*;*
   (b) *in* **Part 1.2** of **Step 2**, $\frac{1}{\sqrt{T}} \|\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}} + T^{a_{I,j}}\right)$*, where* $\mathbf{H}_{I,j} = \frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j} \cdot \frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j} \cdot (\mathbf{V}^{\dagger}_{I,j})^{-1}$*, and* $\mathbf{V}^{\dagger}_{I,j}$ *is the* $r_{I,j} \times r_{I,j}$ *leading principal submatrix of* $\mathbf{V}_{I,j}$*.*

2. *For* $i = 1, \dots, M$*,*
   (a) *in* **Part 2.1** of **Step 2**, $\Pr(\widehat{r}_{E,i} = r_{E,i}) \to 1$*;*
   (b) *in* **Part 2.2** of **Step 2**, $\frac{1}{\sqrt{T}} \|\widehat{\mathbf{F}}_{E,i} - \mathbf{F}_{E,i}\mathbf{H}_{E,i}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}} + T^{a_{E,i}}\right)$*, where* $\mathbf{H}_{E,i} = \frac{1}{N}\mathbf{\Lambda}'_{E,i\bullet}\mathbf{\Lambda}_{E,i\bullet} \cdot \frac{1}{T}\mathbf{F}'_{E,i}\widehat{\mathbf{F}}_{E,i} \cdot (\mathbf{V}^{\dagger}_{E,i})^{-1}$*, and* $\mathbf{V}^{\dagger}_{E,i}$ *is the* $r_{E,i} \times r_{E,i}$ *leading principal submatrix of* $\mathbf{V}_{E,i}$*.*

Theorem 2.2 shows that $r_{E,i}$ and $r_{I,j}$ can be estimated consistently. Moreover, $\widehat{\mathbf{F}}_{I,j}$ and $\widehat{\mathbf{F}}_{E,i}$ respectively recover $\mathbf{F}_{I,j}$ and $\mathbf{F}_{E,i}$ up to rotation matrices.

Till now, we conclude that we have successfully recovered the network presented by (2.4). To establish inferences for the estimation approach, we study the asymptotic distributions associated with **Step 1** and **Step 2** in the next subsection.

2.4 Asymptotic Distribution

In order to establish the asymptotic distributions, the following assumptions are necessary to facilitate the development.

Assumption 5.

1. *Let* $\frac{1}{\sqrt{MN}}\|\mathbf{\Gamma}'\mathbf{\Lambda}_E\|_F = O_P(1)$ *and* $\frac{1}{\sqrt{MN}}\|\mathbf{\Gamma}'\mathbf{\Lambda}_I\|_F = O_P(1)$.

2. $\frac{1}{T}\mathbf{G}'\mathbf{G} = \mathbf{I}_{r_g}$ *and* $\mathbf{\Gamma}'\mathbf{\Gamma}$ *is a diagonal matrix with distinct entries.*

3. *Suppose that* $\frac{1}{\sqrt{MN}}\sum_{i=1}^{M}\sum_{j=1}^{N}\boldsymbol{\gamma}_{ij}\upsilon_{ijt} \to_D N(\mathbf{0},\mathbf{\Phi}_t)$ *for* $t = 1,\ldots,T$, *where* $\upsilon_{ijt} = \boldsymbol{\lambda}_{E,ij}'\boldsymbol{f}_{E,it} + \boldsymbol{\lambda}_{I,ij}'\boldsymbol{f}_{I,jt} + u_{ijt}$.

Assumption 6.

1. *Suppose that* $\frac{1}{T}\|\mathbf{F}_{E,i}'\mathbf{F}_{I,j}\|_F = O_P(T^{b_{EI,ij}})$, *where* $\max_{i\geq 1,j\geq 1} b_{EI,ij} < 0$.

2. (a) $\frac{1}{T}\mathbf{F}_{I,j}'\mathbf{F}_{I,j} = \mathbf{I}_{r_{I,j}}$ *and* $\mathbf{\Lambda}_{I,\bullet j}'\mathbf{\Lambda}_{I,\bullet j}$ *is a diagonal matrix with distinct entries;*

   (b) $\frac{1}{T}\mathbf{F}_{E,i}'\mathbf{F}_{E,i} = \mathbf{I}_{r_{E,i}}$ *and* $\mathbf{\Lambda}_{E,i\bullet}'\mathbf{\Lambda}_{E,i\bullet}$ *is a diagonal matrix with distinct entries.*

3. (a) $\frac{1}{\sqrt{M}}\sum_{i=1}^{M}\boldsymbol{\lambda}_{I,ij}(\boldsymbol{\lambda}_{E,ij}'\boldsymbol{f}_{E,it} + u_{ijt}) \to_D N(\mathbf{0},\mathbf{\Omega}_{I,jt})$ *for each pair of* (j,t);

   (b) $\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\boldsymbol{\lambda}_{E,ij}(\boldsymbol{\lambda}_{I,ij}'\boldsymbol{f}_{I,jt} + u_{ijt}) \to_D N(\mathbf{0},\mathbf{\Omega}_{E,it})$ *for each pair of* (i,t).

Assumption 5.1 requires certain orthogonality between global factor loadings and country-specific factor loadings, which is not unusual in the literature. For instance, Lam and Yao (2012) explain the rational behind such a setting at length. Assumption 5.2 further imposes conditions for the purpose of identification, which has been extensively discussed in Bai and Ng (2013) and Fan et al. (2016). In view of Remark 2.1, Assumption 5.3 is fairly standard. We further explain Assumption 5.3 together with Assumption 6.3 below.

Similar to Assumption 5.1, Assumption 6.1 requires certain orthogonality but focusing on the export factors and importer factors, while Assumption 6.2 is for the purpose of identification. Assumption 6.3 is somewhat interesting. Take

$$\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\boldsymbol{\lambda}_{E,ij}(\boldsymbol{\lambda}_{I,ij}'\boldsymbol{f}_{I,jt} + u_{ijt}) \to_D N(\mathbf{0},\mathbf{\Omega}_{E,it})$$

as an example, which says the asymptotic distribution associated with the i-th exporter factor at time t is not only driven by the error component, but also is driven by its entire importer network. The same argument applies to the importer factor. In this way, the networks of export and import are entangled with each other. Mathematically, it requires country-specific shocks to have mean 0, which is ensured by (2.3). See Assumption 1.ii of Choi et al. (2018) and Assumption 1.a of Han (2019) for similar settings.

To close our theoretical investigation, we summarize the asymptotic distributions associated with the global and country-specific factors in the next theorem.

Theorem 2.3. *Under Assumptions 1-5, Let* $(M, N, T) \to (\infty, \infty, \infty)$.

1. *If* $\sqrt{MN}(\frac{1}{T} + \Delta^*_{g,MNT}) \to 0$, *then* $\sqrt{MN}(\widehat{\boldsymbol{g}}_t - \boldsymbol{g}_t) \to_D N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}^{-1}\boldsymbol{\Phi}_t\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}^{-1})$ *for each* t.

In addition, let Assumption 6 also hold.

2. *If* $\sqrt{M}(\frac{1}{\sqrt{T}} + \Delta^*_{Ij,MNT}) \to 0$, *then* $\sqrt{M}(\widehat{\boldsymbol{f}}_{I,jt} - \boldsymbol{f}_{I,jt}) \to_D N(\mathbf{0}, \boldsymbol{\Sigma}^{-1}_{\boldsymbol{\Lambda}_{I,\bullet j}}\boldsymbol{\Omega}_{I,jt}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\Lambda}_{I,\bullet j}})$ *for each* (j, t);
3. *If* $\sqrt{N}(\frac{1}{\sqrt{T}} + \Delta^*_{Ei,MNT}) \to 0$, *then* $\sqrt{N}(\widehat{\boldsymbol{f}}_{E,it} - \boldsymbol{f}_{E,it}) \to_D N(\mathbf{0}, \boldsymbol{\Sigma}^{-1}_{\boldsymbol{\Lambda}_{E,i\bullet}}\boldsymbol{\Omega}_{E,it}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\Lambda}_{E,i\bullet}})$ *for each* (i, t).

In the above, $\Delta^*_{g,MNT}$, $\Delta^*_{Ij,MNT}$ *and* $\Delta^*_{Ei,MNT}$ *are defined as follows.*

$$\begin{aligned}
\Delta^*_{g,MNT} &= \frac{T^{\max_i a_{E,i}}}{\sqrt{N}} + \frac{T^{\max_j a_{I,j}}}{\sqrt{M}} + \frac{\sqrt{\ln(MN)} \cdot (T^{\max_i a_{E,i}} + T^{\max_j a_{I,j}})}{\sqrt{T}} \\
&\quad + \ln(MN) \cdot (T^{2\max_i a_{E,i}} + T^{2\max_j a_{I,j}}); \\
\Delta^*_{Ij,MNT} &= \frac{\ln(MN) \cdot T^{\max_j a_{I,j}}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}} + \sqrt{\ln(MN)} \cdot (T^{\max_i a_{E,i}} + T^{\max_i b_{EI,ij}}) + T^{a_{I,j}}; \\
\Delta^*_{Ei,MNT} &= \frac{\ln(MN) \cdot T^{\max_i a_{E,i}}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}} + \sqrt{\ln(MN)} \cdot (T^{\max_j a_{I,j}} + T^{\max_j b_{EI,ij}}) + T^{a_{E,i}}.
\end{aligned}$$

The condition $\frac{\sqrt{MN}}{T} \to 0$ in the first result of Theorem 2.3 is equivalent to $\frac{\sqrt{N}}{T} \to 0$ in Theorem 1 of Bai and Ng (2013) in which a two dimension model is considered. The condition $\sqrt{MN} \cdot \Delta^*_{g,MNT} \to 0$ requires the orthogonality between the global and country-specific factor structures are strong enough in order to achieve the optimal rate $\sqrt{MN}$. If we adopt the orthogonality as in Ando and Bai (2017) and Andreou et al. (2019), then this condition will completely vanish.

In order to achieve asymptotic normality for the country-specific factors, slightly stronger restrictions (such as $\frac{M}{T} \to 0$ and $\frac{N}{T} \to 0$) are imposed in the body of this theorem on top of Assumption 6, which is due to the fact that we need to account for the estimation bias caused by Step 1 of the estimation approach. It is noteworthy that $\frac{M}{T} \to 0$ and $\frac{N}{T} \to 0$ imply $\frac{MN}{T^2} \to 0$, which has been discussed above. Therefore, we claim the newly imposed conditions are reasonable, and are only slightly stronger than those used in traditional two dimensional analysis. The conditions $\sqrt{M} \cdot \Delta^*_{Ij,MNT} \to 0$ and $\sqrt{N} \cdot \Delta^*_{Ei,MNT} \to 0$ require the orthogonality between the global and country-specific factor structures are strong enough in order to achieve the optimal rates $\sqrt{M}$ and $\sqrt{N}$. Again, if orthogonality is adopted, these conditions will disappear automatically.

3 Simulation

In this section, we examine the finite sample performance of the methodology proposed in Section 2. Specifically, the data generating process (DGP) is as follows.

$$y_{ijt} = \boldsymbol{\gamma}_{ij}' \boldsymbol{g}_t + \boldsymbol{\lambda}_{E,ij}' \boldsymbol{f}_{E,it} + \boldsymbol{\lambda}_{I,ij}' \boldsymbol{f}_{I,jt} + u_{ijt}, \tag{3.1}$$

where $i = 1, \ldots, M$, $j = 1, \ldots, N$, and $t = 1, \ldots, T$. The global factors, country-specific factors and idiosyncratic errors are generated by the following AR(1) processes

$$\begin{aligned}
\boldsymbol{g}_t &= \phi_g \boldsymbol{g}_{t-1} + \boldsymbol{v}_{g,t} \quad \text{with} \quad \boldsymbol{v}_{g,t} \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{r_g}), \\
\boldsymbol{f}_{E,it} &= \phi_{E,i} \boldsymbol{f}_{E,i,t-1} + \boldsymbol{v}_{E,it} \quad \text{with} \quad \boldsymbol{v}_{E,it} \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{r_{E,i}}), \\
\boldsymbol{f}_{I,jt} &= \phi_{I,j} \boldsymbol{f}_{I,j,t-1} + \boldsymbol{v}_{I,jt} \quad \text{with} \quad \boldsymbol{v}_{I,jt} \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{r_{I,j}}), \\
u_{ijt} &= \phi_u u_{ij,t-1} + e_{ijt} \quad \text{with} \quad e_{ijt} \sim i.i.d.\ N(0, 1),
\end{aligned}$$

where $i.i.d.$ stands for independent and identically distributed. The factor loadings are generated as: $\boldsymbol{\gamma}_{ij} \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{r_g})$, $\boldsymbol{\lambda}_{E,ij} \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{r_{E,i}})$, and $\boldsymbol{\lambda}_{I,ij} \sim i.i.d. N(\mathbf{0}, \mathbf{I}_{r_{I,j}})$.

We consider the following two cases.

DGP 1: Let $\phi_g = \phi_{E,i} = \phi_{I,j} = \phi_u = 0$, $r_g = 3$, $r_{E,i} = 2$ for $i = 1, \ldots, M$, and $r_{I,j} = 1$ for $j = 1, \ldots, N$;

DGP 2: Let $\phi_g = \phi_{E,i} = \phi_{I,j} = \phi_u = 0.5$, and the rest values are the same as those in DGP 1.

For each DGP, we conduct the estimation approach of Section 2 by letting $M, N, T \in \{20, 40, 60, 80\}$, and implement 1000 replications for each given sample size.

To measure the performance of the proposed estimation approach, we define a few criteria below. First, we measure the detection on different factors, and start from the global factor structure.

$$P_{g,c} = \frac{1}{1000} \sum_{\ell=1}^{1000} \mathbb{I}(\widehat{r}_g^{\ell} = r_g), \quad P_{g,u} = \frac{1}{1000} \sum_{\ell=1}^{1000} \mathbb{I}(\widehat{r}_g^{\ell} < r_g), \quad P_{g,o} = \frac{1}{1000} \sum_{\ell=1}^{1000} \mathbb{I}(\widehat{r}_g^{\ell} > r_g),$$

where $\widehat{r}_g^\ell$ defines the estimated of r_g at the ℓ^{th} replication. It is clear that $P_{g,c}$, $P_{g,u}$ and $P_{g,o}$ define the probabilities of correctly, under and over select the number of global factors. For the export factors, we define

$$P_{E,c} = \frac{1}{1000}\sum_{\ell=1}^{1000}\frac{1}{M}\sum_{i=1}^{M}\mathbb{I}(\widehat{r}_{E,i}^\ell = r_{E,i}), \quad P_{E,u} = \frac{1}{1000}\sum_{\ell=1}^{1000}\frac{1}{M}\sum_{i=1}^{M}\mathbb{I}(\widehat{r}_{E,i}^\ell < r_{E,i}),$$
$$P_{E,o} = \frac{1}{1000}\sum_{\ell=1}^{1000}\frac{1}{M}\sum_{i=1}^{M}\mathbb{I}(\widehat{r}_{E,i}^\ell > r_{E,i}),$$

where $\widehat{r}_{E,i}^\ell$ stands for the estimated $\widehat{r}_{E,i}$ at the ℓ^{th} replication. Also, it is obvious that $P_{E,c}$, $P_{E,u}$ and $P_{E,o}$ define the probabilities of correctly, under and over select the number of export factors. Similarly, we can define $P_{I,c}$, $P_{I,u}$ and $P_{I,o}$ for the import factors. The details are omitted for the sake of conciseness.

Second, we measure the estimation on different factors. Recall that we have defined $\mathbf{G}$, $\mathbf{F}_{E,i}$ and $\mathbf{F}_{I,j}$ under (2.5), and then further define

$$\text{RMSE}_{\mathbf{G}} = \sqrt{\frac{1}{1000}\sum_{\ell=1}^{1000}\|\mathbf{P}_{\widehat{\mathbf{G}}^\ell} - \mathbf{P}_{\mathbf{G}^\ell}\|_F^2},$$
$$\text{RMSE}_{\mathbf{E}} = \sqrt{\frac{1}{1000}\sum_{\ell=1}^{1000}\frac{1}{M}\sum_{i=1}^{M}\|\mathbf{P}_{\widehat{\mathbf{F}}_{E,i}^\ell} - \mathbf{P}_{\mathbf{F}_{E,i}^\ell}\|_F^2},$$
$$\text{RMSE}_{\mathbf{I}} = \sqrt{\frac{1}{1000}\sum_{\ell=1}^{1000}\frac{1}{N}\sum_{j=1}^{N}\|\mathbf{P}_{\widehat{\mathbf{F}}_{I,j}^\ell} - \mathbf{P}_{\mathbf{F}_{I,j}^\ell}\|_F^2}.$$

In the above formulas, we let $\widehat{\mathbf{G}}^\ell$ and $\mathbf{G}^\ell$ include the estimated and true global factors from the ℓ^{th} replication. Similarly, we define $\widehat{\mathbf{F}}_{E,i}^\ell$ and $\mathbf{F}_{E,i}^\ell$ for the exporter factors, and define $\widehat{\mathbf{F}}_{I,j}^\ell$ and $\mathbf{F}_{I,j}^\ell$ for the importer factors.

We summarize the simulation results in Table 1 to Table 3. Note that due to the limit of space, the results of some combinations of (M, N, T) are dropped in all tables. In Table 1, it is clear that as the sample size goes up, the values of $P_{g,c}$, $P_{E,c}$ and $P_{I,c}$ converge to 1. When the sample size is relatively small, it seems that we tend to under select the number of factors. Once all M, N, T are greater than and equal to 40, the selection on the factors is quite accurate. In Table 2, we consider a DGP with more time series correlation, and the pattern is almost identical to those presented in Table 1. Table 3 reports the results of $\text{RMSE}_{\mathbf{G}}$, $\text{RMSE}_{\mathbf{E}}$ and $\text{RMSE}_{\mathbf{I}}$. It is not surprising that all values of RMSE converge to 0, as the sample size goes up. Moreover, the values of $\text{RMSE}_{\mathbf{E}}$ and $\text{RMSE}_{\mathbf{I}}$ are larger than $\text{RMSE}_{\mathbf{G}}$ in general, which should be expected. The reason is that

Step 2 includes the estimation bias associated with Step 1, although the bias is negligible in the asymptotic sense under certain restrictions. It is noteworthy that the values of $\mathrm{RMSE}_{\mathbf{E}}$ are larger than those of $\mathrm{RMSE}_{\mathbf{I}}$, which is due to the fact that more unobservable factors are included for the exporters.

Having justified the validity of the proposed estimation approach through simulations, we are now ready to move on to the empirical study in the next section.

4 Empirical Study

In this section, we use the proposed methodology to investigate the international trade flows.

4.1 The Data

We use monthly bilateral export volumes of commodity goods among 23 countries/region over the period of 1982-2019. The export flows data are collected from the Direction of Trade Statistics (DOTS) of International Monetary Fund (IMF) available at `https://www.imf.org/external/index.htm`. We use the FOB (free on board) value of exports of goods denominated in U.S. dollars and restrict the sample to 506 country-pairs of 23 countries/regions from two major trading groups over a 456-month period from January, 1982 to December, 2019.

- Asia-Pacific Economic Cooperation (APEC): Australia (AUS), China Mainland (CHN), Hong Kong (HKG), Indonesia (IDN), Japan (JPN), Korea (KOR), Malaysia (MYS), New Zealand (NZL), Singapore (SGP), Thailand (THA), Canada (CAN), Mexico (MEX), United States (USA)
- European Union (EU): Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Ireland (IRL), Italy (ITA), Netherlands (NLD), Spain (ESP), Sweden (SWE), United Kingdom (GBR)

Canada, Mexico and United States are also the members of North American Free Trade Agreement (NAFTA). As they are already included in APEC, we no longer specifically mention NAFTA in this study. It is worth pointing out that a similar dataset is considered in Chen and Chen (2019) to investigate the patterns in the dynamic network of international trade. The difference between their study and our paper lies on the setting of factor structure. While we consider multiple layers of the factor structure, their study focuses on one layer only with a different presentation. As a consequence, the two models and the corresponding estimation approaches are not directly comparable.

In what follows, the combination of an export country/region and one of its import partner is referred to as a country pair. For example, the export flow from the United States to Australia and the export flow from Australia to the United States are the bilateral export flows for two different country pairs.

4.2 Estimation Results

We first report the estimated numbers of global and country-specific factors. Specifically, only one global factor is identified from the sample. The estimated numbers of exporter factors and importer factors are summarized in Table 4. As shown in the table, majorities have only 1 or 2 factors with the importer factors of IDN being the only exception.

Figure 1 shows the estimated global factor which has a clear upward trend. First, let's explain why such a behaviour can be captured under the proposed framework. Note that Assumption 1 requires $\frac{1}{T}\mathbf{G}'\mathbf{G} \to_P \mathbf{\Sigma_G}$ only. As a special case, it may possess a form like

$$\frac{1}{T}\sum_{t=1}^{T}[g(\tau_t)]^2 \to \int_0^1 [g(w)]^2 dw, \tag{4.1}$$

where $\tau_t = t/T$, and $g(\cdot)$ can be functions such as $g(w) = w$, $g(w) = w^2$, etc. Therefore, the upward trending is obviously included. Detailed discussions on trending behaviour like (4.1) can be seen in Yan et al. (2020). As explained in Wang (2010) and Breitung and Eickmeier (2016), the global factor may be interpreted as global shocks on the entire network of international trade, e.g., the Global Financial Crisis. Our finding is somewhat consistent with their arguments. For example, there is a sudden and severe drop around 2009 which captures the so-called "great trade collapse", a consequence of the 2008 financial crisis, occurred between the third quarter of 2008 and the second quarter of 2009. We refer interested readers to Bems et al. (2012) for more details on great trade collapse. In addition, we note that the global factor becomes more volatile over the sample period, which may indicate the increasing vulnerability of countries to shocks on trade due to globalization over the past couple of decades.

Figures 2 - 5 show the estimated exporter factors and importer factors. Specifically, Figure 2 and Figure 3 present he exporter factors associated with the countries of APEC and EU respectively. Figure 4 and Figure 5 show the importer factors associated with the countries of APEC and EU respectively. The exporter factors can be interpreted as country-specific shocks of export countries which affect the trade volumes from the exporters to the import partners. Similarly, the importer factors can be interpreted as country-specific shocks of import countries which affect the trade volumes from the importers to the export partners. As mentioned in Section 2.1, the exporter and importer factors may capture the unobservable outward and inward multilateral trade resistances (MTRs) for different exporters and importers respectively, which can be seen as measures of outward and inward bilateral trade costs for different exporters and importers. The detailed discussions on the connection between multilateral resistances and country-specific factors can be found in Kapetanios et al. (2020), where the exporter and importer factors are always referred to as source and destination country factors. For almost all country-specific factors, we can observe the increase of the volatility, especially from the beginning of the 21st century, indicating the increasing instability of the inward and outward bilateral trade costs for most of the countries in our sample. Under the assumption of bilateral trade costs symmetry, it follows that the inward and outward multilateral resistances are the same for the same country (Anderson and Van Wincoop, 2003). By comparing the estimated exporter and importer factors for the same country, it can be seen that this symmetry in the multilateral resistances is partially supported by the data. For example, the exporter and importer factors for USA share the similar trend.

Figure 6 presents the heat map of the global factor loadings for different country pairs. Since the global factor loadings are positive for all country pairs, we rescale them to $[0, 1]$ for better presentation. The global factor loading can be interpreted as the responses of the trade volumes for different country pairs to the global shocks. The colour of each cell reflects the sensitivity of the trade volume between two countries to the global shocks. For example, in Figure 6, the darkest cell corresponding to the export flow from CAN to USA indicates that the export volume from CAN to USA is the most sensitive relationship among all bilateral export flows in the sample. Also, the country pairs like CHN and HKG, CHN and USA, MEX and USA also show strong sensitivity to the global shocks. The relationship among USA, MEX and CAN partially can be explained by the fact that all three of them are the members of NAFTA, which eliminates some trade barriers among the three parties and promotes the trading activities. The similar patterns can also be observed among countries from EU and Asia respectively. Overall, by comparing the values in different rows and columns of the plot, it can be seen that the trade flows involving USA, CHN and DEU show relatively strong sensitivity to global shocks, which indicates that, in general, they are leading export and import countries worldwide.

Figure 7 and Figure 8 present the heat maps of the exporter factor loadings and the importer factor loadings, respectively. Each column in the plots represents a country-specific factor loading corresponding to an exporter or importer factor. Similar to the global factor loading, the exporter and importer factor loadings are rescaled to have values between -1 and 1. The exporter factor loadings corresponding to different export countries measure the responses of their import partners to the shocks on those export countries. The importer factor loadings can be interpreted in the same manner. As shown in Figure 7, the export flow from CAN to USA is relatively sensitive to the exporter shocks of CAN. This is also the case for the export flow from JPN to USA which is shown to be sensitive to the exporter shocks of JPN. On the other hand, Figure 8 shows that the export flows from both CAN and JPN to USA are also sensitive to the country-specific importer shocks of USA. The country-specific factor loadings for other countries can be interpreted similarly.

5 Conclusion

In this study, we specifically consider a three-dimensional panel data model, which has been exposed in the literature but has not been fully solved to the best of the authors' knowledge. On theory, our contributions are the following three-fold: (1). under the scenario that all three dimensions can diverge to infinity, we propose an estimation approach to identify the number of global shocks and country-specific shocks sequentially; (2). the newly proposed approach is easy to implement, and the asymptotic theories are established accordingly; (3). we further conduct intensive numerical studies to examine the finite sample performance of the newly proposed approach using both simulated and real datasets. In the empirical study, we then apply the approach to decompose the network of bilateral trade using country level data from two major trading groups (APEC and EU) over the period 1982-2019. We find that the country-specific shocks become more volatile in recent years, which may indicate the increasing instability of the inward and outward bilateral trade costs over the past couple of decades. In addition, we show that the trade flows involving China mainland, Germany and the United States show relatively strong sensitivity to global shocks, which reflects the fact that, in general, they are leading export and import countries worldwide. We note that the relationship among Canada, Mexico, and the United States is also highly sensitive to different shocks, which somewhat reflects the fact that all three of them are highly economically related through NAFTA that eliminates some trade barriers and promotes the trading activities.

References

Ahn, S. C. and Horenstein, A. R. (2013), 'Eigenvalue ratio test for the number of factors', *Econometrica* **81**(3), 1203–1227.

Anderson, J. E. and Van Wincoop, E. (2003), 'Gravity with gravitas: A solution to the border puzzle', *American economic review* **93**(1), 170–192.

Ando, T. and Bai, J. (2017), 'Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures', *Journal of the American Statistical Association* **112**(519), 1182–1198.

Andreou, E., Gagliardini, P., Ghysels, E. and Rubin, M. (2019), 'Inference in group factor models with an application to mixed-frequency data', *Econometrica* **87**(4), 1267–1305.

Bai, J. and Ng, S. (2008), 'Large dimensional factor analysis', *Foundations and Trends(R) in Econometrics* **3**(2), 89–163.

Bai, J. and Ng, S. (2013), 'Principal components estimation and identification of static factors', *Journal of Econometrics* **176**(1), 18–99.

Baltagi, B. H., Egger, P. and Pfaffermayr, M. (2015), 'Panel data gravity models of international trade', *The Oxford Handbook of Panel Data* pp. 608–641.

Beck, G. W., Hubrich, K. and Marcellino, M. (2016), 'On the importance of sectoral and regional shocks for price-setting', *Journal of Applied Econometrics* **31**(7), 1234–1253.

Bems, R., Johnson, R. C. and Yi, K. (2012), The great trade collapse. NBER Working Paper No. 18632.

Bertoli, S. and Fernández-Huertas Moraga, J. (2013), 'Multilateral resistance to migration', *Journal of Development Economics* **102**, 79 – 100.

Breitung, J. and Eickmeier, S. (2016), 'Analyzing international business and financial cycles using Multi-Level factor models: A comparison of alternative approaches', *Advances in Econometrics* **35**, 177–214.

Chen, E. Y. and Chen, R. (2019), 'Modeling dynamic transport network with matrix factor models: with an application to international trade flow', *arXiv:1901.00769* .

Choi, I., Kim, D., Kim, Y. J. and Kwark, N.-S. (2018), 'A multilevel factor model: Identification, asymptotic theory and applications', *Journal of Applied Econometrics* **33**(3), 355–377.

Choi, I., Lin, R. and Shin, Y. (2020), 'Canonical correlation-based model selection for the multilevel factors', *Available at SSRN 3590109* .

Connor, G., Hagmann, M. and Linton, O. (2012), 'Efficient semiparametric estimation of the fama-french model and extensions', *Econometrica* **80**(2), 713–754.

Ergemen, Y. E. and Rodríguez-Caballero, C. V. (2017), Estimation of a dynamic multi-level factor model with possible long-range dependence. UC3M Working Papers, Statistics and Econometrics, 17-08, ISSN 2387-0303.

Fan, J., Liao, Y. and Mincheva, M. (2013), 'Large covariance estimation by thresholding principal orthogonal complements', *Journal of the Royal Statistical Society: Series B* **75**(4), 603–680.

Fan, J., Liao, Y. and Wang, W. (2016), 'Projected principal component analysis in factor models', *Annals of Statistics* **44**, 219–254.

Gregory, A. W., Head, A. C. and Raynauld, J. (1997), 'Measuring world business cycles', *International Economic Review* **38**(3), 677–701.

Han, X. (2019), 'Shrinkage estimation of factor models with global and group-specific factors', *Journal of Business & Economic Statistics* **0**(0), 1–17.

Kapetanios, G., Serlenga, L. and Shin, Y. (2020), 'Estimation and inference for multi-dimensional heterogeneous panel datasets with hierarchical multi-factor error structure', *Journal of Econometrics* .

Kose, M. A., Otrok, C. and Whiteman, C. H. (2003), 'International business cycles: World, region, and country-specific factors', *American Economic Review* **93**(4), 1216–1239.

Lam, C. and Yao, Q. (2012), 'Factor modeling for high-dimensional time series: Inference for the number of factors', *Annals of Statistics* **40**(2), 694–726.

Lam, C., Yao, Q. and Bathia, N. (2011), 'Estimation of latent factors for high-dimensional time series', *Biometrika* **98**(4), 901–918.

Li, D., Qian, J. and Su, L. (2016), 'Panel data models with interactive fixed effects and multiple structural breaks', *Journal of the American Statistical Association* **111**(516), 1804–1819.

Lu, X. and Su, L. (2016), 'Shrinkage estimation of dynamic panel data models with interactive fixed effects', *Journal of Econometrics* **190**(1), 148–175.

Magnus, J. R. and Neudecker, H. (2019), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons.

Moench, E., Ng, S. and Potter, S. (2013), 'Dynamic hierarchical factor models', *The Review of Economics and Statistics* **95**(5), 1811–1817.

Moon, H. R. and Weidner, M. (2015), 'Linear regression for panel with unknown number of factors as interactive fixed effects', *Econometrica* **83**(4), 1543–1579.

Pelger, M. and Xiong, R. (2019), State-varying factor models of large dimensions. arXiv:1807.02248.

Wang, P. (2010), Large dimensional factor models with a multi-level factor structure: Identification, estimation and inference. Working Paper, Hong Kong University of Science and Technology, Hong Kong.

Yan, Y., Gao, J. and Peng, B. (2020), Nonparametric time-varying vector moving average (∞) models. https://www.monash.edu/business/econometrics-and-business-statistics/research/publications.

Table 1: DGP 1 – The percentages of correctly, under and over selecting factors. Specifically, $P_{g,c}$, $P_{g,u}$ and $P_{g,o}$ are for the global factors; $P_{E,c}$, $P_{E,u}$ and $P_{E,o}$ are for the export factors; $P_{I,c}$, $P_{I,u}$ and $P_{I,o}$ are for the import factors. For the sake of space, some combinations of (M, N, T) are omitted in the table.

M	N	T	$P_{g,c}$	$P_{g,u}$	$P_{g,o}$	$P_{E,c}$	$P_{E,u}$	$P_{E,o}$	$P_{I,c}$	$P_{I,u}$	$P_{I,o}$
20	20	20	0.650	0.349	0.001	0.439	0.418	0.144	0.682	0.049	0.269
		40	0.978	0.022	0.000	0.655	0.287	0.059	0.844	0.046	0.110
		60	0.995	0.005	0.000	0.710	0.238	0.051	0.878	0.035	0.087
		80	1.000	0.000	0.000	0.734	0.211	0.055	0.890	0.033	0.077
40	40	20	0.836	0.164	0.000	0.629	0.306	0.065	0.854	0.062	0.084
		40	0.997	0.003	0.000	0.883	0.107	0.010	0.963	0.022	0.015
		60	0.999	0.001	0.000	0.936	0.054	0.010	0.980	0.010	0.009
		80	1.000	0.000	0.000	0.958	0.033	0.010	0.986	0.006	0.008
60	60	20	0.884	0.116	0.000	0.731	0.227	0.042	0.906	0.053	0.041
		40	0.998	0.002	0.000	0.955	0.041	0.003	0.988	0.009	0.002
		60	1.000	0.000	0.000	0.985	0.013	0.002	0.996	0.003	0.002
		80	1.000	0.000	0.000	0.993	0.005	0.002	0.998	0.001	0.001
80	80	20	0.865	0.135	0.000	0.759	0.185	0.056	0.915	0.041	0.044
		40	1.000	0.000	0.000	0.980	0.019	0.001	0.995	0.004	0.001
		60	1.000	0.000	0.000	0.996	0.003	0.001	0.999	0.001	0.000
		80	1.000	0.000	0.000	0.999	0.001	0.000	1.000	0.000	0.000
20	80	20	0.759	0.241	0.000	0.691	0.193	0.116	0.711	0.043	0.246
		40	0.990	0.010	0.000	0.970	0.022	0.008	0.853	0.041	0.106
		60	1.000	0.000	0.000	0.996	0.003	0.001	0.884	0.033	0.084
		80	1.000	0.000	0.000	0.999	0.001	0.000	0.899	0.027	0.074
80	20	20	0.819	0.181	0.000	0.473	0.420	0.107	0.886	0.044	0.070
		40	0.998	0.002	0.000	0.650	0.294	0.056	0.995	0.005	0.000
		60	1.000	0.000	0.000	0.714	0.233	0.053	0.999	0.001	0.000
		80	1.000	0.000	0.000	0.743	0.205	0.051	1.000	0.000	0.000

Table 2: DGP 2 – The percentages of correctly, under and over selecting factors. Specifically, $P_{g,c}$, $P_{g,u}$ and $P_{g,o}$ are for the global factors; $P_{E,c}$, $P_{E,u}$ and $P_{E,o}$ are for the export factors; $P_{I,c}$, $P_{I,u}$ and $P_{I,o}$ are for the import factors. For the sake of space, some combinations of (M, N, T) are omitted in the table.

M	N	T	$P_{g,c}$	$P_{g,u}$	$P_{g,o}$	$P_{E,c}$	$P_{E,u}$	$P_{E,o}$	$P_{I,c}$	$P_{I,u}$	$P_{I,o}$
20	20	20	0.240	0.705	0.055	0.301	0.400	0.299	0.505	0.014	0.481
		40	0.700	0.300	0.000	0.467	0.356	0.178	0.694	0.005	0.301
		60	0.920	0.080	0.000	0.615	0.300	0.085	0.814	0.010	0.176
		80	0.990	0.010	0.000	0.700	0.242	0.058	0.870	0.008	0.123
40	40	20	0.280	0.660	0.060	0.339	0.380	0.280	0.543	0.023	0.435
		40	0.820	0.180	0.000	0.668	0.240	0.092	0.881	0.010	0.109
		60	0.990	0.010	0.000	0.858	0.126	0.016	0.970	0.006	0.024
		80	1.000	0.000	0.000	0.916	0.073	0.011	0.988	0.001	0.011
60	60	20	0.220	0.750	0.030	0.329	0.342	0.329	0.562	0.010	0.428
		40	0.800	0.200	0.000	0.738	0.148	0.114	0.895	0.011	0.094
		60	0.990	0.010	0.000	0.935	0.054	0.011	0.985	0.003	0.012
		80	1.000	0.000	0.000	0.974	0.025	0.001	0.997	0.001	0.002
80	80	20	0.290	0.700	0.010	0.342	0.347	0.311	0.613	0.009	0.378
		40	0.850	0.150	0.000	0.794	0.110	0.096	0.919	0.006	0.075
		60	0.990	0.010	0.000	0.964	0.028	0.008	0.994	0.001	0.005
		80	1.000	0.000	0.000	0.991	0.008	0.001	1.000	0.000	0.000
20	80	20	0.239	0.698	0.063	0.340	0.344	0.316	0.494	0.013	0.493
		40	0.720	0.280	0.000	0.676	0.117	0.207	0.716	0.006	0.278
		60	0.990	0.010	0.000	0.965	0.026	0.009	0.845	0.007	0.149
		80	1.000	0.000	0.000	0.990	0.010	0.000	0.878	0.005	0.117
80	20	20	0.270	0.670	0.060	0.305	0.394	0.301	0.582	0.025	0.393
		40	0.750	0.250	0.000	0.485	0.366	0.149	0.851	0.008	0.141
		60	0.970	0.030	0.000	0.625	0.307	0.068	0.981	0.002	0.017
		80	1.000	0.000	0.000	0.691	0.254	0.055	0.999	0.001	0.000

Table 3: The results of $\mathrm{RMSE}_{\mathbf{G}}$, $\mathrm{RMSE}_{\mathbf{E}}$ and $\mathrm{RMSE}_{\mathbf{I}}$. Specifically, $\mathrm{RMSE}_{\mathbf{G}}$, $\mathrm{RMSE}_{\mathbf{E}}$ and $\mathrm{RMSE}_{\mathbf{I}}$ measure the estimation on global factors, export factors and importer factors, respectively. For the sake of space, some combinations of (M, N, T) are omitted in the table.

			DGP 1			DGP 2		
M	N	T	$\mathrm{RMSE}_{\mathbf{G}}$	$\mathrm{RMSE}_{\mathbf{E}}$	$\mathrm{RMSE}_{\mathbf{I}}$	$\mathrm{RMSE}_{\mathbf{G}}$	$\mathrm{RMSE}_{\mathbf{E}}$	$\mathrm{RMSE}_{\mathbf{I}}$
20	20	20	0.778	1.172	1.034	1.189	1.370	1.321
		40	0.368	0.989	0.793	0.805	1.174	1.146
		60	0.314	0.910	0.733	0.496	1.038	0.961
		80	0.292	0.884	0.695	0.359	0.947	0.891
40	40	20	0.519	1.015	0.761	1.164	1.266	1.165
		40	0.191	0.778	0.574	0.581	0.960	0.790
		60	0.168	0.683	0.525	0.275	0.784	0.610
		80	0.154	0.637	0.499	0.222	0.716	0.549
60	60	20	0.423	0.941	0.680	1.144	1.234	1.189
		40	0.141	0.687	0.512	0.582	0.953	0.718
		60	0.118	0.595	0.454	0.261	0.715	0.546
		80	0.108	0.546	0.423	0.177	0.652	0.500
80	80	20	0.440	0.915	0.664	1.110	1.243	1.130
		40	0.110	0.655	0.470	0.539	0.888	0.679
		60	0.094	0.557	0.421	0.231	0.698	0.500
		80	0.086	0.504	0.394	0.163	0.603	0.458
20	80	20	0.618	0.978	1.009	1.186	1.251	1.337
		40	0.232	0.671	0.792	0.711	0.956	1.062
		60	0.180	0.571	0.719	0.295	0.678	0.818
		80	0.168	0.519	0.689	0.236	0.612	0.739
80	20	20	0.560	1.131	0.694	1.136	1.427	1.088
		40	0.191	0.961	0.492	0.657	1.131	0.723
		60	0.164	0.890	0.432	0.307	0.958	0.534
		80	0.154	0.866	0.403	0.217	0.925	0.468

Table 4: Estimated number of exporter factors and importer factors.

	APEC							
	CAN	MEX	USA	AUS	CHN	HKG	IDN	JPN
Exporter	1	1	1	1	2	2	2	1
Importer	1	2	1	1	1	1	3	1
	KOR	MYS	NZL	SGP	THA			
Exporter	1	1	1	1	1			
Importer	1	1	1	1	2			
	EU							
	DEU	DNK	ESP	FIN	FRA	GBR	IRL	ITA
Exporter	1	1	1	1	1	1	2	1
Importer	2	1	1	2	1	1	1	1
	NLD	SWE						
Exporter	1	1						
Importer	2	1						

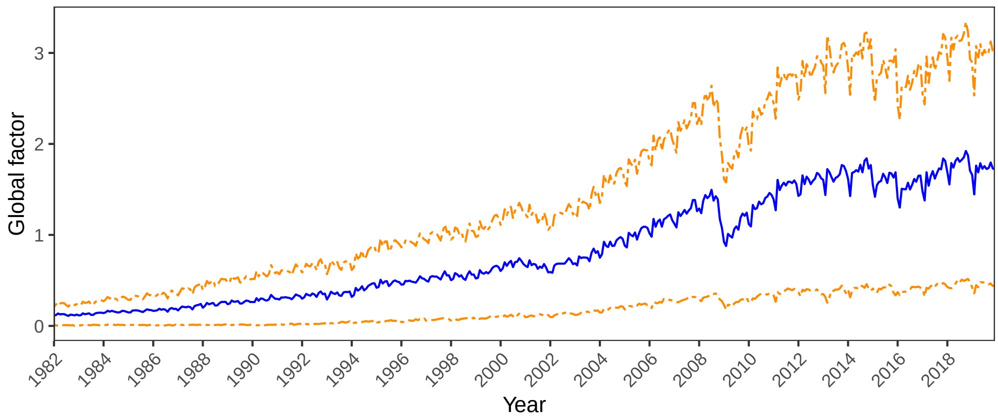

Figure 1: Global factor for bilateral export flows from January, 1982 to December, 2019. The blue solid line represents the estimated global factor and the orange dashed lines provide a 95% confidence interval.

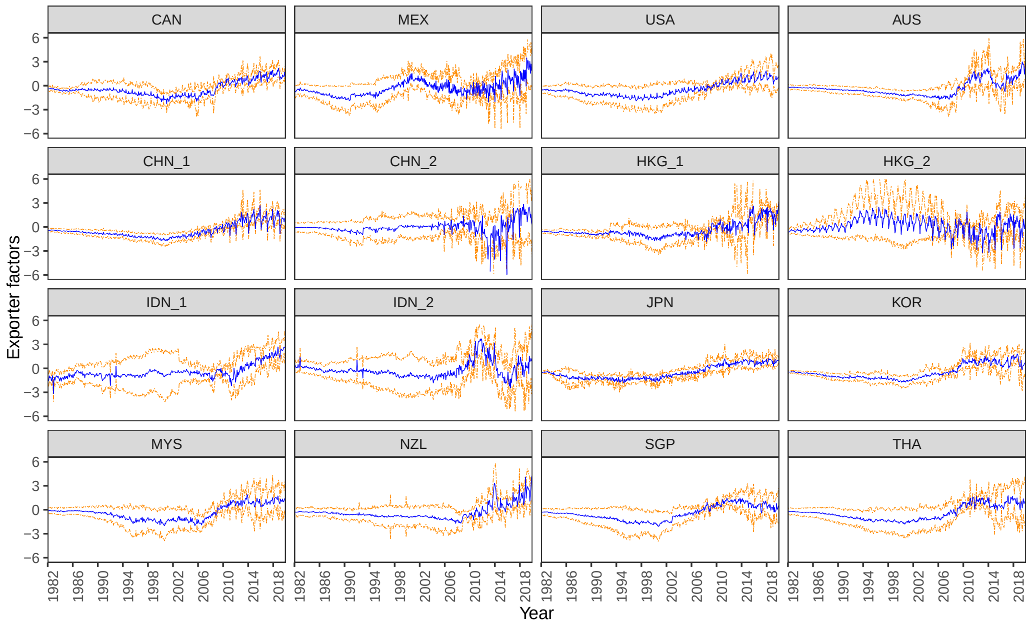

Figure 2: Exporter factors for 13 APEC countries/region in the sample from January, 1982 to December, 2019. The blue solid line represents the estimated global factor and the orange dashed lines provide a 95% confidence interval.

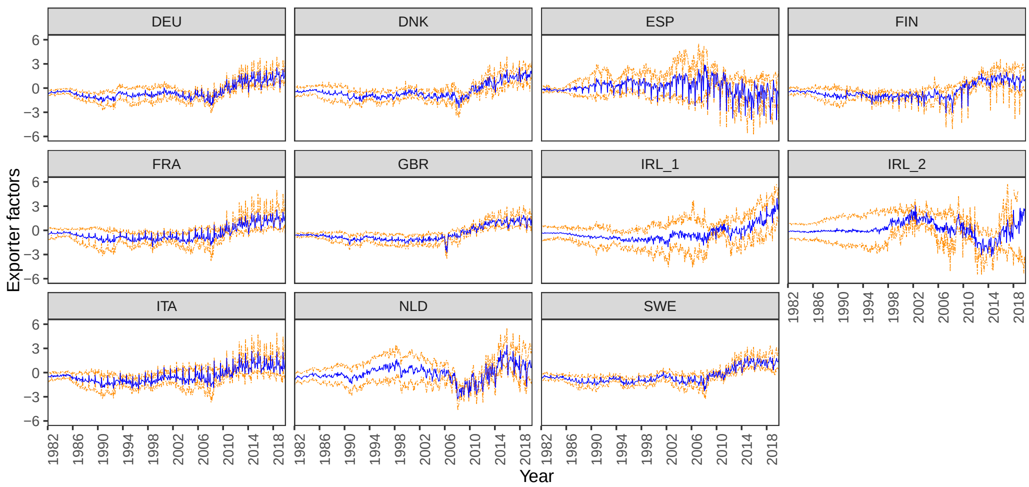

Figure 3: Exporter factors for 10 EU countries in the sample from January, 1982 to December, 2019. The blue solid line represents the estimated global factor and the orange dashed lines provide a 95% confidence interval.

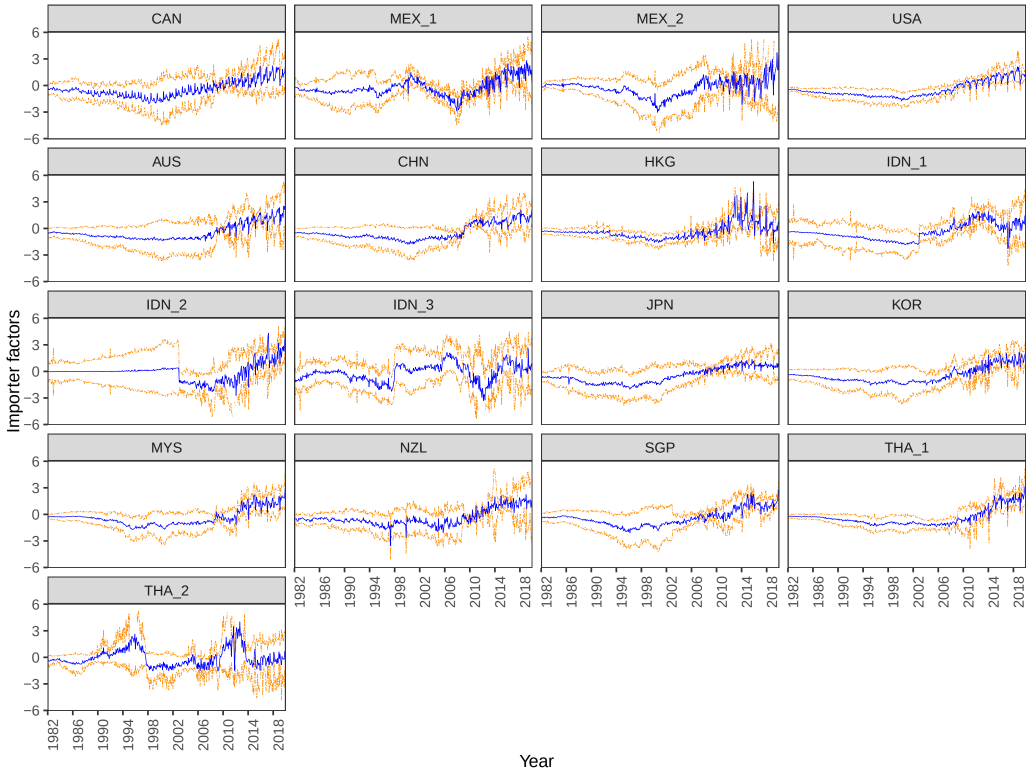

Figure 4: Importer factors for 13 APEC countries/region in the sample from January, 1982 to December, 2019. The blue solid line represents the estimated global factor and the orange dashed lines provide a 95% confidence interval.

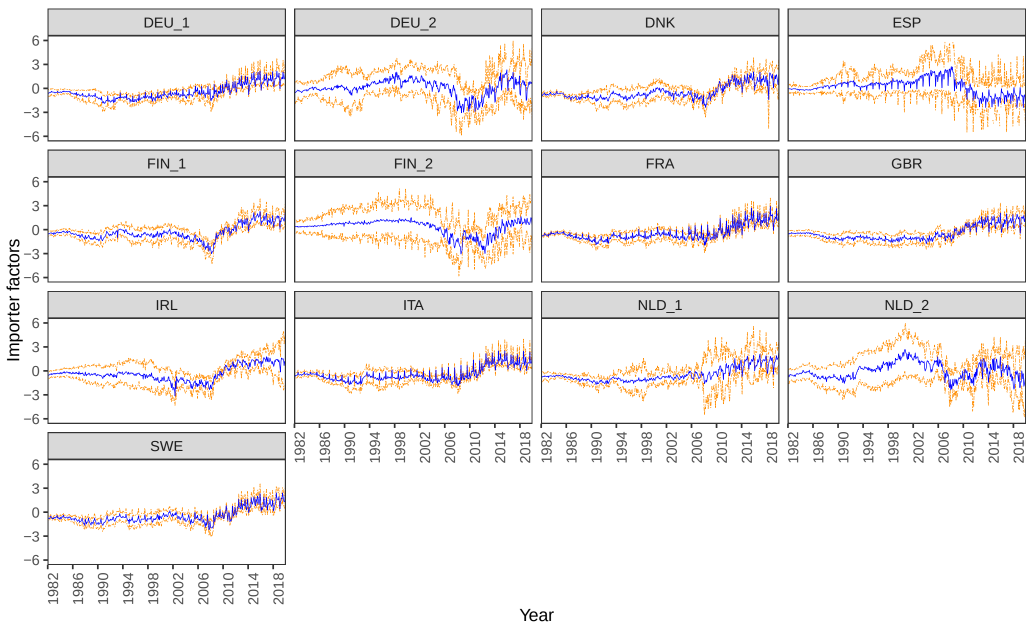

Figure 5: Importer factors for 10 EU countries in the sample from January, 1982 to December, 2019. The blue solid line represents the estimated global factor and the orange dashed lines provide a 95% confidence interval.

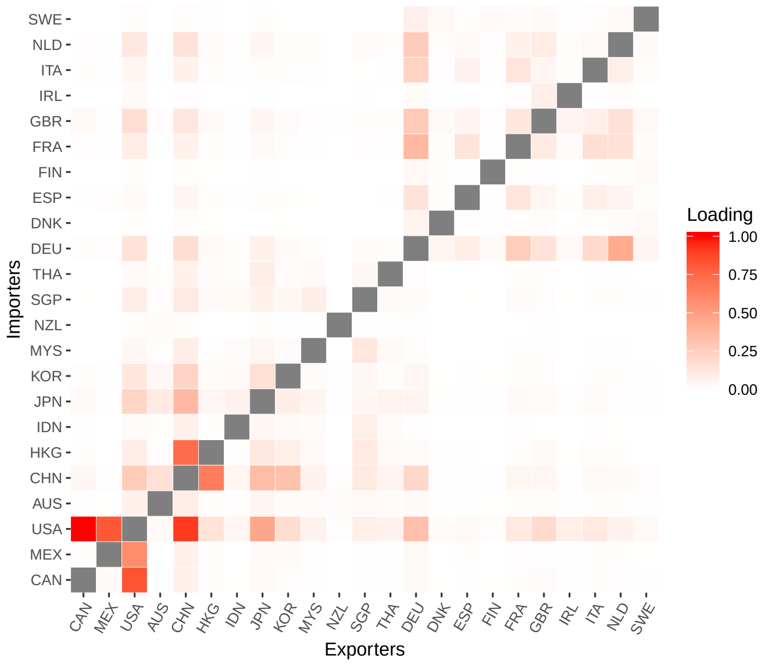

Figure 6: Global factor loadings for bilateral export flows. The corresponding figures are rescaled to have values between 0 and 1.

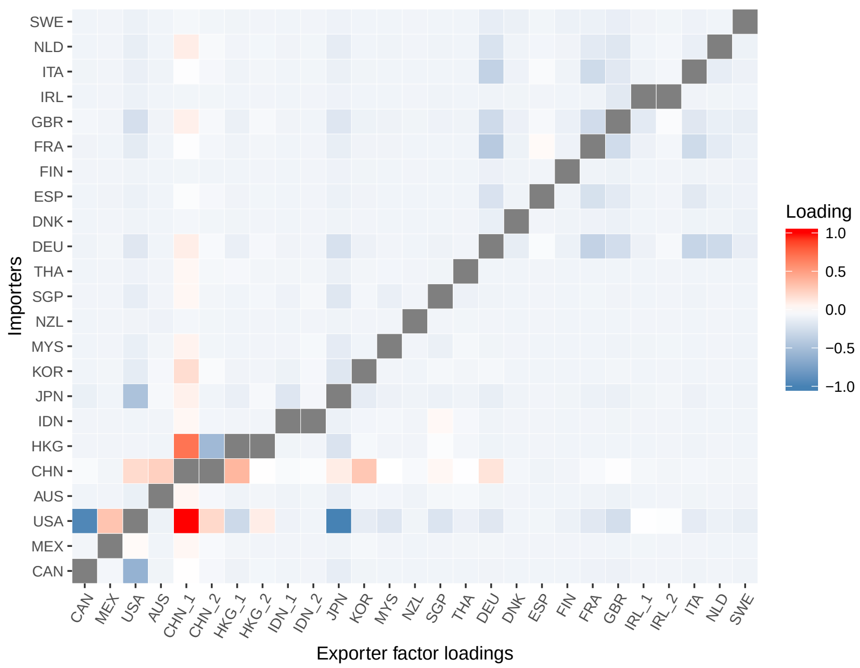

Figure 7: Exporter factor loadings for bilateral export flows. The corresponding figures are rescaled to have values between −1 and 1.

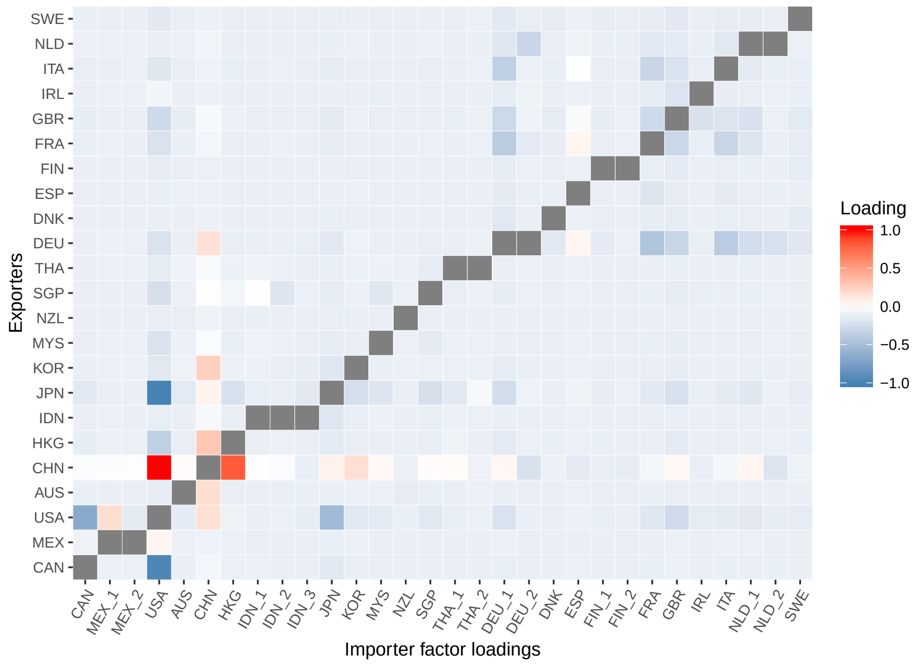

Figure 8: Importer factor loadings for bilateral export flows. The corresponding figures are rescaled to have values between −1 and 1.

Online Supplementary Appendices to
"Decomposition of Bilateral Trade Flows Using a Three-Dimensional Panel Data Model"

Yufeng Mao[♯], Bin Peng[♯], Mervyn Silvapulle[♯], Param Silvapulle[♯] and Yanrong Yang[*]

[♯]Monash University and [*]Australian National University

This file includes two appendices. Appendix A presents the preliminary lemmas, and the proofs of the main results. We relegate the secondary results and the associated proofs to Appendix B. Specifically, Appendix A.1 scratches the outline of the proofs, while Appendix A.2 presents the preliminary lemmas, which facilitate the development of the main results. Appendix A.3 summaries the proofs for each step. In Appendix B, Appendix B.1 states the secondary lemmas, while Appendix B.2 includes all the corresponding proofs.

Appendix A

A.1 Outline of the Proofs

In this section, we provide the outline of the proofs. Lemma A.1 presents the relevant results about the eigenvalues associated with $\frac{1}{MNT}\mathbf{Y}'\mathbf{Y}$ of (2.9), which yields the estimation of r_g, and thus leads to Theorem 2.1 of the main text. After establishing Theorem 2.1, we take the value of r_g as granted and focus on the estimation of country-specific factors. Specifically, we derive Lemma A.2 and Lemma A.3 in the same manner as Lemma A.1 by making use of the newly imposed Assumption 4, which allows us to further estimate the number of exporter factors and importer factors, and thus establish Theorem 2.2 of the main text.

After identifying the number of factors successfully, Lemma A.4-A.6 provide the further results in order to establish the asymptotic distributions associated with the global and country-specific factors. Specifically, Lemma A.4 improve the convergence rate of the global factor presented in Theorem 2.1 by making use of Assumption 4 and Assumption 5. Then we establish the first result of Theorem 2.3 in the main text which provides the asymptotic distribution for the global factor. Based on the results in Lemma A.4.1, Lemma A.5 updates the convergence rate of the country-specific factors presented in Theorem 2.2 of the main text. Similar to the development of the asymptotic distribution for the global factor, we then establish the asymptotic distributions for the exporter and importer factors in the second and third results of Theorem 2.3 in the main text after imposing Assumption 6.

A.2 Preliminary Lemmas

Lemma A.1. *Let Assumptions 1-3 hold, and let* $(M, N, T) \to (\infty, \infty, \infty)$.

1. $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}^{\dagger} - \mathbf{G}\mathbf{H}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$, *where* $\widehat{\mathbf{G}}^{\dagger}$ *includes the first* r_g *columns of* $\widehat{\mathbf{G}}$ *only,* $\mathbf{H} = \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^{\dagger} \cdot (\mathbf{V}_g^{\dagger})^{-1}$ *is an* $r_g \times r_g$ *invertible matrix,* $\mathbf{V}_g^{\dagger}$ *is the* $r_g \times r_g$ *leading principal submatrix of* $\mathbf{V}_g$ *and* $\|\mathbf{H}\|_F = O_P(1)$;
2. $|\widehat{\rho}_{g,k} - \rho_{g,k}^{\nabla}| = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$ *for* $k = 1, \ldots, r_g$, *where* $\widehat{\rho}_{g,k} = \frac{1}{T}\widehat{\mathbf{G}}_k^{\dagger\prime}\widehat{\mathbf{\Sigma}}\widehat{\mathbf{G}}_k^{\dagger}$ *with* $\widehat{\mathbf{G}}_k^{\dagger}$ *being the* k*-th column of* $\widehat{\mathbf{G}}^{\dagger}$, $\rho_{g,k}^{\nabla} = \frac{1}{T}\mathbf{h}_k'\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k$, $\widehat{\mathbf{\Sigma}} = \frac{1}{MNT}\mathbf{Y}'\mathbf{Y}$, $\mathbf{\Sigma} = \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{\Gamma}\mathbf{G}'$, *and* $\mathbf{h}_k$ *is the* k*-th column of* $\mathbf{H}$;
3. $|\widehat{\rho}_{g,k}| = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}}\right)$ *for* $k = r_g + 1, \ldots, k_{\max}$.

Lemma A.2. *Let Assumptions 1-4 hold, and let* $(M, N, T) \to (\infty, \infty, \infty)$. *For* $j = 1, \ldots, N$,

1. $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}_{I,j}^{\dagger} - \mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right)$, *where* $\widehat{\mathbf{F}}_{I,j}^{\dagger}$ *includes the first* $r_{I,j}$ *columns of* $\widehat{\mathbf{F}}_{I,j}$ *only,* $\mathbf{H}_{I,j} = \frac{1}{M}\mathbf{\Lambda}_{I,\bullet j}'\mathbf{\Lambda}_{I,\bullet j} \cdot \frac{1}{T}\mathbf{F}_{I,j}'\widehat{\mathbf{F}}_{I,j}^{\dagger} \cdot (\mathbf{V}_{I,j}^{\dagger})^{-1}$ *is an* $r_{I,j} \times r_{I,j}$ *invertible matrix,* $\mathbf{V}_{I,j}^{\dagger}$ *is the* $r_{I,j} \times r_{I,j}$ *leading principal submatrix of* $\mathbf{V}_{I,j}$ *and* $\|\mathbf{H}_{I,j}\|_F = O_P(1)$;
2. $|\widehat{\rho}_{Ij,k} - \rho_{Ij,k}^{\nabla}| = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right)$ *for* $k = 1, \ldots, r_{I,j}$, *where* $\widehat{\rho}_{Ij,k} = \frac{1}{T}\widehat{\mathbf{F}}_{Ij,k}^{\dagger\prime}\widehat{\mathbf{\Sigma}}_{I,j}\widehat{\mathbf{F}}_{Ij,k}^{\dagger}$ *with* $\widehat{\mathbf{F}}_{Ij,k}^{\dagger}$ *being the* k*-th column of* $\widehat{\mathbf{F}}_{I,j}^{\dagger}$, $\rho_{Ij,k}^{\nabla} = \frac{1}{T}\mathbf{h}_{Ij,k}'\mathbf{F}_{I,j}'\mathbf{\Sigma}_{I,j}\mathbf{F}_{I,j}\mathbf{h}_{Ij,k}$, $\widehat{\mathbf{\Sigma}}_{I,j} = \frac{1}{MT}(\mathbf{Y}_{I,j} - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{Y}_{I,j} - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')$, $\mathbf{\Sigma}_{I,j} = \frac{1}{MT}\mathbf{F}_{I,j}\mathbf{\Lambda}_{I,\bullet j}'\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'$, *and* $\mathbf{h}_{Ij,k}$ *is the* k*-th column of* $\mathbf{H}_{I,j}$;
3. $|\widehat{\rho}_{Ij,k}| = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}} + T^{2a_{I,j}}\right)$ *for* $k = r_{I,j} + 1, \ldots, k_{\max}$.

Lemma A.3. *Let Assumptions 1-4 hold, and let* $(M, N, T) \to (\infty, \infty, \infty)$. *For* $i = 1, \ldots, M$,

1. $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}_{E,i}^{\dagger} - \mathbf{F}_{E,i}\mathbf{H}_{E,i}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{E,i}}\right)$, *where* $\widehat{\mathbf{F}}_{E,i}^{\dagger}$ *includes the first* $r_{E,i}$ *columns of* $\widehat{\mathbf{F}}_{E,i}$ *only,* $\mathbf{H}_{E,i} = \frac{1}{N}\mathbf{\Lambda}_{E,i\bullet}'\mathbf{\Lambda}_{E,i\bullet} \cdot \frac{1}{T}\mathbf{F}_{E,i}'\widehat{\mathbf{F}}_{E,i}^{\dagger} \cdot (\mathbf{V}_{E,i}^{\dagger})^{-1}$ *is an* $r_{E,i} \times r_{E,i}$ *invertible matrix,* $\mathbf{V}_{E,i}^{\dagger}$ *is the* $r_{E,i} \times r_{E,i}$ *leading principal submatrix of* $\mathbf{V}_{E,i}$ *and* $\|\mathbf{H}_{E,i}\|_F = O_P(1)$;
2. $|\widehat{\rho}_{Ei,k} - \rho_{Ei,k}^{\nabla}| = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{E,i}}\right)$ *for* $k = 1, \ldots, r_{E,i}$, *where* $\widehat{\rho}_{Ei,k} = \frac{1}{T}\widehat{\mathbf{F}}_{Ei,k}^{\dagger\prime}\widehat{\mathbf{\Sigma}}_{E,i}\widehat{\mathbf{F}}_{Ei,k}^{\dagger}$ *with* $\widehat{\mathbf{F}}_{Ei,k}^{\dagger}$ *being the* k*-th column of* $\widehat{\mathbf{F}}_{E,i}^{\dagger}$, $\rho_{Ei,k}^{\nabla} = \frac{1}{T}\mathbf{h}_{Ei,k}'\mathbf{F}_{E,i}'\mathbf{\Sigma}_{E,i}\mathbf{F}_{E,i}\mathbf{h}_{Ei,k}$, $\widehat{\mathbf{\Sigma}}_{E,i} = \frac{1}{NT}(\mathbf{Y}_{E,i} - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}')'(\mathbf{Y}_{E,i} - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}')$, $\mathbf{\Sigma}_{E,i} = \frac{1}{NT}\mathbf{F}_{E,i}\mathbf{\Lambda}_{E,i\bullet}'\mathbf{\Lambda}_{E,i\bullet}\mathbf{F}_{E,i}'$, *and* $\mathbf{h}_{Ei,k}$ *is the* k*-th column of* $\mathbf{H}_{E,i}$;
3. $|\widehat{\rho}_{Ei,k}| = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}} + T^{2a_{E,i}}\right)$ *for* $k = r_{E,i} + 1, \ldots, k_{\max}$.

Lemma A.4. *Let Assumptions 1-3 and Assumption 5.1 hold. Denote that*

$$\Delta_{g,MNT} = \frac{\ln(MN) \cdot (T^{\max_i a_{E,i}} + T^{\max_j a_{I,j}})}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}$$
$$\text{and} \quad \Delta^*_{g,MNT} = \frac{T^{\max_i a_{E,i}}}{\sqrt{N}} + \frac{T^{\max_j a_{I,j}}}{\sqrt{M}} + \frac{\sqrt{\ln(MN)} \cdot (T^{\max_i a_{E,i}} + T^{\max_j a_{I,j}})}{\sqrt{T}} + \ln(MN) \cdot (T^{2\max_i a_{E,i}} + T^{2\max_j a_{I,j}})$$

for notational simplicity. As $(M,N,T) \to (\infty,\infty,\infty)$,

1. $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{T,\sqrt{MT},\sqrt{NT}\}} + \frac{1}{\sqrt{MN}} + \Delta_{g,MNT}\right)$;
2. $\frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F = O_P\left(\frac{1}{MN} + \frac{1}{T} + \Delta^*_{g,MNT}\right)$.

In addition, suppose that Assumption 5.2 holds. Then

3. $\mathbf{H} = \mathbf{I}_{r_g} + O_P\left(\frac{1}{MN} + \frac{1}{T} + \Delta^*_{g,MNT}\right)$.

Lemma A.5. *Let Assumptions 1-5 hold. Denote that*

$$\Delta_{Ij,MNT} = \Delta_{g,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_i a_{E,i}} + T^{a_{I,j}}$$
$$\text{and} \quad \Delta_{Ei,MNT} = \Delta_{g,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_j a_{I,j}} + T^{a_{E,i}}$$

for notational simplicity, where $\Delta_{g,MNT}$ *is defined in Lemma A.4. As* $(M,N,T) \to (\infty,\infty,\infty)$,

1. *for* $j = 1,\ldots,N$, $\frac{1}{\sqrt{MT}}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_F = O_P\left(\frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT}\right)$, *where* $\mathbf{\Gamma}_{I,j}$ *is defined in the same way as* $\widehat{\mathbf{\Gamma}}_{I,j}$;
2. *for* $i = 1,\ldots,M$, $\frac{1}{\sqrt{NT}}\|\mathbf{\Gamma}_{E,i}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}'\|_F = O_P\left(\frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ei,MNT}\right)$, *where* $\mathbf{\Gamma}_{E,i}$ *is defined in the same way as* $\widehat{\mathbf{\Gamma}}_{E,i}$.
3. *for* $j = 1,\ldots,N$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{T}\}} + \Delta_{Ij,MNT}\right)$;
4. *for* $i = 1,\ldots,M$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}_{E,i} - \mathbf{F}_{E,i}\mathbf{H}_{E,i}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{N},\sqrt{T}\}} + \Delta_{Ei,MNT}\right)$.

Lemma A.6. *Let Assumptions 1-5 and Assumption 6.1 hold, as* $(M,N,T) \to (\infty,\infty,\infty)$,

1. $\frac{1}{T}\|\mathbf{F}'_{I,j}(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_F = O_P\left(\frac{\ln(MN)}{M} + \frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_i b_{EI,ij}}\right)$;
2. $\frac{1}{T}\|\mathbf{F}'_{E,i}(\widehat{\mathbf{F}}_{E,i} - \mathbf{F}_{E,i}\mathbf{H}_{E,i})\|_F = O_P\left(\frac{\ln(MN)}{N} + \frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ei,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_j b_{EI,ij}}\right)$.

In addition, let Assumption 6.2 hold.

3. $\mathbf{H}_{I,j} = \mathbf{I}_{r_{I,j}} + O_P\left(\frac{\ln(MN)}{M} + \frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_i b_{EI,ij}}\right)$;
4. $\mathbf{H}_{E,j} = \mathbf{I}_{r_{E,j}} + O_P\left(\frac{\ln(MN)}{N} + \frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ei,MNT} + \sqrt{\ln(MN)} \cdot T^{\max_j b_{EI,ij}}\right)$.

A.3 Proofs

We now start presenting the proofs..

A.3.1 On Consistency

Proof of Lemma A.1:

(1). Expanding (2.9), we write

$$\begin{aligned}\widehat{\mathbf{G}}\mathbf{V}_g &= \frac{1}{MNT}(\mathbf{\Gamma}\mathbf{G}' + \mathbf{\Lambda}_E\mathbf{F}_E' + \mathbf{\Lambda}_I\mathbf{F}_I' + \mathbf{U})'(\mathbf{\Gamma}\mathbf{G}' + \mathbf{\Lambda}_E\mathbf{F}_E' + \mathbf{\Lambda}_I\mathbf{F}_I' + \mathbf{U})\widehat{\mathbf{G}} \\ &= (\mathbf{A}_1 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}}\end{aligned}$$

where

$$\begin{aligned}&\mathbf{A}_1 = \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{\Gamma}\mathbf{G}', \quad \mathbf{A}_2 = \frac{1}{MNT}\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_E\mathbf{F}_E', \quad \mathbf{A}_3 = \frac{1}{MNT}\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Lambda}_I\mathbf{F}_I', \\ &\mathbf{A}_4 = \frac{1}{MNT}\mathbf{U}'\mathbf{U}, \quad \mathbf{A}_5 = \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_E\mathbf{F}_E', \quad \mathbf{A}_6 = \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_I\mathbf{F}_I', \\ &\mathbf{A}_7 = \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{U}, \quad \mathbf{A}_8 = \frac{1}{MNT}\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Gamma}\mathbf{G}', \quad \mathbf{A}_9 = \frac{1}{MNT}\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_I\mathbf{F}_I', \\ &\mathbf{A}_{10} = \frac{1}{MNT}\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{U}, \quad \mathbf{A}_{11} = \frac{1}{MNT}\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Gamma}\mathbf{G}', \quad \mathbf{A}_{12} = \frac{1}{MNT}\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Lambda}_E\mathbf{F}_E', \\ &\mathbf{A}_{13} = \frac{1}{MNT}\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{U}, \quad \mathbf{A}_{14} = \frac{1}{MNT}\mathbf{U}'\mathbf{\Gamma}\mathbf{G}', \quad \mathbf{A}_{15} = \frac{1}{MNT}\mathbf{U}'\mathbf{\Lambda}_E\mathbf{F}_E', \\ &\mathbf{A}_{16} = \frac{1}{MNT}\mathbf{U}'\mathbf{\Lambda}_I\mathbf{F}_I'.\end{aligned}$$

Thus, the following equality holds immediately.

$$\widehat{\mathbf{G}}\mathbf{V}_g - \mathbf{G}\cdot\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\cdot\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}} = (\mathbf{A}_2 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}}, \tag{A.1}$$

in which

$$\|\mathbf{A}_2 + \cdots + \mathbf{A}_{16}\|_2 \le \|\mathbf{A}_2\|_2 + \cdots + \|\mathbf{A}_{16}\|_2 = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right)$$

by Lemma B.2.

We left multiply (A.1) by $\frac{1}{T}\mathbf{G}'$ to yield

$$\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}\cdot\mathbf{V}_g - \frac{1}{T}\mathbf{G}'\mathbf{G}\cdot\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\cdot\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}} = \frac{1}{T}\mathbf{G}'(\mathbf{A}_2 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}} = o_P(1),$$

because $\frac{1}{\sqrt{T}}\|\mathbf{G}\|_F = O_P(1)$, $\|\mathbf{A}_2 + \cdots + \mathbf{A}_{16}\|_2 = o_P(1)$ and $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}\|_F = O_P(1)$. Furthermore, $\frac{1}{T}\mathbf{G}'\mathbf{G}\cdot\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{\Sigma}_{\mathbf{G}}\mathbf{\Sigma}_{\mathbf{\Gamma}} + o_P(1)$ by Assumptions 1 and 2. It follows that

$$\mathbf{\Sigma}_{\mathbf{G}}\mathbf{\Sigma}_{\mathbf{\Gamma}}\cdot\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}} = \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}\cdot\mathbf{V}_g + o_P(1). \tag{A.2}$$

Note that $\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}$ is of rank r_g, which then further indicates that $\mathbf{V}_g$ has at least r_g non-zero elements on the main diagonal which converge to the eigenvalues of $\mathbf{\Sigma_G \Sigma_\Gamma}$. We now can conclude that $\mathbf{V}_g$ is of rank r_g in the limit.

We are now ready to investigate $\widehat{\rho}_{g,k}$ with $k \leq r_g$. Because here $k \leq r_g$, we then focus on $\widehat{\mathbf{G}}^\dagger$ and $\mathbf{V}_g^\dagger$, which are defined in this Lemma. In particular, by the argument under (A.2), we can conclude that $\mathbf{V}_g^\dagger \to_P \mathbf{V}_{0g}$, where $\mathbf{V}_{0g} = \mathrm{diag}\{\rho_{g,1}, \ldots, \rho_{g,r_g}\}$ consisting of the eigenvalues of $\mathbf{\Sigma_G \Sigma_\Gamma}$. By (A.1), we can write

$$\widehat{\mathbf{G}}^\dagger \mathbf{V}_g^\dagger - \mathbf{G} \cdot \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger = (\mathbf{A}_2 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}}^\dagger. \tag{A.3}$$

Left multiplying (A.3) by $\frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}$ and using the fact that $\frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}\widehat{\mathbf{G}}^\dagger = \mathbf{I}_{r_g}$, we have

$$\mathbf{V}_g^\dagger - \frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}\mathbf{G} \cdot \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger = \frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}(\mathbf{A}_2 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}}^\dagger = o_P(1), \tag{A.4}$$

because $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}^\dagger\|_F = O(1)$ and $\|\mathbf{A}_2 + \cdots + \mathbf{A}_{16}\|_2 = o_P(1)$ by Lemma B.2. By (A.4) and $\mathbf{V}_g^\dagger \to_P \mathbf{V}_{0g}$, it follows that

$$\frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}\mathbf{G} \cdot \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger \to_P \mathbf{V}_{0g}. \tag{A.5}$$

Hence, $\frac{1}{T}\widehat{\mathbf{G}}^{\dagger\prime}\mathbf{G} \cdot \frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger$ is invertible with probability approaching one (w.p.a.1), which implies that $\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger$ is invertible w.p.a.1.

Right multiplying $(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger)^{-1}(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma})^{-1}$ on each side of (A.3) yields that

$$\begin{aligned}&\widehat{\mathbf{G}}^\dagger \mathbf{V}_g^\dagger \left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1} - \mathbf{G}\\ &= (\mathbf{A}_2 + \cdots + \mathbf{A}_{16})\widehat{\mathbf{G}}^\dagger \left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1}.\end{aligned} \tag{A.6}$$

By examining each term on the right hand side of (A.6), we obtain that

$$\begin{aligned}&\frac{1}{\sqrt{T}}\left\|\widehat{\mathbf{G}}^\dagger \mathbf{V}_g^\dagger \left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1} - \mathbf{G}\right\|_F\\ &\leq O(1)\cdot(\|\mathbf{A}_2\|_2 + \cdots + \|\mathbf{A}_{16}\|_2)\cdot\left\|\frac{1}{\sqrt{T}}\widehat{\mathbf{G}}^\dagger \cdot \left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1}\right\|_F\\ &\leq O_P(1)\cdot(\|\mathbf{A}_2\|_2 + \cdots + \|\mathbf{A}_{16}\|_2) = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right).\end{aligned}$$

where the first inequality follows from (A.6) and $\|\mathbf{\Omega}\|_F \leq \sqrt{\mathrm{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the second inequality follows from $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}^\dagger\|_F = O_P(1)$ and $\|(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}^\dagger)^{-1}(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma})^{-1}\|_F = O_P(1)$, and the equality follows from Lemma B.2.

Moreover, we have $\|\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\|_F = O_P(1)$ by Assumption 2. We have shown that $\|(\mathbf{V}_g^\dagger)^{-1}\|_F = O_P(1)$ since $\mathbf{V}_g^\dagger$ converges to a full rank matrix, and

$$\frac{1}{T}\|\mathbf{G}'\widehat{\mathbf{G}}^{\dagger}\|_F \leq \frac{1}{\sqrt{T}}\|\mathbf{G}\|_F \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}^{\dagger}\|_F = O_P(1).$$

Thus, it follows that

$$\|\mathbf{H}\|_F \leq \frac{1}{MN}\|\mathbf{\Gamma}'\mathbf{\Gamma}\|_F \cdot \frac{1}{T}\|\mathbf{G}'\widehat{\mathbf{G}}^{\dagger}\|_F \cdot \|(\mathbf{V}_g^{\dagger})^{-1}\|_F = O_P(1).$$

(2). Let us now consider $\widehat{\rho}_{g,k} - \rho_{g,k}^{\triangledown}$. By construction of $\widehat{\rho}_{g,k}$ and $\rho_{g,k}^{\triangledown}$, write

$$\begin{aligned}
\widehat{\rho}_{g,k} - \rho_{g,k}^{\triangledown} &= \frac{1}{T}\widehat{\mathbf{G}}_k^{\dagger\prime}\widehat{\mathbf{\Sigma}}\widehat{\mathbf{G}}_k^{\dagger} - \frac{1}{T}\mathbf{h}_k'\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k \\
&= \frac{1}{T}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k + \mathbf{G}\mathbf{h}_k)'(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma} + \mathbf{\Sigma})(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k + \mathbf{G}\mathbf{h}_k) - \frac{1}{T}\mathbf{h}_k'\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k \\
&= \frac{1}{T}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma})(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k) \\
&\quad + \frac{2}{T}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma})\mathbf{G}\mathbf{h}_k \\
&\quad + \frac{1}{T}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'\mathbf{\Sigma}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k) \\
&\quad + \frac{2}{T}(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k \\
&\quad + \frac{1}{T}\mathbf{h}_k'\mathbf{G}'(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma})\mathbf{G}\mathbf{h}_k \\
&:= J_1 + 2J_2 + J_3 + 2J_4 + J_5
\end{aligned}$$

Start our analysis from J_1, and write

$$\begin{aligned}
|J_1| &= \frac{1}{T}|(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma})(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)| \\
&\leq \frac{1}{T}\|\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k\|_F \cdot \|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_2 \cdot \|\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k\|_F \\
&= o_P(\|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_2) = o_P(|J_5|).
\end{aligned}$$

Similarly, we have $|J_2| = o_P(|J_5|)$ and $|J_3| = o_P(|J_4|)$. It remains to consider the orders of J_4 and J_5.

For $|J_4|$, write

$$\begin{aligned}
|J_4| &= \frac{1}{T}|(\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k)'\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k| \leq \frac{1}{T}\|\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k\|_F \cdot \|\mathbf{\Sigma}\mathbf{G}\mathbf{h}_k\|_F \\
&= \frac{1}{T}\|\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k\|_F \cdot \frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{\Gamma}\mathbf{G}'\mathbf{G}\mathbf{h}_k\|_F \\
&\leq \frac{1}{T}\|\widehat{\mathbf{G}}_k^{\dagger} - \mathbf{G}\mathbf{h}_k\|_F \cdot \frac{1}{MNT}\|\mathbf{G}\|_F \cdot \|\mathbf{\Gamma}'\mathbf{\Gamma}\|_F \cdot \|\mathbf{G}'\mathbf{G}\|_F \cdot \|\mathbf{h}_k\|_F \\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right),
\end{aligned}$$

where the last equality follows from Lemma A.1.

For $|J_5|$, write

$$\|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_2 = \left\|\frac{1}{MNT}\mathbf{Y}'\mathbf{Y} - \frac{1}{MNT}\mathbf{G}\mathbf{\Gamma}'\mathbf{\Gamma}\mathbf{G}'\right\|_2 \leq \|\mathbf{A}_2\|_2 + \cdots + \|\mathbf{A}_{16}\|_2$$

$$
= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right),
$$

where $\mathbf{A}_2,\ldots,\mathbf{A}_{16}$ have been defined in the proof of Lemma A.1.

Thus,

$$
\begin{aligned}
|J_5| &= \frac{1}{T}|\mathbf{h}_k'\mathbf{G}'(\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma})\mathbf{G}\mathbf{h}_k| \le \frac{1}{T}\|\mathbf{h}_k\|_F^2\cdot\|\mathbf{G}\|_F^2\cdot\|\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}\|_2 \\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right).
\end{aligned}
$$

Hence, based on the above development,

$$
|\widehat{\rho}_{g,k}-\rho_{g,k}^{\triangledown}| = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right),
$$

which completes the proof of the first result.

(3). This proof is based on Lemma B.1. Let us denote $\mathbf{G}^\perp$ as a $T\times(k_{\max}-r_g)$ matrix such that $\frac{1}{T}(\mathbf{G}^\perp,\mathbf{G}\mathbf{R})'(\mathbf{G}^\perp,\mathbf{G}\mathbf{R})=\mathrm{diag}\{\mathbf{I}_{k_{\max}-r_g},\mathbf{I}_{r_g}\}$, where $\mathbf{R}$ is an $r_g\times r_g$ rotation matrix. The matrices $\frac{1}{\sqrt{T}}\mathbf{G}^\perp$, $\frac{1}{\sqrt{T}}\mathbf{G}\mathbf{R}$, $\mathbf{\Sigma}$ and $\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}$ correspond to $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{A}$ and $\mathbf{E}$ of Lemma B.1. The counterpart of the matrix $\mathbf{Q}_1^0$ of Lemma B.1 is given by

$$
\widehat{\mathbf{G}}^\perp = \frac{1}{\sqrt{T}}(\mathbf{G}^\perp+\mathbf{G}\mathbf{R}\mathbf{P})(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2},
$$

where

$$
\|\mathbf{P}\|_2 \le \frac{4\|\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}\|_2}{\mathrm{sep}(0,\frac{1}{T}\mathbf{G}'\mathbf{\Sigma}\mathbf{G})} \le O_P(1)\cdot\|\widehat{\mathbf{\Sigma}}-\mathbf{\Sigma}\|_2 = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right).
$$

Since $\widehat{\mathbf{G}}^\perp$ is an orthonormal basis for a subspace that is invariant for $\widehat{\mathbf{\Sigma}}$, we have $\rho_{r_g+k}=\widehat{\mathbf{G}}_k^{\perp\prime}\widehat{\mathbf{\Sigma}}\widehat{\mathbf{G}}_k^\perp$, where $k=1,\ldots,k_{\max}-r$, and $\widehat{\mathbf{G}}_k^\perp$ is the k-th column of $\widehat{\mathbf{G}}^\perp$. Consider $\|\widehat{\mathbf{G}}^\perp-\frac{1}{\sqrt{T}}\mathbf{G}^\perp\|_2$, and write

$$
\begin{aligned}
\left\|\widehat{\mathbf{G}}^\perp-\frac{1}{\sqrt{T}}\mathbf{G}^\perp\right\|_2 &= \frac{1}{\sqrt{T}}\|[\mathbf{G}^\perp+\mathbf{G}\mathbf{R}\mathbf{P}-\mathbf{G}^\perp(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{1/2}](\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\
&\le \frac{1}{\sqrt{T}}\|\mathbf{G}^\perp(\mathbf{I}_{k_{\max}-r_g}-(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{1/2})(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\
&\quad +\frac{1}{\sqrt{T}}\|\mathbf{G}\mathbf{R}\mathbf{P}(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\
&\le \|(\mathbf{I}_{k_{\max}-r_g}-(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{1/2})(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\
&\quad +\|\mathbf{P}(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\
&\le \|\mathbf{I}_{k_{\max}-r_g}-(\mathbf{I}_{k_{\max}-r_g}+\mathbf{P}'\mathbf{P})^{1/2}\|_2+\|\mathbf{P}\|_2 \\
&\le 2\|\mathbf{P}\|_2 = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right).
\end{aligned} \tag{A.7}
$$

where the second and third inequalities follow from Exercise 1 on page 231 and Magnus and Neudecker (2019). Thus, for $k = 1, \ldots, k_{\max} - r_g$,

$$
\begin{aligned}
|\widehat{\rho}_{r+k}| &= |\widehat{\mathbf{G}}_k^{\perp\prime}\widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{G}}_k^{\perp}| \\
&= |(\widehat{\mathbf{G}}_k^{\perp} - T^{-1/2}\mathbf{G}_k^{\perp} + T^{-1/2}\mathbf{G}_k^{\perp})'(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma} + \boldsymbol{\Sigma})(\widehat{\mathbf{G}}_k^{\perp} - T^{-1/2}\mathbf{G}_k^{\perp} + T^{-1/2}\mathbf{G}_k^{\perp})| \\
&\leq \|\widehat{\mathbf{G}}_k^{\perp} - T^{-1/2}\mathbf{G}_k^{\perp}\|_F^2 \cdot \|\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_2 \\
&\quad + 2\|\widehat{\mathbf{G}}_k^{\perp} - T^{-1/2}\mathbf{G}_k^{\perp}\|_F^2 \cdot \|\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_2 \cdot \frac{1}{\sqrt{T}}\|\mathbf{G}_k^{\perp}\|_F \\
&\quad + \|\widehat{\mathbf{G}}_k^{\perp} - T^{-1/2}\mathbf{G}_k^{\perp}\|_F^2 \cdot \|\boldsymbol{\Sigma}\|_2 \\
&= O_P\left(\frac{\ln(MN)}{\min\{M, N, T\}}\right),
\end{aligned}
$$

where $\mathbf{G}_k^{\perp}$ is the k-th column of $\mathbf{G}^{\perp}$, and the last equality follows from (A.7) and the proof of the first result of this lemma. The proof is complete. ■

Proof of Theorem 2.1:

(1). First, consider the case where $r_g = 0$. By Lemma A.1.3, we have $\widehat{\rho}_{g,k} = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}}\right)$ for $k = 1, \ldots, k_{\max}$, which are less than ω_{MNT} w.p.a.1 for $k = 1, \ldots, k_{\max}$. In connection with the fact that $\widehat{\rho}_{g,0} = 1$ which is larger than ω_{MNT} and $\widehat{\rho}_{g,1}/\widehat{\rho}_{g,0} = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}}\right)$ which is less than 1 w.p.a.1, it follows that when $r_g = 0$, we have $P(\widehat{r}_g = 0) \to 1$.

Next, we consider the case where $0 < r_g < k_{\max}$. Note that for $k = 1, \ldots, r_g$,

$$
\begin{aligned}
\rho_{g,k}^{\triangledown} &= \frac{1}{T}\mathbf{h}_k'\mathbf{G}'\boldsymbol{\Sigma}\mathbf{G}\mathbf{h}_k = \frac{1}{T}\mathbf{h}_k'\mathbf{G}'\mathbf{G} \cdot \frac{1}{MN}\boldsymbol{\Gamma}'\boldsymbol{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\mathbf{G}\mathbf{h}_k \\
&= \frac{1}{T}\widehat{\mathbf{G}}_k^{\dagger\prime}\mathbf{G} \cdot \frac{1}{MN}\boldsymbol{\Gamma}'\boldsymbol{\Gamma} \cdot \frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}_k^{\dagger} \cdot (1 + o_P(1)) \asymp 1,
\end{aligned}
$$

where $\mathbf{h}_k$, $\boldsymbol{\Sigma}$ and $\widehat{\mathbf{G}}_k^{\dagger}$ have been defined in the proof of Lemma B.1, the third equality follows from Lemma A.1.1, and the last step follows from (A.5).

Note that $\widehat{\rho}_{g,0} = 1$ which is larger than ω_{MNT}. Furthermore, by Lemma A.1.2, $\widehat{\rho}_{g,k} \asymp \rho_{g,k}^{\triangledown}$ for $k = 1, \ldots, r_g$, which are larger than ω_{MNT} w.p.a.1. Moreover, for $k = 0, \ldots, r_g - 1$ we can conclude that

$$
\frac{\widehat{\rho}_{g,k+1}}{\widehat{\rho}_{g,k}} \asymp 1
$$

since $\widehat{\rho}_{g,0} = 1$, $\widehat{\rho}_{g,k} \asymp \rho_{g,k}^{\triangledown}$ and $\rho_{g,k}^{\triangledown} \asymp 1$ for $k = 1, \ldots, r_g$.

For $k = r_g + 1, \ldots, k_{\max}$, by Lemma A.1.3, $\widehat{\rho}_{g,k} = O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}}\right)$, which are less than ω_{MNT} w.p.a.1. Thus, in connection with the fact that $\widehat{\rho}_{g,r_g} \asymp 1$, we obtain that

$$
\frac{\widehat{\rho}_{g,r_g+1}}{\widehat{\rho}_{g,r_g}} = O_P\left(\frac{\ln(MN)}{\min\{M, N, T\}}\right),
$$

which is less than 1 w.p.a.1.

Based on the above and (2.10), we are readily to conclude that $P(\widehat{r}_g = r_g) \to 1$. The first result then follows.

(2). Having establishing the first result, the second step follows from almost an identical procedure of the proof for Lemma B.1. Thus, omitted. The proof is complete. ■

Proof of Lemma A.2:

Before proceeding further, we denote a few notations for simplicity. Let $\mathbf{\Gamma}_{I,j}$ be denoted in the same way as $\widehat{\mathbf{\Gamma}}_{I,j}$, $\mathbf{U}_{\bullet j \bullet} = \{u_{ijt}\}_{M\times T}$.

By expanding (2.11), we obtain that

$$\begin{aligned}
\widehat{\mathbf{F}}_{I,j}\mathbf{V}_{I,j} \;=\; & \frac{1}{MT}\left(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}' + \mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E + \mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j} + \mathbf{U}_{\bullet j\bullet}\right)' \\
& \cdot\left(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}' + \mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E + \mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j} + \mathbf{U}_{\bullet j\bullet}\right)\widehat{\mathbf{F}}_{I,j} \\
:= & (\mathbf{B}_1 + \cdots + \mathbf{B}_{16})\widehat{\mathbf{F}}_{I,j} \qquad \text{(A.8)}
\end{aligned}$$

where

$$\begin{aligned}
&\mathbf{B}_1 = \frac{1}{MT}\mathbf{F}_{I,j}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}, \quad \mathbf{B}_2 = \frac{1}{MT}(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E, \\
&\mathbf{B}_3 = \frac{1}{MT}(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}, \quad \mathbf{B}_4 = \frac{1}{MT}(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{U}_{\bullet j\bullet}, \\
&\mathbf{B}_5 = \mathbf{B}'_2, \quad \mathbf{B}_6 = \frac{1}{MT}\mathbf{F}_E\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E, \quad \mathbf{B}_7 = \frac{1}{MT}\mathbf{F}_E\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j} \\
&\mathbf{B}_8 = \frac{1}{MT}\mathbf{F}_E\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{U}_{\bullet j\bullet}, \quad \mathbf{B}_9 = \mathbf{B}'_3, \quad \mathbf{B}_{10} = \mathbf{B}'_7, \\
&\mathbf{B}_{11} = \frac{1}{MT}(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'), \quad \mathbf{B}_{12} = \frac{1}{MT}\mathbf{F}_{I,j}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{U}_{\bullet j\bullet}, \\
&\mathbf{B}_{13} = \mathbf{B}'_4, \quad \mathbf{B}_{14} = \mathbf{B}'_8, \quad \mathbf{B}_{15} = \mathbf{B}'_{12}, \quad \mathbf{B}_{16} = \frac{1}{MT}\mathbf{U}'_{\bullet j\bullet}\mathbf{U}_{\bullet j\bullet}.
\end{aligned}$$

We can then immediately obtain that

$$\widehat{\mathbf{F}}_{I,j}\mathbf{V}_{I,j} - \mathbf{F}_{I,j}\cdot\frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}\cdot\frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j} = (\mathbf{B}_2 + \cdots + \mathbf{B}_{16})\widehat{\mathbf{F}}_{I,j}. \qquad \text{(A.9)}$$

By Lemma B.5, we obtain that

$$\|\mathbf{B}_2 + \cdots + \mathbf{B}_{16}\|_2 \le \|\mathbf{B}_2\|_2 + \cdots + \|\mathbf{B}_{16}\|_2 = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right).$$

Then left multiplying (A.9) by $\frac{1}{T}\mathbf{F}'_{I,j}$ immediately yields that

$$\begin{aligned}
&\left\|\frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j}\mathbf{V}_{I,j} - \frac{1}{T}\mathbf{F}'_{I,j}\mathbf{F}_{I,j}\cdot\frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}\cdot\frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j}\right\|_F \\
&\le O(1)\frac{1}{\sqrt{T}}\|\mathbf{F}'_{I,j}\|_F\cdot\|\mathbf{B}_2 + \cdots + \mathbf{B}_{16}\|_2\cdot\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}_{I,j}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right),
\end{aligned}$$

where the first inequality follows from (A.9) and $\|\mathbf{\Omega}\|_F \leq \sqrt{\text{rank}(\mathbf{\Omega})} \cdot \|\mathbf{\Omega}\|_2$.

Furthermore, $\frac{1}{T}\mathbf{F}'_{I,j}\mathbf{F}_{I,j} \cdot \frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j} = \mathbf{\Sigma}_{\mathbf{F}_{I,j}}\mathbf{\Sigma}_{\mathbf{\Lambda}_{I,\bullet j}} + o_P(1)$ by Assumptions 4. It follows that

$$\mathbf{\Sigma}_{\mathbf{F}_{I,j}}\mathbf{\Sigma}_{\mathbf{\Lambda}_{I,\bullet j}} \cdot \frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j} = \frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j}\mathbf{V}_{I,j} + o_P(1).$$

Then the first result follows in a manner similar to the first result of Lemma A.1.

(2)-(3). The second and third results follow the proofs similar to the second and third results of Lemma A.1, so the details are omitted. The proof is now complete. ■

Proof of Lemma A.3:

The proofs for the results of Lemma A.3 are identical to those for Lemma A.2. Thus, omitted. ■

Proof of Theorem 2.2:

(1). Having established Lemma A.2, the proof for the first result of this theorem is identical to the proof of Theorem 2.1

(2). The proof for the second result of this theorem can be achieved in a manner similar to the first result, so omitted as well. ■

A.3.2 On Asymptotic Distribution

Proof of Lemma A.4:

(1). The proof is largely the same as Lemma A.1, but we improve the rates of convergence on the following terms. We first improve the convergence rates of the terms $\frac{1}{\sqrt{T}}\|\mathbf{A}_5\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_6\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_8\widehat{\mathbf{G}}\|_2$ and $\frac{1}{\sqrt{T}}\|\mathbf{A}_{11}\widehat{\mathbf{G}}\|_2$ of Lemma A.1 using both Assumption 4.1.(a) and Assumption 5.1.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_5\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{G}\|_F \cdot \|\mathbf{\Gamma}'\mathbf{\Lambda}_E\|_F \cdot \left(\|\mathbf{F}'_E(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}'_E\mathbf{G}\mathbf{H}\|_F\right) \\
&\leq O_P(1)\frac{1}{MNT\sqrt{T}} \cdot \sqrt{T} \cdot \sqrt{MN} \cdot \left((\sqrt{M} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + T\sqrt{M} \cdot T^{\max_i a_{E,i}}\right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + O_P\left(\frac{T^{\max_i a_{E,i}}}{\sqrt{N}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a), Assumption 5.1. and $\|\mathbf{H}\|_F = O_P(1)$.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_6\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{G}\|_F \cdot \|\mathbf{\Gamma}'\mathbf{\Lambda}_I\|_F \cdot \left(\|\mathbf{F}'_I(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}'_I\mathbf{G}\mathbf{H}\|_F\right) \\
&\leq O_P(1)\frac{1}{MNT\sqrt{T}} \cdot \sqrt{T} \cdot \sqrt{MN} \cdot \left((\sqrt{N} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + T\sqrt{N} \cdot T^{\max_j a_{I,j}}\right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + O_P\left(\frac{T^{\max_j a_{I,j}}}{\sqrt{M}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a), Assumption 5.1 and $\|\mathbf{H}\|_F = O_P(1)$.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_8\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{F}_E\|_2 \cdot \|\mathbf{\Lambda}_E'\mathbf{\Gamma}\|_F \cdot \|\mathbf{G}'\widehat{\mathbf{G}}\|_F \\
&\leq O_P(1)\frac{1}{MNT\sqrt{T}}(\sqrt{M} \vee \sqrt{T}) \cdot \sqrt{MN} \cdot T \\
&= O_P(1)\left(\frac{1}{\sqrt{NT}} + \frac{1}{\sqrt{MN}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1 and Assumption 5.1.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_{11}\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{F}_I\|_2 \cdot \|\mathbf{\Lambda}_I'\mathbf{\Gamma}\|_F \cdot \|\mathbf{G}'\widehat{\mathbf{G}}\|_F \\
&\leq O_P(1)\frac{1}{MNT\sqrt{T}}(\sqrt{N} \vee \sqrt{T}) \cdot \sqrt{MN} \cdot T \\
&= O_P(1)\left(\frac{1}{\sqrt{MT}} + \frac{1}{\sqrt{MN}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1 and Assumption 5.1.

Next, we improve the convergence rates of the terms $\frac{1}{\sqrt{T}}\|\mathbf{A}_2\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_3\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_9\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_{12}\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_{15}\widehat{\mathbf{G}}\|_2$ and $\frac{1}{\sqrt{T}}\|\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2$ of Lemma A.1 by making use of Assumption 4.1.(a).

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_2\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{F}_E\|_2 \cdot \|\mathbf{\Lambda}_E\|_2^2 \cdot \left(\|\mathbf{F}_E'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}_E'\mathbf{G}\mathbf{H}\|_F\right) \\
&\leq \frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{M} \vee \sqrt{T}) \cdot N\ln(MN) \\
&\quad \cdot \left((\sqrt{M} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + T\sqrt{M} \cdot T^{\max_i a_{E,i}}\right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + \ln(MN) \cdot O_P\left(\frac{T^{\max_i a_{E,i}}}{\sqrt{M} \wedge \sqrt{T}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a) and the facts that $\|\mathbf{H}\|_F = O_P(1)$ and $\|\mathbf{\Lambda}_E\|_2^2 = O_P(N\ln(MN))$.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_3\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{F}_I\|_2 \cdot \|\mathbf{\Lambda}_I\|_2^2 \cdot \left(\|\mathbf{F}_I'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}_I'\mathbf{G}\mathbf{H}\|_F\right) \\
&\leq \frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{N} \vee \sqrt{T}) \cdot M\ln(MN) \\
&\quad \cdot \left((\sqrt{N} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + T\sqrt{N} \cdot T^{\max_j a_{I,j}}\right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + \ln(MN) \cdot O_P\left(\frac{T^{\max_j a_{I,j}}}{\sqrt{N} \wedge \sqrt{T}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a) and the facts that $\|\mathbf{H}\|_F = O_P(1)$ and $\|\mathbf{\Lambda}_I\|_2^2 = O_P(M\ln(MN))$.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_9\widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_I\|_2 \cdot \left(\|\mathbf{F}_I'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}_I'\mathbf{G}\mathbf{H}\|_F\right) \\
&\leq \frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{M} \vee \sqrt{T}) \cdot \sqrt{M\ln(MN)} \cdot \sqrt{N\ln(MN)}
\end{aligned}$$

$$
\begin{aligned}
&\cdot\Big((\sqrt{N}\vee\sqrt{T})\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+T\sqrt{N}\cdot T^{\max_j a_{I,j}}\Big)\\
&=o_P(1)\cdot\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\ln(MN)\cdot O_P\left(\frac{T^{\max_j a_{I,j}}}{\sqrt{M}\wedge\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a) and the facts that $\|\mathbf{H}\|_F=O_P(1)$, $\|\mathbf{\Lambda}_E\|_2=O_P(\sqrt{N\ln(MN)})$ and $\|\mathbf{\Lambda}_I\|_2=O_P(\sqrt{M\ln(MN)})$.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_{12}\widehat{\mathbf{G}}\|_2\le\frac{1}{MNT\sqrt{T}}\|\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Lambda}_E\|_2\cdot\Big(\|\mathbf{F}_E'(\widehat{\mathbf{G}}-\mathbf{GH})\|_2+\|\mathbf{F}_E'\mathbf{GH}\|_F\Big)\\
&\le\frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{N}\vee\sqrt{T})\cdot\sqrt{N\ln(MN)}\cdot\sqrt{M\ln(MN)}\\
&\quad\cdot\Big((\sqrt{M}\vee\sqrt{T})\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+T\sqrt{M}\cdot T^{\max_i a_{E,i}}\Big)\\
&=o_P(1)\cdot\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\ln(MN)\cdot O_P\left(\frac{T^{\max_i a_{E,i}}}{\sqrt{N}\wedge\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_E\|_2=O_P(\sqrt{N\ln(MN)})$, $\|\mathbf{\Lambda}_I\|_2=O_P(\sqrt{M\ln(MN)})$ and $\|\mathbf{H}\|_F=O_P(1)$.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_{15}\widehat{\mathbf{G}}\|_2\le\frac{1}{MNT\sqrt{T}}\|\mathbf{U}'\mathbf{\Lambda}_E\|_2\cdot\Big(\|\mathbf{F}_E'(\widehat{\mathbf{G}}-\mathbf{GH})\|_2+\|\mathbf{F}_E'\mathbf{GH}\|_F\Big)\\
&\le\frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{MN}\vee\sqrt{T})\cdot\sqrt{N\ln(MN)}\\
&\quad\cdot\Big((\sqrt{M}\vee\sqrt{T})\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+T\sqrt{M}\cdot T^{\max_i a_{E,i}}\Big)\\
&=o_P(1)\cdot\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\sqrt{\ln(MN)}\cdot O_P\left(\frac{T^{\max_i a_{E,i}}}{\sqrt{MN}\wedge\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a) and the facts that $\|\mathbf{U}\|_2=O_P(\sqrt{MN}\vee\sqrt{T})$, $\|\mathbf{\Lambda}_E\|_2=O_P(\sqrt{N\ln(MN)})$ and $\|\mathbf{H}\|_F=O_P(1)$.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2\le\frac{1}{MNT\sqrt{T}}\|\mathbf{U}'\mathbf{\Lambda}_I\|_2\cdot\Big(\|\mathbf{F}_I'(\widehat{\mathbf{G}}-\mathbf{GH})\|_2+\|\mathbf{F}_I'\mathbf{GH}\|_F\Big)\\
&\le\frac{1}{MNT\sqrt{T}}O_P(1)(\sqrt{MN}\vee\sqrt{T})\cdot\sqrt{M\ln(MN)}\\
&\quad\cdot\Big((\sqrt{N}\vee\sqrt{T})\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+T\sqrt{N}\cdot T^{\max_j a_{I,j}}\Big)\\
&=o_P(1)\cdot\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\sqrt{\ln(MN)}\cdot O_P\left(\frac{T^{\max_j a_{I,j}}}{\sqrt{MN}\wedge\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_I\|_2=O_P(\sqrt{M\ln(MN)})$, $\|\mathbf{U}\|_2=O_P(\sqrt{MN}\vee\sqrt{T})$ and $\|\mathbf{H}\|_F=O_P(1)$.

Then we improve the convergence rates of the terms $\frac{1}{\sqrt{T}}\|\mathbf{A}_4\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_7\widehat{\mathbf{G}}\|_2$, $\frac{1}{\sqrt{T}}\|\mathbf{A}_{10}\widehat{\mathbf{G}}\|_2$ and $\frac{1}{\sqrt{T}}\|\mathbf{A}_{13}\widehat{\mathbf{G}}\|_2$ of Lemma A.1 by bounding $\|\mathbf{A}_k\widehat{\mathbf{G}}\|_2$ directly for $k=4,7,10$ and 13.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\|\mathbf{A}_4\widehat{\mathbf{G}}\|_2\le\frac{1}{MNT\sqrt{T}}\|\mathbf{U}\|_2\cdot\Big(\|\mathbf{U}\|_2\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\|\mathbf{UGH}\|_F\Big)\\
&\le O_P(1)\frac{1}{MNT\sqrt{T}}(\sqrt{MN}\vee\sqrt{T})\Big((\sqrt{MN}\vee\sqrt{T})\cdot\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F+\sqrt{MNT}\Big)
\end{aligned}
$$

$$= o_P(1) \cdot \frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + O_P\left(\frac{1}{T} + \frac{1}{\sqrt{MNT}}\right),$$

where the second inequality follows from $\|\mathbf{U}\|_2 = O_P(\sqrt{MN} \vee \sqrt{T})$ and the fact that

$$\frac{1}{M^2N^2T^2} E\|\mathbf{G}'\mathbf{U}'\|_F^2 = \frac{1}{M^2N^2T^2} \sum_{i=1}^{M} \sum_{j=1}^{N} E \left\| \sum_{t=1}^{T} \boldsymbol{g}_t u_{ijt} \right\|_F^2$$
$$= \frac{1}{M^2N^2T^2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{t,s=1}^{T} E[u_{ijt}u_{ijs}\boldsymbol{g}_t'\boldsymbol{g}_s]$$
$$\leq O(1) \frac{1}{MNT^2} \sum_{t,s=1}^{T} |E[u_{ijt}u_{ijs}]| = O\left(\frac{1}{MNT}\right),$$

in which the last equality follows from the mixing condition of Assumption 3.1.

$$\frac{1}{\sqrt{T}} \|\mathbf{A}_7 \widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}} \|\mathbf{G}\|_F \|\mathbf{\Gamma}'\mathbf{U}\widehat{\mathbf{G}}\|_F$$
$$\leq \frac{1}{MNT\sqrt{T}} \|\mathbf{G}\|_F \cdot \left(\|\mathbf{\Gamma}'\mathbf{U}\|_F \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \|\mathbf{\Gamma}'\mathbf{UG}\|_F \cdot \|\mathbf{H}\|_F \right)$$
$$= O_P(1) \frac{1}{MNT} \left(\sqrt{MNT} \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \sqrt{MNT} \right)$$
$$= o_P(1) \cdot \frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + O_P(1) \left(\frac{1}{\sqrt{MNT}} \right),$$

where the first equality follows from $\|\mathbf{\Gamma}'\mathbf{U}\|_F = O_P(\sqrt{MNT})$, $\|\mathbf{H}\|_F = O_P(1)$ and the fact that

$$\frac{1}{M^2N^2T^2} E\|\mathbf{G}'\mathbf{U}'\mathbf{\Gamma}\|_F^2 = \frac{1}{M^2N^2T^2} E \left\| \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{t=1}^{T} \boldsymbol{g}_t u_{ijt} \boldsymbol{\gamma}_{ij}' \right\|_F^2$$
$$\leq O(1) \frac{1}{M^2N^2T^2} \sum_{i,m=1}^{M} \sum_{j,n=1}^{N} \sum_{t,s=1}^{T} |E[u_{ijt}u_{mns}]| = O\left(\frac{1}{MNT}\right),$$

where the last equality follows from the mixing condition of Assumption 3.1.

$$\frac{1}{\sqrt{T}} \|\mathbf{A}_{10} \widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}} \|\mathbf{F}_E \mathbf{\Lambda}_E'\|_2 \cdot \left(\|\mathbf{U}\|_2 \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \|\mathbf{UGH}\|_F \right)$$
$$\leq \frac{1}{MNT\sqrt{T}} O_P(1)(\sqrt{M} \vee \sqrt{T}) \cdot \sqrt{N \ln(MN)}$$
$$\cdot \left((\sqrt{MN} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \sqrt{MNT} \right)$$
$$= o_P(1) \cdot \frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \sqrt{\ln(MN)} \cdot O_P\left(\frac{1}{T} + \frac{1}{\sqrt{MT}} \right),$$

where the second inequality follows from Assumption 1 and the facts that $\|\mathbf{\Lambda}_E\|_2 = \sqrt{N \ln(MN)}$, $\|\mathbf{UG}\|_F = O_P(\sqrt{MNT})$ and $\|\mathbf{H}\|_F = O_P(1)$.

$$\frac{1}{\sqrt{T}} \|\mathbf{A}_{13} \widehat{\mathbf{G}}\|_2 \leq \frac{1}{MNT\sqrt{T}} \|\mathbf{F}_I \mathbf{\Lambda}_I'\|_2 \cdot \left(\|\mathbf{U}\|_2 \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \|\mathbf{UGH}\|_F \right)$$
$$\leq \frac{1}{MNT\sqrt{T}} O_P(1)(\sqrt{N} \vee \sqrt{T}) \cdot \sqrt{M \ln(MN)}$$

$$
\begin{aligned}
&\cdot \left((\sqrt{MN} \vee \sqrt{T}) \cdot \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \sqrt{MNT} \right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + \sqrt{\ln(MN)} \cdot O_P\left(\frac{1}{T} + \frac{1}{\sqrt{NT}} \right),
\end{aligned}
$$

where the second inequality follows from Assumption 1 and the facts that $\|\mathbf{\Lambda}_I\|_2 = \sqrt{M\ln(MN)}$, $\|\mathbf{UG}\|_F = O_P(\sqrt{MNT})$ and $\|\mathbf{H}\|_F = O_P(1)$.

Finally, given $\frac{1}{\sqrt{T}}\|\mathbf{A}_{14}\widehat{\mathbf{G}}\|_2 = O_P\left(\frac{1}{\sqrt{MN}}\right)$ from Lemma A.1, we have

$$
\begin{aligned}
&\frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F \leq O_P(1) \cdot \frac{1}{\sqrt{T}} \left(\|\mathbf{A}_2\widehat{\mathbf{G}}\|_2 + \cdots + \|\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2 \right) \\
&= o_P(1) \cdot \frac{1}{\sqrt{T}} \|\widehat{\mathbf{G}} - \mathbf{GH}\|_F + O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{T, \sqrt{MT}, \sqrt{NT}\}} + \frac{1}{\sqrt{MN}} + \Delta_{g,MNT} \right),
\end{aligned}
$$

where $\Delta_{g,MNT}$ is defined in the body of this lemma. The first result then follows.

(2). Denote $\delta_{g,MNT} = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{T,\sqrt{MT},\sqrt{NT}\}} + \frac{1}{\sqrt{MN}} + \Delta_{g,MNT} \right)$ for notational simplicity. Note that

$$
\frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}} - \mathbf{GH})\|_F \leq \frac{1}{T}(\|\mathbf{G}'\mathbf{A}_2\widehat{\mathbf{G}}\|_2 + \cdots + \|\mathbf{G}'\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2) \cdot \|\mathbf{V}_g^{-1}\|_F,
$$

where $\mathbf{H} = (\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN})(\frac{\mathbf{G}'\widehat{\mathbf{G}}}{T})\mathbf{V}_g^{-1}$. Below, we reconsider the terms on the right hand side, and provide an improved rate under extra conditions of Assumption 5.1. Note that $\|\mathbf{V}_g^{-1}\|_F = O_P(1)$, so we focus on $\frac{1}{T}\|\mathbf{G}'\mathbf{A}_2\widehat{\mathbf{G}}\|_2, \ldots, \frac{1}{T}\|\mathbf{G}'\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2$.

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_2\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_E\|_F \cdot \|\mathbf{\Lambda}_E\|_2^2 \cdot \left(\|\mathbf{F}_E(\widehat{\mathbf{G}} - \mathbf{GH})\|_2 + \|\mathbf{F}_E\mathbf{GH}\|_F \right) \\
&\leq O_P(1)\frac{1}{T}\frac{1}{MNT} \cdot T\sqrt{M} \cdot T^{\max_i a_{E,i}} \cdot N\ln(MN) \\
&\quad \cdot \left((\sqrt{M} \vee \sqrt{T})\sqrt{T} \cdot \delta_{g,MNT} + T\sqrt{M} \cdot T^{\max_i a_{E,i}} \right) \\
&= \ln(MN) \cdot O_P\left(\frac{\delta_{g,MNT} \cdot T^{\max_i a_{E,i}}}{\sqrt{M} \wedge \sqrt{T}} + T^{2\max_i a_{E,i}} \right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_E\|_2^2 = N\ln(MN)$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{GH}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$.

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_3\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_I\|_F \cdot \|\mathbf{\Lambda}_I\|_2^2 \cdot \left(\|\mathbf{F}_I'(\widehat{\mathbf{G}} - \mathbf{GH})\|_2 + \|\mathbf{F}_I'\mathbf{GH}\|_F \right) \\
&\leq O_P(1)\frac{1}{T}\frac{1}{MNT} \cdot T\sqrt{N} \cdot T^{\max_j a_{I,j}} \cdot M\ln(MN) \\
&\quad \cdot \left((\sqrt{N} \vee \sqrt{T})\sqrt{T} \cdot \delta_{g,MNT} + T\sqrt{N} \cdot T^{\max_j a_{I,j}} \right) \\
&= \ln(MN) \cdot O_P\left(\frac{\delta_{g,MNT} \cdot T^{\max_j a_{I,j}}}{\sqrt{N} \wedge \sqrt{T}} + T^{2\max_j a_{I,j}} \right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_I\|_2^2 = M\ln(MN)$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{GH}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$.

$$
\begin{aligned}
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_9\widehat{\mathbf{G}}\|_2 &\le \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_E\|_F\cdot\|\mathbf{\Lambda}_E'\mathbf{\Lambda}_I\|_2\cdot\left(\|\mathbf{F}_I'(\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H})\|_2+\|\mathbf{F}_I'\mathbf{G}\mathbf{H}\|_F\right)\\
&\le O_P(1)\frac{1}{T}\frac{1}{MNT}\cdot T\sqrt{M}\cdot T^{\max_i a_{E,i}}\cdot\ln(MN)\sqrt{MN}\\
&\quad\cdot\left((\sqrt{N}\vee\sqrt{T})\sqrt{T}\cdot\delta_{g,MNT}+T\sqrt{N}\cdot T^{\max_j a_{I,j}}\right)\\
&=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}\cdot T^{\max_i a_{E,i}}}{\sqrt{N}\wedge\sqrt{T}}+T^{\max_i a_{E,i}+\max_j a_{I,j}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_I\|_2=\sqrt{M\ln(MN)}$, $\|\mathbf{\Lambda}_E\|_2=\sqrt{N\ln(MN)}$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F=\delta_{g,MNT}$ and $\|\mathbf{H}\|_F=O_P(1)$.

$$
\begin{aligned}
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{12}\widehat{\mathbf{G}}\|_2 &\le \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_I\|_F\cdot\|\mathbf{\Lambda}_I'\mathbf{\Lambda}_E\|_2\cdot\left(\|\mathbf{F}_E'(\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H})\|_2+\|\mathbf{F}_E'\mathbf{G}\mathbf{H}\|_F\right)\\
&\le O_P(1)\frac{1}{T}\frac{1}{MNT}\cdot T\sqrt{N}\cdot T^{\max_j a_{I,j}}\cdot\ln(MN)\sqrt{MN}\\
&\quad\cdot\left((\sqrt{M}\vee\sqrt{T})\sqrt{T}\cdot\delta_{g,MNT}+T\sqrt{M}\cdot T^{\max_i a_{E,i}}\right)\\
&=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}\cdot T^{\max_j a_{I,j}}}{\sqrt{M}\wedge\sqrt{T}}+T^{\max_i a_{E,i}+\max_j a_{I,j}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_I\|_2=\sqrt{M\ln(MN)}$, $\|\mathbf{\Lambda}_E\|_2=\sqrt{N\ln(MN)}$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F=\delta_{g,MNT}$ and $\|\mathbf{H}\|_F=O_P(1)$.

$$
\begin{aligned}
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{10}\widehat{\mathbf{G}}\|_2 &\le \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_E\|_F\cdot\|\mathbf{\Lambda}_E\|_2\cdot\left(\|\mathbf{U}\|_2\cdot\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F+\|\mathbf{U}\mathbf{G}\mathbf{H}\|_F\right)\\
&\le O_P(1)\frac{1}{T}\frac{1}{MNT}\cdot T\sqrt{M}\cdot T^{\max_i a_{E,i}}\cdot\sqrt{N\ln(MN)}\\
&\quad\cdot\left((\sqrt{MN}\vee\sqrt{T})\sqrt{T}\cdot\delta_{g,MNT}+\sqrt{MNT}\right)\\
&=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}\cdot T^{\max_i a_{E,i}}}{\sqrt{MN}\wedge\sqrt{T}}+\frac{T^{\max_i a_{E,i}}}{\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_E\|_2=\sqrt{N\ln(MN)}$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F=\delta_{g,MNT}$, $\|\mathbf{U}\|_2=O_P(\sqrt{MN}\vee\sqrt{T})$, $\|\mathbf{U}\mathbf{G}\|_F=O_P(\sqrt{MNT})$ and $\|\mathbf{H}\|_F=O_P(1)$.

$$
\begin{aligned}
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{13}\widehat{\mathbf{G}}\|_2 &\le \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_I\|_F\cdot\|\mathbf{\Lambda}_I\|_2\cdot\left(\|\mathbf{U}\|_2\cdot\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F+\|\mathbf{U}\mathbf{G}\mathbf{H}\|_F\right)\\
&\le O_P(1)\frac{1}{T}\frac{1}{MNT}\cdot T\sqrt{N}\cdot T^{\max_j a_{I,j}}\cdot\sqrt{M\ln(MN)}\\
&\quad\cdot\left((\sqrt{MN}\vee\sqrt{T})\sqrt{T}\cdot\delta_{g,MNT}+\sqrt{MNT}\right)\\
&=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}\cdot T^{\max_j a_{I,j}}}{\sqrt{MN}\wedge\sqrt{T}}+\frac{T^{\max_j a_{I,j}}}{\sqrt{T}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and the facts that $\|\mathbf{\Lambda}_I\|_2=\sqrt{M\ln(MN)}$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}\|_F=\delta_{g,MNT}$, $\|\mathbf{U}\|_2=O_P(\sqrt{MN}\vee\sqrt{T})$, $\|\mathbf{U}\mathbf{G}\|_F=O_P(\sqrt{MNT})$ and $\|\mathbf{H}\|_F=O_P(1)$.

For $k=5$ and 6, we apply the inequality $\frac{1}{T}\|\mathbf{G}'\mathbf{A}_k\widehat{\mathbf{G}}\|_2\le\frac{1}{T}\|\mathbf{G}\|_F\cdot\|\mathbf{A}_k\widehat{\mathbf{G}}\|_2$ and then use the results in Lemma A.4.1. For $k=8$ and 11, we bound $\frac{1}{T}\|\mathbf{G}'\mathbf{A}_k\widehat{\mathbf{G}}\|_2$ directly and use Assumption

4.1.(a) and Assumption 5.1.

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_8\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_E\|_F \cdot \|\mathbf{\Lambda}_E'\mathbf{\Gamma}\|_F \cdot \|\mathbf{G}'\widehat{\mathbf{G}}\|_F \\
&\leq O_P(1)\frac{1}{T}\frac{1}{MNT}T\sqrt{M} \cdot T^{\max_i a_{E,i}} \cdot \sqrt{MN} \cdot T \\
&= O_P\left(\frac{T^{\max_i a_{E,i}}}{\sqrt{N}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and Assumption 5.1.

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{11}\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{F}_I\|_F \cdot \|\mathbf{\Lambda}_I'\mathbf{\Gamma}\|_F \cdot \|\mathbf{G}'\widehat{\mathbf{G}}\|_F \\
&= O_P(1)\frac{1}{T}\frac{1}{MNT}T\sqrt{N} \cdot T^{\max_j a_{I,j}} \cdot \sqrt{MN} \cdot T \\
&= O_P\left(\frac{T^{\max_j a_{I,j}}}{\sqrt{M}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a) and Assumption 5.1.

Next, for $k = 4, 7, 14, 15$ and 16, we bound $\frac{1}{T}\|\mathbf{G}'\mathbf{A}_k\widehat{\mathbf{G}}\|_2$ directly and use the convergence rate of $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|$ achieved in the first result of this lemma.

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_4\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{U}'\|_F \cdot \Big(\|\mathbf{U}\|_2 \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + \|\mathbf{U}\mathbf{G}\mathbf{H}\|_F\Big) \\
&\leq O_P(1)\frac{\sqrt{MNT}}{MNT^2}\left((\sqrt{MN} \vee \sqrt{T}) \cdot \sqrt{T} \cdot \delta_{g,MNT} + \sqrt{MNT}\right) \\
&= O_P\left(\frac{\delta_{g,MNT}}{T \wedge \sqrt{MNT}} + \frac{1}{T}\right),
\end{aligned}
$$

where the second inequality follows from the facts that $\|\mathbf{U}\mathbf{G}\|_F = O_P(\sqrt{MNT})$, $\|\mathbf{U}\|_2 = O_P(\sqrt{MN} \vee \sqrt{T})$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$.

Consider the term involving $\mathbf{A}_7$, write

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{G}'\mathbf{A}_7\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{G}\|_F\|\mathbf{\Gamma}'\mathbf{U}\widehat{\mathbf{G}}\|_F \\
&\leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}\|_F^2 \cdot \Big(\|\mathbf{\Gamma}'\mathbf{U}\|_F \cdot \|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F + \|\mathbf{\Gamma}'\mathbf{U}\mathbf{G}\|_F \cdot \|\mathbf{H}\|_F\Big) \\
&\leq O_P(1)\frac{1}{MNT}\left(\sqrt{MNT}\sqrt{T} \cdot \delta_{g,MNT} + \sqrt{MNT}\right) \\
&= O_P\left(\frac{\delta_{g,MNT}}{\sqrt{MN}} + \frac{1}{\sqrt{MNT}}\right),
\end{aligned}
$$

where the third inequality follows from the facts that $\|\mathbf{\Gamma}'\mathbf{U}\|_F = O_P(\sqrt{MNT})$, $\|\mathbf{G}'\mathbf{U}'\mathbf{\Gamma}\|_F = O_P(\sqrt{MNT})$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$. Similarly,

$$
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{14}\widehat{\mathbf{G}}\|_2 = O_P\left(\frac{1}{\sqrt{MNT}}\right).
$$

For the terms involving $\mathbf{A}_{15}$ and $\mathbf{A}_{16}$, write

$$
\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{15}\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{U}'\|_F \cdot \|\mathbf{\Lambda}_E\|_2 \cdot \Big(\|\mathbf{F}_E'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{F}_E'\mathbf{G}\mathbf{H}\|_F\Big)
$$

$$\leq \frac{1}{T}\frac{1}{MNT}O_P(1)\sqrt{MNT}\cdot\sqrt{N\ln(MN)}$$
$$\cdot\left((\sqrt{M}\vee\sqrt{T})\cdot\sqrt{T}\cdot\delta_{g,MNT}+T\sqrt{M}\cdot T^{\max_i a_{E,i}}\right)$$
$$=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}}{T\wedge\sqrt{MT}}+\frac{T^{\max_i a_{E,i}}}{\sqrt{T}}\right),$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a), and the facts that $\|\mathbf{UG}\|_F = O_P(\sqrt{MNT})$, $\|\mathbf{\Lambda}_E\|_2 = O_P(\sqrt{N\ln(MN)})$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$.

$$\frac{1}{T}\|\mathbf{G}'\mathbf{A}_{16}\widehat{\mathbf{G}}\|_2 \leq \frac{1}{T}\frac{1}{MNT}\|\mathbf{G}'\mathbf{U}'\|_F\cdot\|\mathbf{\Lambda}_I\|_2\cdot\left(\|\mathbf{F}_I'(\widehat{\mathbf{G}}-\mathbf{GH})\|_2+\|\mathbf{F}_I'\mathbf{GH}\|_F\right)$$
$$=\frac{1}{T}\frac{1}{MNT}O_P(1)\sqrt{MNT}\cdot\sqrt{M\ln(MN)}$$
$$\cdot\left((\sqrt{N}\vee\sqrt{T})\cdot\sqrt{T}\cdot\delta_{g,MNT}+T\sqrt{N}\cdot T^{\max_j a_{I,j}}\right)$$
$$=\sqrt{\ln(MN)}\cdot O_P\left(\frac{\delta_{g,MNT}}{T\wedge\sqrt{NT}}+\frac{T^{\max_j a_{I,j}}}{\sqrt{T}}\right),$$

where the second inequality follows from Assumption 1, Assumption 4.1.(a), and the facts that $\|\mathbf{UG}\|_F = O_P(\sqrt{MNT})$, $\|\mathbf{\Lambda}_I\|_2 = O_P(\sqrt{M\ln(MN)})$, $\frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}}-\mathbf{GH}\|_F = \delta_{g,MNT}$ and $\|\mathbf{H}\|_F = O_P(1)$.

Based on the above development, we obtain that

$$\frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}}-\mathbf{GH})\|_F = O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right),$$

where $\Delta^*_{g,MNT}$ is defined in the body of this lemma. The second result then follows.

(3). By the second result of this lemma and the condition $\frac{1}{T}\mathbf{G}'\mathbf{G} = \mathbf{I}_{r_g}$, we immediately obtain that

$$\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}-\mathbf{H} = O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right). \tag{A.10}$$

Left multiplying both sides of (A.10) by $\mathbf{H}$ and using $\|\mathbf{H}\|_F = O_P(1)$, we have

$$\frac{1}{T}\mathbf{H}'\mathbf{G}'\widehat{\mathbf{G}}-\mathbf{H}'\mathbf{H} = O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right). \tag{A.11}$$

Furthermore, (B.3) of Lemma B.3 becomes

$$\frac{1}{T}\widehat{\mathbf{G}}'\widehat{\mathbf{G}}-\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{GH} = \mathbf{I}_{r_g}-\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{GH} = O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right). \tag{A.12}$$

Summing up (A.11) and (A.12),

$$\mathbf{I}_{r_g}-\mathbf{H}'\mathbf{H} = O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right). \tag{A.13}$$

(A.13) shows that $\mathbf{H}$ is an orthogonal matrix w.p.a.1 and hence its eigenvalues are either 1 or -1. By the definition of $\mathbf{H}$, we have

$$\mathbf{H}\cdot\mathbf{V}_g=\left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)\mathbf{H}+O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right), \tag{A.14}$$

which implies that $\mathbf{H}$ is a matrix consisting of eigenvectors of $\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}$ in the limit. By Assumption 5.2, $\mathbf{\Gamma}'\mathbf{\Gamma}$ is diagonal with distinct eigenvalues. It follows that each eigenvalue is associated with a unique unitary eigenvector (up to a sign change) and each eigenvector has a single non-zero element. This implies that $\mathbf{H}$ is a diagonal matrix up to the order $O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right)$. In connection with the fact that the eigenvalues of $\mathbf{H}$ are 1 or -1, $\mathbf{H}$ is a diagonal matrix with 1 or -1 as its elements. Without loss of generality, we assume all elements are 1. This implies that

$$\mathbf{H}=\mathbf{I}_{r_g}+O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^*_{g,MNT}\right).$$

Moreover, in connection with (A.14), we have $\mathbf{V}_g=\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}+o_P(1)$. The proof is now complete.

$\blacksquare$

Proof of Lemma A.5:

(1). Based on Assumptions 1-5 and Lemma A.4, we provide faster rates for $\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})\mathbf{\Gamma}'_{I,j}$, $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ below.

First consider $\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})\mathbf{\Gamma}'_{I,j}$ and write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})\mathbf{\Gamma}'_{I,j}\|_F\leq O(1)\cdot\frac{1}{\sqrt{MT}}\|\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{A}_2+\cdots+\mathbf{A}_{16})\widehat{\mathbf{G}}\|_2\cdot\|\mathbf{K}\|_F\cdot\|\mathbf{\Gamma}_{I,j}\|_F,\\
&\leq O_P(1)\frac{1}{\sqrt{MT}}\cdot\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\min\{T,\sqrt{MT},\sqrt{NT}\}}+\frac{1}{\sqrt{MN}}+\Delta_{g,MNT}\right)\cdot\sqrt{M}\\
&=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{T,\sqrt{MT},\sqrt{NT}\}}+\frac{1}{\sqrt{MN}}+\Delta_{g,MNT}\right),
\end{aligned}$$

where the first inequality follows from $\|\mathbf{\Omega}\|_F\leq\sqrt{\mathrm{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the second inequality follows from Lemma A.4, $\|\widehat{\mathbf{G}}\|_F=O(\sqrt{T})$ and $\|\mathbf{\Gamma}_{I,j}\|_F=O_P(\sqrt{M})$. $\Delta_{g,MNT}$ is defined in the body of Lemma A.4.

Consider $\mathbf{D}_1$, and write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_1\|_F=\frac{1}{\sqrt{MT}}\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\mathbf{P}_{\widehat{\mathbf{G}}}\|_F\leq O(1)\frac{1}{\sqrt{MT}}\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\mathbf{P}_{\widehat{\mathbf{G}}}\|_2\\
&\leq O(1)\frac{1}{\sqrt{MT}T}\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2\cdot\|\mathbf{F}'_E\widehat{\mathbf{G}}\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\leq O(1)\frac{1}{\sqrt{MT}T}\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2\cdot\left(\|\mathbf{F}'_E(\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H})\|_F+\|\mathbf{F}'_E\mathbf{G}\mathbf{H}\|_F\right)\cdot\|\widehat{\mathbf{G}}\|_F\\
&=O_P\left(\frac{\sqrt{\ln(MN)}}{\sqrt{T}\wedge\sqrt{M}}\right)\left(\frac{\sqrt{\ln(MN)}}{\min\{T,\sqrt{MT},\sqrt{NT}\}}+\frac{1}{\sqrt{MN}}+\Delta_{g,MNT}\right)+O_P\left(\sqrt{\ln(MN)}\cdot T^{\max_i a_{E,i}}\right),
\end{aligned}$$

where the first inequality follows from $\|\mathbf{\Omega}\|_F\leq\sqrt{\mathrm{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the second inequality follows from $\mathbf{P}_{\widehat{\mathbf{G}}} = \frac{1}{T}\widehat{\mathbf{G}}\widehat{\mathbf{G}}'$, and the last equality follows from Lemma A.4 and Assumption 4.1.(a).

Next, consider $\mathbf{D}_2$, and write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_2\|_F = \frac{1}{\sqrt{MT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{P}_{\widehat{\mathbf{G}}}\|_F \leq \frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\widehat{\mathbf{G}}\|_F \cdot \|\widehat{\mathbf{G}}\|_F \\
&\leq \frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \cdot \|\widehat{\mathbf{G}}\|_F + \frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{G}\mathbf{H}\|_F \cdot \|\widehat{\mathbf{G}}\|_F.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \cdot \|\widehat{\mathbf{G}}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{T, \sqrt{MT}, \sqrt{NT}\}} + \frac{1}{\sqrt{MN}} + \Delta_{g,MNT}\right), \\
&\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{G}\mathbf{H}\|_F \cdot \|\widehat{\mathbf{G}}\|_F = O_P(T^{a_{I,j}}),
\end{aligned}$$

where we have used Assumption 2, Assumption 4.1.(a), Lemma A.4, and $\|\mathbf{F}_{I,j}\|_F = O_P(\sqrt{T})$.

Then we consider $\mathbf{D}_3$ and write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_3\|_F = \frac{1}{\sqrt{MT}}\|\mathbf{U}_{\bullet j \bullet}\mathbf{P}_{\widehat{\mathbf{G}}}\|_F \leq O(1)\frac{1}{\sqrt{MT}}\|\mathbf{U}_{\bullet j \bullet}\mathbf{P}_{\widehat{\mathbf{G}}}\|_2 \\
&\leq O(1)\frac{1}{\sqrt{MTT}}\|\mathbf{U}_{\bullet j \bullet}\widehat{\mathbf{G}}\| \cdot \|\widehat{\mathbf{G}}\|_F \leq O(1)\frac{1}{\sqrt{MTT}}\left(\|\mathbf{U}_{\bullet j \bullet}(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_2 + \|\mathbf{U}_{\bullet j \bullet}\mathbf{G}\mathbf{H}\|_F\right) \cdot \|\widehat{\mathbf{G}}\|_F \\
&= O_P(1)\left(\frac{1}{\sqrt{M}} \vee \frac{1}{\sqrt{T}}\right)\left(\frac{\sqrt{\ln(MN)}}{\min\{T, \sqrt{MT}, \sqrt{NT}\}} + \frac{1}{\sqrt{MN}} + \Delta_{g,MNT}\right) + O_P\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}$$

where the first inequality follows from $\|\mathbf{\Omega}\|_F \leq \sqrt{\text{rank}(\mathbf{\Omega})} \cdot \|\mathbf{\Omega}\|_2$, the second inequality follows from $\mathbf{P}_{\widehat{\mathbf{G}}} = \frac{1}{T}\widehat{\mathbf{G}}\widehat{\mathbf{G}}'$, and the last equality follows from $\|\mathbf{U}_{\bullet j \bullet}\mathbf{G}\|_F = O_P(\sqrt{MT})$, which is similar to the result of $\|\mathbf{G}'\mathbf{U}'\|_F$ in the proof of Lemma A.4(i).

Based on the above development, we obtain that

$$\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\right\|_F = O_P\left(\frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT}\right),$$

where $\Delta_{Ij,MNT}$ is defined in the body of this lemma. This completes the proof of the first result.

(2). The second result can be proved in exactly the same way as the first result. Thus, omitted.

(3). Consider (1) in Lemma B.5. Write

$$\begin{aligned}
&\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\,\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'\|_2 \\
&\leq \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2 \cdot \|\,\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}\|_2 \cdot \|\mathbf{F}_E\|_2 \\
&= \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2 \cdot \sqrt{\max_{i\geq 1, j\geq 1}\|\boldsymbol{\lambda}_{E,ij}\|_F^2} \cdot \|\mathbf{F}_E\|_2 \\
&= O_P(1)\frac{1}{\sqrt{MT}} \cdot \left(\frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT}\right) \cdot \sqrt{\ln(MN)} \cdot (\sqrt{T} \vee \sqrt{M})
\end{aligned}$$

$$= O_P(1)\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\cdot\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right),$$

where the second equality follows from the first result of this lemma, Assumption 1 and Assumption 2.

Consider (2) in Lemma B.5. Write

$$\begin{aligned}
&\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\|_2\\
&\le \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2\cdot\|\mathbf{\Lambda}_{I,\bullet j}\|_2\cdot\|\mathbf{F}_{I,j}\|_F\\
&= O_P(1)\frac{1}{\sqrt{MT}}\cdot\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right)\cdot\sqrt{M}\cdot\sqrt{T}\\
&= O_P\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right),
\end{aligned}$$

where the first equality follows from the first result of this lemma.

Consider (3) in Lemma B.5. Write

$$\begin{aligned}
&\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{U}_{\bullet j\bullet}\|_2\\
&\le \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2\cdot\|\mathbf{U}_{\bullet j\bullet}\|_2\\
&= O_P(1)\frac{1}{\sqrt{MT}}\cdot\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right)\cdot(\sqrt{M}\vee\sqrt{T})\\
&= O_P(1)\frac{1}{\sqrt{M}\wedge\sqrt{T}}\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right),
\end{aligned}$$

where the first equality follows from the first result of this lemma.

Consider (7) in Lemma B.5. Write

$$\begin{aligned}
&\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')\|_2\le\frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2^2\\
&= O_P\left(\frac{1}{MN}+\frac{1}{T}+\Delta^2_{Ij,MNT}\right),
\end{aligned}$$

where the first equality follows from the first result of this lemma.

The results for other terms are the same as those in Lemma B.5. The rest of the proofs are in the same spirit as Lemma A.2. Thus, omitted.

(4). The result can be proved in exactly the same way as in the third result. The proof is now complete. ■

Proof of Lemma A.6:

(1). We left and right multiply (A.9) by $\frac{1}{T}\mathbf{F}'_{I,j}$ and $\mathbf{V}^{-1}_{I,j}$ respectively to obtain that

$$\frac{1}{T}\mathbf{F}'_{I,j}(\widehat{\mathbf{F}}_{I,j}-\mathbf{F}_{I,j}\mathbf{H}_{I,j})=\frac{1}{T}\mathbf{F}'_{I,j}(\mathbf{B}_2+\cdots+\mathbf{B}_{16})\widehat{\mathbf{F}}_{I,j}\mathbf{V}^{-1}_{I,j}, \tag{A.15}$$

where $\mathbf{H}_{I,j} = \frac{1}{M}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j} \cdot \frac{1}{T}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j} \cdot \mathbf{V}_{I,j}^{-1}$. Therefore,

$$\frac{1}{T}\|\mathbf{F}'_{I,j}(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_F \leq O(1) \cdot \frac{1}{T}\left(\|\mathbf{F}'_{I,j}\mathbf{B}_2\widehat{\mathbf{F}}_{I,j}\|_2 + \cdots + \|\mathbf{F}'_{I,j}\mathbf{B}_{16}\widehat{\mathbf{F}}_{I,j}\|_2\right) \cdot \|\mathbf{V}_{I,j}^{-1}\|_F,$$

where the first inequality follows from (A.15) and $\|\mathbf{\Omega}\|_F \leq \sqrt{\text{rank}(\mathbf{\Omega})} \cdot \|\mathbf{\Omega}\|_2$.

Note that $\|\mathbf{V}_{I,j}^{-1}\|_F = O_P(1)$, so we focus on $\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_2\widehat{\mathbf{F}}_{I,j}\|_2, \ldots, \frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_{16}\widehat{\mathbf{F}}_{I,j}\|_2$. For $k = 6-8, 10, 12, 14-16$, we carefully investigate $\|\mathbf{F}'_{I,j}\mathbf{B}_k\widehat{\mathbf{F}}_{I,j}\|_2$ in order to achieve sharper bounds. For other terms, we apply the inequality $\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_k\widehat{\mathbf{F}}_{I,j}\|_2 \leq \frac{1}{T}\|\mathbf{F}_{I,j}\|_F \cdot \|\mathbf{B}_k\|_2 \cdot \|\widehat{\mathbf{F}}_{I,j}\|_F$, and use the updated results for $\mathbf{B}_k$ in the proof of Lemma A.5.

That said, we now start the investigation. For the term involving $\mathbf{B}_6$, write

$$\begin{aligned}
&\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_6\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{F}_E \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\} \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\widehat{\mathbf{F}}_{I,j}\|_2 \\
&\leq \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{F}_E\|_F \cdot \|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2^2 \cdot \left(\|\mathbf{F}'_E(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_2 + \|\mathbf{F}'_E\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right) \\
&\leq O_P(1)\frac{1}{MT}\sqrt{M} \cdot T^{\max_i b_{EI,ij}} \cdot \ln(MN) \\
&\quad \cdot \left((\sqrt{M} \vee \sqrt{T})\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M} \wedge \sqrt{T}} + \Delta_{Ij,MNT}\right) + T\sqrt{M} \cdot T^{\max_i b_{EI,ij}}\right),
\end{aligned}$$

where the second inequality follows from Assumption 6.1, Lemma A.5.3, $\|\mathbf{F}_E\|_2 = O_P(\sqrt{M} \vee \sqrt{T})$, $\|\mathbf{H}_{I,j}\|_F = O_P(1)$ and $\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 = O_P(\sqrt{\ln(MN)})$.

For the term involving $\mathbf{B}_7$, write

$$\begin{aligned}
&\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_7\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{F}_E \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j}\|_2 \\
&\leq \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{F}_E\|_F \cdot \|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 \cdot \|\mathbf{\Lambda}_{I,\bullet j}\|_F \cdot \|\mathbf{F}_{I,j}\|_F \cdot \|\widehat{\mathbf{F}}_{I,j}\|_F \\
&\leq O_P(1)\frac{1}{MT}\sqrt{M} \cdot T^{\max_i b_{EI,ij}} \cdot \sqrt{\ln(MN)} \cdot \sqrt{M} \cdot T \\
&= O_P(1)\sqrt{\ln(MN)} \cdot T^{\max_i b_{EI,ij}},
\end{aligned}$$

where the second inequality follows from Assumption 6.1, $\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 = O_P(\sqrt{\ln(MN)})$, $\|\mathbf{\Lambda}_{I,\bullet j}\|_F = O_P(\sqrt{M})$, $\|\mathbf{F}_{I,j}\|_F = O_P(\sqrt{T})$ and $\|\widehat{\mathbf{F}}_{I,j}\|_F = O_P(\sqrt{T})$.

Consider the term involving $\mathbf{B}_8$, first write

$$\begin{aligned}
&\frac{1}{M^2T^2}E\|\mathbf{U}_{\bullet j \bullet}\mathbf{F}_{I,j}\|_F^2 = \frac{1}{M^2T^2}\sum_{i=1}^{M} E\left\|\sum_{t=1}^{T} \boldsymbol{f}_{I,jt}u_{ijt}\right\|_F^2 \\
&= \frac{1}{M^2T^2}\sum_{i=1}^{M}\sum_{t,s=1}^{T} E[u_{ijt}u_{ijs}\boldsymbol{f}'_{I,jt}\boldsymbol{f}_{I,js}] \leq O(1)\frac{1}{MT^2}\sum_{t,s=1}^{T}|E[u_{ijt}u_{ijs}]| = O\left(\frac{1}{MT}\right).
\end{aligned}$$

Thus, $\frac{1}{MT}\|\mathbf{U}_{\bullet j \bullet}\mathbf{F}_{I,j}\|_F = O_P\left(\frac{1}{\sqrt{MT}}\right)$. It follows that

$$\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_8\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{F}_E \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{U}_{\bullet j \bullet}\widehat{\mathbf{F}}_{I,j}\|_2$$

$$\le \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}'\mathbf{F}_E\|_F \cdot \|\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}\|_2 \cdot \left(\|\mathbf{U}_{\bullet j\bullet}(\widehat{\mathbf{F}}_{I,j}-\mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_2 + \|\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right)$$
$$\le O_P(1)\frac{1}{MT}\sqrt{M}\cdot T^{\max_i b_{EI,ij}} \cdot \sqrt{\ln(MN)}$$
$$\cdot\left((\sqrt{M}\vee\sqrt{T})\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}+\Delta_{Ij,MNT}\right)+\sqrt{MT}\right),$$

where the second inequality follows from Assumption 6.1, Lemma A.5.3, $\|\mathbf{U}_{\bullet j\bullet}\|_2 = O_P(\sqrt{M}\vee\sqrt{T})$, $\|\mathbf{H}_{I,j}\|_F = O_P(1)$ and $\|\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}\|_2 = O_P(\sqrt{\ln(MN)})$.

For the term involving $\mathbf{B}_{10}$, write

$$\frac{1}{T}\|\mathbf{F}_{I,j}'\mathbf{B}_{10}\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}'\mathbf{F}_{I,j}\mathbf{\Lambda}_{I,\bullet j}'\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'\widehat{\mathbf{F}}_{I,j}\|_2$$
$$\le \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}\|_F^2\cdot\|\mathbf{\Lambda}_{I,\bullet j}\|_F\cdot\|\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}\|_2\cdot\left(\|\mathbf{F}_E'(\widehat{\mathbf{F}}_{I,j}-\mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_2+\|\mathbf{F}_E'\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right)$$
$$\le O_P(1)\frac{1}{MT}\cdot\sqrt{M}\cdot\sqrt{\ln(MN)}$$
$$\cdot\left((\sqrt{M}\vee\sqrt{T})\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}+\Delta_{Ij,MNT}\right)+T\sqrt{M}\cdot T^{\max_i b_{EI,ij}}\right)$$
$$= O_P(1)\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\cdot\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}+\Delta_{Ij,MNT}\right)+O_P(1)\sqrt{\ln(MN)}\cdot T^{\max_i b_{EI,ij}}$$

where the second inequality follows from Lemma A.5.3 and Assumption 6.1.

For the terms involving $\mathbf{B}_{12}$ and $\mathbf{B}_{15}$, write

$$\frac{1}{M^2T^2}E\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F^2 = \frac{1}{M^2T^2}E\left\|\sum_{i=1}^M\sum_{t=1}^T \boldsymbol{f}_{I,jt}u_{ijt}\boldsymbol{\lambda}_{I,ij}'\right\|_F^2$$
$$= \frac{1}{M^2T^2}\sum_{i,m=1}^M\sum_{t,s=1}^T E\|\boldsymbol{\lambda}_{I,ij}\boldsymbol{f}_{I,jt}'\boldsymbol{f}_{I,js}\boldsymbol{\lambda}_{I,mj}'u_{ijt}u_{mjs}\|_F = O\left(\frac{1}{MT}\right),$$

where the last equality follows from the mixing condition of Assumption 3.1. Thus, $\frac{1}{MT}\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F = O_P\left(\frac{1}{\sqrt{MT}}\right)$. It follows that

$$\frac{1}{T}\|\mathbf{F}_{I,j}'\mathbf{B}_{12}\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}'\mathbf{F}_{I,j}\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\widehat{\mathbf{F}}_{I,j}\|_2$$
$$\le \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}\|_F^2\cdot\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\widehat{\mathbf{F}}_{I,j}\|_2$$
$$\le \frac{1}{MT}\frac{1}{T}\|\mathbf{F}_{I,j}\|_F^2\cdot\left(\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\|_F\cdot\|\widehat{\mathbf{F}}_{I,j}-\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_2+\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right)$$
$$\le O_P(1)\frac{1}{MT}\left(\sqrt{MT}\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}+\Delta_{Ij,MNT}\right)+\sqrt{MT}\right)$$
$$= O_P(1)\left(\frac{\sqrt{\ln(MN)}}{M}+\frac{\sqrt{\ln(MN)}}{\sqrt{MT}}\right)+O_P\left(\frac{\Delta_{Ij,MNT}}{\sqrt{M}}\right),$$

where the last inequality follows from $\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\|_F = O_P(\sqrt{MT})$ and $\|\mathbf{\Lambda}_{I,\bullet j}'\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F = O_P\left(\sqrt{MT}\right)$. Similarly,

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_{15}\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\widehat{\mathbf{F}}_{I,j}\|_2 \\
&\leq \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\mathbf{\Lambda}_{I,\bullet j}\|_F \cdot \|\mathbf{F}_{I,j}\|_F \cdot \|\widehat{\mathbf{F}}_{I,j}\|_F = O_P\left(\frac{1}{\sqrt{MT}}\right),
\end{aligned}
$$

where the first inequality follows from $\|\mathbf{\Lambda}'_{I,\bullet j}\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F = O_P\left(\sqrt{MT}\right)$.

For the term involving $\mathbf{B}_{14}$, write

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_{14}\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\widehat{\mathbf{F}}_{I,j}\|_2 \\
&\leq \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\|_F \cdot \|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 \cdot \left(\|\mathbf{F}'_E(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_2 + \|\mathbf{F}'_E\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right) \\
&\leq O_P(1)\frac{1}{MT}\frac{1}{T}\sqrt{MT}\cdot\sqrt{\ln(MN)} \\
&\cdot\left((\sqrt{M}\vee\sqrt{T})\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}} + \Delta_{Ij,MNT}\right) + T\sqrt{M}\cdot T^{\max_i b_{EI,ij}}\right),
\end{aligned}
$$

where the second inequality follows from Assumption 6.1, Lemma A.5.3, $\|\mathbf{U}_{\bullet j\bullet}\|_2 = O_P(\sqrt{M}\vee\sqrt{T})$, $\|\mathbf{H}_{I,j}\|_F = O_P(1)$, $\|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 = O_P(\sqrt{\ln(MN)})$ and $\|\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F = O_P(\sqrt{MT})$.

For the term involving $\mathbf{B}_{16}$, write

$$
\begin{aligned}
&\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{B}_{16}\widehat{\mathbf{F}}_{I,j}\|_2 = \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\mathbf{U}_{\bullet j\bullet}\widehat{\mathbf{F}}_{I,j}\|_2 \\
&\leq \frac{1}{MT}\frac{1}{T}\|\mathbf{F}'_{I,j}\mathbf{U}'_{\bullet j\bullet}\|_F \cdot \left(\|\mathbf{U}_{\bullet j\bullet}(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_2 + \|\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\mathbf{H}_{I,j}\|_F\right) \\
&\leq O_P(1)\frac{1}{MT}\frac{1}{T}\sqrt{MT}\cdot\left((\sqrt{M}\vee\sqrt{T})\sqrt{T}\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}} + \Delta_{Ij,MNT}\right) + \sqrt{MT}\right) \\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{M\sqrt{T}} + \frac{1}{T} + \frac{\Delta_{Ij,MNT}}{T\wedge\sqrt{MT}}\right),
\end{aligned}
$$

where the second inequality follows from Lemma A.5.3, $\|\mathbf{U}_{\bullet j\bullet}\|_2 = O_P(\sqrt{M}\vee\sqrt{T})$, $\mathbf{H}_{I,j} = O_P(1)$ and $\|\mathbf{U}_{\bullet j\bullet}\mathbf{F}_{I,j}\|_F = O_P(\sqrt{MT})$.

Based on the above development, we obtain that

$$
\frac{1}{T}\|\mathbf{F}'_{I,j}(\widehat{\mathbf{F}}_{I,j} - \mathbf{F}_{I,j}\mathbf{H}_{I,j})\|_F = O_P\left(\frac{\ln(MN)}{M} + \frac{1}{\sqrt{MN}} + \frac{1}{\sqrt{T}} + \Delta_{Ij,MNT} + \sqrt{\ln(MN)}\cdot T^{\max_i b_{EI,ij}}\right).
$$

(2). The second result can be proved in exactly the same way as the first result. Thus, omitted.

(3)-(4). Given the first result of this lemma, the proof can be done in exactly the same way as in Lemma A.4, thus omitted. ■

Proof of Theorem 2.3:

(1). Let $\upsilon_{ijt} = \boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it} + \boldsymbol{\lambda}'_{I,ij}\boldsymbol{f}_{I,jt} + u_{ijt}$. In matrix notation, we have $\mathbf{\Upsilon} = \mathbf{\Lambda}_E\mathbf{F}'_E + \mathbf{\Lambda}_I\mathbf{F}'_I + \mathbf{U}$. Then $\mathbf{Y} = \mathbf{\Gamma}\mathbf{G}' + \mathbf{\Upsilon}$. If $\frac{\sqrt{MN}}{T} \to 0$, we have

$$\sqrt{MN}(\widehat{\boldsymbol{g}}_t - \mathbf{H}'\boldsymbol{g}_t) = \mathbf{V}_g^{-1} \cdot \frac{\widehat{\mathbf{G}}'\mathbf{G}}{T} \cdot \frac{1}{\sqrt{MN}} \sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{ij} \upsilon_{ijt} + o_P(1)$$

$$= \mathbf{H}' \left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)^{-1} \frac{1}{\sqrt{MN}} \sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{ij} \upsilon_{ijt} + o_P(1)$$

$$= \left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)^{-1} \frac{1}{\sqrt{MN}} \sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{ij} \upsilon_{ijt} + o_P(1)$$

where the second equality follows from the fact that

$$\mathbf{V}_g^{-1}\left(\frac{\widehat{\mathbf{G}}'\mathbf{G}}{T}\right) = \mathbf{V}_g^{-1}\left(\frac{\widehat{\mathbf{G}}'\mathbf{G}}{T}\right)\left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)\left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)^{-1} = \mathbf{H}'\left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)^{-1},$$

and the third equality follows from Lemma A.4.

Note that

$$\sqrt{MN}(\widehat{\boldsymbol{g}}_t - \boldsymbol{g}_t) = \sqrt{MN}(\widehat{\boldsymbol{g}}_t - \mathbf{H}'\boldsymbol{g}_t) + \sqrt{MN}(\mathbf{H}' - \mathbf{I}_{r_g})\boldsymbol{g}_t$$

$$= \sqrt{MN}(\widehat{\boldsymbol{g}}_t - \mathbf{H}'\boldsymbol{g}_t) + \sqrt{MN} \cdot O_P\left(\frac{1}{MN} + \frac{1}{T} + \Delta^*_{g,MNT}\right)$$

$$= \left(\frac{\mathbf{\Gamma}'\mathbf{\Gamma}}{MN}\right)^{-1} \frac{1}{\sqrt{MN}} \sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{ij} \upsilon_{ijt} + o_P(1),$$

if $\frac{\sqrt{MN}}{T} \to 0$ and $\sqrt{MN}\cdot\Delta^*_{g,MNT} \to 0$. Then by Assumption 2.1 and Assumption 5.3, the result follows immediately.

(2). Let $\upsilon_{I,ijt} = \boldsymbol{\gamma}'_{ij}\boldsymbol{g}_t - \widehat{\boldsymbol{\gamma}}'_{ij}\widehat{\boldsymbol{g}}_t + \boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it} + u_{ijt}$ for $i = 1, \ldots, M$. In matrix notation, we have $\mathbf{\Upsilon}_{I,j} = \mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}' + \mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E + \mathbf{U}_{\bullet j \bullet}$. Then $\mathbf{Y}_{I,j} = \mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j} + \mathbf{\Upsilon}_{I,j}$. If $\frac{\sqrt{M}}{T} \to 0$, we have

$$\sqrt{M}(\widehat{\boldsymbol{f}}_{I,jt} - \mathbf{H}'_{I,j}\boldsymbol{f}_{I,jt}) = \mathbf{V}_{I,j}^{-1}\left(\frac{\widehat{\mathbf{F}}'_{I,j}\mathbf{F}_{I,j}}{T}\right)\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}\upsilon_{I,ijt} + o_P(1)$$

$$= \mathbf{H}'_{I,j}\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}\upsilon_{I,ijt} + o_P(1)$$

$$= \left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}\upsilon_{I,ijt} + o_P(1)$$

$$= \left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}(\boldsymbol{\gamma}'_{ij}\boldsymbol{g}_t - \widehat{\boldsymbol{\gamma}}'_{ij}\widehat{\boldsymbol{g}}_t + \boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it} + u_{ijt}) + o_P(1)$$

$$= \left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\left(\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}(\boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it} + u_{ijt}) + \sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}(\boldsymbol{\gamma}'_{ij}\boldsymbol{g}_t - \widehat{\boldsymbol{\gamma}}'_{ij}\widehat{\boldsymbol{g}}_t)\right) + o_P(1)$$

$$= \left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \boldsymbol{\lambda}_{I,ij}(\boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it} + u_{ijt})$$

$$+\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\left(M\cdot O_P\left(\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}\right)\right)+o_P(1),$$

where the second equality follows from the fact that

$$\mathbf{V}_{I,j}^{-1}\left(\frac{\widehat{\mathbf{F}}'_{I,j}\mathbf{F}_{I,j}}{T}\right)=\mathbf{V}_{I,j}^{-1}\left(\frac{\widehat{\mathbf{F}}'_{I,j}\mathbf{F}_{I,j}}{T}\right)\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}=\mathbf{H}'_{I,j}\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1},$$

the third equality follows from $\mathbf{H}_{I,j}=\mathbf{I}_{r_{I,j}}+o_P(1)$, and the last equality follows from Lemma A.5.1. Thus,

$$\begin{aligned}
&\sqrt{M}(\widehat{\boldsymbol{f}}_{I,jt}-\boldsymbol{f}_{I,jt})=\sqrt{M}(\widehat{\boldsymbol{f}}_{I,jt}-\mathbf{H}'_{I,j}\boldsymbol{f}_{I,jt})+\sqrt{M}(\mathbf{H}'_{I,j}-\mathbf{I}_{r_{I,j}})\boldsymbol{f}_{I,jt}\\
&=\sqrt{M}(\widehat{\boldsymbol{f}}_{I,jt}-\mathbf{H}'_{I,j}\boldsymbol{f}_{I,jt})\\
&\quad+\sqrt{M}\cdot O_P\left(\frac{\ln(MN)}{M}+\frac{1}{\sqrt{MN}}+\frac{1}{\sqrt{T}}+\Delta_{Ij,MNT}+\sqrt{\ln(MN)}\cdot T^{\max_i c_{EI,ij}}\right)\\
&=\left(\frac{\mathbf{\Lambda}'_{I,\bullet j}\mathbf{\Lambda}_{I,\bullet j}}{M}\right)^{-1}\frac{1}{\sqrt{M}}\sum_{i=1}^{M}\boldsymbol{\lambda}_{I,ij}(\boldsymbol{\lambda}'_{E,ij}\boldsymbol{f}_{E,it}+u_{ijt})+o_P(1),
\end{aligned}$$

if $\frac{\ln(MN)}{\sqrt{M}}\to 0$, $\frac{\sqrt{M}}{\sqrt{T}}\to 0$ and $\sqrt{M}\left(\Delta_{Ij,MNT}+\sqrt{\ln(MN)}\cdot T^{\max_i c_{EI,ij}}\right)\to 0$. Then by Assumption 4.1(c) and Assumption 6.4(a), the result follows immediately.

(3). The second result can be proved in exactly the same way as the first result. Thus, omitted. ■

Appendix B

In this Appendix, we provide the secondary lemmas and the corresponding proofs.

B.1 Preliminary Lemmas

Lemma B.1. *Suppose that* $\mathbf{A}$ *and* $\mathbf{A}+\mathbf{E}$ *are* $n\times n$ *symmetric matrices and that* $\mathbf{Q}=(\mathbf{Q}_1,\mathbf{Q}_2)$*, where* $\mathbf{Q}_1$ *is* $n\times r$ *and* $\mathbf{Q}_2$ *is* $n\times(n-r)$*, is an orthogonal matrix such that* $\mathrm{span}(\mathbf{Q}_1)$ *is an invariant subspace for* $\mathbf{A}$*; that is,* $\mathbf{A}\times\mathrm{span}(\mathbf{Q}_1)\subset\mathrm{span}(\mathbf{Q}_1)$*. Decompose* $\mathbf{Q}'\mathbf{A}\mathbf{Q}$ *and* $\mathbf{Q}'\mathbf{E}\mathbf{Q}$ *as* $\mathbf{Q}'\mathbf{A}\mathbf{Q}=\mathrm{diag}(\mathbf{D}_1,\mathbf{D}_2)$ *and*

$$\mathbf{Q}'\mathbf{E}\mathbf{Q}=\begin{pmatrix}\mathbf{E}_{11}&\mathbf{E}'_{21}\\ \mathbf{E}_{21}&\mathbf{E}_{22}\end{pmatrix}.$$

Let

$$\mathrm{sep}(\mathbf{D}_1,\mathbf{D}_2)=\min_{\lambda_1\in\lambda(\mathbf{D}_1),\ \lambda_2\in\lambda(\mathbf{D}_2)}|\lambda_1-\lambda_2|,$$

where $\lambda(\mathbf{B})$ *denotes the set of eigenvalues of the matrix* $\mathbf{B}$*. If* $\text{sep}(\mathbf{D}_1, \mathbf{D}_2) > 0$ *and* $\|\mathbf{E}\|_2 \leq \text{sep}(\mathbf{D}_1, \mathbf{D}_2)/5$*, then there exists a* $(n-r)\times r$ *matrix* $\mathbf{P}$ *with* $\|\mathbf{P}\|_2 \leq 4\|\mathbf{E}_{21}\|_2/\text{sep}(\mathbf{D}_1, \mathbf{D}_2)$*, such that the columns of* $\mathbf{Q}_1^0 = (\mathbf{Q}_1 + \mathbf{Q}_2\mathbf{P})(\mathbf{I}_r + \mathbf{P}'\mathbf{P})^{-1/2}$ *define an orthonormal basis for a subspace that is invariant for* $\mathbf{A}+\mathbf{E}$*.*

Lemma B.2. *Under Assumptions 1-3, as* $(M, N, T) \to (\infty, \infty, \infty)$*,*

1. $\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_E\mathbf{F}_E'\|_2 = O_P\left(\ln(MN)\cdot(M^{-1}\vee T^{-1})\right)$,
2. $\frac{1}{MNT}\|\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 = O_P\left(\ln(MN)\cdot(N^{-1}\vee T^{-1})\right)$,
3. $\frac{1}{MNT}\|\mathbf{U}'\mathbf{U}\|_2 = O_P\left(T^{-1/2}\vee(MN)^{-1/2}\right)$,
4. $\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_E\mathbf{F}_E'\|_2 = O_P\left(\sqrt{\ln(MN)}\cdot(M^{-1/2}\vee T^{-1/2})\right)$,
5. $\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 = O_P\left(\sqrt{\ln(MN)}\cdot(N^{-1/2}\vee T^{-1/2})\right)$,
6. $\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{U}\|_2 = O_P\left(\frac{1}{\sqrt{MN}}\right)$,
7. $\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 = O_P\left(\frac{(\sqrt{T}\vee\sqrt{M})\cdot(\sqrt{T}\vee\sqrt{N})\cdot\ln(MN)}{\sqrt{MNT}}\right)$,
8. $\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{U}\|_2 = O_P\left(\frac{(\sqrt{T}\vee\sqrt{M})\cdot(\sqrt{T}\vee\sqrt{MN})\cdot\sqrt{\ln(MN)}}{M\sqrt{NT}}\right)$,
9. $\frac{1}{MNT}\|\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{U}\|_2 = O_P\left(\frac{(\sqrt{T}\vee\sqrt{N})\cdot(\sqrt{T}\vee\sqrt{MN})\cdot\sqrt{\ln(MN)}}{\sqrt{M}NT}\right)$.

Having established Theorem 2.1, we suppose that the number of global factors has been successfully identified. In order to keep the notation simple, for the following lemmas we suppress the dagger superscript in $\widehat{\mathbf{G}}^\dagger$ and $\mathbf{V}_g^\dagger$.

Lemma B.3. *Under Assumptions 1-3, as* $(M, N, T) \to (\infty, \infty, \infty)$*,*

1. $\frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$,
2. $\frac{1}{T}\|\widehat{\mathbf{G}}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$,
3. $\|\mathbf{P}_{\widehat{\mathbf{G}}} - \mathbf{P}_{\mathbf{G}}\|_F^2 = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)$, *where* $\mathbf{H}$ *has been defined in Lemma A.1.*

Lemma B.4. *Under Assumptions 1-4, as* $(M, N, T) \to (\infty, \infty, \infty)$*,*

1. *for* $j = 1, \ldots, N$*,* $\frac{1}{\sqrt{MT}}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right)$*, where* $\mathbf{\Gamma}_{I,j}$ *is denoted in the same way as* $\widehat{\mathbf{\Gamma}}_{I,j}$*;*
2. *for* $i = 1, \ldots, M$*,* $\frac{1}{\sqrt{NT}}\|\mathbf{\Gamma}_{E,i}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{E,i}\widehat{\mathbf{G}}'\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{E,i}}\right)$*, where* $\mathbf{\Gamma}_{E,i}$ *be denoted in the same way as* $\widehat{\mathbf{\Gamma}}_{E,i}$*.*

Lemma B.5. *Under Assumptions 1-4, as* $(M,N,T)\to(\infty,\infty,\infty)$,

1. $\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\|_2=O_P(1)\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\cdot\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}+T^{a_{I,j}}\right)$,
2. $\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\|_2=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}+T^{a_{I,j}}\right)$,
3. $\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{U}_{\bullet j\bullet}\|_2=O_P(1)\left(\frac{1}{\sqrt{M}\wedge\sqrt{T}}\right)\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}+T^{a_{I,j}}\right)$,
4. $\frac{1}{MT}\|\mathbf{F}_E\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\|_2=O_P\left(\frac{\ln(MN)}{M\wedge T}\right)$,
5. $\frac{1}{MT}\|\mathbf{F}_E\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\|_2=O_P\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\right)$,
6. $\frac{1}{MT}\|\mathbf{F}_E\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{U}_{\bullet j\bullet}\|_2=O_P\left(\frac{\sqrt{\ln(MN)}}{M\wedge T}\right)$,
7. $\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')\|_2=O_P\left(\frac{\ln(MN)}{\min\{M,N,T\}}+T^{2a_{I,j}}\right)$,
8. $\frac{1}{MT}\|\mathbf{F}_{I,j}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{U}_{\bullet j\bullet}\|_2=O_P\left(\frac{1}{\sqrt{M}}\right)$,
9. $\frac{1}{MT}\|\mathbf{U}'_{\bullet j\bullet}\mathbf{U}_{\bullet j\bullet}\|_2=O_P\left(\frac{1}{\sqrt{M}}\vee\frac{1}{\sqrt{T}}\right)$.

B.2 Proofs of Lemma B.2 to Lemma B.5

Proof of Lemma B.1:

This is Lemma 3 of Lam et al. (2011). The proof is therefore omitted. ■

Proof of Lemma B.2:

(1). Write

$$
\begin{aligned}
&\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}'_E\mathbf{\Lambda}_E\mathbf{F}'_E\|_2\le\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\|\mathbf{\Lambda}_E\|_2^2\\
=&\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\lambda_{\max}\left\{\mathbf{\Lambda}'_E\mathbf{\Lambda}_E\right\}=\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\lambda_{\max}\left\{\sum_{j=1}^N\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\right\}\\
\le&\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\sum_{j=1}^N\lambda_{\max}\left\{\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\right\}\\
=&\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\sum_{j=1}^N\max_{1\le i\le M}\lambda_{\max}\left\{\boldsymbol{\lambda}_{E,ij}\boldsymbol{\lambda}'_{E,ij}\right\}=\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\sum_{j=1}^N\max_{1\le i\le M}\lambda_{\max}\left\{\boldsymbol{\lambda}'_{E,ij}\boldsymbol{\lambda}_{E,ij}\right\}\\
=&\frac{1}{MNT}\|\mathbf{F}_E\|_2^2\cdot\sum_{j=1}^N\max_{1\le i\le M}\|\boldsymbol{\lambda}_{E,ij}\|_F^2\le\frac{1}{MT}\|\mathbf{F}_E\|_2^2\cdot\max_{1\le i\le M,1\le j\le N}\|\boldsymbol{\lambda}_{E,ij}\|_F^2\\
=&\frac{1}{MT}\cdot O_P(T\vee M)\cdot O_P(\ln(MN))=\ln(MN)\cdot O_P\left(M^{-1}\vee T^{-1}\right),
\end{aligned}
$$

where the second equality follows from the definitions under (2.5), and the last equality follows from Assumption 1.2 and Assumption 2.2.

(2). Write

$$
\begin{aligned}
&\frac{1}{MNT}\|\mathbf{F}_I\mathbf{\Lambda}_I'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 \le \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\|\mathbf{\Lambda}_I\|_2^2 \\
&= \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\lambda_{\max}\{\mathbf{\Lambda}_I'\mathbf{\Lambda}_I\} = \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\lambda_{\max}\left\{\mathrm{diag}\{\mathbf{\Lambda}_{I,\bullet 1}'\mathbf{\Lambda}_{I,\bullet 1},\ldots,\mathbf{\Lambda}_{I,\bullet N}'\mathbf{\Lambda}_{I,\bullet N}\}\right\} \\
&= \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le j\le N}\lambda_{\max}\left\{\mathbf{\Lambda}_{I,\bullet j}'\mathbf{\Lambda}_{I,\bullet j}\right\} = \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le j\le N}\lambda_{\max}\left\{\sum_{i=1}^M\boldsymbol{\lambda}_{I,ij}\boldsymbol{\lambda}_{I,ij}'\right\} \\
&\le \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le j\le N}\sum_{i=1}^M\lambda_{\max}\left\{\boldsymbol{\lambda}_{I,ij}\boldsymbol{\lambda}_{I,ij}'\right\} = \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le j\le N}\sum_{i=1}^M\lambda_{\max}\left\{\boldsymbol{\lambda}_{I,ij}'\boldsymbol{\lambda}_{I,ij}\right\} \\
&= \frac{1}{MNT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le j\le N}\sum_{i=1}^M\|\boldsymbol{\lambda}_{I,ij}\|_F^2 \le \frac{1}{NT}\|\mathbf{F}_I\|_2^2\cdot\max_{1\le i\le M,1\le j\le N}\|\boldsymbol{\lambda}_{I,ij}\|_F^2 \\
&= \ln(MN)\cdot O_P\left(N^{-1}\vee T^{-1}\right),
\end{aligned}
$$

where the second equality follows from the definitions under (2.5), and the last equality follows from Assumption 1.2 and Assumption 2.2.

(3) Write

$$
\begin{aligned}
&\frac{1}{M^2N^2T^2}E\|\mathbf{U}'\mathbf{U}\|_F^2 = \frac{1}{M^2N^2T^2}\sum_{t,s=1}^T\sum_{i,m=1}^M\sum_{j,n=1}^N E[u_{ijt}u_{mnt}u_{ijs}u_{mns}] \\
&= \frac{1}{M^2N^2T^2}\sum_{t,s=1}^T\left(\sum_{i=1}^M\sum_{j=1}^N E[u_{ijt}^2u_{ijs}^2] + \sum_{(i,j)\neq(m,n)}E[(u_{ijt}u_{mnt}-\sigma_{ijmn})(u_{ijs}u_{mns}-\sigma_{ijmn})]\right) \\
&\quad+\frac{1}{M^2N^2}\sum_{(i,j)\neq(m,n)}\sigma_{ijmn}^2 \\
&= \frac{1}{M^2N^2T^2}\sum_{t=1}^T\left(\sum_{i=1}^M\sum_{j=1}^N E[u_{ijt}^4] + \sum_{(i,j)\neq(m,n)}E[(u_{ijt}u_{mnt}-\sigma_{ijmn})^2]\right) \\
&\quad+\frac{1}{M^2N^2T^2}\sum_{t\neq s}\left(\sum_{i=1}^M\sum_{j=1}^N E[u_{ijt}^2u_{ijs}^2] + \sum_{(i,j)\neq(m,n)}E[(u_{ijt}u_{mnt}-\sigma_{ijmn})(u_{ijs}u_{mns}-\sigma_{ijmn})]\right) \\
&\quad+\frac{1}{M^2N^2}\sum_{(i,j)\neq(m,n)}\sigma_{ijmn}^2 \\
&= O\left(\frac{1}{T}+\frac{1}{MN}\right),
\end{aligned}
$$

where the last equality follows from Assumption 3.

(4). Recall that in the first result of this lemma we have shown that $\|\mathbf{\Lambda}_E\|_2^2 = O_P(N\ln(MN))$. Then write

$$
\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_E\mathbf{F}_E'\|_2 \le \frac{1}{MNT}\|\mathbf{G}\|_F\cdot\|\mathbf{\Gamma}\|_F\cdot\|\mathbf{\Lambda}_E\|_2\cdot\|\mathbf{F}_E\|_2
$$

$$
\begin{aligned}
&= \frac{1}{MNT} \cdot O_P(\sqrt{T}) \cdot O_P(\sqrt{MN}) \cdot O_P(\sqrt{N \ln(MN)}) \cdot O_P(\sqrt{T} \vee \sqrt{M}) \\
&= O_P\left(\sqrt{\ln(MN)} \cdot (M^{-1/2} \vee T^{-1/2})\right),
\end{aligned}
$$

where the first equality uses the following facts that $\|\mathbf{G}\|_F = O_P(\sqrt{T})$, $\|\mathbf{\Gamma}\|_F = O_P(\sqrt{MN})$, and $\|\mathbf{F}_E\|_2 = O_P(\sqrt{T} \vee \sqrt{M})$.

(5). Similar to the fourth result of this lemma, we write

$$
\begin{aligned}
&\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 \leq \frac{1}{MNT}\|\mathbf{G}\|_F \cdot \|\mathbf{\Gamma}\|_F \cdot \|\mathbf{\Lambda}_I\|_2 \cdot \|\mathbf{F}_I\|_2 \\
&= \frac{1}{MNT} \cdot O_P(\sqrt{T}) \cdot O_P(\sqrt{MN}) \cdot O_P(\sqrt{M \ln(MN)}) \cdot O_P(\sqrt{T} \vee \sqrt{N}) \\
&= O_P\left(\sqrt{\ln(MN)} \cdot (N^{-1/2} \vee T^{-1/2})\right),
\end{aligned}
$$

where the first equality uses the following facts that $\|\mathbf{G}\|_F = O_P(\sqrt{T})$, $\|\mathbf{\Gamma}\|_F = O_P(\sqrt{MN})$, $\|\mathbf{\Lambda}_I\|_2 = O_P(\sqrt{M \ln(MN)})$ and $\|\mathbf{F}_I\|_2 = O_P(\sqrt{T} \vee \sqrt{N})$.

(6). Write

$$
\begin{aligned}
&\frac{1}{M^2N^2T^2} E\|\mathbf{\Gamma}'\mathbf{U}\|_F^2 = \frac{1}{M^2N^2T^2} \sum_{t=1}^{T} E\left\|\sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{ij} u_{ijt}\right\|_F^2 \\
&= \frac{1}{M^2N^2T^2} \sum_{i,m=1}^{M} \sum_{j,n=1}^{N} \sum_{t=1}^{T} E[u_{ijt}u_{mnt}\gamma_{ij}'\gamma_{mn}] \leq O(1)\frac{1}{M^2N^2T} \sum_{i,m=1}^{M} \sum_{j,n=1}^{N} |\sigma_{ijmn}| \\
&= O\left(\frac{1}{MNT}\right),
\end{aligned}
$$

where the last line follows from Assumption 3. Thus, we further write

$$
\frac{1}{MNT}\|\mathbf{G}\mathbf{\Gamma}'\mathbf{U}\|_F \leq \frac{1}{MNT}\|\mathbf{G}\|_F \cdot \|\mathbf{\Gamma}'\mathbf{U}\|_F = O_P\left(\frac{1}{\sqrt{MN}}\right).
$$

(7). Recall that in the first two results of this lemma we have shown that $\|\mathbf{\Lambda}_E\|_2^2 = O_P(N \ln(MN))$ and $\|\mathbf{\Lambda}_I\|_2^2 = O_P(M \ln(MN))$. Write

$$
\begin{aligned}
&\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{\Lambda}_I\mathbf{F}_I'\|_2 \leq \frac{1}{MNT}\|\mathbf{F}_E\|_2 \cdot \|\mathbf{\Lambda}_E\|_2 \cdot \|\mathbf{\Lambda}_I\|_2 \cdot \|\mathbf{F}_I\|_2 \\
&= O_P(1)\frac{1}{MNT} \cdot (\sqrt{T} \vee \sqrt{M}) \cdot \sqrt{N \ln(MN)} \cdot \sqrt{M \ln(MN)} \cdot (\sqrt{T} \vee \sqrt{N}) \\
&= O_P\left(\frac{(\sqrt{T} \vee \sqrt{M}) \cdot (\sqrt{T} \vee \sqrt{N}) \cdot \ln(MN)}{\sqrt{MN}T}\right),
\end{aligned}
$$

where the first equality follows from Assumption 1 and Assumption 2.

(8). Similar to the seventh result, we have

$$
\frac{1}{MNT}\|\mathbf{F}_E\mathbf{\Lambda}_E'\mathbf{U}\|_2 \leq \frac{1}{MNT}\|\mathbf{F}_E\|_2 \cdot \|\mathbf{\Lambda}_E\|_2 \cdot \|\mathbf{U}\|_2
$$

$$
\begin{aligned}
&= O_P(1)\frac{1}{MNT} \cdot (\sqrt{T} \vee \sqrt{M}) \cdot \sqrt{N \ln(MN)} \cdot (\sqrt{T} \vee \sqrt{MN}) \\
&= O_P\left(\frac{(\sqrt{T} \vee \sqrt{M}) \cdot (\sqrt{T} \vee \sqrt{MN}) \cdot \sqrt{\ln(MN)}}{M\sqrt{N}T}\right),
\end{aligned}
$$

where the first equality follows from Assumption 1, Assumption 2 and the development in the third result of this lemma.

(9) The proof is similar to the eighth result of this lemma. Thus, omitted. ■

Proof of Lemma B.3:

(1). We have

$$
\frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \le \frac{1}{\sqrt{T}}\|\mathbf{G}\|_F \cdot \frac{1}{\sqrt{T}}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right),
$$

where the equality follows from Assumption 1 and Lemma A.1.

(2). We have

$$
\begin{aligned}
&\frac{1}{T}\|\widehat{\mathbf{G}}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F = \frac{1}{T}\|(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H} + \mathbf{G}\mathbf{H})'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \\
&= \frac{1}{T}\|(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}) + (\mathbf{G}\mathbf{H})'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \\
&\le \frac{1}{T}\|\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H}\|_F^2 + \|\mathbf{H}\|_F \cdot \frac{1}{T}\|\mathbf{G}'(\widehat{\mathbf{G}} - \mathbf{G}\mathbf{H})\|_F \\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right),
\end{aligned}
$$

where the third equality follows from the first result of this lemma and the fact that $\|\mathbf{H}\|_F = O_P(1)$.

(3). By the first result of this lemma, we obtain that

$$
\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}} - \frac{1}{T}\mathbf{G}'\mathbf{G}\mathbf{H} = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right), \tag{B.1}
$$

which in connection with $\|\mathbf{H}\|_F = O_P(1)$ implies

$$
\frac{1}{T}\mathbf{H}'\mathbf{G}'\widehat{\mathbf{G}} - \frac{1}{T}\mathbf{H}'\mathbf{G}'\mathbf{G}\mathbf{H} = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right). \tag{B.2}
$$

By the second result of this lemma, we obtain that

$$
\frac{1}{T}\widehat{\mathbf{G}}'\widehat{\mathbf{G}} - \frac{1}{T}\widehat{\mathbf{G}}'\mathbf{G}\mathbf{H} = \mathbf{I}_{r_g} - \frac{1}{T}\widehat{\mathbf{G}}'\mathbf{G}\mathbf{H} = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right), \tag{B.3}
$$

Summing up (B.2) and (B.3) yields

$$
\mathbf{I}_{r_g} - \frac{1}{T}\mathbf{H}'\mathbf{G}'\mathbf{G}\mathbf{H} = O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M}, \sqrt{N}, \sqrt{T}\}}\right). \tag{B.4}
$$

Furthermore, write

$$
\begin{aligned}
&\|\mathbf{P}_{\widehat{\mathbf{G}}}-\mathbf{P}_{\mathbf{G}}\|_F^2=\operatorname{tr}\left\{(\mathbf{P}_{\widehat{\mathbf{G}}}-\mathbf{P}_{\mathbf{G}})'(\mathbf{P}_{\widehat{\mathbf{G}}}-\mathbf{P}_{\mathbf{G}})\right\}\\
&=\operatorname{tr}\left\{\mathbf{P}_{\widehat{\mathbf{G}}}-\mathbf{P}_{\widehat{\mathbf{G}}}\mathbf{P}_{\mathbf{G}}-\mathbf{P}_{\mathbf{G}}\mathbf{P}_{\widehat{\mathbf{G}}}+\mathbf{P}_{\mathbf{G}}\right\}\\
&=\operatorname{tr}\left\{\frac{1}{T}\widehat{\mathbf{G}}\widehat{\mathbf{G}}'-\mathbf{P}_{\widehat{\mathbf{G}}}\mathbf{P}_{\mathbf{G}}-\mathbf{P}_{\mathbf{G}}\mathbf{P}_{\widehat{\mathbf{G}}}+\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\right\}\\
&=\operatorname{tr}\left\{\frac{1}{T}\widehat{\mathbf{G}}\widehat{\mathbf{G}}'\right\}-2\cdot\operatorname{tr}\left\{\mathbf{P}_{\widehat{\mathbf{G}}}\mathbf{P}_{\mathbf{G}}\right\}+\operatorname{tr}\left\{\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\right\}\\
&=\operatorname{tr}\left\{\frac{1}{T}\widehat{\mathbf{G}}'\widehat{\mathbf{G}}\right\}-2\cdot\operatorname{tr}\left\{\frac{1}{T}\widehat{\mathbf{G}}\widehat{\mathbf{G}}'\mathbf{P}_{\mathbf{G}}\right\}+\operatorname{tr}\left\{\mathbf{G}'\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\right\}\\
&=\operatorname{tr}\left\{\mathbf{I}_{r_g}\right\}-2\cdot\operatorname{tr}\left\{\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{P}_{\mathbf{G}}\widehat{\mathbf{G}}\right\}+\operatorname{tr}\left\{\mathbf{I}_{r_g}\right\}=2\cdot\operatorname{tr}\left\{\mathbf{I}_{r_g}-\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{P}_{\mathbf{G}}\widehat{\mathbf{G}}\right\}.
\end{aligned}
$$

Using the results in (B.1), it can be shown that

$$
\begin{aligned}
&\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{P}_{\mathbf{G}}\widehat{\mathbf{G}}=\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\widehat{\mathbf{G}}\\
&=\left(\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{G}\right)\left(\frac{1}{T}\mathbf{G}'\mathbf{G}\right)^{-1}\left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}\right)\\
&=\frac{1}{T}\mathbf{H}'\mathbf{G}'\mathbf{G}\mathbf{H}+O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right),
\end{aligned}
$$

which in connection with (B.4) yields

$$
\frac{1}{T}\widehat{\mathbf{G}}'\mathbf{P}_{\mathbf{G}}\widehat{\mathbf{G}}=\mathbf{I}_{r_g}+O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right).
$$

It follows that

$$
\|\mathbf{P}_{\widehat{\mathbf{G}}}-\mathbf{P}_{\mathbf{G}}\|_F^2=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right).
$$

The proof is now complete. ■

It is worth noting that we do not impose any orthogonality conditions between global and local factors (loadings) in proving the above results.

Proof of Lemma B.4:

(1). First, using $\widehat{\mathbf{\Gamma}}_{I,j}=\frac{1}{T}\mathbf{Y}_{I,j}\widehat{\mathbf{G}}$, we note that

$$
\begin{aligned}
&\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\right\|_F=\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\mathbf{Y}_{I,j}\mathbf{P}_{\widehat{\mathbf{G}}}\right\|_F\\
&=\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\left(\mathbf{\Gamma}_{I,j}\mathbf{G}'+\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'+\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'+\mathbf{U}_{\bullet j\bullet}\right)\mathbf{P}_{\widehat{\mathbf{G}}}\right\|_F\\
&=\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}'\mathbf{M}_{\widehat{\mathbf{G}}}-\left(\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'+\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'+\mathbf{U}_{\bullet j\bullet}\right)\mathbf{P}_{\widehat{\mathbf{G}}}\right\|_F\\
&=\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})'\mathbf{M}_{\widehat{\mathbf{G}}}-\left(\operatorname{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'+\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'+\mathbf{U}_{\bullet j\bullet}\right)\mathbf{P}_{\widehat{\mathbf{G}}}\right\|_F.
\end{aligned}
$$

For the first term on the right hand side above, we have

$$\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})\mathbf{\Gamma}_{I,j}' = -\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{A}_2+\cdots+\mathbf{A}_{16})\widehat{\mathbf{G}}\left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}_{I,j}', \quad \text{(B.5)}$$

where $\mathbf{A}_2,\ldots,\mathbf{A}_{16}$ are defined in Lemma A.1 already. In addition, for notational simplicity, let $\mathbf{K}=\left(\frac{1}{T}\mathbf{G}'\widehat{\mathbf{G}}\right)^{-1}\left(\frac{1}{MN}\mathbf{\Gamma}'\mathbf{\Gamma}\right)^{-1}$ in what follows. It is easy to know that $\|\mathbf{K}\|_F=O_P(1)$.

Write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{G}-\widehat{\mathbf{G}}\mathbf{H}^{-1})\mathbf{\Gamma}_{I,j}'\|_F \le O(1)\cdot\frac{1}{\sqrt{MT}}\|\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{A}_2+\cdots+\mathbf{A}_{16})\widehat{\mathbf{G}}\mathbf{K}\mathbf{\Gamma}_{I,j}'\|_2\\
&\le O(1)\cdot\frac{1}{\sqrt{MT}}\|\mathbf{M}_{\widehat{\mathbf{G}}}(\mathbf{A}_2+\cdots+\mathbf{A}_{16})\widehat{\mathbf{G}}\|_2\cdot\|\mathbf{K}\|_F\cdot\|\mathbf{\Gamma}_{I,j}\|_F,\\
&\le O_P(1)\frac{1}{\sqrt{MT}}\cdot\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\cdot\sqrt{T}\cdot\sqrt{M}\\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right),
\end{aligned}$$

where the first inequality follows from (B.5) and $\|\mathbf{\Omega}\|_F\le\sqrt{\text{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the third inequality follows from Lemma B.2 and the facts that $\|\widehat{\mathbf{G}}\|_F=O(\sqrt{T})$ and $\|\mathbf{\Gamma}_{I,j}\|_F=O_P(\sqrt{M})$.

Next, we consider the second expansion.

$$\begin{aligned}
(\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'+\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'+\mathbf{U}_{\bullet j\bullet})\mathbf{P}_{\widehat{\mathbf{G}}} &= \text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'\mathbf{P}_{\widehat{\mathbf{G}}}+\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{P}_{\widehat{\mathbf{G}}}+\mathbf{U}_{\bullet j\bullet}\mathbf{P}_{\widehat{\mathbf{G}}}\\
&:= \mathbf{D}_1+\mathbf{D}_2+\mathbf{D}_3.
\end{aligned}$$

For $\mathbf{D}_1$, write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_1\|_F=\frac{1}{\sqrt{MT}}\|\,\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'\mathbf{P}_{\widehat{\mathbf{G}}}\|_F\\
&\le O(1)\frac{1}{\sqrt{MT}}\|\,\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}'\mathbf{F}_E'\mathbf{P}_{\widehat{\mathbf{G}}}\|_2\le O(1)\frac{1}{\sqrt{MT}}\|\,\text{diag}\{\mathbf{\Lambda}_{E,\bullet j}'\}\|_2\cdot\|\mathbf{F}_E\|_2\\
&\le O_P(1)\frac{1}{\sqrt{MT}}\cdot\sqrt{\ln(MN)}\cdot(\sqrt{M}\vee\sqrt{T})=O_P\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\right),
\end{aligned}$$

where the first inequality follows from $\|\mathbf{\Omega}\|_F\le\sqrt{\text{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the second inequality follows from $\|\mathbf{P}_{\widehat{\mathbf{G}}}\|_2=1$, and the last inequality follows from Assumption 1 and Assumption 2.

Consider $\mathbf{D}_2$, and write

$$\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_2\|_F=\frac{1}{\sqrt{MT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{P}_{\widehat{\mathbf{G}}}\|_F\le\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\widehat{\mathbf{G}}\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\le\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'(\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H}+\mathbf{G}\mathbf{H})\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\le\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'(\widehat{\mathbf{G}}-\mathbf{G}\mathbf{H})\|_F\cdot\|\widehat{\mathbf{G}}\|_F+\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}_{I,j}'\mathbf{G}\mathbf{H}\|_F\cdot\|\widehat{\mathbf{G}}\|_F.
\end{aligned}$$

Note that

$$
\begin{aligned}
&\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}(\widehat{\mathbf{G}}-\mathbf{GH})\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\le O_P(1)\frac{1}{\sqrt{M}}\cdot\sqrt{M}\cdot\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right),
\end{aligned}
$$

where the inequality follows from Assumption 2, Lemma A.1, and the fact that $\|\mathbf{F}_{I,j}\|_F = O_P(\sqrt{T})$. Note further that

$$
\begin{aligned}
&\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\mathbf{GH}\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\le\frac{1}{\sqrt{MTT}}\|\mathbf{\Lambda}_{I,\bullet j}\|_F\cdot\|\mathbf{F}'_{I,j}\mathbf{G}\|_F\cdot\|\mathbf{H}\|_F\cdot\|\widehat{\mathbf{G}}\|_F\\
&\le O_P(1)\frac{1}{\sqrt{M}}\cdot\sqrt{M}\cdot T^{a_{I,j}}=O_P(T^{a_{I,j}}),
\end{aligned}
$$

where the second inequality follows from Assumption 4.1.(a).

Consider $\mathbf{D}_3$ and write

$$
\begin{aligned}
&\frac{1}{\sqrt{MT}}\|\mathbf{D}_3\|_F=\frac{1}{\sqrt{MT}}\|\mathbf{U}_{\bullet j\bullet}\mathbf{P}_{\widehat{\mathbf{G}}}\|_F\le O(1)\frac{1}{\sqrt{MT}}\|\mathbf{U}_{\bullet j\bullet}\mathbf{P}_{\widehat{\mathbf{G}}}\|_2\\
&\le O(1)\frac{1}{\sqrt{MT}}\|\mathbf{U}_{\bullet j\bullet}\|_2=O_P\left(\frac{1}{\sqrt{M}}\vee\frac{1}{\sqrt{T}}\right),
\end{aligned}
$$

where the first inequality follows from $\|\mathbf{\Omega}\|_F\le\sqrt{\mathrm{rank}(\mathbf{\Omega})}\cdot\|\mathbf{\Omega}\|_2$, the second inequality follows from $\|\mathbf{P}_{\widehat{\mathbf{G}}}\|_2=1$, and the last equality follows from a development similar to Lemma B.2.3 using Assumption 4.

Based on the above development, we obtain that

$$
\begin{aligned}
&\frac{1}{\sqrt{MT}}\left\|\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\right\|_F\\
&=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)+O_P\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}}\right)+O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}\right)\\
&\quad+O_P(T^{a_{I,j}})+O_P\left(\frac{1}{\sqrt{M}}\vee\frac{1}{\sqrt{T}}\right)\\
&=O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}}+T^{a_{I,j}}\right).
\end{aligned}
$$

The proof of the first result is then complete.

(2). The second result can be proved in exactly the same way as the first result. Thus, omitted. ■

Proof of Lemma B.5:

(1). Write

$$
\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}'-\widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\,\mathrm{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\|_2
$$

$$
\begin{aligned}
&\le \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2 \cdot \|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 \cdot \|\mathbf{F}_E\|_2 \\
&= \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2 \cdot \sqrt{\max_{i\ge 1, j\ge 1}\|\boldsymbol{\lambda}_{E,ij}\|_F^2} \cdot \|\mathbf{F}_E\|_2 \\
&= O_P(1)\frac{1}{\sqrt{MT}} \cdot \left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right) \cdot \sqrt{\ln(MN)} \cdot (\sqrt{T} \vee \sqrt{M}) \\
&= O_P(1)\frac{\sqrt{\ln(MN)}}{\sqrt{M}\wedge\sqrt{T}} \cdot \left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right),
\end{aligned}
$$

where the second equality follows from Lemma B.4, Assumption 1 and Assumption 2.

(2). Write

$$
\begin{aligned}
&\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\|_2 \\
&\le \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2 \cdot \|\mathbf{\Lambda}_{I,\bullet j}\|_2 \cdot \|\mathbf{F}_{I,j}\|_2 \\
&= O_P(1)\frac{1}{\sqrt{MT}} \cdot \left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right) \cdot \sqrt{M} \cdot \sqrt{T} \\
&= O_P\left(\frac{\sqrt{\ln(MN)}}{\min\{\sqrt{M},\sqrt{N},\sqrt{T}\}} + T^{a_{I,j}}\right),
\end{aligned}
$$

where the first equality follows from Lemma B.4.

(3). The proof is similar to the second result of this lemma, the result follows from $\|\mathbf{U}_{\bullet j \bullet}\|_2 = O_P(\sqrt{M} \vee \sqrt{T})$.

(4). Write

$$
\begin{aligned}
&\frac{1}{MT}\|\mathbf{F}_E \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\} \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}'\mathbf{F}'_E\|_2 \\
&\le \frac{1}{MT}\|\mathbf{F}_E\|_2^2 \cdot \lambda_{\max}\{\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}' \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\} \\
&\le \frac{1}{MT}\|\mathbf{F}_E\|_2^2 \cdot \max_{i\ge 1, j\ge 1}\|\boldsymbol{\lambda}_{E,ij}\|_F^2 \\
&= O_P(1)\frac{1}{MT} \cdot (T \vee M) \cdot \ln(MN) = O_P\left(\frac{\ln(MN)}{M \wedge T}\right),
\end{aligned}
$$

where the first equality follows from Assumption 1 and Assumption 2.

(5). Write

$$
\begin{aligned}
&\frac{1}{MT}\|\mathbf{F}_E \operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\mathbf{\Lambda}_{I,\bullet j}\mathbf{F}'_{I,j}\|_2 \\
&\le \frac{1}{MT}\|\mathbf{F}_E\|_2 \cdot \|\operatorname{diag}\{\mathbf{\Lambda}'_{E,\bullet j}\}\|_2 \cdot \|\mathbf{\Lambda}_{I,\bullet j}\|_2 \cdot \|\mathbf{F}'_{I,j}\|_2 \\
&\le \frac{1}{MT}\|\mathbf{F}_E\|_2 \cdot \sqrt{\max_{i\ge 1, j\ge 1}\|\boldsymbol{\lambda}_{E,ij}\|_F^2} \cdot \|\mathbf{\Lambda}_{I,\bullet j}\|_2 \cdot \|\mathbf{F}'_{I,j}\|_2 \\
&= O_P(1)\frac{1}{MT} \cdot (\sqrt{T} \vee \sqrt{M}) \cdot \sqrt{\ln(MN)} \cdot \sqrt{M} \cdot \sqrt{T}
\end{aligned}
$$

$$= O_P\left(\frac{\sqrt{\ln(MN)}}{\sqrt{M} \wedge \sqrt{T}}\right),$$

where the second inequality follows from a procedure similar to the third result of this lemma, and the first equality follows from Assumption 1 and Assumption 2.

(6). Similar to the proof of the fifth result, the result follows from $\|\mathbf{U}_{\bullet j \bullet}\|_2 = O_P(\sqrt{M} \vee \sqrt{T})$.

(7). Write

$$\frac{1}{MT}\|(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')'(\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}')\|_2 \leq \frac{1}{MT}\|\mathbf{\Gamma}_{I,j}\mathbf{G}' - \widehat{\mathbf{\Gamma}}_{I,j}\widehat{\mathbf{G}}'\|_2^2$$
$$= O_P\left(\frac{\ln(MN)}{\min\{M, N, T\}} + T^{2a_{I,j}}\right),$$

where the first equality follows from Lemma B.4.

(8). Write

$$\frac{1}{MT}\|\mathbf{F}_{I,j}\mathbf{\Lambda}'_{I,\bullet j}\mathbf{U}_{\bullet j \bullet}\|_2 \leq \frac{1}{MT}\|\mathbf{F}_{I,j}\|_2 \cdot \|\mathbf{\Lambda}_{I,\bullet j}\mathbf{U}_{\bullet j \bullet}\|_F$$
$$= O_P(1)\frac{1}{MT} \cdot \sqrt{T} \cdot \sqrt{MT} = O_P\left(\frac{1}{\sqrt{M}}\right),$$

where the first equality follows from a development similar to Lemma B.2.6 using Assumption 3 and Assumption 4.

(9). Write

$$\frac{1}{MT}\|\mathbf{U}'_{\bullet j \bullet}\mathbf{U}_{\bullet j \bullet}\|_2 = O_P\left(\frac{1}{\sqrt{M}} \vee \frac{1}{\sqrt{T}}\right),$$

where the result follows from a development similar to Lemma B.2.3 using Assumption 3. ■