Time-Varying Matrix Factor Model*

Bin Chen[†1], Elynn Y. Chen [‡2], and Rong Chen[§3]

[1]Department of Economics, University of Rochester
[2]Stern School of Business, New York University
[3]Department of Statistics, Rutgers University

February 28, 2023

Abstract

High-dimensional matrix-valued data are frequently encountered in finance and economics, such as international trade flow data covering multiple countries over an extended period. To account for potential structural changes and identify the underlying information context in the matrix structure, we propose a time-varying matrix factor model that allows for smooth changes in factor loadings over time. Our nonparametric principal component analysis (PCA) method uncovers the latent time-varying dynamic structure and achieves dimension reduction. We establish the consistency and asymptotic normality of our estimator under general conditions that account for correlations across time, rows, or columns of the noise. Our simulation study demonstrates the effectiveness of our proposed estimator. Using international trade flow data, we apply our model to investigate trading hubs, centrality, patterns, and trends in the trading network.

JEL Classifications: *C13, C14, C32, C55*
Key words: Factor models, High dimension, Kernel, Latent low rank, Matrix value, Time-varying.

---

*We thank Zeyu Wang and Yucheng Wong for their excellent research assistance. We also thank Atsushi Inoue, Mahrad Sharifvaghefi, Mark Watson, and seminar participants at Vanderbilt University, the 2022 International Symposium on Econometric Theory and Applications (SETA) and the 15th Greater New York Metropolitan Area Econometrics Colloquium for their useful comments and discussions. Any remaining errors are solely ours.

†binchen@rochester.edu
‡elynn.chen@stern.nyu.edu
§rongchen@stat.rutgers.edu

1 Introduction

Matrix-valued data are commonly encountered in various fields such as economics, finance, health care, and social networks. Two notable categories of such data are dynamic transportation networks and dynamic panel data. The former refers to monthly import-export volumes among countries, which naturally form a dynamic sequence of matrix variates representing weighted directional transportation networks. The latter pertains to investors' information collection for portfolio selection, where observations are matrices with rows representing different firms and columns representing stock prices and various firm characteristics such as book-to-market ratio, dividend-to-price ratio, earnings-to-price ratio, cash flow-to-price ratio, and others.

Statistical methods for analyzing matrix-valued data are still in their early stages of development. One of the most popular models for achieving dimension reduction while exploring matrix structure is the matrix factor model. Wang et al. (2019) first considered a matrix factor model in time series, where both factors and factor loadings are unknown matrices. By exploring the serial correlation of the latent factors, they estimated the model via an eigenanalysis of the auto-cross-covariance matrix. The model was further extended by Chen et al. (2019) to the constrained version and Liu and Chen (2019) to the threshold matrix factor model. Different from the above works, Chen and Fan (2021) proposed an estimation, which is based on the eigenanalysis of an aggregate of the sample mean matrix and the contemporary covariance matrices.

Despite the success of this class of matrix factor models, the maintained assumption is that factor loadings are fixed over time. However, there exist various driving forces, such as policy switches, preference changes, technological progress, and environmental issues that might affect the relationship between economic variables. Some empirical evidence (see, e.g., Chen et al. (2019)) suggests that loading instability is a concern for the economic applications of matrix factor models. Failing to take into account potential structural changes in factor loadings could lead to inconsistent estimation and unreliable inference.

Recently, there has been a growing interest in smooth structural changes, prompting the development of novel time-varying time series and panel data models to capture the evolutionary behavior of economic variables. Among these models, the nonparametric time-varying parameter time series or panel data model has gained considerable attention. First introduced by Robinson (1989) and further studied by Cai (2007), Chen and Hong (2012), Zhang and Wu (2012), Inoue et al. (2017), Chen and Maung (2022), and others, this model's advantage lies in its flexibility, with little restrictions imposed on the functional forms of the time-varying intercept and slope, except for the condition that they evolve smoothly over time. In this paper, we adopt this framework in our matrix factor models.

In this paper, we introduce a more parsimonious representation of matrix-valued data by generalizing the constant loading matrix factor models. Specifically, we model the loading matrices as an unknown function of time and estimate them, along with the latent factor matrix, using local PCA. We develop inferential theory, including establishing consistency, rate of convergence, and limiting distributions under general conditions that allow for time-varying loading matrices, as well as correlations across time, rows, or columns of the noise. Furthermore, we generalize the eigenvalue ratio-based estimator (Ahn and Horenstein (2013)) for latent dimensions to time-varying matrix factor models and show its consistency. Additionally, we consider a projection-based estimation approach, where we project the observation matrix onto the row or column factor space, and conduct a local eigenanalysis on a lower-dimensional matrix. We demonstrate that the projection-based estimator achieves faster convergence by increasing the signal-to-noise ratio (see, e.g., Yu et al. (2022)).

In the classical framework of high-dimensional factor models, Su and Wang (2017) employs a local PCA to estimate time-varying factor loadings. While this method can handle matrix-valued observations by "flattening" them into vectors or separately modeling each dimension, it destroys the natural matrix structure and fails to capture important patterns in large-scale data with complex structures, leading to sub-optimal results. We would further compare our matrix-based estimator with its vectorized counterpart via simulation.

1.1 Notations and Organization

Let lowercase letter y, boldface letter $\boldsymbol{y}$, and boldface capital letter $\boldsymbol{Y}$ represent scalar, vector, and matrix, respectively. For any matrix $\boldsymbol{Y}$, we use $\boldsymbol{Y}_{i\cdot}$, $\boldsymbol{y}_{\cdot j}$, and y_{ij} to denote its i-th row, j-th column, and ij-th entry, respectively. All vectors are column vectors and row vectors are written as $\boldsymbol{y}^\top$ for any vector $\boldsymbol{y}$. We use the following matrix norms: maximum norm $\|\boldsymbol{Y}\|_{max} \triangleq \max_{ij} |y_{ij}|$, ℓ_1-norm $\|\boldsymbol{Y}\|_1 \triangleq \max_{j} \sum_i |y_{ij}|$, ℓ_∞-norm $\|\boldsymbol{Y}\|_\infty \triangleq \max_{i} \sum_j |y_{ij}|$, and ℓ_2-norm $\|\boldsymbol{Y}\|_2 \triangleq \sigma_1$, where σ_1 is the largest singular value σ_i of $\boldsymbol{Y}$ with σ_i be the i-th largest square root of eigenvalues of $\boldsymbol{Y}^\top \boldsymbol{Y}$. When $\boldsymbol{Y}$ is a square matrix, we denote by $\mathrm{Tr}(\boldsymbol{Y})$, $\lambda_{max}(\boldsymbol{y})$, and $\lambda_{min}(\boldsymbol{y})$ the trace, maximum and minimum singular value of $\boldsymbol{Y}$, respectively.

The remaining sections of this paper are organized as follows. Section 2 introduces high-dimensional matrix factor models with time-varying loadings. In Section 3, we present a local PCA estimation procedure and a generalized eigenvalue ratio-based estimator for the latent dimensions. Section 4 establishes the consistency and asymptotic normality of the estimated loading matrices and the generalized eigenvalue ratio-based estimator. Additionally, Section 5 discusses the projection-based estimator, which achieves a faster convergence rate. Section 6 presents a study of the finite sample performance of our estimation through simulation. Section 7 provides an empirical study, and Section 8 concludes the paper. All mathematical proofs are included in the Appendix.

2 Matrix Factor Models with Time-Varying Factor Loadings

Let $\boldsymbol{Y}_t$ $(1 \le t \le T)$ be a matrix-valued time series, where each $\boldsymbol{Y}_t$ is a matrix of size $p \times q$,

$$\boldsymbol{Y}_t = \begin{bmatrix} Y_{t,11} & \cdots & Y_{t,1q} \\ \vdots & \ddots & \vdots \\ Y_{t,p1} & \cdots & Y_{t,pq} \end{bmatrix}.$$

We consider the following matrix factor models with time-varying factor loadings,

$$\boldsymbol{Y}_t = \boldsymbol{R}_t \boldsymbol{F}_t \boldsymbol{C}_t^\top + \boldsymbol{E}_t, \quad t = 1, 2, \cdots, T, \tag{2.1}$$

where $\boldsymbol{F}_t$ is the $k \times r$ common factor matrix, $\boldsymbol{R}_t$ $(\boldsymbol{C}_t)$ is a $p \times k$ $(q \times r)$ time-varying row (column) loading matrix and the sequence of matrices $\boldsymbol{E}_t$ is the noise matrix. The noise term $\boldsymbol{E}_t$ is assumed to be uncorrelated with $\boldsymbol{F}_t$, yet is allowed to be weakly correlated across rows, columns, and times.

Our model generalizes Wang et al. (2019) and Chen and Fan (2021) matrix factor models by allowing for structural changes in factor loadings $\boldsymbol{R}$ and $\boldsymbol{C}$. To cover a wide range of potential time variation, we follow the literature on smooth time-varying parameter models (e.g., Robinson (1989), Cai (2007), Su and Wang (2017)) and model $\boldsymbol{R}_{t,i.}$ $(\boldsymbol{C}_{t,j.})$ as a non-stochastic function of t/T, that is,

$$\begin{aligned} \boldsymbol{R}_{t,i.} &= \boldsymbol{R}_{i.}(t/T), \\ \boldsymbol{C}_{t,j.} &= \boldsymbol{C}_{j.}(t/T), \end{aligned}$$

where $\boldsymbol{R}_{t,i.}(\cdot)$ $(\boldsymbol{C}_{t,j.}(\cdot))$ is an unknown smooth function of t/T on $[0,1]$ for each i (j). The specification that loading matrices $\boldsymbol{R}_{t,i.}(\cdot)$ and $\boldsymbol{C}_{t,j.}(\cdot)$ are some functions of ratio t/T rather than time t only is a common scaling scheme in the literature (see, e.g., Phillips (1990), Robinson (1991) and Cai (2007)). The reason for this specification is that nonparametric loading estimators for $\boldsymbol{R}_{t,i.}(\cdot)$ and $\boldsymbol{C}_{t,j.}(\cdot)$ will not be consistent unless the amount of data on which they depend increases, and merely increasing the sample size will not necessarily improve the estimation of $\boldsymbol{R}_{t,i.}(\cdot)$ and $\boldsymbol{C}_{t,j.}(\cdot)$ at some fixed point t, even if some smoothness conditions are imposed. The amount of local information must increase suitably if the variance and bias of nonparametric estimators of $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$ are to decrease suitably.

Now take the international trade flow as an example. The observed matrix $\mathbf{Y}_t$ is a square matrix, where $p = q = n$. A general entry of $\mathbf{Y}_t$, denoted as $y_{t,ij}$, where $i, j = 1, 2, ..., n$, represents the volume of trade flow from country i to country j at time t. Thus the i-th row represents data for which country i is the exporter and the column j represents data for which country j is the importer. Figure 1 shows a time series plot of $\mathbf{Y}_t$ from January 1982 to December 2018. Model (2.1) identifies $k \times r$ latent factors, which can be viewed as k export hubs and r import hubs. Each country exports to the k exporting hubs in certain distributions (determined by the loading matrix $\boldsymbol{R}_t$) and imports from the r importing hubs in certain distributions (determined by the loading matrix $\boldsymbol{C}_t$). The entry of $\mathbf{R}_t$, denoted as $R_{t,il}$, where $i = 1, ..., n$ and $l = 1, ..., k$, can be interpreted as the export contribution of country i to the export hub l at time t; the entry of $\mathbf{C}_t$, denoted as $C_{t,jm}$, where $j = 1, ..., n$ and $m = 1, ..., r$, can be interpreted as the import contribution of country j to the import hub m at time t. The hubs trade, on behave of the participating countries, among themselves and also within the hubs. The trading volume among the hubs is reflected by the factor matrix $\mathbf{F}_t$, which is changing over time. The (l, m)-th element $F_{t,lm}$, where $l = 1, ..., k$ and $m = 1, ..., r$, reflects the export trading volume from hub l to hub m at time point t. We allow for time-varying export and import contributions and factors, but the numbers of factors are fixed over time.

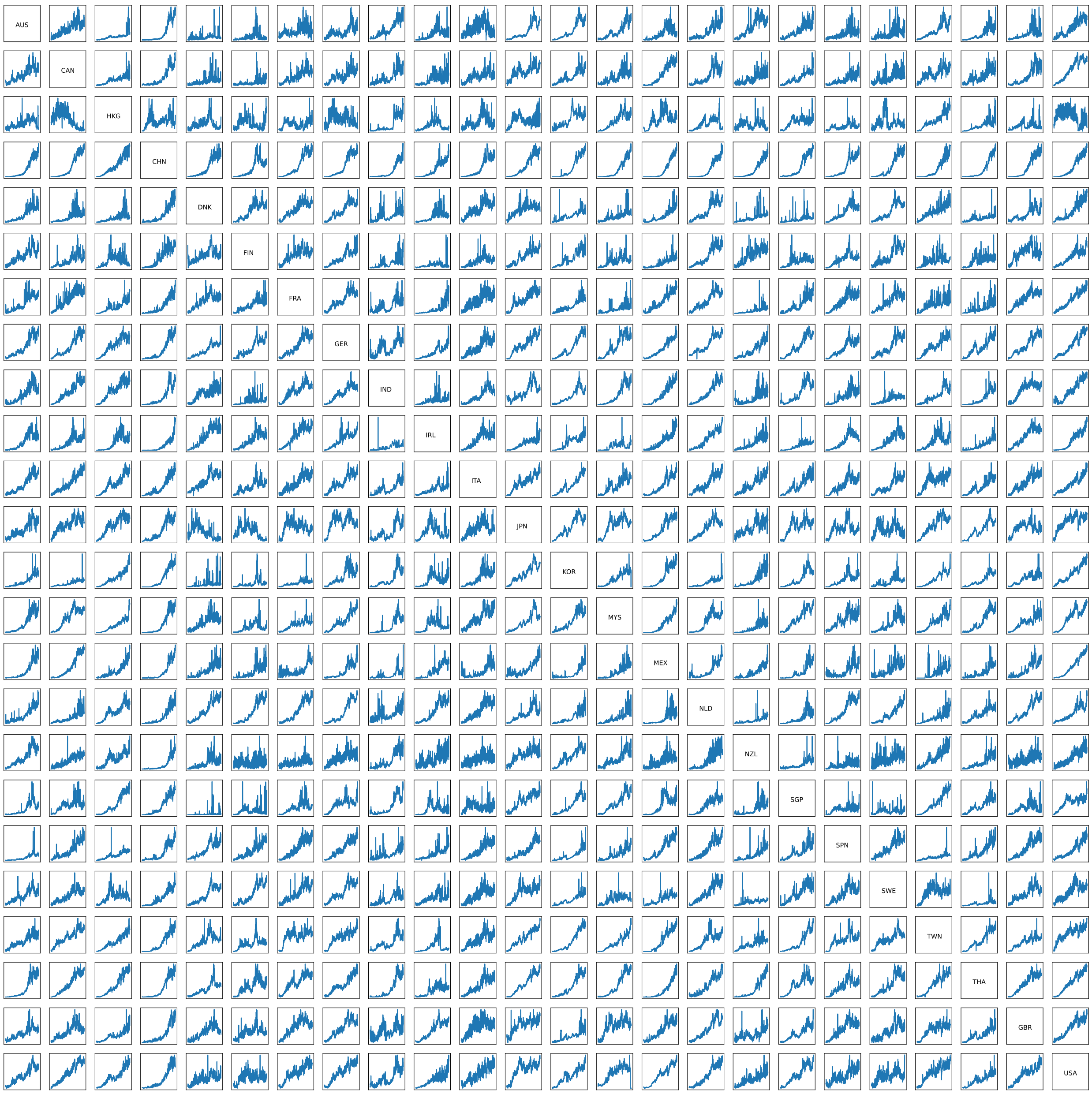
3 Estimation

Fix $s \in \{1, 2, \ldots, T\}$. Under the assumption that $\boldsymbol{R}_{t,i.} = \boldsymbol{R}_{i}.(t/T)$ ($\boldsymbol{C}_{t,j.} = \boldsymbol{C}_{j}.(t/T)$) and $\boldsymbol{R}_{i}.(\boldsymbol{C}_{i}.)$ is a smooth function, we have, when $\frac{t}{T} \approx \frac{s}{T}$

$$\boldsymbol{R}_{t,i.} = \boldsymbol{R}_{i}.(t/T) \approx \boldsymbol{R}_{i}.(s/T) = \boldsymbol{R}_{s,i.}$$

$$\boldsymbol{C}_{t,j.} = \boldsymbol{C}_{j}.(t/T) \approx \boldsymbol{C}_{j}.(s/T) = \boldsymbol{C}_{s,j.}$$

Figure 1: Time series plots of the value of goods traded among 24 countries.



Notes: (1) sample period: January 1982- December 2018. (2) The plots only show the patterns of the time series while the magnitudes are not comparable between plots because the ranges of the y-axis are different.

It follows that

$$\boldsymbol{Y}_s \approx \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top + \boldsymbol{E}_s, \tag{3.1}$$

when $\frac{t}{T} \approx \frac{s}{T}$. To estimate $\boldsymbol{R}_t$, $\boldsymbol{C}_t$, and $\boldsymbol{F}_t$, we minimize the local unexplained variation and bias jointly:

$$\begin{aligned} \underset{\boldsymbol{R}_t, \boldsymbol{C}_t, \{\boldsymbol{F}_t\}_{t=1}^T}{\text{minimize}} \quad & \underbrace{\frac{1}{pq} \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,st} \left(\boldsymbol{Y}_s - \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \right) \right\|_F^2}_{\text{local sample bias}} + \underbrace{\frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} \left\| \boldsymbol{Y}_s - \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \right\|_F^2}_{\text{local sample variance}} \\ \text{subject to} \quad & \frac{1}{p} \boldsymbol{R}_t^\top \boldsymbol{R}_t = \boldsymbol{I}, \ \ \frac{1}{q} \boldsymbol{C}_t^\top \boldsymbol{C}_t = \boldsymbol{I}, \end{aligned} \tag{3.2}$$

where $K_{h,st}$ is a boundary-modified kernel function:

$$K_{h,st} = \begin{cases} h^{-1} k(\frac{s-t}{Th}) / \int_{-t/(Th)}^{1} k(u)\, du & if\, t \in [1, \lfloor Th \rfloor], \\ h^{-1} k(\frac{s-t}{Th}) & if\, t \in [\lfloor Th \rfloor, T - \lfloor Th \rfloor], \\ h^{-1} k(\frac{s-t}{Th}) / \int_{-1}^{(1-t/T)/h} k(u)\, du & otherwise, \end{cases} \tag{3.3}$$

the kernel $k(\cdot) : [-1, 1] \mapsto \mathbb{R}^+$ is a prespecified symmetric probability density, $h = h(p, q, T)$ is a bandwidth parameter for row (column) loading matrix $\boldsymbol{R}_t$ ($\boldsymbol{C}_t$), and $\lfloor Th \rfloor$ denotes the integer part of Th. Examples of $k(\cdot)$ include the uniform, Epanechnikov, and quartic kernels. This boundary-modified kernel function has been used in Hong and Li (2005) and Su and Wang (2017).

Equation (3.2) is non-convex. Thus, we consider an approximate solution by maximizing row and column local variances, respectively, after projection, following Chen and Fan (2021). First, $\{\boldsymbol{Y}_t\}_{1 \le t \le T}$ are projected onto $\boldsymbol{R}_t$ and maximize the row variances of $\boldsymbol{R}_t^\top \boldsymbol{Y}_t$.

$$\begin{aligned} \widehat{\boldsymbol{R}}_t &= \underset{\boldsymbol{R}_t}{\arg\max} \ Tr \left\{ \mathbb{E} \left[\left(\boldsymbol{R}_t^\top \boldsymbol{Y}_t \right) \left(\boldsymbol{R}_t^\top \boldsymbol{Y}_t \right)^\top + \left(\boldsymbol{R}_t^\top \boldsymbol{Y}_t - \mathbb{E}\left[\boldsymbol{R}_t^\top \boldsymbol{Y}_t \right] \right) \left(\boldsymbol{R}_t^\top \boldsymbol{Y}_t - \mathbb{E}\left[\boldsymbol{R}_t^\top \boldsymbol{Y}_t \right] \right)^\top \right] \right\}, \\ &= \underset{\boldsymbol{R}_t}{\arg\max} \ Tr \left\{ pq \boldsymbol{R}_t^\top \boldsymbol{M}_{R,t} \boldsymbol{R}_t \right\}, \\ & \text{subject to} \ \ \frac{1}{p} \boldsymbol{R}_t^\top \boldsymbol{R}_t = \boldsymbol{I}, \end{aligned}$$

where $\boldsymbol{M}_{R,t} \triangleq p^{-1} q^{-1} \mathbb{E}\left[\boldsymbol{Y}_t \boldsymbol{Y}_t^\top \right]$.

Alternatively, $\{\boldsymbol{Y}_t\}_{1 \le t \le T}$ are projected onto $\boldsymbol{C}_t$ and maximize the column variances of

$\boldsymbol{Y}_t\boldsymbol{C}_t$.

$$
\begin{aligned}
\widehat{\boldsymbol{C}}_t &= \underset{\boldsymbol{C}_t}{\arg\max}\ Tr\Big\{\mathbb{E}\Big[(\boldsymbol{Y}_t\boldsymbol{C}_t)^\top(\boldsymbol{Y}_t\boldsymbol{C}_t)+(\boldsymbol{Y}_t\boldsymbol{C}_t-\mathbb{E}[\boldsymbol{Y}_t\boldsymbol{C}_t])^\top(\boldsymbol{Y}_t\boldsymbol{C}_t-\mathbb{E}[\boldsymbol{Y}_t\boldsymbol{C}_t])\Big]\Big\}\\
&= \underset{\boldsymbol{C}_t}{\arg\max}\ Tr\Big\{pq\boldsymbol{C}_t^\top\boldsymbol{M}_{C,t}\boldsymbol{C}_t\Big\},\\
&\text{subject to } \frac{1}{p}\boldsymbol{C}_t^\top\boldsymbol{C}_t=\boldsymbol{I},
\end{aligned}
$$

where $\boldsymbol{M}_{C,t}\triangleq p^{-1}q^{-1}\mathbb{E}\Big[\boldsymbol{Y}_t^\top\boldsymbol{Y}_t\Big]$.

With T noisy observations, we define sample scatter matrix $\widehat{\boldsymbol{M}}_{R,t}$ and $\widehat{\boldsymbol{M}}_{C,t}$ as

$$
\begin{aligned}
\widehat{\boldsymbol{M}}_{R,t} &= \frac{1}{pqT}\sum_{s=1}^{T}K_{h,ts}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top\\
\widehat{\boldsymbol{M}}_{C,t} &= \frac{1}{pqT}\sum_{s=1}^{T}K_{h,ts}\boldsymbol{Y}_s^\top\boldsymbol{Y}_s
\end{aligned}
\tag{3.4}
$$

Furthermore, an estimator of F_t can be obtained by

$$
\widehat{\boldsymbol{F}}_t=\frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top\boldsymbol{Y}_t\widehat{\boldsymbol{C}}_t
$$

and the signal part $\boldsymbol{S}_t=\boldsymbol{R}_tF_t\boldsymbol{C}_t^\top$ can be estimated by

$$
\widehat{\boldsymbol{S}}_t=\frac{1}{pq}\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{R}}_t^\top\boldsymbol{Y}_t\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{C}}_t^\top.
$$

The above local PCA assumes that the latent dimensions $k\times r$ are known. However, in practice, we need to estimate k and r as well. We extend the eigenvalue ratio-based estimator, proposed by Ahn and Horenstein (2013). Let $\widehat{\lambda}_{1,t}\geq\widehat{\lambda}_{2,t}\geq\cdots\geq\widehat{\lambda}_{k,t}\geq 0$ be the ordered eigenvalues of $\widehat{\boldsymbol{M}}_{R,t}$. And the generalized eigenvalue ratio-based estimator for k can be defined as

$$
\widehat{k}=\underset{1\leq j\leq k_{max}}{\operatorname{argmax}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\widehat{\lambda}_{j,t}}{\widehat{\lambda}_{j+1,t}},
\tag{3.5}
$$

and k_{max} is a selected upper bound. We could choose $k_{max}=\lfloor p/2\rfloor$ or $k_{max}=\lfloor p/3\rfloor$, following Ahn and Horenstein (2013) and Chen and Fan (2021). Ratio estimator $\widehat{r}$ can be defined analogously with respect to $\widehat{\boldsymbol{M}}_{C,t}$. Other estimators such as a BIC-based estimator can be applied as well.

4 Theory

Let $\boldsymbol{R}_t^{(1)}\left(\boldsymbol{C}_t^{(1)}\right)$ and $\boldsymbol{R}_t^{(2)}\left(\boldsymbol{C}_t^{(2)}\right)$ denote the first and second order derivative of $\boldsymbol{R}_t$ ($\boldsymbol{C}_t$) respectively, and $\boldsymbol{R}_t^{(0)} = \boldsymbol{R}_t$ and $\boldsymbol{C}_t^{(0)} = \boldsymbol{C}_t$. To derive the asymptotic properties of our estimators, we impose the following regularity conditions.

Assumption 1. ***α-mixing.*** *The vector-valued process $\{\textsc{vec}(\boldsymbol{F}_t)^\top, \textsc{vec}(\boldsymbol{E}_t)^\top\}^\top$ is a stationary α-mixing process with mixing coefficients $\alpha(h) = \sup_{\tau} \sup_{A\in\mathcal{F}_{-\infty}^{\tau}, B\in\mathcal{F}_{\tau+h}^{\infty}} |P(A\cap B) - P(A)P(B)|$, where $\mathcal{F}_\tau^s$ is the σ-field generated by $\{\textsc{vec}(\boldsymbol{F}_t)^\top, \textsc{vec}(\boldsymbol{E}_t)^\top\}^\top : \tau \le t \le s\}$, and the mixing coefficients satisfy the condition that*

$$\sum_{h=1}^{\infty} \alpha(h)^{1-2/\gamma} < \infty,$$

for some $\gamma > 2$.

Assumption 2. ***Factor and noise matrices.*** *There exists a positive constant $C < \infty$ such that*

- *(a) Factor matrix $\boldsymbol{F}_t$ is of fixed dimension $k \times r$, and $\mathbb{E}\|\boldsymbol{F}_t\|^4 \le C$.*
- *(b) For all $i \in [p]$, $j \in [q]$ and $t \in [T]$, $\mathbb{E}\left[e_{t,ij}\right] = 0$ and $\mathbb{E}|e_{t,ij}|^8 \le C$.*
- *(c) Factor and noise are uncorrelated, that is, $\mathbb{E}\left[e_{t,ij} f_{s,lh}\right] = 0$ for any $t, s \in [T]$, $i \in [p]$, $j \in [q]$, $l \in [k]$, and $h \in [r]$.*

Assumption 3. ***Loading matrix.*** *For each row of $\boldsymbol{R}_t$, $\left\|\boldsymbol{R}_{t,i\cdot}\right\| = \mathcal{O}(1)$, and, as $p, q \to \infty$, we have $\left\|p^{-1}\boldsymbol{R}_t{}^\top \boldsymbol{R}_t - \boldsymbol{\Omega}_{R,t}\right\| \longrightarrow 0$ for some $k \times k$ positive definite matrix $\boldsymbol{\Omega}_{R,t}$. For each row of $\boldsymbol{C}_t$, $\left\|\boldsymbol{C}_{t,i\cdot}\right\| = \mathcal{O}(1)$, and, as $p, q \to \infty$, $\left\|q^{-1}\boldsymbol{C}_t{}^\top \boldsymbol{C}_t - \boldsymbol{\Omega}_{C,t}\right\| \longrightarrow 0$ for some $r \times r$ positive definite matrix $\boldsymbol{\Omega}_{C,t}$.*

Assumption 4. ***Cross row (column) correlation of noise $\boldsymbol{E}_t$.*** *There exists some positive constant $C < \infty$ such that,*

- *(a) Let $\boldsymbol{U}_E = \mathbb{E}\left[\boldsymbol{E}_t \boldsymbol{E}_t^T / q\right]$ and $\boldsymbol{V}_E = \mathbb{E}\left[\boldsymbol{E}_t^T \boldsymbol{E}_t / p\right]$, we assume $\|\boldsymbol{U}_E\|_1 \le C$ and $\|\boldsymbol{V}_E\|_1 \le C$.*
- *(b) For all row $i \in [p]$ and column $j \in [q]$ and $t \in [T]$, we assume $\sum_{\substack{l\in[p]\\ l\neq i}} \sum_{\substack{h\in[q]\\ h\neq j}} |\mathbb{E}\left[e_{t,ij} e_{t,lh}\right]| \le C$.*
- *(c) For any row $i, l \in [p]$, any time $t \in [T]$, and any column $j \in [q]$,*

$$\sum_{m\in[p]} \sum_{s\in[T]} \sum_{\substack{h\in[q]\\ h\neq j}} \left|\mathrm{Cov}\left[\left(e_{t,ij} e_{t,lj},\ e_{s,ih} e_{s,mh}\right)\right]\right| \le C$$

Similar, for any column $j,h \in [q]$, any time $t \in [T]$, and any row $i \in [p]$,

$$\sum_{m \in [q]} \sum_{s \in [T]} \sum_{\substack{l \in [p] \\ l \neq i}} \left|\mathrm{Cov}\left[\left(e_{t,ij} e_{t,ih},\ e_{s,lj} e_{s,lm}\right)\right]\right| \leq C$$

Assumption 5. *$\boldsymbol{R}_{i\cdot}(t/T)$ and $\boldsymbol{C}_{j\cdot}(t/T)$ have continuous derivatives up to the second order. Moreover, there exists $m > 2$, $1 < a, b < \infty$, $1/a + 1/b = 1$, $c = 0, 1, 2$ such that, for some positive $C < \infty$,*

(a) For any $l \in [k]$, $i \in [p]$, and $t \in [T]$, $\mathbb{E}\left[\left|\frac{1}{\sqrt{q}} \sum_{j=1}^{q} e_{t,ij}\right|^{mb}\right] = \mathcal{O}(1)$, $\mathbb{E}\left[\left\|\frac{1}{\sqrt{q}} \sum_{j=1}^{q} \boldsymbol{C}_{t,j\cdot}^{(c)} e_{t,ij}\right\|^{mb}\right] = \mathcal{O}(1)$, and $\mathbb{E}\left[\left\|\boldsymbol{f}_{t,l\cdot}\right\|^{ma}\right] \leq C$.

(b) For any $h \in [r]$, $j \in [q]$, and $t \in [T]$, $\mathbb{E}\left[\left|\frac{1}{\sqrt{p}} \sum_{i=1}^{p} e_{t,ij}\right|^{mb}\right] = \mathcal{O}(1)$, $\mathbb{E}\left[\left\|\frac{1}{\sqrt{p}} \sum_{i=1}^{p} \boldsymbol{R}_{t,i\cdot}^{(c)} e_{t,ij}\right\|^{mb}\right] = \mathcal{O}(1)$, and $\mathbb{E}\left[\left\|\boldsymbol{f}_{t,\cdot h}\right\|^{ma}\right] \leq C$.

(c) For any $t \in [T]$, $\mathbb{E}\left[\left|\frac{1}{\sqrt{pq}} \sum_{i=1}^{p} \sum_{j=1}^{q} e_{t,ij}\right|^{mb}\right] = \mathcal{O}(1)$ and $\mathbb{E}\left[\left\|\frac{1}{\sqrt{pq}} \sum_{i=1}^{p} \sum_{j=1}^{q} \boldsymbol{R}_{t,i\cdot}^{(c)} \boldsymbol{C}_{t,j\cdot}^{(c)\top} e_{t,ij}\right\|^{mb}\right] = \mathcal{O}(1)$.

Assumption 6. *$k : [-1,1] \to \mathbb{R}^{+}$ is a symmetric and Lipschitz continuous probability density function such that $\int_{-1}^{1} k(u)du = 1$, $\int_{-1}^{1} uk(u)du = 0$, and $\int_{-1}^{1} u^2 k(u)du < \infty$.*

The α-mixing condition in Assumption 1 allows weak temporal correlations for both the factors and noises. Assumption 2 mainly imposes moment conditions on the errors, factors, and their interactions. Assumption 3 is an extension of the pervasive assumption (Stock and Watson, 2002) to the matrix variate data. It ensures that each row and column of the factor matrix $\boldsymbol{F}_t$ has a nontrivial contribution to the variance of rows and columns of $\boldsymbol{Y}_t$. Assumptions 1-3 are standard in the literature on factor models and similar assumptions have been imposed in Su and Wang (2017), Chen and Fan (2021) and Yu et al. (2022).

Assumption 1 only deals with temporal dependence. The diverging matrix dimensions p and q also determine the convergence rates of our estimator, which is affected by the cross-row and cross-column dependence. Thus Assumptions 4 and 5 are imposed so that the information accumulated over rows (p) or columns (q) is also useful. Assumptions 4 and 5 are obviously satisfied if the errors $\boldsymbol{E}_t$ are i.i.d. over rows and columns for any t and are independent of the factor $\boldsymbol{F}_t$, with assumed moment conditions. We include them to allow for weakly cross-row (-column) and temporal dependence. Assumptions 6 is a standard assumption for kernel regressions. This condition is satisfied by commonly used second-order kernels, such as the Epanechnikov, uniform, and quartic kernels.

Let $\widehat{\boldsymbol{V}}_{R,t} \in \mathbb{R}^{k\times k}$ and $\widehat{\boldsymbol{V}}_{C,t} \in \mathbb{R}^{r\times r}$ be the time-varying diagonal matrices consisting of the first k and r largest eigenvalues of $\widehat{\boldsymbol{M}}_{R,t}$ ($\widehat{\boldsymbol{M}}_{C,t}$) in a descending order respectively. By the definition of our estimators $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$, we have

$$\begin{aligned} \widehat{\boldsymbol{R}}_t &= \widehat{\boldsymbol{M}}_{R,t}\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1} \\ \widehat{\boldsymbol{C}}_t &= \widehat{\boldsymbol{M}}_{C,t}\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{V}}_{C,t}^{-1}. \end{aligned}$$

As in the conventional factor models, the latent factor matrix $\boldsymbol{F}_t$ and loading matrices $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$ are not separately identifiable. However, they can be estimated up to an invertible matrix transformation. We can show that there exist invertible matrices $\boldsymbol{H}_{R,t}$ and $\boldsymbol{H}_{C,t}$ such that $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$ are consistent estimators of $\boldsymbol{R}_t\boldsymbol{H}_{R,t}$ and $\boldsymbol{R}_t\boldsymbol{H}_{R,t}$ respectively. Specifically, define

$$\begin{aligned} \boldsymbol{H}_{R,t} &= \frac{1}{Tpq}\sum_{s=1}^{T} K_{h,ts}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1} \\ \boldsymbol{H}_{C,t} &= \frac{1}{Tpq}\sum_{s=1}^{T} K_{h,ts}\boldsymbol{F}_s\boldsymbol{R}_t^\top\boldsymbol{R}_t\boldsymbol{F}_s^\top\boldsymbol{C}_t^\top\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{V}}_{C,t}^{-1}. \end{aligned} \tag{4.1}$$

Theorem 1. *Suppose Assumptions 1-6 are satisfied. As k,r fixed, $h \to 0$, $Th^3 \to \infty$, and $T,p,q \to \infty$, we have*

$$\begin{aligned} \frac{1}{p}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|_F^2 &= \mathcal{O}_p\left(\frac{1}{p^2} + \frac{1}{Tqh} + h^4\right), \\ \frac{1}{q}\left\|\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right\|_F^2 &= \mathcal{O}_p\left(\frac{1}{q^2} + \frac{1}{Tph} + h^4\right). \end{aligned}$$

Consequently,

$$\begin{aligned} \frac{1}{p}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}\boldsymbol{H}_{R,t}\right\|^2 &= \mathcal{O}_p\left(\frac{1}{p^2} + \frac{1}{Tqh} + h^4\right), \\ \frac{1}{q}\left\|\widehat{\boldsymbol{C}}_t - \boldsymbol{C}\boldsymbol{H}_{C,t}\right\|^2 &= \mathcal{O}_p\left(\frac{1}{q^2} + \frac{1}{Tph} + h^4\right). \end{aligned}$$

Theorem 1 establishes the convergence rate of our nonparametric estimators $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$. An alternative way to study the dynamics of matrix-valued data is to apply the local vector factor model (Su and Wang (2017)) to vectorized observations:

$$\text{VEC}(\boldsymbol{Y}_t) = (\boldsymbol{C}_t \otimes \boldsymbol{R}_t)\cdot\text{VEC}(\boldsymbol{F}_t) + \text{VEC}(\boldsymbol{E}_t), \tag{4.2}$$

where $\text{VEC}(\boldsymbol{Y}_t) \in \mathbb{R}^{pq}$ and $\text{VEC}(\boldsymbol{F}_t) \in \mathbb{R}^{kr}$. Applying results in Su and Wang (2017) without adopting the tensor structure in the loading matrix, we obtain $\frac{1}{pq}\left\|\widehat{\Xi}_t - \Xi_t \boldsymbol{H}_t\right\|^2 = \mathcal{O}_p\left(\frac{1}{p^2q^2} + \frac{1}{Th} + h^4\right)$, where $\Xi_t = \boldsymbol{C}_t \otimes \boldsymbol{R}_t$, and $\boldsymbol{H}_t \in \mathbb{R}^{kr\times kr}$ is an orthonormal matrix.[1]. Hence, in a high-dimensional setting where $p, q \geq T$, Theorem 1 establishes faster ℓ_2 convergence rate for both $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$. In addition, to recover the tensor structure in the factor loadings, one needs to carry out a second step to estimate $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$ from $\widehat{\Xi}_t$ which amounts to noisy Kronecker production decomposition. This may incur further errors (Cai et al., 2019). We shall further compare these two approaches via simulation in Section 6.

To study the limiting distribution of our nonparametric estimators $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$, we define

$$
\begin{aligned}
\Sigma_{FR,t}^{(c_1,c_2)} &= \boldsymbol{E}\left(\boldsymbol{F}_t^\top \frac{\boldsymbol{R}_t^{(c_1)\top}\boldsymbol{R}_t^{(c_2)}}{p}\boldsymbol{F}_t\right), \\
\Sigma_{FC,t}^{(c_1,c_2)} &= \boldsymbol{E}\left(\boldsymbol{F}_t \frac{\boldsymbol{C}_t^{(c_1)\top}\boldsymbol{C}_t^{(c_2)}}{q}\boldsymbol{F}_t^\top\right), \\
\Sigma_{R,t}^{(c)} &= \frac{\boldsymbol{R}_t^\top \boldsymbol{R}_t^{(c)}}{p}, \\
\Sigma_{C,t}^{(c)} &= \frac{\boldsymbol{C}_t^\top \boldsymbol{C}_t^{(c)}}{q},
\end{aligned}
\tag{4.3}
$$

where c_1, c_2 and $c = 0, 1, 2$ and $\boldsymbol{R}_t^{(1)}\left(\boldsymbol{C}_t^{(1)}\right)$ and $\boldsymbol{R}_t^{(2)}\left(\boldsymbol{C}_t^{(2)}\right)$ denote the first and second order derivative of $\boldsymbol{R}_t$ ($\boldsymbol{C}_t$) respectively, and $\boldsymbol{R}_t^{(0)} = \boldsymbol{R}_t$, $\boldsymbol{C}_t^{(0)} = \boldsymbol{C}_t$, $\Sigma_{FR,t}^{(0,0)} = \Sigma_{FR,t}$ and $\Sigma_{FC,t}^{(0,0)} = \Sigma_{FC,t}$

Assumption 7. *The matrices $\boldsymbol{\Omega}_{R,t}\Sigma_{FC,t}$ and $\boldsymbol{\Omega}_{C,t}\Sigma_{FR,t}$ are positive definite with distinct eigenvalues for each t.*

Theorem 2. *Suppose Assumptions 1-7 are satisfied. As k, r fixed, $h \to 0$, $Th^3 \to \infty$, and $T, p, q \to \infty$,*

(i) For row loading matrix $\boldsymbol{R}_t$, if $\sqrt{qTh}/p \to 0$, for each $i = 1, \cdots, p$ and $t = 1, \cdots, T$, we have

$$
\sqrt{qTh}\left(\widehat{\boldsymbol{R}}_{t,i.} - \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_{t,i.} - \boldsymbol{B}_{R_{t,i}}\right) \to^d N(0, \nu_0 \Sigma_{R_{t,i}}),
$$

where

$$
\begin{aligned}
\boldsymbol{B}_{R_{t,i}} &= \frac{1}{2}\boldsymbol{V}_{R,t}^{-1} h^2 \mu_2 \sum_{c_1+c_2+c_3+c_4=2} \boldsymbol{H}_{R,t}^\top \Sigma_{R,t}^{(c_1)} \Sigma_{FC,t}^{(c_2,c_3)} \boldsymbol{R}_{t,i.}^{(c_4)} + o_P\left(h^2\right), \\
\Sigma_{R_{t,i}} &= \boldsymbol{V}_{R,t}^{-1} \boldsymbol{Q}_{R,t} \boldsymbol{\Phi}_{R,t,i} \boldsymbol{Q}_{R,t}^\top \boldsymbol{V}_{R,t}^{-1}, \\
\boldsymbol{\Phi}_{R,t,i} &= \operatorname*{plim}_{q,T\to\infty} \frac{1}{qT}\sum_{r=1}^{T}\sum_{s=1}^{T} \mathbb{E}\left[\boldsymbol{F}_r \boldsymbol{C}_t^\top \boldsymbol{e}_{r,i.} \boldsymbol{e}_{s,i.}^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top\right],
\end{aligned}
\tag{4.4}
$$

[1]The bias terms are missing in Su and Wang (2017) but have been corrected in Su and Wang (2020).

$\nu_0 = \int_{-1}^{1} k^2(u)du$, $\mu_2 = \int_{-1}^{1} u^2 k(u)du$, $\boldsymbol{Q}_{R,t} = \boldsymbol{V}_{R,t}^{1/2}\boldsymbol{\Psi}_{R,t}^{\top}\Sigma_{FC,t}^{-1/2}$, $\boldsymbol{V}_{R,t}$ is a diagonal matrix whose entries are the eigenvalues of $\Sigma_{FC,t}$ in decreasing order, $\Psi_{R,t}$ is the corresponding eigenvector matrix such that $\Psi_{R,t}^{\top}\Psi_{R,t} = \boldsymbol{I}$, and $\Sigma_{FC,t}$ is defined in (4.3).

(ii) For column row loading matrix $\boldsymbol{C}_t$, if $\sqrt{pTh}/q \to 0$, for each $j = 1, \cdots, q$ and $t = 1, \cdots, T$, we have

$$\sqrt{pTh}\left(\widehat{\boldsymbol{C}}_{t,j.} - \boldsymbol{H}_{C,t}^{\top}\boldsymbol{C}_{t,j.} - \boldsymbol{B}_{C_{t,j}}\right) \to^d N(0, \nu_0\Sigma_{C_{t,j}}),$$

where

$$\boldsymbol{B}_{C_{t,j}} = \frac{1}{2}\boldsymbol{V}_{C,t}^{-1}h^2\mu_2 \sum_{c_1+c_2+c_3+c_4=2} \boldsymbol{H}_{C,t}^{\top}\Sigma_{C,t}^{(c_1)}\Sigma_{FR,t}^{(c_2,c_3)}\boldsymbol{C}_{t,j.}^{(c_4)} + o_P\left(h^2\right),$$

and

$$\Sigma_{C_{t,j}} = \boldsymbol{V}_{C,t}^{-1}\boldsymbol{Q}_{C,t}\boldsymbol{\Phi}_{C,t,j}\boldsymbol{Q}_{C,t}^{\top}\boldsymbol{V}_{C,t}^{-1}, \tag{4.5}$$

$\boldsymbol{\Phi}_{C,t,j}$ is defined similarly to $\boldsymbol{\Phi}_{R,t,i}$,$\boldsymbol{Q}_{C,t} = \boldsymbol{V}_{C,t}^{1/2}\Psi_{C,t}^{\top}\Sigma_{FR,t}^{-1/2}$, $\boldsymbol{V}_{C,t}$ is a diagonal matrix whose entries are the eigenvalues of $\Sigma_{FR,t}$ in decreasing order, $\Psi_{C,t}$ is the corresponding eigenvector matrix such that $\Psi_{C,t}^{\top}\Psi_{C,t} = \boldsymbol{I}$, and $\Sigma_{FR,t}$ is defined in (4.3).

Theorem 2 establishes the asymptotic normality of the nonparametric estimators $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{C}}_t$. The asymptotic bias $\boldsymbol{B}_{R_{t,i}}(\boldsymbol{B}_{C_{t,j}})$ is proportional to h^2 in both interior region and boundary regions, thanks to the use of the boundary-modified kernel function in (3.3). It is easy to verify that the pointwise asymptotic mean squared error (AMSE) of $\widehat{\boldsymbol{R}}_{t,i.}$ is

$$AMSE_R = \frac{h^4\mu_2^2}{4}\left\|\sum_{c_1+c_2+c_3+c_4=2} \boldsymbol{V}_{R,t}^{-1}\boldsymbol{H}_{R,t}^{\top}\Sigma_{R,t}^{(c_1)}\Sigma_{FC,t}^{(c_2,c_3)}\boldsymbol{R}_{t,i.}^{(c_4)}\right\|^2 + \frac{\mathrm{Tr}\left(\nu_0\Sigma_{R_{t,i}}\right)}{qTh}, \tag{4.6}$$

and hence the optimal bandwidth for $\widehat{\boldsymbol{R}}_{t,i.}$ is

$$h_R^{opt} = (qT)^{-\frac{1}{5}}\left(\frac{\mathrm{Tr}\left(\nu_0\Sigma_{R_{t,i}}\right)}{\mu_2^2\left\|\sum_{c_1+c_2+c_3+c_4=2} \boldsymbol{V}_{R,t}^{-1}\boldsymbol{H}_{R,t}^{\top}\Sigma_{R,t}^{(c_1)}\Sigma_{FC,t}^{(c_2,c_3)}\boldsymbol{R}_{t,i.}^{(c_4)}\right\|^2}\right)^{\frac{1}{5}}. \tag{4.7}$$

The optimal bandwidth for $\widehat{\boldsymbol{C}}_{t,i.}$ can be derived similarly.

Theorem 3. *Suppose Assumptions 1-7 are satisfied. As k, r fixed, $h \to 0$, $Th^3 \to \infty$, and*

$T, p, q \to \infty$, *we have*

$$\widehat{\boldsymbol{F}}_t - \boldsymbol{H}_{R,t}^{-1}\boldsymbol{F}_t\boldsymbol{H}_{C,t}^{-1} = \mathcal{O}_p\left(\frac{1}{\min(p,q)} + h^2\right),$$
$$\widehat{\boldsymbol{S}}_{t,ij} - \boldsymbol{S}_{t,ij} = \mathcal{O}_p\left(\frac{1}{\min\left(p, q, \sqrt{pTh}, \sqrt{qTh}\right)} + h^2\right),$$

for any $1 \le i \le p$ *and* $1 \le j \le q$.

Theorem 3 derives the convergence rate of the estimated latent factor $\widehat{\boldsymbol{F}}_t$ and signal $\widehat{\boldsymbol{S}}_t$. To achieve consistency, we require dimensions p and q approach infinity. That is because we need sufficient information accumulation to distinguish the signal $\boldsymbol{S}_t$ from the noise $\boldsymbol{E}_t$ at each time point t. Theorems 2 and 3 present the asymptotic properties when the dimension of the latent matrix factor $k \times r$ is assumed to be known. In practice we can estimate k and r via the generalized eigenvalue ratio-based estimator defined in (3.5). The asymptotic validity of $\widehat{k}$ and $\widehat{r}$ is justified in Theorem 4 below.

Theorem 4. *Suppose Assumptions 1-7 are satisfied and* k_{max} *is a predetermined constant no smaller than* k *or* r. *Then*

$$P\left(\widehat{k} = k\right) \to 1$$

and

$$P\left(\widehat{r} = r\right) \to 1$$

as $h \to 0$, $Th^3 \to \infty$, $T, p, q \to \infty$, *where* $\widehat{k}$ *and* $\widehat{r}$ *are defined in* (3.5).

Theorem 4 derives the consistency of the ratio-based estimator $\widehat{k}$ and $\widehat{r}$. This can be viewed as a generalization of Theorem 1 of Ahn and Horenstein (2013) from constant parameter factor models to time-varying matrix factor models.

5 Projection-based Estimation

Given $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$, the optimal $\boldsymbol{F}_t$ is achieved by $\widetilde{\boldsymbol{F}}_t = \boldsymbol{R}_t^\top \boldsymbol{Y}_t \boldsymbol{C}_t$. By concentrating out $\boldsymbol{F}_t$, we can reformulate the optimization problem (**??**) as:

$$\begin{aligned} \underset{\boldsymbol{R}_t, \boldsymbol{C}_t}{\text{maximize}} \quad & \sum_{s=1}^{T} \left\| \boldsymbol{R}_s \boldsymbol{R}_s^\top \boldsymbol{Y}_s^{(t)} \boldsymbol{C}_s \boldsymbol{C}_s^\top \right\|_F^2 \\ \text{subject to} \quad & \frac{1}{p}\boldsymbol{R}_t^\top \boldsymbol{R}_t = \boldsymbol{I}, \ \frac{1}{q}\boldsymbol{C}_t^\top \boldsymbol{C}_t = \boldsymbol{I}. \end{aligned} \tag{5.1}$$

Problem (5.1) has no closed form solutions. Ye (2005) solves the problem by an iterative algorithm. For a given $\widehat{\boldsymbol{C}}_t$, find the optimal $\boldsymbol{R}_t$ as the first k eigenvectors corresponding to the largest k eigenvalues of the matrix

$$\widehat{\boldsymbol{G}}_{R,t} \triangleq \frac{1}{Tpq^2} \sum_{s=1}^{T} \boldsymbol{Y}_s^{(t)} \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{Y}_s^{(t)\top}. \tag{5.2}$$

For a given $\widehat{\boldsymbol{R}}_t$, find the optimal $\boldsymbol{C}_t$ as the first r eigenvectors corresponding to the largest r eigenvalues of the matrix

$$\widehat{\boldsymbol{G}}_{C,t} \triangleq \frac{1}{Tp^2q} \sum_{s=1}^{T} \boldsymbol{Y}_s^{(t)\top} \widehat{\boldsymbol{R}}_t \widehat{\boldsymbol{R}}_t^\top \boldsymbol{Y}_s^{(t)}. \tag{5.3}$$

The projection-based optimal estimator $\widehat{\boldsymbol{R}}_t^p$ ($\widehat{\boldsymbol{C}}_t^p$) is comprised of $\sqrt{p}$ ($\sqrt{q}$) times top k (r) eigenvectors of $\widehat{\boldsymbol{G}}_{R,t}$ ($\widehat{\boldsymbol{G}}_{C,t}$) in descending order by corresponding eigenvalues. Estimation from local PCA can be used as an initial value for GLRAM. Ye (2005) proves the convergence of the algorithm. But the statistical properties of the estimators are unknown. Zhang and Xia (2018) proved statistical properties for different iterative algorithms for Tucker tensor decomposition. Their results are for homogeneous tensors where each mode can be treated equally, while in our setting, time mode is treated differently. Alternatively, we could solve (5.1) by direct optimization on the Stiefel manifold. However, our experiments show that GLRAM performs similarly to optimization on the Stiefel manifold. We next derive the asymptotic properties of the projection-based estimators $\widehat{\boldsymbol{R}}_t^p$ and $\widehat{\boldsymbol{C}}_t^p$.

Theorem 5. *Suppose Assumptions 1-6 are satisfied. As k, r fixed, $h \to 0$, $Th^3 \to \infty$, and $T, p, q \to \infty$, we have*

$$\frac{1}{p}\left\|\widehat{\boldsymbol{R}}_t^p - \boldsymbol{R}_t \boldsymbol{H}_{R,t}^p\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{p^2q^2} + \frac{1}{Tqh} + \frac{1}{T^2p^2h^2} + h^4\right),$$
$$\frac{1}{q}\left\|\widehat{\boldsymbol{C}}_t^p - \boldsymbol{C}_t \boldsymbol{H}_{C,t}^p\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{p^2q^2} + \frac{1}{Tph} + \frac{1}{T^2q^2h^2} + h^4\right),$$

where

$$\boldsymbol{H}_{R,t}^p = \tfrac{1}{Tpq^2} \textstyle\sum_{s=1}^T K_{h,ts} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top \widehat{\boldsymbol{R}}_t \widehat{\boldsymbol{V}}_{R,t}^{-1}$$
$$\boldsymbol{H}_{C,t}^p = \tfrac{1}{Tp^2q} \textstyle\sum_{s=1}^T K_{h,ts} \boldsymbol{F}_s \boldsymbol{R}_t^\top \widehat{\boldsymbol{R}}_t \widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t \boldsymbol{F}_s^\top \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{V}}_{C,t}^{-1}.$$

Theorem 5 shows that the projection-based estimator $\widehat{\boldsymbol{R}_t}^p$ ($\widehat{\boldsymbol{C}_t}^p$) indeed performs no worse than the local PCA estimator $\widehat{\boldsymbol{R}_t}$ ($\widehat{\boldsymbol{C}_t}$), while yields faster convergence rate than the local PCA estimator $\widehat{\boldsymbol{R}_t}$ ($\widehat{\boldsymbol{C}_t}$) when $p = o(Tqh)$ and $ph^4 = o(1)$ ($q = o(Tph)$ and $qh^4 = o(1)$). The improvement is mainly due to the increase in the signal-to-noise ratio as local eigenanalysis is now conducted on a lower-dimension matrix. We shall further compare the finite sample performance of $\widehat{\boldsymbol{R}_t}^p$ ($\widehat{\boldsymbol{C}_t}^p$) and $\widehat{\boldsymbol{R}_t}$ ($\widehat{\boldsymbol{C}_t}$) in Section 6.

6 Simulation

In this section, we study the numerical performance of the proposed time-varying matrix-valued approach. Throughout, the matrix observations $\boldsymbol{Y}_t$'s are generated according to model (2.1). The dimension of the latent factor matrix $\boldsymbol{F}_t$ is fixed at $k \times r = 2 \times 2$. The values of p, q, and T vary in different settings. Following Chen and Fan (2021), we simulate $\text{VEC}(\boldsymbol{F}_t)$ from the following Vector Auto-Regressive model of order one (VAR(1) model):

$$\text{VEC}(\boldsymbol{F}_t) = \boldsymbol{\Phi} \cdot \text{VEC}(\boldsymbol{F}_{t-1}) + \boldsymbol{\varepsilon}_t,$$

where the AR coefficient matrix $\boldsymbol{\Phi} = 0.1 \cdot \boldsymbol{I}_4$ and $\text{Var}[(\boldsymbol{\varepsilon}_t)] = 0.99 \cdot \boldsymbol{I}_4$. Thus, $\text{Var}[\text{VEC}(\boldsymbol{F}_t)] = \boldsymbol{I}_4$. We simulate noise $\boldsymbol{E}_t$ also from VAR(1),

$$\text{VEC}(\boldsymbol{E}_t) = \boldsymbol{\Gamma} \cdot \text{VEC}(\boldsymbol{E}_{t-1}) + \boldsymbol{u}_t,$$

where $\boldsymbol{\Gamma} = \psi \cdot \boldsymbol{I}_{pq}$ and $\text{Var}[\boldsymbol{u}_t] = (1-\psi^2)\boldsymbol{I}_{pq}$. Thus, $\text{Var}[\text{VEC}(\boldsymbol{E}_t)] = \boldsymbol{I}_{pq}$. We choose $\psi = 0.1$ and then increase to $\psi = 0.5$ to examine how temporal dependence may affect our results. The true loading matrices $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$ are generated as:

- *DGP1: The first column of $\boldsymbol{R}_t$: $\boldsymbol{R}_{t,.1} = \boldsymbol{R}_{.1}$, the second column of $\boldsymbol{R}_t$: $\boldsymbol{R}_{t,.2} = \boldsymbol{R}_{.2} + \boldsymbol{G}_t$; the first column of $\boldsymbol{C}_t$: $\boldsymbol{C}_{t,.1} = \boldsymbol{C}_{.1}$, the second column of $\boldsymbol{C}_t$: $\boldsymbol{C}_{t,.2} = \boldsymbol{C}_{.2} + \boldsymbol{H}_t$. The time-invariant loading matrices $\boldsymbol{R}_{.j}$ and $\boldsymbol{C}_{.j}$ are independently sampled from $\mathcal{U}(-1,1)$ for $j = 1, 2$, $\boldsymbol{G}_t = \boldsymbol{G}\left(\frac{t}{T}\right) = 2\left(\frac{t}{T}\right) + \exp\left[-16\left(\frac{t}{T} - 0.5\right)^2\right] - 1$ and $\boldsymbol{H}_t = 0.2\exp\left[-0.7 + 3.5(\frac{t}{T})\right]$, following Cai (2007).*

- *DGP2: The first column of $\boldsymbol{R}_t$: $\boldsymbol{R}_{t,.1} = \boldsymbol{R}_{.1} + \boldsymbol{G}_t$, the second column of $\boldsymbol{R}_t$: $R_{t,i2} = \boldsymbol{F}(10t/T; 2; 5i/p + 2)$, $i = 1, \ldots, p$; the first column of $\boldsymbol{C}_t$: $\boldsymbol{C}_{t,.1} = \boldsymbol{C}_{.1} + \boldsymbol{H}_t$, the second column of $\boldsymbol{C}_t$: $C_{t,j2} = \boldsymbol{F}(10t/T; 2; 5j/q + 2)$, $j = 1, \ldots, q$. The time-invariant loading matrices $\boldsymbol{R}_{.j}$ and $\boldsymbol{C}_{.j}$ are independently sampled from $\mathcal{N}(0,1)$ for $j = 1, 2$, $b = 2$, and $F(\tau; \kappa; \gamma) = \{1 + \exp[-\kappa(\tau - \gamma)]\}^{-1}$, following Su and Wang (2017).*

We first study the performance of our proposed approach on estimating the time-varying loading matrices $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$. We consider three pairs of (p,q) combinations: $(20,20)$, $(100,20)$, or $(100,100)$. The sample size T is selected as $100, 200$, and 400, and simulation results are based on 100 repetitions. For nonparametric estimation, we use the Epanechnikov kernel and Silverman's rule-of-thumb (RoT) bandwidth $h_R = c(qT)^{-\frac{1}{5}}$ and $h_C = c(pT)^{-\frac{1}{5}}$, where $c = \frac{2.345}{\sqrt{12}}$, for estimating $\boldsymbol{R}_t$ and $\boldsymbol{C}_t$ respectively. This simple RoT bandwidth attains the optimal rate for local smoothing as shown in Theorem 2. We also try the RoT bandwidth with different scaling parameters, and the simulation shows that our estimation results are not very sensitive to the bandwidth selection. To evaluate the accuracy of our estimator, we use the average column space distance

$$\overline{\mathcal{D}}_{(\widehat{\boldsymbol{A}},\boldsymbol{A})} = \frac{1}{T}\sum_{t=1}^{T}\left\|\widehat{\boldsymbol{A}}_t\left(\widehat{\boldsymbol{A}}_t^\top\widehat{\boldsymbol{A}}_t\right)^{-1}\widehat{\boldsymbol{A}}_t^\top - \boldsymbol{A}_t\left(\boldsymbol{A}_t^\top\boldsymbol{A}_t\right)^{-1}\boldsymbol{A}_t^\top\right\|, \tag{6.1}$$

for any rank k matrices $\widehat{\boldsymbol{A}}_t, \boldsymbol{A}_t \in \mathbb{R}^{p\times k}$, following Chen and Fan (2021).

Table 1: Estimation of loading matrices

DGP	(p,q)	$T=100$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{R}},\boldsymbol{R})}$	$T=100$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{C}},\boldsymbol{C})}$	$T=200$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{R}},\boldsymbol{R})}$	$T=200$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{C}},\boldsymbol{C})}$	$T=400$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{R}},\boldsymbol{R})}$	$T=400$: $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{C}},\boldsymbol{C})}$
1				$\psi=0.1$			
	$(20,20)$	1.09(0.21)	0.98(0.15)	0.98(0.12)	0.55(0.07)	0.58(0.08)	0.50(0.04)
	$(100,20)$	0.98(0.12)	0.55(0.07)	0.73(0.08)	0.40(0.05)	0.41(0.04)	0.35(0.02)
	$(100,100)$	0.58(0.08)	0.50(0.04)	0.52(0.04)	0.29(0.04)	0.29(0.02)	0.26(0.02)
				$\psi=0.5$			
	$(20,20)$	1.14(0.22)	1.04(0.17)	1.01(0.12)	0.58(0.08)	0.59(0.08)	0.52(0.04)
	$(100,20)$	1.01(0.12)	0.58(0.08)	0.75(0.08)	0.42(0.06)	0.42(0.04)	0.37(0.02)
	$(100,100)$	0.59(0.08)	0.52(0.04)	0.54(0.05)	0.30(0.04)	0.30(0.03)	0.27(0.02)
2				$\psi=0.1$			
	$(20,20)$	2.46(0.20)	2.15(0.20)	2.27(0.18)	1.94(0.16)	2.11(0.14)	1.79(0.16)
	$(100,20)$	2.35(0.20)	1.72(0.17)	2.14(0.14)	1.57(0.13)	1.95(0.11)	1.42(0.10)
	$(100,100)$	1.91(0.20)	1.47(0.13)	1.73(0.13)	1.29(0.08)	1.61(0.10)	1.15(0.09)
				$\psi=0.5$			
	$(20,20)$	2.48(0.20)	2.24(0.20)	2.30(0.18)	2.05(0.15)	2.13(0.14)	1.88(0.16)
	$(100,20)$	2.38(0.19)	1.80(0.17)	2.17(0.14)	1.64(0.13)	1.97(0.11)	1.50(0.11)
	$(100,100)$	1.93(0.19)	1.55(0.13)	1.75(0.13)	1.37(0.08)	1.62(0.10)	1.23(0.09)

Note: Mean and standard deviation in parentheses of $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{R}},\boldsymbol{R})}$ and $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{C}},\boldsymbol{C})}$ from 100 iterations.

Table 1 reports the mean and standard deviation of the average column space distance. For both DGPs considered, $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{R}},\boldsymbol{R})}$ and $\overline{\mathcal{D}}_{(\widehat{\boldsymbol{C}},\boldsymbol{C})}$ decrease with the increase of p,q and T. The estimation results are robust to the the temporal dependence of noise $\boldsymbol{E}_t$.

With the same simulated data, we compare the proposed time-varying matrix-valued approach and the time-varying vector-valued approach in Su and Wang (2017) through the estimation accuracy of the total loading matrix $\boldsymbol{\Xi}_t = \boldsymbol{C}_t \otimes \boldsymbol{R}_t$. In what follows, the subscripts mat and vec denote our approach and Su and Wang (2017)'s method, respectively. The loading space $\widehat{\boldsymbol{\Xi}}_{mat,t}$ is computed as $\widehat{\boldsymbol{\Xi}}_{mat,t} = \widehat{\boldsymbol{C}}_t \otimes \widehat{\boldsymbol{R}}_t$, where $\widehat{\boldsymbol{C}}_t$ and $\widehat{\boldsymbol{R}}_t$ are our nonparametric estimators. For the vector-valued approach, we apply Su and Wang (2017)'s method to the vectorized observations $\text{VEC}(\boldsymbol{Y}_t), t = 1, 2, \cdots, T$ as in (4.2) to obtain $\widehat{\boldsymbol{\Xi}}_{vec,t}$. The results for the estimation accuracies of $\boldsymbol{\Xi}_t$ measured by $\overline{\mathcal{D}}_{mat}$ and $\overline{\mathcal{D}}_{vec}$ are reported in Table 2, which shows that the matrix approach efficiently improves the estimation accuracy over the vector-valued approach. We only consider three pairs of small (p, q) combinations: $(10, 10)$, $(20, 10)$, or $(20, 20)$ because the estimation of time-varying vectorized factor models is extremely time-consuming. Nevertheless we expect the same pattern of estimation comparison would continue to hold for large p and q.

Table 2: Comparison of estimation accuracy

		$T = 100$		$T = 200$	
DGP	(p, q)	$\overline{\mathcal{D}}_{max}$	$\overline{\mathcal{D}}_{vec}$	$\overline{\mathcal{D}}_{max}$	$\overline{\mathcal{D}}_{vec}$
			$\psi = 0.1$		
	$(10, 10)$	0.23(0.06)	0.78(0.10)	0.13(0.02)	0.62(0.10)
	$(20, 10)$	0.17(0.04)	0.65(0.11)	0.15(0.02)	0.75(0.07)
1	$(20, 20)$	0.19(0.03)	0.76(0.09)	0.11(0.02)	0.60(0.08)
			$\psi = 0.5$		
	$(10, 10)$	0.25(0.07)	0.90(0.09)	0.14(0.03)	0.79(0.12)
	$(20, 10)$	0.18(0.05)	0.79(0.13)	0.16(0.03)	0.91(0.06)
	$(20, 20)$	0.20(0.04)	0.91(0.06)	0.12(0.02)	0.78(0.12)
			$\psi = 0.1$		
	$(10, 10)$	0.34(0.03)	0.66(0.11)	0.31(0.02)	0.58(0.12)
	$(20, 10)$	0.32(0.02)	0.64(0.10)	0.29(0.01)	0.54(0.08)
2	$(20, 20)$	0.29(0.02)	0.61(0.09)	0.27(0.02)	0.51(0.07)
			$\psi = 0.5$		
	$(10, 10)$	0.35(0.03)	0.70(0.12)	0.32(0.03)	0.61(0.13)
	$(20, 10)$	0.33(0.02)	0.69(0.11)	0.29(0.02)	0.58(0.09)
	$(20, 20)$	0.30(0.02)	0.65(0.10)	0.27(0.02)	0.55(0.09)

Note: Mean and standard deviation in parentheses of $\overline{\mathcal{D}}_{max}$ and $\overline{\mathcal{D}}_{vec}$ from 100 iterations.

We next evaluate the performance of the matrix-valued approach on estimating the number of factors k and r. Table 3 shows the relative frequencies of estimated rank pairs over 100 iterations. The three pairs $(2, 1), (1, 2)$ and $(2, 2)$ have high appearances in all of the combinations of (p, q, T). The row for the true rank pair $(2, 2)$ is highlighted. It shows that the relative frequency of correctly estimating the true rank pair increases with the increase of (p, q, T) and approaches unity.

Table 3: Estimation of latent dimensions

		$p,q = 20,20$			$p,q = 100,20$			$p,q = 100,100$		
DGP	$(\widehat{k},\widehat{r})$	$T=100$	$T=200$	$T=400$	$T=100$	$T=200$	$T=400$	$T=100$	$T=200$	$T=400$
					$\psi = 0.1$					
1	(2,1)	0.13	0.06	0.04	0.02	0.03	0.01	0.00	0.00	0.00
	(1,2)	0.05	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00
	(2,2)	0.80	0.90	0.93	0.98	0.97	0.99	1.00	1.00	1.00
	other	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					$\psi = 0.5$					
	(2,1)	0.12	0.08	0.03	0.02	0.02	0.00	0.00	0.00	0.00
	(1,2)	0.09	0.04	0.03	0.00	0.00	0.00	0.00	0.00	0.00
	(2,2)	0.77	0.87	0.94	0.98	0.98	1.00	1.00	1.00	1.00
	other	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					$\psi = 0.1$					
2	(2,1)	0.21	0.16	0.14	0.21	0.11	0.06	0.00	0.00	0.00
	(1,2)	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(2,2)	0.74	0.81	0.86	0.79	0.89	0.94	1.00	1.00	1.00
	other	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					$\psi = 0.5$					
	(2,1)	0.22	0.16	0.15	0.20	0.12	0.07	0.00	0.00	0.00
	(1,2)	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(2,2)	0.73	0.80	0.85	0.80	0.88	0.93	1.00	1.00	1.00
	other	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: Relative frequency of estimated rank pair $(\widehat{k},\widehat{r})$ from 100 iterations. The row with the true rank pair $(2,2)$ is highlighted.

7 Empirical Application

Understanding the pattern and evolution of international trade flow is essential for a broad range of economic activities including policy-making, economic forecast, and firm-level optimization. International trade data naturally forms a dynamic sequence of matrix variates and its structure intrinsically resembles that of a network. Those detailed data are readily available and provide a wide variety of information on the patterns of interactions, the evolution of the relative importance, and the natural grouping of actors in the network. Research on networks in international trade has made substantial progress theoretically and empirically (Chaney (2016)). Among them, the Global Vector Autoregressive (GVAR) approach has been quite popular in analyzing interactions in international trade networks (Chudik and Pesaran (2016)). For example, Bussiére et al. (2012) examines the effects of demand shocks and shocks to relative prices on global imbalances with a GVAR model of global trade flows. Greenwood-Nimmo et al. (2012) apply a GVAR model to forecast trade imbalances among a group of 33 countries. One undesirable feature of the GVAR approach is that a pre-specified weighting matrix has to be determined to handle the "curse-of-dimensionality" problem. In this study, we apply our time-varying matrix factor model to the international trade flow data, where time-varying factor loadings and latent factors are estimated simultaneously.

7.1 Data and sample

We use monthly multilateral import and export volumes of commodity goods among 24 countries and regions from January 1982 to December 2018. The data come from the International Monetary Fund Direction of Trade Statistics (IMF-DOTS). The countries and regions included in the study are Australia, Canada, China Mainland, Denmark, Finland, France, Germany, Hong Kong, Indonesia, Ireland, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, New Zealand, Singapore, Spain, Sweden, Taiwan, Thailand, United Kingdom, and the United States.

We apply our time-varying matrix factor model in (2.1) to the trade flow data. The observed Y_t is a 24×24 square matrix, whose rows (columns) represent export (import) countries. We use the import CIF (cost, insurance, and freight) data based on U.S. dollar prices as it is generally believed that they are more accurate than the export data (Linnemann (1966)). Note that the trade data for Taiwan as a reporting region is not published in the IMF-DOTS and import data for Taiwan are imputed from the export data reported by its partner countries in this study. To reduce the impact of incidental transactions of unusual size and incidental difficulties in trade contracts, trade flows were measured as three-month averages, rather than as direct observations of a particular month.

7.2 Estimation Result

Recall that model (2.1) identifies k latent factors, which can be interpreted as k export hubs and r import hubs in our application. To determine the number of dimensions of the latent hubs, we use the generalized ratio-based method in (3.5) and the screen plot, which leads us to select $(4, 4)$. As discussed in Section 4, the rotational indeterminacy implies that only the column spaces of the loading matrices are identified. However, this indeterminacy provides some flexibility for better economic interpretation. To achieve this, we apply the Varimax method (Kaiser (1958)) to the columns of the loading matrices after scaling them to unit length, which ensures their mutual orthogonality. We then further standardize the columns of the loading matrices so that they sum to 1s. Therefore, the columns of the estimated loading matrices represent the percentage of contribution that each country makes to the trading hub.

Figure 2 (3) plots the heat maps of the row export (column import) loading matrix R_t (C_t) from the year 1982 to 2018. The color intensity of the cells changes over time, suggesting the time-varying evolution in a country's participation in the four hubs. The variation is particularly pronounced during periods such as China's accession to the WTO and the financial crisis.

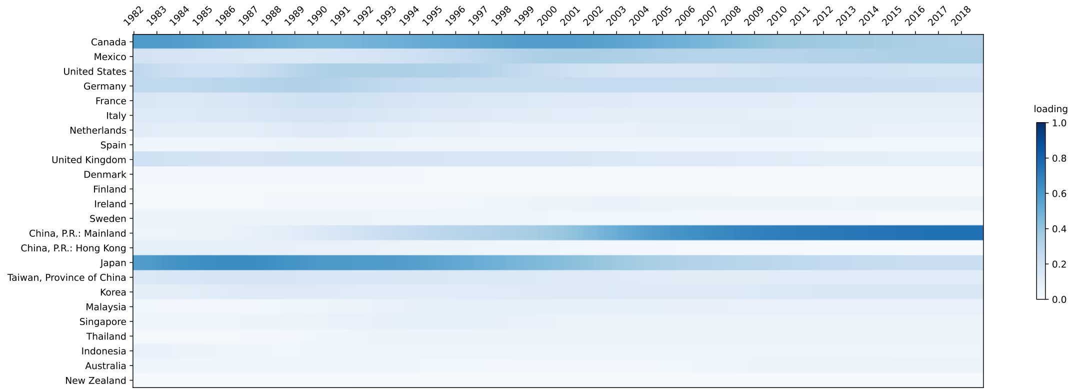
The latent hub shown in Figure 2(a) can be interpreted as a China-dominated export hub, given that China has the highest loading overall. However, prior to 2001, Japan had a higher loading than China. China's loading on this hub increased steadily throughout the sample period, particularly after its accession to the World Trade Organization (WTO) in 2001. This suggests a shift in the exporting centrality of the Asian economy, although other important participants in international trade include Japan, Hong Kong, and Taiwan. The latent hub corresponding to Figure 2(b) can be viewed as a USA-dominated export hub, as the USA's loading on this hub is much larger than that of other countries. Throughout the sample period, the United States dominated this export hub, although its contribution gradually decreased around 2008, likely due to the financial crisis. However, this decrease from the USA was offset by an increase from countries such as Germany, the Netherlands, Mainland China, Japan, Taiwan, Korea, and Australia. The latent hub corresponding to Figure 2(c) can be interpreted as a hub dominated by European countries, including Germany, France, Italy, the Netherlands, and Spain. Lastly, the composition of the latent hub shown in Figure 2(d) changed significantly over time. Before 1994, Germany dominated this hub, but after that, APEC countries became the dominant contributors, although NAFTA countries also significantly contributed to this hub.

Figure 3 shows the latent import hubs. The main difference between Figure 1 and Figure 2 is the order of the first two hub switches. The first import hub shown in Figure 3(a) can be interpreted as a USA-dominated hub, while the second import hub shown in Figure 3(b) can be viewed as a China-dominated hub. Figures 3(c) and 3(d) represent the latent hubs of European countries and APEC countries, respectively, which are similar to those in Figure 1. The time-varying patterns in Figure 3 are also comparable to the patterns in the heat maps shown in Figure 2.

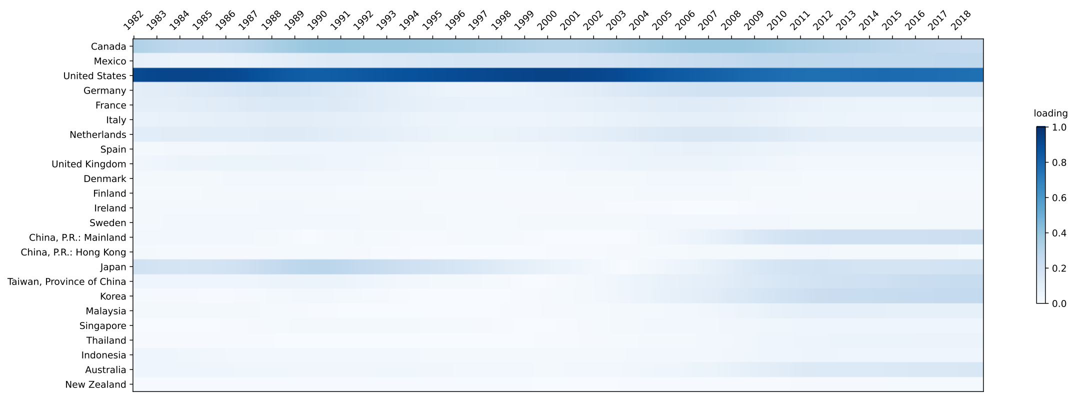
8 Conclusion

Modeling high-dimensional matrix-valued time series has received increasing attention recently, and in this paper, we have made contributions by establishing the asymptotic

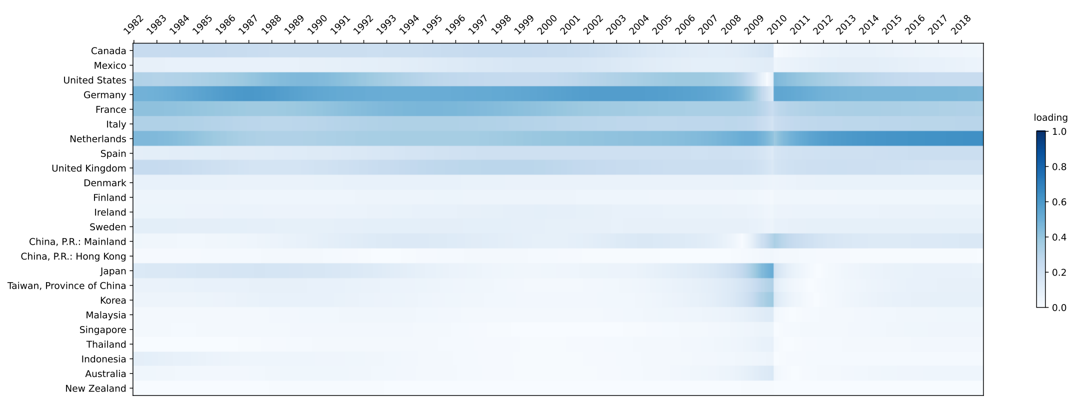
Figure 2: Latent export loadings for the trading volume.



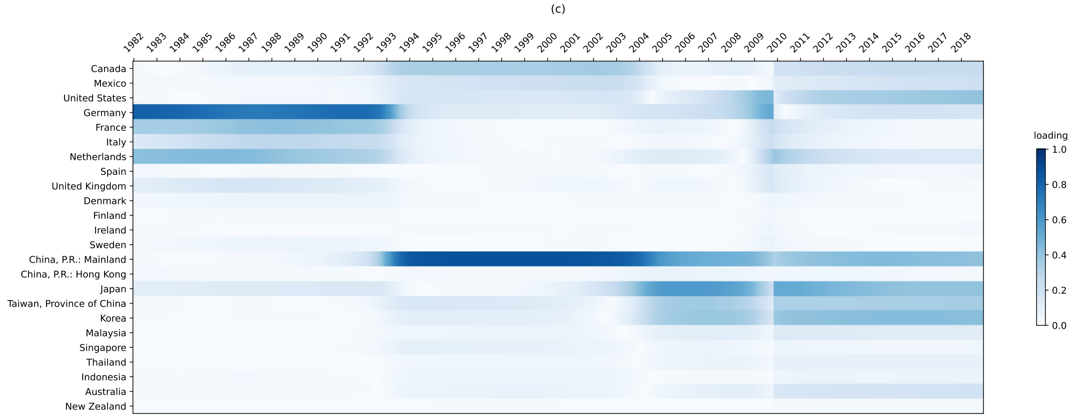
(a)



(b)



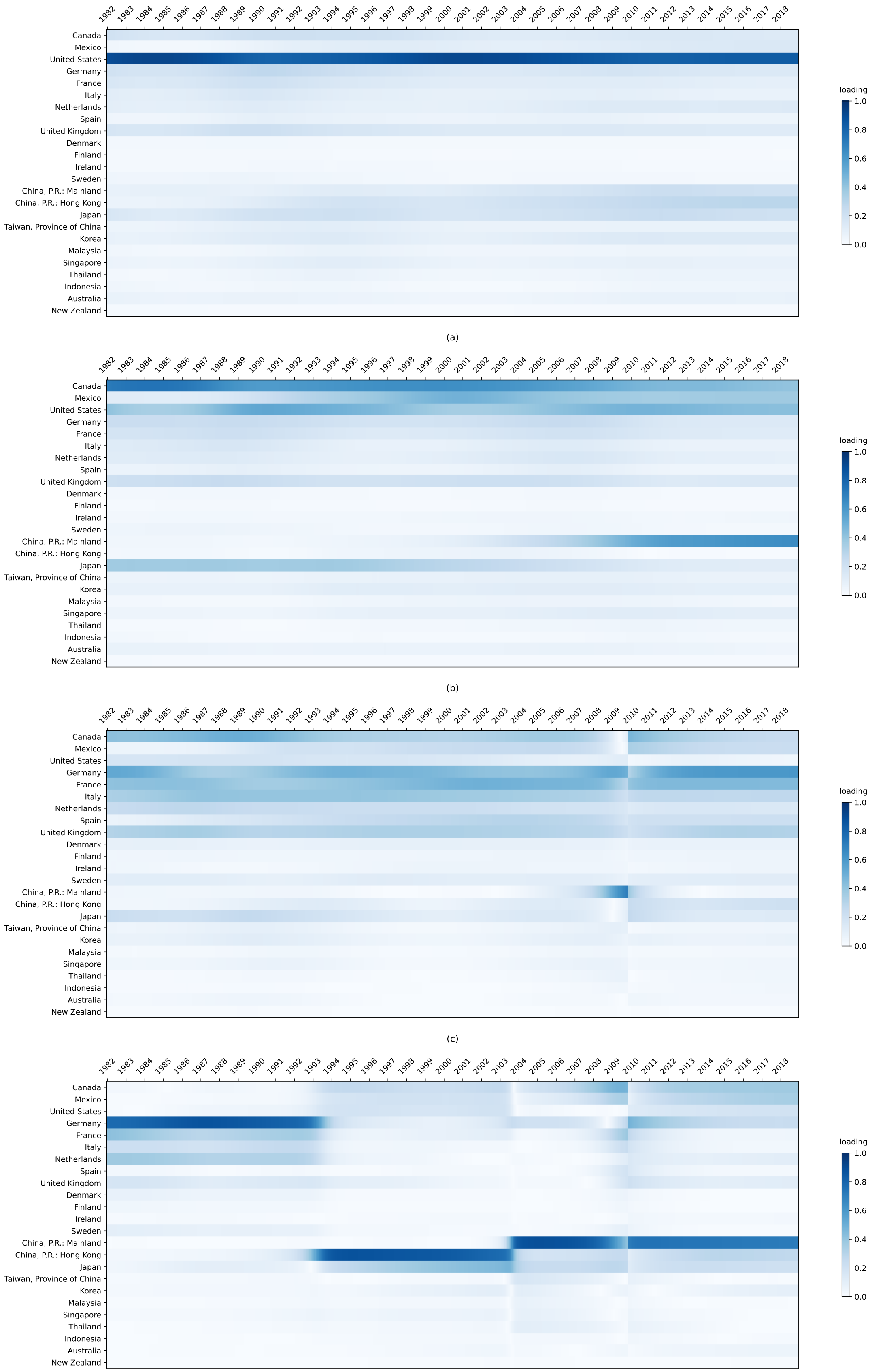
(c)



(d)

Notes: (1) sample period: January 1982- December 2018. (2) The generalized eigenvalue ratio-based estimator selects 4 export hubs.

Figure 3: Latent import loadings for the trading volume.



(a)

(b)

(c)

(d)

Notes: (1) sample period: January 1982- December 2018. (2) The generalized eigenvalue ratio-based estimator selects 4 import hubs.

properties of a local PCA estimator for time-varying matrix factor models and proposing a generalized eigenvalue-ratio estimator for latent dimensions. Our proposed estimators are not only intuitive and straightforward to compute, but they also fully preserve the matrix structure and explore the local variation of the loading matrix. Moreover, our results are obtained under very general conditions that allow for correlations across time, rows, and columns.

The simulations demonstrate that the nonparametric estimator for factor loadings and the generalized eigenvalue ratio-based estimator perform well in finite samples. Additionally, our empirical application concerning international trade flows reveals some interesting time-varying patterns. Our method identified four export hubs and four import hubs, which allowed us to analyze the changing contributions of countries to these hubs over time. Specifically, we found evidence of a transition in exporting centrality from Japan to China and the dominance of the USA and European countries in importing. Our results provide insight into the evolving dynamics of international trade.

References

Ahn, S. and A. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*(5), 1203–1227.

Airoldi, E. M., D. M. Blei, S. E. Fienberg, and E. P. Xing (2008). Mixed membership stochastic blockmodels. *Journal of Machine Learning Research 9*(9), 1981–2014.

Antognini, J. and J. Sohl-Dickstein (2018). Pca of high dimensional random walks with comparison to neural network training. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.), *Advances in Neural Information Processing Systems 31*, pp. 10328–10337. Curran Associates, Inc.

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

Bai, J., K. Li, et al. (2012). Statistical analysis of factor models of high dimension. *The Annals of Statistics 40*(1), 436–465.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bussiére, M., A. Chudik, and G. Sestieri (2012). Modelling global trade flows: results from a gvar model. *Globalization and Monetary Policy Institute Working Paper 119*.

Cai, C., R. Chen, and H. Xiao (2019). Kopa: Automated kronecker product approximation. *arXiv preprint arXiv:1912.02392*.

Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics 136*(1), 163–188.

Chaney, T. (2016). Networks in international trade. *The Oxford Handbook of the Economics of Networks*, 754–775.

Chang, J., B. Guo, and Q. Yao (2015). High dimensional stochastic regression with latent factors, endogeneity and nonlinearity. *Journal of Econometrics 189*(2), 297–312.

Chen, B. and Y. Hong (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica 80*(3), 1157–1183.

Chen, B. and K. Maung (2022). Time-varying forecast combination for high-dimensional data.

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

Chen, E. Y. and J. Fan (2021). Statistical inference for high-dimensional matrix-variate factor models. *Journal of the American Statistical Association*, 1–18.

Chen, E. Y., R. S. Tsay, and R. Chen (2019). Constrained factor models for high-dimensional matrix-variate time series. *Journal of the American Statistical Association*.

Chen, E. Y., Q. Yao, and R. Chen (2017). Multivariate spatial-temporal prediction on latent low-dimensional functional structure with non-stationarity. *arXiv preprint arXiv:1710.06351*.

Chen, R., D. Yang, and C.-H. Zhang (2019). Factor Models for High-Dimensional Tensor Time Series. *arXiv e-prints*.

Chudik, A. and M. H. Pesaran (2016). Theory and practice of gvar modelling. *Journal of Economic Surveys*, 165–197.

Fan, J., Y. Liao, and W. Wang (2016). Projected principal component analysis in factor models. *Annals of statistics 44*(1), 219.

Fan, J., K. Wang, Y. Zhong, and Z. Zhu (2018). Robust high dimensional factor models with applications to statistical machine learning. *arXiv e-prints*, arXiv:1808.03889.

Fan, J. and L. Yuan (2019). Learning latent factors from diversified projections and its applications to over-estimated and weak factors. *working paper*.

Fan, J. and Y. Zhong (2019). Optimal spectral aggregation via generalized method of moments. *working paper*.

Greenwood-Nimmo, M., V. Nguyen, and Y. Shin (2012). Probabilistic forecasting of output, growth, inflation and the balance of trade in a gvar framework. *Journal of Applied Econometrics*, 554–573.

Hong, Y. and H. Li (2005). Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. *The Review of Financial Studies 18*(1), 37–84.

Inoue, A., J. Lu, and B. Rossi (2017). Window selection for out-of-sample forecasting with time-varying parameters. *Journal of econometrics 196*(1), 55–67.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika 23*(3), 187–200.

Kolda, T. G. (2006). Multilinear operators for higher-order decompositions. Technical report, Sandia National Laboratories.

Kolda, T. G. and B. W. Bader (2009). Tensor decompositions and applications. *SIAM review 51*(3), 455–500.

Kong, H., L. Wang, E. K. Teoh, X. Li, J.-G. Wang, and R. Venkateswarlu (2005). Generalized 2d principal component analysis for face image representation and recognition. *Neural Networks 18*(5-6), 585–594.

Kwak, N. (2008). Principal component analysis based on l1-norm maximization. *IEEE transactions on pattern analysis and machine intelligence 30*(9), 1672–1680.

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

Lam, C., Q. Yao, and N. Bathia (2011). Estimation of latent factors for high-dimensional time series. *Biometrika 98*(4), 901–18.

Li, C., H. Farkhoor, R. Liu, and J. Yosinski (2018). Measuring the intrinsic dimension of objective landscapes. *arXiv preprint arXiv:1804.08838*.

Li, X., Y. Pang, and Y. Yuan (2010). L1-norm-based 2dpca. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 40*(4), 1170–1175.

Linnemann, H. (1966). An econometric study of international trade flows.

Liu, J., S. Chen, Z.-H. Zhou, and X. Tan (2010). Generalized low-rank approximations of matrices revisited. *IEEE Transactions on Neural Networks 21*(4), 621–632.

Liu, X. and E. Chen (2019). Helping effects against curse of dimensionality in threshold factor models for matrix time series. *arXiv e-prints*, arXiv:1904.07383.

Mao, X., P. Sarkar, and D. Chakrabarti (2018). Overlapping clustering models, and one (class) svm to bind them all. In *Advances in Neural Information Processing Systems*, pp. 2130–2140.

Meng, D., Q. Zhao, and Z. Xu (2012). Improve robustness of sparse pca by l1-norm maximization. *Pattern Recognition 45*(1), 487–497.

Nie, F., H. Zhang, R. Zhang, and X. Li (2018). Robust multiple rank-k bilinear projections for unsupervised learning. *IEEE Transactions on Image Processing*.

Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics 168*(2), 244–258.

Pang, Y., D. Tao, Y. Yuan, and X. Li (2008). Binary two-dimensional pca. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 38*(4), 1176–1180.

Peña, D. and P. Poncela (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference 136*(4), 1237–1257.

Phillips, P.C.B., H. B. (1990). Statistical inference in instrumental variables regression with i(1) processes. *Rev. Econ. Stud 57*, 99–125.

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In *Statistical Analysis and Forecasting of Economic Structural Change*, pp. 253–264. Springer.

Robinson, P. M. (1991). Time-varying nonlinear regression. In *Economic Structural Change Analysis and Forecasting*, pp. 179–190. Springer.

Sra, S. and I. S. Dhillon (2006). Generalized nonnegative matrix approximations with bregman divergences. In *Advances in neural information processing systems*, pp. 283–290.

Srebro, N. and T. Jaakkola (2003). Weighted low-rank approximations. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 720–727.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*(460), 1167–1179.

Su, L. and X. Wang (2017). On time-varying factor models: Estimation and testing. *Journal of econometrics 198*(1), 84–101.

Su, L. and X. Wang (2020). Corrigendum to "on time-varying factor models: Estimation and testing". *working paper*.

Uhlenbeck, G. E. and L. S. Ornstein (1930). On the theory of the brownian motion. *Physical Review 36*(5), 823.

Wang, D., X. Liu, and R. Chen (2019). Factor models for matrix-valued high-dimensional time series. *Journal of Econometrics 208*(1), 231–248.

Wang, R., F. Nie, X. Yang, F. Gao, and M. Yao (2015). Robust 2dpca with non-greedy ℓ_1-norm maximization for image analysis. *IEEE transactions on cybernetics 45*(5), 1108–1112.

Wang, W. and J. Fan (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. *Annals of statistics 45*(3), 1342.

Xing, E. P., W. Fu, and L. Song (2010). A state-space mixed membership blockmodel for dynamic network tomography. *The Annals of Applied Statistics 4*(2), 535–566.

Yang, J., D. D. Zhang, A. F. Frangi, and J.-y. Yang (2004). Two-dimensional pca: a new approach to appearance-based face representation and recognition. *IEEE transactions on pattern analysis and machine intelligence*.

Ye, J. (2005). Generalized low rank approximations of matrices. *Machine Learning 61*(1-3), 167–191.

Yu, L., Y. He, X. Kong, and X. Zhang (2022). Projected estimation for large-dimensional matrix factor models. *Journal of Econometrics*.

Zhang, A. and D. Xia (2018). Tensor svd: Statistical and computational limits. *IEEE Transactions on Information Theory 64*(11), 7311–7338.

Zhang, D. and Z.-H. Zhou (2005). (2d) 2pca: Two-directional two-dimensional pca for efficient face representation and recognition. *Neurocomputing 69*(1-3), 224–231.

Zhang, R., P. Robinson, and Q. Yao (2018). Identifying cointegration by eigenanalysis. *Journal of the American Statistical Association*, 1–12.

Zhang, T. and W. B. Wu (2012). Inference of time-varying regression models. *The Annals of Statistics 40*(3), 1376–1402.

Appendix

Appendix A Proofs

Proof of Theorem 1

Proof. Recall that our estimator $\widehat{\boldsymbol{R}}_t$ ($\widehat{\boldsymbol{C}}_t$) is given by the matrix of $\sqrt{p}$ ($\sqrt{q}$) times the top k (r) eigenvectors of $\widehat{\boldsymbol{M}}_{R,t}$ ($\widehat{\boldsymbol{M}}_{C,t}$) defined in equation (3.4) in descending order by corresponding eigenvalues. By the definition of eigenvectors and eigenvalues, we have

$$\frac{1}{pqT}\sum_{s=1}^{T}\boldsymbol{Y}_s^{(t)}\boldsymbol{Y}_s^{(t)\top}\widehat{\boldsymbol{R}}_t=\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t},\quad \text{or}\quad \widehat{\boldsymbol{R}}_t=\frac{1}{pqT}\sum_{s=1}^{T}\boldsymbol{Y}_s^{(t)}\boldsymbol{Y}_s^{(t)\top}\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1},$$

$$\frac{1}{pqT}\sum_{s=1}^{T}\boldsymbol{Y}_s^{(t)}\boldsymbol{Y}_s^{(t)\top}\widehat{\boldsymbol{C}}_t=\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{V}}_{C,t},\quad \text{or}\quad \widehat{\boldsymbol{C}}_t=\frac{1}{pqT}\sum_{s=1}^{T}\boldsymbol{Y}_s^{(t)\top}\boldsymbol{Y}_s^{(t)}\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{V}}_{C,t}^{-1}.$$

Then we have the following decomposition:

$$\begin{aligned}
\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}&=\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\mathbf{Y}_s\mathbf{Y}_s^\top\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1}-\boldsymbol{R}_t\boldsymbol{H}_t\\
&=\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\Big(\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top+\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{E}_s\boldsymbol{E}_s^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\\
&+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\\
&+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\\
&+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top\\
&+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top+\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{E}_s^\top\\
&+\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{D}_{C,st}^\top\boldsymbol{E}_s^\top+\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top+\boldsymbol{E}_s\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top+\boldsymbol{E}_s\boldsymbol{D}_{C,st}\boldsymbol{F}_s^\top\boldsymbol{D}_{R,st}^\top)\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1}\\
&\triangleq\sum_{j=1}^{24}\mathrm{I_j}\widehat{\boldsymbol{R}}_{\mathrm{t}}\widehat{\boldsymbol{V}}_{\mathrm{R,t}}^{-1},
\end{aligned}\tag{A.1}$$

where $\boldsymbol{D}_{R,st}=\boldsymbol{R}_s-\boldsymbol{R}_t$ and $\boldsymbol{D}_{C,st}=\boldsymbol{C}_s-\boldsymbol{C}_t$. Theorem 1 follows by Lemma 1 below. □

Lemma 1. *Under Assumption 1 - 6, we have for any* $1\le t\le T$,

(a) $\frac{1}{p}\left\|\mathrm{I}_1\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2=\mathcal{O}_p\left(\frac{1}{qTh}\right)$;

(b) $\frac{1}{p}\left\|\mathrm{I}_2\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2=\mathcal{O}_p\left(\frac{1}{qTh}\right)$;

(c) $\frac{1}{p}\left\|\mathrm{I}_3\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh}\right) + \mathcal{O}_p\left(\frac{1}{p}\right)$;

(d) $\frac{1}{p}\left\|\mathrm{I}_{\mathrm{j}}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 = \mathcal{O}_p\left(h^4\right), j = 4, \cdots, 13$;

(e) $\frac{1}{p}\left\|\mathrm{I}_{\mathrm{j}}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 = \mathrm{o}_p\left(h^4\right), j = 14, \cdots, 24$.

Proof of Lemma 1

Proof. (a) First note that

$$\begin{aligned}
\left\|\sum_{s=1}^{T} K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F^2 &= \left\|\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_{t,j\cdot}\boldsymbol{e}_{s,\cdot j}^\top\right\|_F^2 \\
&= qT\left\|\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}\left(\boldsymbol{e}_{s,\cdot j}\otimes\boldsymbol{F}_s\right)\boldsymbol{C}_{t,j\cdot}\right\|^2 \\
&= qT\sum_{i=1}^{p}\left\|\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}e_{s,ij}\boldsymbol{F}_s\boldsymbol{C}_{t,j\cdot}\right\|^2 \\
&= \mathcal{O}_p\left(pqTh^{-1}\right),
\end{aligned}$$

where we have used Lemma 11 and Markov inequality. Thus, we have

$$\frac{1}{p}\left\|\mathrm{I}_1\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 \leq \frac{1}{p^3q^2T^2}\|\boldsymbol{R}_t\|_F^2\left\|\sum_{s=1}^{T} K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh}\right).$$

(b) Similarly, we have

$$\frac{1}{p}\left\|\mathrm{I}_2\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 \leq \frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T} K_{h,st}\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\right\|_F^2\left\|\boldsymbol{R}_t^\top\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh}\right).$$

(c) First note that

$$\begin{aligned}
\left\|\sum_{s=1}^{T} K_{h,st} \boldsymbol{E}_s \boldsymbol{E}_s^\top\right\|_F^2 &= \left\|\sqrt{qT}\frac{\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}\boldsymbol{e}_{s,\cdot j}\boldsymbol{e}_{s,\cdot j}^\top}{\sqrt{qT}}\right\|_F^2 \\
&\leq p^2 qT\cdot\frac{1}{p^2}\sum_{i=1}^{p}\sum_{l=1}^{p}\left\|\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}\left(e_{s,ij}e_{s,lj}-\mathbb{E}\left[e_{s,ij}e_{s,lj}\right]\right)\right\|^2 \\
&\quad + pq^2T^2\cdot\frac{1}{p}\sum_{i=1}^{p}\left(\sum_{l=1}^{p}\left\|\mathbb{E}\left[\frac{\sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st}e_{s,ij}e_{s,lj}}{qT}\right]\right\|^2\right) \\
&= \mathcal{O}_p\left(p^2qTh^{-1}\right)+\mathcal{O}\left(pq^2T^2\right),
\end{aligned}$$

where we have used Lemma 13 (a). It follows that

$$\frac{1}{p}\left\|\mathrm{I}_3\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 \leq \frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh}\right)+\mathcal{O}_p\left(\frac{1}{p}\right).$$

(d) The proofs for $\frac{1}{p}\left\|\mathrm{I}_{\mathrm{j}}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 = \mathcal{O}_p\left(h^4\right)$, $j=4,\cdots,13$, are tedious. To save space, we only provide the proof for $j=4$. Other proofs are rather similar. Define $M_{R,s}=\boldsymbol{R}_s-\boldsymbol{R}_t-\left(\frac{s-t}{T}\right)\boldsymbol{R}_t^{(1)}-\frac{1}{2}\left(\frac{s-t}{T}\right)^2\boldsymbol{R}_t^{(2)}$.

$$\begin{aligned}
\frac{1}{p}\left\|\mathrm{I}_4\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 \leq& \frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\boldsymbol{R}_t^{(1)}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2 \\
&+\frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2K_{st}\boldsymbol{R}_t^{(2)}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2 \\
&+\frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}M_{R,s}K_{st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2 \\
\triangleq& \frac{1}{p^3q^2T^2}\sum_{j=1}^{3}\left\|\mathrm{I}_{4\mathrm{j}}\right\|_{\boldsymbol{F}}^2. \qquad (A.2)
\end{aligned}$$

For the first term in (A.2), we have

$$\begin{aligned}
\|\mathrm{I}_{41}\|_F^2 &\leq \left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\boldsymbol{R}_t^{(1)}\left(\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t\right\|_F^2+\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\boldsymbol{R}_t^{(1)}\Sigma_{FC,t}\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t\right\|_F^2\\
&\leq \left\|\boldsymbol{R}_t^{(1)}\right\|_F^2\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\left(\boldsymbol{F}_sC_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right\|_F^2\left\|\boldsymbol{R}_t^{\top}\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2\\
&+\left\|\boldsymbol{R}_t^{(1)}\right\|_F^2\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\right\|_F^2\left\|\Sigma_{FC,t}\right\|_F^2\left\|\boldsymbol{R}_t^{\top}\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2\\
&\triangleq \mathrm{I}_{411}+\mathrm{I}_{412}. \qquad (\mathrm{A.3})
\end{aligned}$$

Without loss of generality we assume $k=r=1$. For I_{411}, we have

$$\begin{aligned}
&E\left[\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\left(\boldsymbol{F}_sC_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right]^2\\
&=\frac{1}{T^2}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^2K_{st}^2E\left(\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)^2\\
&+\frac{1}{T^2}\sum_{s_1=1}^{T}\sum_{s_2=1}^{T}\left(\frac{s_1-t}{T}\right)\left(\frac{s_2-t}{T}\right)K_{s_1t}K_{s_2t}E\left(\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\left(\boldsymbol{F}_{s_2}\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_{s_2}^{\top}/q-\Sigma_{FC,t}\right)\\
&=T^{-1}h\int u^2k^2(u)\,du\,E\left(\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)^2\\
&+T^{-1}h\sum_{j=1}^{T}\int(u+\frac{j}{Th})uk(u)k\left(u+\frac{j}{Th}\right)du\,\alpha^{1-\frac{2}{\gamma}}(j)\left[E\left(\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)^{\gamma}\right]^{1/\gamma}+T^{-1}h\\
&=\mathcal{O}\left(T^{-1}h\right),
\end{aligned}$$

and hence $\mathrm{I}_{411}=\mathcal{O}\left(\mathrm{p}^3\mathrm{q}^2\mathrm{Th}\right)$. For I_{412}, we have

$$\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}=h\int uk(u)\,du+\mathcal{O}\left(\frac{1}{T}\right)=\mathcal{O}\left(\frac{1}{T}\right),$$

and hence $\mathrm{I}_{412}=\mathcal{O}\left(\mathrm{p}^3\mathrm{q}^2\right)$, where we have used the uniform approximation of Riemann summation to a definite integral. Therefore, $\frac{1}{p^3q^2T^2}\|\mathrm{I}_{41}\|_{\boldsymbol{F}}^2=\mathcal{O}_p\left(\frac{h}{T}\right)=\mathrm{o}_p\left(h^4\right)$

as $Th^3 \to \infty$. For the second term in (A.2), we have

$$\begin{aligned}
\|\mathrm{I}_{42}\|_F^2 &\le \left\| \sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} \boldsymbol{R}_t^{(2)} \left(\boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right) \boldsymbol{R}_t^\top \widehat{\boldsymbol{R}}_t \right\|_F^2 + \left\| \sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} \boldsymbol{R}_t^{(2)} \Sigma_{FC,t} \boldsymbol{R}_t^\top \widehat{\boldsymbol{R}}_t \right\|_F^2 \\
&\le \left\| \boldsymbol{R}_t^{(2)} \right\|_F^2 \left\| \sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} \left(\boldsymbol{F}_s C_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right) \right\|_F^2 \left\| \boldsymbol{R}_t^\top \right\|_F^2 \left\| \widehat{\boldsymbol{R}}_t \right\|_F^2 \\
&+ \left\| \boldsymbol{R}_t^{(2)} \right\|_F^2 \left\| \sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} \right\|_F^2 \left\| \Sigma_{FC,t} \right\|_F^2 \left\| \boldsymbol{R}_t^\top \right\|_F^2 \left\| \widehat{\boldsymbol{R}}_t \right\|_F^2 \\
&\triangleq \mathrm{I}_{421} + \mathrm{I}_{422}.
\end{aligned} \tag{A.4}$$

Similar to I_{411}, we have

$$\begin{aligned}
&E\left[\frac{1}{T}\sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} \left(\boldsymbol{F}_s C_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right)\right]^2 \\
&= \frac{1}{4T^2} \sum_{s=1}^{T} \left(\frac{s-t}{T}\right)^4 K_{st}^2 E\left(\boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right)^2 \\
&+ \frac{1}{4T^2} \sum_{s_1=1}^{T} \sum_{s_2=1}^{T} \left(\frac{s_1-t}{T}\right)^2 \left(\frac{s_2-t}{T}\right)^2 K_{s_1t} K_{s_2t} E\left(\boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right)\left(\boldsymbol{F}_{s_2} \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_{s_2}^\top / q - \Sigma_{FC,t}\right) \\
&= \frac{h^3}{4T} \int u^4 k^2(u)\, du E\left(\boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right)^2 \\
&+ \frac{h^3}{4T} \sum_{j=1}^{T} \int (u + \frac{j}{Th})^2 u^2 k(u) k\left(u + \frac{j}{Th}\right) du\, \alpha^{1-\frac{2}{\gamma}}(j) \left[E\left(\boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top / q - \Sigma_{FC,t}\right)^\gamma\right]^{1/\gamma} + T^{-1}h \\
&= \mathcal{O}\left(T^{-1}h^3\right),
\end{aligned}$$

and hence $\mathrm{I}_{421} = \mathcal{O}\left(\mathrm{p}^3\mathrm{q}^2\mathrm{Th}^3\right)$. For I_{422}, we have

$$\frac{1}{T}\sum_{s=1}^{T} \frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{st} = h^2 \int u^2 k(u)\, du + \mathcal{O}\left(\frac{h}{T}\right) = \mathcal{O}\left(h^2\right),$$

and hence $\mathrm{I}_{422} = \mathcal{O}\left(\mathrm{p}^3\mathrm{q}^2\mathrm{T}^2\mathrm{h}^4\right)$, where we have used the uniform approximation of Riemann summation to a definite integral. Therefore, $\frac{1}{p^3q^2T^2}\|\mathrm{I}_{42}\|_{\boldsymbol{F}}^2 = \mathcal{O}_p\left(h^4\right)$ as $Th^3 \to \infty$. Finally, by the fact that $p^{-1}\left\|h^{-2}M_{R,s}\right\|_{\boldsymbol{F}}^2 = p^{-1}\left\|h^{-2}M_R\left(\frac{t}{T} + uh\right)\right\|_{\boldsymbol{F}}^2 \to 0$ for any u, we have $\frac{1}{p^3q^2T^2}\|\mathrm{I}_{43}\|_{\boldsymbol{F}}^2 = \mathrm{o}_p\left(h^4\right)$. Combining these three terms, we have

$\frac{1}{p}\left\|\mathrm{I}_4\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2=\mathcal{O}_p\left(h^4\right)$.

(e) First, by the Taylor expansion, we have

$$\begin{aligned}\frac{1}{p}\|\mathrm{I}_{14}\|_F^2 \le& \frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^3 K_{st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(1)\top}\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2\\ &+\frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(2)\top}\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2\\ &+\frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(2)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(1)\top}\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2\\ &+\frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^{(2)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(1)\top}\widehat{\boldsymbol{R}}_t\right\|_{\boldsymbol{F}}^2+\mathrm{o}_p\left(h^8\right)\\ =&\,\mathcal{O}_p\left(h^5T^{-1}\right)+\mathcal{O}_p\left(h^8\right)+\mathcal{O}_p\left(h^8\right)+\mathcal{O}_p\left(h^8\right)+\mathrm{o}_p\left(h^8\right)\\ =&\,\mathrm{o}_p\left(h^4\right).\end{aligned}$$

The order of $\frac{1}{p}\left\|\mathrm{I}_{\mathrm{j}}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2$, $j=15,16,17,18$ can be derived in a similar way. Next we consider I_{19}. By the Taylor expansion, we have

$$\begin{aligned}\left\|\sum_{s=1}^{T}K_{h,st}\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F^2\le&\, qT\left\|\boldsymbol{R}_t^{(1)}\right\|_F^2\left\|\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}\left(\frac{s-t}{T}\right)K_{h,st}\left(\boldsymbol{e}_{s,\cdot j}\otimes\boldsymbol{F}_s\right)\boldsymbol{C}_{t,j\cdot}\right\|^2\\ &+qT\left\|\boldsymbol{R}_t^{(2)}\right\|_F^2\left\|\frac{1}{2\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}\left(\frac{s-t}{T}\right)^2K_{h,st}\left(\boldsymbol{e}_{s,\cdot j}\otimes\boldsymbol{F}_s\right)\boldsymbol{C}_{t,j\cdot}\right\|^2\\ &+qT\left\|\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}\boldsymbol{M}_{R,s}K_{h,st}\left(\boldsymbol{e}_{s,\cdot j}\otimes\boldsymbol{F}_s\right)\boldsymbol{C}_{t,j\cdot}\right\|^2\\ =&\,\mathcal{O}_p\left(p^2qTh\right)+\mathcal{O}_p\left(p^2qTh^3\right)+\mathrm{o}_p\left(p^2qTh^3\right),\end{aligned}$$

where we have used Lemma 11 and Markov inequality. Thus, we have

$$\frac{1}{p}\left\|\mathrm{I}_{19}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2 \le \frac{1}{p^3q^2T^2}\left\|\sum_{s=1}^{T}K_{h,st}\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F^2\left\|\widehat{\boldsymbol{R}}_t\right\|_F^2=\mathcal{O}_p\left(\frac{h}{qT}\right)=\mathrm{o}_p\left(h^4\right).$$

The order of $\frac{1}{p}\left\|\mathrm{I}_{\mathrm{j}}\widehat{\boldsymbol{R}}_{\mathrm{t}}\right\|_F^2$, $j=20,21,22,23,24$ can be derived in a similar way as I_{19}.

$\square$

Before proving Theorem 2, we first study the asymptotic behavior of $\widehat{\boldsymbol{V}}_{R,t}$, $\widehat{\boldsymbol{V}}_{C,t}$, $\boldsymbol{H}_{C,t}$ and $\boldsymbol{H}_{R,t}$.

Proposition 1. *Under Assumption 1-6, we have, as $p,q,T \to \infty$:*

$$\begin{aligned} \widehat{\boldsymbol{V}}_{R,t} &\xrightarrow{\mathcal{P}} \boldsymbol{V}_{R,t} \\ \widehat{\boldsymbol{V}}_{C,t} &\xrightarrow{\mathcal{P}} \boldsymbol{V}_{C,t}, \end{aligned}$$

$\left\|\widehat{\boldsymbol{V}}_{R,t}\right\|_2 = \mathcal{O}_p(1)$ and $\left\|\widehat{\boldsymbol{V}}_{R,t}^{-1}\right\|_2 = \mathcal{O}_p(1)$, where $\boldsymbol{V}_{R,t}$ is the diagonal matrix consisting of the eigenvalues of $\Sigma_{FC,t}^{1/2}\,\boldsymbol{\Omega}_{R,t}\,\Sigma_{FC,t}^{1/2}$ and $\boldsymbol{V}_{C,t}$ is the diagonal matrix consisting of the eigenvalues of $\Sigma_{FR}^{1/2}\,\boldsymbol{\Omega}_{C,t}\,\Sigma_{FR}^{1/2}$. Covariance $\Sigma_{FC,t} = \mathbb{E}\left[\boldsymbol{F}_t \frac{\boldsymbol{C}_t{}^\top \boldsymbol{C}_t}{q} \boldsymbol{F}_t^\top\right]$, $\Sigma_{FR,t} = \mathbb{E}\left[\boldsymbol{F}_t^\top \frac{\boldsymbol{R}_t{}^\top \boldsymbol{R}_t}{p} \boldsymbol{F}_t\right]$. Matrices $\boldsymbol{\Omega}_{R,t}$ and $\boldsymbol{\Omega}_{C,t}$ are defined in Assumption 3.

Proof. From $\frac{1}{pqT}\sum_{s=1}^T K_{t,s}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top\widehat{\boldsymbol{R}}_t = \widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}$ and $\frac{1}{p}\widehat{\boldsymbol{R}}_t^\top\widehat{\boldsymbol{R}}_t = \boldsymbol{I}_k$, we have

$$\widehat{\boldsymbol{V}}_{R,t} = \frac{1}{p}\widehat{\boldsymbol{R}}_t^\top\left(\frac{1}{pqT}\sum_{s=1}^T K_{t,s}\boldsymbol{Y}_s^\top\boldsymbol{Y}_s\right)\widehat{\boldsymbol{R}}_t$$

$\widehat{\boldsymbol{V}}_{R,t}$ is the $k\times k$ diagonal matrix of the first k largest eigenvalues of $\widehat{\boldsymbol{M}}_{R,t} \triangleq \frac{1}{pqT}\sum_{s=1}^T K_{t,s}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top$ in decreasing order. By definition of $\widehat{\boldsymbol{M}}_{R,t}$, we have

$$\widehat{\boldsymbol{M}}_{R,t} = \frac{1}{pqT}\sum_{s=1}^T K_{h,st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top + \sum_{j=1}^{24}\mathrm{I_j},$$

where $\mathrm{I_j}$, $j = 1,2,\ldots,24$ are defined in equation A.1. Applying Lemma 1, we have

$$\begin{aligned}
\|\mathrm{I}_1\| &\le \frac{1}{pqT}\|\boldsymbol{R}_t\|\left\|\sum_{s=1}^T K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F = \mathcal{O}_p\left(\frac{1}{\sqrt{qTh}}\right), \\
\|\mathrm{I}_2\| &\le \frac{1}{pqT}\|\boldsymbol{R}_t\|\left\|\sum_{s=1}^T K_{h,st}\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\right\|_F = \mathcal{O}_p\left(\frac{1}{\sqrt{qTh}}\right), \\
\|\mathrm{I}_3\| &\le \frac{1}{pqT}\left\|\sum_{s=1}^T K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right\|_F = \mathcal{O}_p\left(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{qTh}}\right), \\
\|\mathrm{I}_4\| &\le \frac{1}{pqT}\left\|\sum_{s=1}^T K_{h,st}\boldsymbol{D}_{R,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\right\|\|\boldsymbol{R}_t\| = \mathcal{O}_p\left(h^4\right).
\end{aligned} \tag{A.5}$$

And similarly, $\left\|\mathrm{I_j}\right\| = \mathcal{O}_p\left(h^4\right)$, for $j = 5, \ldots, 13$, and $\left\|\mathrm{I_j}\right\| = \mathrm{o}_p\left(h^4\right)$, for $j = 14, \ldots, 24$. On the other hand, we have

$$
\begin{aligned}
&\left\|\frac{1}{pqT}\boldsymbol{R}_t\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top - \mathbb{E}\left[\frac{1}{pqT}\boldsymbol{R}_t\sum_{s=1}^{T}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\right]\right\| \\
&\leq \left\|\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\left(\boldsymbol{F}_s(\boldsymbol{C}_t^\top\boldsymbol{C}_t/q)\boldsymbol{F}_s^\top - \mathbb{E}\left[\boldsymbol{F}_s(\boldsymbol{C}_t^\top\boldsymbol{C}_t/q)\boldsymbol{F}_s^\top\right]\right)\right\| \cdot \|\boldsymbol{R}_t\|^2/p \\
&= \mathcal{O}_p\left(\frac{1}{\sqrt{Th}}\right).
\end{aligned}
$$

Together, we have

$$
\left\|\widehat{\boldsymbol{M}}_{R,t} - \mathbb{E}\left[\frac{1}{pqT}\boldsymbol{R}_t\sum_{s=1}^{T}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\right]\right\| = \mathrm{o}_p(1).
$$

Using the inequality that for the i-th eigenvalue, $|\lambda_i\left(\widehat{\boldsymbol{A}}\right) - \lambda_i(\boldsymbol{A})| \leq \left\|\widehat{\boldsymbol{A}} - \boldsymbol{A}\right\|_2$, we have

$$
|\widehat{V}_{R,t,i} - V_{R,t,i}| = \mathrm{o}_p(1), \quad \text{for } 1 \leq i \leq k,
$$

and $\widehat{\boldsymbol{V}}_{R,t} \xrightarrow{\mathcal{P}} \boldsymbol{V}_{R,t}$. Further we have the first k eigenvalues of $\frac{1}{pqT}\boldsymbol{R}_t\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top$ are bounded away from both zero and infinity. Thus, $\left\|\widehat{\boldsymbol{V}}_{R,t}\right\|_2 = \mathcal{O}_p(1)$ and $\left\|\widehat{\boldsymbol{V}}_{R,t}^{-1}\right\|_2 = \mathcal{O}_p(1)$. Results for $\widehat{\boldsymbol{V}}_{C,t}$ are obtained in a similar fashion. □

Proposition 2. *Under Assumption 1-6, we have, for any* $1 \leq t \leq T$

$$
\left\|\boldsymbol{H}_{R,t}\right\| = \mathcal{O}_p(1), \quad \textit{and} \quad \left\|\boldsymbol{H}_{C,t}\right\| = \mathcal{O}_p(1).
$$

Proof. Applying results from Proposition 1 and Lemma 1, we obtain

$$
\begin{aligned}
\left\|\boldsymbol{H}_{R,t}\right\| &= \left\|\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\widehat{\boldsymbol{R}}_t\widehat{\boldsymbol{V}}_{R,t}^{-1}\right\| \leq \left\|\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s(\boldsymbol{C}_t^\top\boldsymbol{C}_t/q)\boldsymbol{F}_s^\top\right\|\|\boldsymbol{R}_t\|\left\|\widehat{\boldsymbol{R}}_t\right\|/p\left\|\widehat{\boldsymbol{V}}_{R,t}^{-1}\right\| = \mathcal{O}_p(1), \\
\left\|\boldsymbol{H}_{C,t}\right\| &= \left\|\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\widehat{\boldsymbol{C}}_t\widehat{\boldsymbol{V}}_{C,t}^{-1}\right\| \leq \left\|\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s^\top(\boldsymbol{R}_t^\top\boldsymbol{R}_t/p)\boldsymbol{F}_s\right\|\|\boldsymbol{C}_t\|\left\|\widehat{\boldsymbol{C}}_t\right\|/q\left\|\widehat{\boldsymbol{V}}_{C,t}^{-1}\right\| = \mathcal{O}_p(1).
\end{aligned}
$$

□

Proof of Theorem 2

Proof. We make use of the following equality for each row of equation (A.1): for each row vector $\boldsymbol{R}_{t,i\cdot} \in \mathbb{R}^k$, $i \in [p]$, $t \in [T]$, we have

$$
\begin{aligned}
\widehat{\boldsymbol{R}}_{t,i\cdot} - \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_{t,i\cdot} = \widehat{\boldsymbol{V}}_{R,t}^{-1} \sum_{s=1}^{T} \frac{1}{pqT} K_{h,st} \widehat{\boldsymbol{R}}_t^\top \Big(& \boldsymbol{E}_s \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{E}_{s,i\cdot} + \boldsymbol{E}_s \boldsymbol{E}_{s,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} \\
& + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} \\
& + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{C}_t \boldsymbol{F}_s \boldsymbol{D}_{R,st,i\cdot} \\
& + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} \\
& + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} + \boldsymbol{E}_s \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} \\
& + \boldsymbol{E}_s \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_{t,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} \\
& + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} + \boldsymbol{E}_s \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{E}_{s,i\cdot} \\
& + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{E}_{s,i\cdot} + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{E}_{s,i\cdot} \Big) \\
\triangleq & \sum_{j=1}^{24} \mathrm{II_j} \widehat{\boldsymbol{V}}_{\mathrm{R,t}}^{-1}, \qquad \text{(A.6)}
\end{aligned}
$$

By Lemma 2 below, we have $\sum_{j=4}^{13} \mathrm{II_j} \widehat{\boldsymbol{V}}_{\mathrm{R,t}}^{-1} = \boldsymbol{B}_{\mathrm{R_{t,i}}} + \mathrm{o_P}\left(\mathrm{h}^2\right)$, which determines the asymptotic bias. If $\sqrt{qT}/p \to 0$, then we have

$$
\begin{aligned}
\sqrt{qTh}\Big(\widehat{\boldsymbol{R}}_{t,i\cdot} - \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_{t,i\cdot} - \boldsymbol{B}_{R_{t,i}}\Big) \quad &= \quad \widehat{\boldsymbol{V}}_{R,t}^{-1} \frac{\widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t}{p} \frac{1}{\sqrt{qTh}} \sum_{s=1}^{T} K_{h,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \boldsymbol{e}_{s,i\cdot} + o_p(1) \\
&\xrightarrow{\mathcal{D}} \quad \mathcal{N}\left(\mathbf{0}, \nu_0 \Sigma_{R_{t,i}}\right),
\end{aligned}
$$

by Lemma 14 and continuous mapping theorem. Note that Here $\nu_0 \Sigma_{R_{t,i}}$ is defined in Theorem 2. The proof for the estimated column loading matrix $\widehat{\boldsymbol{C}}_t$ can be derived in a similar way. □

Lemma 2. *Define* $\delta_{pqTh} = \min\left\{\sqrt{p}, \sqrt{qTh}\right\}$. *Under Assumption 1 - 6, we have*

(a) $\mathrm{II}_1 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\delta_{\mathrm{pqTh}}\sqrt{\mathrm{qTh}}}\right)$,

(b) $\mathrm{II}_2 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\sqrt{\mathrm{qTh}}}\right)$

(c) $\mathrm{II}_3 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\delta_{\mathrm{pqTh}}\sqrt{\mathrm{qTh}}}\right) + \mathcal{O}_\mathrm{p}\left(\frac{1}{\delta_{\mathrm{pqTh}}\sqrt{\mathrm{p}}}\right)$

(d) $\sum_{j=4}^{13} \mathrm{II_j} = \frac{1}{2}\mathrm{h}^2 \mu_2 \sum_{\mathrm{c_1+c_2+c_3+c_4=2}} \boldsymbol{H}_{\mathrm{R,t}}^\top \Sigma_{\mathrm{R,t}}^{(\mathrm{c_1})} \Sigma_{\mathrm{FC,t}}^{(\mathrm{c_2,c_3})} \boldsymbol{R}_{\mathrm{t,i.}}^{(\mathrm{c_4})} + \mathrm{o_P}\left(\mathrm{h}^2\right)$;

(e) $\mathrm{II_j} = \mathrm{o_P}\left(\mathrm{h}^2\right), \mathrm{j} = 14, \cdots, 24.$

Proof. (a)

$$
\begin{aligned}
\mathrm{II}_1 &= \frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q} K_{h,st}\widehat{\boldsymbol{R}}_{t,l\cdot}e_{s,lj}\boldsymbol{C}_{t,j\cdot}^{\top}\boldsymbol{F}_s^{\top}\boldsymbol{R}_{t,i\cdot} \\
&= \frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q} K_{h,st}(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot})e_{s,lj}\boldsymbol{C}_{t,j\cdot}^{\top}\boldsymbol{F}_s^{\top}\boldsymbol{R}_{t,i\cdot} \\
&+ \boldsymbol{H}_{R,t}^{\top}\frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q} K_{h,st}\boldsymbol{R}_{t,l\cdot}e_{s,lj}\boldsymbol{C}_{t,j\cdot}^{\top}\boldsymbol{F}_s^{\top}\boldsymbol{R}_{t,i\cdot} \\
&= \mathrm{II}_{11}+\mathrm{II}_{12}.
\end{aligned}
$$

We bound each term as follows.

$$
\begin{aligned}
\|\mathrm{II}_{11}\| &\le \frac{1}{\sqrt{qTh}}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\sqrt{\frac{h}{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}e_{s,lj}\boldsymbol{C}_{t,j\cdot}^{\top}\boldsymbol{F}_s^{\top}\right\|^2\right)^{1/2}\left\|\boldsymbol{R}_{t,i\cdot}\right\| \\
&= \frac{1}{\sqrt{qTh}}\cdot\mathcal{O}_p\left(\frac{1}{\delta_{pqTh}}\right)\cdot\mathcal{O}_p(1),
\end{aligned}
$$

where the last equality results from Theorem 1 and Lemma 11. We also have

$$
\|\mathrm{II}_{12}\| = \frac{1}{\sqrt{pqTh}}\left\|\boldsymbol{H}_{R,t}^{\top}\left(\sqrt{\frac{h}{pqT}}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\boldsymbol{R}_{t,l\cdot}e_{s,lj}\boldsymbol{C}_{t,j\cdot}^{\top}\boldsymbol{F}_s^{\top}\right)\boldsymbol{R}_{t,i\cdot}\right\| = \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right),
$$

where the last equality results from by Markov Theorem, Lemma 12 (b) and Proposition 2.

(b)

$$
\begin{aligned}
\mathrm{II}_2 &= \frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}K_{h,st}\widehat{\boldsymbol{R}}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\boldsymbol{F}_s\sum_{j=1}^{q}\boldsymbol{C}_{t,j\cdot}e_{s,ij} \\
&= \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{l\cdot}^{\top}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_{t,j\cdot}e_{s,ij} \\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_{t,j\cdot}e_{s,ij} \\
&= \mathrm{II}_{21}+\mathrm{II}_{22}.
\end{aligned}
$$

We bound each term as follows.

$$
\begin{aligned}
\|\mathrm{II}_{21}\| &\leq \frac{1}{\sqrt{qT}}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^{2}\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}^{\top}\frac{1}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t,j\cdot}e_{s,ij}\right\|^{2}\right)^{1/2} \\
&\leq \frac{1}{\sqrt{qTh}}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{t,R}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^{2}\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}\right\|^{2}\cdot\left\|\sqrt{\frac{h}{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t,j\cdot}e_{s,ij}\right\|^{2}\right)^{1/2} \\
&= \mathcal{O}_{p}\left(\frac{1}{\delta_{pqTh}\sqrt{qTh}}\right) \qquad \text{by Theorem 1 and Lemma 11.}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\|\mathrm{II}_{22}\| &= \frac{1}{\sqrt{qTh}}\left\|\frac{1}{p}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\right\|\cdot\left\|\sqrt{\frac{h}{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t,j\cdot}e_{s,ij}\right\| \\
&= \mathcal{O}_{p}\left(\frac{1}{\sqrt{qTh}}\right).
\end{aligned}
$$

Combing all the terms, we have

$$
\|\mathrm{II}_{2}\| = \mathcal{O}_{p}\left(\frac{1}{\sqrt{qTh}}\right).
$$

(c)

$$
\begin{aligned}
\mathrm{II}_{3} &= \frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\left(e_{s,lj}e_{s,ij}-\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right) \\
&\quad+\frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\mathbb{E}\left[e_{s,lj}e_{s,ij}\right] \\
&\quad+\frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\left(e_{s,lj}e_{s,ij}-\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right) \\
&\quad+\frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\mathbb{E}\left[e_{s,lj}e_{s,ij}\right] \\
&= \mathrm{II}_{31}+\mathrm{II}_{32}+\mathrm{II}_{33}+\mathrm{II}_{34}
\end{aligned}
$$

We bound each term as follows.

$$\begin{aligned}\|\mathrm{II}_{31}\| &\leq \frac{1}{\sqrt{qTh}}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^{2}\right)^{1/2}\cdot\left(\frac{1}{p}\sum_{l=1}^{p}\left(\sqrt{\frac{h}{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\left(e_{s,lj}e_{s,ij}-\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right)\right)^{2}\right)^{1/2}\\ &= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{qTh}}\right) \qquad \text{by Lemma 13 and Markov inequality.}\end{aligned}$$

$$\begin{aligned}\|\mathrm{II}_{32}\| &\leq \frac{1}{\sqrt{p}}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^{2}\right)^{1/2}\cdot\left(\frac{1}{p}\sum_{l=1}^{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right)^{2}\right)^{1/2}\\ &= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{p}}\right)\end{aligned}$$

$$\begin{aligned}\|\mathrm{II}_{33}\| &\leq \left\|\frac{1}{pqT}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,st}\boldsymbol{R}_{t,l\cdot}\left(e_{s,lj}e_{s,ij}-\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right)\right\|\cdot\left\|\boldsymbol{H}_{R,t}\right\|\\ &= \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right) \qquad \text{by Lemma 12 (a) and Proposition 2.}\end{aligned}$$

$$\begin{aligned}\|\mathrm{II}_{34}\| &\leq \frac{1}{p}\cdot\left\|\boldsymbol{H}_{R,t}^{\top}\right\|\cdot\sum_{l=1}^{p}\left|\frac{1}{qT}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right|\cdot\max_{l\in[p]}\left\|\boldsymbol{R}_{t,l\cdot}\right\|\\ &= \mathcal{O}\left(\frac{1}{p}\right) \qquad \text{by Assumption 3 and Proposition 2.}\end{aligned}$$

Combing the result on each term, we obtain

$$\|\mathrm{II}_3\| = \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{qTh}}\right)+\mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{p}}\right).$$

(d)

$$
\begin{aligned}
\mathrm{II}_4 &= \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s\boldsymbol{R}_{t,i\cdot}^{(1)}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s\boldsymbol{R}_{t,i\cdot}^{(2)}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s M_{R,s,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s\boldsymbol{R}_{t,i\cdot}^{(1)}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s\boldsymbol{R}_{t,i\cdot}^{(2)}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{\top}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s M_{R,s,i\cdot}\\
&= \sum_{j=1}^{6}\mathrm{II}_{4j}. \qquad (A.8)
\end{aligned}
$$

We bound each term as follows.

$$
\begin{aligned}
\|\mathrm{II}_{41}\| &\le \left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left\|\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s\right\|\left\|\boldsymbol{R}_{t,i\cdot}^{(1)}\right\|\\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}}\right)\mathcal{O}_p(1)\,\mathcal{O}_p\left(\sqrt{\frac{h}{T}}\right)\mathcal{O}_p(1)\\
&= o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{II}_{42}\| &\le \left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left\|\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^2 K_{h,st}\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s\right\|\left\|\boldsymbol{R}_{t,i\cdot}^{(2)}\right\|\\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}}\right)\mathcal{O}_p(1)\,\mathcal{O}_p\left(h^2\right)\mathcal{O}_p(1)\\
&= o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{II}_{43}\| &\leq \left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left\|\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s M_{R,s,i\cdot}\right\| \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}}\right)\mathcal{O}_p(1)\,o_P\left(h^2\right)\mathcal{O}_p(1) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{II}_{44} &= \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(1)} \\
&+ \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\left[\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s-\Sigma_{FC,t}^{(0,0)}\right]\boldsymbol{R}_{t,i\cdot}^{(1)} \\
&= \mathcal{O}_p(1)\mathcal{O}_p\left(\frac{1}{T}\right)+\mathcal{O}_p(1)\mathcal{O}_p\left(\sqrt{\frac{h}{T}}\right) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{II}_{45} &= \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^2K_{h,st}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(2)} \\
&+ \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^2K_{h,st}\left[\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s-\Sigma_{FC,t}^{(0,0)}\right]\boldsymbol{R}_{t,i\cdot}^{(2)} \\
&= \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^{\top}\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(2)}+\mathcal{O}_p\left(\frac{h^2}{Th}\right)+\mathcal{O}_p\left(\frac{h^2}{\sqrt{Th}}\right) \\
&= \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^{\top}\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(2)}+o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{II}_{46} &= \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\Sigma_{FC,t}^{(0,0)}M_{R,s,i\cdot} \\
&+ \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\left[\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{\top}\boldsymbol{C}_t}{q}\boldsymbol{F}_s-\Sigma_{FC,t}^{(0,0)}\right]M_{R,s,i\cdot} \\
&= h^2\boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\right)\int k(u)\Sigma_{FC,t}^{(0,0)}h^{-2}M_{R,uhT+t,i\cdot}du+o_P\left(h^2\right) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

Combing the result on each term, we obtain

$$\mathrm{II}_4 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(2)} + o_P\left(h^2\right).$$

Similarly, we obtain

$$\mathrm{II}_5 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(0,2)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_6 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(2)}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_7 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(2,0)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_8 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(1)}\Sigma_{FC,t}^{(0,0)}\boldsymbol{R}_{t,i\cdot}^{(1)} + o_P\left(h^2\right),$$

$$\mathrm{II}_9 = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(1,1)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_{10} = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(1,0)}\boldsymbol{R}_{t,i\cdot}^{(1)} + o_P\left(h^2\right),$$

$$\mathrm{II}_{11} = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(1)}\Sigma_{FC,t}^{(0,1)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_{12} = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(1)}\Sigma_{FC,t}^{(1,0)}\boldsymbol{R}_{t,i\cdot} + o_P\left(h^2\right),$$

$$\mathrm{II}_{13} = \frac{1}{2}h^2\mu_2\boldsymbol{H}_{R,t}^\top\Sigma_{R,t}^{(0)}\Sigma_{FC,t}^{(0,1)}\boldsymbol{R}_{t,i\cdot}^{(1)} + o_P\left(h^2\right).$$

Combining all these terms, we have

$$\sum_{j=4}^{13}\mathrm{II_j} = \frac{1}{2}\mathrm{h}^2\mu_2\sum_{\mathrm{c_1+c_2+c_3+c_4=2}}\boldsymbol{H}_{\mathrm{R,t}}^\top\boldsymbol{\Sigma}_{\mathrm{R,t}}^{(\mathrm{c_1})}\boldsymbol{\Sigma}_{\mathrm{FC,t}}^{(\mathrm{c_2,c_3})}\boldsymbol{R}_{\mathrm{t,i.}}^{(\mathrm{c_4})} + \mathrm{o_P}\left(\mathrm{h}^2\right).$$

(e)

$$
\begin{aligned}
\mathrm{II}_{14} &= \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^{3}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(2)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(2)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\left(\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right)\boldsymbol{R}_{t,l\cdot}^{(2)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^{3}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(2)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{(1)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(2)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}\\
&+ \frac{1}{pqT}\sum_{l=1}^{p}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\boldsymbol{R}_{t,l\cdot}^{(2)\top}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^{4}K_{h,st}\boldsymbol{F}_{s}\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(1)}\boldsymbol{F}_{s}\boldsymbol{R}_{t,i\cdot}+o_{p}\left(h^{4}\right)\\
&= \sum_{j=1}^{8}\mathrm{II}_{14j}+o_{p}\left(h^{4}\right).
\end{aligned}
$$

We bound each term as follows.

$$
\begin{aligned}
\|\mathrm{II}_{141}\| &\leq \left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^{2}\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}^{(1)}\right\|^{2}\right)^{1/2}\left\|\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^{3}K_{h,st}\boldsymbol{F}_{s}\frac{\boldsymbol{C}_{t}^{(1)\top}\boldsymbol{C}_{t}^{(1)}}{q}\boldsymbol{F}_{s}\right\|\left\|\boldsymbol{R}_{t,i\cdot}\right\|\\
&= \mathcal{O}_{p}\left(\frac{1}{\delta_{pqTh}}\right)\mathcal{O}_{p}(1)\,\mathcal{O}_{p}\left(\sqrt{\frac{h^{5}}{T}}\right)\mathcal{O}_{p}(1)\\
&= o_{P}\left(h^{2}\right).
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{II}_{142}\| &\leq \left(\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2}\left(\frac{1}{p}\sum_{l=1}^{p}\left\|\boldsymbol{R}_{t,l\cdot}^{(1)}\right\|^2\right)^{1/2}\left\|\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{h,st}\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{(2)\top}\boldsymbol{C}_t^{(1)}}{q}\boldsymbol{F}_s\right\|\left\|\boldsymbol{R}_{t,i\cdot}\right\| \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}}\right)\mathcal{O}_p(1)\mathcal{O}_p\left(\sqrt{\frac{h^7}{T}}\right)\mathcal{O}_p(1) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

Similarly, we have $\|\mathrm{II}_{143}\| = o_P\left(h^2\right)$ and $\|\mathrm{II}_{144}\| = o_P\left(h^2\right)$. Then

$$
\begin{aligned}
\mathrm{II}_{145} &= \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t^{(1)}}{p}\right)\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^3 K_{h,st}\Sigma_{FC,t}^{(1,1)}\boldsymbol{R}_{t,i\cdot} \\
&+ \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t^{(1)}}{p}\right)\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^3 K_{h,st}\left[\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(1)}}{q}\boldsymbol{F}_s-\Sigma_{FC,t}^{(1,1)}\right]\boldsymbol{R}_{t,i\cdot} \\
&= \mathcal{O}_p(1)\mathcal{O}_p\left(\frac{h^2}{T}\right)+\mathcal{O}_p(1)\mathcal{O}_p\left(\sqrt{\frac{h^5}{T}}\right) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{II}_{146} &= \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t^{(1)}}{p}\right)\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{h,st}\Sigma_{FC,t}^{(2,1)}\boldsymbol{R}_{t,i\cdot} \\
&+ \boldsymbol{H}_{R,t}^{\top}\left(\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t^{(1)}}{p}\right)\frac{1}{2T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^4 K_{h,st}\left[\boldsymbol{F}_s\frac{\boldsymbol{C}_t^{(2)\top}\boldsymbol{C}_t^{(1)}}{q}\boldsymbol{F}_s-\Sigma_{FC,t}^{(2,1)}\right]\boldsymbol{R}_{t,i\cdot} \\
&= \frac{1}{2}h^4\mu_4\boldsymbol{H}_{R,t}^{\top}\Sigma_{R,t}^{(1)}\Sigma_{FC,t}^{(2,1)}\boldsymbol{R}_{t,i\cdot}+\mathcal{O}_p\left(\frac{h^4}{Th}\right)+\mathcal{O}_p\left(\frac{h^4}{\sqrt{Th}}\right) \\
&= o_P\left(h^2\right).
\end{aligned}
$$

Similarly, we can obtain $\mathrm{II}_{147} = \mathrm{o_P}\left(\mathrm{h}^2\right)$ and $\mathrm{II}_{148} = \mathrm{o_P}\left(\mathrm{h}^2\right)$. Combining these term, we have $\mathrm{II}_{14} = \mathrm{o_P}\left(\mathrm{h}^2\right)$. By very similar but tedious derivation, we can obtain $\mathrm{II_j} = \mathrm{o_P}\left(\mathrm{h}^2\right)$ for $j = 15,\dots,24$.

□

Proposition 3. *Under Assumption 1-6,*

$$
\operatorname*{plim}_{p,q,T\to\infty}\frac{\widehat{\boldsymbol{R}}_t^{\top}\boldsymbol{R}_t}{p} = \boldsymbol{Q}_{R,t}, \quad \textit{and} \quad \operatorname*{plim}_{p,q,T\to\infty}\frac{\widehat{\boldsymbol{C}}_t^{\top}\boldsymbol{C}_t}{q} = \boldsymbol{Q}_{C,t}.
$$

The matrix $\boldsymbol{Q}_{R,t} \in \mathbb{R}^{k\times k}$ and $\boldsymbol{Q}_{C,t} \in \mathbb{R}^{r\times r}$ are given, respectively, by

$$\boldsymbol{Q}_{R,t} = \boldsymbol{V}_{R,t}^{1/2}\boldsymbol{\Psi}_{R,t}^{\top}\,\Sigma_{FC,t}^{-1/2} \quad \textit{and} \quad \boldsymbol{Q}_{C,t} = \boldsymbol{V}_{C,t}^{1/2}\,\boldsymbol{\Psi}_{C,t}^{\top}\,\Sigma_{FR,t}^{-1/2},$$

where $\boldsymbol{V}_{R,t}$ ($\boldsymbol{V}_{C,t}$) is a diagonal matrix with diagonal entries being the the eigenvalues of $\Sigma_{FC,t}^{1/2}\;\boldsymbol{\Omega}_{R,t}\;\Sigma_{FC,t}^{1/2}$ ($\Sigma_{FR,t}^{1/2}\;\boldsymbol{\Omega}_{C,t}\;\Sigma_{FR,t}^{1/2}$) in decreasing order, $\boldsymbol{\Psi}_{R,t}$ ($\boldsymbol{\Psi}_{C,t}$) is the corresponding eigenvector matrix such that $\boldsymbol{\Psi}_{R,t}^{\top}\boldsymbol{\Psi}_{R,t} = \boldsymbol{I}$ ($\boldsymbol{\Psi}_{C,t}^{\top}\boldsymbol{\Psi}_{C,t} = \boldsymbol{I}$), and $\boldsymbol{\Omega}_{R,t}$ ($\boldsymbol{\Omega}_{C,t}$) is defined in Assumption 3.

Proof. Let $\boldsymbol{X}_{s,t} = K_{h,st}^{1/2}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}$, multiply the identify $\frac{1}{pqT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^{\top}\widehat{\boldsymbol{R}}_t = \widehat{\boldsymbol{R}}_t\boldsymbol{V}_{R,t}$ on both sides by $\frac{1}{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\boldsymbol{R}_t^{\top}$ to obtain:

$$\frac{1}{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\boldsymbol{R}_t^{\top}\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^{\top}\widehat{\boldsymbol{R}}_t = \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t}{p}\boldsymbol{V}_{R,t}.$$

Expanding $\boldsymbol{Y}_s\boldsymbol{Y}_s^{\top}$, we can rewrite the above as

$$\left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t}{p}\boldsymbol{V}_{R,t} = \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)\frac{\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t}{p} + d_t, \tag{A.9}$$

where

$$d_t = \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{1}{p}\boldsymbol{R}_t^{\top}\left(\sum_{j=1}^{24}\mathrm{I}_j\widehat{\boldsymbol{R}}_t\right).$$

We have

$$\begin{aligned}\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top} &= \frac{1}{T}\sum_{s=1}^{T}K_{h,st}\left(\boldsymbol{F}_s(\boldsymbol{C}_t^{\top}\boldsymbol{C}_t/q)\boldsymbol{F}_s^{\top} - \mathbb{E}\left[\boldsymbol{F}_s(\boldsymbol{C}_t^{\top}\boldsymbol{C}_t/q)\boldsymbol{F}_s^{\top}\right]\right) + \frac{1}{T}\sum_{s=1}^{T}K_{h,st}\mathbb{E}\left[\boldsymbol{F}_s(\boldsymbol{C}_t^{\top}\boldsymbol{C}_t/q)\boldsymbol{F}_s^{\top}\right]\\ &= \mathcal{O}_p(1),\end{aligned}$$

and $d_t = \mathrm{o}_p(1)$ following the proof of Lemma 2.

We rewrite equation (A.9) as

$$\boldsymbol{B}_t\boldsymbol{V}_{R,t} = \left(\boldsymbol{A}_t + d_t\boldsymbol{B}_t^{-1}\right)\boldsymbol{B}_t,$$

where

$$
\begin{aligned}
\boldsymbol{A}_t &= \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{\boldsymbol{R}_t^{\top}\boldsymbol{R}_t}{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}, \\
\boldsymbol{B}_t &= \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{1/2}\frac{\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t}{p}.
\end{aligned}
$$

Each column of $\boldsymbol{B}_t$ is an eigenvector of the matrix $\left(\boldsymbol{A}_t + d_t\boldsymbol{B}_t^{-1}\right)$. By Proposition 1, we have

$$
\boldsymbol{B}_t^{\top}\boldsymbol{B}_t \xrightarrow{\mathcal{P}} \boldsymbol{V}_{R,t}, \tag{A.10}
$$

and $\boldsymbol{V}_{R,t}$ is the diagonal matrix consisting of the eigenvalues of $\Sigma_{FC,t}^{1/2}\,\boldsymbol{\Omega}_{R,t}\,\Sigma_{FC,t}^{1/2}$. Thus the eigenvalues of $\boldsymbol{B}_t^{\top}\boldsymbol{B}_t$ are asymptotically bounded away from infinity and zero, and $\boldsymbol{B}_t^{-1} = \mathcal{O}_p(1)$.

Let $\widehat{\boldsymbol{V}}_{R,t}^{*}$ be a diagonal matrix consisting of the diagonal elements of $\boldsymbol{B}_t^{\top}\boldsymbol{B}_t$. From (A.10), we have

$$
\widehat{\boldsymbol{V}}_{R,t}^{*} \xrightarrow{\mathcal{P}} \boldsymbol{V}_{R,t}. \tag{A.11}
$$

Let

$$
\widehat{\boldsymbol{\Psi}}_{R,t} = \boldsymbol{B}_t\widehat{\boldsymbol{V}}_{R,t}^{*\,-1/2}, \tag{A.12}
$$

then $\left\|\widehat{\boldsymbol{\Psi}}_{R,t}\right\| = 1$ and

$$
\widehat{\boldsymbol{\Psi}}_{R,t}\widehat{\boldsymbol{V}}_{R,t} = \left(\boldsymbol{A}_t + d_t\boldsymbol{B}_t^{-1}\right)\widehat{\boldsymbol{\Psi}}_{R,t},
$$

that is, each column of $\widehat{\boldsymbol{\Psi}}_{R,t}$ is an eigenvector of $\boldsymbol{A}_t+d_t\boldsymbol{B}_t^{-1}$. And we have $\boldsymbol{A}_t \xrightarrow{p} \Sigma_{FC,t}^{1/2}\,\boldsymbol{\Omega}_{R,t}\,\Sigma_{FC,t}^{1/2}$ and $d_t\boldsymbol{B}_t^{-1} = o_p(1)$. By eigenvector perturbation theory and Assumption 7, there exists a unique eigenvector matrix $\boldsymbol{\Psi}_{R,t}$ of $\Sigma_{FC,t}^{1/2}\,\boldsymbol{\Omega}_{R,t}\,\Sigma_{FC,t}^{1/2}$ such that $\left\|\widehat{\boldsymbol{\Psi}}_{R,t} - \boldsymbol{\Psi}_{R,t}\right\| = o_p(1)$.

From (A.12) and (A.11), we have

$$
\frac{\boldsymbol{R}_t^{\top}\widehat{\boldsymbol{R}}_t}{p} = \left(\frac{1}{qT}\sum_{s=1}^{T}\boldsymbol{X}_{s,t}\boldsymbol{X}_{s,t}^{\top}\right)^{-1/2}\widehat{\boldsymbol{\Psi}}_{R,t}\widehat{\boldsymbol{V}}_{R,t}^{*\,1/2} \xrightarrow{p} \Sigma_{FC,t}^{-1/2}\,\boldsymbol{\Psi}_{R,t}\,\boldsymbol{V}_{R,t}^{1/2}.
$$

□

Proof of Theorem 3

Proof. Under the assumption that $\frac{1}{p}\widehat{\boldsymbol{R}}_t^\top \widehat{\boldsymbol{R}}_t = \boldsymbol{I}_p$ and $\frac{1}{q}\widehat{\boldsymbol{C}}_t^\top \widehat{\boldsymbol{C}}_t = \boldsymbol{I}_q$, we have

$$
\begin{aligned}
\widehat{\boldsymbol{F}}_t - \boldsymbol{H}_{R,t}^{-1}\boldsymbol{F}_t\boldsymbol{H}_{C,t}^{-1\top} &= \frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top\left(\boldsymbol{R}_t - \widehat{\boldsymbol{R}}_t\boldsymbol{H}_{R,t}^{-1}\right)\boldsymbol{F}_t\left(\boldsymbol{C}_t - \widehat{\boldsymbol{C}}_t\boldsymbol{H}_{C,t}^{-1}\right)^\top \widehat{\boldsymbol{C}}_t \\
&+ \frac{1}{p}\widehat{\boldsymbol{R}}_t^\top\left(\boldsymbol{R}_t - \widehat{\boldsymbol{R}}_t\boldsymbol{H}_{R,t}^{-1}\right)\boldsymbol{F}_t\boldsymbol{H}_{C,t}^{-1\top} \\
&+ \frac{1}{q}\boldsymbol{H}_{R,t}^{-1}\boldsymbol{F}_t\left(\boldsymbol{C}_t - \widehat{\boldsymbol{C}}_t\boldsymbol{H}_{C,t}^{-1}\right)^\top \widehat{\boldsymbol{C}}_t \\
&+ \frac{1}{pq}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_t\left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right) \\
&+ \frac{1}{pq}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_t\boldsymbol{C}_t\boldsymbol{H}_{C,t} \\
&+ \frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{E}_t\left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right) \\
&+ \frac{1}{pq}\boldsymbol{H}_{R,t}^\top\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\boldsymbol{H}_{C,t} \\
&= \sum_{i=1}^{7}\mathrm{III_i}.
\end{aligned}
$$

Since $\frac{1}{\sqrt{p}}\left\|\boldsymbol{R}_t - \widehat{\boldsymbol{R}}_t\boldsymbol{H}_{R,t}^{-1}\right\| = o_p(1)$ and $\frac{1}{\sqrt{q}}\left\|\boldsymbol{C}_t - \widehat{\boldsymbol{C}}_t\boldsymbol{H}_{C,t}^{-1}\right\| = o_p(1)$, term III_1 is dominated by III_2 and III_3, and term III_4 is dominated by III_5 and III_6. Now we bound III_2, III_3, III_5, III_6 and III_7.

$$
\begin{aligned}
\mathrm{III}_2 &= \frac{1}{p}(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t})^\top\left(\boldsymbol{R}_t - \widehat{\boldsymbol{R}}_t\boldsymbol{H}_{R,t}^{-1}\right)\boldsymbol{F}_t\boldsymbol{H}_{C,t}^{-1\top} + \frac{1}{p}\boldsymbol{H}_{R,t}^\top\boldsymbol{R}_t^\top\left(\boldsymbol{R}_t - \widehat{\boldsymbol{R}}_t\boldsymbol{H}_{R,t}^{-1}\right)\boldsymbol{F}_t\boldsymbol{H}_{C,t}^{-1\top} \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}^2} + h^2\right),
\end{aligned}
$$

by Theorem 1 and Proposition 2, and 3. Similarly, using results in Theorem 1, Proposition 2 and Lemma 3, and 4, we have

$$
\mathrm{III}_3 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\gamma_{\mathrm{pqTh}}^2} + \mathrm{h}^2\right), \quad \mathrm{III}_5 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\delta_{\mathrm{pqTh}}^2} + \mathrm{h}^2\right), \quad \mathrm{III}_6 = \mathcal{O}_\mathrm{p}\left(\frac{1}{\gamma_{\mathrm{pqTh}}^2} + \mathrm{h}^2\right),
$$

where $\gamma_{pqTh} = \min\left\{\sqrt{q}, \sqrt{pTh}\right\}$. Finally, we have

$$
\widehat{\boldsymbol{F}}_t - \boldsymbol{H}_R^{-1}\boldsymbol{F}_t\boldsymbol{H}_C^{-1\top} = \mathcal{O}_p\left(\frac{1}{\min(p,q)} + h^2\right),
$$

where we uses results in Lemma 5 below. Next we consider the convergence of $\widehat{\boldsymbol{S}}_{t,ij}$. Define $\widetilde{\boldsymbol{R}}_{t,i\cdot} = \boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,i\cdot}$, $\widetilde{\boldsymbol{C}}_{t,j\cdot} = \boldsymbol{H}_{C,t}^{\top}\boldsymbol{C}_{t,j\cdot}$, and $\widetilde{\boldsymbol{F}}_t = \boldsymbol{H}_{R,t}^{-1}\boldsymbol{F}_t{\boldsymbol{H}_{C,t}^{\top}}^{-1}$, we have

$$
\begin{aligned}
\widehat{\boldsymbol{S}}_{t,ij} - \boldsymbol{S}_{t,ij} = & \left(\widehat{\boldsymbol{R}}_{t,i\cdot} - \widetilde{\boldsymbol{R}}_{t,i\cdot}\right)^{\top}\left(\widehat{\boldsymbol{F}}_t - \widetilde{\boldsymbol{F}}_t\right)\left(\widehat{\boldsymbol{C}}_{t,j\cdot} - \widetilde{\boldsymbol{C}}_{t,j\cdot}\right) + \left(\widehat{\boldsymbol{R}}_{t,i\cdot} - \widetilde{\boldsymbol{R}}_{t,i\cdot}\right)^{\top}\widetilde{\boldsymbol{F}}_t\left(\widehat{\boldsymbol{C}}_{t,j\cdot} - \widetilde{\boldsymbol{C}}_{t,j\cdot}\right) \\
& + \widetilde{\boldsymbol{R}}_{t,i\cdot}^{\top}\left(\widehat{\boldsymbol{F}}_t - \widetilde{\boldsymbol{F}}_t\right)\left(\widehat{\boldsymbol{C}}_{t,j\cdot} - \widetilde{\boldsymbol{C}}_{t,j\cdot}\right) + \left(\widehat{\boldsymbol{R}}_{t,i\cdot} - \widetilde{\boldsymbol{R}}_{t,i\cdot}\right)^{\top}\left(\widehat{\boldsymbol{F}}_t - \widetilde{\boldsymbol{F}}_t\right)\widetilde{\boldsymbol{C}}_{t,j\cdot} \\
& + \widetilde{\boldsymbol{R}}_{t,i\cdot}^{\top}\widetilde{\boldsymbol{F}}_t\left(\widehat{\boldsymbol{C}}_{t,j\cdot} - \widetilde{\boldsymbol{C}}_{t,j\cdot}\right) + \left(\widehat{\boldsymbol{R}}_{t,i\cdot} - \widetilde{\boldsymbol{R}}_{t,i\cdot}\right)^{\top}\widetilde{\boldsymbol{F}}_t\widetilde{\boldsymbol{C}}_{t,j\cdot} + \widetilde{\boldsymbol{R}}_{t,i\cdot}^{\top}\left(\widehat{\boldsymbol{F}}_t - \widetilde{\boldsymbol{F}}_t\right)\widetilde{\boldsymbol{C}}_{t,j\cdot}.
\end{aligned}
$$

Dominant terms are the last three terms. Note that $\left\|\widetilde{\boldsymbol{F}}_t\right\| = \mathcal{O}_p(1)$, $\left\|\widetilde{\boldsymbol{C}}_{t,j\cdot}\right\| = \mathcal{O}_p(1)$ and $\left\|\widetilde{\boldsymbol{R}}_{t,i\cdot}^{\top}\right\| = \mathcal{O}_p(1)$. From Theorem 2, we have

$$
\widehat{\boldsymbol{R}}_{t,i\cdot} - \widetilde{\boldsymbol{R}}_{t,i\cdot} = \mathcal{O}_p\left(\frac{1}{\min\left(p, \sqrt{qTh}\right)} + h^2\right), \quad \text{and } \widehat{\boldsymbol{C}}_{t,j\cdot} - \widetilde{\boldsymbol{C}}_{t,j\cdot} = \mathcal{O}_p\left(\frac{1}{\min\left(q, \sqrt{pTh}\right)} + h^2\right).
$$

Then using the convergence rate of $\widehat{\boldsymbol{F}}_t$, we have

$$
\widehat{\boldsymbol{S}}_{t,ij} - \boldsymbol{S}_{t,ij} = \mathcal{O}_p\left(\frac{1}{\min\left(p, q, \sqrt{qTh}, \sqrt{pTh}\right)} + h^2\right).
$$

$\square$

Lemma 3. *Under Assumption 1-7, the $k \times k$ matrix*

$$
\frac{1}{p}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^{\top}\boldsymbol{R}_t = \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}^2} + h^2\right);
$$

The $r \times r$ matrix

$$
\frac{1}{q}\left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right)^{\top}\boldsymbol{C}_t = \mathcal{O}_p\left(\frac{1}{\gamma_{pqTh}^2} + h^2\right).
$$

Proof. Using the identity (A.1), we have

$$
\frac{1}{p}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^{\top}\boldsymbol{R}_t = \widehat{\boldsymbol{V}}_{R,t}^{-1}\widehat{\boldsymbol{R}}_t^{\top}\left(\sum_{j=1}^{24}\mathrm{I_j}\right)^{\top}\boldsymbol{R}_t, \tag{A.13}
$$

where $\mathrm{I_j}, \mathrm{j} = 1, \ldots, 24$ are defined in (A.1).

We begin with the first term in (A.1),

$$\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_1^\top \boldsymbol{R}_\mathrm{t} = \frac{1}{\mathrm{p}^2\mathrm{qT}} \sum_{\mathrm{s}=1}^{\mathrm{T}} \mathrm{K}_{\mathrm{h,st}} \left(\widehat{\boldsymbol{R}}_\mathrm{t} - \boldsymbol{R}_\mathrm{t}\boldsymbol{H}_{\mathrm{R,t}}\right)^\top \boldsymbol{E}_\mathrm{s}\boldsymbol{C}_\mathrm{t}\boldsymbol{F}_\mathrm{s}^\top \boldsymbol{R}_\mathrm{t}^\top \boldsymbol{R}_\mathrm{t} + \frac{1}{\mathrm{p}^2\mathrm{qT}} \sum_{\mathrm{s}=1}^{\mathrm{T}} \mathrm{K}_{\mathrm{h,st}} \boldsymbol{H}_{\mathrm{R,t}}^\top \boldsymbol{R}_\mathrm{t}^\top \boldsymbol{E}_\mathrm{s}\boldsymbol{C}_\mathrm{t}\boldsymbol{F}_\mathrm{s}^\top \boldsymbol{R}_\mathrm{t}^\top \boldsymbol{R}_\mathrm{t} = \mathrm{I}_{11} + \mathrm{I}_{12}.$$

Note

$$\begin{aligned}
\|\mathrm{I}_{11}\| &\le \frac{1}{\sqrt{qTh}} \left(\frac{1}{p}\sum_{l=1}^{p} \left\|\widehat{\boldsymbol{R}}_{t,l\cdot} - \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_{t,l\cdot}\right\|^2\right)^{1/2} \cdot \left(\frac{1}{p}\sum_{l=1}^{p} \left\|\frac{\sqrt{h}}{\sqrt{qT}} \sum_{s=1}^{T}\sum_{j=1}^{q} K_{h,st} \boldsymbol{C}_{t,j\cdot}^\top e_{s,lj} \boldsymbol{F}_s^\top\right\|^2\right)^{1/2} \cdot \frac{1}{p}\left\|\boldsymbol{R}_t^\top \boldsymbol{R}_t\right\| \\
&= \frac{1}{\sqrt{qTh}} \cdot \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}} + h^2\right) \cdot \mathcal{O}_p(1),
\end{aligned}$$

and

$$\begin{aligned}
\|\mathrm{I}_{12}\| &\le \frac{1}{\sqrt{pqTh}} \left\|\boldsymbol{H}_{R,t}\right\| \cdot \left\|\frac{\sqrt{h}}{\sqrt{pqT}} \sum_{s=1}^{T} K_{h,st} \boldsymbol{R}_t^\top \boldsymbol{E}_s \boldsymbol{C}_t \boldsymbol{F}_s^\top\right\| \cdot \frac{1}{p}\left\|\boldsymbol{R}_t^\top \boldsymbol{R}_t\right\| \\
&= \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right).
\end{aligned}$$

Second, we deal with the second term in (A.1),

$$\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_2^\top \boldsymbol{R}_\mathrm{t} = \frac{1}{\sqrt{\mathrm{pqTh}}} \cdot \frac{1}{\mathrm{p}} \widehat{\boldsymbol{R}}_\mathrm{t}^\top \boldsymbol{R}_\mathrm{t} \cdot \frac{\sqrt{\mathrm{h}}}{\sqrt{\mathrm{pqT}}} \sum_{\mathrm{s}=1}^{\mathrm{T}} \mathrm{K}_{\mathrm{h,st}} \boldsymbol{F}_\mathrm{s} \boldsymbol{C}_\mathrm{t}^\top \boldsymbol{E}_\mathrm{s}^\top \boldsymbol{R}_\mathrm{t} = \mathcal{O}_\mathrm{p}\left(\frac{1}{\sqrt{\mathrm{pqTh}}}\right),$$

where we use Proposition 3 and Lemma 12.

Third, we deal with the third term in (A.1),

$$\begin{aligned}
\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_3^\top \boldsymbol{R}_\mathrm{t} &= \frac{1}{p^2qT} \sum_{s=1}^{T} K_{h,st} \left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t \boldsymbol{H}_{R,t}\right)^\top \left(\boldsymbol{E}_s\boldsymbol{E}_s^\top - \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right) \boldsymbol{R}_t + \frac{1}{p^2qT} \sum_{s=1}^{T} K_{h,st} \left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t \boldsymbol{H}_{R,t}\right)^\top \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right] \boldsymbol{R}_t \\
&+ \frac{1}{p^2qT} \sum_{s=1}^{T} K_{h,st} \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \left(\boldsymbol{E}_s\boldsymbol{E}_s^\top - \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right) \boldsymbol{R}_t + \frac{1}{p^2qT} \sum_{s=1}^{T} K_{h,st} \boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right] \boldsymbol{R}_t \\
&= \mathrm{I}_{31} + \mathrm{I}_{32} + \mathrm{I}_{33} + \mathrm{I}_{34}.
\end{aligned}$$

Note

$$\|\mathrm{I}_{31}\|^2 \leq \frac{1}{qTh}\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\frac{1}{p}\sum_{l=1}^{p}\frac{1}{p}\sum_{i=1}^{p}\left(\frac{\sqrt{h}}{\sqrt{qT}}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\left(e_{s,lj}e_{s,ij}-\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right)\right)^2\frac{1}{p}\sum_{i=1}^{p}\left\|\boldsymbol{R}_{t,i\cdot}^{\top}\right\|$$

$$= \mathcal{O}_p\left(\frac{1}{qTh}\right)\cdot\mathcal{O}_p\left(\frac{1}{\delta_{pqTh}^2}+h^4\right)\cdot\mathcal{O}_p(1), \qquad \text{using Theorem 1 and Lemma 13.}$$

$$\|\mathrm{I}_{32}\|^2 \leq \frac{1}{p}\frac{1}{p}\sum_{l=1}^{p}\left\|\widehat{\boldsymbol{R}}_{t,l\cdot}-\boldsymbol{H}_{t,R}^{\top}\boldsymbol{R}_{t,l\cdot}\right\|^2\frac{1}{p}\sum_{l=1}^{p}\sum_{i=1}^{p}\left(\frac{1}{qT}\sum_{s=1}^{T}\sum_{j=1}^{q}K_{h,st}\mathbb{E}\left[e_{s,lj}e_{s,ij}\right]\right)^2\frac{1}{p}\sum_{i=1}^{p}\left\|\boldsymbol{R}_{t,i\cdot}^{\top}\right\|^2$$

$$= \frac{1}{p}\cdot\mathcal{O}_p\left(\frac{1}{\delta_{pqTh}^2}+h^4\right)\cdot\mathcal{O}(1) \qquad \text{using Theorem 1.}$$

$$\|\mathrm{I}_{33}\| \leq \frac{1}{\sqrt{pqTh}}\left(\frac{1}{\sqrt{p}}\left\|\frac{\sqrt{h}}{\sqrt{qT}}\sum_{s=1}^{T}K_{h,st}\left(\boldsymbol{E}_s\boldsymbol{E}_s^{\top}-\mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^{\top}\right]\right)\right\|\right)\cdot\left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{p}\|\boldsymbol{R}_t\|^2$$
$$= \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right),$$

$$\|\mathrm{I}_{34}\|^2 \leq \frac{1}{p^2}\left\|\mathbb{E}\left[\frac{1}{qT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^{\top}\right]\right\|^2\frac{1}{p^2}\|\boldsymbol{R}_t\|^4\left\|\boldsymbol{H}_{R,t}\right\|^2$$
$$= \mathcal{O}_p\left(\frac{1}{p^2}\right).$$

Fourth, we deal with the fourth term in (A.1),

$$\widehat{\boldsymbol{R}}_t^{\top}\mathrm{I}_4^{\top}\boldsymbol{R}_t = \frac{1}{pT}\cdot\frac{1}{p}\widehat{\boldsymbol{R}}_t^{\top}\boldsymbol{R}_t\sum_{s=1}^{T}K_{h,st}\left(\frac{1}{q}\boldsymbol{F}_s\boldsymbol{C}_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}-\Sigma_{FC,t}^{(0,0)}\right)\boldsymbol{D}_{R,st}^{\top}\boldsymbol{R}_t$$
$$+\frac{1}{pT}\cdot\frac{1}{p}\widehat{\boldsymbol{R}}_t^{\top}\boldsymbol{R}_t\sum_{s=1}^{T}K_{h,st}\Sigma_{FC,t}^{(0,0)}\boldsymbol{D}_{R,st}^{\top}\boldsymbol{R}_t$$
$$= \mathcal{O}_p\left(\sqrt{\frac{h}{T}}\right)+\mathcal{O}_p\left(h^2\right).$$

Similarly, we can obtain $\widehat{\boldsymbol{R}}_t^{\top}\mathrm{I}_j^{\top}\boldsymbol{R}_t = \mathcal{O}_p\left(h^2\right)+\mathcal{O}_p\left(\sqrt{\frac{h}{T}}\right), j = 5,\ldots,13$. Next, we consider the fourteenth term in (A.1),

$$
\begin{aligned}
\left\|\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_{14}^\top \boldsymbol{R}_\mathrm{t}\right\|^2 \leq & \frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)^3 K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t^{(1)}\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\right\|^2 \\
&+\frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^4 K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t^{(1)}\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(2)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\right\|^2 \\
&+\frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^4 K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t^{(1)}\boldsymbol{F}_s\boldsymbol{C}_t^{(2)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\right\|^2 \\
&+\frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^4 K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{R}_t^{(2)}\boldsymbol{F}_s\boldsymbol{C}_t^{(1)\top}\boldsymbol{C}_t^{(1)}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\right\|^2 + o_P\left(h^8\right) \\
&= \mathcal{O}_p\left(h^5T^{-1}\right)+\mathcal{O}_p\left(h^8\right)+\mathcal{O}_p\left(h^8\right)+\mathcal{O}_p\left(h^8\right)+o_P\left(h^8\right) \\
&= o_P\left(h^4\right).
\end{aligned}
$$

The order of $\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_\mathrm{j}^\top \boldsymbol{R}_\mathrm{t}, \mathrm{j} = 15, \ldots, 18$ can be derived in a similar fashion. Last, we consider the nineteenth term in (A.1),

$$
\begin{aligned}
\left\|\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_{19}^\top \boldsymbol{R}_\mathrm{t}\right\|^2 \leq & \frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\left(\frac{s-t}{T}\right) K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(1)\top}\boldsymbol{R}_t\right\|^2 \\
&+\frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2 K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^{(2)\top}\boldsymbol{R}_t\right\|^2 \\
&+\frac{1}{p^4q^2T^2}\left\|\sum_{s=1}^{T} K_{h,st}\widehat{\boldsymbol{R}}_t^\top \boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{M}_{R,s}^\top\boldsymbol{R}_t\right\|^2 \\
&= \mathcal{O}_p\left(hp^{-2}q^{-1}T^{-1}\right)+\mathcal{O}_p\left(h^3p^{-2}q^{-1}T^{-1}\right)+o_P\left(hp^{-2}q^{-1}T^{-1}\right) \\
&= o_P\left(h^4\right).
\end{aligned}
$$

The order of $\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_\mathrm{j}^\top \boldsymbol{R}_\mathrm{t}, \mathrm{j} = 20, \ldots, 24$ can be derived in a similar fashion as I_{19}. Combining all these together, we obtain $\frac{1}{p}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{R}_t = \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}^2} + h^2\right)$. The order of $\frac{1}{q}\left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right)^\top \boldsymbol{C}_t$ can be derived analogously. □

Lemma 4. *Under Assumption 1 - 6, the $k \times r$ matrix*

$$
\frac{1}{pq}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_t\boldsymbol{C}_t = \mathcal{O}_p\left(\frac{1}{\delta_{pqT}^2} + h^2\right)
$$

$$\frac{1}{pq}\boldsymbol{R}_t^\top \boldsymbol{E}_t\left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t\boldsymbol{H}_{C,t}\right) = \mathcal{O}_p\left(\frac{1}{\gamma_{pqT}^2} + h^2\right)$$

Proof. Using the identity (A.1), we have

$$\frac{1}{pq}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_t\boldsymbol{C}_t = \widehat{\boldsymbol{V}}_{R,t}^{-1}\widehat{\boldsymbol{R}}_t^\top\left(\sum_{j=1}^{24}\mathrm{I}_\mathrm{j}\right)^\top \boldsymbol{E}_t\boldsymbol{C}_t, \tag{A.14}$$

where $\mathrm{I}_\mathrm{j}, \mathrm{j} = 1, \ldots, 24$ are defined in (A.1).

We begin with the first term in (A.1):

$$\begin{aligned}
\left\|\frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_1^\top \boldsymbol{E}_t\boldsymbol{C}_\mathrm{t}\right\| &\le \left\|\frac{1}{p^2q^2T}\sum_{s=1}^T K_{h,st}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&+ \left\|\frac{1}{p^2q^2T}\sum_{s=1}^T K_{h,st}\boldsymbol{H}_{R,t}^\top\boldsymbol{R}_t^\top\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&\le \frac{1}{\sqrt{pqTh}}\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\left\|\sqrt{\frac{h}{qT}}\sum_{s=1}^T K_{h,st}\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\right\|_F\left\|\frac{1}{pq}\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&+ \frac{1}{\sqrt{pqTh}}\left\|\boldsymbol{H}_{R,t}\right\|\left\|\sqrt{\frac{h}{pqT}}\sum_{s=1}^T K_{h,st}\boldsymbol{R}_t^\top\boldsymbol{E}_s\boldsymbol{C}_t\boldsymbol{F}_s^\top\right\|\left\|\frac{1}{pq}\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{qTh}}\right) + \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right),
\end{aligned}$$

where we used Theorem 1, Lemma 11 and Lemma 13. Next we consider the second term in (A.1):

$$\begin{aligned}
\left\|\frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_2^\top \boldsymbol{E}_t\boldsymbol{C}_\mathrm{t}\right\| &\le \left\|\frac{1}{p^2q^2T}\sum_{s=1}^T K_{h,st}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&+ \left\|\frac{1}{p^2q^2T}\sum_{s=1}^T K_{h,st}\boldsymbol{H}_{R,t}^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&\le \frac{1}{pq\sqrt{qTh}}\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\left\|\sqrt{\frac{h}{pqT}}\sum_{s=1}^T K_{h,st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|_F\|\boldsymbol{E}_t\boldsymbol{C}_t\|_F \\
&+ \frac{1}{pq\sqrt{pqTh}}\left\|\boldsymbol{H}_{R,t}\right\|\left\|\sqrt{\frac{h}{pqT}}\sum_{s=1}^T K_{h,st}\boldsymbol{R}_t\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{E}_s^\top\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\|_F \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{pqTh}}\right) + \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left\|\frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_3^\top \boldsymbol{E}_\mathrm{t}\boldsymbol{C}_\mathrm{t}\right\| &\leq \left\|\frac{1}{p^2q^2T}\sum_{s=1}^{T} K_{h,st}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{E}_s\boldsymbol{E}_s^\top \boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&+ \left\|\frac{1}{p^2q^2T}\sum_{s=1}^{T} K_{h,st}\boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \boldsymbol{E}_s\boldsymbol{E}_s^\top \boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
&\leq \frac{1}{pq\sqrt{qTh}}\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\left\|\sqrt{\frac{h}{pqT}}\sum_{s=1}^{T} K_{h,st}\left(\boldsymbol{E}_s\boldsymbol{E}_s^\top - \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right)\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
&+ \frac{1}{pq}\frac{1}{\sqrt{p}}\cdot\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\cdot\left\|\mathbb{E}\left[\frac{1}{qT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right\|\cdot\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
&+ \frac{1}{pq\sqrt{qTh}}\left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\sqrt{\frac{h}{pqT}}\sum_{s=1}^{T} K_{h,st}\left(\boldsymbol{E}_s\boldsymbol{E}_s^\top - \mathbb{E}\left[\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right)\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
&+ \frac{1}{pq\sqrt{p}}\left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\mathbb{E}\left[\frac{1}{qT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
&= \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}\sqrt{pqTh}}\right) + \mathcal{O}_p\left(\frac{1}{\delta_{pqTh}p}\right) + \mathcal{O}_p\left(\frac{1}{\sqrt{pqTh}}\right) + \mathcal{O}_p\left(\frac{1}{p}\right).
\end{aligned}$$

Next we consider the fourth term in (A.1):

$$\begin{aligned}
\frac{1}{pq}\widehat{\boldsymbol{R}}_t^\top \mathrm{I}_4^\top \boldsymbol{E}_\mathrm{t}\boldsymbol{C}_\mathrm{t} &= \frac{1}{p^2qT}\sum_{s=1}^{T} K_{h,st}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{R}_t\left(\frac{\boldsymbol{F}_s\boldsymbol{C}_t^\top \boldsymbol{C}_t\boldsymbol{F}_s^\top}{q} - \Sigma_{FC,t}\right)\boldsymbol{D}_{R,st}^\top \boldsymbol{E}_t\boldsymbol{C}_t \\
&+ \frac{1}{p^2qT}\sum_{s=1}^{T} K_{h,st}\left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t\boldsymbol{H}_{R,t}\right)^\top \boldsymbol{R}_t\Sigma_{FC,t}\boldsymbol{D}_{R,st}^\top \boldsymbol{E}_t\boldsymbol{C}_t \\
&+ \frac{1}{p^2qT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \boldsymbol{R}_t\left(\frac{\boldsymbol{F}_s\boldsymbol{C}_t^\top \boldsymbol{C}_t\boldsymbol{F}_s^\top}{q} - \Sigma_{FC,t}\right)\boldsymbol{D}_{R,st}^\top \boldsymbol{E}_t\boldsymbol{C}_t \\
&+ \frac{1}{p^2qT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \boldsymbol{R}_t\Sigma_{FC,t}\boldsymbol{D}_{R,st}^\top \boldsymbol{E}_t\boldsymbol{C}_t \\
&= \mathrm{I}_{41} + \mathrm{I}_{42} + \mathrm{I}_{43} + \mathrm{I}_{44}.
\end{aligned}$$

Then we bound each term.

$$
\begin{aligned}
\|\mathrm{I}_{41}\| \leq & \frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(1)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\frac{1}{T}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(2)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\frac{1}{pq}\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\frac{1}{T}\sum_{s=1}^{T}K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\boldsymbol{M}_{R,s}^{\top}\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
& =\mathcal{O}_p\left(\frac{\sqrt{h}}{\delta_{pqTh}\sqrt{T}}\right)+\mathcal{O}_p\left(\frac{\sqrt{h^3}}{\delta_{pqTh}\sqrt{T}}\right)+o_p\left(\frac{\sqrt{h^3}}{\delta_{pqTh}\sqrt{T}}\right).
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{I}_{42}\| \leq & \frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left(\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{h,st}\right)\|\Sigma_{FC,t}\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(1)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left(\frac{1}{T}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2K_{h,st}\right)\|\Sigma_{FC,t}\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(2)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\frac{1}{pq}\frac{1}{\sqrt{p}}\left\|\widehat{\boldsymbol{R}}_t-\boldsymbol{R}_t\boldsymbol{H}_{R,t}\right\|\frac{1}{\sqrt{p}}\|\boldsymbol{R}_t\|\left\|\frac{1}{T}\sum_{s=1}^{T}K_{h,st}\Sigma_{FC,t}\boldsymbol{M}_{R,s}^{\top}\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
& =\mathcal{O}_p\left(\frac{1}{\delta_{pqTh}T}\right)+\mathcal{O}_p\left(\frac{h^2}{\delta_{pqTh}}\right)+o_p\left(\frac{h^2}{\delta_{pqTh}}\right).
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{I}_{43}\| \leq & \left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{p}\left\|\boldsymbol{R}_t^{\top}\boldsymbol{R}_t\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\left(\frac{s-t}{T}\right)K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(1)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{p}\left\|\boldsymbol{R}_t^{\top}\boldsymbol{R}_t\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\frac{1}{2}\left(\frac{s-t}{T}\right)^2K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\right\|\left\|\frac{1}{pq}\boldsymbol{R}_t^{(2)\top}\boldsymbol{E}_t\boldsymbol{C}_t\right\| \\
& +\frac{1}{pq}\left\|\boldsymbol{H}_{R,t}\right\|\frac{1}{p}\left\|\boldsymbol{R}_t^{\top}\boldsymbol{R}_t\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}K_{st}\left(\boldsymbol{F}_s C_t^{\top}\boldsymbol{C}_t\boldsymbol{F}_s^{\top}/q-\Sigma_{FC,t}\right)\boldsymbol{M}_{R,s}^{\top}\right\|\|\boldsymbol{E}_t\boldsymbol{C}_t\| \\
& =\mathcal{O}_p\left(\sqrt{\frac{h}{T}}\right)+\mathcal{O}_p\left(\sqrt{\frac{h^3}{T}}\right)+o_p\left(\sqrt{\frac{h^3}{T}}\right).
\end{aligned}
$$

Combining these together, we have $\frac{1}{pq}\widehat{\boldsymbol{R}}_t^{\top}\mathrm{I}_4^{\top}\boldsymbol{E}_t\boldsymbol{C}_t=\mathcal{O}_p\left(h^2\right)$. Similarly, we can show that $\frac{1}{pq}\widehat{\boldsymbol{R}}_t^{\top}\mathrm{I}_j^{\top}\boldsymbol{E}_t\boldsymbol{C}_t=\mathcal{O}_p\left(h^2\right)$ for $j=5,\dots,13$ and $\frac{1}{pq}\widehat{\boldsymbol{R}}_t^{\top}\mathrm{I}_j^{\top}\boldsymbol{E}_t\boldsymbol{C}_t=o_p\left(h^2\right)$ for $j=14,\dots,24$. $\square$

Lemma 5. *Under Assumption 1-7,*

$$\frac{1}{pq}\boldsymbol{H}_{R,t}^{\top}\boldsymbol{R}_t^{\top}\boldsymbol{E}_t\boldsymbol{C}_t\boldsymbol{H}_{C,t}=\mathcal{O}_p\left(\frac{1}{\sqrt{pq}}\right).$$

Proof. Firstly, by Lemma 13 (b) and Markov inequality, we have $\frac{1}{\sqrt{pq}}\boldsymbol{R}_t^\top \boldsymbol{E}_t \boldsymbol{C}_t = \mathcal{O}_p(1)$. Combining results in Proposition 2, we have

$$\frac{1}{pq}\boldsymbol{H}_{R,t}^\top \boldsymbol{R}_t^\top \boldsymbol{E}_t \boldsymbol{C}_t \boldsymbol{H}_{C,t} = \frac{1}{\sqrt{pq}}\boldsymbol{H}_{R,t}^\top \left(\frac{1}{\sqrt{pq}}\boldsymbol{R}_t^\top \boldsymbol{E}_t \boldsymbol{C}_t\right)\boldsymbol{H}_{C,t} = \mathcal{O}_p\left(\frac{1}{\sqrt{pq}}\right).$$

□

Proof of Theorem 4

Proof. Let $\Gamma_i = \frac{1}{T}\sum_{t=1}^T \frac{\widehat{\lambda}_{it}}{\widehat{\lambda}_{i+1,t}}$. By Lemma 6 below, we have $\Gamma_i = \mathcal{O}_p(1)$ uniformly for $i = 1,\cdots,k-1$. We also have $\Gamma_i = \mathcal{O}_p(1)$, uniformly for $i = k+1,\cdots,p-1$ and $\Gamma_k \to \infty$ by Lemma 7. Then

$$\begin{aligned}
\mathbb{P}\left(\widehat{k} \le k\right) &= \mathbb{P}\left(\Gamma_k > \max\left(\Gamma_{k+1},\Gamma_{k+2},\cdots,\Gamma_{kmax}\right)\right) \to 1\\
\mathbb{P}\left(\widehat{k} \ge k\right) &= \mathbb{P}\left(\Gamma_k > \max\left(\Gamma_1,\Gamma_2,\cdots,\Gamma_{k-1}\right)\right) \to 1.
\end{aligned}$$

Therefore, we have $\mathbb{P}\left(\widehat{k} = k\right) \to 1$. □

Let $\widehat{\lambda}_{j,t} = \lambda_j\left(\frac{1}{pqT}\sum_{s=1}^T K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top\right)$.

Lemma 6. *For $j = 1,\cdots,k-1$,*

$$\frac{1}{T}\sum_{t=1}^T \frac{\widehat{\lambda}_{j,t}}{\widehat{\lambda}_{j+1,t}} = \mathcal{O}_p(1).$$

Proof. The result follows immediate from Proposition 1. □

Lemma 7. *, For $j = k,\cdots,p-1$,*

$$c + o_p(1) \le \min\left(\sqrt{qTh}, p\right)\widehat{\lambda}_{j,t} \le C + o_p(1),$$

where c, C, and $o_p(1)$ are uniform in $k \le j \le p-1$ and $1 \le t \le T$.

Proof. Let $\boldsymbol{X}_s = \boldsymbol{F}_s\boldsymbol{C}_s^\top$.

Then we have

$$\begin{aligned}
\frac{1}{pqT}\sum_{s=1}^T K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top &= \frac{1}{pqT}\sum_{s=1}^T K_{h,st}\left(\boldsymbol{R}_s + \boldsymbol{E}_s\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\right)\boldsymbol{X}_s\boldsymbol{X}_s^\top\left(\boldsymbol{R}_s + \boldsymbol{E}_s\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\right)^\top\\
&+ \frac{1}{pqT}\sum_{s=1}^T K_{h,st}\boldsymbol{E}_s\left(\boldsymbol{I} - \boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\boldsymbol{X}_s\right)\boldsymbol{E}_s^\top
\end{aligned}$$

From Weyl inequality and the fact that $\boldsymbol{X}_s$ is a $k \times q$ matrix,

$$\begin{aligned}
\lambda_{k+j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top\right) &\geq \lambda_{k+j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\left(\boldsymbol{I}-\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\boldsymbol{X}_s\right)\boldsymbol{E}_s^\top\right)\\
&= \lambda_{k+j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_t\left(\boldsymbol{I}-\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\boldsymbol{X}_s\right)\boldsymbol{E}_s^\top\right)\\
&\quad + \lambda_{k+1}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{X}_s^\top\left(\boldsymbol{X}_{\boldsymbol{X}_s^\top}\right)^{-1}\boldsymbol{X}_s\boldsymbol{E}_s^\top\right)\\
&\geq \lambda_{2k+j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right).
\end{aligned}$$

From Weyl inequality and the fact that $\boldsymbol{X}_s$ is a $k \times q$ matrix,

$$\begin{aligned}
\lambda_{k+j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{Y}_s\boldsymbol{Y}_s^\top\right) &\leq \lambda_{k+1}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\left(\boldsymbol{R}_s+\boldsymbol{E}_s\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\right)\boldsymbol{X}_s\boldsymbol{X}_s^\top\left(\boldsymbol{R}_s+\boldsymbol{E}_s\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\right)^\top\right)\\
&\quad + \lambda_{j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\left(\boldsymbol{I}-\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\boldsymbol{X}_s\right)\boldsymbol{E}_s^\top\right)\\
&\leq \lambda_{j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\left(\boldsymbol{I}-\boldsymbol{X}_s^\top\left(\boldsymbol{X}_s\boldsymbol{X}_s^\top\right)^{-1}\boldsymbol{X}_s\right)\boldsymbol{E}_s^\top\right)\\
&\leq \lambda_{j}\left(\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right).
\end{aligned}$$

By Lemma 13 and Assumption 4, we have

$$\left\|\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top-\mathbb{E}\left[\frac{1}{pqT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]\right\|=\mathcal{O}_p\left(\sqrt{qTh}\right).$$

and the minimal and maximal eigenvalues of $\mathbb{E}\left[\frac{1}{qT}\sum_{s=1}^{T}K_{h,st}\boldsymbol{E}_s\boldsymbol{E}_s^\top\right]$ are bounded away from above and below. So we have for $j=k,\cdots,p-1$,

$$c+o_p(1)\leq\min\left(\sqrt{qTh},p\right)\widehat{\lambda}_{j,t}\leq C+o_p(1).$$

□

Proof of Theorem 5

Proof. Recall that our estimator $\widehat{\boldsymbol{R}}_t$ ($\widehat{\boldsymbol{C}}_t$) is given by the matrix of $\sqrt{p}$ ($\sqrt{q}$) times the top k (r)

eigenvectors of $\widehat{\boldsymbol{G}}_{R,t}$ ($\widehat{\boldsymbol{G}}_{C,t}$) defined in equation (5.2) in descending order by corresponding eigenvalues. Similar to A.1, we have the following decomposition:

$$
\begin{aligned}
\widehat{\boldsymbol{R}}_t^p - \boldsymbol{R}_t \boldsymbol{H}_{R,t}^p &= \frac{1}{pq^2T} \sum_{s=1}^{T} K_{h,st} \mathbf{Y}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \mathbf{Y}_s^\top \widehat{\boldsymbol{R}}_t^p \widehat{\boldsymbol{V}}_{R,t}^{p\,-1} - \boldsymbol{R}_t \boldsymbol{H}_t^p \\
&= \frac{1}{pq^2T} \sum_{s=1}^{T} K_{h,st} \Big(\boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{E}_s^\top + \boldsymbol{E}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{E}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{E}_s^\top \\
&+ \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top \\
&+ \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top \\
&+ \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top \\
&+ \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \boldsymbol{D}_{C,st} \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top \\
&+ \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top \\
&+ \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{C}_t^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{E}_s^\top + \boldsymbol{R}_t \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{E}_s^\top + \boldsymbol{D}_{R,st} \boldsymbol{F}_s \boldsymbol{D}_{C,st}^\top \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{E}_s^\top \\
&+ \boldsymbol{E}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{C}_t \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top + \boldsymbol{E}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{R}_t^\top + \boldsymbol{E}_s \widehat{\boldsymbol{C}}_t \widehat{\boldsymbol{C}}_t^\top \boldsymbol{D}_{C,st} \boldsymbol{F}_s^\top \boldsymbol{D}_{R,st}^\top) \widehat{\boldsymbol{R}}_t^p \widehat{\boldsymbol{V}}_{R,t}^{p\,-1} \\
&\triangleq \sum_{j=1}^{24} \mathrm{I}_{\mathrm{j}}^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \widehat{\boldsymbol{V}}_{\mathrm{R,t}}^{\mathrm{p}\,-1}, \qquad \text{(A.15)}
\end{aligned}
$$

where $\boldsymbol{D}_{R,st} = \boldsymbol{R}_s - \boldsymbol{R}_t$ and $\boldsymbol{D}_{C,st} = \boldsymbol{C}_s - \boldsymbol{C}_t$. Following the proof of Proposition 1, we can show $\left\| \widehat{\boldsymbol{V}}_{R,t}^{p\,-1} \right\|_F^2 = \mathcal{O}_p(1)$. Then Theorem 5 follows by Lemma 8 below. □

Lemma 8. *Under Assumption 1 - 6, we have for any* $1 \le t \le T$,

(a) $\frac{1}{q} \left\| \frac{1}{Tp} \sum_{s=1}^{T} K_{h,st} \boldsymbol{E}_s^\top \left(\widehat{\boldsymbol{R}}_t - \boldsymbol{R}_t \boldsymbol{H}_{R,t} \right) \boldsymbol{F}_s \right\|_F^2 = \mathcal{O}_p\left(m_{pqTh} \right)$, *and* $\frac{1}{p} \left\| \frac{1}{Tq} \sum_{s=1}^{T} K_{h,st} \boldsymbol{E}_s \left(\widehat{\boldsymbol{C}}_t - \boldsymbol{C}_t \boldsymbol{H}_{C,t} \right) \boldsymbol{F}_s^\top \right\|_F^2 = \mathcal{O}_p\left(w_{pqTh} \right)$, *where* $m_{pqTh} = \frac{1}{T^2p^2h^2} + \frac{1}{Tq^2h} + \frac{1}{T^2pqh^2} + \frac{h^3}{Tpq}$ *and* $w_{pqTh} = \frac{1}{T^2q^2h^2} + \frac{1}{Tp^2h} + \frac{1}{T^2pqh^2} + \frac{h^3}{Tpq}$;

(b) $\frac{1}{p} \left\| \mathrm{I}_1^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh} + w_{pqTh} \right)$;

(c) $\frac{1}{p} \left\| \mathrm{I}_2^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \right\|_F^2 = \mathcal{O}_p\left(\frac{1}{qTh} + w_{pqTh} \right)$;

(d) $\frac{1}{p} \left\| \mathrm{I}_3^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \right\|_F^2 = \mathcal{O}_p\left(\frac{1}{pqTh} + \frac{1}{p^2q^2} + \frac{\gamma_{pqTh}^4}{Thp} + \frac{\gamma_{pqTh}^4}{p^2} \right) + \mathrm{o}_p(1) \times \frac{1}{p} \left\| \widehat{\boldsymbol{R}}_t^p - \boldsymbol{R}_t \boldsymbol{H}_{R,t}^p \right\|$;

(e) $\frac{1}{p} \left\| \mathrm{I}_{\mathrm{j}}^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \right\|_F^2 = \mathcal{O}_p\left(h^4 \right), j = 4, \cdots, 13$;

(f) $\frac{1}{p} \left\| \mathrm{I}_{\mathrm{j}}^{\mathrm{p}} \widehat{\boldsymbol{R}}_{\mathrm{t}}^{\mathrm{p}} \right\|_F^2 = \mathrm{o}_p\left(h^4 \right), j = 14, \cdots, 24.$

Appendix B Technical Lemmas

Lemma 9. *Under Assumption 1 - 6, we have that, as $T \to \infty$,*

$$\frac{1}{qT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{C}_t\boldsymbol{F}_s^\top \xrightarrow{\mathcal{P}} \Sigma_{FC,t}, \quad \text{and} \quad \frac{1}{pT}\sum_{s=1}^{T} K_{h,st}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{R}_t\boldsymbol{F}_s \xrightarrow{\mathcal{P}} \Sigma_{FR,t}.$$

where $\Sigma_{FC,t}$ and $\Sigma_{FR,t}$ are defined in (4.3).

Lemma 10. *Under Assumption 1 - 6, we have*

$$\sum_{s=1}^{T}\left\|\mathbb{E}\left[\frac{1}{pq}K_{h,st}\boldsymbol{R}^\top\boldsymbol{E}_s\boldsymbol{E}_t^\top\boldsymbol{R}\right]\right\| = \mathcal{O}(1), \qquad \sum_{s=1}^{T}\left\|\mathbb{E}\left[\frac{1}{pq}K_{h,st}\boldsymbol{C}^\top\boldsymbol{E}_s^\top\boldsymbol{E}_t\boldsymbol{C}\right]\right\| = \mathcal{O}(1).$$

Lemma 11. *Under Assumption 1 - 6, For any $i \in [p]$ and $j \in [q]$,*

$$\mathbb{E}\left[\left\|\frac{\sqrt{h}}{\sqrt{qT}}\sum_{t=1}^{T}K_{h,st}\boldsymbol{F}_t\boldsymbol{C}_s^\top\boldsymbol{e}_{t,i\cdot}\right\|^2\right] = \mathcal{O}(1), \qquad \mathbb{E}\left[\left\|\frac{\sqrt{h}}{\sqrt{pT}}\sum_{t=1}^{T}K_{h,st}\boldsymbol{F}_t^\top\boldsymbol{R}_s^\top\boldsymbol{e}_{t,\cdot j}\right\|^2\right] = \mathcal{O}(1),$$

Lemma 12. *Under Assumption 1 - 6, we have*

(a) For any row $i \in [p]$,

$$\mathbb{E}\left(\frac{\sqrt{h}}{\sqrt{pqT}}\sum_{t=1}^{T}\sum_{l=1}^{p}\sum_{j=1}^{q}K_{h,ts}\left(e_{t,lj}e_{t,ij} - \mathbb{E}\left[e_{t,lj}e_{t,ij}\right]\right)\right)^2 = \mathcal{O}(1).$$

Similarly, for any column $j \in [q]$,

$$\mathbb{E}\left(\frac{\sqrt{h}}{\sqrt{pqT}}\sum_{t=1}^{T}\sum_{i=1}^{p}\sum_{h=1}^{q}K_{h,ts}\left(e_{t,ih}e_{t,ij} - \mathbb{E}\left[e_{t,ih}e_{t,ij}\right]\right)\right)^2 = \mathcal{O}(1).$$

(b) The $k \times k$ matrix satisfies

$$\mathbb{E}\left\|\frac{\sqrt{h}}{\sqrt{pqT}}\sum_{t=1}^{T}K_{h,ts}\boldsymbol{R}_s^\top\boldsymbol{E}_t\boldsymbol{C}_s\boldsymbol{F}_t^\top\right\|^2 = \mathcal{O}(1).$$

Similarly, the $r \times r$ matrix satisfies

$$\mathbb{E}\left\|\frac{\sqrt{h}}{\sqrt{pqT}}\sum_{t=1}^{T}K_{h,ts}\boldsymbol{F}_t^\top\boldsymbol{R}_s^\top\boldsymbol{E}_t\boldsymbol{C}_s\right\|^2 = \mathcal{O}(1).$$

Lemma 13. *Under Assumption 1 - 6, we have for all p, q and T,*

(a) For any $i,l \in [p]$ and $j,h \in [q]$,

$$\mathbb{E}\left(\frac{\sqrt{h}}{\sqrt{qT}}\sum_{t=1}^{T}\sum_{j=1}^{q}K_{h,ts}\left(e_{t,ij}e_{t,lj}-\mathbb{E}\left[e_{t,ij}e_{t,lj}\right]\right)\right)^2 = \mathcal{O}(1),$$

$$\mathbb{E}\left(\frac{\sqrt{h}}{\sqrt{pT}}\sum_{t=1}^{T}\sum_{i=1}^{p}K_{h,ts}\left(e_{t,ij}e_{t,ih}-\mathbb{E}\left[e_{t,ij}e_{t,ih}\right]\right)\right)^2 = \mathcal{O}(1),$$

$$\mathbb{E}\left\|\frac{\sqrt{h}}{p\sqrt{qT}}\sum_{t=1}^{T}K_{h,ts}\left(\boldsymbol{E}_t\boldsymbol{E}_t^\top-\mathbb{E}\left[\boldsymbol{E}_t\boldsymbol{E}_t^\top\right]\right)\right\|_F^2 = \mathcal{O}(1),$$

$$\mathbb{E}\left\|\frac{\sqrt{h}}{p\sqrt{qT}}\sum_{t=1}^{T}K_{h,ts}\left(\boldsymbol{E}_t^\top\boldsymbol{E}_t-\mathbb{E}\left[\boldsymbol{E}_t^\top\boldsymbol{E}_t\right]\right)\right\|_F^2 = \mathcal{O}(1).$$

(b) For all $1 \le t,s \le T$, $\mathbb{E}\left\|\frac{1}{\sqrt{pq}}\boldsymbol{R}_t^\top\boldsymbol{E}_t\boldsymbol{C}_t\right\|^2 = \mathcal{O}(1)$.

Lemma 14. *Under Assumption 1 - 6, we have*

(a) For each row i, as $q,T \to \infty$,

$$\frac{\sqrt{h}}{\sqrt{qT}}\sum_{s=1}^{T}K_{h,ts}\boldsymbol{F}_s\boldsymbol{C}_t^\top\boldsymbol{e}_{s,i\cdot} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Phi}_{R,t,i}\right).$$

(b) For each column j, as $p,T \to \infty$,

$$\frac{\sqrt{h}}{\sqrt{pT}}\sum_{s=1}^{T}K_{h,ts}\boldsymbol{F}_s^\top\boldsymbol{R}_t^\top\boldsymbol{e}_{t,\cdot j} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Phi}_{C,t,j}\right).$$

The $\boldsymbol{\Phi}_{R,t,i}$'s and $\boldsymbol{\Phi}_{C,t,j}$'s are defined in Theorem 2.

The proofs of Lemmas 9-14 are collected in the online appendix.