Time-Varying Matrix Factor Model*

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

1Department of Economics, University of Rochester
2Stern School of Business, New York University
3Department 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
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 $\mathbf{y}$, and boldface capital letter $\mathbf{Y}$ represent scalar, vector, and matrix, respectively. For any matrix $\mathbf{Y}$, we use $\mathbf{Y}_{i\cdot}$, $\mathbf{y}_{\cdot j}$, and $\mathbf{y}_{ij}$ 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 $\mathbf{y}^\top$ for any vector $\mathbf{y}$. We use the following matrix norms: maximum norm $\|\mathbf{Y}\|_{\text{max}} \triangleq \max_{ij} |y_{ij}|$, $\ell_1$-norm $\|\mathbf{Y}\|_1 \triangleq \max_j \sum_i |y_{ij}|$, $\ell_\infty$-norm $\|\mathbf{Y}\|_\infty \triangleq \max_i \sum_j |y_{ij}|$, and $\ell_2$-norm $\|\mathbf{Y}\|_2 \triangleq \sigma_1$, where $\sigma_1$ is the largest singular value $\sigma_i$ of $\mathbf{Y}$ with $\sigma_i$ be the $i$-th largest square root of eigenvalues of $\mathbf{Y}^\top \mathbf{Y}$. When $\mathbf{Y}$ is a square matrix, we denote by $\text{Tr}(\mathbf{Y})$, $\lambda_{\text{max}}(\mathbf{y})$, and $\lambda_{\text{min}}(\mathbf{y})$ the trace, maximum and minimum singular value of $\mathbf{Y}$, 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 $Y_t (1 \leq t \leq T)$ be a matrix-valued time series, where each $Y_t$ is a matrix of size $p \times q$, 

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

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

$$
Y_t = R_tF_tC_t^T + E_t, \quad t = 1, 2, \cdots, T, 
$$

where $F_t$ is the $k \times r$ common factor matrix, $R_t$ ($C_t$) is a $p \times k$ ($q \times r$) time-varying row (column) loading matrix and the sequence of matrices $E_t$ is the noise matrix. The noise term $E_t$ is assumed to be uncorrelated with $F_t$, 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 $R$ and $C$. 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 $R_{t,i}$ ($C_{t,j}$) as a non-stochastic function of $t/T$, that is,

$$
R_{t,i} = R_i(t/T),
C_{t,j} = C_j(t/T),
$$

where $R_{t,i}(\cdot)$ ($C_{t,j}(\cdot)$) is an unknown smooth function of $t/T$ on $[0, 1]$ for each $i$ ($j$). The
specification that loading matrices $R_{t,i}(\cdot)$ and $C_{t,j}(\cdot)$ 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 $R_{t,i}(\cdot)$ and $C_{t,j}(\cdot)$ 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 $R_{t,i}(\cdot)$ and $C_{t,j}(\cdot)$ 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 $R_t$ and $C_t$ are to decrease suitably.

Now take the international trade flow as an example. The observed matrix $Y_t$ is a square matrix, where $p = q = n$. A general entry of $Y_t$, denoted as $y_{t,ij}$, where $i, j = 1, 2, \ldots, n$, 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 $Y_t$ 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 $R_t$) and imports from the $r$ importing hubs in certain distributions (determined by the loading matrix $C_t$). The entry of $R_t$, denoted as $R_{t,il}$, where $i = 1, \ldots, n$ and $l = 1, \ldots, k$, can be interpreted as the export contribution of country $i$ to the export hub $l$ at time $t$; the entry of $C_t$, denoted as $C_{t,jm}$, where $j = 1, \ldots, n$ and $m = 1, \ldots, r$, can be interpreted as the import contribution of country $j$ to the import hub $m$ at time $t$. The hubs trade, on behalf of the participating countries, among themselves and also within the hubs. The trading volume among the hubs is reflected by the factor matrix $F_t$, which is changing over time. The $(l,m)$-th element $F_{t,lm}$, where $l = 1, \ldots, k$ and $m = 1, \ldots, r$, 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 $R_{t,i} = R_i(t/T)$ ($C_{t,j} = C_j(t/T)$) and $R_i(C_j)$ is a smooth function, we have, when $\frac{t}{T} \approx \frac{s}{T}$

$$R_{t,i} \approx R_i(t/T) \approx R_i(s/T) = R_{s,i}.$$  

$$C_{t,j} \approx C_j(t/T) \approx C_j(s/T) = C_{s,j}.$$
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

$$Y_s \approx R_t F_s C_t^\top + E_s, \tag{3.1}$$

when \( \frac{t}{T} \approx \frac{s}{T} \). To estimate \( R_t, C_t, \) and \( F_t \), we minimize the local unexplained variation and bias jointly:

$$\min_{R_t, C_t, \{F_t\}_{t=1}^T} \frac{1}{pqT} \sum_{s=1}^T K_{h,st} \|Y_s - R_t F_s C_t^\top\|_F^2 + \frac{1}{pqT} \sum_{s=1}^T K_{h,st} \|Y_s - R_t F_s C_t^\top\|_F^2$$

subject to \( \frac{1}{p} R_t^\top R_t = I, \frac{1}{q} C_t^\top C_t = I \),

where \( K_{h,st} \) is a boundary-modified kernel function:

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

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 \( R_t \) (\( C_t \)), and \([Th]\) 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, \( \{Y_t\}_{1 \leq t \leq T} \) are projected onto \( R_t \) and maximize the row variances of \( R_t^\top Y_t \).

$$\tilde{R}_t = \arg \max_{R_t} Tr \left\{ \mathbb{E} \left( (R_t^\top Y_t)(R_t^\top Y_t)^\top \right) + \left( R_t^\top Y_t - \mathbb{E} \left[ R_t^\top Y_t \right] \right) \left( R_t^\top Y_t - \mathbb{E} \left[ R_t^\top Y_t \right] \right)^\top \right\} ,

= \arg \max_{R_t} Tr \left\{ pq R_t^\top M_{R,t} R_t \right\} ,

\text{subject to } \frac{1}{p} R_t^\top R_t = I,$$

where \( M_{R,t} \equiv p^{-1} q^{-1} \mathbb{E} \left[ Y_t Y_t^\top \right] \).

Alternatively, \( \{Y_t\}_{1 \leq t \leq T} \) are projected onto \( C_t \) and maximize the column variances of \( Y_t \).
\( Y_t C_t \).

\[
\widehat{C}_t = \arg\max_{C_t} \text{Tr} \left\{ \mathbb{E} \left[ (Y_t C_t)^\top (Y_t C_t) + (Y_t C_t - \mathbb{E}[Y_t C_t])^\top (Y_t C_t - \mathbb{E}[Y_t C_t]) \right] \right\}
\]

\[
= \arg\max_{C_t} \text{Tr} \left\{ pq C_t^\top M_{C,t} C_t \right\},
\]

subject to \( \frac{1}{p} C_t^\top C_t = I \),

where \( M_{C,t} \equiv p^{-1} q^{-1} \mathbb{E} \left[ Y_t^\top Y_t \right] \).

With \( T \) noisy observations, we define sample scatter matrix \( \widetilde{M}_{R,t} \) and \( \widetilde{M}_{C,t} \) as

\[
\begin{align*}
\widetilde{M}_{R,t} & = \frac{1}{pqT} \sum_{s=1}^{T} K_{h,ts} Y_s Y_s^\top \\
\widetilde{M}_{C,t} & = \frac{1}{pqT} \sum_{s=1}^{T} K_{h,ts} Y_s^\top Y_s
\end{align*}
\]

(3.4)

Furthermore, an estimator of \( F_t \) can be obtained by

\[
\widehat{F}_t = \frac{1}{pq} \widehat{R}_t^\top Y_t \widehat{C}_t
\]

and the signal part \( S_t = R_t F_t C_t^\top \) can be estimated by

\[
\widehat{S}_t = \frac{1}{pq} \widehat{R}_t \widehat{R}_t^\top Y_t \widehat{C}_t \widehat{C}_t^\top.
\]

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 \textit{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 \( \widetilde{M}_{R,t} \). And the generalized eigenvalue ratio-based estimator for \( k \) can be defined as

\[
\widehat{k} = \arg\max_{1 \leq j \leq k_{\text{max}}} \frac{1}{T} \sum_{t=1}^{T} \frac{\widehat{\lambda}_{j,t}}{\widehat{\lambda}_{j+1,t}},
\]

(3.5)

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

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

**Assumption 1. \( \alpha \)-mixing.** The vector-valued process \( \{ \text{vec}(F_t)^\top, \text{vec}(E_t)^\top \}^\top \) is a stationary \( \alpha \)-mixing process with mixing coefficients \( \alpha(h) = \sup \sup |P(A \cap B) - P(A)P(B)| \), where \( \mathcal{F}_t^\tau \) is the \( \sigma \)-field generated by \( \{ \text{vec}(F_t)^\top, \text{vec}(E_t)^\top \}^\top: \tau \leq t \leq s \), 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 \( F_t \) is of fixed dimension \( k \times r \), and \( \mathbb{E} \|F_t\|^4 \leq C \).

(b) For all \( i \in [p] \), \( j \in [q] \) and \( t \in [T] \), \( \mathbb{E}[e_{t,ij}] = 0 \) and \( \mathbb{E}|e_{t,ij}|^8 \leq C \).

(c) Factor and noise are uncorrelated, that is, \( \mathbb{E}[e_{t,ij}f_{s,ih}] = 0 \) 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 \( R_t \), \( \|R_{t,i}\| = O(1) \), and, as \( p, q \to \infty \), we have \( \|p^{-1}R_t^\top R_t - \Omega_{R,t}\| \to 0 \) for some \( k \times k \) positive definite matrix \( \Omega_{R,t} \). For each row of \( C_t \), \( \|C_{t,i}\| = O(1) \), and, as \( p, q \to \infty \), \( \|q^{-1}C_t^\top C_t - \Omega_{C,t}\| \to 0 \) for some \( r \times r \) positive definite matrix \( \Omega_{C,t} \).

**Assumption 4. Cross row (column) correlation of noise \( E_t \).** There exists some positive constant \( C < \infty \) such that,

(a) Let \( U_E = \mathbb{E}\left[ E_tE_t^\top / q \right] \) and \( V_E = \mathbb{E}\left[ E_tE_t^\top / p \right] \), we assume \( \|U_E\|_1 \leq C \) and \( \|V_E\|_1 \leq C \).

(b) For all row \( i \in [p] \) and column \( j \in [q] \) and \( t \in [T] \), we assume \( \sum_{l \in [p]} \sum_{h \in [q]} \mathbb{E}|e_{t,ij}e_{t,lih}| \leq C \).

(c) For any row \( i, l \in [p] \), any time \( t \in [T] \), and any column \( j \in [q] \),

\[
\sum_{m \in [p]} \sum_{s \in [T]} \sum_{h \in [q]} \left| \text{Cov}\left[ (e_{t,ij}, e_{t,lih}, e_{s,ih}, e_{s,mb}) \right] \right| \leq C
\]
Similar, for any column $j, h \in [q]$, any time $t \in [T]$, and any row $i \in [p]$,

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

**Assumption 5.** $R_i(t/T)$ and $C_{j}(t/T)$ 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 $t \in [T]$, 
$$
\mathbb{E} \left[ \left| \frac{1}{\sqrt{q}} \sum_{j=1}^{q} e_{t,ij} \right|^m \right] = O(1), \quad \mathbb{E} \left[ \left| \frac{1}{\sqrt{q}} \sum_{j=1}^{q} C_{t,j}^{(c)} e_{t,ij} \right|^m \right] = O(1), \text{ and } \mathbb{E} \left[ \left\| f_{t,l} \right\|^m \right] \leq C.
$$

(b) For any $h \in [r], j \in [q]$, and $t \in [T]$, 
$$
\mathbb{E} \left[ \left| \frac{1}{\sqrt{p}} \sum_{i=1}^{p} e_{t,ij} \right|^m \right] = O(1), \quad \mathbb{E} \left[ \left| \frac{1}{\sqrt{p}} \sum_{i=1}^{p} R_{t,i}^{(c)} e_{t,ij} \right|^m \right] = O(1), \text{ and } \mathbb{E} \left[ \left\| f_{t,h} \right\|^m \right] \leq C.
$$

(c) For any $t \in [T]$, 
$$
\mathbb{E} \left[ \left| \frac{1}{\sqrt{pq}} \sum_{i=1}^{p} \sum_{j=1}^{q} e_{t,ij} \right|^m \right] = O(1) \text{ and } \mathbb{E} \left[ \left\| \frac{1}{\sqrt{pq}} \sum_{i=1}^{p} \sum_{j=1}^{q} R_{t,i}^{(c)} C_{t,j}^{(c)\top} e_{t,ij} \right\|^m \right] = O(1).
$$

**Assumption 6.** $k : [-1, 1] \rightarrow \mathbb{R}^+$ 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 $\alpha$-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 $F_t$ has a nontrivial contribution to the variance of rows and columns of $Y_t$. 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 $E_t$ are i.i.d. over rows and columns for any $t$ and are independent of the factor $F_t$, 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 $\tilde{V}_{R,t} \in \mathbb{R}^{k \times k}$ and $\tilde{V}_{C,t} \in \mathbb{R}^{r \times r}$ be the time-varying diagonal matrices consisting of the first $k$ and $r$ largest eigenvalues of $\tilde{M}_{R,t}$ $(\tilde{M}_{C,t})$ in a descending order respectively. By the definition of our estimators $\tilde{R}_t$ and $\tilde{C}_t$, we have

$$
\tilde{R}_t = \tilde{M}_{R,t} \tilde{R}_t \tilde{V}_{R,t}^{-1},
\tilde{C}_t = \tilde{M}_{C,t} \tilde{C}_t \tilde{V}_{C,t}^{-1}.
$$

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

$$
H_{R,t} = \frac{1}{Tpq} \sum_{s=1}^{T} K_{h,ts} F_s C_t^T C_t F_s^T R_t^T R_t \tilde{R}_t \tilde{V}_{R,t}^{-1},
$$

$$
H_{C,t} = \frac{1}{Tpq} \sum_{s=1}^{T} K_{h,ts} F_s R_t^T R_t F_s C_t^T C_t \tilde{C}_t \tilde{V}_{C,t}^{-1}.
$$

**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

$$
\frac{1}{p} \left\| \tilde{R}_t - R_t H_{R,t} \right\|^2_F = O_p \left( \frac{1}{p^2} + \frac{1}{Tqh} + h^4 \right),
$$

$$
\frac{1}{q} \left\| \tilde{C}_t - C_t H_{C,t} \right\|^2_F = O_p \left( \frac{1}{q^2} + \frac{1}{Tp^2} + h^4 \right).
$$

Consequently,

$$
\frac{1}{p} \left\| \tilde{R}_t - RH_{R,t} \right\|^2 = O_p \left( \frac{1}{p^2} + \frac{1}{Tqh} + h^4 \right),
$$

$$
\frac{1}{q} \left\| \tilde{C}_t - CH_{C,t} \right\|^2 = O_p \left( \frac{1}{q^2} + \frac{1}{Tp^2} + h^4 \right).
$$

Theorem 1 establishes the convergence rate of our nonparametric estimators $\tilde{R}_t$ and $\tilde{C}_t$. 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:

$$
\text{vec}(Y_t) = (C_t \otimes R_t) \cdot \text{vec}(F_t) + \text{vec}(E_t),
$$

(4.2)

where $\text{vec}(Y_t) \in \mathbb{R}^{pq}$ and $\text{vec}(F_t) \in \mathbb{R}^{kr}$. Applying results in Su and Wang (2017) with-
out adopting the tensor structure in the loading matrix, we obtain $O(pq, b, \Xi^t - \Xi^t H^t, 2) = O_p \left( \frac{1}{p^2 q^2 T^2 h + h^4} \right)$, where $\Xi^t = C_t \otimes R_t$, and $H_t \in \mathbb{R}^{k_r \times k_r}$ is an orthonormal matrix.\(^1\). Hence, in a high-dimensional setting where $p, q \geq T$, Theorem 1 establishes faster $\ell_2$ convergence rate for both $\hat{R}_t$ and $\hat{C}_t$. In addition, to recover the tensor structure in the factor loadings, one needs to carry out a second step to estimate $R_t$ and $C_t$ from $\hat{\Xi}_t$ 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 $\hat{R}_t$ and $\hat{C}_t$, we define

$$
\Sigma_{FR,t}^{(c_1,c_2)} = E \left( F_t^T R_t^{(c_1)} R_t^{(c_2)} F_t \right),
$$

$$
\Sigma_{FC,t}^{(c_1,c_2)} = E \left( F_t^T C_t^{(c_1)} C_t^{(c_2)} F_t^T \right),
$$

$$
\Sigma_{R,t}^{(c)} = \frac{R_t^{(c)} R_t^{(c)}}{p},
$$

$$
\Sigma_{C,t}^{(c)} = \frac{C_t^{(c)} C_t^{(c)}}{q},
$$

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

**Assumption 7.** The matrices $\Omega_{R,t} \Sigma_{FC,t}$ and $\Omega_{C,t} \Sigma_{FR,t}$ 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 $R_t$, if $\sqrt{qT h}/p \to 0$, for each $i = 1, \cdots, p$ and $t = 1, \cdots, T$, we have

$$
\sqrt{qT h} \left( \hat{R}_{t,i} - H_{R,t}^T R_{t,i} - B_{R,t,i} \right) \to^d N(0, v_0 \Sigma_{R,t,i}),
$$

where

$$
B_{R,t,i} = \frac{1}{2} V_{R,t}^{-1} h^2 \mu_2 \sum_{c_1+c_2+c_3+c_4=2} \Sigma_{R,t}^{(c_1)} \Sigma_{FC,t}^{(c_2,c_3)} R_{t,i}^{(c_4)} + o_p \left( h^2 \right),
$$

$$
\Sigma_{R,t,i} = V_{R,t}^{-1} Q_{R,t,i} \Phi_{R,t,i} Q_{R,t,i}^T V_{R,t}^{-1},
$$

$$
\Phi_{R,t,i} = \text{plim} \frac{1}{qT} \sum_{r=1}^T \sum_{s=1}^T \mathbb{E} \left[ F_t C_t^T e_{r,i} e_{s,i}^T C_t F_s^T \right],
$$

\(^1\)The bias terms are missing in Su and Wang (2017) but have been corrected in Su and Wang (2020).
\( \nu_0 = \int_{-1}^{1} k^2(u)du, \mu_2 = \int_{-1}^{1} u^2 k(u)du, Q_{R,t} = V_{R,t}^{1/2} \Psi_{R,t}^{\top} \Sigma_{FC,t}^{-1/2}, \) \( V_{R,t} \) 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 \( \Psi_{R,t}^{\top} \Psi_{R,t} = I, \) and \( \Sigma_{FC,t} \) is defined in (4.3).

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

\[
\sqrt{pTh} \left( \bar{C}_{t,j} - H_{C,t}^{\top} C_{t,j} - B_{C,t,j} \right) \to^{d} N(0, \nu_0 \Sigma_{C,t,j}),
\]

where

\[
B_{C,t,j} = \frac{1}{2} V_{C,t}^{-1} h^2 \mu_2 \sum_{c_1+c_2+c_3+c_4=2} H_{C,t}^{\top} \Sigma_{C,t}^{(c_1)} \Sigma_{FC,t}^{(c_2,c_3)} \Sigma_{FR,t}^{(c_4)} C_{t,j} + o_p(h^2),
\]

and

\[
\Sigma_{C,t,j} = V_{C,t}^{-1} Q_{C,t} \Phi_{C,t,j} Q_{C,t}^{\top} V_{C,t}^{-1}
\]  \( (4.5) \)

\( \Phi_{C,t,j} \) is defined similarly to \( \Phi_{R,t,i} = V_{C,t}^{1/2} \Psi_{C,t}^{\top} \Sigma_{FR,t}^{-1/2}, \) \( V_{C,t} \) 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 \( \Psi_{C,t}^{\top} \Psi_{C,t} = I, \) and \( \Sigma_{FR,t} \) is defined in (4.3).

Theorem 2 establishes the asymptotic normality of the nonparametric estimators \( \tilde{R}_t \) and \( \tilde{C}_t. \) The asymptotic bias \( B_{R,t,i}(B_{C,t,j}) \) 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 \( \tilde{R}_{t,i}. \) is

\[
AMSE_R = \frac{h^4}{4} \mu_2^2 \sum_{c_1+c_2+c_3+c_4=2} V_{R,t}^{-1} H_{R,t}^{\top} \Sigma_{R,t}^{(c_1)} \Sigma_{FC,t}^{(c_2,c_3)} R_{t,i}^{(c_4)} + \frac{\text{Tr}(\nu_0 \Sigma_{R,t,i})}{qTh},
\]  \( (4.6) \)

and hence the optimal bandwidth for \( \tilde{R}_{t,i}. \) is

\[
h_{R}^{opt} = (qT)^{-\frac{1}{2}} \left( \frac{\text{Tr}(\nu_0 \Sigma_{R,t,i})}{\mu_2^2 \sum_{c_1+c_2+c_3+c_4=2} V_{R,t}^{-1} H_{R,t}^{\top} \Sigma_{R,t}^{(c_1)} \Sigma_{FC,t}^{(c_2,c_3)} R_{t,i}^{(c_4)} \right)^{\frac{1}{2}}
\]  \( (4.7) \)

The optimal bandwidth for \( \tilde{C}_{t,i}. \) 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

\[
\hat{F}_t - H_{R,t}^{-1}F_t H_{C,t}^{-1} = O_p\left( \frac{1}{\min(p, q) + h^2} \right),
\]
\[
\hat{S}_{t,ij} - S_{t,ij} = O_p\left( \frac{1}{\min(p, q, \sqrt{pTh}, \sqrt{qTh}) + h^2} \right),
\]

for any \( 1 \leq i \leq p \) and \( 1 \leq j \leq q \).

Theorem 3 derives the convergence rate of the estimated latent factor \( \hat{F}_t \) and signal \( \hat{S}_t \). To achieve consistency, we require dimensions \( p \) and \( q \) approach infinity. That is because we need sufficient information accumulation to distinguish the signal \( S_t \) from the noise \( E_t \) 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 \( \hat{k} \) and \( \hat{r} \) is justified in Theorem 4 below.

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

\[
P(\hat{k} = k) \to 1
\]

and

\[
P(\hat{r} = r) \to 1
\]

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

Theorem 4 derives the consistency of the ratio-based estimator \( \hat{k} \) and \( \hat{r} \). 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 \( R_t \) and \( C_t \), the optimal \( F_t \) is achieved by \( \bar{F}_t = R_t^T Y_t C_t \). By concentrating out \( F_t \), we can reformulate the optimization problem (5.7) as:

\[
\begin{align*}
\text{maximize} & \quad \sum_{s=1}^{T} \left\| R_s^T Y_s^{(t)} C_s C_s^T \right\|_F^2 \\
\text{subject to} & \quad \frac{1}{p} R_t^T R_t = I, \quad \frac{1}{q} C_t^T C_t = I.
\end{align*}
\]

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

$$\tilde{\mathbf{G}}_{R,t} \triangleq \frac{1}{Tpq^2} \sum_{s=1}^T Y_s^{(t)} \mathbf{C}_t \mathbf{C}_t^\top Y_s^{(t)} \mathbf{C}^\top_t.$$

(5.2)

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

$$\tilde{\mathbf{G}}_{C,t} \triangleq \frac{1}{Tp^2q} \sum_{s=1}^T Y_s^{(t)} \mathbf{R}_t \mathbf{R}_t^\top Y_s^{(t)} \mathbf{R}_t^\top \mathbf{C}_t.$$

(5.3)

The projection-based optimal estimator $\mathbf{R}_t^p$ ($\mathbf{C}_t^p$) is comprised of $\sqrt{p}$ ($\sqrt{q}$) times top $k$ ($r$) eigenvectors of $\tilde{\mathbf{G}}_{R,t}$ ($\tilde{\mathbf{G}}_{C,t}$) 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 $\mathbf{R}_t^p$ and $\mathbf{C}_t^p$.

**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

$$\frac{1}{p} \left\| \mathbf{R}_t^p - \mathbf{R}_t \mathbf{H}_{R,t}^p \right\|_F^2 = O_p \left( \frac{1}{p^2 q^2} + \frac{1}{Tqh} + \frac{1}{T^2 p^2 h^2} + h^4 \right),$$

$$\frac{1}{q} \left\| \mathbf{C}_t^p - \mathbf{C}_t \mathbf{H}_{C,t}^p \right\|_F^2 = O_p \left( \frac{1}{p^2 q^2} + \frac{1}{Tph} + \frac{1}{T^2 q^2 h^2} + h^4 \right),$$

where

$$\mathbf{H}_{R,t}^p = \frac{1}{Tpq^2} \sum_{s=1}^T K_h s F_s C_t^\top \mathbf{C}_t \mathbf{C}_t^\top C_t F_s^\top \mathbf{R}_t^\top \mathbf{R}_t \mathbf{V}_{R,t}^{-1},$$

$$\mathbf{H}_{C,t}^p = \frac{1}{Tp^2q} \sum_{s=1}^T K_h s F_s R_t^\top \mathbf{R}_t \mathbf{R}_t^\top R_t F_s^\top C_t^\top \mathbf{C}_t \mathbf{V}_{C,t}^{-1}.$$
Theorem 5 shows that the projection-based estimator $\widehat{R}_i^p$ ($\widehat{C}_i^p$) indeed performs no worse than the local PCA estimator $\hat{R}_i$ ($\hat{C}_i$), while yields faster convergence rate than the local PCA estimator $\hat{R}_i$ ($\hat{C}_i$) when $p = o(Tqh)$ and $qh^4 = o(1)$ (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 $\widehat{R}_i^p$ ($\widehat{C}_i^p$) and $\widehat{R}_i$ ($\hat{C}_i$) 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 $Y_t\text{'s}$ are generated according to model (2.1). The dimension of the latent factor matrix $F_t$ 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 $\text{vec}(F_t)$ from the following Vector Auto-Regressive model of order one (VAR(1) model):

$$\text{vec}(F_t) = \Phi \cdot \text{vec}(F_{t-1}) + \epsilon_t,$$

where the AR coefficient matrix $\Phi = 0.1 \cdot I_4$ and $\text{Var}[(\epsilon_t)] = 0.99 \cdot I_4$. Thus, $\text{Var}[	ext{vec}(F_t)] = I_4$. We simulate noise $E_t$ also from VAR(1),

$$\text{vec}(E_t) = \Gamma \cdot \text{vec}(E_{t-1}) + u_t,$$

where $\Gamma = \psi \cdot I_{pq}$ and $\text{Var}[u_t] = (1 - \psi^2)I_{pq}$. Thus, $\text{Var}[	ext{vec}(E_t)] = I_{pq}$. 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 $R_t$ and $C_t$ are generated as:

- **DGP1**: The first column of $R_t$: $R_{t,1} = R_1$, the second column of $R_t$: $R_{t,2} = R_2 + G_t$; the first column of $C_t$: $C_{t,1} = C_1$, the second column of $C_t$: $C_{t,2} = C_2 + H_t$. The time-invariant loading matrices $R_j$ and $C_j$ are independently sampled from $\mathcal{U}(-1,1)$ for $j = 1, 2$, $G_t = G\left(\frac{t}{T}\right) = 2\left(\frac{t}{T}\right)^2 - 1$ and $H_t = 0.2 \exp\left[-0.7 + 3.5\left(\frac{t}{T}\right)\right]$, following Cai (2007).

- **DGP2**: The first column of $R_t$: $R_{t,1} = R_1 + G_t$, the second column of $R_t$: $R_{t,2} = F(10t/T; 2; 5i/p + 2)$, $i = 1, ..., p$; the first column of $C_t$: $C_{t,1} = C_1 + H_t$, the second column of $C_t$: $C_{t,2} = F(10t/T; 2; 5j/q + 2)$, $j = 1, ..., q$. The time-invariant loading matrices $R_j$ and $C_j$ 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 $R_t$ and $C_t$. 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 $c = \frac{2.345}{\sqrt{k}}$, for estimating $R_t$ and $C_t$ 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

$$D(\hat{A}_t, A_t) = \frac{1}{T} \sum_{t=1}^{T} \left\| \hat{A}_t (A_t^T \hat{A}_t)^{-1} A_t^T - A_t (A_t^T A_t)^{-1} A_t^T \right\|,$$

(6.1)

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

<table>
<thead>
<tr>
<th>DGP</th>
<th>$(p,q)$</th>
<th>$T = 100$</th>
<th>$T = 200$</th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\bar{D}(\hat{R},R)$</td>
<td>$\bar{D}(\hat{C},C)$</td>
<td>$\bar{D}(\hat{R},R)$</td>
</tr>
<tr>
<td>1</td>
<td>(20,20)</td>
<td>1.09(0.21)</td>
<td>0.98(0.12)</td>
<td>0.55(0.07)</td>
</tr>
<tr>
<td></td>
<td>(100,20)</td>
<td>0.98(0.12)</td>
<td>0.55(0.07)</td>
<td>0.40(0.05)</td>
</tr>
<tr>
<td></td>
<td>(100,100)</td>
<td>0.58(0.08)</td>
<td>0.50(0.04)</td>
<td>0.29(0.04)</td>
</tr>
<tr>
<td></td>
<td>(20,20)</td>
<td>1.14(0.22)</td>
<td>1.04(0.17)</td>
<td>1.01(0.12)</td>
</tr>
<tr>
<td></td>
<td>(100,20)</td>
<td>1.01(0.12)</td>
<td>0.58(0.08)</td>
<td>0.75(0.08)</td>
</tr>
<tr>
<td></td>
<td>(100,100)</td>
<td>0.59(0.08)</td>
<td>0.52(0.04)</td>
<td>0.30(0.04)</td>
</tr>
<tr>
<td>2</td>
<td>(20,20)</td>
<td>2.46(0.20)</td>
<td>2.15(0.20)</td>
<td>2.27(0.18)</td>
</tr>
<tr>
<td></td>
<td>(100,20)</td>
<td>2.35(0.20)</td>
<td>1.72(0.17)</td>
<td>2.14(0.14)</td>
</tr>
<tr>
<td></td>
<td>(100,100)</td>
<td>1.91(0.20)</td>
<td>1.47(0.13)</td>
<td>1.73(0.13)</td>
</tr>
<tr>
<td></td>
<td>(20,20)</td>
<td>2.48(0.20)</td>
<td>2.24(0.20)</td>
<td>2.30(0.18)</td>
</tr>
<tr>
<td></td>
<td>(100,20)</td>
<td>2.38(0.19)</td>
<td>1.80(0.17)</td>
<td>2.17(0.14)</td>
</tr>
<tr>
<td></td>
<td>(100,100)</td>
<td>1.93(0.19)</td>
<td>1.55(0.13)</td>
<td>1.75(0.13)</td>
</tr>
</tbody>
</table>

Note: Mean and standard deviation in parentheses of $\bar{D}(\hat{R},R)$ and $\bar{D}(\hat{C},C)$ from 100 iterations.

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

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

17
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 $\Xi_t = C_t \otimes R_t$. In what follows, the subscripts mat and vec denote our approach and Su and Wang (2017)'s method, respectively. The loading space $\widehat{\Xi}_{mat,t}$ is computed as $\widehat{\Xi}_{mat,t} = \widehat{C}_t \otimes \widehat{R}_t$, where $\widehat{C}_t$ and $\widehat{R}_t$ are our nonparametric estimators. For the vector-valued approach, we apply Su and Wang (2017)'s method to the vectorized observations $vec(Y_t), t = 1, 2, \cdots, T$ as in (4.2) to obtain $\widehat{\Xi}_{vec,t}$. The results for the estimation accuracies of $\Xi_t$ measured by $D_{max}$ and $D_{vec}$ 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

<table>
<thead>
<tr>
<th>DGP</th>
<th>$(p,q)$</th>
<th>$D_{max}$</th>
<th>$D_{vec}$</th>
<th>$D_{max}$</th>
<th>$D_{vec}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(10,10)$</td>
<td>0.23(0.06)</td>
<td>0.78(0.10)</td>
<td>0.13(0.02)</td>
<td>0.62(0.10)</td>
</tr>
<tr>
<td></td>
<td>$(20,10)$</td>
<td>0.17(0.04)</td>
<td>0.65(0.11)</td>
<td>0.15(0.02)</td>
<td>0.75(0.07)</td>
</tr>
<tr>
<td></td>
<td>$(20,20)$</td>
<td>0.19(0.03)</td>
<td>0.76(0.09)</td>
<td>0.11(0.02)</td>
<td>0.60(0.08)</td>
</tr>
<tr>
<td></td>
<td>$(10,10)$</td>
<td>0.25(0.07)</td>
<td>0.90(0.09)</td>
<td>0.14(0.03)</td>
<td>0.79(0.12)</td>
</tr>
<tr>
<td></td>
<td>$(20,10)$</td>
<td>0.18(0.05)</td>
<td>0.79(0.13)</td>
<td>0.16(0.03)</td>
<td>0.91(0.06)</td>
</tr>
<tr>
<td></td>
<td>$(20,20)$</td>
<td>0.20(0.04)</td>
<td>0.91(0.06)</td>
<td>0.12(0.02)</td>
<td>0.78(0.12)</td>
</tr>
<tr>
<td>2</td>
<td>$(10,10)$</td>
<td>0.34(0.03)</td>
<td>0.66(0.11)</td>
<td>0.31(0.02)</td>
<td>0.58(0.12)</td>
</tr>
<tr>
<td></td>
<td>$(20,10)$</td>
<td>0.32(0.02)</td>
<td>0.64(0.10)</td>
<td>0.29(0.01)</td>
<td>0.54(0.08)</td>
</tr>
<tr>
<td></td>
<td>$(20,20)$</td>
<td>0.29(0.02)</td>
<td>0.61(0.09)</td>
<td>0.27(0.02)</td>
<td>0.51(0.07)</td>
</tr>
<tr>
<td></td>
<td>$(10,10)$</td>
<td>0.35(0.03)</td>
<td>0.70(0.12)</td>
<td>0.32(0.03)</td>
<td>0.61(0.13)</td>
</tr>
<tr>
<td></td>
<td>$(20,10)$</td>
<td>0.33(0.02)</td>
<td>0.69(0.11)</td>
<td>0.29(0.02)</td>
<td>0.58(0.09)</td>
</tr>
<tr>
<td></td>
<td>$(20,20)$</td>
<td>0.30(0.02)</td>
<td>0.65(0.10)</td>
<td>0.27(0.02)</td>
<td>0.55(0.09)</td>
</tr>
</tbody>
</table>

Note: Mean and standard deviation in parentheses of $\overline{D}_{max}$ and $\overline{D}_{vec}$ 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
Table 3: Estimation of latent dimensions

<table>
<thead>
<tr>
<th>DGP (k, r)</th>
<th>T = 100</th>
<th>T = 200</th>
<th>T = 400</th>
<th>T = 100</th>
<th>T = 200</th>
<th>T = 400</th>
<th>T = 100</th>
<th>T = 200</th>
<th>T = 400</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2, 1)</td>
<td>0.13</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>0.05</td>
<td>0.02</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(2, 2)</td>
<td>0.80</td>
<td>0.90</td>
<td>0.93</td>
<td>0.98</td>
<td>0.97</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>other</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(2, 1)</td>
<td>0.12</td>
<td>0.08</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>0.09</td>
<td>0.04</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(2, 2)</td>
<td>0.77</td>
<td>0.87</td>
<td>0.94</td>
<td>0.98</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>other</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: Relative frequency of estimated rank pair (\(\hat{k}, \hat{r}\)) 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, Bussière 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 \( 24 \times 24 \) 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 (Linne-mann (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.

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.

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


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


Appendix

Appendix A  Proofs

Proof of Theorem 1

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

\[
\frac{1}{pqT} \sum_{s=1}^{T} Y_s^T Y_s^T \hat{R}_t = \hat{R}_t \hat{V}_{R,t}, \quad \text{or} \quad \hat{R}_t = \frac{1}{pqT} \sum_{s=1}^{T} Y_s^T Y_s^T \hat{V}_{R,t}^{-1},
\]

\[
\frac{1}{pqT} \sum_{s=1}^{T} Y_s^T Y_s^T \hat{C}_t = \hat{C}_t \hat{V}_{C,t}, \quad \text{or} \quad \hat{C}_t = \frac{1}{pqT} \sum_{s=1}^{T} Y_s^T Y_s^T \hat{V}_{C,t}^{-1}.
\]

Then we have the following decomposition:

\[
\hat{R}_t - R_t H_{R,t} = \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} Y_s Y_s^T \hat{V}_{R,t}^{-1} - R_t H_t
\]

\[
= \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} \left( R_t F_s C_t^T E_t^T + E_t C_t F_t^T R_t^T + E_t E_t^T + D_{R,st} F_s C_t^T C_s F_s^T C_t^T + R_t F_s D_{C,st} C_s F_s^T C_t^T \right)
\]

\[
+ R_t F_s C_t^T C_t F_t^T D_{R,st} + R_t F_s C_t^T D_{C,st} F_s^T R_t^T + D_{R,st} F_s C_t^T C_s F_s^T D_{R,st} + R_t F_s D_{C,st} C_t F_t^T D_{R,st}
\]

\[
+ D_{R,st} F_s D_{C,st} C_t^T F_t^T R_t^T + R_t F_s D_{C,st} C_t F_t^T D_{R,st} + D_{R,st} F_s D_{C,st} C_s F_s^T D_{R,st}
\]

\[
+ D_{R,st} F_s D_{C,st} C_t F_t^T D_{R,st} + D_{R,st} F_s D_{C,st} C_s F_s^T D_{R,st}
\]

\[
= \sum_{j=1}^{24} I_j \hat{R}_t \hat{V}_{R,t}^{-1}
\]

(A.1)

where \( D_{R,st} = R_s - R_t \) and \( D_{C,st} = C_s - C_t \). Theorem 1 follows by Lemma 1 below. \( \square \)

Lemma 1. Under Assumption 1 - 6, we have for any \( 1 \leq t \leq T \),

(a) \( \frac{1}{p} \| I_1 \hat{R}_t \|_F^2 = O_p \left( \frac{1}{q Th} \right) \);

(b) \( \frac{1}{p} \| I_2 \hat{R}_t \|_F^2 = O_p \left( \frac{1}{q Th} \right) \);
Proof of Lemma 1

Proof. (a) First note that

\[
\frac{1}{p} \left\| \sum_{s=1}^{T} K_{h,st} F_s C_t^T E_s^T \right\|_F^2 = \frac{1}{p} \left\| \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} F_s C_t, j e_s, j \right\|_F^2
\]

\[
= q T \left\| \frac{1}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} (e_s, j \otimes F_s) C_t, j \right\|_F^2
\]

\[
= q T \sum_{i=1}^{p} \left\| \frac{1}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} e_s, ij F_s C_t, j \right\|_F^2
\]

\[= \mathcal{O}_p \left( pq Th^{-1} \right), \]

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

\[
\frac{1}{p} \left\| \tilde{R}_1 \right\|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| R_1 \right\|_F^2 \left\| \sum_{s=1}^{T} K_{h,st} F_s C_t^T E_s^T \right\|_F^2 \left\| \bar{R}_1 \right\|_F^2 = \mathcal{O}_p \left( \frac{1}{q Th} \right). \]

(b) Similarly, we have

\[
\frac{1}{p} \left\| \tilde{R}_2 \right\|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} K_{h,st} E_s C_t F_s^T \right\|_F^2 \left\| R_1^+ \right\|_F^2 \left\| \bar{R}_1 \right\|_F^2 = \mathcal{O}_p \left( \frac{1}{q Th} \right). \]
(c) First note that

\[ \left\| \sum_{s=1}^{T} K_{h,st} E_s E_s^\top \right\|_F^2 = \left\| \sqrt{qT} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} e_{s,j} e_{s,j}^\top \right\|_F^2 \]

\[ \leq \frac{p^2 q T}{p^2} \sum_{i=1}^{p} \sum_{l=1}^{p} \left\| \frac{1}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} (e_{s,ij} e_{s,lj} - \mathbb{E}[e_{s,ij} e_{s,lj}]) \right\|^2 \]

\[ + pq^2 T^2 \cdot \frac{1}{p} \sum_{i=1}^{p} \left( \left\| \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} e_{s,ij} e_{s,lj} \right\|_F \right)^2 \]

\[ = O_p \left( p^2 q T h^{-1} \right) + O \left( pq^2 T^2 \right), \]

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

\[ \frac{1}{p} \left\| I_{13} \tilde{R}_t \right\|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} K_{h,st} E_s E_s^\top \right\|_F^2 \left\| \tilde{R}_t \right\|_F^2 = O_p \left( \frac{1}{q T h} \right) + O \left( \frac{1}{p} \right). \]

(d) The proofs for \( \frac{1}{p} \left\| I_{1j} \tilde{R}_t \right\|_F^2 = O_p \left( h^j \right), j = 4, \cdots, 13, \) are tedious. To save space, we only provide the proof for \( j = 4. \) Other proofs are rather similar. Define \( M_{R,s} = R_s - R_t - \left( \tilde{z}_{-t}^\top \right) R_t^{(1)} - \frac{1}{2} \left( \tilde{z}_{-t}^\top \right) R_t^{(2)}. \)

\[ \frac{1}{p} \left\| I_{14} \tilde{R}_t \right\|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} R_t^{(1)} F_s C_t^\top C_t F_s^\top R_t^\top \tilde{R}_t \right\|_F^2 \]

\[ + \frac{1}{p^3 q^2 T^2} \left( \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^2 \right) K_{st} R_t^{(2)} F_s C_t^\top C_t F_s^\top R_t^\top \tilde{R}_t \right\|_F^2 \]

\[ + \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} M_{R,s} K_{st} F_s C_t^\top C_t F_s^\top R_t^\top \tilde{R}_t \right\|_F^2 \]

\[ = \frac{1}{p^3 q^2 T^2} \sum_{j=1}^{3} \left\| I_{14j} \right\|_F^2. \]  

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

\[
\|I_{41}\|_F^2 \leq \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} R_{1s}^{(1)} (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t}) R_t^T \overline{R}_t \right\|_F^2 + \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} R_{1s}^{(1)} \Sigma_{FC,t} R_t^T \overline{R}_t \right\|_F^2
\]

\[
\leq \left\| R_{1s}^{(1)} \right\|_F^2 \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t}) \right\|_F^2 \left\| R_t^T \right\|_F^2 \| \overline{R}_t \|_F^2
\]

\[
+ \left\| R_{1s}^{(1)} \right\|_F^2 \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} \right\|_F \left\| \Sigma_{FC,t} \right\|_F \left\| R_t^T \right\|_F \| \overline{R}_t \|_F^2
\]

\[
\triangleq I_{411} + I_{412}.
\]

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

\[
E \left[ \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t}) \right]^2
\]

\[
= \frac{1}{T^2} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^2 K_{st}^2 E (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t})^2
\]

\[
+ \frac{1}{T^2} \sum_{s_1=1}^{T} \sum_{s_2=1}^{T} \left( \frac{s_1-t}{T} \right) \left( \frac{s_2-t}{T} \right) K_{s_1 t} K_{s_2 t} E (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t}) (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t})
\]

\[
= T^{-1} h \int u^2 k^2 (u) du E (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t})^2
\]

\[
+ T^{-1} h \sum_{j=1}^{T} \int (u + \frac{j}{Th}) u k(u) k\left(u + \frac{j}{Th}\right) du \alpha^{1/2} (j) \left[ E (F_s C_t^T C_t F_s^T/q - \Sigma_{FC,t})^\gamma \right]^{1/\gamma} + T^{-1} h
\]

\[
= \mathcal{O}(T^{-1} h),
\]

and hence \( I_{411} = \mathcal{O}(p^3 q^2 T h) \). For \( I_{412} \), we have

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

and hence \( I_{412} = \mathcal{O}(p^3 q^2) \), where we have used the uniform approximation of Riemann summation to a definite integral. Therefore, \( \frac{1}{p^3 q^2 T} \|I_{41}\|_F^2 = \mathcal{O}_p\left(\frac{h}{T}\right) = o_p\left(h^4\right) \).
as $Th^3 \to \infty$. For the second term in (A.2), we have

$$
\|I_{421}\|^2_F \leq \left\| \frac{1}{T} \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^2 K_{st} R_i^{(2)} (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) R_i^T \bar{R}_i \right\|^2_F + \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^2 K_{st} R_i^{(2)} \Sigma_{FC,t} R_i^T \bar{R}_i \right\|^2_F \\
\leq \left\| R_i^{(2)} \right\|^2_F \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^2 K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) \right\|^2_F + \left\| R_i^{(2)} \right\|^2_F \left\| \Sigma_{FC,t} \right\|_F \left\| R_i^T \right\|^2_F \left\| \bar{R}_i \right\|^2_F \\
\triangleq I_{421} + I_{422}. \quad (A.4)
$$

Similar to $I_{411}$, we have

$$
E \left[ \frac{1}{T} \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^2 K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) \right]^2 \\
= \frac{1}{4T^2} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{st}^2 E (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t})^2 \\
+ \frac{1}{4T^2} \sum_{s_1=1}^{T} \sum_{s_2=1}^{T} \left( \frac{s_1-t}{T} \right)^2 \left( \frac{s_2-t}{T} \right)^2 K_{s_1} K_{s_2} E (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) \\
= \frac{h^3}{4T} \int u^4 k^2(u) du E (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t})^2 \\
+ \frac{h^3}{4T} \sum_{j=1}^{T} \left( u + \frac{j}{Th} \right)^2 u^2 k(u) k \left( u + \frac{j}{Th} \right) du \alpha^{1-\frac{2}{\gamma}} (j) \left[ E (F_s C_t^T C_t F_s^T / q - \Sigma_{FC,t}) \right]^{1/\gamma} + T^{-1} h \\
= O(T^{-1} h^3),
$$

and hence $I_{421} = O\left(p^3 q^2 T^3 h^3\right)$. For $I_{422}$, we have

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

and hence $I_{422} = O\left(p^3 q^2 T^2 h^4\right)$, where we have used the uniform approximation of the Riemann summation to a definite integral. Therefore, $\frac{1}{p^3 q^2 T^2} \|I_{42}\|^2_F = O\left(h^4\right)$ as $Th^3 \to \infty$. Finally, by the fact that $p^{-1} \|h^{-2} M_{RS}\|^2_F = p^{-1} \|h^{-2} M_R \left( \frac{1}{Th} + uh \right)\|^2_F \to 0$ for any $u$, we have $\frac{1}{p^3 q^2 T^2} \|I_{43}\|^2_F = o_p\left(h^4\right)$. Combining these three terms, we have
\[ \frac{1}{p} \| I_{14} \hat{R}_t \|_F^2 = O_p(h^4). \]

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

\[ \frac{1}{p} \| I_{14} \|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{st} R_t F_s C_t (1)^\tau C_t (1)^\top F_s^\top R_t^\top R_t \right\|_F^2 + \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{st} R_t F_s C_t (1)^\tau C_t (1)^\top F_s^\top R_t^\top R_t \right\|_F^2 + \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{st} R_t F_s C_t (2)^\tau C_t (1)^\top F_s^\top R_t^\top R_t \right\|_F^2 + o_p \left( h^8 \right) \]

The order of \( \frac{1}{p} \| I_{14} \|_F^2 \), \( j = 15, 16, 17, 18 \) can be derived in a similar way. Next we consider \( I_{19} \). By the Taylor expansion, we have

\[ \left\| \sum_{s=1}^{T} K_{h,st} D_{R,st} F_s C_t E_s^\top \right\|_F^2 \leq q T \left\| R_t^{(1)} \right\|_F^2 \left\| \frac{1}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} \left( \frac{s-t}{T} \right) K_{h,st} \left( e_{s,j} \otimes F_s \right) C_t \right\|_F^2 + q T \left\| R_t^{(2)} \right\|_F^2 \left\| \frac{1}{2 \sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} \left( \frac{s-t}{T} \right)^2 K_{h,st} \left( e_{s,j} \otimes F_s \right) C_t \right\|_F^2 + q T \left\| \frac{1}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} M_{R,s} K_{h,st} \left( e_{s,j} \otimes F_s \right) C_t \right\|_F^2 = O_p \left( p^2 q Th \right) + O_p \left( p^3 q Th^3 \right) + o_p \left( p^3 q Th^3 \right), \]

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

\[ \frac{1}{p} \| I_{19} \hat{R}_t \|_F^2 \leq \frac{1}{p^3 q^2 T^2} \left\| \sum_{s=1}^{T} K_{h,st} D_{R,st} F_s C_t E_s^\top \right\|_F^2 \left\| \hat{R}_t \right\|_F^2 = O_p \left( \frac{h}{\sqrt{qT}} \right) = o_p \left( h^4 \right). \]

The order of \( \frac{1}{p} \| I_{19} \hat{R}_t \|_F^2 \), \( j = 20, 21, 22, 23, 24 \) can be derived in a similar way as \( I_{19} \).
Before proving Theorem 2, we first study the asymptotic behavior of \( \tilde{V}_{R,t}, \tilde{V}_{C,t}, H_{C,t} \) and \( H_{R,t} \).

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

\[
\begin{align*}
\tilde{V}_{R,t} & \xrightarrow{p} V_{R,t} \\
\tilde{V}_{C,t} & \xrightarrow{p} V_{C,t},
\end{align*}
\]

\[
\|\tilde{V}_{R,t}\|_2 = O_p(1) \quad \text{and} \quad \|\tilde{V}_{R,t}^{-1}\|_2 = O_p(1),
\]

where \( V_{R,t} \) is the diagonal matrix consisting of the eigenvalues of \( \Sigma_{FC,t}^{1/2} \Omega_{R,t} \Sigma_{FC,t}^{1/2} \) and \( V_{C,t} \) is the diagonal matrix consisting of the eigenvalues of \( \Sigma_{FR,t}^{1/2} \Omega_{C,t} \Sigma_{FR,t}^{1/2} \). Covariance \( \Sigma_{FC,t} = E h_{Ft} C_t^{\top} C_t q F_t^{\top} \), \( \Sigma_{FR,t} = E h_{Ft} R_t^{\top} R_t F_t^{\top} \). Matrices \( \Omega_{R,t} \) and \( \Omega_{C,t} \) are defined in Assumption 3.

**Proof.** From \( \frac{1}{pqT} \sum_{s=1}^{T} K_{t,s} Y_s Y_s^{\top} \tilde{R}_t = \tilde{R}_t \tilde{V}_{R,t} \) and \( \frac{1}{p} \tilde{R}_t \tilde{R}_t^{\top} = I_k \), we have

\[
\tilde{V}_{R,t} = \frac{1}{p} \tilde{R}_t^{\top} \left( \frac{1}{pqT} \sum_{s=1}^{T} K_{t,s} Y_s Y_s^{\top} \right) \tilde{R}_t
\]

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

\[
\tilde{M}_{R,t} = \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} R_s F_s C_t^{\top} C_t F_s^{\top} R_t^{\top} + \sum_{j=1}^{24} I_j,
\]

where \( I_j, j = 1, 2, ..., 24 \) are defined in equation A.1. Applying Lemma 1, we have

\[
\begin{align*}
\|I_1\| & \leq \frac{1}{pqT} \|R_t\| \left\| \sum_{s=1}^{T} K_{h,st} F_s C_t^{\top} F_s^{\top} \right\|_F = O_p \left( \frac{1}{\sqrt{qT}} \right), \\
\|I_2\| & \leq \frac{1}{pqT} \|R_t\| \left\| \sum_{s=1}^{T} K_{h,st} E_s C_t^{\top} F_s^{\top} \right\|_F = O_p \left( \frac{1}{\sqrt{qT}} \right), \\
\|I_3\| & \leq \frac{1}{pqT} \left\| \sum_{s=1}^{T} K_{h,st} E_s F_s^{\top} \right\|_F = O_p \left( \frac{1}{\sqrt{T}} + \frac{1}{\sqrt{qT}} \right), \\
\|I_4\| & \leq \frac{1}{pqT} \left\| \sum_{s=1}^{T} K_{h,st} D_{R,st} F_s C_t^{\top} C_t F_s^{\top} \right\|_F = O_p \left( h^4 \right).
\end{align*}
\]

(A.5)
And similarly, $\|l_j\| = O_p(h^4)$, for $j = 5, \ldots, 13$, and $\|l_j\| = o_p(h^4)$, for $j = 14, \ldots, 24$. On the other hand, we have

$$
\left\| \frac{1}{pqT} R_t \sum_{s=1}^{T} K_{h,st} F_s C_t^T C_t F_s^T R_t^T - \mathbb{E} \left[ \frac{1}{pqT} R_t \sum_{s=1}^{T} F_s C_t^T C_t F_s^T R_t^T \right] \right\| \\
\leq \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,st} \left( F_s (C_t^T C_t/q) F_s^T - \mathbb{E} \left[ F_s (C_t^T C_t/q) F_s^T \right] \right) \right\| \cdot \|R_t\|^2 / p \\
= O_p \left( \frac{1}{\sqrt{Th}} \right).
$$

Together, we have

$$
\left\| \tilde{M}_{R,t} - \mathbb{E} \left[ \frac{1}{pqT} R_t \sum_{s=1}^{T} F_s C_t^T C_t F_s^T R_t^T \right] \right\| = o_p(1). 
$$

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

$$
|\tilde{V}_{R,t,i} - V_{R,t,i}| = o_p(1), \quad \text{for } 1 \leq i \leq k,
$$

and $\tilde{V}_{R,t} \xrightarrow{p} V_{R,t}$. Further we have the first $k$ eigenvalues of $\frac{1}{pqT} R_t \sum_{s=1}^{T} K_{h,st} F_s C_t^T C_t F_s^T R_t^T$ are bounded away from both zero and infinity. Thus, $\|\tilde{V}_{R,t}\|_2 = O_p(1)$ and $\|\tilde{V}_{R,t}^{-1}\|_2 = O_p(1)$. Results for $\tilde{V}_{C,t}$ are obtained in a similar fashion.

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

$$
\|H_{R,t}\| = O_p(1), \quad \text{and} \quad \|H_{C,t}\| = O_p(1).
$$

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

$$
\|H_{R,t}\| = \left\| \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} F_s C_t^T C_t F_s^T R_t R_t^{-1} \tilde{V}_{R,t}^{-1} \right\| \\
\leq \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,st} F_s (C_t^T C_t/q) F_s^T \right\| \cdot \|R_t\| \|\tilde{R}_t\| / p \|\tilde{V}_{R,t}^{-1}\| = O_p(1),
$$

$$
\|H_{C,t}\| = \left\| \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} F_s^T R_t R_t F_s C_t^T \tilde{C}_t \tilde{V}_{C,t}^{-1} \right\| \\
\leq \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,st} F_s^T (R_t^T R_t/p) F_s \right\| \cdot \|C_t\| \|\tilde{C}_t\| / q \|\tilde{V}_{C,t}^{-1}\| = O_p(1).
$$

**Proof of Theorem 2**
Proof. We make use of the following equality for each row of equation (A.1): for each row vector $R_{t,i} \in \mathbb{R}^k$, $i \in [p]$, $t \in [T]$, we have

$$\begin{align*}
\overline{R}_{t,i} - H_{R_{t,i}}^T R_{t,i} &= \overline{V}_{R,t}^{-1} \sum_{s=1}^{T} \frac{1}{pqT} K_{h,st} \overline{R}_{t}^T \left( E_s C_t F^T R_{t,i} + R_t F_s C_t^T E_{s,i} + E_s E_{s,i} + R_t F_s C_t^T F_s^T D_{R,st,i} \right) \\
&+ R_t F_s C_t^T D_{C,st} F_s^T R_{t,i} + D_{R,st} F_s C_t^T C_t F_s^T R_{t,i} + R_t F_s D_{C,st}^T C_t F_s^T R_{t,i} \\
&+ D_{R,st} F_s C_t^T D_{C,st} F_s^T R_{t,i} + R_t F_s D_{C,st}^T C_t F_s^T R_{t,i} + R_t F_s D_{C,st}^T C_t F_s^T D_{R,st,i} \\
&+ D_{R,st} F_s D_{C,st}^T C_t F_s^T D_{R,st,i} + R_t F_s D_{C,st}^T C_t F_s^T D_{R,st,i} \\
&+ E_s D_{C,st} F_s^T R_{t,i} + D_{R,st} F_s C_t^T D_{C,st} F_s^T D_{R,st,i} + R_t F_s D_{C,st}^T D_{R,st,i} + D_{R,st} F_s C_t^T E_{s,i} \\
&+ D_{R,st} F_s D_{C,st}^T C_t E_{s,i} + D_{R,st} F_s D_{C,st}^T E_{s,i} \\
&\Delta \sum_{j=1}^{24} \Pi_j \overline{V}_{R,t}^{-1}
\end{align*}$$

(A.6)

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

$$\sqrt{qT/h} \left( \overline{R}_{t,i} - H_{R_{t,i}}^T R_{t,i} - B_{R_{t,i}} \right) = \overline{V}_{R,t}^{-1} \frac{\overline{R}_{t}^T R_t}{p} \sum_{s=1}^{T} K_{h,st} F_s C_t^T e_{s,i} + o_p\left(1\right) \xrightarrow{D} \mathcal{N}\left(0, \nu_0 \Sigma_{R_{t,i}}\right),$$

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 $\widehat{C}_t$ can be derived in a similar way.

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

(a) $\Pi_1 = O_p\left(\frac{1}{\delta_{pqTh} \sqrt{qT/h}}\right)$,

(b) $\Pi_2 = O_p\left(\frac{1}{\sqrt{qT/h}}\right)$,

(c) $\Pi_3 = O_p\left(\frac{1}{\delta_{pqTh} \sqrt{qT/h}}\right) + O_p\left(\frac{1}{\delta_{pqTh} \sqrt{p}}\right)$,

(d) $\sum_{j=1}^{13} \Pi_j = \frac{1}{2} h^2 \mu_2 \sum_{c_1+c_2+c_3+c_4=2} H_{R_{t,i}}^T \Sigma_{F_{C,t}} (c_1) \Sigma_{F_{C,t}} (c_2,c_3) R_{t,i}^{(c_4)} + o_p(h^2);$

\[9\]
(e) \( \Pi_j = o_p\left(h^2\right), j = 14, \ldots, 24. \)

Proof.  
(a) 
\[
\Pi_1 = \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,sl} \hat{R}_{t,l} e_{s,lj} C_{t,j}^T F_s^T R_{t,i}.
\]
\[
= \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,sl} (\hat{R}_{t,l} - H_{R,t} R_{t,l} e_{s,lj} C_{t,j}^T F_s^T R_{t,i}.
+ H_{R,t}^T \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,sl} R_{t,l} e_{s,lj} C_{t,j}^T F_s^T R_{t,i}.
\]
\[
= \Pi_{11} + \Pi_{12}.
\]

We bound each term as follows.
\[
\|\Pi_{11}\| \leq \frac{1}{\sqrt{qTh}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \hat{R}_{t,l} - H_{R,t} R_{t,l} \right\|^2 \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,sl} e_{s,lj} C_{t,j}^T F_s^T \right\| \right)^{1/2} \left\| R_{t,i} \right\|
\]
\[
= \frac{1}{\sqrt{qTh}} \cdot O_p\left( \frac{1}{\delta_{pqTh}} \right) \cdot O_p(1),
\]
where the last equality results from Theorem 1 and Lemma 11. We also have
\[
\|\Pi_{12}\| = \frac{1}{\sqrt{pqTh}} \left\| H_{R,t}^T \left( \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,sl} R_{t,l} e_{s,lj} C_{t,j}^T F_s^T \right) R_{t,i} \right\| = O_p\left( \frac{1}{\sqrt{pqTh}} \right),
\]
where the last equality results from by Markov Theorem, Lemma 12 (b) and Proposition 2.

(b) 
\[
\Pi_2 = \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} K_{h,sl} \hat{R}_{t,l} e_{s,ij} F_s C_{t,j} e_{s,ij}
\]
\[
= \frac{1}{pqT} \sum_{l=1}^{p} \left( \hat{R}_{t,l} - H_{R,t} R_{t,l} \right) R_{t,l}^T \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,sl} F_s C_{t,j} e_{s,ij}
+ \frac{1}{pqT} \sum_{l=1}^{p} H_{R,t} R_{t,l} R_{t,l}^T \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,sl} F_s C_{t,j} e_{s,ij}
\]
\[
= \Pi_{21} + \Pi_{22}.
\]
We bound each term as follows.

\[
\|\Pi_{21}\| \leq \frac{1}{\sqrt{qT}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l} - H_{R,t}^\top R_{t,l} \right\| \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l} \right\|^2 \right)^{1/2}
\]

\[
\leq \frac{1}{\sqrt{qTh}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l} - H_{t,R}^\top R_{t,l} \right\|^2 \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l} \right\|^2 \cdot \left\| \sqrt{\frac{h}{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} F_s C_{t,j} e_{s,ij} \right\| \right)^{2/1}
\]

\[
= O_p \left( \frac{1}{\delta_{pqTh}\sqrt{qTh}} \right) \quad \text{by Theorem 1 and Lemma 11.}
\]

Similarly,

\[
\|\Pi_{22}\| = \frac{1}{\sqrt{qTh}} \left\| \frac{1}{p} \sum_{l=1}^{p} H_{R,t}^\top R_{t,l} R_{t,l}^\top \right\| \cdot \left\| \sqrt{\frac{h}{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} F_s C_{t,j} e_{s,ij} \right\|
\]

\[
= O_p \left( \frac{1}{\sqrt{qTh}} \right).
\]

Combing all the terms, we have

\[
\|\Pi_2\| = O_p \left( \frac{1}{\sqrt{qTh}} \right).
\]

(c)

\[
\Pi_3 = \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,st} (\tilde{R}_{t,l} - H_{R,t}^\top R_{t,l}) (e_{s,ij}e_{s,ij} - \mathbb{E}[e_{s,ij}])
\]

\[
+ \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,st} (\tilde{R}_{l,l} - H_{R,t}^\top R_{t,l}) \mathbb{E}[e_{s,ij}]
\]

\[
+ \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,st} H_{R,t}^\top R_{l,l} (e_{s,ij}e_{s,ij} - \mathbb{E}[e_{s,ij}])
\]

\[
+ \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,st} H_{R,l}^\top R_{t,l} \mathbb{E}[e_{s,ij}]
\]

\[
= \Pi_{31} + \Pi_{32} + \Pi_{33} + \Pi_{34}
\]
We bound each term as follows.

\[
\|I_{31}\| \leq \frac{1}{\sqrt{qT}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right\|^2 \right)^{1/2} \cdot \left( \frac{1}{p} \sum_{l=1}^{p} \left( \frac{1}{qT} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} (e_{s,lj} e_{s,ij} - \mathbb{E}[e_{s,lj} e_{s,ij}]) \right)^2 \right)^{1/2} \\
= O_p \left( \frac{1}{\delta_{pqTh} \sqrt{qT}} \right) \text{ by Lemma 13 and Markov inequality.}
\]

\[
\|I_{32}\| \leq \frac{1}{\sqrt{p}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right\|^2 \right)^{1/2} \cdot \left( \frac{1}{p} \sum_{l=1}^{p} \left( \frac{1}{qT} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} \mathbb{E}[e_{s,lj} e_{s,ij}] \right)^2 \right)^{1/2} \\
= O_p \left( \frac{1}{\delta_{pqTh} \sqrt{p}} \right)
\]

\[
\|I_{33}\| \leq \left\| \frac{1}{pqT} \sum_{s=1}^{T} \sum_{l=1}^{p} \sum_{j=1}^{q} K_{h,st} R_{t,l} (e_{s,lj} e_{s,ij} - \mathbb{E}[e_{s,lj} e_{s,ij}]) \right\| \cdot \|H_{R,t}\| \\
= O_p \left( \frac{1}{\sqrt{pqT}} \right) \text{ by Lemma 12 (a) and Proposition 2.}
\]

\[
\|I_{34}\| \leq \frac{1}{p} \cdot \|H_{R,t}^T\| \cdot \sum_{l=1}^{p} \left( \frac{1}{qT} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} \mathbb{E}[e_{s,lj} e_{s,ij}] \right) \cdot \max_{l \in [p]} \|R_{t,l}\| \\
= O(\frac{1}{p}) \text{ by Assumption 3 and Proposition 2.}
\]

Combining the result on each term, we obtain

\[
\|I_3\| = O_p \left( \frac{1}{\delta_{pqTh} \sqrt{qT}} \right) + O_p \left( \frac{1}{\delta_{pqTh} \sqrt{p}} \right).
\]
\[
\Pi_4 = \frac{1}{pqT} \sum_{l=1}^{p} (\bar{R}_{t,l} - H_{R,t}^T R_{t,l}^T) R_{t,l}^T \sum_{s=1}^{T} \left(\frac{s-t}{T}\right) K_{h,st} F_s C_t^T C_t F_s R_{t,i}^{(1)} \\
+ \frac{1}{pqT} \sum_{l=1}^{p} (\bar{R}_{t,l} - H_{R,t}^T R_{t,l}^T) R_{t,l}^T \sum_{s=1}^{T} \frac{1}{2} \left(\frac{s-t}{T}\right)^2 K_{h,st} F_s C_t^T C_t F_s R_{t,i}^{(2)} \\
+ \frac{1}{pqT} \sum_{l=1}^{p} H_{R,t}^T R_{t,l} R_{t,l}^T \sum_{s=1}^{T} \left(\frac{s-t}{T}\right) K_{h,st} F_s C_t^T C_t F_s R_{t,i}^{(1)} \\
+ \frac{1}{pqT} \sum_{l=1}^{p} H_{R,t}^T R_{t,l} R_{t,l}^T \sum_{s=1}^{T} \frac{1}{2} \left(\frac{s-t}{T}\right)^2 K_{h,st} F_s C_t^T C_t F_s R_{t,i}^{(2)} \\
+ \frac{1}{pqT} \sum_{l=1}^{p} H_{R,t}^T R_{t,l} R_{t,l}^T \sum_{s=1}^{T} K_{h,st} F_s C_t^T C_t F_s M_{R,s,i} \\
= \sum_{j=1}^{6} \Pi_{4j}.
\] (A.8)

We bound each term as follows.

\[
\|\Pi_{41}\| \leq \left\{ \frac{1}{p} \sum_{l=1}^{p} \| \bar{R}_{t,l} - H_{R,t}^T R_{t,l} \|^{2} \right\}^{1/2} \left\{ \frac{1}{p} \sum_{l=1}^{p} \| R_{t,l} \|^{2} \right\}^{1/2} \left\{ \frac{1}{T} \sum_{s=1}^{T} \left(\frac{s-t}{T}\right) K_{h,st} F_s C_t^T C_t F_s \right\} \| R_{t,i}^{(1)} \| \\
= O_p \left( \frac{1}{\delta_{pqTh}} \right) O_p \left( 1 \right) O_p \left( \sqrt{h/T} \right) O_p \left( 1 \right) \\
= o_p \left( h^2 \right).
\]

\[
\|\Pi_{42}\| \leq \left\{ \frac{1}{p} \sum_{l=1}^{p} \| \bar{R}_{t,l} - H_{R,t}^T R_{t,l} \|^{2} \right\}^{1/2} \left\{ \frac{1}{p} \sum_{l=1}^{p} \| R_{t,l} \|^{2} \right\}^{1/2} \left\{ \frac{1}{2T} \sum_{s=1}^{T} \left(\frac{s-t}{T}\right)^2 K_{h,st} F_s C_t^T C_t F_s \right\} \| R_{t,i}^{(2)} \| \\
= O_p \left( \frac{1}{\delta_{pqTh}} \right) O_p \left( 1 \right) O_p \left( h^2 \right) O_p \left( 1 \right) \\
= o_p \left( h^2 \right).
\]
\[ \| \Pi_{43} \| \leq \left( \frac{1}{p} \sum_{l=1}^{p} \| R_{t,l} - H_{R,t}^{\top} R_{t,l} \|^{2} \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \| R_{t,l} \|^{2} \right)^{1/2} \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,s} F_{s} C_{t}^{\top} C_{t} \frac{F_{s} M_{R,s,i}}{q} \right\| \\
= O_p \left( \frac{1}{\delta_{pq} \theta h} \right) O_p(h^2) O_p(1) \\
= o_p(h^2). \]

\[ \Pi_{44} = \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{h,s} \Sigma_{FC,t}^{(0,0)} R_{t,i}^{(1)} \\
+ \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{h,s} \left[ F_{s} C_{t}^{\top} C_{t} \frac{F_{s} - \Sigma_{FC,t}^{(0,0)}}{q} \right] R_{t,i}^{(1)} \\
= O_p(1) O_p \left( \frac{1}{T} \right) + O_p(1) O_p \left( \sqrt{\frac{h}{T}} \right) \\
= o_p(h^2). \]

\[ \Pi_{45} = \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{2T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^{2} K_{h,s} \Sigma_{FC,t}^{(0,0)} R_{t,i}^{(2)} \\
+ \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{2T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^{2} K_{h,s} \left[ F_{s} C_{t}^{\top} C_{t} \frac{F_{s} - \Sigma_{FC,t}^{(0,0)}}{q} \right] R_{t,i}^{(2)} \\
= \frac{1}{2} h^2 M_{h^2} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^{2} \left[ F_{s} C_{t}^{\top} C_{t} \frac{F_{s} - \Sigma_{FC,t}^{(0,0)}}{q} \right] R_{t,i}^{(2)} + O_p \left( \frac{h^2}{TH} \right) + O_p \left( \frac{h^2}{\sqrt{TH}} \right) \\
= \frac{1}{2} h^2 M_{h^2} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^{2} R_{t,i}^{(2)} + o_p(h^2). \]

\[ \Pi_{46} = \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} K_{h,s} \Sigma_{FC,t}^{(0,0)} M_{R,s,i} \\
+ \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} K_{h,s} \left[ F_{s} C_{t}^{\top} C_{t} \frac{F_{s} - \Sigma_{FC,t}^{(0,0)}}{q} \right] M_{R,s,i} \\
= h^2 M_{h^2} \left( \frac{R_{t}^{\top} R_{t}}{p} \right) \int k(u) \Sigma_{FC,t}^{(0,0)} h^{-2} M_{R,uhT+t,i} \ d u + o_p(h^2) \\
= o_p(h^2). \]
Combing the result on each term, we obtain

\[ II_4 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (0,0) R_{t,i}^{(2)} + o_P\left(h^2\right). \]

Similarly, we obtain

\[ II_5 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (0,2) R_{t,i} + o_P\left(h^2\right), \]

\[ II_6 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (2) \sum_{FC,t} (0,0) R_{t,i} + o_P\left(h^2\right), \]

\[ II_7 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (2,0) R_{t,i} + o_P\left(h^2\right), \]

\[ II_8 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (1) \sum_{FC,t} (0,0) R_{t,i}^{(1)} + o_P\left(h^2\right), \]

\[ II_9 = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (1,0) R_{t,i}^{(1)} + o_P\left(h^2\right), \]

\[ II_{10} = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (1,1) R_{t,i}^{(1)} + o_P\left(h^2\right), \]

\[ II_{11} = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (1) \sum_{FC,t} (0,1) R_{t,i} + o_P\left(h^2\right), \]

\[ II_{12} = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (1) \sum_{FC,t} (1,0) R_{t,i}^{(1)} + o_P\left(h^2\right), \]

\[ II_{13} = \frac{1}{2} h^2 \mu_2 H_{R,t}^\top \sum_{c} (0) \sum_{FC,t} (0,1) R_{t,i}^{(1)} + o_P\left(h^2\right). \]

Combining all these terms, we have

\[ \sum_{j=4}^{13} II_j = \frac{1}{2} h^2 \mu_2 \sum_{c_1+c_2+c_3+c_4=2} H_{R,t}^\top \sum_{c_1} (c_1) \sum_{c_2} (c_2, c_3) \sum_{c_4} (c_4) R_{t,i}^{(1)} + o_P\left(h^2\right). \]
We bound each term as follows.

\[
\|I_{14}\| = \frac{1}{pqT} \sum_{l=1}^{p+1} \left( \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right) R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{h,s} F_s C_t^{(1)T} C_t^{(1)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} \left( \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right) R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(2)T} C_t^{(1)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} \left( \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right) R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(1)T} C_t^{(2)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} \left( \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \right) R_{t,l}^{(2)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(1)T} C_t^{(1)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} H_{R,t}^T R_{t,l} R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{h,s} F_s C_t^{(1)T} C_t^{(1)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} H_{R,t}^T R_{t,l} R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(2)T} C_t^{(1)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} H_{R,t}^T R_{t,l} R_{t,l}^{(1)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(1)T} C_t^{(2)} F_s R_{t,i} \\
+ \frac{1}{pqT} \sum_{l=1}^{p+1} H_{R,t}^T R_{t,l} R_{t,l}^{(2)T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,s} F_s C_t^{(1)T} C_t^{(1)} F_s R_{t,i} + o_p(h^4) \\
= \sum_{j=1}^{8} I_{14j} + o_p(h^4).
\]

We bound each term as follows.

\[
\|I_{14j}\| \leq \left( \frac{1}{p} \sum_{l=1}^{p} \| \tilde{R}_{t,l} - H_{R,t}^T R_{t,l} \| \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \| R_{t,l}^{(1)} \| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{h,s} F_s C_t^{(1)T} C_t^{(1)} F_s \right) \| R_{t,i} \| \\
= O_p \left( \frac{1}{pqTh} \right) O_p(1) O_p \left( \sqrt{\frac{h^5}{T}} \right) O_p(1) \\\n= o_p(h^2).
\]
\[ \|I_{142}\| \leq \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l}^\top - H_{R,t}^\top R_{t,l} \right\|^2 \right)^{1/2} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| R_{t,l}^{(1)} \right\|^2 \right)^{1/2} \left\| \frac{1}{2T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,st} F_s \frac{C_t^{(2)} C_t^{(1)}}{q} F_s \right\| R_{t,l} \right\| \\
= O_p \left( \frac{1}{\delta_{pqTh}} \right) O_p (1) O_p \left( \sqrt{\frac{h^4}{T}} \right) O_p (1) \\
= o_p (h^2). \]

Similarly, we have \( \|I_{143}\| = o_p (h^2) \) and \( \|I_{144}\| = o_p (h^2) \). Then

\[ II_{145} = H_{R,t}^\top \left( \frac{R_t^\top R_t^{(1)}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{h,st} \Sigma_{FC,t}^{(1)} R_{t,l}. \]

\[ + H_{R,t}^\top \left( \frac{R_t^\top R_t^{(1)}}{p} \right) \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^3 K_{h,st} \left[ F_s \frac{C_t^{(1)} C_t^{(1)}}{q} F_s - \Sigma_{FC,t}^{(1)} \right] R_{t,l}. \]

\[ = O_p (1) O_p \left( \frac{h^2}{T} \right) + O_p (1) O_p \left( \sqrt{\frac{h^4}{T}} \right) \\
= o_p (h^2). \]

\[ II_{146} = H_{R,t}^\top \left( \frac{R_t^\top R_t^{(1)}}{p} \right) \frac{1}{2T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,st} \Sigma_{FC,t}^{(2)} R_{t,l}. \]

\[ + H_{R,t}^\top \left( \frac{R_t^\top R_t^{(1)}}{p} \right) \frac{1}{2T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^4 K_{h,st} \left[ F_s \frac{C_t^{(2)} C_t^{(1)}}{q} F_s - \Sigma_{FC,t}^{(2)} \right] R_{t,l}. \]

\[ = \frac{1}{2} h^4 \mu_4 H_{R,t}^\top \Sigma_{FC,t}^{(1)} \Sigma_{FC,t}^{(2)} R_{t,l} + O_p \left( \frac{h^4}{T} \right) + O_p \left( \frac{h^4}{\sqrt{T h}} \right) \\
= o_p (h^2). \]

Similarly, we can obtain \( II_{147} = o_p (h^2) \) and \( II_{148} = o_p (h^2) \). Combining these term, we have \( II_{14} = o_p (h^2) \). By very similar but tedious derivation, we can obtain \( II_j = o_p (h^2) \) for \( j = 15, \ldots, 24 \).

\[ \Box \]

**Proposition 3.** Under Assumption 1-6,

\[ \text{plim}_{p,q,T \to \infty} \frac{R_t^\top R_t}{p} = Q_{R,t}, \quad \text{and} \quad \text{plim}_{p,q,T \to \infty} \frac{C_t^\top C_t}{q} = Q_{C,t}. \]
The matrix $Q_{R,t} \in \mathbb{R}^{k \times k}$ and $Q_{C,t} \in \mathbb{R}^{r \times r}$ are given, respectively, by

$$Q_{R,t} = V_{R,t}^{1/2} \Psi_{R,t}^{T} \Sigma_{FC,t}^{1/2} \text{ and } Q_{C,t} = V_{C,t}^{1/2} \Psi_{C,t}^{T} \Sigma_{FR,t}^{1/2},$$

where $V_{R,t}$ ($V_{C,t}$) is a diagonal matrix with diagonal entries being the eigenvalues of $\Sigma_{FC,t}^{1/2} \Omega_{R,t} \Sigma_{FC,t}^{1/2}$ (in decreasing order), $\Psi_{R,t}$ ($\Psi_{C,t}$) is the corresponding eigenvector matrix such that $\Psi_{R,t}^{T} \Psi_{R,t} = I$ ($\Psi_{C,t}^{T} \Psi_{C,t} = I$), and $\Omega_{R,t}$ ($\Omega_{C,t}$) is defined in Assumption 3.

**Proof.** Let $X_{s,t} = K_{h,st} F_{s} C_{t}^{T}$, multiply the identify $\frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} Y_{s} Y_{s}^{T} \widetilde{R}_{t} = \widetilde{R}_{t} V_{R,t}$ on both sides by $\frac{1}{p} \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} R_{t}^{T}$ to obtain:

$$\frac{1}{p} \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} R_{t}^{T} \frac{1}{pqT} \sum_{s=1}^{T} K_{h,st} Y_{s} Y_{s}^{T} \widetilde{R}_{t} = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} \frac{R_{t}^{T} \widetilde{R}_{t}}{p} V_{R,t}.$$

Expanding $Y_{s} Y_{s}^{T}$, we can rewrite the above as

$$\left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} \frac{R_{t}^{T} \widetilde{R}_{t}}{p} V_{R,t} = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} \frac{R_{t}^{T} \widetilde{R}_{t}}{p} \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} \frac{R_{t}^{T} \widetilde{R}_{t}}{p} + d_{t},$$

(A.9)

where

$$d_{t} = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} \right)^{1/2} \frac{1}{p} \left( \sum_{j=1}^{24} I_{j} \widetilde{R}_{t} \right).$$

We have

$$\frac{1}{qT} \sum_{s=1}^{T} X_{s,t} X_{s,t}^{T} = \frac{1}{T} \sum_{s=1}^{T} K_{h,st} \left( F_{s}(C_{t}^{T} C_{t}/q) F_{s}^{T} - \mathbb{E} \left[ F_{s}(C_{t}^{T} C_{t}/q) F_{s}^{T} \right] \right) + \frac{1}{T} \sum_{s=1}^{T} K_{h,st} \mathbb{E} \left[ F_{s}(C_{t}^{T} C_{t}/q) F_{s}^{T} \right] = O_{p}(1),$$

and $d_{t} = o_{p}(1)$ following the proof of Lemma 2.

We rewrite equation (A.9) as

$$B_{t} V_{R,t} = (A_{t} + d_{t} B_{t}^{-1}) B_{t},$$

18
where

$$A_t = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t}X_{s,t}^\top \right)^{1/2} \frac{R_t^\top R_t}{p} \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t}X_{s,t}^\top \right)^{1/2},$$

$$B_t = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t}X_{s,t}^\top \right)^{1/2} \frac{R_t^\top R_t}{p}.$$

Each column of $B_t$ is an eigenvector of the matrix $(A_t + d_t B_t^{-1})$. By Proposition 1, we have

$$B_t^\top B_t \xrightarrow{p} V_{R,t},$$

(A.10)

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

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

$$\mathcal{V}_{R,t} \xrightarrow{p} V_{R,t}.$$  

(A.11)

Let

$$\mathcal{A}_{R,t} = B_t \mathcal{V}_{R,t}^{-1/2},$$

(A.12)

then $\|\mathcal{A}_{R,t}\| = 1$ and

$$\mathcal{A}_{R,t} \mathcal{V}_{R,t} = (A_t + d_t B_t^{-1}) \mathcal{A}_{R,t},$$

that is, each column of $\mathcal{A}_{R,t}$ is an eigenvector of $A_t + d_t B_t^{-1}$. And we have $A_t \xrightarrow{p} \Sigma_{FC,t}^{1/2} \Omega_{R,t} \Sigma_{FC,t}^{1/2}$ and $d_t B_t^{-1} = \mathcal{O}_p(1)$. By eigenvector perturbation theory and Assumption 7, there exists a unique eigenvector matrix $\mathcal{A}_{R,t}$ of $\Sigma_{FC,t}^{1/2} \Omega_{R,t} \Sigma_{FC,t}^{1/2}$ such that $\|\mathcal{A}_{R,t} - \mathcal{A}_{R,t}\| = \mathcal{O}_p(1)$. From (A.12) and (A.11), we have

$$\frac{R_t^\top R_t}{p} = \left( \frac{1}{qT} \sum_{s=1}^{T} X_{s,t}X_{s,t}^\top \right)^{-1/2} \mathcal{A}_{R,t}^\top \mathcal{A}_{R,t}^* \xrightarrow{p} \Sigma^{-1/2}_{FC,t} \mathcal{A}_{R,t} \mathcal{V}_{R,t}^{1/2}.$$  

Proof of Theorem 3

□

19
Proof. Under the assumption that $\frac{1}{p} \bar{R}_t^\top \bar{R}_t = I_p$ and $\frac{1}{q} \bar{C}_t^\top \bar{C}_t = I_q$, we have

$$\begin{align*}
\bar{F}_t - H_{R,t}^{-1} F_t H_{C,t}^{-1}^\top &= \frac{1}{pq} \bar{R}_t^\top (R_t - \bar{R}_t H_{R,t}^{-1}) F_t (C_t - \bar{C}_t H_{C,t}^{-1})^\top \bar{C}_t \\
&+ \frac{1}{p} \bar{R}_t^\top (R_t - \bar{R}_t H_{R,t}^{-1}) F_t H_{C,t}^{-1}^\top \\
&+ \frac{1}{q} H_{R,t}^{-1} F_t (C_t - \bar{C}_t H_{C,t}^{-1})^\top \bar{C}_t \\
&+ \frac{1}{pq} (\bar{R}_t - R_t H_{R,t})^\top E_t (\bar{C}_t - C_t H_{C,t}) \\
&+ \frac{1}{pq} (\bar{R}_t - R_t H_{R,t})^\top E_t C_t H_{C,t} \\
&+ \frac{1}{pq} \bar{R}_t^\top E_t (\bar{C}_t - C_t H_{C,t}) \\
&+ \frac{1}{pq} H_{R,t}^{-1} R_t^\top E_t C_t H_{C,t} \\
&= \sum_{i=1}^7 \text{III}_i.
\end{align*}$$

Since $\frac{1}{\sqrt{pq}} \| R_t - \bar{R}_t H_{R,t}^{-1} \| = o_p(1)$ and $\frac{1}{\sqrt{pq}} \| C_t - \bar{C}_t H_{C,t}^{-1} \| = o_p(1)$, 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$_7$.

$$\text{III}_2 = \frac{1}{p} (\bar{R}_t - R_t H_{R,t})^\top (R_t - \bar{R}_t H_{R,t}^{-1}) F_t H_{C,t}^{-1}^\top + \frac{1}{p} H_{R,t}^{-1} R_t^\top (R_t - \bar{R}_t H_{R,t}^{-1}) F_t H_{C,t}^{-1}^\top$$

$$= O_p \left( \frac{1}{\delta_{pq Th}^2} + h^2 \right),$$

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

$$\text{III}_3 = O_p \left( \frac{1}{\gamma_{pq Th}^2} + h^2 \right), \quad \text{III}_5 = O_p \left( \frac{1}{\delta_{pq Th}^2} + h^2 \right), \quad \text{III}_6 = O_p \left( \frac{1}{\gamma_{pq Th}^2} + h^2 \right),$$

where $\gamma_{pq Th} = \min \{ \sqrt{q}, \sqrt{p Th} \}$. Finally, we have

$$\bar{F}_t - H_R^{-1} F_t H_C^{-1}^\top = O_p \left( \frac{1}{\min(p, q)} + h^2 \right),$$

20
where we uses results in Lemma 5 below. Next we consider the convergence of $\tilde{S}_{t,ij}$. Define $\tilde{R}_{t,i} = H_{R,t} R_{t,i}$, $\tilde{C}_{t,j} = H_{C,t} C_{t,j}$, and $\tilde{F}_t = H_{R,t}^{-1} F_t H_{C,t}^{-1}$, we have

\[
\tilde{S}_{t,ij} - S_{t,ij} = (R_{t,i} - \tilde{R}_{t,i})^\top (F_t - \tilde{F}_t) (C_{t,j} - \tilde{C}_{t,j}) + (R_{t,i} - \tilde{R}_{t,i})^\top \tilde{F}_t (C_{t,j} - \tilde{C}_{t,j}) \\
+ \tilde{R}_{t,i}^\top (F_t - \tilde{F}_t) (C_{t,j} - \tilde{C}_{t,j}) + (R_{t,i} - \tilde{R}_{t,i})^\top (F_t - \tilde{F}_t) \tilde{C}_{t,j} \\
+ \tilde{R}_{t,i}^\top \tilde{F}_t (C_{t,j} - \tilde{C}_{t,j}) + (R_{t,i} - \tilde{R}_{t,i})^\top \tilde{F}_t \tilde{C}_{t,j} + \tilde{R}_{t,i}^\top (F_t - \tilde{F}_t) \tilde{C}_{t,j}.
\]

Dominant terms are the last three terms. Note that $e_F t = O_p(1)$, $e_C t, j = O_p(1)$ and $e_R t, i = O_p(1)$. From Theorem 2, we have

\[
\tilde{R}_{t,i} - R_{t,i} = O_p \left( \frac{1}{\min(p, \sqrt{qTh})} + h^2 \right), \quad \text{and} \quad \tilde{C}_{t,j} - C_{t,j} = O_p \left( \frac{1}{\min(q, \sqrt{pTh})} + h^2 \right).
\]

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

\[
\tilde{S}_{t,ij} - S_{t,ij} = O_p \left( \frac{1}{\min(p, q, \sqrt{qTh}, \sqrt{pTh})} + h^2 \right).
\]

\[
\text{Lemma 3. Under Assumption 1-7, the } k \times k \text{ matrix}
\]

\[
\frac{1}{p} (\tilde{R}_t - R_t H_{R,t})^\top R_t = O_p \left( \frac{1}{\delta^2_{pqTh}} + h^2 \right);
\]

\[
The \ r \times r \ matrix
\]

\[
\frac{1}{q} (\tilde{C}_t - C_t H_{C,t})^\top C_t = O_p \left( \frac{1}{\gamma^2_{pqTh}} + h^2 \right).
\]

\[
\text{Proof. Using the identity (A.1), we have}
\]

\[
\frac{1}{p} (\tilde{R}_t - R_t H_{R,t})^\top R_t = \tilde{V}_{R,t}^{-1} \tilde{R}_t \left( \sum_{j=1}^{24} I_j \right)^\top R_t,
\]

where $I_{t,j} = 1, \ldots, 24$ are defined in (A.1).
We begin with the first term in (A.1),

$$
\bar{R}_t I_t^T R_t = \frac{1}{p^2 q T} \sum_{s=1}^{T} K_{h,s} \left( \bar{R}_t - R_t H_{R,t} \right)^T E_s C_t F_s^T R_t^T + \frac{1}{p^2 q T} \sum_{s=1}^{T} K_{h,s} H_{R,t}^T R_t^T E_s C_t F_s^T R_t^T R_t = I_{11} + I_{12}.
$$

Note

$$
\|I_{11}\| \leq \frac{1}{\sqrt{q T h}} \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \bar{R}_{t,l} - H_{R,t,l}^T \right\|^2 \right)^{1/2} \cdot \left( \frac{1}{p} \sum_{l=1}^{p} \left\| \sqrt{h} E_s C_t F_s^T \right\| \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,s} C_{t,j}^T e_{s,l,j} F_s \right)^{1/2} \cdot \frac{1}{p} \|R_t^T R_t\|
$$

and

$$
\|I_{12}\| \leq \frac{1}{\sqrt{p q T h}} \left\| H_{R,t} \right\| \left\| \sqrt{h} E_s C_t F_s^T \right\| \sum_{s=1}^{T} K_{h,s} F_s C_t^T E_s^T R_t \cdot \frac{1}{p} \|R_t^T R_t\|
$$

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

$$
\bar{R}_t I_2^T R_t = \frac{1}{\sqrt{p q T h}} \cdot \frac{1}{p} \bar{R}_t^T R_t \cdot \sqrt{h} E_s C_t F_s^T \sum_{s=1}^{T} K_{h,s} F_s C_t^T E_s R_t = O_p \left( \frac{1}{\sqrt{p q T h}} \right),
$$

where we use Proposition 3 and Lemma 12.

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

$$
\bar{R}_t I_3^T R_t = \frac{1}{p^2 q T} \sum_{s=1}^{T} K_{h,s} \left( \bar{R}_t - R_t H_{R,t} \right)^T E_s E_s^T - \mathbb{E} \left[ E_s E_s^T \right] R_t + \frac{1}{p^2 q T} \sum_{s=1}^{T} K_{h,s} H_{R,t}^T R_t^T E_s E_s^T - \mathbb{E} \left[ E_s E_s^T \right] R_t + \frac{1}{p^2 q T} \sum_{s=1}^{T} K_{h,s} H_{R,t}^T R_t^T E_s E_s^T - \mathbb{E} \left[ E_s E_s^T \right] R_t
$$

$$
= I_{31} + I_{32} + I_{33} + I_{34}.
$$
Note

\[
\|I_{31}\|^2 \leq \frac{1}{qT} \frac{1}{p} \sum_{l=1}^{p} \left\| \mathbf{R}_{t,l} - \mathbf{H}_{R,t} \mathbf{R}_{t,l} \right\| \| \frac{1}{p} \sum_{l=1}^{p} \frac{1}{p} \sum_{i=1}^{p} \left( \frac{\sqrt{h}}{\sqrt{qT}} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} (e_{s,lj} e_{s,ij} - \mathbb{E}[e_{s,lj} e_{s,ij}]) \right) \right\|^2 \frac{1}{p} \sum_{i=1}^{p} \| \mathbf{R}_{t,i} \| \| \mathbf{R}_{t,i} \| \\
= \mathcal{O}_p \left( \frac{1}{qT} \right) \cdot \mathcal{O}_p \left( \frac{1}{\delta_{pqT}} + h^4 \right) \cdot \mathcal{O}_p (1), \quad \text{using Theorem 1 and Lemma 13.}
\]

\[
\|I_{32}\|^2 \leq \frac{1}{p} \sum_{l=1}^{p} \left\| \mathbf{R}_{t,l} - \mathbf{H}_{t,R} \mathbf{R}_{t,l} \right\| \| \frac{1}{p} \sum_{l=1}^{p} \frac{1}{p} \sum_{i=1}^{p} \left( \frac{1}{qT} \sum_{s=1}^{T} \sum_{j=1}^{q} K_{h,st} \mathbb{E}[e_{s,lj} e_{s,ij}] \right) \right\|^2 \frac{1}{p} \sum_{i=1}^{p} \| \mathbf{R}_{t,i} \| \| \mathbf{R}_{t,i} \| \\
= \frac{1}{p} \cdot \mathcal{O}_p \left( \frac{1}{\delta_{pqT}} + h^4 \right) \cdot \mathcal{O}(1) \quad \text{using Theorem 1.}
\]

\[
\|I_{33}\| \leq \frac{1}{\sqrt{pqT}} \left( \frac{1}{\sqrt{p}} \left\| \left\| \frac{\sqrt{h}}{\sqrt{qT}} \sum_{s=1}^{T} K_{h,st} (\mathbb{E}[E_s E_s^T - \mathbb{E}[E_s E_s^T]]) \right\| \right\| \right) \cdot \| \mathbf{H}_{R,t} \| \| \frac{1}{p} \| \mathbf{R}_{t} \|^2 \\
= \mathcal{O}_p \left( \frac{1}{\sqrt{pqT}} \right),
\]

\[
\|I_{34}\|^2 \leq \frac{1}{p^2} \left\| \mathbb{E} \left[ \frac{1}{qT} \sum_{s=1}^{T} K_{h,st} E_s E_s^T \right] \right\|^2 \frac{1}{p^2} \| \mathbf{R}_{t} \|^4 \| \mathbf{H}_{R,t} \|^2 \\
= \mathcal{O}_p \left( \frac{1}{p^2} \right).
\]

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

\[
\mathbf{R}_t^\top \mathbf{I}_4^\top \mathbf{R}_t = \frac{1}{pT} \cdot \frac{1}{p^2} \mathbf{R}_t^\top \mathbf{R}_t \sum_{s=1}^{T} K_{h,st} \left( \frac{1}{q} F_s C_{t}^\top C_{t} F_s^\top - \Sigma_{FC,t}^{(0,0)} \right) \mathbf{D}_{R,st}^\top \mathbf{R}_t \\
+ \frac{1}{pT} \cdot \frac{1}{p} \mathbf{R}_t^\top \mathbf{R}_t \sum_{s=1}^{T} K_{h,st} \Sigma_{FC,t}^{(0,0)} \mathbf{D}_{R,st}^\top \mathbf{R}_t \\
= \mathcal{O}_p \left( \frac{1}{p} \right) \cdot \mathcal{O}_p \left( \frac{1}{p^2} \right).
\]

Similarly, we can obtain \( \mathbf{R}_t^\top \mathbf{I}_4^\top \mathbf{R}_t = \mathcal{O}_p (h^2) + \mathcal{O}_p \left( \sqrt{\frac{h}{T}} \right), j = 5, \ldots, 13. \) Next, we consider the
The order of the fourteenth term in (A.1),

\[
\left\| \hat{R}_i^\top I_{14}^\top R_i \right\|^2 \leq \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{(s-t)^3}{T} K_{h,s,t} \hat{R}_i^\top R_i^{(1)} F_s C_t^{(1)} C_t^{(1)} F_s^\top R_i^\top R_i \right\|^2 \\
+ \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^4 K_{h,s,t} \hat{R}_i^\top R_i^{(1)} F_s C_t^{(2)} C_t^{(1)} F_s^\top R_i^\top R_i \right\|^2 \\
+ \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^4 K_{h,s,t} \hat{R}_i^\top R_i^{(2)} F_s C_t^{(1)} C_t^{(1)} F_s^\top R_i^\top R_i \right\|^2 \\
+ \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^4 K_{h,s,t} \hat{R}_i^\top R_i^{(2)} F_s C_t^{(1)} C_t^{(1)} F_s^\top R_i^\top R_i \right\|^2 + o_p \left( h^8 \right) \\
= O_p \left( h^5 T^{-1} \right) + O_p \left( h^8 \right) + O_p \left( h^8 \right) + o_p \left( h^8 \right) \\
= o_p \left( h^4 \right).
\]

The order of \( \hat{R}_i^\top I_{19}^\top R_t, j = 15, \ldots, 18 \) can be derived in a similar fashion. Last, we consider the nineteenth term in (A.1),

\[
\left\| \hat{R}_i^\top I_{19}^\top R_i \right\|^2 \leq \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{(s-t)}{T} K_{h,s,t} \hat{R}_i^\top E_s C_t F_s^\top R_i^{(1)} R_t \right\|^2 \\
+ \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} \frac{1}{2} \left( \frac{s-t}{T} \right)^2 K_{h,s,t} \hat{R}_i^\top E_s C_t F_s^\top R_i^{(2)} R_t \right\|^2 \\
+ \frac{1}{p^4 q^2 T^2} \left\| \sum_{s=1}^{T} K_{h,s,t} \hat{R}_i^\top E_s C_t F_s^\top M_{R,s}^\top R_t \right\|^2 \\
= O_p \left( h^p q^{-1} T^{-1} \right) + O_p \left( h^3 p^{-2} q^{-1} T^{-1} \right) + o_p \left( h^3 p^{-2} q^{-1} T^{-1} \right) \\
= o_p \left( h^4 \right).
\]

The order of \( \hat{R}_i^\top I_{19}^\top R_t, j = 20, \ldots, 24 \) can be derived in a similar fashion as \( I_{19} \). Combining all these together, we obtain \( \frac{1}{p} \left( \hat{R}_t - R_t H_{R,t} \right)^\top R_t = O_p \left( \frac{1}{\delta^2_{pq T h}} + h^2 \right) \). The order of \( \frac{1}{q} \left( \hat{C}_t - C_t H_{C,t} \right)^\top C_t \) can be derived analogously.

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

\[
\frac{1}{pq} \left( \hat{R}_t - R_t H_{R,t} \right)^\top E_t C_t = O_p \left( \frac{1}{\delta^2_{pq T h}} + h^2 \right)
\]

24
\[ \frac{1}{pq} R_t^\top E_t \left( \bar{C}_t - C_t H_{C,t} \right) = \mathcal{O}_p \left( \frac{1}{\sqrt{pqT}} + h^2 \right) \]

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

\[ \frac{1}{pq} \left( R_t - R_t H_{R,t} \right)^\top E_t C_t = \tilde{V}_{R,t}^{-1} R_t^\top \left( \sum_{j=1}^{24} I_j \right)^\top E_t C_t, \quad (A.14) \]

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

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

\[ \left\| \frac{1}{pq} R_t^\top I_1^\top E_t C_t \right\| \leq \left\| \frac{1}{p^2q^2T} \sum_{s=1}^{T} K_{h,st} \left( R_t - R_t H_{R,t} \right)^\top E_s C_t F_s^\top R_t^\top E_t C_t \right\| 
+ \left\| \frac{1}{p^2q^2T} \sum_{s=1}^{T} K_{h,st} H_{R,t} R_t^\top E_s C_t F_s^\top R_t^\top E_t C_t \right\| 
\leq \frac{1}{\sqrt{pqTh}} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h,st} \left\| R_t - R_t H_{R,t} \right\| \left\| E_s C_t F_s^\top \right\| \frac{1}{pq} \right) 
+ \frac{1}{\sqrt{pqTh}} \left\| H_{R,t} \right\| \left\| \sum_{s=1}^{T} K_{h,st} R_t^\top E_s C_t F_s^\top \left\| R_t^\top E_t C_t \right\| \right\| 
= \mathcal{O}_p \left( \frac{1}{\delta_{pqTh} \sqrt{qTh}} \right) + \mathcal{O}_p \left( \frac{1}{\sqrt{pqTh}} \right), \]

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

\[
\left\| \frac{1}{pq} \tilde{R}_i^T I_3^T E_t C_t \right\| \leq \left\| \frac{1}{p^2 q^2 T} \sum_{s=1}^{T} K_{h,st} \left( \tilde{R}_i - R_i H_{R,t} \right)^T E_s E_s^T E_t C_t \right\| \\
+ \left\| \frac{1}{p^2 q^2 T} \sum_{s=1}^{T} K_{h,st} H_{R,t}^T R_i^T E_s E_s^T E_t C_t \right\| \\
\leq \frac{1}{pq} \frac{1}{\sqrt{qTh}} \left\| \tilde{R}_i - R_i H_{R,t} \right\| \left\| \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} K_{h,st} \left( E_s E_s^T - \mathbb{E}\left[ E_s E_s^T \right] \right) \right\| \left\| E_t C_t \right\| \\
+ \frac{1}{pq} \frac{1}{\sqrt{p}} \left\| \tilde{R}_i - R_i H_{R,t} \right\| \left\| \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} K_{h,st} \left( E_s E_s^T - \mathbb{E}\left[ E_s E_s^T \right] \right) \right\| \left\| E_t C_t \right\| \\
+ \frac{1}{pq} \left\| H_{R,t} \right\| \frac{1}{\sqrt{p}} \left\| R_i \right\| \left\| \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} K_{h,st} \left( E_s E_s^T - \mathbb{E}\left[ E_s E_s^T \right] \right) \right\| \left\| E_t C_t \right\| \\
+ \frac{1}{pq} \left\| H_{R,t} \right\| \frac{1}{\sqrt{p}} \left\| R_i \right\| \left\| \sqrt{\frac{h}{pqT}} \sum_{s=1}^{T} K_{h,st} \left( E_s E_s^T - \mathbb{E}\left[ E_s E_s^T \right] \right) \right\| \left\| E_t C_t \right\| \\
= O_p\left( \frac{1}{\delta_{pqTh} \sqrt{pqTh}} \right) + O_p\left( \frac{1}{\delta_{pqTh}} \right) + O_p\left( \frac{1}{\sqrt{pqTh}} \right) + O_p\left( \frac{1}{p} \right).
\]

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

\[
\frac{1}{pq} \tilde{R}_i^T I_4^T E_t C_t = \frac{1}{p^2 q^2 T} \sum_{s=1}^{T} K_{h,st} \left( \tilde{R}_i - R_i H_{R,t} \right)^T R_i \left( \frac{F_s C_i^T C_s^T F_i^T}{q} - \Sigma_{FC,t} \right) D_{R,st}^T E_t C_t \\
+ \frac{1}{p^2 q^2 T} \sum_{s=1}^{T} K_{h,st} H_{R,t}^T R_i^T R_i \left( \frac{F_s C_i^T C_s^T F_i^T}{q} - \Sigma_{FC,t} \right) D_{R,st}^T E_t C_t \\
+ \frac{1}{p^2 q^2 T} \sum_{s=1}^{T} K_{h,st} H_{R,t}^T R_i^T \left( \frac{F_s C_i^T C_s^T F_i^T}{q} - \Sigma_{FC,t} \right) D_{R,st}^T E_t C_t \\
= I_{41} + I_{42} + I_{43} + I_{44}.
\]
Then we bound each term.

\[
\|I_{41}\| \leq \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t R_{H_{R,t}} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) \right\| \left\| \frac{1}{pq} R_t^{(1)^T} E_t C_t \right\|
\]

\[
+ \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t R_{H_{R,t}} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^2 K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) \right\| \left\| \frac{1}{pq} R_t^{(2)^T} E_t C_t \right\|
\]

\[
+ \frac{1}{pq} \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t H_{R,t} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) M_{R,s}^T \right\| \| E_t C_t \|
\]

\[
= O_p \left( \frac{\sqrt{h}}{\delta_{pqTh} \sqrt{T}} \right) + O_p \left( \frac{\sqrt{h^3}}{\delta_{pqTh} \sqrt{T}} \right) + o_p \left( \frac{\sqrt{h^3}}{\delta_{pqTh} \sqrt{T}} \right).
\]

\[
\|I_{42}\| \leq \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t R_{H_{R,t}} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{h,st} \right\| \left\| \Sigma_{F_{C,t}} \right\| \left\| \frac{1}{pq} R_t^{(1)^T} E_t C_t \right\|
\]

\[
+ \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t H_{R,t} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^2 K_{h,st} \right\| \left\| \Sigma_{F_{C,t}} \right\| \left\| \frac{1}{pq} R_t^{(2)^T} E_t C_t \right\|
\]

\[
+ \frac{1}{pq} \frac{1}{\sqrt{p}} \left\| \tilde{R}_t - R_t H_{R,t} \right\| \frac{1}{\sqrt{p}} \| R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} K_{h,st} \Sigma_{F_{C,t}} M_{R,s}^T \right\| \| E_t C_t \|
\]

\[
= O_p \left( \frac{1}{\delta_{pqTh} T} \right) + O_p \left( \frac{h^2}{\delta_{pqTh}} \right) + o_p \left( \frac{h^2}{\delta_{pqTh}} \right).
\]

\[
\|I_{43}\| \leq \| H_{R,t} \| \frac{1}{p} \| R_t^T R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right) K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) \right\| \left\| \frac{1}{pq} R_t^{(1)^T} E_t C_t \right\|
\]

\[
+ \| H_{R,t} \| \frac{1}{p} \| R_t^T R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} \left( \frac{s-t}{T} \right)^2 K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) \right\| \left\| \frac{1}{pq} R_t^{(2)^T} E_t C_t \right\|
\]

\[
+ \frac{1}{pq} \| H_{R,t} \| \frac{1}{p} \| R_t^T R_t \| \left\| \frac{1}{T} \sum_{s=1}^{T} K_{st} (F_s C_t^T C_t F_s^T / q - \Sigma_{F_{C,t}}) M_{R,s}^T \right\| \| E_t C_t \|
\]

\[
= O_p \left( \sqrt{\frac{h}{T}} \right) + O_p \left( \sqrt{\frac{h^3}{T}} \right) + o_p \left( \sqrt{\frac{h^3}{T}} \right).
\]

Combining these together, we have \( \frac{1}{pq} \tilde{R}_t^T I_{41}^T E_t C_t = O_p \left( h^2 \right) \). Similarly, we can show that \( \frac{1}{pq} \tilde{R}_t^T I_j^T E_t C_t = O_p \left( h^2 \right) \) for \( j = 5, \ldots, 13 \) and \( \frac{1}{pq} \tilde{R}_t^T I_j^T E_t C_t = o_p \left( h^2 \right) \) for \( j = 14, \ldots, 24 \).

\[\square\]

**Lemma 5.** Under Assumption 1-7,

\[ \frac{1}{pq} H_{R,t}^T R_t^T E_t C_t H_{C,t} = O_p \left( \frac{1}{\sqrt{pq}} \right). \]
Proof. Firstly, by Lemma 13 (b) and Markov inequality, we have \( \frac{1}{\sqrt{pq}} R_t^T E_t C_t = O_p(1) \). Combining results in Proposition 2, we have
\[
\frac{1}{pq} H_{R,t}^T R_t^T E_t C_t H_{C,t} = \frac{1}{\sqrt{pq}} H_{K,t}^T \left( \frac{1}{\sqrt{pq}} R_t^T E_t C_t \right) H_{C,t} = O_p \left( \frac{1}{\sqrt{pq}} \right).
\]

Proof of Theorem 4

Proof. Let \( \Gamma_i = \frac{1}{T} \sum_{t=1}^{T} \frac{x_{i,t}}{\binom{i_{i+1,t}}{1}} \). By Lemma 6 below, we have \( \Gamma_i = O_p(1) \) uniformly for \( i = 1, \cdots, k-1 \). We also have \( \Gamma_i = O_p(1) \), uniformly for \( i = k+1, \cdots, p-1 \) and \( \Gamma_k \to \infty \) by Lemma 7. Then
\[
\mathbb{P}(k \leq k) = \mathbb{P}(\Gamma_k > \max(\Gamma_{k+1}, \Gamma_{k+2}, \cdots, \Gamma_{k_{max}})) \to 1
\]
\[
\mathbb{P}(k \geq k) = \mathbb{P}(\Gamma_k > \max(\Gamma_{1}, \Gamma_{2}, \cdots, \Gamma_{k-1})) \to 1.
\]
Therefore, we have \( \mathbb{P}(\hat{k} = k) \to 1 \).

Let \( \tilde{\lambda}_{j,t} = \lambda_j \left( \frac{1}{pq} T \sum_{s=1}^{T} k_{h,st} Y_s Y_s^T \right) \).

Lemma 6. For \( j = 1, \cdots, k-1 \),
\[
\frac{1}{T} \sum_{t=1}^{T} \tilde{\lambda}_{j,t} = O_p(1).
\]
Proof. The result follows immediate from Proposition 1.

Lemma 7. For \( j = k, \cdots, p-1 \),
\[
c + o_p(1) \leq \min \left( \sqrt{qT h, p} \right) \tilde{\lambda}_{j,t} \leq C + o_p(1),
\]
where \( c, C, \) and \( o_p(1) \) are uniform in \( k \leq j \leq p-1 \) and \( 1 \leq t \leq T \).

Proof. Let \( X_s = F_s C_s^T \).

Then we have
\[
\frac{1}{pqT} \sum_{s=1}^{T} k_{h,st} Y_s Y_s^T = \frac{1}{pqT} \sum_{s=1}^{T} k_{h,st} \left( R_s + E_s X_s^T \left( X_s X_s^T \right)^{-1} \right) X_s X_s^T \left( R_s + E_s X_s^T \left( X_s X_s^T \right)^{-1} \right)^T
\]
\[
+ \frac{1}{pqT} \sum_{s=1}^{T} k_{h,st} E_s \left( I - X_s^T \left( X_s X_s^T \right)^{-1} X_s \right) E_s^T.
\]
From Weyl inequality and the fact that $X_s$ is a $k \times q$ matrix,

$$
\lambda_{k+j} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} Y_s Y_s^T \right) \geq \lambda_{k+j} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_s \left( I - X_s^T \left( X_s X_s^T \right)^{-1} X_s \right) E_s^T \right) \\
= \lambda_{k+j} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_t \left( I - X_s^T \left( X_s X_s^T \right)^{-1} X_s \right) E_s^T \right) \\
+ \lambda_{k+1} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_s X_s \left( X_s X_s^T \right)^{-1} X_s E_s^T \right) \\
\geq \lambda_{2k+j} \left( \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_s E_s^T \right).
$$

By Lemma 13 and Assumption 4, we have

$$
\left\| \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_s E_s^T - \mathbb{E} \left[ \frac{1}{pq} \sum_{s=1}^{T} K_{h, st} E_s E_s^T \right] \right\| = O_p \left( \sqrt{q T h} \right),
$$

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

$$
c + o_p(1) \leq \min \left( \sqrt{q T h}, p \right) \tilde{\lambda}_{j,t} \leq C + o_p(1).
$$

**Proof of Theorem 5**

**Proof.** Recall that our estimator $\hat{R}_i (\hat{C}_i)$ is given by the matrix of $\sqrt{p} (\sqrt{q})$ times the top $k (r)$
eigenvectors of $\hat{G}_{R,t}$ ($\hat{G}_{C,t}$) defined in equation (5.2) in descending order by corresponding eigenvalues. Similar to A.1, we have the following decomposition:

$$\hat{R}_t^p - R_tH_{R,t}^p = \frac{1}{pq^2}T \sum_{s=1}^{T} K_{h,st} Y_s \hat{C}_t \hat{C}_t^T Y_s^T \hat{R}_t^p \hat{V}_{R,t}^{-1} - R_tH_t^p$$

$$\begin{align*}
&= \frac{1}{pq^2}T \sum_{s=1}^{T} K_{h,st} \left( R_t F_s^T \hat{C}_t \hat{C}_t^T E_s^T + E_s \hat{C}_t \hat{C}_t^T C_t F_s^T + E_s \hat{C}_t \hat{C}_t^T E_s \right) \\
&\quad + D_{R,s,t} F_s^T \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s D_{C,s,t} \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s \hat{C}_t \hat{C}_t^T C_t F_s^T D_{R,s,t} \\
&\quad + R_t F_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T R_t + R_t F_s D_{C,s,t} \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T D_{R,s,t} \\
&\quad + D_{C,s,t} F_s^T \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s D_{C,s,t} \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T D_{R,s,t} \\
&\quad + D_{C,s,t} F_s^T \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s D_{C,s,t} \hat{C}_t \hat{C}_t^T C_t F_s^T + R_t F_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T D_{R,s,t} \\
&\quad + E_s \hat{C}_t \hat{C}_t^T C_t F_s^T D_{R,s,t} + E_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T R_t + E_s \hat{C}_t \hat{C}_t^T D_{C,s,t} F_s^T D_{R,s,t} \hat{V}_{R,t}^{-1}
\end{align*}$$

where $D_{R,s,t} = R_s - R_t$ and $D_{C,s,t} = C_s - C_t$. Following the proof of Proposition 1, we can show $\|\hat{V}_{R,t}^{-1}\|_F^2 = O_p(1)$. Then Theorem 5 follows by Lemma 8 below.

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

\begin{enumerate}
  \item[(a)] $\frac{1}{q} \frac{1}{p} \| \sum_{s=1}^{T} K_{h,st} E_s^T (\hat{R}_t - R_t H_{R,t}) F_s \|_F^2 = O_p \left( m_{pqTh} \right)$, and $\frac{1}{p} \| \sum_{s=1}^{T} K_{h,st} E_s (\hat{C}_t - C_t H_{C,t}) F_s \|_F^2 = O_p \left( w_{pqTh} \right)$, where $m_{pqTh} = \frac{1}{T^2 p^2 h^2} + \frac{1}{T^2 q^2 h^2} + \frac{1}{T^2 pq h^2} + \frac{h^3}{T^2 pq^2 h^2} + \frac{1}{T^2 pq h^2} + \frac{h^3}{T pq^2 h}$.
  \item[(b)] $\frac{1}{p} \left\| I_1^P \hat{R}_t^p \right\|_F^2 = O_p \left( \frac{1}{qTh} + w_{pqTh} \right)$;
  \item[(c)] $\frac{1}{p} \left\| I_2^P \hat{R}_t^p \right\|_F^2 = O_p \left( \frac{1}{qTh} + w_{pqTh} \right)$;
  \item[(d)] $\frac{1}{p} \left\| I_3^P \hat{R}_t^p \right\|_F^2 = O_p \left( \frac{1}{pqTh} + \frac{1}{p^2 q^2} + \frac{\gamma_{pqTh}}{Thp} + \frac{\gamma_{pqTh}}{p^2} \right) + o_p(1) \times \frac{1}{p} \left\| \hat{R}_t^p - R_t H_{R,t}^p \right\|_F$;
  \item[(e)] $\frac{1}{p} \left\| I_j^P \hat{R}_t^p \right\|_F^2 = O_p(h^4), j = 4, \cdots, 13$;
  \item[(f)] $\frac{1}{p} \left\| I_j^P \hat{R}_t^p \right\|_F^2 = o_p(h^4), j = 14, \cdots, 24$.
\end{enumerate}
Appendix B  Technical Lemmas

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

$$
\frac{1}{qT} \sum_{s=1}^{T} K_{h,st} F_s C^T_t C_t F^T_s \to \Sigma_{FC,t}
$$

and

$$
\frac{1}{pT} \sum_{s=1}^{T} K_{h,st} F^T_s R_t R_t F_s \to \Sigma_{FR,t}.
$$

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

Lemma 10. Under Assumption 1 - 6, we have

$$
\sum_{s=1}^{T} \left\| \mathbb{E} \left[ \frac{1}{pq} K_{h,st} R^T_s E_s E^T_t R \right] \right\| = O(1),
$$

$$
\sum_{s=1}^{T} \left\| \mathbb{E} \left[ \frac{1}{pq} K_{h,st} C^T_s E_s E^T_t C \right] \right\| = O(1).
$$

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

$$
\mathbb{E} \left[ \frac{\sqrt{\eta}}{\sqrt{pqT}} \left\| \sum_{t=1}^{T} K_{h,tst} F_t C^T_s e_{t,ij} \right\|^2 \right] = O(1),
$$

$$
\mathbb{E} \left[ \frac{\sqrt{\eta}}{\sqrt{pqT}} \left\| \sum_{t=1}^{T} K_{h,tst} F^T_s R_t e_{t,ij} \right\|^2 \right] = O(1).
$$

Lemma 12. Under Assumption 1 - 6, we have

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

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

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

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

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

$$
\mathbb{E} \left\| \frac{\sqrt{\eta}}{\sqrt{pqT}} \sum_{t=1}^{T} K_{h,tst} R^T_s E_t C_s F^T_t \right\|^2 = O(1).
$$
Similarly, the $r \times r$ matrix satisfies
\[
\mathbb{E} \left\| \frac{\sqrt{h}}{\sqrt{pqT}} \sum_{t=1}^{T} K_{h,ts} F_t^T R_s^T E_t C_s \right\|^2 = O(1).
\]

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

(a) For any $i,l \in [p]$ and $j,h \in [q]$,
\[
\mathbb{E} \left( \frac{\sqrt{h}}{\sqrt{qT}} \sum_{t=1}^{T} \sum_{j=1}^{q} K_{h,ts} (e_{t,ij}e_{t,lj} - \mathbb{E}[e_{t,ij}e_{t,lj}]) \right)^2 = O(1),
\]
\[
\mathbb{E} \left( \frac{\sqrt{h}}{\sqrt{pT}} \sum_{t=1}^{T} \sum_{i=1}^{p} K_{h,ts} (e_{t,ij}e_{t,ih} - \mathbb{E}[e_{t,ij}e_{t,ih}]) \right)^2 = O(1),
\]
\[
\mathbb{E} \left\| \frac{\sqrt{h}}{p\sqrt{qT}} \sum_{t=1}^{T} K_{h,ts} (E_tE_t^T - \mathbb{E}[E_tE_t^T]) \right\|_F^2 = O(1),
\]
\[
\mathbb{E} \left\| \frac{\sqrt{h}}{p\sqrt{qT}} \sum_{t=1}^{T} K_{h,ts} (E_t^TE_t - \mathbb{E}[E_t^TE_t]) \right\|_F^2 = O(1).
\]

(b) For all $1 \leq t, s \leq T$, \(\mathbb{E} \left\| \frac{1}{\sqrt{pq}} R_t^T E_t C_s \right\|^2 = O(1)\).

Lemma 14. Under Assumption 1 - 6, we have

(a) For each row $i$, as $q, T \to \infty$,
\[
\frac{\sqrt{h}}{\sqrt{qT}} \sum_{s=1}^{T} K_{h,ts} F_s^T C_t e_{s,i} \overset{D}{\longrightarrow} N(\mathbf{0}, \Phi_{R,i}).
\]

(b) For each column $j$, as $p, T \to \infty$,
\[
\frac{\sqrt{h}}{\sqrt{pT}} \sum_{s=1}^{T} K_{h,ts} F_s^T R_t e_{t,j} \overset{D}{\longrightarrow} N(\mathbf{0}, \Phi_{C,j}).
\]

The $\Phi_{R,i}$'s and $\Phi_{C,j}$'s are defined in Theorem 2.

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