

Multilevel Main Effects Matrix Factor Model

Kaixin Liu*, Clifford Lam[†] and Zetai Cen[‡]

Department of Statistics, London School of Economics and Political Science

School of Mathematics, University of Bristol

Abstract

We propose a multilevel main effects matrix factor model (MMEFM) to extract meaningful row and column effects from groups of matrix data. As a generalization of a multilevel matrix factor model, we provide rigorous definitions of MMEFM and identifications, together with iterative algorithms for the estimators of all global and local components in the main effects and common component, including global/local main effects and core ranks, with rates of convergence spelt out. An important feature of the model is that all main effects can be nonstationary and still be consistently estimated at each time point. At the same time, our estimated global common factors are alignment independent, and hence are truly global, and can be presented and given interpretations more easily. We propose a Bootstrap procedure for detecting the existence and identifying the groups of data that are driven by the global common factors. The power of the test, together with the accuracy of various estimators, are demonstrated in extensive simulation settings. Estimation performance is also compared to other state-of-the-art methods for matrix time series by simulations and a set of U.S. equity data. All numerical works are done using the R package **MMEFM**, which is available at **CRAN**.

Key words and phrases: Groups of panels; iterative projection; varying factor strength; Tucker decomposition; nonstationary main effects.

*Kaixin Liu is PhD student, Department of Statistics, London School of Economics. Email: K.Liu31@lse.ac.uk

[†]Clifford Lam is Professor, Department of Statistics, London School of Economics. Email: C.Lam2@lse.ac.uk

[‡]Zetai Cen is Senior Research Associate, School of Mathematics, University of Bristol, and Visiting Fellow, Department of Statistics, London School of Economics. Email: zetai.cen@bristol.ac.uk. Supported by Engineering and Physical Sciences Research Council (EP/Z531327/1).

1 Introduction

With rapid advances in data collection and storage technologies, large numbers of measurements are now routinely recorded over time in matrix form. Such matrix-valued time series arise, for example, as asset-by-indicator panels in finance, country-by-indicator panels in economics, and origin-by-destination traffic matrices. For such data, matrix factor models, which extend factor models for vector time series (Bai, 2003; Lam et al., 2011; Lam and Yao, 2012), provide a natural approach to dimension reduction by preserving the intrinsic row and column structure, rather than treating the vectorized observations through an unrestricted loading structure; see, among others, Wang et al. (2019), Chen et al. (2020), He et al. (2023a), Yu et al. (2022), Chen and Fan (2023) and He et al. (2023b). More recently, Lam and Cen (2025) introduce a main effects matrix factor model (MEFM) that explicitly separates a time-varying grand mean and row and column main effects from the interaction (common) component. This distinction allows different forms of systematic variation within a matrix to be modeled and interpreted separately.

In some applications, several related high-dimensional time series are observed, and their dimensions need not agree across groups. Such data may exhibit another layer of structure: some latent dynamics are shared across groups, while others are group-specific. This distinction has been well studied in the literature on multilevel factor models for vector time series; see, among others, Moench et al. (2013), Choi et al. (2018), Han (2021) and Choi et al. (2023). Related global–local factor structures have also been considered for higher-dimensional data. Lu et al. (2025) study a three-dimensional factor model with global and local factors, while He et al. (2025) consider common and cluster-specific factors for biclustering matrix-valued time series. More generally, Cen (2026) develops a multi-order tensor factor model that separates global and local components across multiple tensor time series. Most closely related to our setting, Zhang et al. (2025) propose a multilevel matrix factor model (MFM) with global and local bilinear components. A global matrix factor process is shared among groups through group-specific row and column loadings, while local matrix factors capture group-specific variation.

The two forms of structure, multilevel variation across groups and additive main effects within each matrix, are complementary and can arise simultaneously in grouped matrix-valued time series. As pointed out by Lam and Cen (2025), additive main effects can be absorbed into an enlarged bilinear representation, but treating such a representation as unstructured obscures their distinct interpretation and rank structure. We therefore introduce a **Multilevel Main Effects matrix Factor Model** (MMEFM), which extends the MFM of Zhang et al. (2025) by incorporating time-varying grand means and row and column main effects, each with global and local components. While Lam and Cen (2025) discuss low-rank factor representations for the main effects, we develop such structures in a multilevel setting by decomposing the row and column main effects into global and local components, respectively. This exploits cross-group information and

provides a parsimonious representation while retaining group-specific dynamics. At the model level, the MFM of Zhang et al. (2025) is a restricted case of the MMEFM, so that the latter provides a more general representation by allowing global and local main-effect structures; see Remark 2.

Using pre-averaging ideas from Chen and Lam (2024), our estimation procedure utilizes cross-group information for the global main effects as well as the global common factors in the common component through projected cross-group moments. Local components in each data group are then estimated and removed from the data, with global factor processes still expressed under different row and column transformations. We therefore align and pool them through a nearest-Kronecker approximation to obtain a global common factor series $\hat{\mathbf{G}}_t$, followed by a refitting the group-specific global loadings. This leads to one of our contributions distinguished from Zhang et al. (2025), that unlike their group-specific representations $\hat{\mathbf{S}}_{mt}$ of the global common factors which are still group-dependent, our global main effects and common factor series are truly common across groups. They can then be easier visualized and compared with other estimated group-specific components to provide further insight.

We further establish convergence rates for the proposed estimators under a class of piecewise-stationary linear factor processes. In particular, the class $\mathcal{G}_T$ accommodates both short- and long-memory dependence as well as changes in dynamics; see Section 4 for details. Within the same setting, we also provide rank estimators of all loading matrices in Section 3.4, with theoretical results presented in Section 4.

Another important contribution is to determine the presence of a global common component. Imposing such a component without first assessing its presence may give a false impression of shared dynamics when no cross-group common factor exists. We construct group-level statistics and show that the order of the statistics for groups driven by the global components are larger than those for the remaining groups. Building on this separation, we develop a practical existence test for the global common components together with a screening rule for identifying the groups that share them.

The remainder of the paper is organized as follows. Section 2 introduces notation. Section 3 presents the MMEFM and its estimation. Section 4 introduces the assumptions for the MMEFM and establishes its theoretical properties, while Section 5 presents the practical detection and screening procedure for global common factors. Section 6 reports simulation results, including comparisons with other state-of-the-art matrix factor models, and an empirical analysis of the U.S. stock market. Proofs and additional results are included in the Supplementary Materials.

2 Notations

Throughout this paper, we use a letter, bold lower-case letter, and bold capital letter, i.e., a or A , $\mathbf{a}$, and $\mathbf{A}$, to denote a scalar, a vector, and a matrix respectively. We also use a_i , $A_{i,j}$, $\mathbf{A}_{i\cdot}$, $\mathbf{A}_{\cdot i}$ to denote, respectively,

the i th element of $\mathbf{a}$, the (i, j) th element of $\mathbf{A}$, the i th row vector (as a column vector) of $\mathbf{A}$, and the i th column vector of $\mathbf{A}$. For any $a, b \in \mathbb{R}$, we write $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. We use $a \asymp b$ to denote $a = O(b)$ and $b = O(a)$, while $a \asymp_P b$ to denote $a = O_P(b)$ and $b = O_P(a)$. We also use $a \gtrsim b$ and $a \lesssim b$ to denote $b = O(a)$ and $a = O(b)$ respectively, and $\gtrsim_P$ and $\lesssim_P$ their respective asymptotic probabilistic versions. We use $\circ$ to denote the Hadamard product and $\otimes$ the Kronecker product.

Given a positive integer m , define $[m] := \{1, 2, \dots, m\}$. The vector $\mathbf{1}_m$ denotes a vector of ones with length m . The i th largest eigenvalue of a matrix $\mathbf{A}$ is denoted by $\lambda_i(\mathbf{A})$. The notation $\mathbf{A} \succcurlyeq 0$ (resp. $\mathbf{A} \succ 0$) means that $\mathbf{A}$ is positive semi-definite (resp. positive definite). We use $\mathbf{A}'$ to denote the transpose of $\mathbf{A}$, and $\text{diag}(\mathbf{A})$ to denote a diagonal matrix with the diagonal elements of $\mathbf{A}$, while $\text{diag}(\{a_1, \dots, a_n\})$ represents the diagonal matrix with $\{a_1, \dots, a_n\}$ on the diagonal. The matrix $\mathbf{A}^\perp$ has orthonormal columns spanning the orthogonal complement of the column space of $\mathbf{A}$, so that $\mathbf{A}'\mathbf{A}^\perp = \mathbf{0}$ and $\mathbf{A}^{\perp'}\mathbf{A}^\perp = \mathbf{I}$. We use $\mathbf{A}^{\otimes 2}$ to denote $\mathbf{A}\mathbf{A}'$.

Some norm notations. For a given set, we denote by $|\cdot|$ its cardinality. We use $\|\cdot\|$ to denote the spectral norm of a matrix or the L_2 norm of a vector, and $\|\cdot\|_F$ to denote the Frobenius norm of a matrix. We use $\|\cdot\|_{\max}$ to denote the maximum absolute value of the elements in a vector or a matrix. The notations $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote the L_1 and L_∞ -norm of a matrix respectively, defined by $\|\mathbf{A}\|_1 := \max_j \sum_i |(\mathbf{A})_{ij}|$ and $\|\mathbf{A}\|_\infty := \max_i \sum_j |(\mathbf{A})_{ij}|$. Without loss of generality, we always assume the eigenvalues of a matrix are arranged by descending orders, and so are their corresponding eigenvectors.

3 Model and Estimation

We propose the MMEFM as follows. For each data group $m \in [M]$ and $t \in [T]$,

$$\begin{aligned} \mathbf{Y}_t^{(m)} &= (\mu_t + \mu_t^{(m)})\mathbf{1}_{p_m}\mathbf{1}_{q_m}' + [(\mathbf{A}_{1m}\boldsymbol{\alpha}_t) + (\mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)})]\mathbf{1}_{q_m}' + \mathbf{1}_{p_m}[(\mathbf{B}_{1m}\boldsymbol{\beta}_t) + (\mathbf{B}_{2m}\boldsymbol{\beta}_t^{(m)})]' \\ &\quad + \mathbf{Q}_m\mathbf{G}_t\mathbf{J}_m' + \mathbf{R}_m\mathbf{F}_t^{(m)}\mathbf{C}_m' + \mathbf{E}_t^{(m)} \\ &= \mu_t\mathbf{1}_{p_m}\mathbf{1}_{q_m}' + (\mathbf{A}_{1m}\boldsymbol{\alpha}_t)\mathbf{1}_{q_m}' + \mathbf{1}_{p_m}(\mathbf{B}_{1m}\boldsymbol{\beta}_t)' + \mathbf{Q}_m\mathbf{G}_t\mathbf{J}_m' \\ &\quad + \mu_t^{(m)}\mathbf{1}_{p_m}\mathbf{1}_{q_m}' + (\mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)})\mathbf{1}_{q_m}' + \mathbf{1}_{p_m}(\mathbf{B}_{2m}\boldsymbol{\beta}_t^{(m)})' + \mathbf{R}_m\mathbf{F}_t^{(m)}\mathbf{C}_m' + \mathbf{E}_t^{(m)}, \end{aligned} \quad (3.1)$$

where $\mathbf{Y}_t^{(m)}$ is the m th group observed matrix of dimension $p_m \times q_m$ at time t , μ_t , $\boldsymbol{\alpha}_t$, $\boldsymbol{\beta}_t$, $\mu_t^{(m)}$, $\boldsymbol{\alpha}_t^{(m)}$ and $\boldsymbol{\beta}_t^{(m)}$ are the global grand mean, row main effects, column main effects, and their corresponding local versions at group m respectively at time t . The matrices $\mathbf{A}_{1m} \in \mathbb{R}^{p_m \times r_1}$ (resp. $\mathbf{A}_{2m} \in \mathbb{R}^{p_m \times r_2^{(m)}}$) and $\mathbf{B}_{1m} \in \mathbb{R}^{q_m \times \ell_1}$ (resp. $\mathbf{B}_{2m} \in \mathbb{R}^{q_m \times \ell_2^{(m)}}$) are the global (resp. local) row main effect and the global (resp. local) column main effect loading matrix respectively at group m . The matrices $\mathbf{Q}_m \in \mathbb{R}^{p_m \times k_r}$, $\mathbf{J}_m \in \mathbb{R}^{q_m \times k_c}$, $\mathbf{R}_m \in \mathbb{R}^{p_m \times k_r^{(m)}}$ and $\mathbf{C}_m \in \mathbb{R}^{q_m \times k_c^{(m)}}$ are the global row, global column, local row and local

column factor loading matrix respectively at group m . Finally, $\mathbf{G}_t$ and $\mathbf{F}_t^{(m)}$ are the global factor and local factor matrices (at group m) respectively.

The upper equality of model (3.1) lists the grand means and the main effects first before the factor components, whereas the lower equality lists the global components first before the local components. This model can be considered a generalization of the one in Zhang et al. (2025), with global and local main effects included in our model. It is clear that the global time series $\{\boldsymbol{\alpha}_t\}$, $\{\boldsymbol{\beta}_t\}$ and $\{\mathbf{G}_t\}$ are all group-independent, whereas all other components of the model are dependent on the group m . It is also clear that at time t , the grand mean, row and column main effects for the group m data are given respectively by

$$c_t^{(m)} := \mu_t + \mu_t^{(m)}, \quad \mathbf{a}_t^{(m)} := \mathbf{A}_{1m}\boldsymbol{\alpha}_t + \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)}, \quad \mathbf{b}_t^{(m)} := \mathbf{B}_{1m}\boldsymbol{\beta}_t + \mathbf{B}_{2m}\boldsymbol{\beta}_t^{(m)}, \quad (3.2)$$

which are all an affine transformation of the global and local grand means/main effects and are naturally unidentifiable without further restrictions. The same applies to the global and local common components $\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m'$ and $\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m'$, which are themselves unidentifiable because they are the product of unknown matrices.

With (3.2), model (3.1) can also be written as, for each $t \in [T]$,

$$\mathbf{Y}_t^{(m)} = c_t^{(m)} \mathbf{1}_{p_m} \mathbf{1}_{q_m}' + \mathbf{a}_t^{(m)} \mathbf{1}_{q_m}' + \mathbf{1}_{p_m} \mathbf{b}_t^{(m)'} + (\mathbf{Q}_m, \mathbf{R}_m) \begin{pmatrix} \mathbf{G}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_t^{(m)} \end{pmatrix} \begin{pmatrix} \mathbf{J}_m' \\ \mathbf{C}_m' \end{pmatrix} + \mathbf{E}_t^{(m)}, \quad m \in [M],$$

which is a main effect matrix factor model in Lam and Cen (2025) for each group m , with a special block diagonal structure for the factor matrix. If we ignore this structure, the total number of factor series to be estimated is then $\sum_{m=1}^M (k_r + k_r^{(m)})(k_c + k_c^{(m)})$, which is much larger than $k_r k_c + \sum_{m=1}^M k_r^{(m)} k_c^{(m)}$, the number of factor series in model (3.1). More importantly, the interplay between these groups of time series and thus their interpretation would be lost.

Remark 1 Model (3.1) has the main effects split into “global” and “local”, apparently increasing the number of parameters and main effect series to estimate. However, as is evident in Section 3.3, the split is mainly there to promote interpretation rather than to increase the difficulty of estimation, since $c_t^{(m)}$, $\mathbf{a}_t^{(m)}$ and $\mathbf{b}_t^{(m)}$ can all be estimated consistently for each $t \in [T]$ from Theorem 2.

3.1 Examples

Consider the case where each data group represents an industry sector. For instance, $M = 11$ and each data matrix from group m , $\mathbf{Y}_t^{(m)}$, records data at time t from one of the eleven sectors classified by the Global Industry Classification Standard (GICS). Within each group, we may consider countries from the

G20 as the rows of $\mathbf{Y}_t^{(m)}$, and various economic indicators of each industry/sub-industry related to the sector at time t as the columns. Hence in this example we have $p_m = 20$ countries, but q_m can be different for each $m \in [11]$, depending on what economic indicators and what industry/sub-industry classifications we go for in each sector.

As for interpretations, the term $(\mathbf{A}_{1m}\boldsymbol{\alpha}_t)\mathbf{1}_{q_m}'$ from model (3.1) represents the country main effects which are *the same* across different economic indicators/ industries/sub-industries. The same vector $\boldsymbol{\alpha}_t$ continues to drive the country main effects across different sectors. Any genuine variation for the country main effects arising from differences of sectors are captured by the terms $(\mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)})\mathbf{1}_{q_m}'$ for $m \in [11]$. Reversing the roles of country and economic indicators/industries/sub-industries while replacing $\boldsymbol{\alpha}_t$ with $\boldsymbol{\beta}_t$ above give interpretations of $\mathbf{1}_{p_m}(\mathbf{B}_{1m}\boldsymbol{\beta}_t)'$ and $\mathbf{1}_{p_m}(\mathbf{B}_{2m}\boldsymbol{\beta}_t^{(m)})'$ respectively. Finally, the term $\mathbf{Q}_m\mathbf{G}_t\mathbf{J}_m'$ represents any *global* interaction effects remaining after accounting for all the global and local main effects, which are all driven by some global factor time series in $\mathbf{G}_t$. Any genuine differences in interaction effects across sectors are captured by the terms $\mathbf{R}_m\mathbf{F}_t^{(m)}\mathbf{C}_m'$.

Another example would be the financial data set used in Zhang et al. (2025). Same as the example described above, each data group still represents an industry sector. However, each row now represents a stock within the sector, and there are $p_m = 20$ stocks included for each sector. For the columns, financial data like return, volume of trade and turnover rate are included, so that q_m are the same for each m .

3.2 Identification

We first present (IC1) for identifying the main effects from the common component and (IC2) for identifying the global effects from the local ones. Condition (IC3) is needed for separating the loading matrices from the time series, so that both can be identified.

(IC1) (Main effects v.s. common component) For $m \in [M]$ and $j = 1, 2$, we have

$$\mathbf{1}_{p_m}'\mathbf{A}_{jm} = \mathbf{0}, \mathbf{1}_{q_m}'\mathbf{B}_{jm} = \mathbf{0}, \mathbf{1}_{p_m}'\mathbf{Q}_m = \mathbf{0}, \mathbf{1}_{q_m}'\mathbf{J}_m = \mathbf{0}, \mathbf{1}_{p_m}'\mathbf{R}_m = \mathbf{0}, \mathbf{1}_{q_m}'\mathbf{C}_m = \mathbf{0}.$$

(IC2) (Global v.s. local) For each $m \in [M]$, the series $\{\mathbf{G}_t\}$ is uncorrelated with $\{\mathbf{F}_t^{(m)}\}$. The series $\{\boldsymbol{\alpha}_t^{(m)}\}$, $\{\boldsymbol{\beta}_t^{(m)}\}$ and $\{\mathbf{F}_t^{(m)}\}$ are uncorrelated to $\{\boldsymbol{\alpha}_t^{(n)}\}$, $\{\boldsymbol{\beta}_t^{(n)}\}$ and $\{\mathbf{F}_t^{(n)}\}$ respectively for all $m \neq n$. Also $\sum_{m=1}^M \mu_t^{(m)} = 0$, $\sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t^{(m)'} = \mathbf{0}$ and $\sum_{t=1}^T \boldsymbol{\beta}_t \boldsymbol{\beta}_t^{(m)'} = \mathbf{0}$ for each $m \in [M]$. The elements of error series $\{\mathbf{E}_t^{(m)}\}$ are uncorrelated with those of $\{\mathbf{E}_t^{(n)}\}$ for $m \neq n$, and are also uncorrelated with those of $\{\boldsymbol{\alpha}_t\}$, $\{\boldsymbol{\beta}_t\}$, $\{\mathbf{G}_t\}$, $\{\boldsymbol{\alpha}_t^{(s)}\}$, $\{\boldsymbol{\beta}_t^{(s)}\}$ and $\{\mathbf{F}_t^{(s)}\}$ for all $s \in [M]$.

(IC3) (Separation of loading matrices and time series) All loading matrices involved in model (3.1) are of

full rank; of full rank are also the matrices $\mathbb{E}(\boldsymbol{\alpha}_t^{\otimes 2})$, $\mathbb{E}(\boldsymbol{\beta}_t^{\otimes 2})$, $\mathbb{E}(\boldsymbol{\alpha}_t^{(m)\otimes 2})$, $\mathbb{E}(\boldsymbol{\beta}_t^{(m)\otimes 2})$, $\mathbb{E}(\text{vec}(\mathbf{G}_t)^{\otimes 2})$, $\mathbb{E}(\text{vec}(\mathbf{F}_t^{(m)})^{\otimes 2})$, $\mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m}$, $\mathbf{B}_{1m}^{\perp'} \mathbf{B}_{2m}$, $\mathbf{Q}_m^{\perp'} \mathbf{R}_m$ and $\mathbf{J}_m^{\perp'} \mathbf{C}_m$.

The assumptions $\sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t^{(m)'} = \mathbf{0}$ and $\sum_{t=1}^T \boldsymbol{\beta}_t \boldsymbol{\beta}_t^{(m)'} = \mathbf{0}$ in (IC2) are not restrictive at all and are only for convenience in relevant proofs. For instance,

$$\begin{aligned} \mathbf{A}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)} &= \left\{ \mathbf{A}_{1m} + \mathbf{A}_{2m} \left(\sum_{s=1}^T \boldsymbol{\alpha}_s^{(m)} \boldsymbol{\alpha}_s' \right) \left(\sum_{s=1}^T \boldsymbol{\alpha}_s \boldsymbol{\alpha}_s' \right)^{-1} \right\} \boldsymbol{\alpha}_t \\ &\quad + \mathbf{A}_{2m} \left\{ \boldsymbol{\alpha}_t^{(m)} - \left(\sum_{s=1}^T \boldsymbol{\alpha}_s^{(m)} \boldsymbol{\alpha}_s' \right) \left(\sum_{s=1}^T \boldsymbol{\alpha}_s \boldsymbol{\alpha}_s' \right)^{-1} \boldsymbol{\alpha}_t \right\} =: \tilde{\mathbf{A}}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \tilde{\boldsymbol{\alpha}}_t^{(m)}, \end{aligned}$$

now with $\sum_{t=1}^T \boldsymbol{\alpha}_t \tilde{\boldsymbol{\alpha}}_t^{(m)'} = \mathbf{0}$. The (asymptotic) invertibility of $\sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t'$ comes from Assumption (IC3), and (S1) in Section 4.1. Meanwhile, if $\{\boldsymbol{\alpha}_t\}$ is uncorrelated with $\{\boldsymbol{\alpha}_t^{(m)}\}$ for $m \in [M]$, then elements in $T^{-1} \sum_{t=1}^T \tilde{\boldsymbol{\alpha}}_t^{(m)} \tilde{\boldsymbol{\alpha}}_t^{(n)'} are going to 0 at the same rate as those in $T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(n)'} for $m \neq n$, rendering $\{\tilde{\boldsymbol{\alpha}}_t^{(m)}\}$ and $\{\tilde{\boldsymbol{\alpha}}_t^{(n)}\}$ asymptotically uncorrelated for $m \neq n$. For simplicity, we do not engage in asymptotic identification in the proof of Theorem 1 and only use (IC2) as above for exact identification.$$

Theorem 1 (Identification) *Under (IC1), (IC2) and (IC3), all main effects, loading matrices, global and local means and common factors in model (3.1) are identified.*

Remark 2 *Following a demonstration in Lam and Cen (2025), using the notations in (3.2), model (3.1) can be written as a multilevel matrix FM as follows:*

$$\begin{aligned} \mathbf{Y}_t^{(m)} &= \tilde{\mathbf{Q}}_m \tilde{\mathbf{W}}_t^{(m)} \tilde{\mathbf{J}}_m' + \mathbf{E}_t^{(m)}, \quad \text{where} \\ \tilde{\mathbf{Q}}_m &:= (\mathbf{I}_{p_m}, \mathbf{Q}_m, \mathbf{R}_m, \mathbf{1}_{p_m}), \quad \tilde{\mathbf{J}}_m := (\mathbf{1}_{q_m}, \mathbf{J}_m, \mathbf{C}_m, \mathbf{I}_{q_m}), \quad \tilde{\mathbf{W}}_t^{(m)} := \begin{pmatrix} \mathbf{a}_t^{(m)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_t & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_t^{(m)} & \mathbf{0} \\ c_t^{(m)} & \mathbf{0} & \mathbf{0} & \mathbf{b}_t^{(m)'} \end{pmatrix}. \end{aligned}$$

Ignoring the above structure means that we are estimating $(p_m + k_r + k_r^{(m)} + 1)(q_m + k_c + k_c^{(m)} + 1)$ time series in $\tilde{\mathbf{W}}_t^{(m)}$, which tends to overfit the data because $p_m q_m$ is usually too large relative to T .

On the other hand, consider a multilevel matrix FM given by $\dot{\mathbf{Y}}_t^{(m)} = \dot{\mathbf{C}}_t + \dot{\mathbf{E}}_t$ for $m \in [M]$, where $\dot{\mathbf{C}}_t := \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m'$. Defining $\mathbf{M}_n := \mathbf{I}_n - n^{-1} \mathbf{1}_n \mathbf{1}_n'$, we construct

$$c_t^{(m)} := (p_m q_m)^{-1} \mathbf{1}_{p_m}' \dot{\mathbf{C}}_t \mathbf{1}_{q_m}, \quad \mathbf{a}_t^{(m)} := q_m^{-1} \mathbf{M}_{p_m} \dot{\mathbf{C}}_t \mathbf{1}_{q_m}, \quad \mathbf{b}_t^{(m)} := p_m^{-1} \mathbf{M}_{q_m} \dot{\mathbf{C}}_t' \mathbf{1}_{p_m}.$$

Then $\mathbf{Y}_t^{(m)}$ can be expressed as

$$\mathbf{Y}_t^{(m)} = c_t^{(m)} \mathbf{1}_{p_m} \mathbf{1}_{q_m}' + \mathbf{a}_t^{(m)} \mathbf{1}_{q_m}' + \mathbf{1}_{p_m} \mathbf{b}_t^{(m)'} + (\mathbf{M}_{p_m} \mathbf{Q}_m) \mathbf{G}_t (\mathbf{M}_{q_m} \mathbf{J}_m)' + (\mathbf{M}_{p_m} \mathbf{R}_m) \mathbf{F}_t^{(m)} (\mathbf{M}_{q_m} \mathbf{C}_m)' + \hat{\mathbf{E}}_t,$$

which is a multilevel MEFM in (3.1), with grand means and main effects given by

$$\begin{aligned} \mu_t &= (p_m q_m)^{-1} \mathbf{1}_{p_m}' \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{1}_{q_m}, \quad \mu_t^{(m)} = (p_m q_m)^{-1} \mathbf{1}_{p_m}' \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{1}_{q_m}, \quad \alpha_t = \mathbf{G}_t \mathbf{J}_m' \mathbf{1}_{q_m}, \\ \beta_t &= \mathbf{G}_t' \mathbf{Q}_m' \mathbf{1}_{p_m}, \quad \alpha_t^{(m)} = \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{1}_{q_m}, \quad \beta_t^{(m)} = \mathbf{F}_t^{(m)'} \mathbf{R}_m' \mathbf{1}_{p_m}, \end{aligned}$$

and the associated main effect loading matrices for α_t , β_t , $\alpha_t^{(m)}$ and $\beta_t^{(m)}$ given respectively by

$$\mathbf{A}_{1m} = q_m^{-1} \mathbf{M}_{p_m} \mathbf{Q}_m, \quad \mathbf{B}_{1m} = p_m^{-1} \mathbf{M}_{q_m} \mathbf{J}_m, \quad \mathbf{A}_{2m} = q_m^{-1} \mathbf{M}_{p_m} \mathbf{R}_m, \quad \mathbf{B}_{2m} = p_m^{-1} \mathbf{M}_{q_m} \mathbf{C}_m.$$

Two important observations are that all factor loading matrices are now one rank less because of the multiplication of $\mathbf{M}_{p_m}$ and $\mathbf{M}_{q_m}$, and $\alpha_t^{(m)}$ and $\beta_t^{(m)}$ are now of length $k_r^{(m)}$ and $k_c^{(m)}$ rather than being $r_2^{(m)}$ and $\ell_2^{(m)}$ respectively in model (3.1). Since $r_2^{(m)}$ and $\ell_2^{(m)}$ can be larger than $k_r^{(m)}$ and $k_c^{(m)}$, a multilevel MEFM is certainly more flexible than a multilevel matrix FM.

3.3 Estimation

The proof of Theorem 1 is instrumental to the estimators to be presented below. We first present the estimators for the grand mean and main effects as a whole. Then we present all global and local estimators for the main effects, before all common component estimators.

Hereafter, whenever an estimated loading matrix is formed from the first k eigenvectors of a symmetric matrix $\mathbf{V}$, these eigenvectors are chosen to be orthonormal and to correspond to the k largest eigenvalues of $\mathbf{V}$.

1. (Grand mean and main effects) The estimators of $c_t^{(m)}$, $\mathbf{a}_t^{(m)}$ and $\mathbf{b}_t^{(m)}$ are given respectively by

$$\hat{c}_t^{(m)} := (p_m q_m)^{-1} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m}, \quad \hat{\mathbf{a}}_t^{(m)} := q_m^{-1} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m}, \quad \hat{\mathbf{b}}_t^{(m)} := p_m^{-1} \mathbf{M}_{q_m} \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m}.$$

We estimate the global and local grand means by respectively

$$\hat{\mu}_t := M^{-1} \sum_{m=1}^M \hat{c}_t^{(m)}, \quad \hat{\mu}_t^{(m)} := \hat{c}_t^{(m)} - \hat{\mu}_t.$$

2. (Global and local main effects) Define

$$\widehat{\Omega}_{A_1,m} := (M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ (Tq_m q_n)^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{Y}_t^{(n)'} \mathbf{M}_{p_n} \right\}^{\otimes 2}, \quad (3.3)$$

$$\widehat{\Omega}_{B_1,m} := (M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ (Tp_m p_n)^{-1} \mathbf{M}_{q_m} \sum_{t=1}^T \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m} \mathbf{1}_{p_n}' \mathbf{Y}_t^{(n)} \mathbf{M}_{q_n} \right\}^{\otimes 2}. \quad (3.4)$$

We set $\widehat{\mathbf{A}}_{1m}$ and $\widehat{\mathbf{B}}_{1m}$ respectively as the first r_1 eigenvectors of $\widehat{\Omega}_{A_1,m}$, and the first ℓ_1 eigenvectors of $\widehat{\Omega}_{B_1,m}$. We take $\widehat{\mathbf{A}}_{2m}$ and $\widehat{\mathbf{B}}_{2m}$ respectively as the first $r_2^{(m)}$ and $\ell_2^{(m)}$ eigenvectors of

$$\widehat{\Omega}_{A_2,m} := \left\{ T^{-1} \sum_{t=1}^T \widehat{\mathbf{a}}_t^{(m) \otimes 2} \widehat{\mathbf{A}}_{1m}^\perp \right\}^{\otimes 2} \quad \text{and} \quad \widehat{\Omega}_{B_2,m} := \left\{ T^{-1} \sum_{t=1}^T \widehat{\mathbf{b}}_t^{(m) \otimes 2} \widehat{\mathbf{B}}_{1m}^\perp \right\}^{\otimes 2}. \quad (3.5)$$

The local main effect time series are estimated as

$$\begin{aligned} \widehat{\alpha}_t^{(m)} &:= q_m^{-1} (\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{A}}_{1m}^\perp \widehat{\mathbf{A}}_{1m}^{\perp'} \widehat{\mathbf{A}}_{2m})^{-1} \widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{A}}_{1m}^\perp \widehat{\mathbf{A}}_{1m}^{\perp'} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m}, \\ \widehat{\beta}_t^{(m)} &:= p_m^{-1} (\widehat{\mathbf{B}}_{2m}' \widehat{\mathbf{B}}_{1m}^\perp \widehat{\mathbf{B}}_{1m}^{\perp'} \widehat{\mathbf{B}}_{2m})^{-1} \widehat{\mathbf{B}}_{2m}' \widehat{\mathbf{B}}_{1m}^\perp \widehat{\mathbf{B}}_{1m}^{\perp'} \mathbf{M}_{q_m} \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m}. \end{aligned} \quad (3.6)$$

Define

$$\begin{aligned} \widehat{\mathbf{Z}}_{\alpha,m} &:= \widehat{\mathbf{A}}_{1m}' (\widehat{\mathbf{y}}_1^{(m)}, \dots, \widehat{\mathbf{y}}_T^{(m)}), \quad \text{where} \quad \widehat{\mathbf{y}}_t^{(m)} := q_m^{-1} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} - \widehat{\mathbf{A}}_{2m} \widehat{\alpha}_t^{(m)}, \\ \widehat{\mathbf{Z}}_{\beta,m} &:= \widehat{\mathbf{B}}_{1m}' (\widehat{\mathbf{w}}_1^{(m)}, \dots, \widehat{\mathbf{w}}_T^{(m)}), \quad \text{where} \quad \widehat{\mathbf{w}}_t^{(m)} := p_m^{-1} \mathbf{M}_{q_m} \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m} - \widehat{\mathbf{B}}_{2m} \widehat{\beta}_t^{(m)}. \end{aligned}$$

We estimate $\widehat{\mathbf{A}}_\alpha := (\widehat{\alpha}_1, \dots, \widehat{\alpha}_T)$ and $\widehat{\mathbf{B}}_\beta := (\widehat{\beta}_1, \dots, \widehat{\beta}_T)$ as $\sqrt{T}$ times the matrices whose rows are, respectively, the first r_1 and ℓ_1 eigenvectors of

$$M^{-1} \sum_{m=1}^M \widehat{\mathbf{Z}}_{\alpha,m}' (\widehat{\mathbf{Z}}_{\alpha,m} \widehat{\mathbf{Z}}_{\alpha,m}')^{-1} \widehat{\mathbf{Z}}_{\alpha,m} \quad \text{and} \quad M^{-1} \sum_{m=1}^M \widehat{\mathbf{Z}}_{\beta,m}' (\widehat{\mathbf{Z}}_{\beta,m} \widehat{\mathbf{Z}}_{\beta,m}')^{-1} \widehat{\mathbf{Z}}_{\beta,m}.$$

We then re-estimate $\mathbf{A}_{1m}$ and $\mathbf{B}_{1m}$ respectively as

$$\widetilde{\mathbf{A}}_{1m} := T^{-1} (\widehat{\mathbf{y}}_1^{(m)}, \dots, \widehat{\mathbf{y}}_T^{(m)}) \widehat{\mathbf{A}}_\alpha' \quad \text{and} \quad \widetilde{\mathbf{B}}_{1m} := T^{-1} (\widehat{\mathbf{w}}_1^{(m)}, \dots, \widehat{\mathbf{w}}_T^{(m)}) \widehat{\mathbf{B}}_\beta'. \quad (3.7)$$

3. (Common component) Define

$$\check{\mathbf{Y}}_t^{(m)} := \mathbf{Y}_t^{(m)} - \widehat{c}_t^{(m)} \mathbf{1}_{p_m} \mathbf{1}_{q_m}' - \widehat{\mathbf{a}}_t^{(m)} \mathbf{1}_{q_m}' - \mathbf{1}_{p_m} \widehat{\mathbf{b}}_t^{(m)'}$$

Using the pre-averaging idea in [Chen and Lam \(2024\)](#), we find from a family of randomly generated

directions the one that aligns the most with the direction of the strongest factors from $\mathbf{Q}_m$, say $\mathbf{v}_a^{(m)}$.

Do the same for $\mathbf{J}_m$, say $\mathbf{v}_b^{(m)}$. See Section A for all the details of this procedure.

Let $\mathbf{v}_a^{(m)}$ and $\mathbf{v}_b^{(m)}$ be standardized as unit vectors. Define

$$\mathbf{X}_{Q,m} := T^{-2}(M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} \right\}^{\otimes 2}, \quad (3.8)$$

$$\mathbf{X}_{J,m} := T^{-2}(M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)'} \mathbf{v}_a^{(m)} \mathbf{v}_a^{(n)'} \check{\mathbf{Y}}_t^{(n)} \right\}^{\otimes 2}. \quad (3.9)$$

We take $\hat{\mathbf{Q}}_m$ as the first k_r eigenvectors of $\mathbf{X}_{Q,m}$, and $\hat{\mathbf{J}}_m$ as the first k_c eigenvectors of $\mathbf{X}_{J,m}$.

We then define

$$\mathbf{X}_{R,m} := T^{-1} \sum_{t=1}^T \{ \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{J}}_m^\perp \}^{\otimes 2}, \quad \mathbf{X}_{C,m} := T^{-1} \sum_{t=1}^T \{ \check{\mathbf{Y}}_t^{(m)'} \hat{\mathbf{Q}}_m^\perp \}^{\otimes 2}. \quad (3.10)$$

We take $\hat{\mathbf{R}}_m$ as the first $k_r^{(m)}$ eigenvectors of $\mathbf{X}_{R,m}$, and $\hat{\mathbf{C}}_m$ as the first $k_c^{(m)}$ eigenvectors of $\mathbf{X}_{C,m}$.

The local common factor $\mathbf{F}_t^{(m)}$ is estimated using least square method:

$$\hat{\mathbf{F}}_t^{(m)} := (\hat{\mathbf{R}}_m' \hat{\mathbf{Q}}_m^\perp \hat{\mathbf{Q}}_m^{\perp'} \hat{\mathbf{R}}_m)^{-1} \hat{\mathbf{R}}_m' \hat{\mathbf{Q}}_m^\perp \hat{\mathbf{Q}}_m^{\perp'} \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{C}}_m.$$

Define $\check{\mathbf{Y}}_{G,t}^{(m)} := \check{\mathbf{Y}}_t^{(m)} - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m'$, and

$$\hat{\mathbf{Z}}_{G,m} := (\text{vec}(\hat{\mathbf{Q}}_m' \check{\mathbf{Y}}_{G,1}^{(m)} \hat{\mathbf{J}}_m), \dots, \text{vec}(\hat{\mathbf{Q}}_m' \check{\mathbf{Y}}_{G,T}^{(m)} \hat{\mathbf{J}}_m)).$$

We first choose a reference group $m_0 \in [M]$. For each $m \neq m_0$, define

$$\hat{\Xi}_m := \hat{\mathbf{Z}}_{G,m_0} \hat{\mathbf{Z}}_{G,m}' (\hat{\mathbf{Z}}_{G,m} \hat{\mathbf{Z}}_{G,m}')^{-1}.$$

We obtain a two-sided alignment by solving $\min_{\Xi_{Q,m} \in \mathbb{R}^{k_r \times k_r}, \Xi_{J,m} \in \mathbb{R}^{k_c \times k_c}} \|\Xi_{J,m} \otimes \Xi_{Q,m} - \hat{\Xi}_m\|_F^2$ (see e.g., Section 6 of [Loan, 2000](#)). We estimate $\Xi_{Q,m}$ and $\Xi_{J,m}$ using $\text{vec}(\hat{\Xi}_{Q,m}) = \hat{\sigma}_{\Xi,m}^{1/2} \hat{\mathbf{u}}_{\Xi,m}$ and $\text{vec}(\hat{\Xi}_{J,m}) = \hat{\sigma}_{\Xi,m}^{1/2} \hat{\mathbf{v}}_{\Xi,m}$ respectively, where $\hat{\sigma}_{\Xi,m}$, $\hat{\mathbf{u}}_{\Xi,m}$ and $\hat{\mathbf{v}}_{\Xi,m}$ are the largest singular value, the corresponding left and right singular vectors of $\mathcal{R}(\hat{\Xi}_m)$, respectively; the matrix $\mathcal{R}(\mathbf{W})$ for a block matrix $\mathbf{W} = (\mathbf{W}_{ij})_{i \in [k_c], j \in [k_r]}$, with each $\mathbf{W}_{ij}$ of size $k_r \times k_r$, is defined as

$$\mathcal{R}(\mathbf{W}) := (\text{vec}(\mathbf{W}_{11}), \dots, \text{vec}(\mathbf{W}_{k_c 1}), \dots, \text{vec}(\mathbf{W}_{1 k_c}), \dots, \text{vec}(\mathbf{W}_{k_c k_c})).$$

For the reference group, set $\widehat{\mathbf{\Xi}}_{Q,m_0} := \mathbf{I}_{k_r}$ and $\widehat{\mathbf{\Xi}}_{J,m_0} := \mathbf{I}_{k_c}$. The global common factor series is then estimated as

$$\widehat{\mathbf{G}}_t := M^{-1} \sum_{m=1}^M \widehat{\mathbf{\Xi}}_{Q,m} \widehat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,t}^{(m)} \widehat{\mathbf{J}}_m \widehat{\mathbf{\Xi}}'_{J,m}.$$

We then re-estimate $\mathbf{J}_m \otimes \mathbf{Q}_m$ by least squares:

$$\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m} := (\text{vec}(\check{\mathbf{Y}}_{G,1}^{(m)}), \dots, \text{vec}(\check{\mathbf{Y}}_{G,T}^{(m)})) (\text{vec}(\widehat{\mathbf{G}}_1), \dots, \text{vec}(\widehat{\mathbf{G}}_T))' [(\text{vec}(\widehat{\mathbf{G}}_1), \dots, \text{vec}(\widehat{\mathbf{G}}_T))^{\otimes 2}]^{-1}.$$

Finally, to solve $\min_{\mathbf{Q}_m, \mathbf{J}_m} \|\mathbf{J}_m \otimes \mathbf{Q}_m - \widehat{\mathbf{J}_m \otimes \mathbf{Q}_m}\|_F^2$, we estimate $\mathbf{Q}_m$ and $\mathbf{J}_m$ by setting $\text{vec}(\tilde{\mathbf{Q}}_m) = \hat{\sigma}_{Q,m}^{1/2} \widehat{\mathbf{u}}_{Q,m}$ and $\text{vec}(\tilde{\mathbf{J}}_m) = \hat{\sigma}_{Q,m}^{1/2} \widehat{\mathbf{v}}_{J,m}$ respectively, where $\hat{\sigma}_{Q,m}$, $\widehat{\mathbf{u}}_{Q,m}$ and $\widehat{\mathbf{v}}_{J,m}$ are the largest singular value, and the corresponding left and right singular vectors of $\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m})$ respectively.

Different from the global time series estimators proposed in [Zhang et al. \(2025\)](#), our proposed $\widehat{\mathbf{a}}_t$, $\widehat{\mathbf{\beta}}_t$ in Step 2 and $\widehat{\mathbf{G}}_t$ in Step 3 are all independent of the index m , with all associated loading matrices re-estimated based on these truly global time series estimators.

Remark 3 We stress that the grand mean time series $\{c_t^{(m)}\}$ and the main effects time series $\{\mathbf{a}_t^{(m)}\}$ and $\{\mathbf{b}_t^{(m)}\}$ themselves can be nonstationary, which is an important feature for MEFM in [Lam and Cen \(2025\)](#). The global or local grand means and main effects can also be nonstationary, as long as Conditions (ME1) and (ME2) in Section 4.1 are satisfied.

Remark 4 The two-sided alignment step is needed because, for each m , $\widehat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,t}^{(m)} \widehat{\mathbf{J}}_m$ is expressed in group-specific row and column coordinates. In particular, if $\widehat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,t}^{(m)} \widehat{\mathbf{J}}_m \approx \mathbf{L}_m \mathbf{G}_t \mathbf{S}'_m$, then direct averaging yields $M^{-1} \sum_{m=1}^M \mathbf{L}_m \mathbf{G}_t \mathbf{S}'_m$, while $M^{-1} \sum_{m=1}^M (\mathbf{S}_m \otimes \mathbf{L}_m)$ need not be a Kronecker product. Hence, the average generally cannot be written as $\mathbf{L} \mathbf{G}_t \mathbf{S}'$ for a common pair $(\mathbf{L}, \mathbf{S})$, and the resulting coordinate change cannot in general be absorbed by the subsequent two-sided least-squares refitting.

At the population level, $\mathbf{G}_t$ is represented in the coordinate systems of group m and m_0 as $\mathbf{L}_m \mathbf{G}_t \mathbf{S}'_m$ and $\mathbf{L}_{m_0} \mathbf{G}_t \mathbf{S}'_{m_0}$, respectively. The vectorized coordinate change from group m to group m_0 is then $(\mathbf{S}_{m_0} \mathbf{S}_m^{-1}) \otimes (\mathbf{L}_{m_0} \mathbf{L}_m^{-1})$, since applying this transformation to $\text{vec}(\mathbf{L}_m \mathbf{G}_t \mathbf{S}'_m)$ yields $\text{vec}(\mathbf{L}_{m_0} \mathbf{G}_t \mathbf{S}'_{m_0})$. Under the rearrangement operator,

$$\mathcal{R}((\mathbf{S}_{m_0} \mathbf{S}_m^{-1}) \otimes (\mathbf{L}_{m_0} \mathbf{L}_m^{-1})) = \text{vec}(\mathbf{L}_{m_0} \mathbf{L}_m^{-1}) \text{vec}(\mathbf{S}_{m_0} \mathbf{S}_m^{-1})'.$$

Hence, the leading left and right singular vectors of $\mathcal{R}(\widehat{\mathbf{\Xi}}_m)$ estimate the row and column alignment matrices, bringing the representation from group m into the coordinate system of group m_0 before averaging.

Remark 5 The reference group in the global alignment can be chosen arbitrarily in principle. In numerical implementation, we choose the group with the largest projected-core norm, for instance the largest value of $\sum_{t=1}^T \left\| \hat{\mathbf{Q}}'_m \tilde{\mathbf{Y}}_{G,t}^{(m)} \hat{\mathbf{J}}_m \right\|_F^2$ to improve numerical stability and to avoid using a group whose global common component is nearly degenerate as the reference. A more formal procedure for detecting whether a group contains nonzero global common factors is given in Section 4.3.

3.4 Rank estimation for all loading matrices

The rank of all loading matrices in model (3.1) are assumed known so far. We propose the following eigenvalue-ratio based estimators for $r_1, r_2^{(m)}, \ell_1, \ell_2^{(m)}, k_r, k_c, k_r^{(m)}$ and $k_c^{(m)}$ respectively:

$$\begin{aligned} \hat{r}_1 &:= \arg \min_{j \in [p_m]} \frac{\max_{m \in [M]} \lambda_{j+1}(\hat{\mathbf{\Omega}}_{A_1,m})}{\max_{m \in [M]} \lambda_j(\hat{\mathbf{\Omega}}_{A_1,m})}, \quad \hat{r}_2^{(m)} := \arg \min_{j \in [p_m]} \frac{\lambda_{j+1}(\hat{\mathbf{\Omega}}_{A_2,m})}{\lambda_j(\hat{\mathbf{\Omega}}_{A_2,m})}, \quad \hat{\ell}_1 := \arg \min_{j \in [q_m]} \frac{\max_{m \in [M]} \lambda_{j+1}(\hat{\mathbf{\Omega}}_{B_1,m})}{\max_{m \in [M]} \lambda_j(\hat{\mathbf{\Omega}}_{B_1,m})}, \\ \hat{\ell}_2^{(m)} &:= \arg \min_{j \in [q_m]} \frac{\lambda_{j+1}(\hat{\mathbf{\Omega}}_{B_2,m})}{\lambda_j(\hat{\mathbf{\Omega}}_{B_2,m})}, \quad \hat{k}_r := \arg \min_{j \in [p_m]} \frac{\max_{m \in [M]} \lambda_{j+1}(\mathbf{X}_{Q,m})}{\max_{m \in [M]} \lambda_j(\mathbf{X}_{Q,m})}, \quad \hat{k}_c := \arg \min_{j \in [q_m]} \frac{\max_{m \in [M]} \lambda_{j+1}(\mathbf{X}_{J,m})}{\max_{m \in [M]} \lambda_j(\mathbf{X}_{J,m})}, \\ \hat{k}_r^{(m)} &:= \arg \min_{j \in [p_m]} \frac{\lambda_{j+1}(\mathbf{X}_{R,m})}{\lambda_j(\mathbf{X}_{R,m})}, \quad \hat{k}_c^{(m)} := \arg \min_{j \in [q_m]} \frac{\lambda_{j+1}(\mathbf{X}_{C,m})}{\lambda_j(\mathbf{X}_{C,m})}, \end{aligned}$$

where $\hat{\mathbf{\Omega}}_{A_1,m}, \hat{\mathbf{\Omega}}_{B_1,m}, \hat{\mathbf{\Omega}}_{A_2,m}, \hat{\mathbf{\Omega}}_{B_2,m}, \mathbf{X}_{Q,m}, \mathbf{X}_{J,m}, \mathbf{X}_{R,m}$ and $\mathbf{X}_{C,m}$ are defined in (3.3), (3.4), (3.5), (3.8), (3.9) and (3.10), respectively.

4 Assumptions and Theoretical Results

4.1 Assumptions

We say $\{\mathbf{w}_t\}_{t \in [T]} \in \mathcal{G}_T(\{d_1, \dots, d_m\} \mid \{\mathbf{x}_i, i \in \mathbb{Z}\})$ if the process $\{\mathbf{w}_t\}_{t \in [T]}$ is piecewise-stationary with m stationary segments, each with length of order T . In particular, for t within the j -th segment, $\mathbf{w}_t$ can be expressed as

$$\mathbf{w}_t = \sum_{w \geq 0} \psi_j(w) \mathbf{x}_{t-w}, \quad \psi_j(0) = 1, \quad \psi_j(w) := w^{d_j-1} l_j(w) \text{ for } w \geq 1,$$

where $d_j < 1/2$, and $l_j(\cdot) : \mathbb{R} \mapsto \mathbb{R}$ is a function such that $l_j(w)$ converges to a constant as $w \rightarrow \infty$. All elements in $\{\mathbf{x}_w\}$ have mean 0 and are independent and identically distributed with bounded 4th order moments. To facilitate notations under the regime of $\mathcal{G}_T(\cdot)$, define

$$R(T, d) := \begin{cases} T^{-1}, & d < 1/4, \\ T^{-1} \log T, & d = 1/4, \\ T^{4d-2}, & 1/4 < d < 1/2. \end{cases} \quad (4.1)$$

(ME1) (Main effect (uniform) factor strength). We assume that for each $m \in [M]$, as $p_m, q_m \rightarrow \infty$,

$$\begin{aligned} \mathbf{A}'_{1m} \mathbf{A}_{1m} / p_m^{\gamma_r} &\rightarrow \Sigma_{A_1, m}, & \mathbf{B}'_{1m} \mathbf{B}_{1m} / q_m^{\gamma_c} &\rightarrow \Sigma_{B_1, m}, \\ \mathbf{A}'_{2m} \mathbf{A}_{2m} / p_m^{\gamma_r^{(m)}} &\rightarrow \Sigma_{A_2, m}, & \mathbf{B}'_{2m} \mathbf{B}_{2m} / q_m^{\gamma_c^{(m)}} &\rightarrow \Sigma_{B_2, m}. \end{aligned}$$

where $\gamma_r, \gamma_r^{(m)}, \gamma_c, \gamma_c^{(m)}$ are all in $(1/2, 1]$. The matrices $\Sigma_{A_1, m}, \Sigma_{B_1, m}, \Sigma_{A_2, m}$, and $\Sigma_{B_2, m}$ are positive definite with eigenvalues bounded away from zero and infinity for all $m \in [M]$.

(ME2) (Time series in main effects). The time series $\{\alpha_t\}, \{\alpha_t^{(m)}\}, \{\beta_t\}$ and $\{\beta_t^{(m)}\}$ for $t \in [T]$ are all in $\mathcal{G}_T$ with potentially different set and number of piecewise-stationary segments, different set of d_j 's and l_j 's, and defined with time series $\{\mathbf{x}_{\alpha, t}\}, \{\mathbf{x}_{\alpha, t}^{(m)}\}, \{\mathbf{x}_{\beta, t}\}$ and $\{\mathbf{x}_{\beta, t}^{(m)}\}$ respectively. Also, $\{\mathbf{x}_{\alpha, t}^{(m)}\}$ and $\{\mathbf{x}_{\beta, t}^{(m)}\}$ are independent of $\{\mathbf{x}_{\alpha, t}^{(n)}\}$ and $\{\mathbf{x}_{\beta, t}^{(n)}\}$ respectively for $m \neq n$.

(E1) (Decomposition of error). We assume that

$$\mathbf{E}_t^{(m)} = \mathbf{Q}_{e, m} \mathbf{G}_{e, t} \mathbf{J}'_{e, m} + \mathbf{R}_{e, m} \mathbf{F}_{e, t}^{(m)} \mathbf{C}'_{e, m} + \Sigma_{\epsilon, m} \circ \epsilon_t^{(m)},$$

where $\mathbf{G}_{e, t}$ and $\mathbf{F}_{e, t}^{(m)}$ are $k_{e, r} \times k_{e, c}$ and $k_{e, r}^{(m)} \times k_{e, c}^{(m)}$ matrices containing independent elements with mean 0 and variance 1, respectively. The matrix $\epsilon_t^{(m)} \in \mathbb{R}^{p_m \times q_m}$ consists of independent entries with mean 0 and variance 1. The series $\{\mathbf{G}_{e, t}\}, \{\mathbf{F}_{e, t}^{(m)}\}$ and $\{\epsilon_t^{(m)}\}$ are mutually independent. The matrix $\Sigma_{\epsilon, m}$ contains standard deviations, uniformly bounded away from 0 and infinity, of the corresponding elements in $\epsilon_t^{(m)}$. Moreover, the error loadings are all (approximately) sparse matrices uniformly across all groups, i.e., for all $m \in [M]$, $\|\mathbf{Q}_{e, m}\|_1 = \|\mathbf{J}_{e, m}\|_1 = \|\mathbf{R}_{e, m}\|_1 = \|\mathbf{C}_{e, m}\|_1 = O(1)$, with $k_{e, r}, k_{e, c}, k_{e, r}^{(m)}, k_{e, c}^{(m)} = O(1)$. Furthermore, assume $\mathbf{Q}_{e, m} \mathbf{Q}'_{e, n} = \mathbf{0}$ or $\mathbf{J}_{e, m} \mathbf{J}'_{e, n} = \mathbf{0}$ for each $m \neq n$.

(E2) (Time series in $\mathbf{E}_t^{(m)}$). There is $\mathbf{X}_{e, t}$ the same dimension as $\mathbf{G}_{e, t}$, $\mathbf{X}_{e, t}^{(m)}$ the same dimension as $\mathbf{F}_{e, t}^{(m)}$ and $\mathbf{X}_{\epsilon, t}^{(m)}$ the same dimension as $\epsilon_t^{(m)}$ such that $\mathbf{G}_{e, t} = \sum_{w \geq 0} a_{e, w} \mathbf{X}_{e, t-w}$, $\mathbf{F}_{e, t}^{(m)} = \sum_{w \geq 0} a_{e, w}^{(m)} \mathbf{X}_{e, t-w}^{(m)}$ and $\epsilon_t^{(m)} = \sum_{w \geq 0} a_{\epsilon, w}^{(m)} \mathbf{X}_{\epsilon, t-w}^{(m)}$, with $\{\mathbf{X}_{e, t}\}, \{\mathbf{X}_{e, t}^{(m_1)}\}, \{\mathbf{X}_{e, t}^{(m_2)}\}, \{\mathbf{X}_{\epsilon, t}^{(n_1)}\}$ and $\{\mathbf{X}_{\epsilon, t}^{(n_2)}\}$ independent of each other for all $m_1, m_2, n_1, n_2 \in [M]$. The series $\{\mathbf{X}_{e, t}\}$ and $\{\mathbf{X}_{e, t}^{(m)}\}$ have independent elements for each $m \in [M]$, while $\{\mathbf{X}_{\epsilon, t}^{(m)}\}$ has i.i.d. elements for each $m \in [M]$, all with mean 0 and variance 1 with uniformly bounded fourth order moments. The time series $\{\mathbf{X}_{e, t}\}, \{\mathbf{X}_{e, t}^{(m)}\}$, and $\{\mathbf{X}_{\epsilon, t}^{(m)}\}$ are independent of all main effect and factor series in model (3.1) for each $m \in [M]$.

The coefficients $a_{e, w}, a_{e, w}^{(m)}$ and $a_{\epsilon, w}^{(m)}$ are such that $\sum_{w \geq 0} a_{e, w}^2 = \sum_{w \geq 0} a_{e, w}^{(m)2} = \sum_{w \geq 0} a_{\epsilon, w}^{(m)2} = 1$ and $\sum_{w \geq 0} |a_{e, w}|, \sum_{w \geq 0} |a_{e, w}^{(m)}|, \sum_{w \geq 0} |a_{\epsilon, w}^{(m)}| \leq c$ for some constant $c > 0$.

(F1) (Time series in $\mathbf{G}_t$ and $\mathbf{F}_t^{(m)}$). For $t \in [T]$, define $\mathbf{z}_{G, t} := \text{vec}(\mathbf{G}'_t)$, and $\mathbf{z}_{F, t}^{(m)} := \text{vec}(\mathbf{F}_t^{(m)'})$ for each

$m \in [M]$. The time series $\{\mathbf{z}_{G,t}\}$ and $\{\mathbf{z}_{F,t}^{(m)}\}$ are in $\mathcal{G}_T$ with potentially different sets and numbers of piecewise-stationary segments, different sets of d_j 's and l_j 's, and defined with independent time series $\{\mathbf{x}_{G,t}\}$ and $\{\mathbf{x}_{F,t}^{(m)}\}$ respectively. Also, $\{\mathbf{x}_{F,t}^{(m)}\}$ is independent of $\{\mathbf{x}_{F,t}^{(n)}\}$ for $m \neq n$. Both $\{\mathbf{G}_t\}$ and $\{\mathbf{F}_t^{(m)}\}$ are independent of $\{\mathbf{E}_t^{(n)}\}$ for $m, n \in [M]$.

(L1) (Factor loading matrices) For each $m \in [M]$, define $\mathbf{Z}_q^{(m)} := \text{diag}(\mathbf{Q}_m' \mathbf{Q}_m)$, $\mathbf{Z}_j^{(m)} := \text{diag}(\mathbf{J}_m' \mathbf{J}_m)$, $\mathbf{Z}_r^{(m)} := \text{diag}(\mathbf{R}_m' \mathbf{R}_m)$ and $\mathbf{Z}_c^{(m)} := \text{diag}(\mathbf{C}_m' \mathbf{C}_m)$. We assume that

$$\begin{aligned} (\mathbf{Z}_q^{(m)})^{-1/2} \mathbf{Q}_m' \mathbf{Q}_m (\mathbf{Z}_q^{(m)})^{-1/2} &\rightarrow \Sigma_{Q,m}, & (\mathbf{Z}_j^{(m)})^{-1/2} \mathbf{J}_m' \mathbf{J}_m (\mathbf{Z}_j^{(m)})^{-1/2} &\rightarrow \Sigma_{J,m}, \\ (\mathbf{Z}_r^{(m)})^{-1/2} \mathbf{R}_m' \mathbf{R}_m (\mathbf{Z}_r^{(m)})^{-1/2} &\rightarrow \Sigma_{R,m}, & (\mathbf{Z}_c^{(m)})^{-1/2} \mathbf{C}_m' \mathbf{C}_m (\mathbf{Z}_c^{(m)})^{-1/2} &\rightarrow \Sigma_{C,m}, \end{aligned}$$

where $\Sigma_{Q,m}$, $\Sigma_{J,m}$, $\Sigma_{R,m}$ and $\Sigma_{C,m}$ have eigenvalues uniformly bounded away from 0 and infinity for all $m \in [M]$. Also, we assume

$$\begin{aligned} (\mathbf{Z}_q^{(m)})_{ii} &\asymp p_m^{\delta_{q,i}^{(m)}} \text{ for } i \in [k_r], & (\mathbf{Z}_j^{(m)})_{ii} &\asymp q_m^{\delta_{j,i}^{(m)}} \text{ for } i \in [k_c], \\ (\mathbf{Z}_r^{(m)})_{ii} &\asymp p_m^{\delta_{r,i}^{(m)}} \text{ for } i \in [k_r^{(m)}], & (\mathbf{Z}_c^{(m)})_{ii} &\asymp q_m^{\delta_{c,i}^{(m)}} \text{ for } i \in [k_c^{(m)}], \end{aligned}$$

where $\delta_{q,i}^{(m)}$, $\delta_{j,i}^{(m)}$, $\delta_{r,i}^{(m)}$ and $\delta_{c,i}^{(m)}$ are all in $(1/2, 1]$, ordered non-increasingly.

(S1) (Non-degeneracy and separation conditions) As $T \rightarrow \infty$, the eigenvalues of $T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t^{\otimes 2})$, $T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\beta}_t^{\otimes 2})$, $T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t^{(m)\otimes 2})$, $T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\beta}_t^{(m)\otimes 2})$, $T^{-1} \sum_{t=1}^T \mathbb{E}(\text{vec}(\mathbf{G}_t)^{\otimes 2})$, and $T^{-1} \sum_{t=1}^T \mathbb{E}(\text{vec}(\mathbf{F}_t^{(m)})^{\otimes 2})$ are all bounded away from zero and infinity over $m \in [M]$. In addition, for sufficiently large p_m and q_m , the singular values of

$$\begin{aligned} \mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m} (\mathbf{A}_{2m}' \mathbf{A}_{2m})^{-1/2}, & \quad \mathbf{B}_{1m}^{\perp'} \mathbf{B}_{2m} (\mathbf{B}_{2m}' \mathbf{B}_{2m})^{-1/2}, \\ \mathbf{Q}_m^{\perp'} \mathbf{R}_m (\mathbf{R}_m' \mathbf{R}_m)^{-1/2}, & \quad \mathbf{J}_m^{\perp'} \mathbf{C}_m (\mathbf{C}_m' \mathbf{C}_m)^{-1/2}. \end{aligned}$$

are bounded away from zero and infinity for each $m \in [M]$.

The overall structure of Assumptions (ME1)–(L1) broadly follows [Lam and Cen \(2025\)](#), with modifications to accommodate the multilevel structure and the broader temporal class $\mathcal{G}_T$. Assumption (ME1) allows global and local main effects to differ in strength, but assumes equal strength within each loading matrix for simplicity. By contrast, (L1) allows heterogeneous strengths across loading directions in the common components. Assumptions (ME2) and (F1) place the main-effect and common-component factors in $\mathcal{G}_T$, allowing changes in dynamics as well as short- and long-memory dependence. Under (F1), the entries of $\{\mathbf{z}_{G,t}\}$ and $\{\mathbf{z}_{F,t}^{(m)}\}$ are independent, while the cross-sectional dependence is accommodated through

$\Sigma_{Q,m}$, $\Sigma_{J,m}$, $\Sigma_{R,m}$ and $\Sigma_{C,m}$ in (L1). Assumptions (E1)–(E2) allow cross-sectional and serial dependence in the errors, while the cross-group orthogonality condition in (E1) prevents the errors from generating spurious global signals. Similar to Conditions 2 and 3 in Zhang et al. (2025), Assumption (S1) ensures non-degenerate factor processes and sufficient separation between the local and global loading spaces, so that the local signals are preserved after projecting out the global loading spaces.

Remark 6 *The orthogonality condition $\mathbf{Q}_{e,m}\mathbf{Q}'_{e,n} = \mathbf{0}$ or $\mathbf{J}_{e,m}\mathbf{J}'_{e,n} = \mathbf{0}$ for each $m \neq n$ in (E1) is imposed to ensure that the error matrices are uncorrelated across groups, as required by (IC2). However, this condition could be relaxed, in which case the model would be asymptotically identified. Such a relaxation would introduce additional terms into the cross-group error products and lead to more complicated proofs and convergence rates. We do not pursue the corresponding theoretical extension here, but in simulation in subsection 6.1, this condition is not imposed.*

4.2 Rate of convergence for estimators

We establish the asymptotic rates of convergence for the estimators depicted in Section 3.3. First, we consider the estimators of grand means and main effects.

Theorem 2 (Consistency of grand mean and main effect estimators) *Let Assumptions (IC1), (IC2), (E1) and (E2) hold. Then, for each $m \in [M]$ and $t \in [T]$,*

$$(\hat{c}_t^{(m)} - c_t^{(m)})^2 = O_P\left(\frac{1}{p_m q_m}\right), \quad \frac{1}{p_m} \|\hat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)}\|^2 = O_P\left(\frac{1}{q_m}\right), \quad \frac{1}{q_m} \|\hat{\mathbf{b}}_t^{(m)} - \mathbf{b}_t^{(m)}\|^2 = O_P\left(\frac{1}{p_m}\right).$$

Moreover, for each $m \in [M]$ and $t \in [T]$,

$$(\hat{\mu}_t - \mu_t)^2 = O_P\left(\max_{m \in [M]} \frac{1}{p_m q_m}\right), \quad (\hat{\mu}_t^{(m)} - \mu_t^{(m)})^2 = O_P\left(\max_{m \in [M]} \frac{1}{p_m q_m}\right).$$

Theorem 2 demonstrates the performance of direct moment estimation and reflects the gain from cross-sectional averaging.

(R1) (Rate conditions for main effects) For each $m \in [M]$, define $\nu_{r,m} := p_m^{\gamma_r^{(m)} - \gamma_r}$ and $\nu_{c,m} := q_m^{\gamma_c^{(m)} - \gamma_c}$.

Let $d_{r,*}$ be the largest memory parameter across all elements of α_t and $\alpha_t^{(m)}$, $m \in [M]$, with $d_{c,*}$ defined analogously for β_t and $\beta_t^{(m)}$. We assume that

$$\begin{aligned} \max_{m \in [M]} \left\{ \frac{p_m^{1-\gamma_r}}{q_m}, \frac{p_m^{1-\gamma_r^{(m)}}}{q_m} \right\} &= o(1), & R(T, d_{r,*}) \max_{\substack{m,n \in [M] \\ m \neq n}} \left\{ (\nu_{r,m} \vee \frac{1}{\nu_{r,m}})(1 \vee \nu_{r,n}) \right\} &= o(1), \\ \max_{m \in [M]} \left\{ \frac{q_m^{1-\gamma_c}}{p_m}, \frac{q_m^{1-\gamma_c^{(m)}}}{p_m} \right\} &= o(1), & R(T, d_{c,*}) \max_{\substack{m,n \in [M] \\ m \neq n}} \left\{ (\nu_{c,m} \vee \frac{1}{\nu_{c,m}})(1 \vee \nu_{c,n}) \right\} &= o(1). \end{aligned}$$

The quantities $\nu_{r,m}$ and $\nu_{c,m}$ measure the local main effect strength relative to the corresponding global effect. Assumption (R1) controls how this relative strength can vary with the matrix dimensions and temporal dependence. In particular, (R1) automatically holds when all main effect factors are pervasive.

Theorem 3 (Consistency of initial global main effect loading estimators) *Let the assumptions of Theorem 2 hold. Further assume Assumptions (ME1), (ME2), (R1) and (S1). For each $m \in [M]$, define $\check{\mathbf{A}}_{1m} := \mathbf{A}_{1m}(\mathbf{A}'_{1m}\mathbf{A}_{1m})^{-1/2}$ and $\check{\mathbf{B}}_{1m} := \mathbf{B}_{1m}(\mathbf{B}'_{1m}\mathbf{B}_{1m})^{-1/2}$. Then there exist asymptotically orthogonal matrices $\mathbf{H}_{A_1,m}$ and $\mathbf{H}_{B_1,m}$ with $\mathbf{H}_{A_1,m}\mathbf{H}'_{A_1,m} = \mathbf{I} + o_P(1)$ and $\mathbf{H}_{B_1,m}\mathbf{H}'_{B_1,m} = \mathbf{I} + o_P(1)$, such that*

$$\begin{aligned}\|\hat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m}\mathbf{H}_{A_1,m}\|_F^2 &= O_P(\eta_{A_1m}), \quad \eta_{A_1m} := R(T, d_{r,\star})\nu_{r,m} \max_{n \in [M] \setminus \{m\}} \nu_{r,n} + \frac{p_m^{1-\gamma_r}}{Tq_m} (1 \vee \max_{n \in [M] \setminus \{m\}} \nu_{r,n}), \\ \|\hat{\mathbf{B}}_{1m} - \check{\mathbf{B}}_{1m}\mathbf{H}_{B_1,m}\|_F^2 &= O_P(\eta_{B_1m}), \quad \eta_{B_1m} := R(T, d_{c,\star})\nu_{c,m} \max_{n \in [M] \setminus \{m\}} \nu_{c,n} + \frac{q_m^{1-\gamma_c}}{Tp_m} (1 \vee \max_{n \in [M] \setminus \{m\}} \nu_{c,n}).\end{aligned}$$

Theorem 4 (Consistency of local main effect loading and factor estimators) *Let the assumptions of Theorem 3 hold. For each $m \in [M]$, define $\check{\mathbf{A}}_{2m} := \mathbf{A}_{2m}(\mathbf{A}'_{2m}\mathbf{A}_{2m})^{-1/2}$ and $\check{\mathbf{B}}_{2m} := \mathbf{B}_{2m}(\mathbf{B}'_{2m}\mathbf{B}_{2m})^{-1/2}$. Then there exist asymptotically orthogonal matrices $\mathbf{H}_{A_2,m}$ and $\mathbf{H}_{B_2,m}$ such that*

$$\begin{aligned}\|\hat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m}\mathbf{H}_{A_2,m}\|_F^2 &= O_P\left(R(T, d_{r,\star}) \frac{\max_{n \in [M] \setminus \{m\}} \nu_{r,n}}{\nu_{r,m}} + \frac{p_m^{1-\gamma_r}}{Tq_m} \frac{(1 \vee \max_{n \in [M]} \nu_{r,n})}{\nu_{r,m}^2} + \frac{p_m^{1-2\gamma_r^{(m)}}}{q_m^2}\right), \\ \|\hat{\mathbf{B}}_{2m} - \check{\mathbf{B}}_{2m}\mathbf{H}_{B_2,m}\|_F^2 &= O_P\left(R(T, d_{c,\star}) \frac{\max_{n \in [M] \setminus \{m\}} \nu_{c,n}}{\nu_{c,m}} + \frac{q_m^{1-\gamma_c}}{Tp_m} \frac{(1 \vee \max_{n \in [M]} \nu_{c,n})}{\nu_{c,m}^2} + \frac{q_m^{1-2\gamma_c^{(m)}}}{p_m^2}\right).\end{aligned}$$

Moreover, define $\check{\boldsymbol{\alpha}}_t^{(m)} := (\mathbf{A}'_{2m}\mathbf{A}_{2m})^{1/2}\boldsymbol{\alpha}_t^{(m)}$ and $\check{\boldsymbol{\beta}}_t^{(m)} := (\mathbf{B}'_{2m}\mathbf{B}_{2m})^{1/2}\boldsymbol{\beta}_t^{(m)}$. Then,

$$\begin{aligned}\frac{1}{p_m^{\gamma_r^{(m)}}} \|\hat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{H}_{A_2,m}^{-1} \check{\boldsymbol{\alpha}}_t^{(m)}\|^2 &= O_P\left(\frac{1}{p_m^{\gamma_r^{(m)}} q_m} + R(T, d_{r,\star}) \frac{\max_{n \in [M] \setminus \{m\}} \nu_{r,n}}{1 \wedge \nu_{r,m}} + \frac{p_m^{1-\gamma_r^{(m)}}}{Tq_m} \frac{1 \vee \max_{n \in [M]} \nu_{r,n}}{1 \wedge \nu_{r,m}}\right), \\ \frac{1}{q_m^{\gamma_c^{(m)}}} \|\hat{\boldsymbol{\beta}}_t^{(m)} - \mathbf{H}_{B_2,m}^{-1} \check{\boldsymbol{\beta}}_t^{(m)}\|^2 &= O_P\left(\frac{1}{q_m^{\gamma_c^{(m)}} p_m} + R(T, d_{c,\star}) \frac{\max_{n \in [M] \setminus \{m\}} \nu_{c,n}}{1 \wedge \nu_{c,m}} + \frac{q_m^{1-\gamma_c^{(m)}}}{Tp_m} \frac{1 \vee \max_{n \in [M]} \nu_{c,n}}{1 \wedge \nu_{c,m}}\right).\end{aligned}$$

Theorem 5 (Consistency of global main effect factor and refitted loading estimators) *Let the assumptions in Theorem 4 hold. Define $\check{\boldsymbol{\alpha}}_t := (T^{-1} \sum_{s=1}^T \boldsymbol{\alpha}_s \boldsymbol{\alpha}'_s)^{-1/2} \boldsymbol{\alpha}_t$ and $\check{\boldsymbol{\beta}}_t := (T^{-1} \sum_{s=1}^T \boldsymbol{\beta}_s \boldsymbol{\beta}'_s)^{-1/2} \boldsymbol{\beta}_t$. Then there exist asymptotically orthogonal matrices $\mathbf{H}_\alpha$ and $\mathbf{H}_\beta$ such that, for each $t \in [T]$,*

$$\begin{aligned}\|\hat{\boldsymbol{\alpha}}_t - \mathbf{H}_\alpha^{-1} \check{\boldsymbol{\alpha}}_t\|^2 &= O_P\left(\max_{m \in [M]} \left\{ \frac{1}{p_m^{\gamma_r} q_m} + R(T, d_{r,\star}) (1 \vee \nu_{r,m}) \max_{n \in [M] \setminus \{m\}} \nu_{r,n} + \frac{p_m^{1-\gamma_r}}{Tq_m} \frac{1 \vee \max_{n \in [M]} \nu_{r,n}}{1 \wedge \nu_{r,m}} \right\}\right), \\ \|\hat{\boldsymbol{\beta}}_t - \mathbf{H}_\beta^{-1} \check{\boldsymbol{\beta}}_t\|^2 &= O_P\left(\max_{m \in [M]} \left\{ \frac{1}{q_m^{\gamma_c} p_m} + R(T, d_{c,\star}) (1 \vee \nu_{c,m}) \max_{n \in [M] \setminus \{m\}} \nu_{c,n} + \frac{q_m^{1-\gamma_c}}{Tp_m} \frac{1 \vee \max_{n \in [M]} \nu_{c,n}}{1 \wedge \nu_{c,m}} \right\}\right).\end{aligned}$$

Moreover, let η_{A_1m} and η_{B_1m} denote the rates in Theorem 3, and η_α and η_β denote the rates in the two

factor bounds above. Define $\check{\check{\mathbf{A}}}_{1m} := \mathbf{A}_{1m}(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t')^{1/2}$ and $\check{\check{\mathbf{B}}}_{1m} := \mathbf{B}_{1m}(T^{-1} \sum_{t=1}^T \boldsymbol{\beta}_t \boldsymbol{\beta}_t')^{1/2}$. Then,

$$\begin{aligned} \frac{1}{p_m^{\gamma_r}} \|\tilde{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_\alpha\|_F^2 &= O_P\left(\eta_{A_{1m}} + \left\{ \frac{p_m^{1-\gamma_r}}{q_m} + R(T, d_{r,\star})(1 \vee \nu_{r,m}) \max_{n \in [M] \setminus \{m\}} \nu_{r,n} \right\} \eta_\alpha\right), \\ \frac{1}{q_m^{\gamma_c}} \|\tilde{\mathbf{B}}_{1m} - \check{\check{\mathbf{B}}}_{1m} \mathbf{H}_\beta\|_F^2 &= O_P\left(\eta_{B_{1m}} + \left\{ \frac{q_m^{1-\gamma_c}}{p_m} + R(T, d_{c,\star})(1 \vee \nu_{c,m}) \max_{n \in [M] \setminus \{m\}} \nu_{c,n} \right\} \eta_\beta\right). \end{aligned}$$

Theorems 3-5 reveal how the relative strengths $\nu_{r,m}$ and $\nu_{c,m}$ induce a trade-off between global and local estimation. The refitted loading $\tilde{\mathbf{A}}_{1m}$ inherits both the error of $\hat{\mathbf{A}}_{1m}$ and that of the estimated global main effect factor series. Although refitting does not necessarily improve the loading convergence rate, it aligns the group-specific global loadings to a common temporal coordinate. This alignment substantially improves the estimation of the global main effect products, as demonstrated in subsection 6.1.2.

A useful benchmark is obtained when all main effect factors are pervasive and their memory parameters are below 1/4, as summarized in the following corollary.

Corollary 6 (Simplified rates of main effect estimators) *Let Assumptions (IC1), (IC2), (E1), (E2) and (S1) hold. Suppose that (ME1) and (ME2) hold with pervasive main effect factors, i.e., $\gamma_r = \gamma_c = \gamma_r^{(m)} = \gamma_c^{(m)} = 1$ for all $m \in [M]$, and that all main effect memory parameters are less than 1/4. Then, for each $m \in [M]$ and $t \in [T]$,*

- $\|\hat{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_{A_{1,m}}\|_F^2 = O_P(1/T)$, $\|\hat{\mathbf{B}}_{1m} - \check{\check{\mathbf{B}}}_{1m} \mathbf{H}_{B_{1,m}}\|_F^2 = O_P(1/T)$.
- $\|\hat{\mathbf{A}}_{2m} - \check{\check{\mathbf{A}}}_{2m} \mathbf{H}_{A_{2,m}}\|_F^2 = O_P(1/T + 1/(p_m q_m^2))$, $\|\hat{\mathbf{B}}_{2m} - \check{\check{\mathbf{B}}}_{2m} \mathbf{H}_{B_{2,m}}\|_F^2 = O_P(1/T + 1/(q_m p_m^2))$.
- $1/p_m \|\hat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{H}_{A_{2,m}}^{-1} \check{\check{\boldsymbol{\alpha}}}_t^{(m)}\|^2 = O_P(1/T + 1/(p_m q_m))$, $1/q_m \|\hat{\boldsymbol{\beta}}_t^{(m)} - \mathbf{H}_{B_{2,m}}^{-1} \check{\check{\boldsymbol{\beta}}}_t^{(m)}\|^2 = O_P(1/T + 1/(p_m q_m))$.
- $\|\hat{\boldsymbol{\alpha}}_t - \mathbf{H}_\alpha^{-1} \check{\check{\boldsymbol{\alpha}}}_t\|^2 = O_P(1/T + \max_{n \in [M]} 1/(p_n q_n))$, $\|\hat{\boldsymbol{\beta}}_t - \mathbf{H}_\beta^{-1} \check{\check{\boldsymbol{\beta}}}_t\|^2 = O_P(1/T + \max_{n \in [M]} 1/(p_n q_n))$.
- $1/p_m \|\tilde{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_\alpha\|_F^2 = O_P(1/T + 1/q_m \max_{n \in [M]} 1/(p_n q_n))$, $1/q_m \|\tilde{\mathbf{B}}_{1m} - \check{\check{\mathbf{B}}}_{1m} \mathbf{H}_\beta\|_F^2 = O_P(1/T + 1/p_m \max_{n \in [M]} 1/(p_n q_n))$.

Before imposing some regularity conditions for Step 3 in Section 3.3, recall $\mathbf{v}_a^{(m)}$ and $\mathbf{v}_b^{(m)}$ therein. Then define $s_{Q,m} := \|\mathbf{Q}_m' \mathbf{v}_a^{(m)}\|$ and $s_{J,m} := \|\mathbf{J}_m' \mathbf{v}_b^{(m)}\|$, which are similarly seen in Chen and Lam (2024). In particular, we would have $s_{Q,m} \asymp p_m^{\delta_{q,1}/2}$ when $\mathbf{v}_a^{(m)}$ aligns with the column of $\mathbf{Q}_m$ corresponding to the strongest factor. In the following, we present some assumptions on the projection vectors $\mathbf{v}_a^{(m)}$ and $\mathbf{v}_b^{(m)}$.

(L2) (Cross-group alignment of projection directions) There exists an absolute constant $c > 0$ such that for all $m, n \in [M]$, with probability tending to one,

$$|\mathbf{v}_b^{(m)'} \mathbf{J}_m \mathbf{J}_n' \mathbf{v}_b^{(n)}| \geq c \|\mathbf{J}_m' \mathbf{v}_b^{(m)}\| \|\mathbf{J}_n' \mathbf{v}_b^{(n)}\|, \quad |\mathbf{v}_a^{(m)'} \mathbf{Q}_m \mathbf{Q}_n' \mathbf{v}_a^{(n)}| \geq c \|\mathbf{Q}_m' \mathbf{v}_a^{(m)}\| \|\mathbf{Q}_n' \mathbf{v}_a^{(n)}\|.$$

(R2) (Signal dominance for global common component) Let $d_{\text{fac},\star}$ be the largest long-memory exponent

among all elements of $\mathbf{G}_t$ and $\mathbf{F}_t^{(m)}$ for $m \in [M]$. Then for all $m, n \in [M]$ with $m \neq n$, assume that

$$\begin{aligned} p_m^{\delta_{r,1}^{(m)} - \delta_{q,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} &= O(s_{J,m}^2), \quad p_n^{\delta_{r,1}^{(n)} - \delta_{q,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} = O(s_{J,n}^2), \quad p_m^{1 - \delta_{q,1}^{(m)}} = O(s_{J,m}^2), \quad p_n^{1 - \delta_{q,1}^{(n)}} = O(s_{J,n}^2), \\ q_m^{\delta_{c,1}^{(m)} - \delta_{j,1}^{(m)}} p_m^{\delta_{r,1}^{(m)}} &= O(s_{Q,m}^2), \quad q_n^{\delta_{c,1}^{(n)} - \delta_{j,1}^{(n)}} p_n^{\delta_{r,1}^{(n)}} = O(s_{Q,n}^2), \quad q_m^{1 - \delta_{j,1}^{(m)}} = O(s_{Q,m}^2), \quad q_n^{1 - \delta_{j,1}^{(n)}} = O(s_{Q,n}^2). \end{aligned}$$

Assumption (L2) formalizes that the selected projection directions align with the same latent global column and row factor directions across different groups. This assumption would hold if $\mathbf{v}_b^{(m)}$ and $\mathbf{v}_b^{(n)}$ are the leading singular directions of $\mathbf{J}_m$ and $\mathbf{J}_n$ correspond to the same strongest global factor. Our numerical results in Section 6.1.3 show that the iterative projection procedure in Section A estimate those vectors appropriately.

Theorem 7 (Consistency of local loading and factor estimators) *Let Assumptions (IC1), (E1), (E2), (F1), (L1), (L2), and (R2) hold. Define $\check{\mathbf{R}}_m := \mathbf{R}_m(\mathbf{R}_m' \mathbf{R}_m)^{-1/2}$, $\check{\mathbf{C}}_m := \mathbf{C}_m(\mathbf{C}_m' \mathbf{C}_m)^{-1/2}$, and*

$$\begin{aligned} \eta_{Q,m} &:= R(T, d_{fac,*}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2, \\ \eta_{J,m} &:= R(T, d_{fac,*}) q_m^{2(\delta_{j,1}^{(m)} - \delta_{j,k_c}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{Q,n}^2 q_n^{\delta_{j,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{Q,n}^2 q_n^{\delta_{j,k_c}^{(n)}}} \right)^2. \end{aligned}$$

Further assume $\eta_{Q,m}, \eta_{J,m} = o(1)$. Then there exist asymptotically orthogonal matrices $\mathbf{H}_{R,m}$ and $\mathbf{H}_{C,m}$ with $\mathbf{H}_{R,m} \mathbf{H}_{R,m}' = \mathbf{I} + o_P(1)$ and $\mathbf{H}_{C,m} \mathbf{H}_{C,m}' = \mathbf{I} + o_P(1)$, such that

$$\begin{aligned} \|\widehat{\mathbf{R}}_m - \check{\mathbf{R}}_m \mathbf{H}_{R,m}\|_F^2 &= O_P(\Delta_{R,m}), \quad \|\widehat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m}\|_F^2 = O_P(\Delta_{C,m}), \quad \text{where} \\ \Delta_{R,m} &:= p_m^{-2\delta_{r,1}^{(m)}} q_m^{-2\delta_{c,1}^{(m)}} \left\{ \left(p_m^{2\delta_{r,1}^{(m)}} q_m^{2\delta_{c,1}^{(m)}} + p_m^{1+\delta_{q,1}^{(m)}} q_m^{1+\delta_{j,1}^{(m)}} \right) \eta_{J,m} + \frac{p_m^{1+\delta_{r,1}^{(m)}} q_m^{1+\delta_{c,1}^{(m)}}}{T} + p_m^{2(1-\delta_{r,1}^{(m)})} + \frac{p_m^2 q_m}{T} \right\}, \\ \Delta_{C,m} &:= q_m^{-2\delta_{c,1}^{(m)}} p_m^{-2\delta_{r,1}^{(m)}} \left\{ \left(q_m^{2\delta_{c,1}^{(m)}} p_m^{2\delta_{r,1}^{(m)}} + q_m^{1+\delta_{j,1}^{(m)}} p_m^{1+\delta_{q,1}^{(m)}} \right) \eta_{Q,m} + \frac{q_m^{1+\delta_{c,1}^{(m)}} p_m^{1+\delta_{r,1}^{(m)}}}{T} + q_m^{2(1-\delta_{c,1}^{(m)})} + \frac{q_m^2 p_m}{T} \right\}. \end{aligned}$$

Further let Assumption (S1) hold, and define $\check{\mathbf{F}}_t^{(m)} := (\mathbf{R}_m' \mathbf{R}_m)^{1/2} \mathbf{F}_t^{(m)} (\mathbf{C}_m' \mathbf{C}_m)^{1/2}$. Then

$$\begin{aligned} \frac{1}{p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}}} \|\widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}\|_F^2 &= O_P(\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}), \quad \text{where} \\ \Delta_{F,m} &:= \min \left(q_m^{1-2\delta_{c,1}^{(m)}} p_m^{1-\delta_{r,1}^{(m)}} \Delta_{R,m} + q_m^{1-2\delta_{c,1}^{(m)}} p_m^{-\delta_{r,1}^{(m)}}, p_m^{1-2\delta_{r,1}^{(m)}} q_m^{1-\delta_{c,1}^{(m)}} \Delta_{C,m} + p_m^{1-2\delta_{r,1}^{(m)}} q_m^{-\delta_{c,1}^{(m)}} \right). \end{aligned}$$

Theorem 7 establishes the consistency of the local factor loading and factor estimators. The terms $\eta_{Q,m}$ and $\eta_{J,m}$ measure the error inherited from the estimated global loading directions, and the condition

$\eta_{Q,m}, \eta_{J,m} = o(1)$ ensures that the projected cross-group moment matrices $\mathbf{X}_{C,m}$ and $\mathbf{X}_{R,m}$ recover the local factor spaces consistently. When all common factors are strong, i.e., all δ 's equal one, and dimensions are comparable, then $\Delta_{R,m}$ and $\Delta_{C,m}$ become $O_P(T^{-1})$. The additional Assumption (S1) then yields the rate for the local factor series $\widehat{\mathbf{F}}_t^{(m)}$, whose leading term combines $\Delta_{R,m}$, $\Delta_{C,m}$, and an additional term $\Delta_{F,m}$ accounting for the asymmetry between the row and column sides in the refitting.

Theorem 8 (Consistency of global loading and factor estimators) *Let all assumptions in Theorem 7 hold. Let $m_0 \in [M]$ be a fixed reference group, and define $\check{\mathbf{G}}_t := (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} \mathbf{G}_t (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2}$. For any $m \in [M]$, define*

$$\begin{aligned} \Delta_{G,m} &:= \min(p_m, q_m) + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + (\eta_{Q,m} + \eta_{J,m}) p_m q_m, \\ S_m &:= \lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m) \asymp p_m^{\delta_{q,kr}^{(m)}} q_m^{\delta_{j,kc}^{(m)}}. \end{aligned}$$

and assume $\Delta_{G,m} = o\{\lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m)\}$. Further assume that the reference group m_0 is chosen such that for any $m \in [M]$, $\lambda_{\min}(\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0}) \lambda_{\min}(\mathbf{J}'_{m_0} \mathbf{J}_{m_0}) = O\{\lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m)\}$. Then there exist asymptotically orthogonal matrices $\mathbf{H}_{Q,m_0}$ and $\mathbf{H}_{J,m_0}$ such that

$$\|\widehat{\mathbf{G}}_t - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0}\|_F^2 = O_P(\Delta_G), \quad \text{where} \quad \Delta_G := \max_{m \in [M]} \Delta_{G,m} + \left(\max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}} \right) \max_{m \in [M]} S_m.$$

Also, we have

$$\|\tilde{\mathbf{Q}}_m - \mathbf{Q}_m (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{-1/2} \mathbf{H}_{Q,m_0}\|_F^2 = O_P(\Delta_{K,m}), \quad \text{where} \quad \Delta_{K,m} := \Delta_G \frac{p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}}}{p_{m_0}^{2\delta_{q,kr}^{(m_0)}} q_{m_0}^{2\delta_{j,kc}^{(m_0)}}} + \frac{p_m q_m}{T S_{m_0}},$$

and similarly $\|\tilde{\mathbf{J}}_m - \mathbf{J}_m (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{-1/2} \mathbf{H}_{J,m_0}\|_F^2 = O_P(\Delta_{K,m})$.

Built upon Theorem 7, Theorem 8 addresses the global common component. The quantity Δ_G has two parts. The first, $\max_{m \in [M]} \Delta_{G,m}$, aggregates the local-component estimation errors across groups. The second term in Δ_G arises from the two-sided alignment and the refitting step, where S_m measures the smallest squared signal in the m -th group's projected global loading matrices. The signal-dominance condition $\Delta_{G,m} = o(S_m)$ ensures that the perturbed Gram matrices in the alignment step remain well-conditioned, and the reference-group condition guarantees the alignment from any group to the reference group does not blow up. Under these conditions, the first part of the theorem shows that $\widehat{\mathbf{G}}_t$ is consistent for the aligned version $\mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0}$ of the global factor, which is infeasible by merely using the group-wise “global” part in Zhang et al. (2025). Then the theorem gives the rates for the refitted loadings $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$. In the strong-factor setting where all δ 's equal one and all groups have comparable dimensions,

both terms reduce to $O_P(T^{-1})$, delivering consistency of the refitted global loadings up to the asymptotic orthogonal matrices $\mathbf{H}_{Q,m_0}$ and $\mathbf{H}_{J,m_0}$.

4.3 Detecting the existence of global common factors

So far we have assumed that the global factors in $\mathbf{G}_t$ drive the dynamics of each data set $\mathbf{Y}_t^{(m)}$ for $m \in [M]$. However, if we are exploring new data sets which may not share common dynamics with others, blindly applying Algorithm 1 in Section A will give us a false impression that there are global factors driving the dynamics of these new data sets, when there may be none. It is important to first investigate if there are global factors driving the dynamics of any data set at all.

To this end, suppose we have $\hat{\mathbf{Q}}_m$ and $\hat{\mathbf{J}}_m$ as the output of Algorithm 1 in Section A. Also, for $m \in [M]$, let $\mathbf{U}_m = (\mathbf{Q}_m, \mathbf{R}_m)$ and $\mathbf{V}_m = (\mathbf{J}_m, \mathbf{C}_m)$, and $\hat{\mathbf{U}}_{m,c_m}$ and $\hat{\mathbf{V}}_{m,d_m}$ be the c_m th and d_m th eigenvectors of

$$T^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} \mathbf{Y}_t^{(m)'} \mathbf{M}_{p_m} \quad \text{and} \quad T^{-1} \sum_{t=1}^T \mathbf{M}_{q_m} \mathbf{Y}_t^{(m)'} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m},$$

respectively (Lam and Cen, 2025). This means that we are using the data $\{\mathbf{Y}_t^{(m)}\}_{t \in [T]}$ only to estimate the MMEFM in (3.1) without distinguishing global and local factors.

Theorem 9 *Let all the conditions in Theorem 7 hold. Define the global active set as*

$$\mathcal{A}_G := \{m \in [M] : \text{there exists } n \neq m \in [M] \text{ such that } \{\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_n'\}_{t \in [T]} \text{ and } \{\mathbf{Q}_n \mathbf{G}_t \mathbf{J}_m'\}_{t \in [T]} \text{ are nonzero}\}, \quad (4.2)$$

and the statistics

$$S_{G,m} := \frac{\sum_{n=1; n \neq m}^M \max_{c_m, c_n \in [k_r], d_m, d_n \in [k_c]} \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{Q}}_{m,c_m}' \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{J}}_{m,d_m}' \hat{\mathbf{J}}_{n,d_n}' \check{\mathbf{Y}}_t^{(n)'} \hat{\mathbf{Q}}_{n,c_n} \right\}^2}{\max_{c_m \in [k_r + k_r^{(m)}], d_m \in [k_c + k_c^{(m)}]} \left(T^{-1} \sum_{t=1}^T \hat{\mathbf{U}}_{m,c_m}' \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{V}}_{m,d_m}' \hat{\mathbf{V}}_{m,d_m}' \check{\mathbf{Y}}_t^{(m)'} \hat{\mathbf{U}}_{m,c_m} \right)^2}, \quad m \in [M]. \quad (4.3)$$

Then for any $m \in \mathcal{A}_G$ and $k \notin \mathcal{A}_G$, assuming

$$\max_{n \neq m, n \in \mathcal{A}_G} p_n^{\delta_{q,k_r}^{(n)}} q_n^{\delta_{j,k_c}^{(n)}} \gtrsim \max \left(\max_{n \neq m, n \in \mathcal{A}_G} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}, \max_{n \notin \mathcal{A}_G} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} \right) R(T, d_{fac}, \star),$$

we have

$$S_{G,m} \gtrsim_P \frac{\max_{n \neq m, n \in \mathcal{A}_G} p_n^{\delta_{q,k_r}^{(n)}} q_n^{\delta_{j,k_c}^{(n)}}}{p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}}}, \quad \text{and}$$

$$S_{G,k} \lesssim_P \frac{\max \left(\max_{n \in \mathcal{A}_G} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}, \max_{n \neq k, n \notin \mathcal{A}_G} p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} \right)}{p_k^{\delta_{r,1}^{(k)}} q_k^{\delta_{c,1}^{(k)}}} R(T, d_{fac, \star}),$$

where $R(\cdot, \cdot)$ is defined in (4.1) and $d_{fac, \star}$ in Assumption (R2). In particular, if $p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \asymp p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}} \asymp p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}$ for any $n, m \in \mathcal{A}_G$, and $\max_{n \in [M]} (p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}, p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}}) R(T, d_{fac, \star}) = o(\min_{k \notin \mathcal{A}_G} p_k^{\delta_{r,1}^{(k)}} q_k^{\delta_{c,1}^{(k)}})$, then

$$S_{G,m} \asymp_P 1, \quad S_{G,k} = o_P(1).$$

The last conditions in the above theorem translate to that all signals from the global common factors from different groups are of the same order, while the maximum global or local signals are of a smaller order than the minimum local ones after multiplying a factor of $R(T, d_{fac, \star})$. In practice, we need to replace $k_r, k_r^{(m)}, k_c, k_c^{(m)}$ in the definition of $S_{G,m}$ by their estimators described in Section 3.4. We stress that over-estimating these numbers of factors is not much of a problem, but under-estimating them can lead to failure in detecting the existence of global common factors potentially. Below we provide a theorem for the rank estimators for all the loading matrices presented in Section 3.4.

Theorem 10 *Let all the conditions in Theorem 7 hold. Then for each $m \in [M]$, as $\min(T, p_m, q_m) \rightarrow \infty$,*

$$P(\hat{r}_1 \geq r_1) \rightarrow 1, \quad P(\hat{\ell}_1 \geq \ell_1) \rightarrow 1, \quad P(\hat{r}_2^{(m)} \geq r_2^{(m)}) \rightarrow 1, \quad P(\hat{\ell}_2^{(m)} \geq \ell_2^{(m)}) \rightarrow 1,$$

if $\eta_{A_{1m}} = o(\nu_{r,m}^2)$ and $\eta_{B_{1m}} = o(\nu_{c,m}^2)$, where $\eta_{A_{1m}}$ and $\eta_{B_{1m}}$ are the rates defined in Theorem 3. Moreover, we also have $P(\hat{k}_r \geq k_r) \rightarrow 1$ and $P(\hat{k}_c \geq k_c) \rightarrow 1$ if

$$\begin{aligned} & \max_{m, n \in [M], n \neq m} p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}} \cdot \max_{m, n \in [M], n \neq m} [T^{-1} p_m q_m p_n q_n + p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}} R(T, d_{fac, \star})] \\ &= o \left(\min \left(\max_{m, n \in [M], n \neq m} p_m^{2\delta_{q,1}^{(m)}} q_m^{2\delta_{j,1}^{(m)}} p_n^{2\delta_{q,1}^{(n)}} q_n^{2\delta_{j,1}^{(n)}}, \max_{m, n \in [M], n \neq m} p_m^{2\delta_{q,1}^{(m)}} q_m^{2\delta_{j,1}^{(m)}} p_n^{2\delta_{q,1}^{(n)}} q_n^{2\delta_{j,1}^{(n)}} \right) \right). \end{aligned}$$

Finally, we also have $P(\hat{k}_r^{(m)} \geq k_r^{(m)}) \rightarrow 1$ and $P(\hat{k}_c^{(m)} \geq k_c^{(m)}) \rightarrow 1$ if

$$\begin{aligned} & q_m^{2(1-\delta_{c,1}^{(m)})} + T^{-1} q_m^2 p_m + p_m^{1+\delta_{j,1}^{(m)}} p_m^{1+\delta_{q,1}^{(m)}} \eta_{Q,m} = o(q_m^{4\delta_{c,k_c}^{(m)}} - 2\delta_{c,1}^{(m)} p_m^{2\delta_{r,1}^{(m)}}), \\ & p_m^{2(1-\delta_{r,1}^{(m)})} + T^{-1} p_m^2 q_m + p_m^{1+\delta_{q,1}^{(m)}} q_m^{1+\delta_{j,1}^{(m)}} \eta_{J,m} = o(p_m^{4\delta_{r,k_r}^{(m)}} - 2\delta_{r,1}^{(m)} q_m^{2\delta_{c,1}^{(m)}}). \end{aligned}$$

If all factors are strong and all processes have short memory, the greater than or equal sign in all asymptotic probabilities above will become the equal sign, if we add the following terms to the numerator and denominator of the eigenvalue ratios defining $\hat{r}_1, \hat{r}_2^{(m)}, \hat{\ell}_1, \hat{\ell}_2^{(m)}, \hat{k}_r, \hat{k}_c, \hat{k}_r^{(m)}$ and $\hat{k}_c^{(m)}$ respectively (C is a

generic constant):

$$\begin{aligned} & \max_{m \neq n \in [M]} CT^{-1}p_m p_n, CT^{-1}p_m^2, \max_{m \neq n \in [M]} CT^{-1}q_m q_n, CT^{-1}q_m^2 \\ & \max_{m \neq n \in [M]} CT^{-1}p_m q_m p_n q_n, \max_{m \neq n \in [M]} CT^{-1}p_m q_m p_n q_n, CT^{-1/2}p_m q_m, CT^{-1/2}p_m q_m. \end{aligned}$$

The addition of the rates in the last part of Theorem 10 stabilizes the eigenvalue ratios to constant order after the dip to $o_P(1)$ at the true number of factors, hence proving the consistency of the rank estimators.

5 Practical Detection and Screening of Global Common Factors

Theorem 9 provides a theoretical grounding for detecting global common factors through the group-level statistics $S_{G,m}$, $m \in [M]$. From this, we derive two practical procedures for detecting which group of data, $\{\mathbf{Y}_t^{(m)}\}_{t \in [T]}$ for $m \in [M]$, share the global common factors. The first one investigates whether any nontrivial global common factors exist, and the second one screens which groups are driven by these factors.

For practical implementation, we standardize $\check{\mathbf{Y}}_t^{(m)}$ and define

$$\check{\mathbf{Z}}_t^{(m)} = \check{\mathbf{Y}}_t^{(m)} / \hat{\sigma}_m, \quad \hat{\sigma}_m^2 = (Tp_m q_m)^{-1} \sum_{t=1}^T \|\check{\mathbf{Y}}_t^{(m)}\|_F^2. \quad (5.1)$$

The statistic $S_{G,m}$ in (4.3) is then computed using $\check{\mathbf{Z}}_t^{(m)}$ in place of $\check{\mathbf{Y}}_t^{(m)}$. This normalization reduces the influence of purely marginal scale differences between groups, preventing signal domination from groups just for their larger scales.

Let $S_{G,(1)} \geq S_{G,(2)} \geq \dots \geq S_{G,(M)}$ denote the decreasing order statistics of $S_{G,1}, \dots, S_{G,M}$. Testing the existence of global factors is equivalent to testing the following hypotheses:

$$H_{0,G} : \mathcal{A}_G = \emptyset \quad \text{v.s.} \quad H_{1,G} : |\mathcal{A}_G| \geq 2.$$

Theorem 9 implies that $S_{G,(1)}$ could be used for this test, since all $S_{G,m}$ decay under $H_{0,G}$, while at least two are of constant order under $H_{1,G}$. Nevertheless, we use $S_{G,(2)}$ because $H_{1,G}$ requires the global common component to be shared by at least two groups, so second-largest statistic directly reflects the alternative. Given an existence cutoff τ_1 , we reject $H_{0,G}$ when $S_{G,(2)} > \tau_1$. After rejecting the null, we may further screen which groups share the global common factors. Let τ_2 be a screening cutoff and define the initial

selected set $\hat{\mathcal{A}}_G^\circ = \{m \in [M] : S_{G,m} > \tau_2\}$. The final selected set is then

$$\hat{\mathcal{A}}_G = \begin{cases} \hat{\mathcal{A}}_G^\circ, & S_{G,(2)} > \tau_1 \text{ and } |\hat{\mathcal{A}}_G^\circ| \geq 2, \\ \emptyset, & \text{otherwise.} \end{cases}$$

The definition of $\hat{\mathcal{A}}_G$ enforces the convention that a global common component must involve at least two groups, cf. Remark 3 of [Cen \(2026\)](#). It also prevents reporting isolated selected groups when the evidence for a global factor is insufficient.

We next discuss how to choose the cutoffs τ_1 and τ_2 . Although Theorem 9 suggests that inactive statistics decay at the rate $R(T, d_{\text{fac}, \star})$, we do not directly use a deterministic rate threshold in the main procedure, since the finite-sample magnitude of $S_{G,m}$ can still depend on constants related to M , the dimensions p_m, q_m , factor strengths, among others. Instead, we use a circular-shift bootstrap to choose the existence and screening cutoffs τ_1 and τ_2 .

The bootstrap is designed to preserve the within-group temporal dependence and marginal residual structure, while breaking the synchronous cross-group alignment that would be induced by global common factors. The procedure is summarized as follows.

Algorithm 1 Circular-shift bootstrap calibration for global common factor detection and screening

1. Given the standardized arrays $\{\check{\mathbf{Z}}_t^{(m)}\}$ from (5.1), choose the bootstrap size B and calibration level α . Fix one reference group $m_0 \in [M]$.
2. For each bootstrap replication $b = 1, \dots, B$, keep the reference group fixed: $\check{\mathbf{Z}}_t^{*(m_0, b)} = \check{\mathbf{Z}}_t^{(m_0)}$ for $t \in [T]$. For each $m \neq m_0$, draw an independent circular shift $\ell_m^{(b)}$ uniformly from $\{0, 1, \dots, T-1\}$, and define $\check{\mathbf{Z}}_t^{*(m, b)} = \check{\mathbf{Z}}_{1+\{(t+\ell_m^{(b)}-1) \bmod T\}}^{(m)}$ for $t \in [T]$.
3. Reapply the detection procedure used to compute $S_{G,m}$ to the shifted arrays $\{\check{\mathbf{Z}}_t^{*(m, b)}\}$, and obtain $S_{G,1}^{*(b)}, \dots, S_{G,M}^{*(b)}$.
4. For each $b \in [B]$, let $S_{G,(1)}^{*(b)} \geq S_{G,(2)}^{*(b)} \geq \dots \geq S_{G,(M)}^{*(b)}$ denote the decreasing order statistics of $S_{G,1}^{*(b)}, \dots, S_{G,M}^{*(b)}$. After B bootstrap replications, set

$$\tau_1 = \tau_2 = \hat{q}_{1-\alpha} \left\{ S_{G,(2)}^{*(b)} : b = 1, \dots, B \right\},$$

where $\hat{q}_{1-\alpha}$ denotes the empirical $(1-\alpha)$ -quantile.

We use $S_{G,(2)}^{*(b)}$ to calibrate both τ_1 and τ_2 , aligning the screening rule with the existence test. Comparisons in Section E.1.2 of the Supplementary Materials show that the analogous $S_{G,(1)}$ -based procedure

has similar existence-test performance but is more conservative in screening and can substantially reduce recovery.

Remark 7 *The circular-shift bootstrap in Algorithm 2 is formulated under the assumption that the factor series have a single stationary segment. For data with known or unknown structural breaks, or more generally with multiple regimes, whole-sample circular shifting may mix observations from different regimes. A possible extension is to first divide the sample into regimes, either using known break points or estimated break points from change point detection methods, and then apply circular shifting within each regime.*

6 Numerical Results

6.1 Simulation

6.1.1 Data generating process

We generate a collection of matrix time series $\{\mathbf{Y}_t^{(m)}\}_{t \in [T]}$, where $\mathbf{Y}_t^{(m)} \in \mathbb{R}^{p_m \times q_m}$, from the MMEFM.

Loadings: For each group m , the entries of $\mathbf{A}_{1m}$, $\mathbf{A}_{2m}$, $\mathbf{B}_{1m}$, $\mathbf{B}_{2m}$, $\mathbf{Q}_m$, $\mathbf{J}_m$, $\mathbf{R}_m$, and $\mathbf{C}_m$ are first generated independently from $N(0, 1)$. We impose the desired factor strengths by columnwise rescaling. Specifically, if $\mathbf{L}$ is a loading matrix with p rows and its j th column has strength parameter $\xi_j \in (1/2, 1]$, we set $\mathbf{L}_{\cdot, j} = p^{-(1-\xi_j)/2} \mathbf{Z}_{\cdot, j}$ with $\mathbf{Z}_{\cdot, j} \sim N(\mathbf{0}, \mathbf{I}_p)$, so that $\|\mathbf{L}_{\cdot, j}\|_2^2 \asymp p^{\xi_j}$. Each column is then demeaned to satisfy (IC1). We set $r_1 = r_2^{(m)} = \ell_1 = \ell_2^{(m)} = 2$ for the main effects, $k_r = k_c = 3$ for the global common component, and $k_r^{(m)} = k_c^{(m)} = 2$ for the local common components.

Factor dynamics: All latent factor series are generated according to the piecewise-stationary construction described in subsection 4.1. This applies to μ_t , $\mu_t^{(m)}$, and the entries of $\boldsymbol{\alpha}_t$, $\boldsymbol{\beta}_t$, $\boldsymbol{\alpha}_t^{(m)}$, $\boldsymbol{\beta}_t^{(m)}$, $\mathbf{G}_t$, and $\mathbf{F}_t^{(m)}$. We use two equal-length segments. For a generic scalar series, we first generate $x_t = \sum_{w \geq 0} \phi^w \eta_{t-w}$ with $\phi = 0.6$ and $\eta_t \sim N(0, 1)$ independently over t . Within segment s , we then set $y_t = (1 - B)^{-d_s} x_t$, where B denotes the backshift operator and $(1 - B)^{-d_s} = \sum_{w \geq 0} \pi_w(d_s) B^w$, with $\pi_0(d_s) = 1$ and $\pi_w(d_s) = \pi_{w-1}(d_s)(w-1+d_s)/w$ for $w \geq 1$. When $d_s = 0$, this reduces to the short-memory AR(1) case. This yields segmentwise stationary dynamics with polynomially decaying dependence when $d_s > 0$, while allowing the structure to change at the midpoint. To enforce the sample orthogonality in (IC2), let $\mathbf{A} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_T]'$ and $\mathbf{A}^{(m)} = [\boldsymbol{\alpha}_1^{(m)}, \dots, \boldsymbol{\alpha}_T^{(m)}]'$. We replace $\mathbf{A}^{(m)}$ by $(\mathbf{I} - \mathbf{P}_\alpha) \mathbf{A}^{(m)}$, where $\mathbf{P}_\alpha = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$. Thus $\mathbf{A}'\mathbf{A}^{(m)} = 0$. We apply the same construction to $\boldsymbol{\beta}_t$ and $\boldsymbol{\beta}_t^{(m)}$. The local grand means are demeaned across groups at each t so that $\sum_{m=1}^M \mu_t^{(m)} = 0$.

Idiosyncratic noise: For $\mathbf{Q}_{e,m}$, $\mathbf{J}_{e,m}$, $\mathbf{R}_{e,m}$, and $\mathbf{C}_{e,m}$, entries are first generated independently from $N(0, 1)$ and are then set to zero independently with probability 0.95. The entries of $\mathbf{G}_{e,t}$ and $\mathbf{F}_{e,t}^{(m)}$ follow

independent AR(1) processes with coefficient 0.2 and standard normal innovations. The entries of $\epsilon_t^{(m)}$ are generated in the same way. Finally, the entries of $\Sigma_{\epsilon,m}$ are generated independently from $|N(0, 1)|$.

Unless otherwise stated, we take $M = 3$ groups and use 500 Monte Carlo replications in each setting. To evaluate the accuracy of our estimators for time-varying quantities, given any series of parameters $\theta = \{\theta_t\}_{t \in [T]}$, where θ_t can be a scalar, vector, or matrix, and its estimator $\hat{\theta} = \{\hat{\theta}_t\}_{t \in [T]}$, we define the (relative) mean squared error as

$$\text{MSE}(\hat{\theta}) := \frac{\sum_{t=1}^T \|\theta_t - \hat{\theta}_t\|_F^2}{\sum_{t=1}^T \|\theta_t\|_F^2}.$$

Moreover, for any given loading matrix $\mathbf{Q}$ and its estimator $\hat{\mathbf{Q}}$, we use the column-space distance

$$\mathcal{D}(\mathbf{Q}, \hat{\mathbf{Q}}) := \left\| \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}' - \hat{\mathbf{Q}}(\hat{\mathbf{Q}}'\hat{\mathbf{Q}})^{-1}\hat{\mathbf{Q}}' \right\|,$$

which is widely used in the literature such as [Chen et al. \(2022\)](#), among others. For group-specific quantities, MSE or $\mathcal{D}$ is first computed within each group and then averaged over m .

6.1.2 Accuracy of main effects

We first study estimation of the grand means and main effects under four designs (Ia)–(Id). Each design is considered at $T \in \{100, 400\}$. The common component part of the DGP follows the baseline design (IIa) in subsection 6.1.3.

(Ia) **Main effect baseline.** Set $(p_m, q_m) = (30, 30)$ for each $m \in [M]$. All loading strength parameters in (ME1) are set to 1. For all latent series, we set memory parameters $d_1 = d_2 = 0$.

(Ib) **Long-memory main effects.** Same as (Ia), except that $d_1 = d_2 = 0.1$ for $\mu_t, \mu_t^{(m)}$, and each entry of $\alpha_t, \beta_t, \alpha_t^{(m)}$, and $\beta_t^{(m)}$.

(Ic) **Weaker main effect factors.** Same as (Ia), except that $\gamma_r = \gamma_c = \gamma_r^{(m)} = \gamma_c^{(m)} = 0.8$ for each $m \in [M]$.

(Id) **Larger imbalanced dimensions.** Same as (Ia), except that $(p_m, q_m) = (30, 60)$ for each $m \in [M]$.

Table 1 reports the results. The direct moment estimators change little with T , since their pointwise accuracy is driven by cross-sectional averaging; the slightly smaller relative MSEs under long memory in (Ib) mainly reflect the larger signal magnitude in the denominator rather than improved estimation accuracy. In contrast, the main effect loading space errors decrease clearly with T , consistent with Theorems 3–5. Long memory mainly increases dispersion, while weaker loadings moderately reduce accuracy. Increasing only q_m in (Id) gives a clearer improvement on the column side. The initial and refitted global loading estimates have

nearly identical errors, indicating that refitting mainly links the loadings to their corresponding estimated global main effect factor series rather than changing their estimated spaces. Accordingly, the global main effect products improve sharply after refitting, while the local products also become more accurate as T increases. Long memory and weaker loadings mainly affect small- T accuracy, while the imbalanced design remains stable.

Table 1: Main effect estimation accuracy. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses.

Settings	Ia		Ib		Ic		Id	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$
$\text{MSE}(\hat{\alpha}_t^{(m)})$	0.05(0.01)	0.05(0.00)	0.04(0.01)	0.04(0.00)	0.05(0.01)	0.05(0.00)	0.02(0.00)	0.02(0.00)
$\text{MSE}(\hat{\mathbf{a}}_t^{(m)})$	0.59(0.07)	0.58(0.06)	0.48(0.07)	0.47(0.04)	1.17(0.13)	1.14(0.11)	0.30(0.04)	0.29(0.03)
$\text{MSE}(\hat{\mathbf{b}}_t^{(m)})$	0.60(0.07)	0.58(0.05)	0.49(0.07)	0.47(0.05)	1.17(0.14)	1.16(0.11)	0.59(0.06)	0.58(0.04)
$\text{MSE}(\hat{\mu}_t)$	0.03(0.01)	0.03(0.00)	0.02(0.01)	0.02(0.00)	0.03(0.01)	0.03(0.00)	0.01(0.00)	0.01(0.00)
$\text{MSE}(\hat{\mu}_t^{(m)})$	0.08(0.01)	0.08(0.01)	0.06(0.01)	0.06(0.01)	0.08(0.02)	0.08(0.01)	0.04(0.01)	0.04(0.00)
$\mathcal{D}(\mathbf{A}_{1m}, \hat{\mathbf{A}}_{1m})$	6.53(4.16)	2.77(1.11)	8.27(7.02)	3.19(1.63)	6.82(4.71)	2.89(1.07)	6.14(4.46)	2.57(1.02)
$\mathcal{D}(\mathbf{A}_{1m}, \hat{\mathbf{A}}_{1m})$	6.51(4.16)	2.76(1.11)	8.26(7.02)	3.19(1.63)	6.79(4.71)	2.88(1.07)	6.14(4.46)	2.57(1.02)
$\mathcal{D}(\mathbf{A}_{2m}, \hat{\mathbf{A}}_{2m})$	5.91(2.37)	2.72(0.86)	7.55(4.45)	3.12(1.21)	6.43(2.71)	2.89(0.91)	5.69(2.72)	2.57(0.88)
$\mathcal{D}(\mathbf{B}_{1m}, \hat{\mathbf{B}}_{1m})$	6.78(5.27)	2.76(1.07)	9.64(9.27)	3.26(1.46)	7.07(4.26)	2.82(0.95)	4.92(4.10)	2.01(0.71)
$\mathcal{D}(\mathbf{B}_{1m}, \hat{\mathbf{B}}_{1m})$	6.76(5.27)	2.76(1.07)	9.62(9.27)	3.26(1.46)	7.05(4.26)	2.80(0.95)	4.91(4.10)	2.00(0.71)
$\mathcal{D}(\mathbf{B}_{2m}, \hat{\mathbf{B}}_{2m})$	6.28(3.54)	2.68(0.86)	8.38(6.53)	3.16(1.15)	6.59(2.57)	2.82(0.78)	4.57(3.22)	1.98(0.59)
$\text{MSE}(\hat{\mathbf{A}}_{1m}\hat{\alpha}_t)$	101.74(10.48)	102.27(10.74)	102.17(9.27)	101.31(8.93)	104.18(14.56)	104.66(14.85)	102.22(9.85)	102.09(10.05)
$\text{MSE}(\hat{\mathbf{A}}_{1m}\hat{\alpha}_t)$	0.44(0.60)	0.10(0.06)	0.89(2.41)	0.12(0.13)	0.60(1.85)	0.13(0.07)	0.45(1.99)	0.07(0.05)
$\text{MSE}(\hat{\mathbf{A}}_{2m}\hat{\alpha}_t^{(m)})$	0.59(0.58)	0.18(0.07)	1.11(2.49)	0.20(0.13)	0.82(1.59)	0.28(0.08)	0.56(1.46)	0.13(0.06)
$\text{MSE}(\hat{\mathbf{B}}_{1m}\hat{\beta}_t)$	101.86(9.91)	102.09(10.26)	102.14(10.31)	101.64(9.40)	103.32(15.62)	104.52(15.73)	101.31(7.17)	101.09(7.34)
$\text{MSE}(\hat{\mathbf{B}}_{1m}\hat{\beta}_t)$	0.62(2.25)	0.10(0.05)	1.41(4.77)	0.12(0.10)	0.56(0.90)	0.13(0.06)	0.31(1.30)	0.05(0.03)
$\text{MSE}(\hat{\mathbf{B}}_{2m}\hat{\beta}_t^{(m)})$	0.76(1.97)	0.18(0.07)	1.60(5.09)	0.21(0.11)	0.78(0.80)	0.27(0.07)	0.43(2.26)	0.09(0.03)

6.1.3 Accuracy of common components

We next study estimation of the global and local common components under four designs (IIa)–(IIId). Each design is considered at $T \in \{100, 400\}$. The main effect part of the DGP follows the baseline design (Ia) in subsection 6.1.2.

(IIa) **Common component baseline.** Set $(p_m, q_m) = (30, 30)$ for each $m \in [M]$. All loading strength parameters in (L1) are set to 1. For all latent series, we set memory parameters $d_1 = d_2 = 0$.

(IIb) **Long-memory common factors.** Same as (IIa), except that $d_1 = d_2 = 0.1$ for each entry of $\mathbf{G}_t$ and $\mathbf{F}_t^{(m)}$.

(IIc) **Weaker common factors.** Same as (IIa), except that for each $m \in [M]$, $\delta_{q,i}^{(m)} = \delta_{j,i}^{(m)} = 0.8$ for $i = 1, 2, 3$ and $\delta_{r,i}^{(m)} = \delta_{c,i}^{(m)} = 0.8$ for $i = 1, 2$.

(IIId) **Larger imbalanced dimensions.** Same as (IIa), except that $(p_m, q_m) = (30, 60)$ for each $m \in [M]$.

Table 2 reports the results. The refitted global loading estimates $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$ are substantially more accurate than the initial estimates $\hat{\mathbf{Q}}_m$ and $\hat{\mathbf{J}}_m$ in every design, showing the benefit of cross-group alignment

and refitting. The loading-space errors also fall clearly as T increases; weaker loadings in (IIc) make estimation more difficult, while the larger imbalanced dimensions in (IId) remain stable and generally improve accuracy. The improvement is even more pronounced for the global common component product, whose error drops sharply after alignment and refitting; the local product also improves with T . Under long memory, several $T = 100$ entries have large Monte Carlo variation, but this largely disappears at $T = 400$.

Table 2: Common component estimation accuracy. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses.

Settings	IIa		IIb		IIc		IId	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$
$\mathcal{D}(\mathbf{Q}_m, \hat{\mathbf{Q}}_m)$	7.98(3.09)	3.90(1.45)	9.84(5.71)	4.52(1.77)	9.21(3.03)	4.46(1.41)	5.75(2.21)	2.73(1.05)
$\mathcal{D}(\mathbf{Q}_m, \tilde{\mathbf{Q}}_m)$	1.77(0.40)	0.79(0.15)	2.13(4.02)	0.79(0.18)	3.17(0.46)	1.44(0.18)	1.13(0.20)	0.50(0.07)
$\mathcal{D}(\mathbf{J}_m, \hat{\mathbf{J}}_m)$	8.22(3.15)	3.93(1.57)	9.69(5.65)	4.51(1.74)	9.27(3.16)	4.24(1.25)	8.24(3.12)	3.73(1.32)
$\mathcal{D}(\mathbf{J}_m, \tilde{\mathbf{J}}_m)$	1.56(0.20)	0.71(0.08)	1.75(3.54)	0.67(0.08)	3.01(0.37)	1.38(0.16)	1.46(0.15)	0.66(0.06)
$\mathcal{D}(\mathbf{R}_m, \hat{\mathbf{R}}_m)$	2.53(0.64)	1.01(0.18)	3.28(4.47)	1.05(0.22)	4.23(0.79)	2.42(0.55)	2.23(0.72)	0.76(0.15)
$\mathcal{D}(\mathbf{C}_m, \hat{\mathbf{C}}_m)$	2.50(0.66)	1.01(0.16)	3.33(4.54)	1.04(0.22)	4.23(0.90)	2.41(0.56)	2.09(0.48)	0.84(0.10)
$\text{MSE}(\hat{\mathbf{Q}}_m, \hat{\mathbf{Q}}_t, \hat{\mathbf{J}}'_m)$	155.20(39.46)	153.60(34.79)	155.72(38.31)	155.31(36.50)	156.11(38.47)	156.13(34.55)	152.32(37.03)	153.54(36.18)
$\text{MSE}(\hat{\mathbf{Q}}_m, \tilde{\mathbf{Q}}_t, \tilde{\mathbf{J}}'_m)$	0.07(0.02)	0.04(0.01)	0.49(9.09)	0.03(0.01)	0.25(0.05)	0.16(0.02)	0.04(0.01)	0.02(0.00)
$\text{MSE}(\hat{\mathbf{R}}_m, \hat{\mathbf{F}}_t^{(m)}, \hat{\mathbf{C}}'_m)$	0.25(0.11)	0.13(0.03)	0.82(10.24)	0.12(0.03)	0.71(0.20)	0.48(0.10)	0.14(0.06)	0.06(0.01)

To assess our common component estimation procedure, we compare it with the basic and augmented MFM estimators of Zhang et al. (2025). For comparability, all three estimators use the same $\check{\mathbf{Y}}_t^{(m)}$ for common component estimation. In addition to setting (IIa), we consider setting (IIc') which is the same as (IIa), except that the global loading strengths are $(1, 0.8, 0.8)$ in both modes, while the local loading strengths remain $(1, 1)$. The row and column designs are symmetric, so we report only the row-side results. Table 3 shows that the augmented MFM improves substantially on the basic estimator, while our estimator yields smaller errors throughout. The advantage is more pronounced when the global loadings are weaker.

Table 3: Global row loading-space estimation accuracy for MMEFM and MFM estimators. MFM Augmented uses 15 iterations. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses.

Settings	IIa		IIc'	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$
MMEFM	1.78(0.42)	0.80(0.14)	5.15(13.91)	1.30(0.22)
MFM Basic	14.96(8.95)	3.84(1.34)	69.33(21.23)	17.08(11.93)
MFM Augmented	2.71(2.08)	1.22(0.16)	10.07(13.33)	2.05(0.30)

6.1.4 Behavior of the statistic $S_{G,m}$

We next examine the finite-sample behavior of the statistic $S_{G,m}$ in Theorem 9. Let $\mathcal{A}_G$ denote the active set of groups containing a shared global common component, and let $\mathcal{A}_G^c = [M] \setminus \mathcal{A}_G$. Since a global common component must be shared by at least two groups, we consider $|\mathcal{A}_G| \in \{0, 2, 3, 4, 5\}$ with $M = 5$.

The active groups are randomly selected in each Monte Carlo replication. We use 400 replications and set $T = 200$ unless otherwise stated.

We consider six settings. The first four settings are inherited from (IIa)–(IId) and are relabelled as (IIIa)–(IIId), respectively. We additionally consider

(IIIe) **Weaker local common component.** Same as (IIIa), except that $\delta_{r,i}^{(m)} = \delta_{c,i}^{(m)} = 0.8$ for $i = 1, 2$ and each $m \in [M]$.

(IIIf) **Weaker global common component.** Same as (IIIa), except that $\delta_{q,i}^{(m)} = \delta_{j,i}^{(m)} = 0.8$ for $i = 1, 2, 3$ and each $m \in [M]$.

Before computing $S_{G,m}$, we standardize $\check{\mathbf{Y}}_t^{(m)}$ using (5.1). Figure 1 reports the within-replication averages of $S_{G,m}$ over the active and inactive groups at $T = 200$. The active and inactive averages are clearly separated in settings (IIIa)–(IIIe). The weaker-global setting (IIIf) is more difficult, especially when only a few groups are active. The separation also tends to increase with $|\mathcal{A}_G|$, as each active group then shares the global component with more groups.

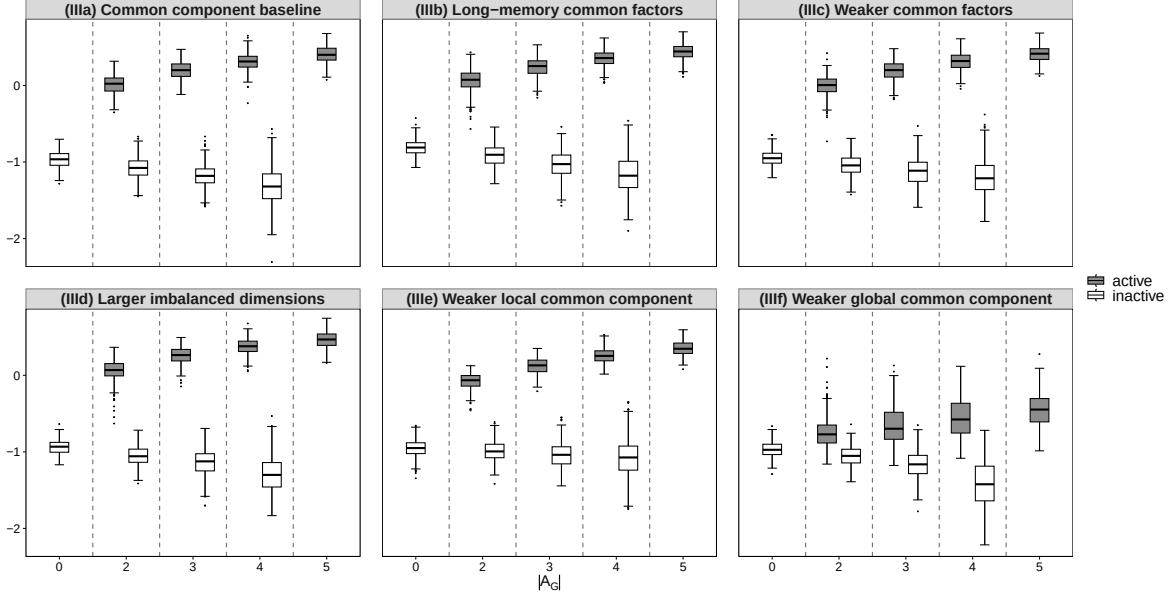


Figure 1: Averages of $S_{G,m}$ over the active and inactive groups at $T = 200$. The y-axis is $\log_{10}$ -transformed.

Figure 2 plots the Monte Carlo medians of these active and inactive averages as T increases. Across all settings, the active averages remain relatively stable, while the inactive averages decrease. The gap therefore widens with T , consistent with Theorem 9. In the weaker-global setting (IIIf), the active averages are lower in finite samples but remain much more stable than the inactive averages.

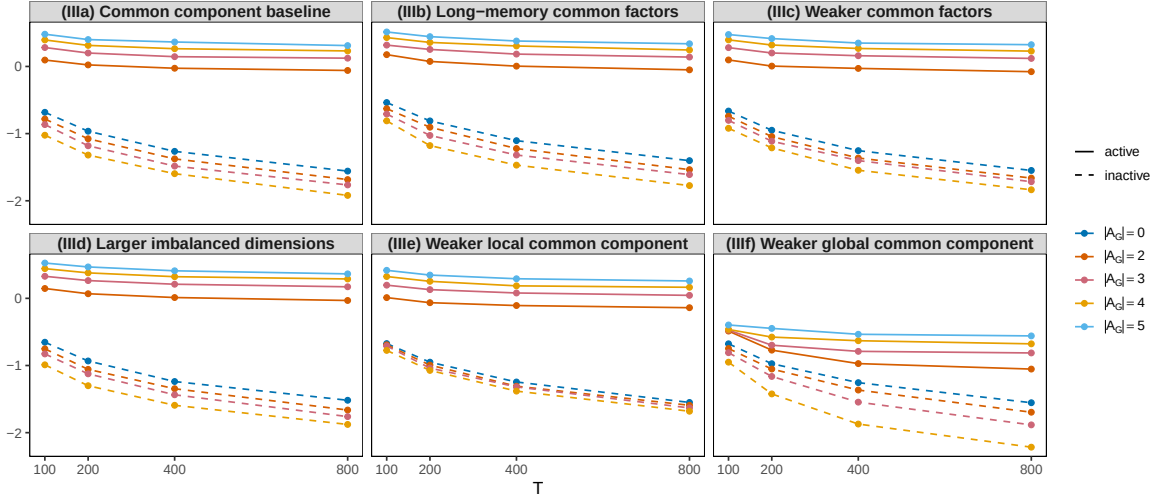


Figure 2: Medians of the averages of $S_{G,m}$ over the active and inactive groups, respectively. The y-axis is $\log_{10}$ -transformed. Solid and dashed lines correspond to active and inactive groups, respectively.

6.1.5 Accuracy of detection and screening

We next examine whether the separation of $S_{G,m}$ translates into reliable detection and screening using the procedure in Section 5. For these experiments, we use 400 Monte Carlo replications and true factor ranks, set the nominal level to 5%, and use $B = 199$ bootstrap replications.

Tables 4 and 5 report the performance of the bootstrap existence test. In the baseline setting, the empirical size remains close to the nominal 5% level across sample sizes, while power increases rapidly with both T and $|\mathcal{A}_G|$. At $T = 200$, the test continues to control size well across settings. Power remains high in settings (IIIa)–(IIIe), with some loss under long memory when only a few groups are active. The weaker-global setting (IIIf) is substantially more difficult, consistent with the lower active statistics observed in Section 6.1.4. Power nevertheless improves as more groups share the global component.

Table 4: Bootstrap existence test in the baseline setting (IIIa). Entries are empirical rejection rates with Monte Carlo standard errors in parentheses; $|\mathcal{A}_G| = 0$ reports empirical size and $|\mathcal{A}_G| \geq 2$ reports power.

T	$ \mathcal{A}_G = 0$	$ \mathcal{A}_G = 2$	$ \mathcal{A}_G = 3$	$ \mathcal{A}_G = 4$	$ \mathcal{A}_G = 5$
100	0.043 (0.010)	0.828 (0.019)	0.993 (0.004)	1.000 (0.000)	1.000 (0.000)
200	0.058 (0.012)	0.995 (0.004)	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)
400	0.050 (0.011)	0.998 (0.002)	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)

Table 5: Bootstrap existence test across settings at $T = 200$. Entries are described in Table 4.

$ \mathcal{A}_G $	(IIIa)	(IIIb)	(IIIc)	(IIId)	(IIIe)	(IIIf)
0	0.058 (0.012)	0.068 (0.013)	0.050 (0.011)	0.055 (0.011)	0.058 (0.012)	0.053 (0.011)
2	0.995 (0.004)	0.930 (0.013)	0.993 (0.004)	0.985 (0.006)	1.000 (0.000)	0.085 (0.014)
4	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)	0.648 (0.024)

Tables 6 and 7 report the screening performance, where $\text{TPR} = |\hat{\mathcal{A}}_G \cap \mathcal{A}_G|/|\mathcal{A}_G|$, $\text{FPR} = |\hat{\mathcal{A}}_G \cap$

$\mathcal{A}_G^c/|\mathcal{A}_G^c|$, and Exact denotes $\mathbb{1}\{\hat{\mathcal{A}}_G = \mathcal{A}_G\}$, averaged over replications. In the baseline setting, screening becomes nearly exact as T increases when global common factors are present, while false positives remain rare. The screening procedure therefore becomes increasingly accurate as the separation in $S_{G,m}$ strengthens. At $T = 200$, screening remains accurate across settings (IIIa)–(IIIe), with high true positive rates and low false positive rates. The weaker-global setting (IIIf) is again the most challenging, where the reduced shared signal leads to substantially lower true positive and exact-selection rates. This pattern closely follows the existence-test results, since the final selected set is gated by the existence decision.

Table 6: Bootstrap screening in the baseline setting (IIIa). Entries are empirical averages computed from the final gated selected set $\hat{\mathcal{A}}_G$.

Metric	$ \mathcal{A}_G = 0$			$ \mathcal{A}_G = 2$			$ \mathcal{A}_G = 3$			$ \mathcal{A}_G = 4$			$ \mathcal{A}_G = 5$		
	100	200	400	100	200	400	100	200	400	100	200	400	100	200	400
TPR	–	–	–	0.826	0.995	0.998	0.951	0.998	1.000	0.971	0.996	0.999	0.985	0.995	1.000
FPR	0.019	0.026	0.021	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–	–
Exact	0.958	0.943	0.950	0.825	0.990	0.998	0.868	0.993	1.000	0.888	0.983	0.998	0.928	0.975	1.000

Table 7: Bootstrap screening across settings at $T = 200$. Entries are described in Table 6.

$ \mathcal{A}_G $	Metric	(IIIa)	(IIIb)	(IIIc)	(IIId)	(IIIe)	(IIIf)
0	FPR	0.026	0.029	0.023	0.024	0.025	0.022
	Exact	0.943	0.933	0.950	0.945	0.943	0.948
2	TPR	0.995	0.930	0.993	0.985	1.000	0.073
	FPR	0.002	0.000	0.001	0.002	0.029	0.009
	Exact	0.990	0.930	0.990	0.980	0.915	0.060
4	TPR	0.996	0.986	0.995	0.999	0.999	0.425
	FPR	0.000	0.000	0.000	0.000	0.005	0.003
	Exact	0.983	0.950	0.980	0.998	0.990	0.085

Finally, as a robustness check, we repeat the estimation experiments and the baseline statistic-separation experiment under Student $t(6)$ innovations, with the corresponding results reported in Section E.1.1 of the Supplementary Materials. The qualitative patterns remain largely unchanged. Heavy tails mainly increase finite-sample variation, especially for smaller T , while having little effect on the overall estimation accuracy or the separation between active and inactive groups.

6.2 Real data analysis

We study daily U.S. equity data from CRSP, grouping stocks into the 11 GICS sectors. The set of stocks is fixed on December 30, 2022 by selecting the 30 largest U.S. companies by market capitalization within each sector. Six stocks with structural exits during 2023–2024 are excluded from the primary sample, leaving 324 stocks in total. We consider nine indicators: daily return, 20-day cumulative log return, log dollar volume, log turnover, 20-day relative turnover, 20-day log realized volatility, high-low range, log bid-ask spread,

and 20-day log Amihud illiquidity (Amihud, 2002). The sample runs from January 3, 2023 to December 31, 2024, with $T = 502$ trading days, giving 11 grouped matrix time series $\{\mathbf{Y}_t^{(m)}\}_{t=1}^T$. Observations in 2025 are reserved for out-of-sample forecasting.

For each indicator, the mean and standard deviation are computed from all observed entries in the 2023–2024 primary sample before imputation and used for standardization. Ordinary missing values are imputed using same-day sector medians, and each series is then demeaned over 2023–2024. We continue to write $\{\mathbf{Y}_t^{(m)}\}$ for the resulting cell-demeaned standardized matrices. Indicator construction and preprocessing are described in subsection E.2.1 of the Supplementary Materials.

Why a MMEFM structure? We first motivate the two features of the MMEFM that matter here: explicit main effects and a shared global structure.

First, under (IC1), we have the decomposition $\|\mathbf{Y}_t^{(m)}\|_F^2 = p_m q_m (\hat{c}_t^{(m)})^2 + q_m \|\hat{\mathbf{a}}_t^{(m)}\|^2 + p_m \|\hat{\mathbf{b}}_t^{(m)}\|^2 + \|\tilde{\mathbf{Y}}_t^{(m)}\|_F^2$. The interaction component accounts only for $\sum_m \sum_t \|\tilde{\mathbf{Y}}_t^{(m)}\|_F^2 / \sum_m \sum_t \|\mathbf{Y}_t^{(m)}\|_F^2 \approx 56.77\%$ of the total squared magnitude, leaving 43.23% attributable to the additive components. This sizable contribution motivates modeling the main effects explicitly.

Second, we apply the global-common-factor detection and screening procedure developed in Section 5. The two largest statistics are $S_{G,(1)} = 8.254$ and $S_{G,(2)} = 7.571$, both exceeding the 95% circular-shift bootstrap cutoff of 4.589, providing evidence for a nontrivial global common component. Sector-level screening retains Industrials, Communication Services, and Real Estate. We nevertheless estimate the MMEFM using all 11 sectors, since the remaining $S_{G,m}$ ’s are also away from zero and may reflect weaker shared market dynamics. For comparison, we also consider a screened specification in which the global common component is estimated using only the three retained sectors, while local components are retained for all sectors. Further evidence on the contributions of the main effects and global component, together with the complete set of sector-level $S_{G,m}$ statistics, is provided in the Supplementary Materials.

Model estimation and interpretation. We next estimate the MMEFM on our stock data. The factor dimensions are selected using the procedure in subsection 3.4. The global main-effect and common-component ranks are both (1,1), while the local ranks vary across sectors, with the complete selection provided in Supplementary Table E.6.

We first examine the global structure. Figure 3 shows that the global indicator-main loadings $\tilde{\mathbf{B}}_{1m}$ are broadly similar across sectors, with turnover and relative turnover loading in the opposite direction to daily and 20-day returns, suggesting a market-wide contrast between trading activity and recent price performance. Figure 4 further shows that the global common component has more heterogeneous indicator-loading profiles across the three selected sectors: 20-day return is most prominent in Industrials, whereas

liquidity and trading measures are more prominent in Communication Services and Real Estate. The shared global factor also shows a large movement around the late-July–early-August 2024 market dislocation, pointing to a market-wide change rather than a sector-specific episode.

We next illustrate the local structure using Financials and Information Technology. Figure 5 shows that the Financials stock-main loadings broadly separate large banks and broker-dealers from insurers and several market-infrastructure firms. Figure 6 shows that the Information Technology common stock loadings distinguish semiconductor and chip-equipment firms from software, payment, and other large technology firms, while returns load in the opposite direction to trading activity and the high-low range.

Additional estimates are reported in the Supplementary Materials.

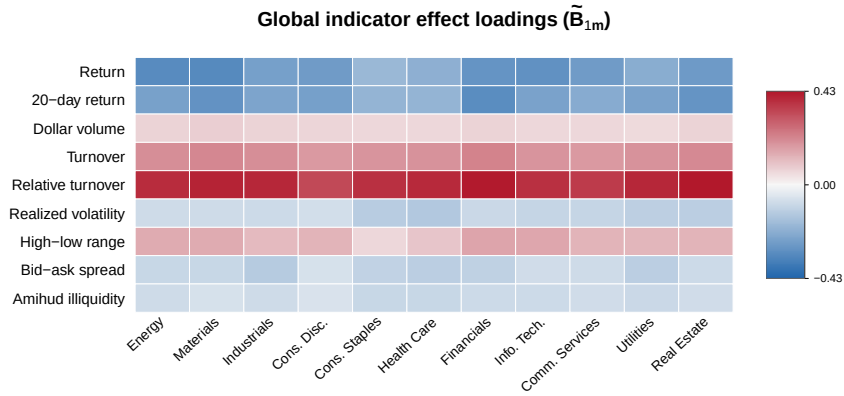


Figure 3: Global indicator main-effect estimates.

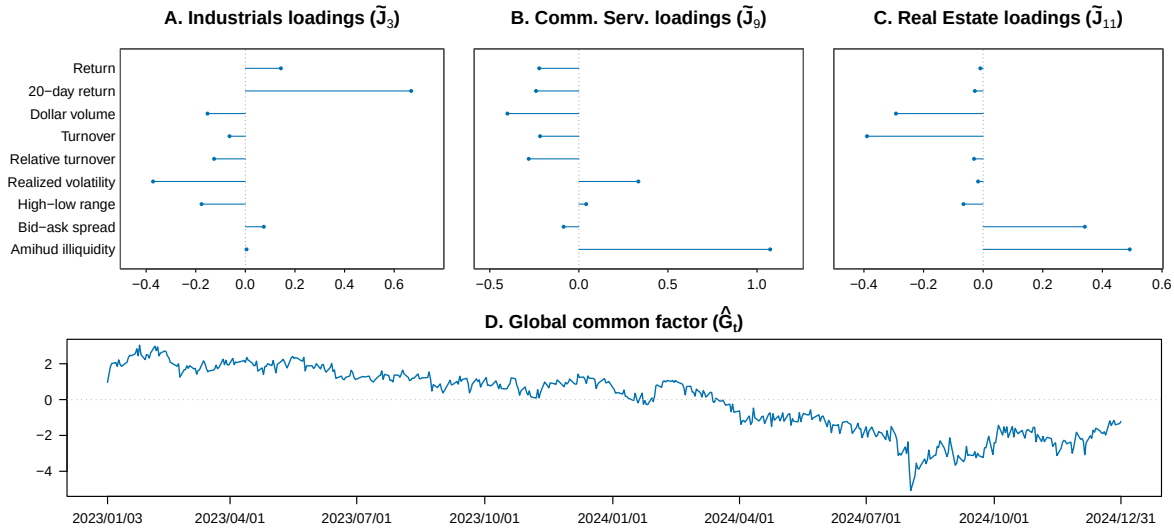


Figure 4: Global common component estimates.

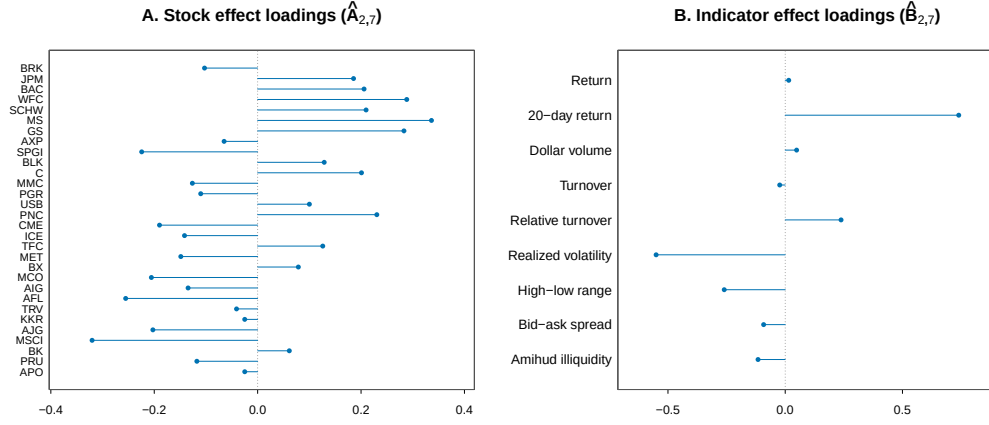


Figure 5: Local main-effect estimates for Financials.

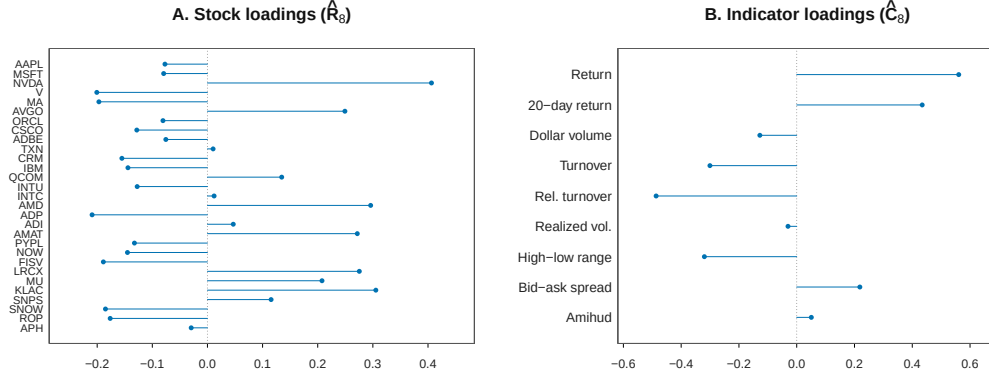


Figure 6: Local common-component estimates for Information Technology.

Model comparison. We next compare the in-sample fit of several MMEFM specifications and MFM. Table 8 reports RSS/TSS together with a simple parameter-count adjustment, $(RSS/TSS) \cdot N/(N - \nu)$, where $N := \sum_{m=1}^M T p_m q_m$ and $\nu := T \cdot \#\{\text{latent series per date}\} + \#\{\text{loading coefficients}\}$.

The local-only and global-only specifications retain only local and global dynamics, respectively. MFM Basic uses the baseline estimator of Zhang et al. (2025), whereas MFM Augmented uses their Section 3.5 iterative estimator with 15 iterations. The MMEFM specifications fit better than the two single-level structured benchmarks, while the screened-global specification achieves essentially the same fit as the all-sector model with fewer loading coefficients. MFM Augmented improves markedly over MFM Basic, but both fit less well than the MMEFM specifications.

We complement the comparison with daily rolling forecasts for 2025. At each origin, ranks and loading matrices are re-estimated on a 252-day window advanced by one trading day; the recovered time-varying quantities are forecast entrywise by $AR(p)$ models with $p \in [5]$ chosen by BIC from the latest 250 observations. We consider $h = 1, 3, 5, 10$, keep the three screened sectors fixed throughout 2025, and define MSPE

Table 8: In-sample reconstruction across MMEFM specifications and MFM benchmarks. The MFM rank estimator selects global rank (1, 1); the local ranks are (3, 3) for Energy, (1, 2) for Information Technology, (2, 3) for Consumer Discretionary, Consumer Staples, and Communication Services, and (1, 3) otherwise.

Model	Per-date dim.	Loading coeffs.	RSS/TSS	Adjusted
MMEFM, all sectors	68	2,074	0.712	0.730
MMEFM, screened global	68	1,737	0.708	0.726
MMEFM, local only	64	1,228	0.742	0.760
MMEFM, global only	4	846	0.869	0.871
MFM Basic	58	1,180	1.001	1.022
MFM Augmented	58	1,180	0.771	0.787

as the total squared Frobenius forecast error divided by the number of evaluated entries.

Table 9 shows that both MMEFM specifications outperform MFM Basic at every horizon. Relative to MFM Augmented, the all-sector MMEFM reduces MSPE by around 10%, and the screened-global specification performs similarly.

Table 9: Daily rolling forecasting MSPEs for 2025.

Model	$h = 1$	$h = 3$	$h = 5$	$h = 10$
MMEFM, all sectors	0.802	0.870	0.894	0.940
MMEFM, screened global	0.810	0.875	0.897	0.940
MFM, Basic	0.983	1.022	1.028	1.038
MFM, Augmented	0.908	0.953	0.973	1.003

References

- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56.
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.
- Cai, X., Kong, X., Wu, X., and Zhao, P. (2025). Matrix-factor-augmented regression. *Journal of Business & Economic Statistics*, 43(4):1145–1157.
- Cen, Z. (2026). Identification and estimation of multi-order tensor factor models. *Journal of Time Series Analysis*. forthcoming.
- Chen, E. Y. and Fan, J. (2023). Statistical inference for high-dimensional matrix-variate factor models. *Journal of the American Statistical Association*, 118(542):1038–1055.
- Chen, E. Y., Tsay, R. S., and Chen, R. (2020). Constrained factor models for high-dimensional matrix-variate time series. *Journal of the American Statistical Association*, 115(530):775–793.

- Chen, R., Yang, D., and Zhang, C.-H. (2022). Factor models for high-dimensional tensor time series. *Journal of the American Statistical Association*, 117(537):94–116.
- Chen, W. and Lam, C. (2024). Rank and factor loadings estimation in time series tensor factor model by pre-averaging. *Annals of Statistics*, 52(1):364 – 391.
- Choi, I., Kim, D., Kim, Y. J., and Kwark, N.-S. (2018). A multilevel factor model: Identification, asymptotic theory and applications. *Journal of Applied Econometrics*, 33(3):355–377.
- Choi, I., Lin, R., and Shin, Y. (2023). Canonical correlation-based model selection for the multilevel factors. *Journal of Econometrics*, 233(1):22–44.
- Han, X. (2021). Shrinkage estimation of factor models with global and group-specific factors. *Journal of Business & Economic Statistics*, 39(1):1–17.
- He, Y., Kong, X., Yu, L., Zhang, X., and Zhao, C. (2023a). Matrix factor analysis: From least squares to iterative projection. *Journal of Business and Economic Statistics*, 0(0):1 – 13.
- He, Y., Kong, X.-B., Liu, D., and Zhao, R. (2023b). Robust statistical inference for large-dimensional matrix-valued time series via iterative huber regression. *arXiv:2306.03317*.
- He, Y., Ma, X., Wang, X., and Wang, Y. (2025). Factor modelling for biclustering large-dimensional matrix-valued time series. *arXiv preprint arXiv:2502.06397*.
- Lam, C. and Cen, Z. (2025). Matrix-valued factor model with time-varying main effects. *Journal of Econometrics*, 252:106105.
- Lam, C. and Yao, Q. (2012). Factor modeling for high-dimensional time series: Inference for the number of factors. *The Annals of Statistics*, 40(2):694–726.
- Lam, C., Yao, Q., and Bathia, N. (2011). Estimation of latent factors for high-dimensional time series. *Biometrika*, 98(4):901–918.
- Loan, C. F. (2000). The ubiquitous kronecker product. *Journal of Computational and Applied Mathematics*, 123(1):85–100. Numerical Analysis 2000. Vol. III: Linear Algebra.
- Lu, X., Jin, S., and Su, L. (2025). Three-dimensional factor models with global and local factors. *Econometric Theory*, pages 1–47.
- Moench, E., Ng, S., and Potter, S. (2013). Dynamic hierarchical factor models. *Review of Economics and Statistics*, 95(5):1811–1817.

- Wang, D., Liu, X., and Chen, R. (2019). Factor models for matrix-valued high-dimensional time series. *Journal of Econometrics*, 208(1):231–248. Special Issue on Financial Engineering and Risk Management.
- Yu, L., He, Y., Kong, X., and Zhang, X. (2022). Projected estimation for large-dimensional matrix factor models. *Journal of Econometrics*, 229(1):201–217.
- Yu, Y., Wang, T., and Samworth, R. J. (2015). A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323.
- Zhang, Y., Hui, Y., Song, J., and Zheng, S. (2025). Multilevel matrix factor model. *Journal of Econometrics*, 251:106033.

Supplementary Materials for the paper “Multilevel Main Effects Matrix Factor Model”

A Iterative Projection for Estimating $\mathbf{Q}_m$ and $\mathbf{J}_m$

For finding the common components in Step 3 in Section 3.3, we assume that the directions that align more to the strongest factors of $\mathbf{Q}_m$ and $\mathbf{J}_m$ are available for all $m \in [M]$, namely $\mathbf{v}_a^{(m)}$ and $\mathbf{v}_b^{(m)}$ for $m \in [M]$ in (3.8) and (3.9), which are of course not the case in reality. In the following, we present an iterative algorithm for finding $\mathbf{v}_a^{(m)}$ and $\mathbf{v}_b^{(m)}$, and more accurate estimators of $\mathbf{Q}_m$ and $\mathbf{J}_m$ in Step 3 in Section 3.3.

Algorithm A.1 Algorithm for choosing good projection directions for estimating $\mathbf{Q}_m$ and $\mathbf{J}_m$

1. Preset K_0 (recommend $K_0 = 20$). Let $\{S_{m,k}\}_{m \in [M], k \in [K_0]}$ be a sequence (in k) of subsets of $[q_m]$ for each $m \in [M]$, which chooses elements in $[q_m]$ uniformly, with $|S_{m,k}| = \lfloor q_m/2 \rfloor$. Denote by $\mathbf{1}_{d,S}$ a vector of ones and zeros of length d , where the positions in $S \subseteq [d]$ are 1 and 0 otherwise.
2. Let $m = 1$, and let $k_n \in [K_0]$ for each $n \neq m$. Define

$$\mathbf{X}_{Q,m,k|k_n} := T^{-2}(M-1)^{-1} \sum_{n \neq m; n=1}^M \left\{ \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{1}_{q_m, S_{m,k}} \mathbf{1}_{q_n, S_{n,k_n}}' \check{\mathbf{Y}}_t^{(n)'} \right\}^{\otimes 2}.$$

Find $\tilde{k}_m$ such that $\mathbf{X}_{Q,m,\tilde{k}_m|k_n}$ has the largest eigenvalue. Take $\mathbf{v}_b^{(m)} = \mathbf{1}_{q_m, S_{m,\tilde{k}_m}}$.

3. For $m \in [M]$ larger than 1, repeat Step 2, replacing k_n by $\tilde{k}_n$ found in previous steps in the definition of $\mathbf{X}_{Q,m,k|k_n}$.
4. With $\mathbf{v}_b^{(m)} = \mathbf{1}_{q_m, S_{m,\tilde{k}_m}}$ for each $m \in [M]$, take $\hat{\mathbf{Q}}_m$ as the first k_r eigenvectors of $\mathbf{X}_{Q,m}$ defined in (3.8).
5. Set $\mathbf{v}_a^{(m)} = \hat{\mathbf{Q}}_{m,1}$ for each $m \in [M]$.

Fix $m = 1$. Replace $\mathbf{v}_a^{(m)} = \hat{\mathbf{Q}}_{m,k}$ and find the eigenvalue of $\mathbf{X}_{J,m}$ defined in (3.9), $k \in [k_r]$. Set $\mathbf{v}_a^{(m)} = \hat{\mathbf{Q}}_{m,k}$ so that $\mathbf{X}_{J,m}$ has the largest eigenvalue. Repeat for $m = 2, \dots, M$, resulting in a new set of $\mathbf{v}_a^{(m)}$ for $m \in [M]$.

Estimate $\hat{\mathbf{J}}_m$ as the first k_c eigenvectors of $\mathbf{X}_{J,m}$.

6. Set $\mathbf{v}_b^{(m)} = \hat{\mathbf{J}}_{m,1}$ for each $m \in [M]$.

Fix $m = 1$. Replace $\mathbf{v}_b^{(m)} = \hat{\mathbf{J}}_{m,k}$ and find the eigenvalue of $\mathbf{X}_{Q,m}$ defined in (3.8), $k \in [k_c]$. Set $\mathbf{v}_b^{(m)} = \hat{\mathbf{J}}_{m,k}$ so that $\mathbf{X}_{Q,m}$ has the largest eigenvalue. Repeat for $m = 2, \dots, M$, resulting in a new set of $\mathbf{v}_b^{(m)}$ for $m \in [M]$.

Estimate $\hat{\mathbf{Q}}_m$ as the first k_r eigenvectors of $\mathbf{X}_{Q,m}$.

7. Iterate Step 5 and 6 until some kind of convergence is reached.

One major difference between Algorithm 1 and the iterative algorithm in Zhang et al. (2025) is that we only use the directions that align more to the strongest factors for iterative projection in Steps 5 and 6 above. This is to avoid projection using weak factors when it is likely that in real data such weak factors exist, which can numerically lead to less accurate estimators if used as projection directions. Our final two-step aligned estimators $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$ in step 3 in Section 3.3 are constructed based on $\hat{\mathbf{Q}}_m$ and $\hat{\mathbf{J}}_m$ and indeed perform better than the augmented estimators from Zhang et al. (2025). See Table 3 in Section 6 for more details.

B Alternative Procedures for Global Common Component Estimation

The estimation procedure in Step 3 of Section 3.3 uses the nearest Kronecker product approximation, described in Section 6 of Loan (2000), implemented through a matrix rearrangement and singular value decomposition, for both the alignment procedure and the re-estimation procedure. For completeness, we describe two alternative iterative procedures. Subsection B.1 presents a two-sided Procrustes alternative for the alignment step, while Subsection B.2 presents an alternating least-squares alternative. These two alternatives concern separate stages of the estimation procedure and may be applied independently.

B.1 Two-sided Procrustes alignment

To align the coordinate systems of the $\hat{\mathbf{Q}}_m$ and $\hat{\mathbf{J}}_m$ across $m \in [M]$, we first choose a reference group $m_0 \in [M]$. For each $m \neq m_0$, we find orthogonal alignment matrices $\hat{\mathbf{O}}_{Q,m}$ and $\hat{\mathbf{O}}_{J,m}$ by the two-sided Procrustes criterion, namely

$$(\hat{\mathbf{O}}_{Q,m}, \hat{\mathbf{O}}_{J,m}) = \arg \min_{\mathbf{O}_Q \in \mathbb{O}(k_r), \mathbf{O}_J \in \mathbb{O}(k_c)} \sum_{t=1}^T \left\| \hat{\mathbf{Q}}_{m_0}' \tilde{\mathbf{Y}}_{G,t}^{(m_0)} \hat{\mathbf{J}}_{m_0} - \mathbf{O}_Q' \hat{\mathbf{Q}}_m' \tilde{\mathbf{Y}}_{G,t}^{(m)} \hat{\mathbf{J}}_m \mathbf{O}_J \right\|_F^2, \quad (\text{B.1})$$

where $\mathbb{O}(a)$ denotes the set of $a \times a$ orthogonal matrices. To this end, given $\mathbf{O}_{J,m}$, define

$$\Psi_{Q,m} := \sum_{t=1}^T \hat{\mathbf{Q}}_m' \tilde{\mathbf{Y}}_{G,t}^{(m)} \hat{\mathbf{J}}_m \mathbf{O}_{J,m} \hat{\mathbf{J}}_{m_0}' \tilde{\mathbf{Y}}_{G,t}^{(m_0)'} \hat{\mathbf{Q}}_{m_0}.$$

The corresponding Procrustes update is

$$\mathbf{O}_{Q,m} = \mathbf{U}_{Q,m} \mathbf{V}_{Q,m}', \quad (\text{B.2})$$

where $\mathbf{U}_{Q,m}$ and $\mathbf{V}_{Q,m}$ are the left- and right-singular vectors of $\Psi_{Q,m}$. Similarly, given $\mathbf{O}_{Q,m}$, define

$$\Psi_{J,m} := \sum_{t=1}^T \hat{\mathbf{J}}_m' \check{\mathbf{Y}}_{G,t}^{(m)'} \hat{\mathbf{Q}}_m \mathbf{O}_{Q,m} \hat{\mathbf{Q}}_{m_0}' \check{\mathbf{Y}}_{G,t}^{(m_0)} \hat{\mathbf{J}}_{m_0},$$

and its corresponding Procrustes update is

$$\mathbf{O}_{J,m} = \mathbf{U}_{J,m} \mathbf{V}_{J,m}', \quad (\text{B.3})$$

where $\mathbf{U}_{J,m}$ and $\mathbf{V}_{J,m}$ are the left- and right-singular vectors of $\Psi_{J,m}$. Starting from $\mathbf{O}_{Q,m} = \mathbf{I}_{k_r}$ and $\mathbf{O}_{J,m} = \mathbf{I}_{k_c}$, we iterate (B.2) and (B.3) to obtain $\hat{\mathbf{O}}_{Q,m}$ and $\hat{\mathbf{O}}_{J,m}$. We then replace $\hat{\mathbf{Q}}_m$ and $\hat{\mathbf{J}}_m$ with $\hat{\mathbf{Q}}_m \hat{\mathbf{O}}_{Q,m}$ and $\hat{\mathbf{J}}_m \hat{\mathbf{O}}_{J,m}$ respectively for $m \neq m_0$. After the alignment, the global common factor $\mathbf{G}_t$ is estimated by least square as

$$\hat{\mathbf{G}}_t := M^{-1} \sum_{m=1}^M \hat{\mathbf{Q}}_m' (\check{\mathbf{Y}}_t^{(m)} - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m') \hat{\mathbf{J}}_m. \quad (\text{B.4})$$

B.2 Alternating least-squares refitting

For $m \in [M]$, we re-estimate $\mathbf{Q}_m$ and $\mathbf{J}_m$ by minimizing $\sum_{t=1}^T \|\check{\mathbf{Y}}_t^{(m)} - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m' - \mathbf{Q}_m \hat{\mathbf{G}}_t \mathbf{J}_m'\|_F^2$ over $\mathbf{Q}_m$ and $\mathbf{J}_m$. To this end, given $\mathbf{J}_m$, we have

$$\tilde{\mathbf{Q}}_m = \left(\sum_{t=1}^T (\check{\mathbf{Y}}_t^{(m)} - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m') \mathbf{J}_m \hat{\mathbf{G}}_t' \right) \left(\sum_{t=1}^T \hat{\mathbf{G}}_t \mathbf{J}_m' \mathbf{J}_m \hat{\mathbf{G}}_t' \right)^{-1}, \quad (\text{B.5})$$

and given $\mathbf{Q}_m$, we have

$$\tilde{\mathbf{J}}_m = \left(\sum_{t=1}^T (\check{\mathbf{Y}}_t^{(m)} - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m')' \mathbf{Q}_m \hat{\mathbf{G}}_t \right) \left(\sum_{t=1}^T \hat{\mathbf{G}}_t' \mathbf{Q}_m' \mathbf{Q}_m \hat{\mathbf{G}}_t \right)^{-1}. \quad (\text{B.6})$$

Using estimators $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$, we can iterate (B.5) and (B.6) to obtain the final estimators $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$ for $\mathbf{Q}_m$ and $\mathbf{J}_m$, respectively. This is in parallel to the algorithm in Section 2.2 of [Cai et al. \(2025\)](#).

Table B.1: Common component estimation accuracy using Procrustes alignment and ALS refitting. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses.

Settings	IIa		IIb		IIc		IIId	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$
$\mathcal{D}(\mathbf{Q}_m, \hat{\mathbf{Q}}_m)$	7.98(3.09)	3.90(1.45)	9.84(5.71)	4.52(1.77)	9.21(3.03)	4.46(1.41)	5.75(2.21)	2.73(1.05)
$\mathcal{D}(\mathbf{Q}_m, \tilde{\mathbf{Q}}_m)$	1.67(0.42)	0.75(0.15)	2.09(4.45)	0.76(0.20)	2.85(0.46)	1.35(0.17)	1.06(0.22)	0.48(0.07)
$\mathcal{D}(\mathbf{J}_m, \hat{\mathbf{J}}_m)$	8.22(3.15)	3.93(1.57)	9.69(5.65)	4.51(1.74)	9.27(3.16)	4.24(1.25)	8.24(3.12)	3.73(1.32)
$\mathcal{D}(\mathbf{J}_m, \tilde{\mathbf{J}}_m)$	1.42(0.20)	0.66(0.07)	1.66(4.15)	0.62(0.07)	2.69(0.34)	1.28(0.13)	1.31(0.15)	0.61(0.05)
$\mathcal{D}(\mathbf{R}_m, \hat{\mathbf{R}}_m)$	2.53(0.64)	1.01(0.18)	3.28(4.47)	1.05(0.22)	4.23(0.79)	2.42(0.55)	2.23(0.72)	0.76(0.15)
$\mathcal{D}(\mathbf{C}_m, \hat{\mathbf{C}}_m)$	2.50(0.66)	1.01(0.16)	3.33(4.54)	1.04(0.22)	4.23(0.90)	2.41(0.56)	2.09(0.48)	0.84(0.10)
$\text{MSE}(\tilde{\mathbf{Q}}_m \hat{\mathbf{G}}_t \tilde{\mathbf{J}}_m')$	0.52(2.44)	0.06(0.64)	1.23(9.15)	0.03(0.00)	0.58(2.27)	0.15(0.57)	0.30(1.87)	0.02(0.00)
$\text{MSE}(\hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m')$	0.25(0.11)	0.13(0.03)	0.82(10.24)	0.12(0.03)	0.71(0.20)	0.48(0.10)	0.14(0.06)	0.06(0.01)

C Proof of Theorems

Proof of Theorem 1. The grand mean $c_t^{(m)}$ (as defined in (3.2)) is identified when we pre- and post-multiply by $\mathbf{1}_{p_m}'$ and $\mathbf{1}_{q_m}$ respectively on both sides of model (3.1) under (IC1). Then for $\{\mathbf{a}_t^{(m)}\}$ and $\{\mathbf{b}_t^{(m)}\}$ we can post-multiply by $\mathbf{1}_{q_m}$ and pre-multiply by $\mathbf{1}_{p_m}'$ on both sides of (3.1) respectively. The remaining common component is then identified as well.

For global and local effects, since $c_t^{(m)} = \mu_t + \mu_t^{(m)}$ is identified, we have $\mu_t = M^{-1} \sum_{m=1}^M c_t^{(m)}$ using (IC2), and hence both $\{\mu_t\}$ and $\{\mu_t^{(m)}\}$ are identified. Next, we in fact identify the row and column main effects at group m respectively as

$$\mathbf{a}_t^{(m)} = q_m^{-1} \mathbf{M}_{p_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m}, \quad \mathbf{b}_t^{(m)} = p_m^{-1} \mathbf{M}_{q_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)})' \mathbf{1}_{p_m}. \quad (\text{C.1})$$

Hence for $m \neq n$, using (3.2) and (IC2), we have

$$\begin{aligned} \mathbf{A}_{1m} \mathbb{E} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{\otimes 2} \right) \mathbf{A}_{1n}' &= \mathbb{E} \left(T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(n)'} \right) \\ &= q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \mathbb{E} \left\{ T^{-1} \sum_{t=1}^T (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m} \mathbf{1}_{q_n}' (\mathbf{Y}_t^{(n)} - \mathbf{E}_t^{(n)})' \right\} \mathbf{M}_{p_n} \\ &= T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \mathbb{E} \{ \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{Y}_t^{(n)'} \} \mathbf{M}_{p_n}, \\ \mathbf{B}_{1m} \mathbb{E} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\beta}_t^{\otimes 2} \right) \mathbf{B}_{1n}' &= \mathbb{E} \left(T^{-1} \sum_{t=1}^T \mathbf{b}_t^{(m)} \mathbf{b}_t^{(n)'} \right) \\ &= p_m^{-1} p_n^{-1} \mathbf{M}_{q_m} \mathbb{E} \left\{ T^{-1} \sum_{t=1}^T (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)})' \mathbf{1}_{p_m} \mathbf{1}_{p_n}' (\mathbf{Y}_t^{(n)} - \mathbf{E}_t^{(n)}) \right\} \mathbf{M}_{q_n} \\ &= T^{-1} p_m^{-1} p_n^{-1} \mathbf{M}_{q_m} \sum_{t=1}^T \mathbb{E} \{ \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m} \mathbf{1}_{p_n}' \mathbf{Y}_t^{(n)} \} \mathbf{M}_{q_n}. \end{aligned} \quad (\text{C.2})$$

If we define

$$\begin{aligned} \boldsymbol{\Omega}_{A_1, m} &:= (M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \mathbb{E} \{ \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{Y}_t^{(n)'} \} \mathbf{M}_{p_n} \right\}^{\otimes 2}, \\ \boldsymbol{\Omega}_{B_1, m} &:= (M-1)^{-1} \sum_{n=1; n \neq m}^M \left\{ T^{-1} p_m^{-1} p_n^{-1} \mathbf{M}_{q_m} \sum_{t=1}^T \mathbb{E} \{ \mathbf{Y}_t^{(m)'} \mathbf{1}_{p_m} \mathbf{1}_{p_n}' \mathbf{Y}_t^{(n)} \} \mathbf{M}_{q_n} \right\}^{\otimes 2}, \end{aligned}$$

then a version of $\mathbf{A}_{1m}$ (resp. $\mathbf{B}_{1m}$) can be identified as the first r_1 eigenvectors of $\boldsymbol{\Omega}_{A_1, m}$ (resp. first ℓ_1 eigenvectors of $\boldsymbol{\Omega}_{B_1, m}$).

As for $\mathbf{A}_{2m}$, since $\mathbf{A}_{1m}^\perp$ have orthonormal columns spanning the orthogonal complement of the column

space of $\mathbf{A}_{1m}$, from (3.2), we have

$$\mathbf{A}_{1m}^{\perp'} \mathbf{a}_t^{(m)} = \mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)}.$$

Hence

$$T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} \mathbf{A}_{1m}^{\perp} = \mathbf{A}_{1m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t^{(m)'} \right) \mathbf{A}_{2m}' \mathbf{A}_{1m}^{\perp} + \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \mathbf{A}_{2m}' \mathbf{A}_{1m}^{\perp}.$$

Hence (IC2) gives

$$\left\{ T^{-1} \sum_{t=1}^T (\mathbf{a}_t^{(m) \otimes 2}) \mathbf{A}_{1m}^{\perp} \right\}^{\otimes 2} = \mathbf{A}_{2m} \left\{ T^{-1} \sum_{t=1}^T (\boldsymbol{\alpha}_t^{(m) \otimes 2}) \mathbf{A}_{2m}' \mathbf{A}_{1m}^{\perp} \right\}^{\otimes 2} \mathbf{A}_{2m}'. \quad (\text{C.3})$$

Therefore, we can take the first $r_2^{(m)}$ eigenvectors of $\left\{ T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m) \otimes 2} \mathbf{A}_{1m}^{\perp} \right\}^{\otimes 2}$ as a version of $\mathbf{A}_{2m}$.

Similarly, we can take the first $\ell_2^{(m)}$ eigenvectors of $\left\{ T^{-1} \sum_{t=1}^T \mathbf{b}_t^{(m) \otimes 2} \mathbf{B}_{1m}^{\perp} \right\}^{\otimes 2}$ as a version of $\mathbf{B}_{2m}$.

To identify $\boldsymbol{\alpha}_t^{(m)}$, since $\mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)} = q_m^{-1} \mathbf{A}_{1m}^{\perp'} \mathbf{M}_{p_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m}$, we have

$$\boldsymbol{\alpha}_t^{(m)} = q_m^{-1} (\mathbf{A}_{2m}' \mathbf{A}_{1m}^{\perp} \mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m})^{-1} \mathbf{A}_{2m}' \mathbf{A}_{1m}^{\perp} \mathbf{A}_{1m}^{\perp'} \mathbf{M}_{p_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m}.$$

Similarly,

$$\boldsymbol{\beta}_t^{(m)} = p_m^{-1} (\mathbf{B}_{2m}' \mathbf{B}_{1m}^{\perp} \mathbf{B}_{1m}^{\perp'} \mathbf{B}_{2m})^{-1} \mathbf{B}_{2m}' \mathbf{B}_{1m}^{\perp} \mathbf{B}_{1m}^{\perp'} \mathbf{M}_{q_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)})' \mathbf{1}_{p_m}.$$

We can then identify $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t$ as

$$\begin{aligned} \boldsymbol{\alpha}_t &= (\mathbf{A}_{1m}' \mathbf{A}_{1m})^{-1} \mathbf{A}_{1m}' \{ q_m^{-1} \mathbf{M}_{p_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m} - \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)} \}, \\ \boldsymbol{\beta}_t &= (\mathbf{B}_{1m}' \mathbf{B}_{1m})^{-1} \mathbf{B}_{1m}' \{ p_m^{-1} \mathbf{M}_{q_m} (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)})' \mathbf{1}_{p_m} - \mathbf{B}_{2m} \boldsymbol{\beta}_t^{(m)} \}. \end{aligned}$$

(A refinement procedure can be carried out to estimate $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t$ as an average over all $m \in [M]$, with $\mathbf{A}_{1m}$ and $\mathbf{B}_{1m}$ re-estimated using least squares estimation. This is indeed integrated in our estimation procedure in Section 3.3.)

Define

$$\widetilde{\mathbf{Y}}_t^{(m)} := \mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)} - c_t \mathbf{1}_{p_m} \mathbf{1}_{q_m}' - \mathbf{a}_t^{(m)} \mathbf{1}_{q_m}' - \mathbf{1}_{p_m} \mathbf{b}_t^{(m)'}.$$

If we know $\mathbf{v}_1^{(m)}$ as a vector aligning more to the direction of the strongest factors in $\mathbf{Q}_m$, and $\mathbf{v}_2^{(m)}$ as such a vector for $\mathbf{J}_m$ (practically this can be done using similar pre-averaging ideas in Chen and Lam (2024), which is also integrated in our estimation procedure), then since for $m \neq n$, by (IC2),

$$\mathbb{E}(\widetilde{\mathbf{Y}}_t^{(m)} \mathbf{v}_2^{(m)} \mathbf{v}_2^{(n)'} \widetilde{\mathbf{Y}}_t^{(n)'}) = \mathbf{Q}_m \mathbb{E}(\mathbf{G}_t \mathbf{J}_m' \mathbf{v}_2^{(m)} \mathbf{v}_2^{(n)'} \mathbf{J}_n' \mathbf{G}_t') \mathbf{Q}_n'$$

$$\begin{aligned}
&= \mathbf{Q}_m \mathbb{E} \left\{ \text{vec}(\mathbf{G}_t \mathbf{J}_m' \mathbf{v}_2^{(m)}) \text{vec}(\mathbf{G}_t \mathbf{J}_n \mathbf{v}_2^{(n)})' \right\} \mathbf{Q}_n' \\
&= \mathbf{Q}_m \mathbb{E} \left\{ (\mathbf{v}_2^{(m)})' \mathbf{J}_m \otimes \mathbf{I}_{k_r} \right\} \text{vec}(\mathbf{G}_t) \text{vec}'(\mathbf{G}_t) (\mathbf{J}_m' \mathbf{v}_2^{(m)} \otimes \mathbf{I}_{k_r}) \mathbf{Q}_n' \\
&= \mathbf{Q}_m (\mathbf{v}_2^{(m)})' \mathbf{J}_m \otimes \mathbf{I}_{k_r} \mathbb{E}(\text{vec}(\mathbf{G}_t))^{\otimes 2} (\mathbf{J}_m' \mathbf{v}_2^{(m)} \otimes \mathbf{I}_{k_r}) \mathbf{Q}_n', \tag{C.4}
\end{aligned}$$

which is of rank k_r through having $\mathbb{E}(\text{vec}(\mathbf{G}_t))^{\otimes 2}$ of full rank by (IC3), we can identify $\mathbf{Q}_m$ as the first k_r eigenvectors of

$$\sum_{n=1; n \neq m}^M \{ \mathbb{E}(\tilde{\mathbf{Y}}_t^{(m)} \mathbf{v}_2^{(m)} \mathbf{v}_2^{(n)})' \tilde{\mathbf{Y}}_t^{(n)} \}^{\otimes 2}.$$

Similarly, we can identify $\mathbf{J}_m$ as the first k_c eigenvectors of

$$\sum_{n=1; n \neq m}^M \{ \mathbb{E}(\tilde{\mathbf{Y}}_t^{(m)})' \mathbf{v}_1^{(m)} \mathbf{v}_1^{(n)} \tilde{\mathbf{Y}}_t^{(n)} \}^{\otimes 2}.$$

We can also identify $\mathbf{R}_m$ as the first $k_r^{(m)}$ eigenvectors of

$$\mathbb{E}(\tilde{\mathbf{Y}}_t^{(m)} \mathbf{J}_m^\perp)^{\otimes 2} = \mathbf{R}_m \mathbb{E}[\mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{J}_m^\perp \mathbf{J}_m^{\perp'} \mathbf{C}_m \mathbf{F}_t^{(m)}] \mathbf{R}_m',$$

which is of full rank since $\mathbb{E}(\text{vec}(\mathbf{F}_t^{(m)}))^{\otimes 2}$ and $\mathbf{J}_m^{\perp'} \mathbf{C}_m$ have full rank by (IC3). Similarly, $\mathbf{C}_m$ can be identified as the first $k_c^{(m)}$ eigenvectors of

$$\mathbb{E}(\tilde{\mathbf{Y}}_t^{(m)})' \mathbf{Q}_m^\perp)^{\otimes 2} = \mathbf{C}_m \mathbb{E}[\mathbf{F}_t^{(m)} \mathbf{R}_m' \mathbf{Q}_m^\perp \mathbf{Q}_m^{\perp'} \mathbf{R}_m \mathbf{F}_t^{(m)}] \mathbf{C}_m'.$$

Then $\mathbf{F}_t^{(m)}$ can be identified as

$$\mathbf{F}_t^{(m)} = (\mathbf{R}_m' \mathbf{Q}_m^\perp \mathbf{Q}_m^{\perp'} \mathbf{R}_m)^{-1} \mathbf{R}_m' \mathbf{Q}_m^\perp \mathbf{Q}_m^{\perp'} \tilde{\mathbf{Y}}_t^{(m)} \mathbf{C}_m.$$

Finally $\mathbf{G}_t$ can be identified as

$$\mathbf{G}_t = \mathbf{Q}_m' (\tilde{\mathbf{Y}}_t^{(m)} - \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m') \mathbf{J}_m.$$

Again, we can consider the average of the above over all $m \in [M]$ as an estimator of $\mathbf{G}_t$, and re-estimate $\mathbf{Q}_m$ and $\mathbf{J}_m$ using least square method, which is indeed integrated into our estimation procedure. $\square$

Proof of Theorem 2. We first derive the convergence rate for grand mean and its corresponding global and local components. By left- and right-multiplying $\mathbf{1}_{p_m}'$ and $\mathbf{1}_{q_m}$ on (3.1), we have

$$\mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} = p_m q_m c_t^{(m)} + \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}. \tag{C.5}$$

Thus,

$$\begin{aligned}\mathbb{E}(\hat{c}_t^{(m)} - c_t^{(m)})^2 &= (p_m q_m)^{-2} \mathbb{E}(\mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m})^2 = (p_m q_m)^{-2} \mathbb{E}\left(\sum_{i=1}^{p_m} \sum_{j=1}^{q_m} E_{t,ij}^{(m)}\right)^2 \\ &= (p_m q_m)^{-2} \sum_{i=1}^{p_m} \sum_{l=1}^{p_m} \sum_{j=1}^{q_m} \sum_{h=1}^{q_m} \mathbb{E}(E_{t,ij}^{(m)} E_{t,lh}^{(m)}) = O(p_m^{-1} q_m^{-1}),\end{aligned}\tag{C.6}$$

where we apply Lemma 1 in the last equality. Hence $|\hat{c}_t^{(m)} - c_t^{(m)}| = O_P((p_m q_m)^{-1/2})$. Taking sum on both sides of (C.5) w.r.t. m , we then have $\mu_t = M^{-1} \sum_{m=1}^M (p_m q_m)^{-1} \mathbf{1}_{p_m}' (\mathbf{Y}_t^{(m)} - \mathbf{E}_t^{(m)}) \mathbf{1}_{q_m}$. Therefore, since $\hat{\mu}_t = M^{-1} \sum_{m=1}^M \hat{c}_t^{(m)}$, we have

$$\begin{aligned}(\hat{\mu}_t - \mu_t)^2 &\lesssim \left(\sum_{m=1}^M (p_m q_m)^{-1} \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\right)^2 \lesssim \sum_{m=1}^M \left((p_m q_m)^{-1} \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\right)^2 \\ &= \sum_{m=1}^M (\hat{c}_t^{(m)} - c_t^{(m)})^2 = \sum_{m=1}^M O_P((p_m q_m)^{-1}) = O_P\left(\max_{m \in [M]} (p_m q_m)^{-1}\right).\end{aligned}$$

Furthermore, as $\hat{\mu}_t^{(m)} = \hat{c}_t^{(m)} - \hat{\mu}_t$ and $\mu_t^{(m)} = c_t^{(m)} - \mu_t$, for each $m \in [M]$,

$$(\hat{\mu}_t^{(m)} - \mu_t^{(m)})^2 \lesssim (\hat{c}_t^{(m)} - c_t^{(m)})^2 + (\hat{\mu}_t - \mu_t)^2 = O_P\left(\max_{n \in [M]} (p_n q_n)^{-1}\right),$$

which is the stated rate.

By (C.1), the definitions of $\hat{\mathbf{a}}_t^{(m)}$ and $\hat{\mathbf{b}}_t^{(m)}$, and (IC1), we have

$$\|\hat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)}\|^2 = q_m^{-2} \|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\|^2 = q_m^{-2} \|\mathbf{E}_t^{(m)} \mathbf{1}_{q_m} - p_m^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\|^2.$$

For the first part, by Lemma 1,

$$\mathbb{E}\|\mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\|^2 = \sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \sum_{h=1}^{q_m} \mathbb{E}(E_{t,ij}^{(m)} E_{t,ih}^{(m)}) \leq \sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \sum_{k=1}^{p_m} \sum_{h=1}^{q_m} |\mathbb{E}(E_{t,ij}^{(m)} E_{t,kh}^{(m)})| = O(p_m q_m).$$

And for the second part, write

$$\begin{aligned}\mathbb{E}\|p_m^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\|^2 &= p_m^{-2} \mathbb{E}(\mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}) \\ &= p_m^{-1} \mathbb{E}(\mathbf{1}_{p_m}' \mathbf{E}_t^{(m)} \mathbf{1}_{q_m})^2 = O(q_m),\end{aligned}$$

where the last equality follows by (C.6). By triangle inequality, combining the two bounds, we obtain

$\mathbb{E}\|\widehat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)}\|^2 = O(p_m q_m^{-1})$. Therefore, by Markov's inequality, we have

$$p_m^{-1}\|\widehat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)}\|^2 = O_P(q_m^{-1}), \quad (\text{C.7})$$

$$q_m^{-1}\|\widehat{\mathbf{b}}_t^{(m)} - \mathbf{b}_t^{(m)}\|^2 = O_P(p_m^{-1}), \quad (\text{C.8})$$

where column effect $\widehat{\mathbf{b}}_t^{(m)}$ follows similarly. This ends the proof. $\square$

For the remaining main effect proofs and lemmas, we use the notation $\gamma_{r,\max}^{(m)} := \gamma_r \vee \gamma_r^{(m)}$ and $\gamma_{r,\min}^{(m)} := \gamma_r \wedge \gamma_r^{(m)}$, and define $\gamma_{c,\max}^{(m)}$ and $\gamma_{c,\min}^{(m)}$ analogously. These quantities are related to the notation in (R1) by

$$\begin{aligned} p_m^{\gamma_{r,\max}^{(m)} - \gamma_r} &= 1 \vee \nu_{r,m}, & p_m^{\gamma_r - \gamma_{r,\min}^{(m)}} &= 1 \vee \frac{1}{\nu_{r,m}}, & p_m^{\gamma_{r,\max}^{(m)} - \gamma_{r,\min}^{(m)}} &= \nu_{r,m} \vee \frac{1}{\nu_{r,m}}, \\ p_m^{\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}} &= \frac{1}{1 \wedge \nu_{r,m}}, & \frac{p_m^{1 - \gamma_{r,\min}^{(m)}}}{q_m} &= \frac{p_m^{1 - \gamma_r}}{q_m(1 \wedge \nu_{r,m})}. \end{aligned}$$

The column-side identities follow similarly.

Proof of Theorem 3. For $m \neq n \in [M]$, define

$$\mathbf{S}_{A_1,mn} := T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \mathbb{E} \left(\mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{Y}_t^{(n)'} \right) \mathbf{M}_{p_n}, \quad (\text{C.9})$$

$$\widehat{\mathbf{S}}_{A_1,mn} := T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \left(\mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{Y}_t^{(n)'} \right) \mathbf{M}_{p_n}. \quad (\text{C.10})$$

The corresponding population matrix and its estimator are

$$\mathbf{\Omega}_{A_1,m} = (M-1)^{-1} \sum_{n=1, n \neq m}^M \mathbf{S}_{A_1,mn}^{\otimes 2}, \quad \widehat{\mathbf{\Omega}}_{A_1,m} = (M-1)^{-1} \sum_{n=1, n \neq m}^M \widehat{\mathbf{S}}_{A_1,mn}^{\otimes 2}.$$

By (C.2), we have $\mathbf{S}_{A_1,mn} = T^{-1} \mathbf{A}_{1m} \left(\sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t') \right) \mathbf{A}_{1n}'$. Hence,

$$\mathbf{\Omega}_{A_1,m} = T^{-2} (M-1)^{-1} \mathbf{A}_{1m} \left(\sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t') \right) \left(\sum_{n \neq m} \mathbf{A}_{1n}' \mathbf{A}_{1n} \right) \left(\sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t') \right) \mathbf{A}_{1m}'. \quad (\text{C.11})$$

Define $\omega_{A_1,m} := p_m^{\gamma_r} \sum_{n \neq m} p_n^{\gamma_r}$. Since (C.11) implies that the column space of $\mathbf{\Omega}_{A_1,m}$ is contained in that of $\mathbf{A}_{1m}$, with $\mathbf{P}_{A_1,m} := \check{\mathbf{A}}_{1m} \check{\mathbf{A}}_{1m}'$, we have

$$(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{\Omega}_{A_1,m} = \mathbf{0}. \quad (\text{C.12})$$

Since $\widehat{\mathbf{A}}_{1m}$ consists of the leading r_1 eigenvectors of $\widehat{\mathbf{\Omega}}_{A_1,m}$,

$$\widehat{\mathbf{\Omega}}_{A_1,m} \widehat{\mathbf{A}}_{1m} = \widehat{\mathbf{A}}_{1m} \widehat{\mathbf{D}}_{A_1,m}, \quad \widehat{\mathbf{D}}_{A_1,m} := \text{diag}(\lambda_1(\widehat{\mathbf{\Omega}}_{A_1,m}), \dots, \lambda_{r_1}(\widehat{\mathbf{\Omega}}_{A_1,m})). \quad (\text{C.13})$$

Define $\mathbf{\Delta}_{A_1,mn} := \widehat{\mathbf{S}}_{A_1,mn} - \mathbf{S}_{A_1,mn}$. Then,

$$\widehat{\mathbf{\Omega}}_{A_1,m} - \mathbf{\Omega}_{A_1,m} = (M-1)^{-1} \sum_{n \neq m} (\mathbf{\Delta}_{A_1,mn} \mathbf{S}'_{A_1,mn} + \mathbf{S}_{A_1,mn} \mathbf{\Delta}'_{A_1,mn} + \mathbf{\Delta}_{A_1,mn} \mathbf{\Delta}'_{A_1,mn}). \quad (\text{C.14})$$

Left-multiplying (C.14) by $\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}$ and using $(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{S}_{A_1,mn} = \mathbf{0}$ gives

$$\begin{aligned} & (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})(\widehat{\mathbf{\Omega}}_{A_1,m} - \mathbf{\Omega}_{A_1,m}) \\ &= (M-1)^{-1} \sum_{n \neq m} \{ (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{\Delta}_{A_1,mn} \mathbf{S}'_{A_1,mn} + (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{\Delta}_{A_1,mn} \mathbf{\Delta}'_{A_1,mn} \}. \end{aligned}$$

By (D.43) and (D.44) in Lemma 5, $\|\mathbf{S}_{A_1,mn}\|_F^2 + \|\mathbf{\Delta}_{A_1,mn}\|_F^2 = O_P(p_m^{\gamma_r} p_n^{\gamma_r})$. Hence, (D.45) gives

$$\begin{aligned} & \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})(\widehat{\mathbf{\Omega}}_{A_1,m} - \mathbf{\Omega}_{A_1,m})\|_F^2 \\ &= O_P\left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)} + \gamma_r} \sum_{n \neq m} p_n^{\gamma_r^{(n)} + \gamma_r} + T^{-1} p_m^{\gamma_r + 1} q_m^{-1} \sum_{n \neq m} p_n^{\gamma_{r,\max}^{(n)} + \gamma_r}\right) \\ &= O_P\left(\omega_{A_1,m}^2 \left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} + T^{-1} p_m^{1 - \gamma_r} q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\right)\right). \quad (\text{C.15}) \end{aligned}$$

Define $\mathbf{H}_{A_1,m} := \check{\mathbf{A}}'_{1m} \widehat{\mathbf{A}}_{1m}$. Since $\check{\mathbf{A}}_{1m}$ and $\widehat{\mathbf{A}}_{1m}$ both have orthonormal columns, Lemma 4 gives

$$\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \widehat{\mathbf{A}}_{1m}. \quad (\text{C.16})$$

Left-multiplying (C.13) by $\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}$, and using (C.12) and (C.16), gives

$$(\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}) \widehat{\mathbf{D}}_{A_1,m} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})(\widehat{\mathbf{\Omega}}_{A_1,m} - \mathbf{\Omega}_{A_1,m}) \widehat{\mathbf{A}}_{1m}. \quad (\text{C.17})$$

Combining (C.15), (C.17), and Lemma 6, together with $\|\widehat{\mathbf{A}}_{1m}\| = 1$, we obtain

$$\begin{aligned} & \|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F^2 \leq \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})(\widehat{\mathbf{\Omega}}_{A_1,m} - \mathbf{\Omega}_{A_1,m})\|_F^2 \|\widehat{\mathbf{A}}_{1m}\|^2 \|\widehat{\mathbf{D}}_{A_1,m}^{-1}\|_F^2 \\ &= O_P\left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} + T^{-1} p_m^{1 - \gamma_r} q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\right). \end{aligned}$$

Using $p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} = 1 \vee \nu_{r,n}$, this is exactly the row-side rate stated in Theorem 3. Both terms above are $o(1)$ under (R1), so that $\|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F = o_P(1)$. Hence, Lemma 4 gives $\mathbf{H}'_{A_1,m} \mathbf{H}_{A_1,m} = \mathbf{I}_{r_1} + o_P(1)$,

$\mathbf{H}_{A_1,m} \mathbf{H}_{A_1,m}' = \mathbf{I}_{r_1} + o_P(1)$, and $\mathbf{H}_{A_1,m}$ is invertible with probability tending to one. Moreover, defining $\widehat{\mathbf{P}}_{A_1,m} := \widehat{\mathbf{A}}_{1m} \widehat{\mathbf{A}}_{1m}'$, we have

$$\|\widehat{\mathbf{P}}_{A_1,m} - \mathbf{P}_{A_1,m}\|_F^2 = 2\|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F^2. \quad (\text{C.18})$$

The proof for the global column-loading estimator $\widehat{\mathbf{B}}_{1m}$ follows similarly, with $\mathbf{H}_{B_1,m} := \check{\mathbf{B}}_{1m}' \widehat{\mathbf{B}}_{1m}$. This ends the proof. $\square$

Proof of Theorem 4. Recall that

$$\mathbf{S}_{A_2,m} := T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} \mathbf{P}_{A_1,m}^\perp, \quad \widehat{\mathbf{S}}_{A_2,m} := T^{-1} \sum_{t=1}^T \widehat{\mathbf{a}}_t^{(m)} \widehat{\mathbf{a}}_t^{(m)'} \widehat{\mathbf{P}}_{A_1,m}^\perp, \quad (\text{C.19})$$

where $\mathbf{P}_{A_1,m}^\perp := \mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}$ and $\widehat{\mathbf{P}}_{A_1,m}^\perp := \mathbf{I}_{p_m} - \widehat{\mathbf{P}}_{A_1,m}$. Since $\widehat{\mathbf{A}}_{1m}^\perp \widehat{\mathbf{A}}_{1m}^{\perp'} = \widehat{\mathbf{P}}_{A_1,m}^\perp$, the sample matrix in the estimation procedure satisfies

$$\widehat{\mathbf{\Omega}}_{A_2,m} = \left(T^{-1} \sum_{t=1}^T \widehat{\mathbf{a}}_t^{(m)} \widehat{\mathbf{a}}_t^{(m)'} \right) \widehat{\mathbf{P}}_{A_1,m}^\perp \left(T^{-1} \sum_{t=1}^T \widehat{\mathbf{a}}_t^{(m)} \widehat{\mathbf{a}}_t^{(m)'} \right) = \widehat{\mathbf{S}}_{A_2,m} \widehat{\mathbf{S}}_{A_2,m}'. \quad (\text{C.20})$$

Define its population counterpart as

$$\mathbf{\Omega}_{A_2,m} := \left(T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} \right) \mathbf{P}_{A_1,m}^\perp \left(T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} \right) = \mathbf{S}_{A_2,m} \mathbf{S}_{A_2,m}'. \quad (\text{C.21})$$

By (IC2),

$$\mathbf{S}_{A_2,m} = \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \mathbf{A}_{2m}' \mathbf{P}_{A_1,m}^\perp.$$

Hence, with $\mathbf{P}_{A_2,m} := \check{\mathbf{A}}_{2m} \check{\mathbf{A}}_{2m}'$,

$$(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}) \mathbf{S}_{A_2,m} = \mathbf{0}, \quad (\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}) \mathbf{\Omega}_{A_2,m} = \mathbf{0}. \quad (\text{C.22})$$

Since $\widehat{\mathbf{A}}_{2m}$ consists of the leading $r_2^{(m)}$ eigenvectors of $\widehat{\mathbf{\Omega}}_{A_2,m}$,

$$\widehat{\mathbf{\Omega}}_{A_2,m} \widehat{\mathbf{A}}_{2m} = \widehat{\mathbf{A}}_{2m} \widehat{\mathbf{D}}_{A_2,m}, \quad \widehat{\mathbf{D}}_{A_2,m} := \text{diag}(\lambda_1(\widehat{\mathbf{\Omega}}_{A_2,m}), \dots, \lambda_{r_2^{(m)}}(\widehat{\mathbf{\Omega}}_{A_2,m})). \quad (\text{C.23})$$

Define $\mathbf{H}_{A_2,m} := \check{\mathbf{A}}_{2m}' \widehat{\mathbf{A}}_{2m}$. Since $\check{\mathbf{A}}_{2m}$ and $\widehat{\mathbf{A}}_{2m}$ both have orthonormal columns, Lemma 4 gives

$$\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}) \widehat{\mathbf{A}}_{2m}. \quad (\text{C.24})$$

Left-multiplying (C.23) by $\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}$, and using (C.22) and (C.24), gives

$$(\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}) \widehat{\mathbf{D}}_{A_2,m} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})(\widehat{\mathbf{\Omega}}_{A_2,m} - \mathbf{\Omega}_{A_2,m}) \widehat{\mathbf{A}}_{2m}. \quad (\text{C.25})$$

Let $\mathbf{\Delta}_{A_2,m} := \widehat{\mathbf{S}}_{A_2,m} - \mathbf{S}_{A_2,m}$. Since

$$\widehat{\mathbf{\Omega}}_{A_2,m} - \mathbf{\Omega}_{A_2,m} = \mathbf{\Delta}_{A_2,m} \mathbf{S}'_{A_2,m} + \mathbf{S}_{A_2,m} \mathbf{\Delta}'_{A_2,m} + \mathbf{\Delta}_{A_2,m} \mathbf{\Delta}'_{A_2,m},$$

using $(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}) \mathbf{S}_{A_2,m} = \mathbf{0}$ gives

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})(\widehat{\mathbf{\Omega}}_{A_2,m} - \mathbf{\Omega}_{A_2,m})\|_F^2 \lesssim \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}) \mathbf{\Delta}_{A_2,m}\|_F^2 (\|\mathbf{S}_{A_2,m}\|_F^2 + \|\mathbf{\Delta}_{A_2,m}\|_F^2). \quad (\text{C.26})$$

By (D.56) and (D.57) in Lemma 7, $\|\mathbf{S}_{A_2,m}\|_F^2 + \|\mathbf{\Delta}_{A_2,m}\|_F^2 = O_P(p_m^{2\gamma_r^{(m)}})$. Combining (C.26), (D.58), (C.25), and Lemma 8, together with $\|\widehat{\mathbf{A}}_{2m}\| = 1$, we obtain

$$\begin{aligned} \|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F^2 &\leq \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})(\widehat{\mathbf{\Omega}}_{A_2,m} - \mathbf{\Omega}_{A_2,m})\|_F^2 \|\widehat{\mathbf{A}}_{2m}\|^2 \|\widehat{\mathbf{D}}_{A_2,m}^{-1}\|_F^2 \\ &= O_P\left(R(T, d_{r,\star}) p_m^{\gamma_r - \gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} + T^{-1} p_m^{1 + \gamma_r - 2\gamma_r^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} + p_m^{1 - 2\gamma_r^{(m)}} q_m^{-2}\right). \end{aligned}$$

Together with the definition of $\nu_{r,m}$, this proves the row-side loading rate in Theorem 4. All three terms above are $o(1)$ under (R1), so that $\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F = o_P(1)$. Hence, Lemma 4 gives $\mathbf{H}'_{A_2,m} \mathbf{H}_{A_2,m} = \mathbf{I}_{r_2^{(m)}} + o_P(1)$, $\mathbf{H}_{A_2,m} \mathbf{H}'_{A_2,m} = \mathbf{I}_{r_2^{(m)}} + o_P(1)$, and $\mathbf{H}_{A_2,m}$ is invertible with probability tending to one. Moreover, defining $\widehat{\mathbf{P}}_{A_2,m} := \widehat{\mathbf{A}}_{2m} \widehat{\mathbf{A}}_{2m}'$, we have

$$\|\widehat{\mathbf{P}}_{A_2,m} - \mathbf{P}_{A_2,m}\|_F^2 = 2\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F^2. \quad (\text{C.27})$$

The proof for the local column-loading estimator $\widehat{\mathbf{B}}_{2m}$ follows similarly, with $\mathbf{H}_{B_2,m} := \check{\mathbf{B}}_{2m}' \widehat{\mathbf{B}}_{2m}$.

We next prove the local main effect factor series. For $m \in [M]$, define

$$\mathcal{L}_{\alpha,m} := \mathbf{H}_{A_2,m}^{-1} (\check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1,m}^\perp \check{\mathbf{A}}_{2m})^{-1} \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1,m}^\perp, \quad \widehat{\mathcal{L}}_{\alpha,m} := (\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1,m}^\perp \widehat{\mathbf{A}}_{2m})^{-1} \widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1,m}^\perp.$$

By (S1) and the asymptotic orthogonality of $\mathbf{H}_{A_2,m}$, we have $\|\mathcal{L}_{\alpha,m}\| = O_P(1)$. Moreover, $\mathcal{L}_{\alpha,m} \mathbf{A}_{1m} = \mathbf{0}$, $\mathcal{L}_{\alpha,m} \check{\mathbf{A}}_{2m} = \mathbf{H}_{A_2,m}^{-1}$, $\widehat{\mathcal{L}}_{\alpha,m} \widehat{\mathbf{A}}_{2m} = \mathbf{I}_{r_2^{(m)}}$, and $\widehat{\mathcal{L}}_{\alpha,m} \widehat{\mathbf{P}}_{A_1,m} = \mathbf{0}$.

We first control $\widehat{\mathcal{L}}_{\alpha,m} - \mathcal{L}_{\alpha,m}$. Note that

$$\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1,m}^\perp \widehat{\mathbf{A}}_{2m} - \mathbf{H}'_{A_2,m} \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1,m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}$$

$$\begin{aligned}
&= (\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m})' \widehat{\mathbf{P}}_{A_1, m}^\perp \widehat{\mathbf{A}}_{2m} + \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' (\widehat{\mathbf{P}}_{A_1, m}^\perp - \mathbf{P}_{A_1, m}^\perp) \widehat{\mathbf{A}}_{2m} \\
&\quad + \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp (\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}).
\end{aligned}$$

Since $\check{\mathbf{A}}_{2m}$ and $\widehat{\mathbf{A}}_{2m}$ have orthonormal columns, $\|\mathbf{H}_{A_2, m}\| = O_P(1)$, and $\|\widehat{\mathbf{P}}_{A_1, m}^\perp - \mathbf{P}_{A_1, m}^\perp\| = \|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|$, it follows that

$$\begin{aligned}
&\|\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1, m}^\perp \widehat{\mathbf{A}}_{2m} - \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\| \\
&= O_P\left(\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\|_F + \|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|\right).
\end{aligned} \tag{C.28}$$

By (S1), the eigenvalues of $\check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m}$ are bounded away from zero. Together with the asymptotic orthogonality of $\mathbf{H}_{A_2, m}$, this implies $\lambda_{r_2^{(m)}}(\mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}) \gtrsim_P 1$. Since the right-hand side of (C.28) is $o_P(1)$, Weyl's inequality further gives

$$\|(\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1, m}^\perp \widehat{\mathbf{A}}_{2m})^{-1}\| = O_P(1), \quad \|(\mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m})^{-1}\| = O_P(1).$$

Therefore, using $\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}$,

$$\begin{aligned}
&\|(\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1, m}^\perp \widehat{\mathbf{A}}_{2m})^{-1} - (\mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m})^{-1}\| \\
&= O_P\left(\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\|_F + \|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|\right).
\end{aligned} \tag{C.29}$$

Moreover,

$$\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1, m}^\perp - \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp = (\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m})' \widehat{\mathbf{P}}_{A_1, m}^\perp + \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' (\widehat{\mathbf{P}}_{A_1, m}^\perp - \mathbf{P}_{A_1, m}^\perp),$$

and hence

$$\|\widehat{\mathbf{A}}_{2m}' \widehat{\mathbf{P}}_{A_1, m}^\perp - \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp\| = O_P\left(\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\|_F + \|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|\right). \tag{C.30}$$

Since $\mathcal{L}_{\alpha, m} = (\mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m})^{-1} \mathbf{H}_{A_2, m}' \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1, m}^\perp$, combining (C.29) and (C.30) yields

$$\|\widehat{\mathcal{L}}_{\alpha, m} - \mathcal{L}_{\alpha, m}\| = O_P\left(\|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\|_F + \|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|\right).$$

Thus, by Lemma 4,

$$\|\widehat{\mathcal{L}}_{\alpha, m} - \mathcal{L}_{\alpha, m}\|^2 = O_P\left(\|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1, m}\|_F^2 + \|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2, m}\|_F^2\right). \tag{C.31}$$

In particular, $\|\widehat{\mathcal{L}}_{\alpha,m}\| = O_P(1)$.

By the definition of $\widehat{\alpha}_t^{(m)}$ and $\widehat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)} = q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$,

$$\begin{aligned} \widehat{\alpha}_t^{(m)} - \mathbf{H}_{A_2,m}^{-1} \check{\alpha}_t^{(m)} &= q_m^{-1} \mathcal{L}_{\alpha,m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} + q_m^{-1} (\widehat{\mathcal{L}}_{\alpha,m} - \mathcal{L}_{\alpha,m}) \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \\ &\quad + \widehat{\mathcal{L}}_{\alpha,m} (\mathbf{P}_{A_1,m} - \widehat{\mathbf{P}}_{A_1,m}) \mathbf{A}_{1m} \alpha_t - \widehat{\mathcal{L}}_{\alpha,m} (\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}) \mathbf{H}_{A_2,m}^{-1} \check{\alpha}_t^{(m)} \\ &=: \mathcal{I}_{\alpha,1}^{(m)} + \mathcal{I}_{\alpha,2}^{(m)} + \mathcal{I}_{\alpha,3}^{(m)} + \mathcal{I}_{\alpha,4}^{(m)}, \end{aligned} \quad (\text{C.32})$$

where the first equality uses $\mathcal{L}_{\alpha,m} \mathbf{A}_{1m} = \mathbf{0}$, $\widehat{\mathcal{L}}_{\alpha,m} \widehat{\mathbf{P}}_{A_1,m} = \mathbf{0}$, $\mathcal{L}_{\alpha,m} \check{\mathbf{A}}_{2m} = \mathbf{H}_{A_2,m}^{-1}$ and $\widehat{\mathcal{L}}_{\alpha,m} \widehat{\mathbf{A}}_{2m} = \mathbf{I}_{r_2}^{(m)}$.

For the first term in (C.32), note that the matrix $(\check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1,m}^\perp \check{\mathbf{A}}_{2m})^{-1} \check{\mathbf{A}}_{2m}' \mathbf{P}_{A_1,m}^\perp$ has bounded Frobenius norm by (S1). Since $\|\mathbf{H}_{A_2,m}^{-1}\| = O_P(1)$, Lemma 1 gives $\|\mathcal{I}_{\alpha,1}^{(m)}\|^2 = O_P(q_m^{-1})$. Moreover, Lemma 1, (C.31), (ME1) and (ME2) yield respectively

$$\begin{aligned} \|\mathcal{I}_{\alpha,2}^{(m)}\|^2 &= O_P(p_m q_m^{-1} (\|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F^2 + \|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F^2)), \\ \|\mathcal{I}_{\alpha,3}^{(m)}\|^2 &= O_P(p_m^{\gamma_r} \|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F^2), \quad \|\mathcal{I}_{\alpha,4}^{(m)}\|^2 = O_P(p_m^{\gamma_r^{(m)}} \|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F^2). \end{aligned}$$

Consequently,

$$\begin{aligned} p_m^{-\gamma_r^{(m)}} \|\widehat{\alpha}_t^{(m)} - \mathbf{H}_{A_2,m}^{-1} \check{\alpha}_t^{(m)}\|^2 &= O_P(p_m^{-\gamma_r^{(m)}} q_m^{-1} + (p_m^{1-\gamma_r^{(m)}} q_m^{-1} + p_m^{\gamma_r - \gamma_r^{(m)}}) \|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1,m}\|_F^2 \\ &\quad + (p_m^{1-\gamma_r^{(m)}} q_m^{-1} + 1) \|\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m} \mathbf{H}_{A_2,m}\|_F^2). \end{aligned}$$

Substituting the rates from Theorems 3 and 4 obtained above, and using $1 + p_m^{\gamma_r - \gamma_r^{(m)}} \lesssim p_m^{\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}}$, and noting that $p_m^{1-2\gamma_r^{(m)}} q_m^{-2} = o(p_m^{-\gamma_r^{(m)}} q_m^{-1})$ under (R1), gives

$$\begin{aligned} p_m^{-\gamma_r^{(m)}} \|\widehat{\alpha}_t^{(m)} - \mathbf{H}_{A_2,m}^{-1} \check{\alpha}_t^{(m)}\|^2 &= O_P(p_m^{-\gamma_r^{(m)}} q_m^{-1} + R(T, d_{r,*}) p_m^{\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \\ &\quad + T^{-1} p_m^{1+\gamma_{r,\max}^{(m)} - 2\gamma_r^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}). \end{aligned} \quad (\text{C.33})$$

Since $p_m^{\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}} = (1 \wedge \nu_{r,m})^{-1}$ and $p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} = 1 \vee \nu_{r,n}$, this is the row-side factor rate stated in Theorem 4. The column-side rate follows by symmetry. This ends the proof. $\square$

Proof of Theorem 5. We prove the row-side results; the column-side results follow by symmetry. Let $\mathbf{A}_\alpha := (\alpha_1, \dots, \alpha_T)$, $\Gamma_{\alpha,T} := T^{-1} \mathbf{A}_\alpha \mathbf{A}_\alpha'$, $\check{\mathbf{A}}_\alpha := \Gamma_{\alpha,T}^{-1/2} \mathbf{A}_\alpha = (\check{\alpha}_1, \dots, \check{\alpha}_T)$. Also write $\widehat{\mathbf{A}}_\alpha := (\widehat{\alpha}_1, \dots, \widehat{\alpha}_T)$ and define $\mathbf{P}_\alpha := T^{-1} \check{\mathbf{A}}_\alpha' \check{\mathbf{A}}_\alpha$. By Lemma 2, (S1), and Weyl's inequality, $\lambda_{r_1}(\Gamma_{\alpha,T}) \asymp_P \lambda_1(\Gamma_{\alpha,T}) \asymp_P 1$. Hence $\|\Gamma_{\alpha,T}^{-1/2}\| = O_P(1)$ and $\|\check{\alpha}_t\| = O_P(1)$ for each $t \in [T]$.

For each $m \in [M]$, define $\mathbf{Z}_{\alpha,m} := \widehat{\mathbf{A}}_{1m}' \mathbf{A}_{1m} \mathbf{A}_\alpha$ and $\Delta_{\alpha,m} := \widehat{\mathbf{Z}}_{\alpha,m} - \mathbf{Z}_{\alpha,m}$, and let $\delta_{\alpha,m,t}$ denote the

t th column of $\mathbf{\Delta}_{\alpha,m}$. Since $\widehat{\mathbf{A}}'_{1m}\mathbf{A}_{1m} = \mathbf{H}'_{A_1,m}(\mathbf{A}'_{1m}\mathbf{A}_{1m})^{1/2}$, Theorem 3 and (ME1) imply that $\widehat{\mathbf{A}}'_{1m}\mathbf{A}_{1m}$ is invertible with probability tending to one and has squared singular values of order $p_m^{\gamma_r}$. Thus,

$$\mathbf{P}_{\mathbf{Z}_{\alpha,m}} = \mathbf{A}'_{\alpha}(\mathbf{A}_{\alpha}\mathbf{A}'_{\alpha})^{-1}\mathbf{A}_{\alpha} = \mathbf{P}_{\alpha}, \quad \sigma_{r_1}^2(\mathbf{Z}_{\alpha,m}) \asymp_P T p_m^{\gamma_r}. \quad (\text{C.34})$$

Write

$$\begin{aligned} \widehat{\mathbf{y}}_t^{(m)} - \mathbf{A}_{1m}\boldsymbol{\alpha}_t &= q_m^{-1}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m} - (\widehat{\mathbf{A}}_{2m}\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)}), \\ \boldsymbol{\delta}_{\alpha,m,t} &= \widehat{\mathbf{A}}'_{1m}\left(q_m^{-1}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m} - (\widehat{\mathbf{A}}_{2m}\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)})\right). \end{aligned}$$

Using $\widehat{\mathbf{A}}_{2m}\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)} = \widehat{\mathbf{A}}_{2m}(\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{H}_{A_2,m}^{-1}\check{\boldsymbol{\alpha}}_t^{(m)}) + (\widehat{\mathbf{A}}_{2m} - \check{\mathbf{A}}_{2m}\mathbf{H}_{A_2,m})\mathbf{H}_{A_2,m}^{-1}\check{\boldsymbol{\alpha}}_t^{(m)}$, the loading and series rates in Theorem 4, the asymptotic orthogonality of $\mathbf{H}_{A_2,m}$, and $\|\check{\boldsymbol{\alpha}}_t^{(m)}\|^2 = O_P(p_m^{\gamma_r^{(m)}})$ give

$$\begin{aligned} \|\widehat{\mathbf{A}}_{2m}\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)}\|^2 &= O_P\left(q_m^{-1} + R(T, d_{r,*})p_m^{\gamma_{r,\max}^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1}p_m^{1+\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} \right). \end{aligned} \quad (\text{C.35})$$

The same rate holds for $T^{-1} \sum_{t=1}^T \|\widehat{\mathbf{A}}_{2m}\widehat{\boldsymbol{\alpha}}_t^{(m)} - \mathbf{A}_{2m}\boldsymbol{\alpha}_t^{(m)}\|^2$. Indeed, the same argument for $\mathcal{I}_{\alpha,1}^{(m)}$ gives $T^{-1} \sum_t \|\mathcal{I}_{\alpha,1}^{(m)}\|^2 = O_P(q_m^{-1})$. Applying the remaining bounds in (C.32), together with

$$\begin{aligned} T^{-1} \sum_t \|q_m^{-1}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m}\|^2 &= O_P(p_m q_m^{-1}), \\ T^{-1} \sum_t \|\mathbf{A}_{1m}\boldsymbol{\alpha}_t\|^2 &= O_P(p_m^{\gamma_r}), \quad T^{-1} \sum_t \|\check{\boldsymbol{\alpha}}_t^{(m)}\|^2 = O_P(p_m^{\gamma_r^{(m)}}), \end{aligned}$$

gives the stated time-average reconstruction rate.

Next, since

$$q_m^{-1}\widehat{\mathbf{A}}'_{1m}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m} = q_m^{-1}\mathbf{H}'_{A_1,m}\check{\mathbf{A}}'_{1m}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m} + q_m^{-1}(\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m}\mathbf{H}_{A_1,m})'\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m},$$

Lemma 1 gives $\|q_m^{-1}\check{\mathbf{A}}'_{1m}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m}\|^2 = O_P(q_m^{-1})$, with the same rate after averaging over t . Using Theorem 3 and (R1), this gives

$$\begin{aligned} &\|q_m^{-1}\widehat{\mathbf{A}}'_{1m}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m}\|^2, \quad T^{-1} \sum_{t=1}^T \|q_m^{-1}\widehat{\mathbf{A}}'_{1m}\mathbf{M}_{p_m}\mathbf{E}_t^{(m)}\mathbf{1}_{q_m}\|^2 \\ &= O_P\left(q_m^{-1} + R(T, d_{r,*})p_m^{\gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} + T^{-1}p_m q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\right), \end{aligned}$$

which is bounded by the right-hand side of (C.35). Since $\|\hat{\mathbf{A}}_{1m}\| = 1$, we therefore obtain

$$\begin{aligned} \|\delta_{\alpha,m,t}\|^2, \quad T^{-1}\|\Delta_{\alpha,m}\|_F^2 = O_P\Big(& q_m^{-1} + R(T, d_{r,\star}) p_m^{\gamma_{r,\max}^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \\ & + T^{-1} p_m^{1+\gamma_{r,\max}^{(m)} - \gamma_r^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\Big). \end{aligned} \quad (\text{C.36})$$

Define $\hat{\mathbf{P}}_{\alpha,m} := \hat{\mathbf{Z}}'_{\alpha,m} (\hat{\mathbf{Z}}_{\alpha,m} \hat{\mathbf{Z}}'_{\alpha,m})^{-1} \hat{\mathbf{Z}}_{\alpha,m}$. By (C.34), (C.36), and (R1), $\|\Delta_{\alpha,m}\|_F^2 = o_P(T p_m^{\gamma_r})$. Hence Weyl's inequality gives $\sigma_{r_1}^2(\hat{\mathbf{Z}}_{\alpha,m}) \asymp_P T p_m^{\gamma_r}$, so that $\hat{\mathbf{Z}}_{\alpha,m} \hat{\mathbf{Z}}'_{\alpha,m}$ is invertible with probability tending to one. Since $\mathbf{P}_\alpha^\perp \mathbf{Z}'_{\alpha,m} = \mathbf{0}$,

$$\mathbf{P}_\alpha^\perp \hat{\mathbf{P}}_{\alpha,m} = \mathbf{P}_\alpha^\perp \Delta'_{\alpha,m} (\hat{\mathbf{Z}}_{\alpha,m} \hat{\mathbf{Z}}'_{\alpha,m})^{-1} \hat{\mathbf{Z}}_{\alpha,m}. \quad (\text{C.37})$$

Using $\|(\hat{\mathbf{Z}}_{\alpha,m} \hat{\mathbf{Z}}'_{\alpha,m})^{-1} \hat{\mathbf{Z}}_{\alpha,m}\| = \sigma_{r_1}^{-1}(\hat{\mathbf{Z}}_{\alpha,m})$ and $\|\hat{\mathbf{P}}_{\alpha,m} - \mathbf{P}_\alpha\|_F^2 = 2\|\mathbf{P}_\alpha^\perp \hat{\mathbf{P}}_{\alpha,m}\|_F^2$, we obtain

$$\begin{aligned} \|\hat{\mathbf{P}}_{\alpha,m} - \mathbf{P}_\alpha\|_F^2 = O_P\Big(& p_m^{-\gamma_r} q_m^{-1} + R(T, d_{r,\star}) p_m^{\gamma_{r,\max}^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \\ & + T^{-1} p_m^{1-\gamma_{r,\min}^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\Big). \end{aligned} \quad (\text{C.38})$$

All terms on the right-hand side are $o(1)$ under (R1).

For the rowwise bound, let $\mathbf{U}_\alpha := T^{-1/2} \check{\mathbf{A}}'_\alpha$, so that $\mathbf{P}_\alpha = \mathbf{U}_\alpha \mathbf{U}'_\alpha$ and $\|\mathbf{e}'_t \mathbf{U}_\alpha\|^2 = T^{-1} \|\check{\boldsymbol{\alpha}}_t\|^2 = O_P(T^{-1})$. Hence

$$\|\mathbf{e}'_t \mathbf{P}_\alpha^\perp \Delta'_{\alpha,m}\|^2 \lesssim \|\delta_{\alpha,m,t}\|^2 + T^{-1} \|\check{\boldsymbol{\alpha}}_t\|^2 \|\Delta_{\alpha,m}\|_F^2.$$

Combining this with (C.37) and (C.36) yields

$$\begin{aligned} T \|\mathbf{e}'_t \mathbf{P}_\alpha^\perp (\hat{\mathbf{P}}_{\alpha,m} - \mathbf{P}_\alpha)\|^2 = O_P\Big(& p_m^{-\gamma_r} q_m^{-1} + R(T, d_{r,\star}) p_m^{\gamma_{r,\max}^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \\ & + T^{-1} p_m^{1-\gamma_{r,\min}^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r}\Big). \end{aligned} \quad (\text{C.39})$$

Now let $\hat{\mathbf{K}}_\alpha := M^{-1} \sum_{m=1}^M \hat{\mathbf{P}}_{\alpha,m}$ and $\hat{\mathbf{U}}_\alpha := T^{-1/2} \hat{\mathbf{A}}'_\alpha$, and define $\mathbf{H}_\alpha := \mathbf{U}'_\alpha \hat{\mathbf{U}}_\alpha = T^{-1} \check{\mathbf{A}}_\alpha \hat{\mathbf{A}}'_\alpha$. By construction, $\hat{\mathbf{K}}_\alpha \hat{\mathbf{U}}_\alpha = \hat{\mathbf{U}}_\alpha \hat{\mathbf{D}}_\alpha$, where $\hat{\mathbf{D}}_\alpha$ contains the leading r_1 eigenvalues of $\hat{\mathbf{K}}_\alpha$. Since M is fixed, (C.38) gives $\|\hat{\mathbf{K}}_\alpha - \mathbf{P}_\alpha\|_F = o_P(1)$. The nonzero eigenvalues of $\mathbf{P}_\alpha$ are all equal to one, so $\lambda_{r_1}(\hat{\mathbf{K}}_\alpha) \asymp_P 1$ and $\|\hat{\mathbf{D}}_\alpha^{-1}\| = O_P(1)$.

By Lemma 4, $\hat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha = \mathbf{P}_\alpha^\perp \hat{\mathbf{U}}_\alpha$. Left-multiplying the preceding eigenequation by $\mathbf{P}_\alpha^\perp$ therefore

gives

$$(\widehat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha) \widehat{\mathbf{D}}_\alpha = \mathbf{P}_\alpha^\perp (\widehat{\mathbf{K}}_\alpha - \mathbf{P}_\alpha) \widehat{\mathbf{U}}_\alpha. \quad (\text{C.40})$$

Equations (C.38) and (C.40) imply $\|\widehat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha\|_F = o_P(1)$. Hence Lemma 4 shows that $\mathbf{H}_\alpha$ is asymptotically orthogonal.

For each $t \in [T]$, left-multiplying (C.40) by $\mathbf{e}'_t$ and using (C.39) gives

$$\begin{aligned} \|\widehat{\boldsymbol{\alpha}}_t - \mathbf{H}'_\alpha \check{\boldsymbol{\alpha}}_t\|^2 = O_P \left(\max_{m \in [M]} \left\{ p_m^{-\gamma_r} q_m^{-1} + R(T, d_{r,*}) p_m^{\gamma_{r,\max}^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \right. \\ \left. \left. + T^{-1} p_m^{1 - \gamma_{r,\min}^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} \right\} \right). \end{aligned}$$

Moreover, $\mathbf{H}'_\alpha \mathbf{H}_\alpha = \mathbf{I} - (\widehat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha)' (\widehat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha)$. Since $\mathbf{H}_\alpha^{-1} = (\mathbf{H}'_\alpha \mathbf{H}_\alpha)^{-1} \mathbf{H}'_\alpha$ and $\mathbf{H}_\alpha$ is asymptotically orthogonal, $\|\mathbf{H}_\alpha^{-1} - \mathbf{H}'_\alpha\| = O_P(\|\widehat{\mathbf{U}}_\alpha - \mathbf{U}_\alpha \mathbf{H}_\alpha\|_F^2)$. Together with $\|\check{\boldsymbol{\alpha}}_t\| = O_P(1)$, replacing $\mathbf{H}'_\alpha$ by $\mathbf{H}_\alpha^{-1}$ contributes only the square of the preceding bound and is therefore negligible under (R1). Therefore,

$$\begin{aligned} \|\widehat{\boldsymbol{\alpha}}_t - \mathbf{H}_\alpha^{-1} \check{\boldsymbol{\alpha}}_t\|^2 = O_P \left(\max_{m \in [M]} \left\{ p_m^{-\gamma_r} q_m^{-1} + R(T, d_{r,*}) p_m^{\gamma_{r,\max}^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \right. \\ \left. \left. + T^{-1} p_m^{1 - \gamma_{r,\min}^{(m)}} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} \right\} \right), \end{aligned}$$

which proves the stated rate for the global row main effect series.

We next consider the refitted global loading $\tilde{\mathbf{A}}_{1m}$. Let $\widehat{\mathbf{Y}}_m := (\widehat{\mathbf{y}}_1^{(m)}, \dots, \widehat{\mathbf{y}}_T^{(m)})$ and $\mathbf{Y}_m := \mathbf{A}_{1m} \mathbf{A}_\alpha$. By (3.7) and the definitions of $\check{\mathbf{A}}_{1m}$ and $\mathbf{H}_\alpha$,

$$\tilde{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_\alpha = T^{-1} (\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \widehat{\mathbf{A}}'_\alpha. \quad (\text{C.41})$$

Define $\boldsymbol{\Pi}_{2|1,m} := \check{\mathbf{A}}_{2m} (\check{\mathbf{A}}'_{2m} \mathbf{P}_{A_1,m}^\perp \check{\mathbf{A}}_{2m})^{-1} \check{\mathbf{A}}'_{2m} \mathbf{P}_{A_1,m}^\perp$, and define $\widehat{\boldsymbol{\Pi}}_{2|1,m}$ analogously using $\widehat{\mathbf{A}}_{2m}$ and $\widehat{\mathbf{P}}_{A_1,m}^\perp$. By (S1) and Theorem 4, $\|\widehat{\boldsymbol{\Pi}}_{2|1,m}\| = O_P(1)$. Since $\widehat{\mathbf{A}}_{2m} \widehat{\boldsymbol{\alpha}}_t^{(m)} = \widehat{\boldsymbol{\Pi}}_{2|1,m} \widehat{\mathbf{a}}_t^{(m)}$ and $\mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)} = \boldsymbol{\Pi}_{2|1,m} \mathbf{a}_t^{(m)}$, we have

$$\widehat{\mathbf{y}}_t^{(m)} - \mathbf{A}_{1m} \boldsymbol{\alpha}_t = q_m^{-1} (\mathbf{I}_{p_m} - \widehat{\boldsymbol{\Pi}}_{2|1,m}) \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} - (\widehat{\boldsymbol{\Pi}}_{2|1,m} - \boldsymbol{\Pi}_{2|1,m}) \mathbf{a}_t^{(m)}.$$

By (IC2), $T^{-1} \sum_t \mathbf{a}_t^{(m)} \check{\boldsymbol{\alpha}}'_t = \mathbf{A}_{1m} \boldsymbol{\Gamma}_{\alpha,T}^{1/2} = \check{\mathbf{A}}_{1m}$ and $\boldsymbol{\Pi}_{2|1,m} \check{\mathbf{A}}_{1m} = \mathbf{0}$. Consequently,

$$T^{-1} (\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \check{\mathbf{A}}'_\alpha = (\mathbf{I}_{p_m} - \widehat{\boldsymbol{\Pi}}_{2|1,m}) \left(T^{-1} \sum_{t=1}^T q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \check{\boldsymbol{\alpha}}'_t \right) - \widehat{\boldsymbol{\Pi}}_{2|1,m} \check{\mathbf{A}}_{1m}.$$

Since $T^{-1} \sum_t q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \check{\alpha}'_t = q_m^{-1} \mathbf{M}_{p_m} (T^{-1} \sum_t \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \alpha'_t) \Gamma_{\alpha, T}^{-1/2}$, (D.33) in Lemma 3 gives its squared Frobenius norm as $O_P(T^{-1} p_m q_m^{-1})$. Also $\mathbf{P}_{A_1, m}^\perp \check{\check{\mathbf{A}}}_{1m} = \mathbf{0}$, and hence

$$\|\widehat{\mathbf{\Pi}}_{2|1, m} \check{\check{\mathbf{A}}}_{1m}\|_F^2 \lesssim_P \|(\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}) \check{\check{\mathbf{A}}}_{1m}\|_F^2 = O_P\left(p_m^{\gamma_r} \|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1, m}\|_F^2\right).$$

Substituting Theorem 3 gives

$$\begin{aligned} \|T^{-1}(\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \check{\mathbf{A}}'_\alpha\|_F^2 &= O_P\left(R(T, d_{r, \star}) p_m^{\gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1} p_m q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r, \max}^{(n)} - \gamma_r}\right). \end{aligned} \quad (\text{C.42})$$

Furthermore, (C.35) and $T^{-1} \sum_t \|q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}\|^2 = O_P(p_m q_m^{-1})$ give

$$T^{-1} \|\widehat{\mathbf{Y}}_m - \mathbf{Y}_m\|_F^2 = O_P\left(p_m q_m^{-1} + R(T, d_{r, \star}) p_m^{\gamma_{r, \max}^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r}\right), \quad (\text{C.43})$$

where the remaining term in (C.35) is absorbed into $p_m q_m^{-1}$ under (R1).

Finally, since $\mathbf{P}_\alpha \widehat{\mathbf{A}}'_\alpha = \check{\mathbf{A}}'_\alpha \mathbf{H}_\alpha$, (C.41) gives

$$\check{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_\alpha = (T^{-1}(\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \check{\mathbf{A}}'_\alpha) \mathbf{H}_\alpha + T^{-1}(\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \mathbf{P}_\alpha^\perp \widehat{\mathbf{A}}'_\alpha.$$

The first term is controlled by (C.42). For the second term, since $\mathbf{P}_\alpha^\perp \widehat{\mathbf{A}}'_\alpha = \sqrt{T} \mathbf{P}_\alpha^\perp \widehat{\mathbf{U}}_\alpha$,

$$\|T^{-1}(\widehat{\mathbf{Y}}_m - \mathbf{Y}_m) \mathbf{P}_\alpha^\perp \widehat{\mathbf{A}}'_\alpha\|_F^2 \leq \{T^{-1} \|\widehat{\mathbf{Y}}_m - \mathbf{Y}_m\|_F^2\} \|\mathbf{P}_\alpha^\perp \widehat{\mathbf{U}}_\alpha\|_F^2.$$

By (C.38) and (C.40), $\|\mathbf{P}_\alpha^\perp \widehat{\mathbf{U}}_\alpha\|_F^2$ satisfies the same bound obtained above for the global row main effect series. Combining this with (C.42) and (C.43), and dividing the resulting bound by $p_m^{\gamma_r}$, gives

$$p_m^{-\gamma_r} \|\check{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_\alpha\|_F^2 = O_P\left(\eta_{A_1, m} + \left\{\frac{p_m^{1-\gamma_r}}{q_m} + R(T, d_{r, \star})(1 \vee \nu_{r, m}) \max_{n \in [M] \setminus \{m\}} \nu_{r, n}\right\} \eta_\alpha\right),$$

which proves Theorem 5. The column-side result follows by symmetry.

Expanding $\eta_{A_1, m}$ and η_α gives the equivalent explicit rate:

$$\begin{aligned} p_m^{-\gamma_r} \|\check{\mathbf{A}}_{1m} - \check{\check{\mathbf{A}}}_{1m} \mathbf{H}_\alpha\|_F^2 &= O_P\left(R(T, d_{r, \star}) p_m^{\gamma_r^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1} p_m^{1-\gamma_r} q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r, \max}^{(n)} - \gamma_r} + \left\{p_m^{1-\gamma_r} q_m^{-1} + R(T, d_{r, \star}) p_m^{\gamma_{r, \max}^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r}\right\} \right. \\ &\quad \left. \cdot \max_{n \in [M]} \left\{p_n^{-\gamma_r} q_n^{-1} + R(T, d_{r, \star}) p_n^{\gamma_{r, \max}^{(n)} - \gamma_r} \max_{s \in [M] \setminus \{n\}} p_s^{\gamma_r^{(s)} - \gamma_r} + T^{-1} p_n^{1-\gamma_{r, \min}^{(n)}} q_n^{-1} \max_{s \in [M]} p_s^{\gamma_{r, \max}^{(s)} - \gamma_r}\right\}\right), \end{aligned}$$

$$\begin{aligned}
q_m^{-\gamma_c} \left\| \widetilde{\mathbf{B}}_{1m} - \widetilde{\widetilde{\mathbf{B}}}_{1m} \mathbf{H}_\beta \right\|_F^2 &= O_P \left(R(T, d_{c,\star}) q_m^{\gamma_c^{(m)} - \gamma_c} \max_{n \in [M] \setminus \{m\}} q_n^{\gamma_c^{(n)} - \gamma_c} \right. \\
&\quad + T^{-1} q_m^{1-\gamma_c} p_m^{-1} \max_{n \in [M] \setminus \{m\}} q_n^{\gamma_{c,\max}^{(n)} - \gamma_c} + \left\{ q_m^{1-\gamma_c} p_m^{-1} + R(T, d_{c,\star}) q_m^{\gamma_{c,\max}^{(m)} - \gamma_c} \max_{n \in [M] \setminus \{m\}} q_n^{\gamma_c^{(n)} - \gamma_c} \right\} \\
&\quad \cdot \max_{n \in [M]} \left\{ q_n^{-\gamma_c} p_n^{-1} + R(T, d_{c,\star}) q_n^{\gamma_{c,\max}^{(n)} - \gamma_c} \max_{s \in [M] \setminus \{n\}} q_s^{\gamma_c^{(s)} - \gamma_c} + T^{-1} q_n^{1-\gamma_{c,\min}^{(n)}} p_n^{-1} \max_{s \in [M]} q_s^{\gamma_{c,\max}^{(s)} - \gamma_c} \right\} \Big),
\end{aligned}$$

where the column-side results follow by symmetry. This completes the proof. $\square$

Proof of Corollary 6. The result follows by substituting $\gamma_r = \gamma_c = \gamma_r^{(m)} = \gamma_c^{(m)} = 1$ and $R(T, d_{r,\star}) = R(T, d_{c,\star}) = T^{-1}$ into Theorems 3-5. $\square$

Proof of Theorem 7. For the local factor loading estimators, it suffices to show the results for $\widehat{\mathbf{C}}_m$, and analogous arguments follow for $\widehat{\mathbf{R}}_m$. To this end, we adapt the decomposition and arguments used in the proof of Lemma 9, with the roles of rows and columns interchanged appropriately. Recall from (D.74) that $\check{\mathbf{Y}}_t^{(m)} = \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}$. Hence

$$\begin{aligned} \check{\mathbf{Y}}_t^{(m)'} \widehat{\mathbf{Q}}_m^\perp &= \mathbf{J}_m \mathbf{G}_t' \mathbf{Q}_m' \widehat{\mathbf{Q}}_m^\perp + \mathbf{C}_m \mathbf{F}_t^{(m)'} \mathbf{R}_m' \widehat{\mathbf{Q}}_m^\perp + \mathbf{M}_{q_m} \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m}' \widehat{\mathbf{Q}}_m^\perp \\ &=: \mathcal{A}_t + \mathcal{B}_t + \mathcal{C}_t. \end{aligned} \quad (\text{C.44})$$

With (C.44), we may expand $\mathbf{X}_{C,m}$ as

$$\begin{aligned} \mathbf{X}_{C,m} &= \frac{1}{T} \sum_{t=1}^T (\check{\mathbf{Y}}_t^{(m)'} \widehat{\mathbf{Q}}_m^\perp) (\check{\mathbf{Y}}_t^{(m)'} \widehat{\mathbf{Q}}_m^\perp)' \\ &= \frac{1}{T} \sum_{t=1}^T \mathcal{B}_t \mathcal{B}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{B}_t \mathcal{C}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{B}_t \mathcal{A}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{C}_t \mathcal{B}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{C}_t \mathcal{C}_t' \\ &\quad + \frac{1}{T} \sum_{t=1}^T \mathcal{C}_t \mathcal{A}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{A}_t \mathcal{B}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{A}_t \mathcal{C}_t' + \frac{1}{T} \sum_{t=1}^T \mathcal{A}_t \mathcal{A}_t' =: \sum_{i=1}^9 \mathcal{J}_i. \end{aligned} \quad (\text{C.45})$$

For convenience, define the matrices $\mathbf{P}_{Q,m}^\perp := \mathbf{Q}_m^\perp \mathbf{Q}_m^{\perp'}$ and $\widehat{\mathbf{P}}_{Q,m}^\perp := \widehat{\mathbf{Q}}_m^\perp \widehat{\mathbf{Q}}_m^{\perp'}$. Then we can write $\mathcal{B}_t \mathcal{B}_t' = \mathbf{C}_m \mathbf{F}_t^{(m)'} \mathbf{R}_m' \widehat{\mathbf{P}}_{Q,m}^\perp \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m'$. Further define

$$\mathcal{B}_t^0 := \mathbf{C}_m \mathbf{F}_t^{(m)'} \mathbf{R}_m' \mathbf{Q}_m^\perp, \quad \mathcal{J}_1^0 := \frac{1}{T} \sum_{t=1}^T \mathcal{B}_t^0 \mathcal{B}_t^{0'}. \quad (\text{C.46})$$

Then by Lemmas 4, 9, and following Lemma S5 in Cen (2026), we have

$$\begin{aligned} \left\| \widehat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp \right\|_F^2 &= \left\| \widehat{\mathbf{P}}_{Q,m} - \mathbf{P}_{Q,m} \right\|_F^2 \\ &= O_P \left\{ R(T, d_{\text{fac},*}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}. \end{aligned} \quad (\text{C.47})$$

Consider $\mathcal{J}_1$. Note that $\mathcal{J}_1 - \mathcal{J}_1^0 = T^{-1} \sum_{t=1}^T \mathbf{C}_m \mathbf{F}_t^{(m)'} \mathbf{R}_m' (\widehat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp) \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m'$, we have

$$\begin{aligned} \|\mathcal{J}_1 - \mathcal{J}_1^0\|_F^2 &\leq \frac{1}{T} \sum_{t=1}^T \left\| \mathbf{C}_m \mathbf{F}_t^{(m)'} \mathbf{R}_m' (\widehat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp) \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \right\|_F^2 \\ &\lesssim \|\mathbf{C}_m\|^4 \|\mathbf{R}_m\|^4 \left\| \widehat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp \right\|_F^2 \cdot \frac{1}{T} \sum_{t=1}^T \|\mathbf{F}_t^{(m)}\|_F^4, \end{aligned} \quad (\text{C.48})$$

where the last inequality used the sub-multiplicativity and the fact that the Frobenius norm of a product is bounded by products of spectral and Frobenius norms. By Assumption (F1), $\mathbf{F}_t^{(m)}$ has uniformly bounded

fourth moments, so $T^{-1} \sum_{t=1}^T \|\mathbf{F}_t^{(m)}\|_F^4 = O_P(1)$. Combining (C.48), (C.47), and Assumption (L1), we thus have

$$\|\mathcal{J}_1 - \mathcal{J}_1^0\|_F^2 \lesssim_P q_m^{2\delta_{c,1}^{(m)}} p_m^{2\delta_{r,1}^{(m)}} \left\{ R(T, d_{\text{fac}, \star}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}. \quad (\text{C.49})$$

Next, by Assumptions (L1) and (S1), the nonzero eigenvalues of $\mathcal{J}_1^0$ satisfy that, for any $i \in [k_c^{(m)}]$,

$$\lambda_i(\mathcal{J}_1^0) \asymp q_m^{\delta_{c,i}^{(m)}} p_m^{\delta_{r,1}^{(m)}}, \quad (\text{C.50})$$

which follows analogous arguments as for (D.97) in the proof of Lemma 9.

It remains to bound $\mathcal{J}_i$, $i = 2, \dots, 9$. Note that these terms all involve at least one of $\mathcal{A}_t$, $\mathcal{C}_t$, or the projection perturbation (cf. (C.47)). We bound their squared Frobenius norms separately. To this end, for $\mathcal{J}_2$ and $\mathcal{J}_4$, by Assumptions (E2), (F1), and the same calculation as in (D.86), we have

$$\|\mathcal{J}_2\|_F^2, \|\mathcal{J}_4\|_F^2 = \left\| \frac{1}{T} \sum_{t=1}^T \mathcal{B}_t \mathcal{C}_t' \right\|_F^2 \lesssim_P \frac{1}{T} q_m^{\delta_{c,1}^{(m)}} p_m^{\delta_{r,1}^{(m)}} q_m p_m = \frac{1}{T} q_m^{1+\delta_{c,1}^{(m)}} p_m^{1+\delta_{r,1}^{(m)}}. \quad (\text{C.51})$$

For $\mathcal{J}_5$, using Lemma 3 and the proof of (D.35), it holds that

$$\|\mathcal{J}_5\|_F^2 \lesssim_P q_m^{2-2\delta_{c,1}^{(m)}} + \frac{1}{T} q_m^2 p_m. \quad (\text{C.52})$$

For $\mathcal{J}_6$ and $\mathcal{J}_8$, it suffices to consider the former and the latter follows in similar arguments. Recall that $\mathcal{A}_t = \mathbf{J}_m \mathbf{G}_t' \mathbf{Q}_m' \hat{\mathbf{Q}}_m^\perp$ from (C.44). Note that

$$\begin{aligned} \|\mathbf{Q}_m' \hat{\mathbf{Q}}_m^\perp\|_F &= \|\mathbf{Q}_m' (\hat{\mathbf{Q}}_m^\perp \hat{\mathbf{Q}}_m^{\perp'} - \mathbf{Q}_m^\perp \mathbf{Q}_m^{\perp'}) \hat{\mathbf{Q}}_m^\perp\| \\ &\leq \|\mathbf{Q}_m\| \cdot \|\hat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp\|_F \cdot \|\hat{\mathbf{Q}}_m^\perp\| = \|\mathbf{Q}_m\| \cdot \|\hat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp\|_F, \end{aligned} \quad (\text{C.53})$$

Also, with $f(\mathbf{X}) = \|\mathbf{X}\|_F^2$ being convex, we have $\|T^{-1} \sum_{t=1}^T \mathbf{X}_t\|_F^2 = f(T^{-1} \sum_{t=1}^T \mathbf{X}_t) \leq T^{-1} \sum_{t=1}^T f(\mathbf{X}_t) = T^{-1} \sum_{t=1}^T \|\mathbf{X}_t\|_F^2$. Altogether, by Assumption (F1), and the sub-multiplicativity of the Frobenius norm, we arrive at

$$\begin{aligned} \|\mathcal{J}_6\|_F^2 &= \left\| \frac{1}{T} \sum_{t=1}^T \mathcal{C}_t \mathcal{A}_t' \right\|_F^2 \leq \frac{1}{T} \sum_{t=1}^T \|\mathcal{C}_t\|_F^2 \|\mathcal{A}_t\|_F^2 \\ &\leq \frac{1}{T} \sum_{t=1}^T \|\mathcal{C}_t\|_F^2 \cdot \|\mathbf{J}_m\|^2 \cdot \|\mathbf{G}_t\|_F^2 \cdot \|\mathbf{Q}_m\|_F^2 \cdot \|\hat{\mathbf{P}}_{Q,m}^\perp - \mathbf{P}_{Q,m}^\perp\|_F^2 \end{aligned}$$

$$\lesssim_P q_m^{1+\delta_{j,1}^{(m)}} p_m^{1+\delta_{q,1}^{(m)}} \left\{ R(T, d_{\text{fac}, \star}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}, \quad (\text{C.54})$$

where the second inequality used (C.47), (C.53), and $\|\mathcal{C}_t\|_F^2 = O_P(p_m q_m)$ which follows from similar arguments for (D.81). Next, $\mathcal{J}_3$ and $\mathcal{J}_7$ can be bounded analogously by the above arguments, with $\mathcal{C}_t$ replaced by $\mathcal{B}_t$ where $\|\mathcal{B}_t\|_F^2 = O_P(q_m^{\delta_{c,1}^{(m)}} p_m^{\delta_{r,1}^{(m)}})$, thus giving

$$\|\mathcal{J}_3\|_F^2, \|\mathcal{J}_7\|_F^2 \lesssim_P q_m^{\delta_{c,1}^{(m)} + \delta_{j,1}^{(m)}} p_m^{\delta_{r,1}^{(m)} + \delta_{q,1}^{(m)}} \left\{ R(T, d_{\text{fac}, \star}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}. \quad (\text{C.55})$$

For $\mathcal{J}_9$, using Assumption (F1), the bound of $\|\mathcal{A}_t\|_F^2$ as in (C.53) and (C.54), we have

$$\|\mathcal{J}_9\|_F^2 \leq \frac{1}{T} \sum_{t=1}^T \|\mathcal{A}_t\|_F^4 \lesssim_P q_m^{2\delta_{j,1}^{(m)}} p_m^{2\delta_{q,1}^{(m)}} \left\{ R(T, d_{\text{fac}, \star}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}^2. \quad (\text{C.56})$$

Finally, recall (C.46). Then combining (C.49), (C.51), (C.52), (C.54), (C.55), and (C.56) give the rate on $\|\mathbf{X}_{C,m} - \mathcal{J}_1^0\|_F^2$, which, together with (C.50) and applying the Davis–Kahan theorem (Yu et al., 2015, Theorem 2), complete the proof for $\|\hat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m}\|_F^2$.

It remains to show the results for the local factor estimators. To this end, define

$$\mathcal{L}_{R,m} := (\check{\mathbf{R}}'_m \mathbf{P}_{Q,m}^\perp \check{\mathbf{R}}_m)^{-1} \check{\mathbf{R}}'_m \mathbf{P}_{Q,m}^\perp, \quad \hat{\mathcal{L}}_{R,m} := (\hat{\mathbf{R}}'_m \hat{\mathbf{P}}_{Q,m}^\perp \hat{\mathbf{R}}_m)^{-1} \hat{\mathbf{R}}'_m \hat{\mathbf{P}}_{Q,m}^\perp.$$

By Assumptions (L1) and (S1), the matrix $\check{\mathbf{R}}'_m \mathbf{P}_{Q,m}^\perp \check{\mathbf{R}}_m$ is positive definite with eigenvalues bounded away from zero, and it holds that

$$\|\mathcal{L}_{R,m}\|, \|\hat{\mathcal{L}}_{R,m}\| = O_P(p_m^{-\delta_{r,1}^{(m)}/2}), \quad \mathcal{L}_{R,m} \check{\mathbf{R}}_m = \mathbf{I}, \quad \mathcal{L}_{R,m} \mathbf{Q}_m = \mathbf{0}. \quad (\text{C.57})$$

Then by the definition of $\hat{\mathbf{F}}_t^{(m)}$ in Step 3 of Section 3.3 and (D.74), we may decompose

$$\begin{aligned} \hat{\mathbf{F}}_t^{(m)} &= \hat{\mathcal{L}}_{R,m} \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{C}}_m \\ &= \hat{\mathcal{L}}_{R,m} \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m \hat{\mathbf{C}}_m + \hat{\mathcal{L}}_{R,m} \check{\mathbf{R}}_m \check{\mathbf{F}}_t^{(m)} \check{\mathbf{C}}'_m \hat{\mathbf{C}}_m + \hat{\mathcal{L}}_{R,m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{C}}_m =: \mathcal{K}_1 + \mathcal{K}_2 + \mathcal{K}_3. \end{aligned} \quad (\text{C.58})$$

For $\mathcal{K}_1$, by (C.57), (C.53), and Assumption (F1), we have

$$\begin{aligned}\|\mathcal{K}_1\|_F^2 &\leq \|\widehat{\mathcal{L}}_{R,m}\|^2 \|\mathbf{Q}'_m \widehat{\mathbf{P}}_{Q,m}^\perp\|_F^2 \|\mathbf{G}_t\|_F^2 \|\mathbf{J}_m\|^2 \|\widehat{\mathbf{C}}_m\|^2 \\ &\lesssim_P p_m^{\delta_{q,1}^{(m)} - \delta_{r,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \eta_{Q,m} \|\mathbf{G}_t\|_F^2 \lesssim_P p_m^{\delta_{q,1}^{(m)} - \delta_{r,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \eta_{Q,m},\end{aligned}\quad (\text{C.59})$$

with $\eta_{Q,m}$ defined in the statement of Theorem 7.

Consider $\mathcal{K}_2$ next. Note that we can write

$$\begin{aligned}\mathcal{K}_2 &= (\mathbf{H}_{R,m}^{-1} + \widehat{\mathcal{L}}_{R,m} \check{\mathbf{R}}_m - \mathbf{H}_{R,m}^{-1}) \check{\mathbf{F}}_t^{(m)} (\mathbf{H}_{C,m} + \check{\mathbf{C}}'_m \widehat{\mathbf{C}}_m - \mathbf{H}_{C,m}) \\ &= \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m} + (\widehat{\mathcal{L}}_{R,m} \check{\mathbf{R}}_m - \mathbf{H}_{R,m}^{-1}) \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m} \\ &\quad + \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} (\check{\mathbf{C}}'_m \widehat{\mathbf{C}}_m - \mathbf{H}_{C,m}) + (\widehat{\mathcal{L}}_{R,m} \check{\mathbf{R}}_m - \mathbf{H}_{R,m}^{-1}) \check{\mathbf{F}}_t^{(m)} (\check{\mathbf{C}}'_m \widehat{\mathbf{C}}_m - \mathbf{H}_{C,m}).\end{aligned}\quad (\text{C.60})$$

First, by the rates of convergence of the local loading matrices, we have

$$\begin{aligned}\|\widehat{\mathcal{L}}_{R,m} \check{\mathbf{R}}_m - \mathbf{H}_{R,m}^{-1}\|_F^2 &= \|(\widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m)^{-1} \widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \check{\mathbf{R}}_m - \mathbf{H}_{R,m}^{-1}\|_F^2 \\ &= \|(\widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m)^{-1} \widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \check{\mathbf{R}}_m \mathbf{H}_{R,m} - \mathbf{I}\|_F^2 \\ &= \|(\widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m)^{-1} \widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp (\check{\mathbf{R}}_m \mathbf{H}_{R,m} - \widehat{\mathbf{R}}_m) + (\widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m)^{-1} \widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m - \mathbf{I}\|_F^2 \\ &\leq \|(\widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp \widehat{\mathbf{R}}_m)^{-1} \widehat{\mathbf{R}}'_m \widehat{\mathbf{P}}_{Q,m}^\perp\|^2 \cdot \|\check{\mathbf{R}}_m \mathbf{H}_{R,m} - \widehat{\mathbf{R}}_m\|_F^2 \lesssim_P \Delta_{R,m},\end{aligned}\quad (\text{C.61})$$

where $\Delta_{R,m}$ is defined in the statement of Theorem 7. Similarly, we have

$$\begin{aligned}\|\check{\mathbf{C}}'_m \widehat{\mathbf{C}}_m - \mathbf{H}_{C,m}\|_F^2 &= \|\check{\mathbf{C}}'_m (\widehat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m}) + \check{\mathbf{C}}'_m \check{\mathbf{C}}_m \mathbf{H}_{C,m} - \mathbf{H}_{C,m}\|_F^2 \\ &= \|\check{\mathbf{C}}'_m (\widehat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m}) + (\mathbf{C}'_m \mathbf{C}_m)^{-1/2} \mathbf{C}'_m \mathbf{C}_m (\mathbf{C}'_m \mathbf{C}_m)^{-1/2} \mathbf{H}_{C,m} - \mathbf{H}_{C,m}\|_F^2 \lesssim_P \Delta_{C,m},\end{aligned}\quad (\text{C.62})$$

Also, recall the definition $\check{\mathbf{F}}_t^{(m)} = (\mathbf{R}'_m \mathbf{R}_m)^{1/2} \mathbf{F}_t^{(m)} (\mathbf{C}'_m \mathbf{C}_m)^{1/2}$, and it is straightforward from Assumptions (F1) and (L1) that $\|\check{\mathbf{F}}_t^{(m)}\|_F^2 = O_P(p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}})$, which, combined with (C.60), (C.61), (C.62), give

$$\|\mathcal{K}_2 - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}\|_F^2 \lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m}).\quad (\text{C.63})$$

Finally, for $\mathcal{K}_3$, by (C.57) and the same argument as for (D.81),

$$\begin{aligned}\|\mathcal{K}_3\|_F^2 &= \|\widehat{\mathcal{L}}_{R,m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \widehat{\mathbf{C}}_m\|_F^2 \\ &\leq \|\widehat{\mathcal{L}}_{R,m}\|^2 \cdot \|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \{\check{\mathbf{C}}_m \mathbf{H}_{C,m} + (\widehat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m})\}\|_F^2 \\ &\lesssim \|\widehat{\mathcal{L}}_{R,m}\|^2 \cdot \|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \check{\mathbf{C}}_m \mathbf{H}_{C,m}\|_F^2 + \|\widehat{\mathcal{L}}_{R,m}\|^2 \cdot \|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} (\widehat{\mathbf{C}}_m - \check{\mathbf{C}}_m \mathbf{H}_{C,m})\|_F^2\end{aligned}$$

$$\lesssim_P p_m^{1-\delta_{r,1}^{(m)}} + p_m^{1-\delta_{r,1}^{(m)}} q_m \Delta_{C,m}. \quad (\text{C.64})$$

By combining (C.58), (C.59), (C.63), and (C.64), we conclude that

$$\frac{1}{\frac{\delta_{r,1}^{(m)}}{p_m} \frac{\delta_{c,1}^{(m)}}{q_m}} \left\| \widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m} \right\|_F^2 \lesssim_P \Delta_{R,m} + \Delta_{C,m} + p_m^{1-2\delta_{r,1}^{(m)}} q_m^{1-\delta_{c,1}^{(m)}} \Delta_{C,m} + p_m^{1-2\delta_{r,1}^{(m)}} q_m^{-\delta_{c,1}^{(m)}},$$

where the rate in (C.59), after scaled by $p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}}$, is omitted as it is dominated by $\Delta_{C,m}$, particularly:

$$q_m^{-2\delta_{c,k_c^{(m)}}^{(m)}} p_m^{-2\delta_{r,1}^{(m)}} \left(q_m^{2\delta_{c,1}^{(m)}} p_m^{2\delta_{r,1}^{(m)}} + q_m^{1+\delta_{j,1}^{(m)}} p_m^{1+\delta_{q,1}^{(m)}} \right) \eta_{Q,m}.$$

Noting that all these arguments hold if we alternatively work on $\widehat{\mathbf{F}}_t^{(m) \prime}$, with the resulting rates of convergence symmetric (i.e., with row and column dimensions swapped), and taking the minimum of these two completes the proof of this theorem. $\square$

Proof of Theorem 8. Recall $\check{\mathbf{G}}_{t,m}$ from Lemma 10 and we can write

$$\begin{aligned} \check{\mathbf{G}}_{t,m} &= \check{\mathbf{Q}}'_m \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m \check{\mathbf{J}}_m \\ &= (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{Q}'_m \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m \mathbf{J}_m (\mathbf{J}'_m \mathbf{J}_m)^{-1/2} = (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \mathbf{G}_t (\mathbf{J}'_m \mathbf{J}_m)^{1/2}, \end{aligned}$$

so that for the reference group m_0 , we have $\check{\mathbf{G}}_{t,m_0} = \check{\mathbf{G}}_t$, where $\check{\mathbf{G}}_t = (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} \mathbf{G}_t (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2}$ as defined in the statement of Theorem 8.

For each group $m \in [M]$ and $t \in [T]$, define the population quantities

$$\begin{aligned} \Xi_{Q,m}^* &:= \mathbf{H}'_{Q,m_0} (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{H}_{Q,m}, \\ \Xi_{J,m}^* &:= \mathbf{H}'_{J,m_0} (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2} (\mathbf{J}'_m \mathbf{J}_m)^{-1/2} \mathbf{H}_{J,m}, \\ \mathbf{G}_{t,m}^* &:= \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \mathbf{G}_t (\mathbf{J}'_m \mathbf{J}_m)^{1/2} \mathbf{H}_{J,m}. \end{aligned}$$

First, note that for all $t \in [T]$,

$$\text{vec}(\mathbf{G}_{t,m_0}^*) = (\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) \text{vec}(\mathbf{G}_{t,m}^*) + o_P \left(p_m^{\delta_{q,1}^{(m)}/2} q_m^{\delta_{j,1}^{(m)}/2} \right). \quad (\text{C.65})$$

As $\mathbf{H}_{Q,m}$ and $\mathbf{H}_{J,m}$ are asymptotically orthogonal by Lemmas 9 and 12, it also holds that

$$\begin{aligned} \Xi_{Q,m}^* \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} &= \mathbf{H}'_{Q,m_0} (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{H}_{Q,m} \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \\ &= \mathbf{H}'_{Q,m_0} (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} + O_P \left(\eta_{Q,m}^{1/2} p_m^{\delta_{q,1}^{(m_0)}/2} \right), \end{aligned}$$

$$\begin{aligned}\Xi_{J,m}^* \mathbf{H}'_{J,m} (\mathbf{J}'_m \mathbf{J}_m)^{1/2} &= \mathbf{H}'_{J,m_0} (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2} (\mathbf{J}'_m \mathbf{J}_m)^{-1/2} \mathbf{H}_{J,m} \mathbf{H}'_{J,m} (\mathbf{J}'_m \mathbf{J}_m)^{1/2} \\ &= \mathbf{H}'_{J,m_0} (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2} + O_P \left(\eta_{J,m}^{1/2} q_{m_0}^{\delta_{j,1}^{(m_0)}/2} \right).\end{aligned}$$

so that together with $\lambda_{\min}(\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0}) \lambda_{\min}(\mathbf{J}'_{m_0} \mathbf{J}_{m_0}) = O\{\lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m)\}$ from the statement of Theorem 8, we arrive at

$$\begin{aligned}\Xi_{Q,m}^* \mathbf{G}_{t,m}^* \Xi_{J,m}^{*'} &= \Xi_{Q,m}^* \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \mathbf{G}_t (\mathbf{J}'_m \mathbf{J}_m)^{1/2} \mathbf{H}_{J,m} \Xi_{J,m}^{*'} \\ &= \mathbf{H}'_{Q,m_0} (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{1/2} \mathbf{G}_t (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{1/2} \mathbf{H}_{J,m_0} + O_P \left(\eta_{Q,m}^{1/2} \eta_{J,m}^{1/2} p_m^{\delta_{q,1}^{(m)}/2} q_m^{\delta_{j,1}^{(m)}/2} \right) \\ &= \mathbf{G}_{t,m_0}^* + O_P \left(\eta_{Q,m}^{1/2} \eta_{J,m}^{1/2} p_m^{\delta_{q,1}^{(m)}/2} q_m^{\delta_{j,1}^{(m)}/2} \right) = \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} + O_P \left(\eta_{Q,m}^{1/2} \eta_{J,m}^{1/2} p_m^{\delta_{q,1}^{(m)}/2} q_m^{\delta_{j,1}^{(m)}/2} \right).\end{aligned}\quad (\text{C.66})$$

Also by Lemma 10, we have

$$\hat{\mathbf{G}}_{t,m} = \mathbf{G}_{t,m}^* + O_P(\Delta_{G,m}^{1/2}) = \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \mathbf{G}_t (\mathbf{J}'_m \mathbf{J}_m)^{1/2} \mathbf{H}_{J,m} + O_P(\Delta_{G,m}^{1/2}), \quad (\text{C.67})$$

where $O_P(\cdot)$ is in terms of Frobenius norm, and $\Delta_{G,m}$ is defined in the statement of Theorem 8 such that

$$\Delta_{G,m} = \min(p_m, q_m) + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + (\eta_{Q,m} + \eta_{J,m}) p_m q_m.$$

By the construction of $\hat{\mathbf{Z}}_{G,m}$ in Step 3 of Section 3.3 and with $\hat{\mathbf{G}}_{t,m}$ defined in Lemma 10, note that

$$\hat{\mathbf{Z}}_{G,m} = \left(\text{vec}(\hat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,1}^{(m)} \hat{\mathbf{J}}_m), \dots, \text{vec}(\hat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,T}^{(m)} \hat{\mathbf{J}}_m) \right) = \left(\text{vec}(\hat{\mathbf{G}}_{1,m}), \dots, \text{vec}(\hat{\mathbf{G}}_{T,m}) \right).$$

Then we may define the population version of $\hat{\mathbf{Z}}_{G,m}$ as $\mathbf{Z}_{G,m}^* := \left(\text{vec}(\mathbf{G}_{1,m}^*), \dots, \text{vec}(\mathbf{G}_{T,m}^*) \right)$, so that by (C.67), we have

$$\|\hat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*\|_F^2 \lesssim_P T \cdot \Delta_{G,m}. \quad (\text{C.68})$$

Next, recall the matrix $\hat{\Xi}_m = \hat{\mathbf{Z}}_{G,m_0} \hat{\mathbf{Z}}'_{G,m} (\hat{\mathbf{Z}}_{G,m} \hat{\mathbf{Z}}'_{G,m})^{-1}$ for the two-sided alignment in Step 3 of Section 3.3. For any group $m \neq m_0$, the population version of $\hat{\Xi}_m$ is $\Xi_m^* := \mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1}$. With (C.65), we can write

$$\mathbf{Z}_{G,m_0}^* = \left(\text{vec}(\mathbf{G}_{1,m_0}^*), \dots, \text{vec}(\mathbf{G}_{T,m_0}^*) \right) = (\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) \mathbf{Z}_{G,m}^* + o_P \left(\|\mathbf{Z}_{G,m}^*\|_F \right),$$

which implies $\Xi_m^* = \mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} = \Xi_{J,m}^* \otimes \Xi_{Q,m}^* + o_P(1)$ given Assumption (S1). Meanwhile,

also by Assumption (S1), we have

$$\left\| (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} \right\| = O_P \left(\frac{1}{TS_m} \right), \quad (\text{C.69})$$

where S_m is defined in the main text. Hence we may expand

$$\begin{aligned} \widehat{\Xi}_m &= \{ \mathbf{Z}_{G,m_0}^* + (\widehat{\mathbf{Z}}_{G,m_0} - \mathbf{Z}_{G,m_0}^*) \} \{ \mathbf{Z}_{G,m}^* + (\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*) \}' \\ &\quad \cdot \left[\{ \mathbf{Z}_{G,m}^* + (\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*) \} \{ \mathbf{Z}_{G,m}^* + (\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*) \}' \right]^{-1} \\ &=: \mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} + \mathcal{R}_{\Xi,m}, \end{aligned} \quad (\text{C.70})$$

where $\mathcal{R}_{\Xi,m}$ denotes all remaining terms in the decomposition and would involve at least one $(\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*)$ or $(\widehat{\mathbf{Z}}_{G,m_0} - \mathbf{Z}_{G,m_0}^*)$. As detailed, we bound all terms in $\mathcal{R}_{\Xi,m}$. To this end, by (C.68) and (C.69), we first have

$$\begin{aligned} &\left\| (\widehat{\mathbf{Z}}_{G,m_0} - \mathbf{Z}_{G,m_0}^*) \mathbf{Z}_{G,m}^{*'} (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} \right\|_F^2 \lesssim_P \left\| \widehat{\mathbf{Z}}_{G,m_0} - \mathbf{Z}_{G,m_0}^* \right\|_F^2 \cdot \left\| \mathbf{Z}_{G,m}^{*'} \right\|_F^2 \cdot \left\| (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} \right\|^2 \\ &= O_P(T \cdot \Delta_{G,m_0}) \cdot O_P(TS_m) \cdot O_P((TS_m)^{-2}) = O_P \left(\frac{\Delta_{G,m_0}}{S_m} \right), \end{aligned} \quad (\text{C.71})$$

where $\left\| \mathbf{Z}_{G,m}^{*'} \right\|_F^2 = O_P(TS_m)$ used Assumption (F1). Similarly,

$$\begin{aligned} &\left\| \mathbf{Z}_{G,m_0}^* (\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*)' (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} \right\|_F^2 \lesssim_P \left\| \mathbf{Z}_{G,m_0}^* \right\|^2 \cdot \left\| \widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^* \right\|_F^2 \cdot \left\| (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} \right\|^2 \\ &= O_P(TS_{m_0}) \cdot O_P(T \cdot \Delta_{G,m}) \cdot O_P((TS_m)^{-2}) = O_P \left(\frac{\Delta_{G,m} S_{m_0}}{S_m^2} \right). \end{aligned} \quad (\text{C.72})$$

To bound the cross term in $\mathcal{R}_{\Xi,m}$ involving both $\widehat{\mathbf{Z}}_{G,m_0} - \mathbf{Z}_{G,m_0}^*$ and $\widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*$, let $\mathbf{E}_{G,m} := \widehat{\mathbf{Z}}_{G,m} - \mathbf{Z}_{G,m}^*$. By the condition $\Delta_{G,m} = o\{\lambda_{\min}(\mathbf{Q}_m' \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}_m' \mathbf{J}_m)\}$ in the statement of Theorem 8, we have $\Delta_{G,m} = o(S_m)$. Then the cross term satisfies

$$\begin{aligned} &\left\| \mathbf{E}_{G,m_0} \mathbf{E}_{G,m}' \left[(\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m}) (\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m})' \right]^{-1} \right\|_F^2 \\ &\lesssim_P \left\| \mathbf{E}_{G,m_0} \right\|_F^2 \left\| \mathbf{E}_{G,m} \right\|_F^2 \left\| \left[(\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m}) (\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m})' \right]^{-1} \right\|^2 \\ &\lesssim_P (T \Delta_{G,m_0}) (T \Delta_{G,m}) (T^{-2} S_m^{-2}) = \frac{\Delta_{G,m_0} \Delta_{G,m}}{S_m^2}, \end{aligned} \quad (\text{C.73})$$

where we used $\left\| \left[(\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m}) (\mathbf{Z}_{G,m}^* + \mathbf{E}_{G,m})' \right]^{-1} \right\|^2 \lesssim_P T^{-2} S_m^{-2}$, which follows from applying Weyl's inequality on $\lambda_{\min}(A_m) = O_P(TS_m)$ (as in (C.74) below), combined with (C.75) and $\Delta_{G,m} = o(S_m)$.

Finally, we bound

$$\mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} \left\{ (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'})^{-1} - (\mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'} + B_m)^{-1} \right\} = \mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} \left\{ A_m^{-1} - (A_m + B_m)^{-1} \right\}, \quad (\text{C.74})$$

where $A_m := \mathbf{Z}_{G,m}^* \mathbf{Z}_{G,m}^{*'}$ and $B_m := \mathbf{E}_{G,m} \mathbf{Z}_{G,m}^{*'} + \mathbf{Z}_{G,m}^* \mathbf{E}_{G,m}' + \mathbf{E}_{G,m} \mathbf{E}_{G,m}'$. Note that $\lambda_{\min}(A_m) = O_P(TS_m)$ by Assumption (S1) and the definition of S_m . Also, by the condition $\Delta_{G,m} = o(S_m)$ from the statement of Theorem 8 and (C.68), we have

$$\|B_m\| \lesssim_P \|\mathbf{E}_{G,m}\|_F \|\mathbf{Z}_{G,m}^*\| + \|\mathbf{E}_{G,m}\|_F^2 = O_P\left(\sqrt{T\Delta_{G,m}} \cdot \sqrt{TS_m}\right) + O_P(T\Delta_{G,m}) = O_P\left(T\sqrt{\Delta_{G,m}S_m}\right), \quad (\text{C.75})$$

where we used $\|\mathbf{Z}_{G,m}^*\|^2 \asymp_P TS_m$ under Assumptions (S1) and (F1).

Using the matrix identity $\mathbf{A}^{-1} - (\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1}\mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}$ for any matrices $\mathbf{A}, \mathbf{B}$, we obtain

$$\begin{aligned} \left\| A_m^{-1} - (A_m + B_m)^{-1} \right\| &\leq \|A_m^{-1}\| \|B_m\| \|(A_m + B_m)^{-1}\| \\ &= O_P\left(T^{-1}S_m^{-1}T\sqrt{\Delta_{G,m}S_m}T^{-1}S_m^{-1}\right) = O_P\left(T^{-1}S_m^{-3/2}\Delta_{G,m}^{1/2}\right). \end{aligned}$$

Hence the corresponding term in $\mathcal{R}_{\Xi,m}$ from (C.74) is bounded by

$$\begin{aligned} \left\| \mathbf{Z}_{G,m_0}^* \mathbf{Z}_{G,m}^{*'} \left[A_m^{-1} - (A_m + B_m)^{-1} \right] \right\|_F^2 &\lesssim \|\mathbf{Z}_{G,m_0}^*\|^2 \|\mathbf{Z}_{G,m}^*\|^2 \left\| A_m^{-1} - (A_m + B_m)^{-1} \right\|^2 \\ &= O_P\left(TS_{m_0}TS_mT^{-2}S_m^{-3}\Delta_{G,m}\right) = O_P\left(\frac{S_{m_0}}{S_m^2}\Delta_{G,m}\right) = O_P\left(\frac{\Delta_{G,m}}{S_m}\right), \end{aligned} \quad (\text{C.76})$$

where the last equality used $\lambda_{\min}(\mathbf{Q}'_{m_0}\mathbf{Q}_{m_0})\lambda_{\min}(\mathbf{J}'_{m_0}\mathbf{J}_{m_0}) = O\{\lambda_{\min}(\mathbf{Q}'_m\mathbf{Q}_m)\lambda_{\min}(\mathbf{J}'_m\mathbf{J}_m)\}$ so that $S_{m_0} = O(S_m)$ from the statement of Theorem 8.

Combining (C.71), (C.72), (C.73), and (C.76), we conclude $\|\mathcal{R}_{\Xi,m}\|_F^2 = O_P(\Delta_{G,m_0}/S_m + \Delta_{G,m}/S_m)$, and thus

$$\|\widehat{\Xi}_m - \Xi_m^*\|_F^2 \lesssim_P \max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}}. \quad (\text{C.77})$$

Next, recall from Step 3 of Section 3.3 the definition of the notation $\mathcal{R}(\mathbf{W})$ for a block matrix $\mathbf{W} = (\mathbf{W}_{ij})_{i \in [k_c], j \in [k_c]}$, with each $\mathbf{W}_{ij}$ of size $k_r \times k_r$, such that

$$\mathcal{R}(\mathbf{W}) = (\text{vec}(\mathbf{W}_{11}), \dots, \text{vec}(\mathbf{W}_{k_c 1}), \dots, \text{vec}(\mathbf{W}_{1k_c}), \dots, \text{vec}(\mathbf{W}_{k_c k_c})).$$

Note that $\mathcal{R}(\cdot)$ is a fixed permutation of the entries of the matrix input and hence preserves the Frobenius

norm. Thus, $\|\mathcal{R}(\widehat{\Xi}_m) - \mathcal{R}(\Xi_m^*)\|_F = \|\widehat{\Xi}_m - \Xi_m^*\|_F$, and hence by (C.77),

$$\|\mathcal{R}(\widehat{\Xi}_m) - \mathcal{R}(\Xi_m^*)\|_F^2 = O_P \left(\max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}} \right). \quad (\text{C.78})$$

By construction in Step 3 of Section 3.3, the matrix $\widehat{\sigma}_{\Xi,m} \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m}$ is the rank-1 singular value decomposition of $\mathcal{R}(\widehat{\Xi}_m)$. Then let $\sigma_{\Xi,m}^* \mathbf{u}_{\Xi,m}^* \mathbf{v}_{\Xi,m}^{*'} be the rank-1 singular value decomposition of $\mathcal{R}(\Xi_m^*)$. From Lemma 11, we have$

$$\mathcal{R}(\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}) = \widehat{\sigma}_{\Xi,m} \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m}.$$

Recall previously that $\Xi_m^* = \Xi_{J,m}^* \otimes \Xi_{Q,m}^*$, and by analogous arguments in the proof of Lemma 11, we also have

$$\sigma_{\Xi,m}^* \mathbf{u}_{\Xi,m}^* \mathbf{v}_{\Xi,m}^{*'} = \mathcal{R}(\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) = \text{vec}(\Xi_{Q,m}^*) \text{vec}(\Xi_{J,m}^*)',$$

which implies

$$\sigma_{\Xi,m}^* = \|\Xi_{Q,m}^*\|_F \cdot \|\Xi_{J,m}^*\|_F = O_P(1), \quad (\text{C.79})$$

where we used the definitions of $\Xi_{Q,m}^*$ and $\Xi_{J,m}^*$ at the beginning of this proof, together with the rate condition $\lambda_{\min}(\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0}) \lambda_{\min}(\mathbf{J}'_{m_0} \mathbf{J}_{m_0}) = O\{\lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m)\}$ in the statement of Theorem 8. Altogether, using the fact that $\mathcal{R}(\cdot)$ preserves the Frobenius norm (as in (C.78)), we have

$$\begin{aligned} & \left\| \widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m} - \Xi_{J,m}^* \otimes \Xi_{Q,m}^* \right\|_F^2 = \left\| \mathcal{R}(\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}) - \mathcal{R}(\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) \right\|_F^2 \\ &= \left\| \widehat{\sigma}_{\Xi,m} \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m} - \sigma_{\Xi,m}^* \mathbf{u}_{\Xi,m}^* \mathbf{v}_{\Xi,m}^{*'} \right\|_F^2 \\ &\lesssim |\widehat{\sigma}_{\Xi,m} - \sigma_{\Xi,m}^*|^2 \cdot \left\| \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m} \right\|_F^2 + (\sigma_{\Xi,m}^*)^2 \left\| \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m} - \mathbf{u}_{\Xi,m}^* \mathbf{v}_{\Xi,m}^{*'} \right\|_F^2 \\ &\lesssim_P \left\| \mathcal{R}(\widehat{\Xi}_m) - \mathcal{R}(\Xi_m^*) \right\|_F^2 = \left\| \widehat{\Xi}_m - \Xi_m^* \right\|_F^2 = O_P \left(\max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}} \right), \end{aligned} \quad (\text{C.80})$$

where the last line used Weyl's inequality and the Davis–Kahan theorem (Yu et al., 2015, Theorem 2), in addition to (C.78) and (C.79).

By the construction of the global factor estimator $\widehat{\mathbf{G}}_t$ in Step 3 of Section 3.3, we have

$$\left\| \widehat{\mathbf{G}}_t - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} \right\|_F^2 \lesssim \max_{m \in [M]} \left\| \widehat{\Xi}_{Q,m} \widehat{\mathbf{G}}_{t,m} \widehat{\Xi}'_{J,m} - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} \right\|_F^2. \quad (\text{C.81})$$

For each fixed m , we decompose

$$\widehat{\Xi}_{Q,m} \widehat{\mathbf{G}}_{t,m} \widehat{\Xi}'_{J,m} - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} = \widehat{\Xi}_{Q,m} (\widehat{\mathbf{G}}_{t,m} - \mathbf{G}_{t,m}^*) \widehat{\Xi}'_{J,m} + (\widehat{\Xi}_{Q,m} \mathbf{G}_{t,m}^* \widehat{\Xi}'_{J,m} - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0})$$

$$=: \mathcal{I}_{\Xi,G,1} + \mathcal{I}_{\Xi,G,2}. \quad (\text{C.82})$$

For $\mathcal{I}_{\Xi,G,1}$, by (C.79) and (C.67), we have $\|\widehat{\Xi}_{Q,m}\|, \|\widehat{\Xi}_{J,m}\| = O_P(1)$, and $\|\widehat{\mathbf{G}}_{t,m} - \mathbf{G}_{t,m}^*\|_F^2 = O_P(\Delta_{G,m})$, so that

$$\|\mathcal{I}_{\Xi,G,1}\|_F^2 \lesssim_P \Delta_{G,m}. \quad (\text{C.83})$$

For $\mathcal{I}_{\Xi,G,2}$, we can write

$$\begin{aligned} \text{vec}(\mathcal{I}_{\Xi,G,2}) &= (\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}) \text{vec}(\mathbf{G}_{t,m}^*) - (\mathbf{H}'_{J,m_0} \otimes \mathbf{H}'_{Q,m_0}) \text{vec}(\check{\mathbf{G}}_t) \\ &= \left[(\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}) - (\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) \right] \text{vec}(\mathbf{G}_{t,m}^*) \\ &\quad + \left[(\Xi_{J,m}^* \otimes \Xi_{Q,m}^*) \text{vec}(\mathbf{G}_{t,m}^*) - (\mathbf{H}'_{J,m_0} \otimes \mathbf{H}'_{Q,m_0}) \text{vec}(\check{\mathbf{G}}_t) \right] =: \mathcal{I}_{\Xi,G,21} + \mathcal{I}_{\Xi,G,22}. \end{aligned} \quad (\text{C.84})$$

In particular, for $\mathcal{I}_{\Xi,G,21}$, by (C.80) and $\|\mathbf{G}_{t,m}^*\|_F^2 = O_P(S_m)$, we have

$$\|\mathcal{I}_{\Xi,G,21}\|_F^2 \lesssim_P \left\| \widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m} - \Xi_{J,m}^* \otimes \Xi_{Q,m}^* \right\|_F^2 \cdot \|\mathbf{G}_{t,m}^*\|_F^2 \lesssim_P \max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}} S_m. \quad (\text{C.85})$$

Also, by (C.66),

$$\|\mathcal{I}_{\Xi,G,22}\|_F^2 = O_P \left(\eta_{Q,m} \eta_{J,m} p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \right). \quad (\text{C.86})$$

Combining (C.81), (C.82), (C.84), (C.83), (C.85), and (C.86), we obtain

$$\left\| \widehat{\mathbf{G}}_t - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} \right\|_F^2 \lesssim_P \max_{m \in [M]} \Delta_{G,m} + \max_{m' \in [M]} \frac{\Delta_{G,m'}}{S_{m'}} \max_{m \in [M]} S_m + \max_{m \in [M]} \eta_{Q,m} \eta_{J,m} p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}},$$

which thus concludes the proof for $\left\| \widehat{\mathbf{G}}_t - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} \right\|_F^2$, by noting that the last term above is dominated by the first term.

It remains to derive the rate for the re-estimated global loading matrices $\widetilde{\mathbf{Q}}_m$ and $\widetilde{\mathbf{J}}_m$ in Step 3 of Section 3.3. First, as $\mathbf{H}_{Q,m_0}$ and $\mathbf{H}_{J,m_0}$ are asymptotically orthogonal by Lemma 12, the population version of the re-estimated Kronecker product matrix is

$$\mathbf{K}_m^0 := \left\{ \mathbf{J}_m (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{-1/2} \mathbf{H}_{J,m_0} \right\} \otimes \left\{ \mathbf{Q}_m (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{-1/2} \mathbf{H}_{Q,m_0} \right\}. \quad (\text{C.87})$$

For each $m \in [M]$, define the matrices

$$\begin{aligned}\mathbf{Y}_m &:= \left(\mathbf{vec}(\check{\mathbf{Y}}_{G,1}^{(m)}), \dots, \mathbf{vec}(\check{\mathbf{Y}}_{G,T}^{(m)}) \right), \quad \mathbf{G}^* := \left(\mathbf{vec}(\mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_1 \mathbf{H}_{J,m_0}), \dots, \mathbf{vec}(\mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_T \mathbf{H}_{J,m_0}) \right), \\ \hat{\mathbf{G}} &:= \left(\mathbf{vec}(\hat{\mathbf{G}}_1), \dots, \mathbf{vec}(\hat{\mathbf{G}}_T) \right), \quad \mathbf{U}_m := \mathbf{Y}_m - \mathbf{K}_m^0 \mathbf{G}^*,\end{aligned}$$

so that $\mathbf{U}_m$ satisfies

$$\mathbf{U}_m = \mathbf{Y}_m - \mathbf{K}_m^0 \mathbf{G}^* = (\mathbf{Y}_m - \mathbf{Y}_m^0) + (\mathbf{Y}_m^0 - \mathbf{K}_m^0 \mathbf{G}^*), \quad (\text{C.88})$$

where $\mathbf{Y}_m^0$ denotes the matrix whose columns are $\mathbf{vec}(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m')$, $t \in [T]$. Consider the first term in (C.88), which has columns $\mathbf{vec}(\check{\mathbf{Y}}_{G,t}^{(m)} - \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m')$. By our already constructed result in Theorem 8 and with the notation Δ_G in the theorem's statement, we have $\|\hat{\mathbf{G}} - \mathbf{G}^*\|_F^2 = O_P(T\Delta_G)$. For each $m \in [M]$, by Theorem 7 and Lemma 3, we have

$$\begin{aligned}\frac{1}{T} \|\mathbf{Y}_m - \mathbf{Y}_m^0\|_F^2 &= \frac{1}{T} \sum_{t=1}^T \|(\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m') + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}\|_F^2 \\ &\lesssim_P \frac{\delta_{r,1}^{(m)}}{p_m} \frac{\delta_{c,1}^{(m)}}{q_m} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + p_m q_m.\end{aligned} \quad (\text{C.89})$$

For the second term in (C.88), for each $t \in [T]$,

$$\begin{aligned}&\left\| (\mathbf{Q}_m (\mathbf{Q}_{m_0}' \mathbf{Q}_{m_0})^{-1/2} \mathbf{H}_{Q,m_0}) \mathbf{H}_{Q,m_0}' \check{\mathbf{G}}_t \mathbf{H}_{J,m_0} (\mathbf{J}_m (\mathbf{J}_{m_0}' \mathbf{J}_{m_0})^{-1/2} \mathbf{H}_{J,m_0})' - \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \right\|_F^2 \\ &\lesssim_P (\eta_{Q,m_0} + \eta_{J,m_0}) \|\mathbf{Q}_m\|^2 \|\mathbf{J}_m\|^2 \|(\mathbf{Q}_{m_0}' \mathbf{Q}_{m_0})^{-1/2}\|^2 \|(\mathbf{J}_{m_0}' \mathbf{J}_{m_0})^{-1/2}\|^2 \|\check{\mathbf{G}}_t\|_F^2 \\ &= O_P \left((\eta_{Q,m_0} + \eta_{J,m_0}) \frac{p_m^{(m)} q_m^{(m)}}{\delta_{q,k_r}^{(m_0)} \delta_{j,k_c}^{(m_0)}} S_{m_0} \right),\end{aligned} \quad (\text{C.90})$$

where the last line used Assumption (S1), the definition of S_{m_0} in Theorem 8, and $\|\check{\mathbf{G}}_t\|_F^2 = O_P(S_{m_0})$.

Combining (C.88), (C.89), and (C.90), we have

$$\frac{1}{T} \|\mathbf{U}_m\|_F^2 \lesssim_P \frac{\delta_{r,1}^{(m)}}{p_m} \frac{\delta_{c,1}^{(m)}}{q_m} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + p_m q_m + (\eta_{Q,m_0} + \eta_{J,m_0}) \frac{p_m^{(m)} q_m^{(m)}}{\delta_{q,k_r}^{(m_0)} \delta_{j,k_c}^{(m_0)}} S_{m_0}. \quad (\text{C.91})$$

We next bound $\|T^{-1} \mathbf{U}_m \hat{\mathbf{G}}'\|_F^2$. By definition, the (i, j) entry of $T^{-1} \mathbf{U}_m \hat{\mathbf{G}}'$ is

$$(T^{-1} \mathbf{U}_m \hat{\mathbf{G}}')_{ij} = \frac{1}{T} \sum_{t=1}^T (\mathbf{U}_m)_{it} (\hat{\mathbf{G}})_{jt}, \quad (\text{C.92})$$

where $(\mathbf{U}_m)_{it}$ denotes the (i, t) entry of $\mathbf{U}_m$ and $(\hat{\mathbf{G}})_{jt}$ denotes the (j, t) entry of $\hat{\mathbf{G}}$. Recall the decompo-

sition $\mathbf{U}_m = (\mathbf{Y}_m - \mathbf{Y}_m^0) + (\mathbf{Y}_m^0 - \mathbf{K}_m^0 \mathbf{G}^*)$, where $\mathbf{Y}_m^0$ is the matrix whose columns are $\text{vec}(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m)$, $t \in [T]$. Thus we have

$$(\mathbf{U}_m)_{it} = \left[\text{vec}(\check{\mathbf{Y}}_{G,t}^{(m)} - \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) \right]_i + \left[\text{vec}(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) - \mathbf{K}_m^0 \mathbf{G}_{\cdot t}^* \right]_i,$$

where $\mathbf{G}_{\cdot t}^*$ denotes the t -th column of $\mathbf{G}^*$. By the Cauchy–Schwarz inequality,

$$\begin{aligned} (T^{-1} \mathbf{U}_m \hat{\mathbf{G}}')_{ij}^2 &\leq 2 \left[\frac{1}{T} \sum_{t=1}^T \left[\text{vec}(\check{\mathbf{Y}}_{G,t}^{(m)} - \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) \right]_i (\hat{\mathbf{G}})_{jt} \right]^2 \\ &\quad + 2 \left[\frac{1}{T} \sum_{t=1}^T \left[\text{vec}(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) - \mathbf{K}_m^0 \mathbf{G}_{\cdot t}^* \right]_i (\hat{\mathbf{G}})_{jt} \right]^2. \end{aligned} \quad (\text{C.93})$$

For the second term on the right-hand side of (C.93), by (C.90) and $\|\hat{\mathbf{G}}\|_F^2 = O_P(T S_{m_0})$, we have

$$\begin{aligned} &\left(\frac{1}{T} \sum_{t=1}^T \left[\text{vec}(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) - \mathbf{K}_m^0 \mathbf{G}_{\cdot t}^* \right]_i (\hat{\mathbf{G}})_{jt} \right)^2 \\ &\lesssim_P \frac{1}{T} \frac{1}{p_m q_m} (\eta_{Q,m_0} + \eta_{J,m_0}) \frac{p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}}}{S_{m_0}} S_{m_0} = \frac{1}{T p_m q_m} (\eta_{Q,m_0} + \eta_{J,m_0}) p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}}. \end{aligned} \quad (\text{C.94})$$

For the first term on the right-hand side of (C.93), we further decompose

$$\left[\text{vec}(\check{\mathbf{Y}}_{G,t}^{(m)} - \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m) \right]_i = \left[\text{vec}(\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}) \right]_i + \left[\text{vec}(\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}'_m) \right]_i.$$

By Assumption (E2), the noise process $\{\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}\}$ is independent of the factor process $\{\mathbf{G}_t\}$. By Lemma 1, the entries of $\{\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}\}$ have summable autocovariance bounded by $C\rho(\cdot)$. By Assumptions (S1), (F1), and (L1), the entries of $\hat{\mathbf{G}}$ have second moment of order S_{m_0} . Altogether, we obtain

$$\left(\frac{1}{T} \sum_{t=1}^T \left[\text{vec}(\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}) \right]_i (\hat{\mathbf{G}})_{jt} \right)^2 \lesssim_P \frac{S_{m_0}}{T}. \quad (\text{C.95})$$

For $\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}'_m$, by Assumption (IC2), the local factor $\mathbf{F}_t^{(m)}$ is uncorrelated with the global factor $\mathbf{G}_t$. Hence by Theorem 7, we have $\|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}'_m\|_F^2 \lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})$, so that

$$\left(\frac{1}{T} \sum_{t=1}^T \left[\text{vec}(\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}'_m) \right]_i (\hat{\mathbf{G}})_{jt} \right)^2 \lesssim_P \frac{1}{T p_m q_m} p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) S_{m_0}. \quad (\text{C.96})$$

Combining (C.93), (C.94), (C.95), and (C.96), summing over $i \in [p_m q_m]$ and $j \in [k_r k_c]$, we conclude

$$\begin{aligned} \left\| \frac{1}{T} \mathbf{U}_m \widehat{\mathbf{G}}' \right\|_F^2 &\lesssim_P \frac{p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})}{T} S_{m_0} \\ &+ \frac{1}{T} (\eta_{Q,m_0} + \eta_{J,m_0}) p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \lesssim \frac{p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})}{T} S_{m_0}. \end{aligned} \quad (\text{C.97})$$

With our notations above, by the construction in Step 3 of Section 3.3, we can write

$$\begin{aligned} \widehat{\mathbf{J}_m \otimes \mathbf{Q}_m} &= \mathbf{Y}_m \widehat{\mathbf{G}}' (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1} = (\mathbf{K}_m^0 \mathbf{G}^* + \mathbf{U}_m) \widehat{\mathbf{G}}' (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1} \\ &= \mathbf{K}_m^0 + \mathbf{K}_m^0 (\mathbf{G}^* \widehat{\mathbf{G}}' - \widehat{\mathbf{G}} \widehat{\mathbf{G}}') (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1} + \mathbf{U}_m \widehat{\mathbf{G}}' (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1}. \end{aligned} \quad (\text{C.98})$$

We bound the two terms in (C.98). First, by Assumption (S1), the smallest eigenvalue of $T^{-1} \mathbf{G}^* \mathbf{G}^*$ is of order S_{m_0} . Then we obtain

$$\|(\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1}\| = O_P(T^{-1} S_{m_0}^{-1}). \quad (\text{C.99})$$

Also, note that $\|\mathbf{G}^*\|_F^2 = O_P(T S_{m_0})$ and $\|\widehat{\mathbf{G}}\|_F^2 = O_P(T S_{m_0})$. Then for the first term in (C.98), we have

$$\begin{aligned} \left\| \mathbf{K}_m^0 (\mathbf{G}^* \widehat{\mathbf{G}}' - \widehat{\mathbf{G}} \widehat{\mathbf{G}}') (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1} \right\|_F^2 &\lesssim \|\mathbf{K}_m^0\|^2 \|\mathbf{G}^* - \widehat{\mathbf{G}}\|_F^2 \|\widehat{\mathbf{G}}\|^2 \|(\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1}\|^2 \\ &= O_P\left(\|\mathbf{K}_m^0\|^2 \cdot T \Delta_G \cdot T S_{m_0} \cdot T^{-2} S_{m_0}^{-2}\right) = O_P\left(\|\mathbf{K}_m^0\|^2 \frac{\Delta_G}{S_{m_0}}\right). \end{aligned} \quad (\text{C.100})$$

For the second term in (C.98), by (C.97) and (C.99) we have

$$\begin{aligned} \left\| \mathbf{U}_m \widehat{\mathbf{G}}' (\widehat{\mathbf{G}} \widehat{\mathbf{G}}')^{-1} \right\|_F^2 &= \left\| \left(\frac{1}{T} \mathbf{U}_m \widehat{\mathbf{G}}' \right) \left(\frac{1}{T} \widehat{\mathbf{G}} \widehat{\mathbf{G}}' \right)^{-1} \right\|_F^2 \lesssim_P \left\| \frac{1}{T} \mathbf{U}_m \widehat{\mathbf{G}}' \right\|_F^2 \cdot \left\| \left(\frac{1}{T} \widehat{\mathbf{G}} \widehat{\mathbf{G}}' \right)^{-1} \right\|_F^2 \\ &= O_P\left(\frac{S_{m_0}}{T} \{p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})\} \cdot S_{m_0}^{-2}\right) \\ &= O_P\left(\frac{p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})}{T S_{m_0}}\right). \end{aligned} \quad (\text{C.101})$$

Combining (C.98), (C.100), and (C.101), we arrive at

$$\left\| \widehat{\mathbf{J}_m \otimes \mathbf{Q}_m} - \mathbf{K}_m^0 \right\|_F^2 \lesssim_P \|\mathbf{K}_m^0\|^2 \frac{\Delta_G}{S_{m_0}} + \frac{p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})}{T S_{m_0}} = \Delta_{K,m}, \quad (\text{C.102})$$

where the last equality used the definition of $\Delta_{K,m}$ in the statement of Theorem 8, the fact that $\|\mathbf{K}_m^0\|_F^2 =$

$O_P(p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} / S_{m_0})$ which is direct from (C.87), and the condition from the statement of Theorem 8 that $\lambda_{\min}(\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0}) \lambda_{\min}(\mathbf{J}'_{m_0} \mathbf{J}_{m_0}) = O\{\lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \lambda_{\min}(\mathbf{J}'_m \mathbf{J}_m)\}$ and thus

$$\frac{p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m})}{TS_{m_0}} \lesssim \frac{\Delta_{G,m}}{TS_{m_0}} \lesssim \frac{\Delta_{G,m}}{S_{m_0}} \frac{1}{T} \frac{p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}}}{S_{m_0}} = o\left(\|\mathbf{K}_m^0\|^2 \frac{\Delta_G}{S_{m_0}}\right).$$

The remaining arguments are similar to the previous ones in this proof on $\|\hat{\mathbf{G}}_t - \mathbf{H}'_{Q,m_0} \check{\mathbf{G}}_t \mathbf{H}_{J,m_0}\|_F^2$. In particular, by the construction in Step 3 of Section 3.3, the matrix $\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m})$ has a rank-1 singular value decomposition

$$\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m}) = \hat{\sigma}_{\Xi,m} \hat{\mathbf{u}}_{\Xi,m} \hat{\mathbf{v}}'_{\Xi,m} + \hat{\Xi}_m^{(\text{rem})}, \quad (\text{C.103})$$

where the remainder $\hat{\Xi}_m^{(\text{rem})}$ includes the remaining singular components. Similarly, write the population version of the above as the rank-1 singular value decomposition and the remaining singular components (which are zero due to the following), such that

$$\mathcal{R}(\mathbf{K}_m^0) = \sigma_{*,m} \mathbf{u}_{*,m} \mathbf{v}'_{*,m} = \text{vec}(\mathbf{Q}_m^*) \text{vec}(\mathbf{J}_m^*)', \quad (\text{C.104})$$

where $\sigma_{*,m}$ denotes the largest singular value of $\mathcal{R}(\mathbf{K}_m^0)$. By Assumptions (S1), (L1), and the asymptotic orthogonality of $\mathbf{H}_{Q,m_0}$ and $\mathbf{H}_{J,m_0}$ from Lemma 12, the leading singular value of $\mathcal{R}(\mathbf{K}_m^0)$ satisfies $\sigma_{*,m} \asymp_P 1$. Then by Weyl's inequality, (C.102), and recall that $\mathcal{R}(\cdot)$ preserves the Frobenius norm,

$$|\hat{\sigma}_{\Xi,m} - \sigma_{*,m}| \lesssim_P \|\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m}) - \mathcal{R}(\mathbf{K}_m^0)\| \lesssim_P \Delta_{K,m}^{1/2}, \quad (\text{C.105})$$

which then implies $\hat{\sigma}_{\Xi,m} \asymp_P 1$. By the Davis–Kahan theorem (Yu et al., 2015, Theorem 2), we have

$$\|\hat{\mathbf{u}}_{\Xi,m} - \mathbf{u}_{*,m}\|^2 \lesssim_P \frac{\|\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m}) - \mathcal{R}(\mathbf{K}_m^0)\|_F^2}{(\sigma_{*,m} - \sigma_{*,m}^{(2)})^2}, \quad \|\hat{\mathbf{v}}_{\Xi,m} - \mathbf{v}_{*,m}\|^2 \lesssim_P \frac{\|\mathcal{R}(\widehat{\mathbf{J}_m \otimes \mathbf{Q}_m}) - \mathcal{R}(\mathbf{K}_m^0)\|_F^2}{(\sigma_{*,m} - \sigma_{*,m}^{(2)})^2},$$

where $\sigma_{*,m}^{(2)}$ denotes the second largest singular value of $\mathcal{R}(\mathbf{K}_m^0)$ such that $\sigma_{*,m}^{(2)} = 0$ due to (C.104), and thus $|\sigma_{*,m} - \sigma_{*,m}^{(2)}| \asymp_P 1$. Then with (C.105), we may conclude

$$\|\hat{\mathbf{u}}_{\Xi,m} - \mathbf{u}_{*,m}\|^2 \lesssim_P \Delta_{K,m}, \quad \|\hat{\mathbf{v}}_{\Xi,m} - \mathbf{v}_{*,m}\|^2 \lesssim_P \Delta_{K,m}. \quad (\text{C.106})$$

Recall from Step 3 of Section 3.3 that $\text{vec}(\tilde{\mathbf{Q}}_m) = \hat{\sigma}_{\Xi,m}^{1/2} \hat{\mathbf{u}}_{\Xi,m}$ and $\text{vec}(\tilde{\mathbf{J}}_m) = \hat{\sigma}_{\Xi,m}^{1/2} \hat{\mathbf{v}}_{\Xi,m}$. Similarly, define their population versions $\text{vec}(\mathbf{Q}_m^*) := \sigma_{*,m}^{1/2} \mathbf{u}_{*,m}$ and $\text{vec}(\mathbf{J}_m^*) = \sigma_{*,m}^{1/2} \mathbf{v}_{*,m}$. Combining (C.105) and

(C.106), we obtain

$$\begin{aligned} \|\mathbf{vec}(\tilde{\mathbf{Q}}_m) - \mathbf{vec}(\mathbf{Q}_m^*)\|^2 &= \|\hat{\sigma}_{\Xi,m}^{1/2} \hat{\mathbf{u}}_{\Xi,m} - \sigma_{*,m}^{1/2} \mathbf{u}_{*,m}\|^2 \\ &\lesssim |\hat{\sigma}_{\Xi,m}^{1/2} - \sigma_{*,m}^{1/2}|^2 \|\hat{\mathbf{u}}_{\Xi,m}\|^2 + \sigma_{*,m} \|\hat{\mathbf{u}}_{\Xi,m} - \mathbf{u}_{*,m}\|^2 = O_P(\Delta_{K,m}), \end{aligned} \quad (\text{C.107})$$

and analogously,

$$\|\mathbf{vec}(\tilde{\mathbf{J}}_m) - \mathbf{vec}(\mathbf{J}_m^*)\|^2 = O_P(\Delta_{K,m}). \quad (\text{C.108})$$

Thus finally, the re-estimated global loading matrices $\tilde{\mathbf{Q}}_m$ and $\tilde{\mathbf{J}}_m$ satisfy

$$\|\tilde{\mathbf{Q}}_m - \mathbf{Q}_m^*\|_F^2 = O_P(\Delta_{K,m}), \quad \|\tilde{\mathbf{J}}_m - \mathbf{J}_m^*\|_F^2 = O_P(\Delta_{K,m}), \quad (\text{C.109})$$

where

$$\mathbf{Q}_m^* := \mathbf{Q}_m (\mathbf{Q}'_{m_0} \mathbf{Q}_{m_0})^{-1/2} \mathbf{H}_{Q,m_0}, \quad \mathbf{J}_m^* := \mathbf{J}_m (\mathbf{J}'_{m_0} \mathbf{J}_{m_0})^{-1/2} \mathbf{H}_{J,m_0}.$$

This completes the proof of the theorem. $\square$

Proof of Theorem 9. For each $m \in [M]$ and any $c_m \in [k_r]$ and $d_m \in [k_c]$, we have

$$\begin{aligned} \hat{\mathbf{Q}}'_{m,\cdot c_m} \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{J}}_{m,\cdot d_m} &= (\hat{\mathbf{J}}'_{m,\cdot d_m} \mathbf{J}_m \otimes \hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{Q}_m) \mathbf{vec}(\mathbf{G}_t) + (\hat{\mathbf{J}}'_{m,\cdot d_m} \mathbf{C}_m \otimes \hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{R}_m) \mathbf{vec}(\mathbf{F}_t^{(m)}) \\ &\quad + \hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{J}}_{m,\cdot d_m}. \end{aligned} \quad (\text{C.110})$$

From the proof of Theorem 7, for some invertible constant order matrices $\mathbf{H}_q$ and $\mathbf{H}_j$, we have

$$\|\hat{\mathbf{Q}}_m - \mathbf{Q}_m (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{H}_{q,m}\| = O_P(\eta_{Q,m}) = o_P(1), \quad \|\hat{\mathbf{J}}_m - \mathbf{J}_m (\mathbf{J}'_m \mathbf{J}_m)^{-1/2} \mathbf{H}_{j,m}\| = O_P(\eta_{J,m}) = o_P(1). \quad (\text{C.111})$$

Hence, defining $\mathbf{e}_{p,k}$ as the real unit vector of length p with 1 at its k th position, we have

$$\begin{aligned} \max_{c_m \in [k_r]} \|\hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{Q}_m\| &= \max_{c_m \in [k_r]} \|\mathbf{e}'_{k_r, c_m} \hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{Q}_m\| \asymp \|\hat{\mathbf{Q}}'_{m,\cdot c_m} \mathbf{Q}_m\| \\ &\geq \|(\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \mathbf{H}_{q,m}\| - \|\mathbf{Q}'_m (\hat{\mathbf{Q}}_m - \mathbf{Q}_m (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{H}_{q,m})\| \asymp_P p_m^{\delta_{q,1}^{(m)}/2}. \end{aligned} \quad (\text{C.112})$$

Similarly,

$$\max_{d_m \in [k_c]} \|\hat{\mathbf{J}}'_{m,\cdot d_m} \mathbf{J}_m\| \asymp_P q_m^{\delta_{j,1}^{(m)}/2}. \quad (\text{C.113})$$

Consider $m \in \mathcal{A}_G$. Then (C.110) implies that

$$\begin{aligned}
T^{-1} \sum_{t=1}^T \hat{\mathbf{Q}}'_{m, \cdot c_m} \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{J}}_{m, \cdot d_m} \hat{\mathbf{J}}'_{n, \cdot d_n} \check{\mathbf{Y}}_t^{(n)'} \hat{\mathbf{Q}}_{m, \cdot c_n} &= \sum_{j=0}^9 I_j, \text{ where} \\
I_0 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{J}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \mathbb{E}(\text{vec}(\mathbf{G}_t))^{\otimes 2} (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{J}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{Q}_n)', \\
I_1 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{J}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{Q}_m) T^{-1} \sum_{t=1}^T (\text{vec}(\mathbf{G}_t)^{\otimes 2} - \mathbb{E}(\text{vec}(\mathbf{G}_t))^{\otimes 2}) (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{J}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{Q}_n)', \\
I_2 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{J}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t) \text{vec}(\mathbf{F}_t^{(n)})' (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{C}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{R}_n)', \\
I_3 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{J}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t) \hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \hat{\mathbf{Q}}_{n, \cdot c_n}, \\
I_4 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{C}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \text{vec}(\mathbf{G}_t)' (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{J}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{Q}_n)', \\
I_5 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{C}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \text{vec}(\mathbf{F}_t^{(n)})' (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{C}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{R}_n)', \\
I_6 &:= (\hat{\mathbf{J}}'_{m, \cdot d_m} \mathbf{C}_m \otimes \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \hat{\mathbf{Q}}_{n, \cdot c_n}, \\
I_7 &:= \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{M}_{p_m} T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{J}}_{m, \cdot d_m} \text{vec}(\mathbf{G}_t)' (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{J}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{Q}_n)', \\
I_8 &:= \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{M}_{p_m} T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{J}}_{m, \cdot d_m} \text{vec}(\mathbf{F}_t^{(n)})' (\hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{C}_n \otimes \hat{\mathbf{Q}}'_{n, \cdot c_n} \mathbf{R}_n)', \\
I_9 &:= \hat{\mathbf{Q}}'_{m, \cdot c_m} \mathbf{M}_{p_m} T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{J}}_{m, \cdot d_m} \hat{\mathbf{J}}'_{n, \cdot d_n} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \hat{\mathbf{Q}}_{n, \cdot c_n},
\end{aligned}$$

with $I_0 = I_1 = I_4 = I_7 = 0$ if $n \notin \mathcal{A}_G$. By Assumption (S1) on $T^{-1} \sum_{t=1}^T \mathbb{E} \text{vec}(\mathbf{G}_t)^{\otimes 2}$, and that the set of vectors $\{\mathbf{e}'_{k_c, d_n} \mathbf{H}'_{j, n} (\mathbf{J}'_n \mathbf{J}_n)^{1/2} \otimes \mathbf{e}'_{k_r, c_n} \mathbf{H}'_{q, n} (\mathbf{Q}'_n \mathbf{Q}_n)^{1/2}\}_{c_n \in [k_r], d_n \in [k_c]}$ is linearly independent by Assumption (L1), for each $n \neq m$ with $n \in \mathcal{A}_G$, there must exist $c_n \in [k_r], d_n \in [k_c]$ (through using (C.111), so that we can replace $\hat{\mathbf{Q}}_{n, \cdot c_n}$ by $\mathbf{Q}_n (\mathbf{Q}'_n \mathbf{Q}_n)^{-1/2} \mathbf{H}_{q, n} \mathbf{e}_{k_r, c_n}$ and similarly $\hat{\mathbf{J}}_{n, \cdot d_n}$ by $\mathbf{J}_n (\mathbf{J}'_n \mathbf{J}_n)^{-1/2} \mathbf{H}_{j, n} \mathbf{e}_{k_c, d_n}$) such that

$$\begin{aligned}
\max_{c_m \in [k_r], d_m \in [k_c]} I_0^2 &\asymp_P \max_{c_m \in [k_r], d_m \in [k_c]} \left\| \mathbf{e}'_{k_c, d_m} \mathbf{H}'_{j, m} (\mathbf{J}'_m \mathbf{J}_m)^{1/2} \otimes \mathbf{e}'_{k_r, c_m} \mathbf{H}'_{q, m} (\mathbf{Q}'_m \mathbf{Q}_m)^{1/2} \right\|^2 \\
&\cdot \left\| \mathbf{e}'_{k_c, d_n} \mathbf{H}'_{j, n} (\mathbf{J}'_n \mathbf{J}_n)^{1/2} \otimes \mathbf{e}'_{k_r, c_n} \mathbf{H}'_{q, n} (\mathbf{Q}'_n \mathbf{Q}_n)^{1/2} \right\|^2 \\
&\gtrsim_P \frac{\delta_{q, 1}^{(m)}}{p_m} \frac{\delta_{j, 1}^{(m)}}{q_m} \cdot \frac{\delta_{q, k_r}^{(n)}}{p_n} \frac{\delta_{j, k_c}^{(n)}}{q_n}, \tag{C.114}
\end{aligned}$$

where the last line used (C.112), (C.113) and Assumption (L1). By Lemma 2-2 and 2-3, we also have

$$\begin{aligned} \max_{c_m \in [k_r], d_m \in [k_c]} I_1^2 &\lesssim_P p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \cdot p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}} R(T, d_{\text{fac}, \star}), & \max_{c_m \in [k_r], d_m \in [k_c]} I_2^2 &\lesssim_P p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \cdot p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} R(T, d_{\text{fac}, \star}), \\ \max_{c_m \in [k_r], d_m \in [k_c]} I_4^2 &\lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} \cdot p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}} R(T, d_{\text{fac}, \star}), & \max_{c_m \in [k_r], d_m \in [k_c]} I_5^2 &\lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} \cdot p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} R(T, d_{\text{fac}, \star}). \end{aligned}$$

Using similar proof for Lemma 3, and the fact that

$$\max_{i \in [p_m], j \in [k_r]} |(\hat{\mathbf{Q}}_m - \mathbf{Q}_m(\mathbf{Q}_m' \mathbf{Q}_m)^{-1/2} \mathbf{H}_{q,m})_{ij}| = O_P(\|\hat{\mathbf{Q}}_m - \mathbf{Q}_m(\mathbf{Q}_m' \mathbf{Q}_m)^{-1/2} \mathbf{H}_{q,m}\|) = o_P(\eta_{Q,m}) = o_P(1),$$

and similarly

$$\max_{i \in [q_m], j \in [k_c]} |(\hat{\mathbf{J}}_m - \mathbf{J}_m(\mathbf{J}_m' \mathbf{J}_m)^{-1/2} \mathbf{H}_{j,m})_{ij}| = O_P(\|\hat{\mathbf{J}}_m - \mathbf{J}_m(\mathbf{J}_m' \mathbf{J}_m)^{-1/2} \mathbf{H}_{j,m}\|) = o_P(\eta_{J,m}) = o_P(1),$$

we can prove the following in parallel to (D.33), (D.34) and (D.35)(details omitted):

$$\|T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \hat{\mathbf{v}}_m \mathbf{u}_t'\|_F^2 = O_P(T^{-1} p_m), \quad (\text{C.115})$$

$$\|T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \hat{\mathbf{v}}_m \hat{\mathbf{v}}_n' \mathbf{E}_t^{(n)'}\|_F^2 = O_P(T^{-1} p_m p_n), \quad (\text{C.116})$$

$$\|T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \hat{\mathbf{v}}_m \hat{\mathbf{v}}_m' \mathbf{E}_t^{(m)'}\|_F^2 = O_P(p_m + T^{-1} p_m^2), \quad (\text{C.117})$$

where $\{\mathbf{u}_t\}$ is independent of $\{\mathbf{E}_t^{(m)}\}$ for each $m \in [M]$, and $\hat{\mathbf{v}}_m, \hat{\mathbf{v}}_n$ are unit vectors potentially dependent on $\{\mathbf{E}_t^{(m)}\}$ for some $m \in [M]$, but are consistent to some non-random unit vectors $\mathbf{v}_m$ and $\mathbf{v}_n$ respectively.

With (C.115) and (C.116), we can show that

$$\begin{aligned} \max_{c_m \in [k_r], d_m \in [k_c]} I_3^2 &\asymp_P p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \cdot T^{-1} p_n, & \max_{c_m \in [k_r], d_m \in [k_c]} I_6^2 &\asymp_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} \cdot T^{-1} p_n, \\ \max_{c_m \in [k_r], d_m \in [k_c]} I_7^2 &\asymp_P T^{-1} p_m \cdot p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}, & \max_{c_m \in [k_r], d_m \in [k_c]} I_8^2 &\asymp_P T^{-1} p_m \cdot p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}}, \\ \max_{c_m \in [k_r], d_m \in [k_c]} I_9^2 &\asymp_P T^{-1} p_m p_n. \end{aligned}$$

Using the rate assumption in the statement of the theorem, together with Assumption (R2), (C.114) and the rates for I_1^2 to I_9^2 result in

$$\sum_{n \neq m} \max_{c_m, c_n \in [k_r]} \max_{d_m, d_n \in [k_c]} \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{Q}}_{m, \cdot c_m}' \check{\mathbf{Y}}_t^{(m)} \hat{\mathbf{J}}_{m, \cdot d_m}' \hat{\mathbf{J}}_{n, \cdot d_n}' \check{\mathbf{Y}}_t^{(n)'} \hat{\mathbf{Q}}_{n, \cdot c_n} \right\}^2 \gtrsim_P p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} \max_{n \neq m, n \in \mathcal{A}_G} p_n^{\delta_{q,kr}^{(n)}} q_n^{\delta_{j,kc}^{(n)}}.$$

On the other hand, if $m = k \notin \mathcal{A}_G$, then $I_0 = I_1 = I_2 = I_3 = 0$, and the rates for I_4^2 to I_9^2 lead to

$$\begin{aligned} & \sum_{n \neq k} \max_{c_k, c_n \in [k_r]} \max_{d_k, d_n \in [k_c]} \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{Q}}'_{k, c_k} \check{\mathbf{Y}}_t^{(k)} \hat{\mathbf{J}}_{k, d_k} \hat{\mathbf{J}}'_{n, d_n} \check{\mathbf{Y}}_t^{(n)'} \hat{\mathbf{Q}}_{n, c_n} \right\}^2 \\ & \lesssim_P p_k^{\delta_{r,1}^{(k)}} q_k^{\delta_{c,1}^{(k)}} \max \left(\max_{n \in \mathcal{A}_G} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}, \max_{n \neq k, n \notin \mathcal{A}_G} p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} \right) R(T, d_{\text{fac}, \star}). \end{aligned}$$

Now we rewrite $\mathbf{Y}_t^{(m)}$ as

$$\check{\mathbf{Y}}_t^{(m)} = \mathbf{U}_m \mathbf{H}_t^{(m)} \mathbf{V}_m' + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m},$$

where $\mathbf{U}_m := (\mathbf{Q}_m, \mathbf{R}_m)$, $\mathbf{V}_m := (\mathbf{J}_m, \mathbf{C}_m)$ and $\mathbf{H}_t^{(m)} := \text{diag}(\mathbf{G}_t, \mathbf{F}_t^{(m)})$. We then have $\text{rank}(\mathbf{U}_m) \leq k_r + k_r^{(m)}$ and $\text{rank}(\mathbf{V}_m) \leq k_c + k_c^{(m)}$. Repeating our estimation methodology for just the local common components for individual groups using data $\{\check{\mathbf{Y}}_t^{(m)}\}$, we can use the same proof as Theorem 7 to arrive at $\|\hat{\mathbf{U}}_m - \mathbf{U}_m(\mathbf{U}_m' \mathbf{U}_m)^{-1/2} \mathbf{H}_{u,m}\|_F^2 = o_P(1)$ and $\|\hat{\mathbf{V}}_m - \mathbf{V}_m(\mathbf{V}_m' \mathbf{V}_m)^{-1/2} \mathbf{H}_{v,m}\|_F^2 = o_P(1)$, where $\mathbf{H}_{u,m}$ and $\mathbf{H}_{v,m}$ are two constant order invertible matrices. Then using (C.117) and Assumption (R2) and (S1), similar to the treatment before, we have for $m \in \mathcal{A}_G$,

$$\begin{aligned} & \max_{c_m \in [k_r + k_r^{(m)}], d_m \in [k_c + k_c^{(m)}]} \left(T^{-1} \sum_{t=1}^T (\hat{\mathbf{U}}'_{m, c_m} \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \hat{\mathbf{V}}_{m, d_m} + \hat{\mathbf{U}}'_{m, c_m} \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \hat{\mathbf{V}}_{m, d_m} \right. \\ & \quad \left. + \hat{\mathbf{U}}'_{m, c_m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{V}}_{m, d_m})^2 \right)^2 \\ & \gtrsim_P p_m^{2\delta_{q,1}^{(m)}} q_m^{2\delta_{j,1}^{(m)}} - p_m^{2\delta_{r,1}^{(m)}} q_m^{2\delta_{c,1}^{(m)}} - p_m - T^{-1} p_m^2 \asymp_P p_m^{2\delta_{q,1}^{(m)}} q_m^{2\delta_{j,1}^{(m)}}. \end{aligned}$$

At the same time, if $m \notin \mathcal{A}_G$,

$$\begin{aligned} & \max_{c_m \in [k_r + k_r^{(m)}], d_m \in [k_c + k_c^{(m)}]} \left(T^{-1} \sum_{t=1}^T (\hat{\mathbf{U}}'_{m, c_m} \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \hat{\mathbf{V}}_{m, d_m} + \hat{\mathbf{U}}'_{m, c_m} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \hat{\mathbf{V}}_{m, d_m})^2 \right)^2 \\ & \gtrsim_P p_m^{2\delta_{r,1}^{(m)}} q_m^{2\delta_{c,1}^{(m)}} - p_m - T^{-1} p_m^2 \asymp_P p_m^{2\delta_{r,1}^{(m)}} q_m^{2\delta_{c,1}^{(m)}}. \end{aligned}$$

This completes the proof of the theorem. $\square$

Proof of Theorem 10. Consider $\hat{r}_1$. We omit the details for $\hat{\ell}_1$ since its proof is parallel to that for $\hat{r}_1$. It is not difficult to see that, through the finiteness of M ,

$$\max_{n \in [M]} \max_{n \neq m} \sigma_{r_1}^2 \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{a}}_t^{(m)} \hat{\mathbf{a}}_t^{(n)'} \right\} \lesssim \max_{m \in [M]} \lambda_i(\hat{\boldsymbol{\Omega}}_{\mathbf{A}_1, m}) \lesssim \max_{m \in [M]} \max_{n \neq m} \sigma_1^2 \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{a}}_t^{(m)} \hat{\mathbf{a}}_t^{(n)'} \right\}, \quad i \in [r_1], \quad (\text{C.118})$$

where $\sigma_i(\mathbf{A})$ denotes the i th largest singular value of the matrix $\mathbf{A}$. At the same time, for $i \in [r_1]$,

$$\sigma_i \left\{ \mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{\otimes 2}) \mathbf{A}'_{1n} \right\} - \sum_{j=2}^5 I_j \leq \sigma_i \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{a}}_t^{(m)} \hat{\mathbf{a}}_t^{(n)'} \right\} \leq \sigma_1 \left\{ \mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{\otimes 2}) \mathbf{A}'_{1n} \right\} + \sum_{j=2}^5 I_j, \quad (\text{C.119})$$

where

$$\begin{aligned} I_2 &:= \sigma_1 \left\{ \mathbf{A}_{2m} (T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(n)'}) \mathbf{A}'_{2n} \right\}, \quad I_3 := \sigma_1 \left\{ T^{-1} \sum_{t=1}^T q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} (\mathbf{A}_{1n} \boldsymbol{\alpha}_t + \mathbf{A}_{2n} \boldsymbol{\alpha}_t^{(n)})' \right\}, \\ I_4 &:= \sigma_1 \left\{ T^{-1} \sum_{t=1}^T (\mathbf{A}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)}) q_n^{-1} \mathbf{1}'_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right\}, \\ I_5 &:= \sigma_1 \left\{ T^{-1} q_m^{-1} q_n^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}'_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right\}. \end{aligned}$$

Firstly, for $i \in [r_1]$, using (D.14) and Assumption (S1),

$$\begin{aligned} \sigma_i \left\{ \mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{\otimes 2}) \mathbf{A}'_{1n} \right\} &\geq \lambda_i^{1/2} (\mathbf{A}'_{1m} \mathbf{A}_{1m}) \lambda_{r_1}^{1/2} (\mathbf{A}'_{1n} \mathbf{A}_{1n}) \lambda_{r_1} (T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{\otimes 2}) \\ &\gtrsim_P p_m^{\gamma_r/2} p_n^{\gamma_r/2} (1 + R(T, d_{r,*})^{1/2}). \end{aligned}$$

Using (ME1), and (D.15) on I_2 , (D.33) on I_3 and I_4 and (D.34) on I_5 ,

$$\begin{aligned} I_2 &:= O_P(p_m^{\gamma_r^{(m)}/2} p_n^{\gamma_r^{(n)}/2} R(T, d_{r,*})^{1/2}) = o_P(p_m^{\gamma_r/2} p_n^{\gamma_r/2}), \\ I_3 &:= O_P(T^{-1/2} p_m^{1/2} q_m^{-1/2} \cdot (p_n^{\gamma_r/2} + p_n^{\gamma_r^{(n)}/2})) = o_P(p_m^{\gamma_r/2} p_n^{\gamma_r/2}), \\ I_4 &:= O_P(T^{-1/2} p_n^{1/2} q_n^{-1/2} \cdot (p_m^{\gamma_r/2} + p_m^{\gamma_r^{(m)}/2})) = o_P(p_m^{\gamma_r/2} p_n^{\gamma_r/2}), \\ I_5 &:= O_P(T^{-1/2} p_m^{1/2} p_n^{1/2} q_m^{-1/2} q_n^{-1/2}) = o_P(p_m^{\gamma_r/2} p_n^{\gamma_r/2}), \end{aligned}$$

where we use (R1) as well for concluding that all the rates above are $o_P(p_m^{\gamma_r/2} p_n^{\gamma_r/2})$. Hence (C.119) implies that for $i \in [r_1]$,

$$\max_{m,n \in [M], n \neq m} \sigma_i^2 \left\{ T^{-1} \sum_{t=1}^T \hat{\mathbf{a}}_t^{(m)} \hat{\mathbf{a}}_t^{(n)'} \right\} \asymp_P \max_{m,n \in [M], n \neq m} p_m^{\gamma_r} p_n^{\gamma_r},$$

so that by (C.118), for $i \in [r_1]$,

$$\max_{m \in [M]} \lambda_i(\hat{\boldsymbol{\Omega}}_{\mathbf{A}_1, m}) \asymp_P \max_{m,n \in [M], n \neq m} p_m^{\gamma_r} p_n^{\gamma_r}. \quad (\text{C.120})$$

On the other hand, we can write

$$\begin{aligned}
\widehat{\Omega}_{\mathbf{A}_1, m} &= (M-1)^{-1} \sum_{n \neq m} \left(\mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) \mathbf{A}'_{1n} + R_{m,n} \right)^{\otimes 2} \\
&= (M-1)^{-1} \sum_{n \neq m} \left(\mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) \mathbf{A}'_{1n} \mathbf{A}_{1n} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) \mathbf{A}'_{1m} + \mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) \mathbf{A}'_{1n} R'_{m,n} \right. \\
&\quad \left. + R_{m,n} \mathbf{A}_{1n} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) \mathbf{A}'_{1m} + R_{m,n} R'_{m,n} \right),
\end{aligned}$$

so that for $i \in \{r_1 + 1, \dots, p_m\}$,

$$\begin{aligned}
\lambda_i(\widehat{\Omega}_{\mathbf{A}_1, m}) &\leq \lambda_1 \left\{ (M-1)^{-1} \sum_{n \neq m} R_{m,n} R'_{m,n} \right\} \leq \max_{n \neq m} \sigma_1^2(R_{m,n}) \\
&\lesssim \max_{n \neq m} \sum_{j=2}^5 I_j^2 = O_P(\max_{n \neq m} p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} R(T, d_{r, \star})),
\end{aligned} \tag{C.121}$$

where the last equality used Assumption (R1). Hence combining (C.120) and (C.121), we have

$$\frac{\max_{m \in [M]} \lambda_{i+1}(\widehat{\Omega}_{\mathbf{A}_1, m})}{\max_{m \in [M]} \lambda_i(\widehat{\Omega}_{\mathbf{A}_1, m})} \begin{cases} \asymp_P 1, & i \in [r_1 - 1]; \\ \lesssim_P \frac{\max_{m \in [M], n \neq m} p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} R(T, d_{r, \star})}{\max_{m \in [M], n \neq m} p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}}} = o_P(1), & i = r_1, \end{cases}$$

where the equation for the case $i = r_1$ used Assumption (R1). This implies $P(\widehat{r}_1 \geq r_1) \rightarrow 1$. It is also clear that using (C.120) and (C.121), the ratio

$$\frac{\max_{m \in [M]} \lambda_{i+1}(\widehat{\Omega}_{\mathbf{A}_1, m}) + \max_{m \neq n \in [M]} C p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} R(T, d_{r, \star})}{\max_{m \in [M]} \lambda_i(\widehat{\Omega}_{\mathbf{A}_1, m}) + \max_{m \neq n \in [M]} C p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} R(T, d_{r, \star})} \begin{cases} \asymp_P 1, & i \in [r_1 - 1]; \\ = o_P(1), & i = r_1, \\ \asymp_P 1, & i = r_1 + 1, \dots, p_m, \end{cases}$$

so that $P(\widehat{r}_1 = r_1) \rightarrow 1$, which aligns with the statement of adding $\max_{m \neq n \in [M]} C T^{-1} p_m p_n$ to the ratio when all factors are strong and all processes have short memory.

For $\widehat{r}_2^{(m)}$ (We omit $\widehat{\ell}_2^{(m)}$ since the proof is parallel), given $i \in [r_2^{(m)}]$, consider

$$\begin{aligned}
J_1 - \sum_{i=2}^6 J_i &\lesssim \sigma_i \left(T^{-1} \sum_{t=1}^T (\widehat{\mathbf{a}}_t^{(m)})^{\otimes 2} \widehat{\mathbf{A}}_{1m}^\perp \right) \lesssim J_1 + \sum_{i=2}^6 J_i, \quad \text{where} \\
J_1 &:= \sigma_i \left(\mathbf{A}_{2m} (T^{-1} \sum_{t=1}^T (\alpha_t^{(m)})^{\otimes 2}) \mathbf{A}'_{2m} \mathbf{A}_{1m}^\perp \right), \\
J_2 &:= \sigma_1 \left(\mathbf{A}_{1m} (T^{-1} \sum_{t=1}^T \alpha_t^{\otimes 2}) (\mathbf{A}'_{1m} \mathbf{A}_{1m})^{1/2} (\mathbf{H}_{\mathbf{A}_1, m}^{-1})' (\check{\mathbf{A}}_{1m} \mathbf{H}_{A_1, m} - \widehat{\mathbf{A}}_{1m})' \widehat{\mathbf{A}}_{1m}^\perp \right),
\end{aligned} \tag{C.122}$$

$$\begin{aligned}
J_3 &:= \sigma_i \left(\mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T (\boldsymbol{\alpha}_t^{(m)})^{\otimes 2} \right) \mathbf{A}_{2m}' (\widehat{\mathbf{A}}_{1m}^\perp - \mathbf{A}_{1m}^\perp) \right), \\
J_4 &:= \sigma_1 \left(T^{-1} \sum_{t=1}^T q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} (\mathbf{A}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)})' \widehat{\mathbf{A}}_{1m}^\perp \right), \\
J_5 &:= \sigma_1 \left(T^{-1} \sum_{t=1}^T (\mathbf{A}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)}) \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} q_m^{-1} \widehat{\mathbf{A}}_{1m}^\perp \right), \\
J_6 &:= \sigma_1 \left(T^{-1} q_m^{-2} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} \widehat{\mathbf{A}}_{1m}^\perp \right).
\end{aligned}$$

For $i \in [r_2^{(m)}]$, using (D.14) and Assumption (S1) and (L1),

$$J_1 = \sigma_i \left(\left(T^{-1} \sum_{t=1}^T (\boldsymbol{\alpha}_t^{(m)})^{\otimes 2} \right) (\mathbf{A}_{2m}' \mathbf{A}_{1m}^\perp)^{\otimes 2} \right) \geq \sigma_i^2 (\mathbf{A}_{2m}' \mathbf{A}_{1m}^\perp) \sigma_{r_2^{(m)}} \left(T^{-1} \sum_{t=1}^T (\boldsymbol{\alpha}_t^{(m)})^{\otimes 2} \right) \asymp_P p_m^{\gamma_r^{(m)}},$$

Also,

$$\begin{aligned}
J_2 &= O_P(p_m^{\gamma_r} \eta_{A_{1m}}^{1/2}) = o_P(p_m^{\gamma_r^{(m)}}), \\
J_3 &= O_P(p_m^{\gamma_r^{(m)}} \|\widehat{\mathbf{A}}_{1m}^\perp - \mathbf{A}_{1m}^\perp\|) = O_P(p_m^{\gamma_r^{(m)}} \|\widehat{\mathbf{A}}_{1m} - \check{\mathbf{A}}_{1m} \mathbf{H}_{A_1, m}\|) = o_P(p_m^{\gamma_r^{(m)}}), \\
J_4, J_5 &= O_P(T^{-1/2} p_m^{1/2} q_m^{-1/2} \cdot (p_m^{\gamma_r/2} + p_m^{\gamma_r^{(m)}/2})) = o_P(p_m^{\gamma_r^{(m)}}), \\
J_6 &= O_P(p_m^{1/2} q_m^{-1} + T^{-1/2} p_m q_m^{-1}) = o_P(p_m^{\gamma_r^{(m)}}),
\end{aligned}$$

where $\eta_{A_{1m}}$ is the rate defined in Theorem 3, and the equation for J_2 used the assumption in the statement of the theorem. We use (D.33) for J_4 and J_5 and that $\eta_{A_{1m}} = o(\nu_{r,m}^2)$ and (R1) to conclude that they are $o_P(p_m^{\gamma_r^{(m)}})$, while J_6 uses (D.35).

On the other hand, for $i \in \{r_2^{(m)+1}, \dots, p_m\}$,

$$\sigma_i \left(T^{-1} \sum_{t=1}^T (\widehat{\mathbf{a}}_t^{(m)})^{\otimes 2} \widehat{\mathbf{A}}_{1m}^\perp \right) \lesssim_P \sum_{i=2}^6 J_i = o_P(p_m^{\gamma_r^{(m)}}). \quad (\text{C.123})$$

The rates for J_1 to J_6 , combined with (C.122) and (C.123), allow us to conclude that

$$\frac{\lambda_{j+1}(\widehat{\boldsymbol{\Omega}}_{A_2, m})}{\lambda_j(\widehat{\boldsymbol{\Omega}}_{A_2, m})} \begin{cases} \asymp_P 1, & j \in [r_2^{(m)} - 1]; \\ = o_P(1), & j = r_2^{(m)}, \end{cases}$$

so that $P(\widehat{r}_2^{(m)} \geq r_2^{(m)}) \rightarrow 1$. It is also clear that when all factors are strong and all processes have short

memory, $\eta_{A_{1m}} = T^{-1}$, so that the rate in (C.123) is $O_P(T^{-1/2}p_m)$, and the ratio

$$\frac{\lambda_{j+1}(\hat{\mathbf{\Omega}}_{A_2,m}) + T^{-1}p_m^2}{\lambda_j(\hat{\mathbf{\Omega}}_{A_2,m}) + T^{-1}p_m^2} \begin{cases} \asymp_P 1, & j \in [r_2^{(m)} - 1]; \\ = o_P(1), & j = r_2^{(m)}; \\ \asymp_P 1 & j = r_2^{(m)} + 1, \dots, p_m, \end{cases}$$

so that $P(\hat{r}_2^{(m)} = r_2^{(m)}) \rightarrow 1$.

Now for $\hat{k}_r$ (we omit the details for $\hat{k}_c$ since it involves parallel arguments), consider for $i \in [k_r]$,

$$\begin{aligned} \max_{m \in [M]} \max_{n \neq m} \sigma_i^2 \left\{ T^{-1} \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} \right\} &\lesssim \max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m}) \\ &\lesssim \max_{m \in [M]} \max_{n \neq m} \sigma_1^2 \left\{ T^{-1} \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} \right\}. \end{aligned} \quad (\text{C.124})$$

At the same time, for $i \in [k_r]$,

$$\begin{aligned} K_{1,i} - \sum_{j=2}^9 K_j &\leq \sigma_i \left\{ T^{-1} \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} \right\} \leq K_{1,i} + \sum_{j=2}^9 K_j, \quad \text{where} \\ K_{1,i} &:= \sigma_i \left((\mathbf{v}_b^{(m)'} \mathbf{J}_m \otimes \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t)^{\otimes 2} (\mathbf{v}_b^{(n)'} \mathbf{J}_n \otimes \mathbf{Q}_n)' \right), \\ K_2 &:= \sigma_1 \left((\mathbf{v}_b^{(m)'} \mathbf{J}_m \otimes \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t) \text{vec}(\mathbf{F}_t^{(n)})' (\mathbf{v}_b^{(n)'} \mathbf{C}_n \otimes \mathbf{R}_n)' \right), \\ K_3 &:= \sigma_1 \left((\mathbf{v}_b^{(m)'} \mathbf{J}_m \otimes \mathbf{Q}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t) \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right), \\ K_4 &:= \sigma_1 \left((\mathbf{v}_b^{(m)'} \mathbf{C}_m \otimes \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \text{vec}(\mathbf{G}_t)' (\mathbf{v}_b^{(n)'} \mathbf{J}_n \otimes \mathbf{Q}_n)' \right), \\ K_5 &:= \sigma_1 \left((\mathbf{v}_b^{(m)'} \mathbf{C}_m \otimes \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \text{vec}(\mathbf{F}_t^{(n)})' (\mathbf{v}_b^{(n)'} \mathbf{C}_n \otimes \mathbf{R}_n)' \right), \\ K_6 &:= \sigma_1 \left((\mathbf{v}_b^{(m)'} \mathbf{C}_m \otimes \mathbf{R}_m) T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{F}_t^{(m)}) \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right), \\ K_7 &:= \sigma_1 \left(\mathbf{M}_{p_m} \cdot T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \text{vec}(\mathbf{G}_t)' (\mathbf{v}_b^{(n)'} \mathbf{J}_n \otimes \mathbf{Q}_n)' \right), \\ K_8 &:= \sigma_1 \left(\mathbf{M}_{p_m} \cdot T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \text{vec}(\mathbf{F}_t^{(m)})' (\mathbf{v}_b^{(n)'} \mathbf{C}_n \otimes \mathbf{R}_n)' \right), \\ K_9 &:= \sigma_1 \left(\mathbf{M}_{p_m} \cdot T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \cdot \mathbf{M}_{p_n} \right). \end{aligned}$$

To bound the above terms, firstly, we have $\|\mathbf{v}_b^{(m)'} \mathbf{J}_m\| \asymp q_m^{\delta_{j,1}^{(m)}/2}$ and $\|\mathbf{v}_a^{(m)'} \mathbf{Q}_m\| \asymp p_m^{\delta_{q,1}/2}$ by construction,

and hence using (D.14) and Assumption (S1), for $i \in [k_r]$,

$$\begin{aligned} K_{1,i} &\gtrsim_P \lambda_1(\mathbf{J}'_n \mathbf{v}_b^{(n)} \mathbf{v}_b^{(m)} \mathbf{J}'_m) \lambda_i(\mathbf{Q}'_n \mathbf{Q}_m) \lambda_{k_r+k_c} \left(T^{-1} \sum_{t=1}^T \text{vec}(\mathbf{G}_t)^{\otimes 2} \right) \\ &\gtrsim_P q_m^{\delta_{j,1}^{(m)}/2} q_n^{\delta_{j,1}^{(n)}/2} p_m^{\delta_{q,i}^{(m)}/2} p_n^{\delta_{q,k_r}^{(m)}/2}. \end{aligned}$$

Also, using (D.15), (C.115) and (C.116),

$$\begin{aligned} K_2 &= O_P(R(T, d_{\text{fac}, \star})^{1/2} q_m^{\delta_{j,1}^{(m)}/2} q_n^{\delta_{c,1}^{(n)}/2} p_m^{\delta_{q,1}^{(m)}/2} p_n^{\delta_{r,1}^{(n)}/2}), & K_3 &= O_P(T^{-1/2} p_n^{1/2} q_n^{1/2} \cdot q_m^{\delta_{j,1}^{(m)}/2} p_m^{\delta_{q,1}^{(m)}/2}), \\ K_4 &= O_P(R(T, d_{\text{fac}, \star})^{1/2} q_m^{\delta_{c,1}^{(m)}/2} q_n^{\delta_{j,1}^{(n)}/2} p_m^{\delta_{r,1}^{(m)}/2} p_n^{\delta_{q,1}^{(n)}/2}), & K_5 &= O_P(R(T, d_{\text{fac}, \star})^{1/2} q_m^{\delta_{c,1}^{(m)}/2} q_n^{\delta_{c,1}^{(n)}/2} p_m^{\delta_{r,1}^{(m)}/2} p_n^{\delta_{r,1}^{(n)}/2}), \\ K_6 &= O_P(T^{-1/2} p_n^{1/2} q_n^{1/2} \cdot q_m^{\delta_{c,1}^{(m)}/2} p_m^{\delta_{r,1}^{(m)}/2}), & K_7 &= O_P(T^{-1/2} p_m^{1/2} q_m^{1/2} \cdot q_n^{\delta_{j,1}^{(n)}/2} p_n^{\delta_{q,1}^{(n)}/2}), \\ K_8 &= O_P(T^{-1/2} p_m^{1/2} q_m^{1/2} \cdot q_n^{\delta_{c,1}^{(n)}/2} p_n^{\delta_{r,1}^{(n)}/2}), & K_9 &= O_P(T^{-1/2} p_m^{1/2} p_n^{1/2} q_m^{1/2} q_n^{1/2}). \end{aligned}$$

By the assumption in the statement of the theorem, we also have for $i \in [k_r]$,

$$\max_{m,n \in [M], n \neq m, j=2, \dots, 9} K_j^2 = O_P \left(\max_{m,n \in [M], n \neq m} q_m^{\delta_{j,1}^{(m)}} q_n^{\delta_{j,1}^{(n)}} p_m^{\delta_{q,i}^{(m)}} p_n^{\delta_{q,k_r}^{(m)}} \right) \lesssim_P \max_{m,n \in [M], n \neq m} K_{1,i}^2 \asymp_P \max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m}),$$

after using (C.124). On the other hand, for $i \in \{k_r + 1, \dots, p_m\}$,

$$\begin{aligned} \max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m}) &\lesssim_P \max_{m,n \in [M], n \neq m, j=2, \dots, 9} K_j^2 \\ &= O_P \left(\max_{m,n \in [M], n \neq m} (T^{-1} p_m p_n q_m q_n + R(T, d_{\text{fac}, \star}) p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}}) \right). \end{aligned}$$

With these intermediate results, we can conclude that for $i \in [k_r - 1]$,

$$\begin{aligned} \frac{\max_{m \in [M]} \lambda_{i+1}(\mathbf{X}_{Q,m})}{\max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m})} &\gtrsim_P \frac{\max_{m,n \in [M], n \neq m} q_m^{\delta_{j,1}^{(m)}} q_n^{\delta_{j,1}^{(n)}} p_m^{\delta_{q,i+1}^{(m)}} p_n^{\delta_{q,k_r}^{(m)}}}{\max_{m,n \in [M], n \neq m} q_m^{\delta_{j,1}^{(m)}} q_n^{\delta_{j,1}^{(n)}} p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}}} \\ &\succ_P \frac{\max_{m,n \in [M], n \neq m} (T^{-1} p_m p_n q_m q_n + R(T, d_{\text{fac}, \star}) p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} q_n^{\delta_{j,1}^{(n)}})}{\max_{m,n \in [M], n \neq m} q_m^{\delta_{j,1}^{(m)}} q_n^{\delta_{j,1}^{(n)}} p_m^{\delta_{q,k_r}^{(m)}} p_n^{\delta_{q,k_r}^{(m)}}} (= o_P(1)) \\ &\gtrsim_P \frac{\max_{m \in [M]} \lambda_{k_r+1}(\mathbf{X}_{Q,m})}{\max_{m \in [M]} \lambda_{k_r}(\mathbf{X}_{Q,m})}, \end{aligned} \tag{C.125}$$

where the second line is by the rate assumption in the statement of the theorem. The above implies $P(\hat{k}_r \geq k_r) \rightarrow 1$. When all factors are strong and all processes have short memory, then for $i \in \{k_r + 1, \dots, p_m\}$,

$$\max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m}) = O_P \left(\max_{m,n \in [M], n \neq m} T^{-1} p_m p_n q_m q_n \right),$$

so we can easily see that

$$\frac{\max_{m \in [M]} \lambda_{i+1}(\mathbf{X}_{Q,m}) + \max_{m,n \in [M], n \neq m} CT^{-1} p_m p_n q_m q_n}{\max_{m \in [M]} \lambda_i(\mathbf{X}_{Q,m}) + \max_{m,n \in [M], n \neq m} CT^{-1} p_m p_n q_m q_n} \asymp_P 1$$

for $i \in \{k_r + 1, \dots, p_m\}$. This, together with (C.125), proves $P(\hat{k}_r = k_r) \rightarrow 1$.

For $\hat{k}_c^{(m)}$ (the proof for $\hat{k}_r^{(m)}$ runs parallel and is omitted), consider the decomposition of $\mathbf{X}_{C,m}$ (C.45), where we can see from the rate in (C.50) that for $i \in [k_c^{(m)}]$,

$$\lambda_i(\mathbf{X}_{C,m}) \asymp_P q_m^{\delta_{c,i}^{(m)}} p_m^{\delta_{r,1}^{(m)}},$$

while for $i \in \{k_c^{(m)} + 1, \dots, q_m\}$, the rates $\|\mathcal{J}_5\|_F^2$ and $\|\mathcal{J}_6\|_F^2$ from (C.52) and (C.54) respectively dominated, and hence we have

$$\lambda_i(\mathbf{X}_{C,m}) = O_P(q_m^{1-\delta_{c,1}^{(m)}} + T^{-1/2} q_m p_m^{1/2} + q_m^{(1+\delta_{j,1}^{(m)})/2} p_m^{(1+\delta_{q,1}^{(m)})/2} \eta_{Q,m}^{1/2}).$$

With these two intermediate results, we can conclude that for $i \in [k_c^{(m)} - 1]$,

$$\begin{aligned} \frac{\lambda_{i+1}(\mathbf{X}_{C,m})}{\lambda_i(\mathbf{X}_{C,m})} &\gtrsim_P \frac{q_m^{\delta_{c,i+1}^{(m)}} p_m^{\delta_{r,1}^{(m)}}}{q_m^{\delta_{c,1}^{(m)}} p_m^{\delta_{r,1}^{(m)}}} \succ_P \frac{q_m^{1-\delta_{c,1}^{(m)}} + T^{-1/2} q_m p_m^{1/2} + q_m^{(1+\delta_{j,1}^{(m)})/2} p_m^{(1+\delta_{q,1}^{(m)})/2} \eta_{Q,m}^{1/2}}{q_m^{\delta_{c,k_c^{(m)}}^{(m)}} p_m^{\delta_{r,1}^{(m)}}} (= o_P(1)) \\ &\gtrsim_P \frac{\lambda_{k_c^{(m)}+1}(\mathbf{X}_{C,m})}{\lambda_{k_c^{(m)}}(\mathbf{X}_{C,m})}, \end{aligned} \quad (\text{C.126})$$

where the second line is by the rate assumption in the statement of the theorem. The above implies $P(\hat{k}_c^{(m)} \geq k_c^{(m)}) \rightarrow 1$. When all factors are strong and all processes have short memory, $\eta_{Q,m} = T^{-1}$, and hence for $i \in \{k_c^{(m)} + 1, \dots, q_m\}$,

$$\lambda_i(\mathbf{X}_{C,m}) = O_P(T^{-1/2} p_m q_m).$$

Hence, we can see that for $i \in \{k_c^{(m)} + 1, \dots, q_m\}$,

$$\frac{\lambda_{i+1}(\mathbf{X}_{C,m}) + CT^{-1/2} p_m q_m}{\lambda_i(\mathbf{X}_{C,m}) + CT^{-1/2} p_m q_m} \asymp_P 1.$$

This, together with (C.126), proves $P(\hat{k}_c^{(m)} = k_c^{(m)}) \rightarrow 1$. $\square$

D Auxiliary Results and Proofs

Proposition 1 (*Contraction properties of orthogonal projectors*) Let $\mathbf{A} \in \mathbb{R}^{a \times b}$ be any matrix, and let $\mathbf{P} \in \mathbb{R}^{a \times a}$ and $\mathbf{Q} \in \mathbb{R}^{b \times b}$ be orthogonal projectors. Then,

$$\|\mathbf{PAQ}\| \leq \|\mathbf{A}\|, \quad \|\mathbf{PAQ}\|_F \leq \|\mathbf{A}\|_F. \quad (\text{D.1})$$

In particular, for any positive integer n , define $\mathbf{M}_n := \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n'$. Then, $\mathbf{M}_n$ is the orthogonal projector onto $\mathbf{1}_n^\perp$, and

$$\mathbf{M}_n' = \mathbf{M}_n, \quad \mathbf{M}_n^2 = \mathbf{M}_n, \quad \mathbf{M}_n\mathbf{1}_n = \mathbf{0}, \quad \|\mathbf{M}_n\| \leq 1. \quad (\text{D.2})$$

Moreover,

$$\|\mathbf{M}_n\|_{\max} \leq 1, \quad \|\mathbf{M}_n\|_1 = \|\mathbf{M}_n\|_\infty = 2(1 - n^{-1}) \leq 2. \quad (\text{D.3})$$

Proof of Proposition 1. Since $\mathbf{P}$ and $\mathbf{Q}$ are orthogonal projectors, $\|\mathbf{P}\| \leq 1$ and $\|\mathbf{Q}\| \leq 1$. Hence,

$$\begin{aligned} \|\mathbf{PAQ}\| &\leq \|\mathbf{P}\| \|\mathbf{A}\| \|\mathbf{Q}\| \leq \|\mathbf{A}\|, \\ \|\mathbf{PAQ}\|_F &\leq \|\mathbf{P}\| \|\mathbf{A}\|_F \|\mathbf{Q}\| \leq \|\mathbf{A}\|_F. \end{aligned}$$

By direct calculation, $\mathbf{M}_n' = \mathbf{M}_n$, $\mathbf{M}_n^2 = \mathbf{M}_n$, and $\mathbf{M}_n\mathbf{1}_n = \mathbf{0}$. Thus, $\mathbf{M}_n$ is the orthogonal projector onto $\mathbf{1}_n^\perp$. Since the eigenvalues of an orthogonal projector belong to $\{0, 1\}$, we have $\|\mathbf{M}_n\| \leq 1$. Finally, its diagonal entries are $1 - 1/n$ and its off-diagonal entries are $-1/n$. Therefore,

$$\|\mathbf{M}_n\|_{\max} \leq 1, \quad \|\mathbf{M}_n\|_1 = \|\mathbf{M}_n\|_\infty = (1 - n^{-1}) + (n - 1)n^{-1} = 2(1 - n^{-1}) \leq 2.$$

This completes the proof. $\square$

Lemma 1 (*Covariance bounds for $\mathbf{E}_t^{(m)}$*) Let Assumptions (E1) and (E2) hold. Then there exist a non-negative sequence $\{\rho(h) : h \in \mathbb{Z}\}$ and a constant $C < \infty$, such that $\sum_{h \in \mathbb{Z}} \rho(h) < \infty$, $\sup_{h \in \mathbb{Z}} \rho(h) \leq 1$, and, uniformly over $h \in \mathbb{Z}$ and all t such that $t, t + h \in [T]$,

$$\max_{m \in [M]} \sup_{i \in [p_m]} \sup_{j \in [q_m]} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} \left| \mathbb{E}(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)}) \right| \leq C\rho(h). \quad (\text{D.4})$$

In particular, uniformly over all t such that $t, t+h \in [T]$,

$$\max_{m \in [M]} \sup_{h \in \mathbb{Z}} \sup_{i \in [p_m]} \sup_{j \in [q_m]} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} \left| \mathbb{E}(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)}) \right| \leq C. \quad (\text{D.5})$$

Moreover, uniformly over $h \in \mathbb{Z}$, all t such that $t, t+h \in [T]$, the covariance of the error products satisfies

$$\max_{m,n \in [M]} \sup_{i \in [p_m]} \sup_{j \in [q_m]} \sup_{k \in [p_n]} \sup_{l \in [q_n]} \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(n)}, E_{t,ia}^{(m)} E_{t,kb}^{(n)} \right) \right| \leq C\rho(h). \quad (\text{D.6})$$

Proof of Lemma 1. For $h \in \mathbb{Z}$, let $\gamma_e(h) := \sum_{w \geq 0} |a_{e,w} a_{e,w+|h|}|$, and for each $m \in [M]$, define analogously $\gamma_e^{(m)}(h) := \sum_{w \geq 0} |a_{e,w}^{(m)} a_{e,w+|h|}^{(m)}|$ and $\gamma_\epsilon(h)^{(m)} := \sum_{w \geq 0} |a_{\epsilon,w}^{(m)} a_{\epsilon,w+|h|}^{(m)}|$. By (E2),

$$\begin{aligned} \sum_{h \in \mathbb{Z}} \gamma_e(h) &= \sum_{w \geq 0} |a_{e,w}|^2 + 2 \sum_{h=1}^{\infty} \sum_{w \geq 0} |a_{e,w}| |a_{e,w+h}| \\ &= \sum_{w \geq 0} |a_{e,w}|^2 + 2 \sum_{0 \leq w < r < \infty} |a_{e,w}| |a_{e,r}| = \left(\sum_{w \geq 0} |a_{e,w}| \right)^2 \leq c^2 < \infty, \end{aligned}$$

and analogously, for every $m \in [M]$, $\sum_{h \in \mathbb{Z}} \gamma_e^{(m)}(h) \leq c^2$ and $\sum_{h \in \mathbb{Z}} \gamma_\epsilon^{(m)}(h) \leq c^2$, where c is the constant in (E2). Moreover, by the Cauchy-Schwarz inequality and (E2),

$$\gamma_e(h) \leq \left(\sum_{w \geq 0} a_{e,w}^2 \right)^{1/2} \left(\sum_{w \geq 0} a_{e,w+|h|}^2 \right)^{1/2} \leq 1, \quad (\text{D.7})$$

so that $\sup_h \gamma_e(h) \leq 1$. The same argument gives $\sup_h \gamma_e^{(m)}(h) \leq 1$ and $\sup_h \gamma_\epsilon^{(m)}(h) \leq 1$. Since M is fixed, define

$$\rho(h) := \frac{1}{3} \left\{ \gamma_e(h) + \max_{m \in [M]} \gamma_e^{(m)}(h) + \max_{m \in [M]} \gamma_\epsilon^{(m)}(h) \right\}.$$

Then $\rho(h) \geq 0$, $\sup_{h \in \mathbb{Z}} \rho(h) \leq 1$, and

$$\sum_{h \in \mathbb{Z}} \rho(h) \leq \frac{1}{3} \left\{ \sum_{h \in \mathbb{Z}} \gamma_e(h) + \sum_{m=1}^M \sum_{h \in \mathbb{Z}} \gamma_e^{(m)}(h) + \sum_{m=1}^M \sum_{h \in \mathbb{Z}} \gamma_\epsilon^{(m)}(h) \right\} \leq \frac{1+2M}{3} c^2 < \infty.$$

We next prove (D.4). Fix $m \in [M]$. By (E1), write

$$E_{t,ij}^{(m)} = (\mathbf{Q}_{e,m} \mathbf{G}_{e,t} \mathbf{J}_{e,m}')_{ij} + (\mathbf{R}_{e,m} \mathbf{F}_{e,t}^{(m)} \mathbf{C}_{e,m}')_{ij} + (\mathbf{\Sigma}_{\epsilon,m} \circ \boldsymbol{\epsilon}_t^{(m)})_{ij} =: U_{t,ij}^{(m)} + V_{t,ij}^{(m)} + W_{t,ij}^{(m)}. \quad (\text{D.8})$$

Since the three component processes are mutually independent and have mean zero under (E1) and (E2),

we have

$$\mathbb{E}(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)}) = \mathbb{E}(U_{t+h,ij}^{(m)} U_{t,kl}^{(m)}) + \mathbb{E}(V_{t+h,ij}^{(m)} V_{t,kl}^{(m)}) + \mathbb{E}(W_{t+h,ij}^{(m)} W_{t,kl}^{(m)}). \quad (\text{D.9})$$

We bound the three terms on the r.h.s. separately. First,

$$U_{t,ij}^{(m)} = \sum_{a=1}^{k_{e,r}} \sum_{b=1}^{k_{e,c}} (\mathbf{Q}_{e,m})_{ia} (\mathbf{G}_{e,t})_{ab} (\mathbf{J}_{e,m})_{jb}.$$

Hence

$$\mathbb{E}(U_{t+h,ij}^{(m)} U_{t,kl}^{(m)}) = \sum_{a=1}^{k_{e,r}} \sum_{b=1}^{k_{e,c}} \sum_{c=1}^{k_{e,r}} \sum_{d=1}^{k_{e,c}} (\mathbf{Q}_{e,m})_{ia} (\mathbf{Q}_{e,m})_{kc} (\mathbf{J}_{e,m})_{jb} (\mathbf{J}_{e,m})_{ld} \mathbb{E}\{(\mathbf{G}_{e,t+h})_{ab} (\mathbf{G}_{e,t})_{cd}\}.$$

By (E2), we have $\mathbb{E}\{(\mathbf{X}_{e,r})_{ab} (\mathbf{X}_{e,s})_{cd}\} = \mathbb{1}\{r = s\} \mathbb{1}\{a = c, b = d\}$, so that

$$\begin{aligned} \mathbb{E}\{(\mathbf{G}_{e,t+h})_{ab} (\mathbf{G}_{e,t})_{cd}\} &= \sum_{u \geq 0} \sum_{v \geq 0} a_{e,u} a_{e,v} \mathbb{E}\{(\mathbf{X}_{e,t+h-u})_{ab} (\mathbf{X}_{e,t-v})_{cd}\} \\ &= \mathbb{1}\{a = c, b = d\} \sum_{u \geq 0} \sum_{v \geq 0} a_{e,u} a_{e,v} \mathbb{1}\{t+h-u = t-v\} \\ &= \mathbb{1}\{a = c, b = d\} \sum_{w \geq 0} a_{e,w+|h|} a_{e,w}, \end{aligned}$$

where the last equality follows by setting $w = v$ when $h \geq 0$ and $w = u$ when $h < 0$. Hence,

$$|\mathbb{E}\{(\mathbf{G}_{e,t+h})_{ab} (\mathbf{G}_{e,t})_{cd}\}| \leq \gamma_e(h) \mathbb{1}\{a = c, b = d\}.$$

Therefore,

$$|\mathbb{E}(U_{t+h,ij}^{(m)} U_{t,kl}^{(m)})| \leq \gamma_e(h) \sum_{a=1}^{k_{e,r}} \sum_{b=1}^{k_{e,c}} |(\mathbf{Q}_{e,m})_{ia}| |(\mathbf{Q}_{e,m})_{ka}| |(\mathbf{J}_{e,m})_{jb}| |(\mathbf{J}_{e,m})_{lb}|.$$

Taking absolute values and summing over k and l , we obtain

$$\begin{aligned} &\sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(U_{t+h,ij}^{(m)} U_{t,kl}^{(m)})| \\ &\leq \gamma_e(h) \sum_{a=1}^{k_{e,r}} \sum_{b=1}^{k_{e,c}} |(\mathbf{Q}_{e,m})_{ia}| |(\mathbf{J}_{e,m})_{jb}| \left(\sum_{k=1}^{p_m} |(\mathbf{Q}_{e,m})_{ka}| \right) \left(\sum_{l=1}^{q_m} |(\mathbf{J}_{e,m})_{lb}| \right) \\ &\leq \gamma_e(h) \left(\sum_{a=1}^{k_{e,r}} |(\mathbf{Q}_{e,m})_{ia}| \right) \left(\max_a \sum_{k=1}^{p_m} |(\mathbf{Q}_{e,m})_{ka}| \right) \left(\sum_{b=1}^{k_{e,c}} |(\mathbf{J}_{e,m})_{jb}| \right) \left(\max_b \sum_{l=1}^{q_m} |(\mathbf{J}_{e,m})_{lb}| \right) \\ &\leq \gamma_e(h) (k_{e,r} \|\mathbf{Q}_{e,m}\|_1) \|\mathbf{Q}_{e,m}\|_1 (k_{e,c} \|\mathbf{J}_{e,m}\|_1) \|\mathbf{J}_{e,m}\|_1 = O(\gamma_e(h)), \end{aligned}$$

where we use $\sum_k |(\mathbf{Q}_{e,m})_{ka}| \leq \|\mathbf{Q}_{e,m}\|_1$ and $\sum_l |(\mathbf{J}_{e,m})_{lb}| \leq \|\mathbf{J}_{e,m}\|_1$ in the first inequality, $\sum_a |(\mathbf{Q}_{e,m})_{ia}| \leq k_{e,r} \|\mathbf{Q}_{e,m}\|_1$ and $\sum_b |(\mathbf{J}_{e,m})_{jb}| \leq k_{e,c} \|\mathbf{J}_{e,m}\|_1$ in the second inequality, and Assumption (E1) in the last equality. By the same argument,

$$\sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(V_{t+h,ij}^{(m)} V_{t,kl}^{(m)})| = O(\gamma_e^{(m)}(h)).$$

Finally, (E2) gives

$$|\mathbb{E}(\epsilon_{t+h,ij}^{(m)} \epsilon_{t,kl}^{(m)})| \leq \gamma_\epsilon^{(m)}(h) \mathbb{1}\{(i, j) = (k, l)\}.$$

Therefore,

$$\begin{aligned} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(W_{t+h,ij}^{(m)} W_{t,kl}^{(m)})| &\leq \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |(\boldsymbol{\Sigma}_{\epsilon,m})_{ij} (\boldsymbol{\Sigma}_{\epsilon,m})_{kl} \gamma_\epsilon^{(m)}(h) \mathbb{1}\{(i, j) = (k, l)\}| \\ &= \gamma_\epsilon^{(m)}(h) (\boldsymbol{\Sigma}_{\epsilon,m})_{ij}^2 = O(\gamma_\epsilon^{(m)}(h)), \end{aligned}$$

by the uniform boundedness of $\boldsymbol{\Sigma}_{\epsilon,m}$ in (E1). Combining the above bounds, there exists an absolute constant C such that

$$\begin{aligned} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(U_{t+h,ij}^{(m)} U_{t,kl}^{(m)})| &\leq \frac{C}{3} \gamma_e(h), \\ \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(V_{t+h,ij}^{(m)} V_{t,kl}^{(m)})| &\leq \frac{C}{3} \gamma_e^{(m)}(h), \\ \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(W_{t+h,ij}^{(m)} W_{t,kl}^{(m)})| &\leq \frac{C}{3} \gamma_\epsilon^{(m)}(h), \end{aligned}$$

uniformly over $m \in [M]$, $h \in \mathbb{Z}$, $i \in [p_m]$, and $j \in [q_m]$. By (D.9) and the triangle inequality,

$$\begin{aligned} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)})| &\leq \frac{C}{3} (\gamma_e(h) + \gamma_e^{(m)}(h) + \gamma_\epsilon^{(m)}(h)) \\ &\leq \frac{C}{3} \left(\gamma_e(h) + \max_{m \in [M]} \gamma_e^{(m)}(h) + \max_{m \in [M]} \gamma_\epsilon^{(m)}(h) \right) = C\rho(h), \end{aligned}$$

which concludes (D.4). Since $\sup_{h \in \mathbb{Z}} \rho(h) \leq 1$, (D.5) follows immediately.

We next show (D.6). For $h \in \mathbb{Z}$, define $\kappa_e(h) := \sum_{w \geq 0} a_{e,w}^2 a_{e,w+|h|}^2$, and for each $m \in [M]$, define $\kappa_e^{(m)}(h) := \sum_{w \geq 0} (a_{e,w}^{(m)})^2 (a_{e,w+|h|}^{(m)})^2$ and $\kappa_\epsilon^{(m)}(h) := \sum_{w \geq 0} (a_{\epsilon,w}^{(m)})^2 (a_{\epsilon,w+|h|}^{(m)})^2$. Since

$$\kappa_e(h) \leq \gamma_e(h)^2 \leq \gamma_e(h) \leq 3\rho(h),$$

and analogously $\kappa_e^{(m)}(h) \leq 3\rho(h)$ and $\kappa_\epsilon^{(m)}(h) \leq 3\rho(h)$ for every $m \in [M]$, we have

$$\kappa_e(h) + \max_{m \in [M]} \kappa_e^{(m)}(h) + \max_{m \in [M]} \kappa_\epsilon^{(m)}(h) \leq 9\rho(h).$$

Moreover, for any two sequences γ_1 and γ_2 chosen from $\{\gamma_e, \gamma_e^{(m)}, \gamma_\epsilon^{(m)} : m \in [M]\}$, we have

$$\gamma_1(h)\gamma_2(h) \leq 9\rho(h)^2 \leq 9\rho(h),$$

where the second inequality used $\rho(h) \leq 1$.

Recall the decomposition in (D.8). Expanding each of the four error entries in the covariance yields finitely many terms. We first consider

$$S_U(h) := \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(U_{t+h,ij}^{(m)}, U_{t+h,kl}^{(n)}, U_{t,ia}^{(m)} U_{t,kb}^{(n)} \right) \right|.$$

By the definition of $U_{t,ij}^{(m)}$,

$$U_{t,ij}^{(m)} = \sum_{c=1}^{k_{e,r}} \sum_{d=1}^{k_{e,c}} (\mathbf{Q}_{e,m})_{ic} (\mathbf{J}_{e,m})_{jd} (\mathbf{G}_{e,t})_{cd}.$$

It follows that

$$\begin{aligned} & \text{Cov} \left(U_{t+h,ij}^{(m)}, U_{t+h,kl}^{(n)}, U_{t,ia}^{(m)} U_{t,kb}^{(n)} \right) \\ &= \sum_{c_1=1}^{k_{e,r}} \sum_{d_1=1}^{k_{e,c}} \sum_{c_2=1}^{k_{e,r}} \sum_{d_2=1}^{k_{e,c}} \sum_{c_3=1}^{k_{e,r}} \sum_{d_3=1}^{k_{e,c}} \sum_{c_4=1}^{k_{e,r}} \sum_{d_4=1}^{k_{e,c}} (\mathbf{Q}_{e,m})_{ic_1} (\mathbf{J}_{e,m})_{jd_1} (\mathbf{Q}_{e,n})_{kc_2} (\mathbf{J}_{e,n})_{ld_2} \\ & \quad \cdot (\mathbf{Q}_{e,m})_{ic_3} (\mathbf{J}_{e,m})_{ad_3} (\mathbf{Q}_{e,n})_{kc_4} (\mathbf{J}_{e,n})_{bd_4} \\ & \quad \cdot \text{Cov} \left((\mathbf{G}_{e,t+h})_{c_1 d_1}, (\mathbf{G}_{e,t+h})_{c_2 d_2}, (\mathbf{G}_{e,t})_{c_3 d_3}, (\mathbf{G}_{e,t})_{c_4 d_4} \right). \end{aligned}$$

By (E2), we have

$$\begin{aligned} & \left| \text{Cov} \left((\mathbf{G}_{e,t+h})_{c_1 d_1}, (\mathbf{G}_{e,t+h})_{c_2 d_2}, (\mathbf{G}_{e,t})_{c_3 d_3}, (\mathbf{G}_{e,t})_{c_4 d_4} \right) \right| \\ & \leq C\gamma_e(h)^2 \mathbb{1}\{(c_1, d_1) = (c_3, d_3), (c_2, d_2) = (c_4, d_4)\} \\ & \quad + C\gamma_e(h)^2 \mathbb{1}\{(c_1, d_1) = (c_4, d_4), (c_2, d_2) = (c_3, d_3)\} \\ & \quad + C\kappa_e(h) \mathbb{1}\{(c_1, d_1) = (c_2, d_2) = (c_3, d_3) = (c_4, d_4)\}. \end{aligned}$$

Substituting the above into the preceding expansion and using the triangle inequality, we obtain

$$\begin{aligned}
S_U(h) &\leq C\gamma_e(h)^2 \sum_{c_1=1}^{k_{e,r}} \sum_{d_1=1}^{k_{e,c}} \sum_{c_2=1}^{k_{e,r}} \sum_{d_2=1}^{k_{e,c}} |(\mathbf{Q}_{e,m})_{ic_1}(\mathbf{J}_{e,m})_{jd_1}(\mathbf{Q}_{e,n})_{kc_2}(\mathbf{J}_{e,n})_{ld_2}| \\
&\quad \cdot \sum_{a=1}^{q_m} |(\mathbf{Q}_{e,m})_{ic_1}(\mathbf{J}_{e,m})_{ad_1}| \sum_{b=1}^{q_n} |(\mathbf{Q}_{e,n})_{kc_2}(\mathbf{J}_{e,n})_{bd_2}| \\
&\quad + C\gamma_e(h)^2 \sum_{c_1=1}^{k_{e,r}} \sum_{d_1=1}^{k_{e,c}} \sum_{c_2=1}^{k_{e,r}} \sum_{d_2=1}^{k_{e,c}} |(\mathbf{Q}_{e,m})_{ic_1}(\mathbf{J}_{e,m})_{jd_1}(\mathbf{Q}_{e,n})_{kc_2}(\mathbf{J}_{e,n})_{ld_2}| \\
&\quad \cdot \sum_{a=1}^{q_m} |(\mathbf{Q}_{e,m})_{ic_2}(\mathbf{J}_{e,m})_{ad_2}| \sum_{b=1}^{q_n} |(\mathbf{Q}_{e,n})_{kc_1}(\mathbf{J}_{e,n})_{bd_1}| \\
&\quad + C\kappa_e(h) \sum_{c=1}^{k_{e,r}} \sum_{d=1}^{k_{e,c}} |(\mathbf{Q}_{e,m})_{ic}(\mathbf{J}_{e,m})_{jd}(\mathbf{Q}_{e,n})_{kc}(\mathbf{J}_{e,n})_{ld}| \\
&\quad \cdot \sum_{a=1}^{q_m} |(\mathbf{Q}_{e,m})_{ic}(\mathbf{J}_{e,m})_{ad}| \sum_{b=1}^{q_n} |(\mathbf{Q}_{e,n})_{kc}(\mathbf{J}_{e,n})_{bd}|.
\end{aligned}$$

Since $k_{e,r}$ and $k_{e,c}$ are fixed, and since $\|\mathbf{Q}_{e,m}\|_1 = \|\mathbf{J}_{e,m}\|_1 = O(1)$ and $\|\mathbf{Q}_{e,n}\|_1 = \|\mathbf{J}_{e,n}\|_1 = O(1)$ for all $m, n \in [M]$, we have, uniformly over $i \in [p_m]$, $j \in [q_m]$, $k \in [p_n]$, and $l \in [q_n]$,

$$S_U(h) \leq C\gamma_e(h)^2 + C\kappa_e(h) \leq C\rho(h). \quad (\text{D.10})$$

Next, define

$$S_V(h) := \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(V_{t+h,ij}^{(m)} V_{t+h,kl}^{(n)}, V_{t,ia}^{(m)} V_{t,kb}^{(n)} \right) \right|.$$

When $m = n$, the above arguments apply to the local error process. When $m \neq n$, the cross-group independence of the local error innovation processes in (E2) yields, uniformly over $i \in [p_m]$, $j \in [q_m]$, $k \in [p_n]$, and $l \in [q_n]$,

$$S_V(h) \leq C\gamma_e^{(m)}(h)\gamma_e^{(n)}(h) + C\mathbb{1}\{m = n\}\kappa_e^{(m)}(h) \leq C\rho(h). \quad (\text{D.11})$$

Similarly, define

$$S_W(h) := \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(W_{t+h,ij}^{(m)} W_{t+h,kl}^{(n)}, W_{t,ia}^{(m)} W_{t,kb}^{(n)} \right) \right|.$$

Using the independence and uniformly bounded fourth moments of the idiosyncratic innovations in (E2),

together with the uniform boundedness of the entries of $\Sigma_{\epsilon, m}$ in (E1), we obtain

$$S_W(h) \leq C\gamma_\epsilon^{(m)}(h)\gamma_\epsilon^{(n)}(h) + C\mathbb{1}\{m = n\}\kappa_\epsilon^{(m)}(h) \leq C\rho(h). \quad (\text{D.12})$$

It remains to bound the terms involving at least two components among $U_{t,ij}^{(m)}$, $V_{t,ij}^{(m)}$, and $W_{t,ij}^{(m)}$. These remaining terms fall into the following two cases.

(i) At least one driving innovation process appears only once in the covariance. Such terms have zero covariance. For example,

$$\begin{aligned} & \text{Cov}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}, U_{t,ia}^{(m)}V_{t,kb}^{(n)}\right) \\ &= \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}U_{t,ia}^{(m)}V_{t,kb}^{(n)}\right) - \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}\right)\mathbb{E}\left(U_{t,ia}^{(m)}V_{t,kb}^{(n)}\right) \\ &= \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}U_{t,ia}^{(m)}\right)\mathbb{E}\left(V_{t,kb}^{(n)}\right) - \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}\right)\mathbb{E}\left(U_{t,ia}^{(m)}\right)\mathbb{E}\left(V_{t,kb}^{(n)}\right) = 0. \end{aligned}$$

The same argument applies whenever one of the driving innovation processes occurs only once. For example,

$$\text{Cov}\left(U_{t+h,ij}^{(m)}V_{t+h,kl}^{(n)}, U_{t,ia}^{(m)}W_{t,kb}^{(n)}\right) = 0.$$

(ii) Two independent driving innovation processes each appear twice. There are two possible arrangements.

(a) *The two occurrences of each process lie on the same side of the covariance.* In this case, the two products are independent, so their covariance is zero. For example,

$$\text{Cov}\left(U_{t+h,ij}^{(m)}U_{t+h,kl}^{(n)}, V_{t,ia}^{(m)}V_{t,kb}^{(n)}\right) = 0.$$

(b) *Each process contributes once on each side of the covariance.* We first consider the arrangements in which each innovation process contributes one factor at time $t+h$ and one factor at time t . As a representative case, write

$$\begin{aligned} S_{UV}(h) &:= \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov}\left(U_{t+h,ij}^{(m)}V_{t+h,kl}^{(n)}, U_{t,ia}^{(m)}V_{t,kb}^{(n)}\right) \right| \\ &= \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t,ia}^{(m)}\right)\mathbb{E}\left(V_{t+h,kl}^{(n)}V_{t,kb}^{(n)}\right) \right| \\ &\leq \left\{ \sum_{a=1}^{q_m} \left| \mathbb{E}\left(U_{t+h,ij}^{(m)}U_{t,ia}^{(m)}\right) \right| \right\} \left\{ \sum_{b=1}^{q_n} \left| \mathbb{E}\left(V_{t+h,kl}^{(n)}V_{t,kb}^{(n)}\right) \right| \right\} \\ &\leq C\gamma_e(h)\gamma_e^{(n)}(h) \leq C\rho(h). \end{aligned}$$

We finally consider the arrangement in which the two component types exchange their group assignments across the two sides of the covariance. For example,

$$\text{Cov} \left(U_{t+h,ij}^{(m)} V_{t+h,kl}^{(n)}, V_{t,ia}^{(m)} U_{t,kb}^{(n)} \right).$$

If $m \neq n$, then the covariance is zero by case (i). If $m = n$, uniformly over $m, n \in [M]$, independence of the global and local innovation processes gives

$$\begin{aligned} & \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(U_{t+h,ij}^{(m)} V_{t+h,kl}^{(n)}, V_{t,ia}^{(m)} U_{t,kb}^{(n)} \right) \right| \\ &= \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \mathbb{E} \left(U_{t+h,ij}^{(m)} U_{t,kb}^{(m)} \right) \mathbb{E} \left(V_{t+h,kl}^{(m)} V_{t,ia}^{(m)} \right) \right| \\ &\leq C \gamma_e(h) \gamma_e^{(m)}(h) \leq C \rho(h). \end{aligned}$$

Cases (i) and (ii) cover all terms involving at least two of $U_{t,ij}^{(m)}$, $V_{t,ij}^{(m)}$, and $W_{t,ij}^{(m)}$, where these terms are not necessarily corresponding to the same group m . Combining these bounds with (D.10), (D.11) and (D.12), we obtain, uniformly over $h \in \mathbb{Z}$ and all t such that $t, t+h \in [T]$:

$$\max_{m,n \in [M]} \sup_{i \in [p_m]} \sup_{j \in [q_m]} \sup_{k \in [p_n]} \sup_{l \in [q_n]} \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(n)}, E_{t,ia}^{(m)} E_{t,kb}^{(n)} \right) \right| \leq C \rho(h).$$

This proves (D.6) and completes the proof. $\square$

Lemma 2 (*Basic properties of $\mathcal{G}_T$ processes*) Let $\{\mathbf{u}_t\}_{t \in [T]} \in \mathcal{G}_T(\{d_{u,1}, \dots, d_{u,m_u}\} | \{\mathbf{x}_{u,i}, i \in \mathbb{Z}\})$ be a p_u -dimensional process, where p_u and m_u are bounded. Define $d_u := \max_{1 \leq j \leq m_u} d_{u,j}$. Then:

1. There exists a constant $C < \infty$ such that the second moments are uniformly bounded as

$$\sup_{t \in [T]} \mathbb{E} \|\mathbf{u}_t\|^2 \leq C. \quad (\text{D.13})$$

2. For any $h \in \{0, \dots, T-1\}$, the centered sample autocovariance satisfies

$$\left\| T^{-1} \sum_{t=1}^{T-h} (\mathbf{u}_t \mathbf{u}_{t+h}' - \mathbb{E}(\mathbf{u}_t \mathbf{u}_{t+h}')) \right\|_F^2 = O_P(R(T, d_u)). \quad (\text{D.14})$$

3. Let $\{\mathbf{v}_t\}_{t \in [T]} \in \mathcal{G}_T(\{d_{v,1}, \dots, d_{v,m_v}\} | \{\mathbf{x}_{v,i}, i \in \mathbb{Z}\})$ be another p_v -dimensional process, independent of $\{\mathbf{u}_t\}_{t \in [T]}$, where p_v and m_v are bounded. Define $d_v := \max_{1 \leq j \leq m_v} d_{v,j}$. For $d_{uv} := (d_u + d_v)/2$, the

sample cross-moment satisfies

$$\left\| T^{-1} \sum_{t=1}^T \mathbf{u}_t \mathbf{v}_t' \right\|_F^2 = O_P(R(T, d_{uv})). \quad (\text{D.15})$$

Proof of Lemma 2. We first show (D.13). Fix $t \in [T]$, and suppose that t belongs to segment $a \in [m_u]$ of $\{\mathbf{u}_t\}$. By the definition of $\mathcal{G}_T(\cdot)$,

$$\mathbf{u}_t = \sum_{w \geq 0} \psi_{u,a}(w) \mathbf{x}_{u,t-w},$$

where $\psi_{u,a}(0) = 1$, and for $w \geq 1$,

$$\psi_{u,a}(w) = w^{d_{u,a}-1} l_{u,a}(w), \quad d_{u,a} < 1/2.$$

Since $l_{u,a}(w)$ converges to a finite constant and m_u is bounded,

$$|\psi_{u,a}(w)| \lesssim w^{d_{u,a}-1}, \quad w \geq 1.$$

Since the entries of the innovation sequence $\{\mathbf{x}_{u,t}\}$ are independent across coordinates and over time and have zero means and uniformly bounded second moments,

$$\begin{aligned} \mathbb{E} \|\mathbf{u}_t\|^2 &= \sum_{w \geq 0} \psi_{u,a}(w)^2 \mathbb{E} \|\mathbf{x}_{u,t-w}\|^2 \\ &\lesssim p_u \sum_{w \geq 0} \psi_{u,a}(w)^2 \lesssim p_u \left(1 + \sum_{w \geq 1} w^{2d_{u,a}-2} \right) \lesssim 1, \end{aligned}$$

because $d_{u,a} < 1/2$, while p_u and m_u are bounded. Taking the supremum over $t \in [T]$ gives (D.13).

We next show (D.14). Fix $h \in \{0, \dots, T-1\}$, and define

$$\tilde{\mathbf{\Gamma}}_u(h) := T^{-1} \sum_{t=1}^{T-h} (\mathbf{u}_t \mathbf{u}_{t+h}' - \mathbb{E}(\mathbf{u}_t \mathbf{u}_{t+h}')).$$

Since p_u is bounded, it is enough to bound the variance of each entry. The proof proceeds in three steps. We first control the covariance between observations associated with arbitrary segment pairs. We then bound the covariance between the corresponding products. Finally, we group the pairs of time indices according to their lag.

First, fix an entry (i, j) . For $t, s \in [T - h]$, let

$$k = J_u(t), \quad r = J_u(t + h), \quad k' = J_u(s), \quad r' = J_u(s + h),$$

where $J_u(t) \in [m_u]$ denotes the segment containing time t , and write $\tau = t - s$. Then

$$\text{Var}(\tilde{\Gamma}_{u,ij}(h)) = T^{-2} \sum_{t=1}^{T-h} \sum_{s=1}^{T-h} \text{Cov}(u_{t,i}u_{t+h,j}, u_{s,i}u_{s+h,j}). \quad (\text{D.16})$$

Because all coordinates use the same filter within each segment and the innovation coordinates are identically distributed, the covariance between two observations from the same coordinate index depends only on their segment labels and time lag. For $k, k' \in [m_u]$ and $\tau \geq 0$, define

$$\gamma_{k,k'}(\tau) := \sigma_u^2 \sum_{w \geq 0} \psi_{u,k}(w + \tau) \psi_{u,k'}(w), \quad (\text{D.17})$$

where $\sigma_u^2 := \mathbb{E}(x_{u,0,1}^2) < \infty$, and, for $\tau < 0$, set $\gamma_{k,k'}(\tau) := \gamma_{k',k}(-\tau)$. By the temporal independence of the innovations, whenever $J_u(t) = k$, $J_u(s) = k'$, and $t - s = \tau$, it holds that $\text{Cov}(u_{t,i}, u_{s,i}) = \gamma_{k,k'}(\tau)$ for every $i \in [p_u]$. For each segment $a \in [m_u]$, since each $l_{u,a}(w)$ converges to a finite constant, we have

$$|\psi_{u,a}(w)| \lesssim w^{d_{u,a}-1}, \quad w \geq 1. \quad (\text{D.18})$$

Moreover, by the Cauchy–Schwarz inequality and the square summability of the filters,

$$\max_{k,k' \in [m_u]} |\gamma_{k,k'}(0)| \lesssim \max_{k,k' \in [m_u]} \left\{ \sum_{w \geq 0} \psi_{u,k}(w)^2 \right\}^{1/2} \left\{ \sum_{w \geq 0} \psi_{u,k'}(w)^2 \right\}^{1/2} \lesssim 1.$$

Suppose first that $d_u > 0$. For $\tau \geq 1$, since $d_{u,a} \leq d_u$ for every $a \in [m_u]$,

$$\begin{aligned} |\gamma_{k,k'}(\tau)| &\lesssim \tau^{d_u-1} + \sum_{w \geq 1} (w + \tau)^{d_u-1} w^{d_u-1} \\ &\lesssim \tau^{d_u-1} + \tau^{d_u-1} \sum_{w=1}^{\tau} w^{d_u-1} + \sum_{w > \tau} w^{2d_u-2}. \end{aligned}$$

Since $0 < d_u < 1/2$, integral comparison gives

$$\sum_{w=1}^{\tau} w^{d_u-1} \leq 1 + \int_1^{\tau} x^{d_u-1} dx = 1 + \frac{\tau^{d_u} - 1}{d_u} \lesssim \tau^{d_u},$$

and

$$\sum_{w>\tau} w^{2d_u-2} \leq \int_{\tau}^{\infty} x^{2d_u-2} dx \lesssim \tau^{2d_u-1}.$$

Combining these bounds yields $|\gamma_{k,k'}(\tau)| \lesssim \tau^{2d_u-1}$ for $\tau \geq 1$. Hence,

$$\max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)|^2 \lesssim (1+|\tau|)^{4d_u-2}, \quad \tau \in \mathbb{Z}. \quad (\text{D.19})$$

Therefore, by the standard partial-sum bounds for p -series,

$$\sum_{\tau=-2(T-1)}^{2(T-1)} \max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)|^2 \lesssim 1 + \sum_{\tau=1}^{2(T-1)} \tau^{4d_u-2} \lesssim \begin{cases} 1, & 0 < d_u < 1/4, \\ \log T, & d_u = 1/4, \\ T^{4d_u-1}, & 1/4 < d_u < 1/2. \end{cases} \quad (\text{D.20})$$

If $d_u = 0$, then (D.18), together with $\psi_{u,a}(0) = 1$, gives

$$|\psi_{u,a}(w)| \lesssim (1+w)^{-1}, \quad a \in [m_u], \quad w \geq 0.$$

Hence, for $\tau \geq 1$, splitting the sum in (D.17) at $w = \tau$ gives

$$\begin{aligned} |\gamma_{k,k'}(\tau)| &\lesssim \sum_{w=0}^{\tau} \frac{1}{(1+w)(1+w+\tau)} + \sum_{w>\tau} \frac{1}{(1+w)(1+w+\tau)} \\ &\lesssim (1+\tau)^{-1} \sum_{w=0}^{\tau} (1+w)^{-1} + \sum_{w>\tau} (1+w)^{-2} \\ &\lesssim (1+\tau)^{-1} \log(2+\tau). \end{aligned}$$

Together with the bound at $\tau = 0$ and the symmetry for negative lags, we have

$$\max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)| \lesssim (1+|\tau|)^{-1} \log(2+|\tau|), \quad \tau \in \mathbb{Z}, \quad (\text{D.21})$$

so that

$$\sum_{\tau \in \mathbb{Z}} \max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)|^2 < \infty. \quad (\text{D.22})$$

If $d_u < 0$, all segment filters are absolutely summable. Extending each filter by zero on negative indices,

we have

$$\sum_{\tau \in \mathbb{Z}} |\gamma_{k,k'}(\tau)| \lesssim \sum_{\tau \in \mathbb{Z}} \sum_{w \in \mathbb{Z}} |\psi_{u,k}(w+\tau)| |\psi_{u,k'}(w)| = \left\{ \sum_{v \in \mathbb{Z}} |\psi_{u,k}(v)| \right\} \left\{ \sum_{w \in \mathbb{Z}} |\psi_{u,k'}(w)| \right\} < \infty.$$

Hence, it follows that

$$\sum_{\tau \in \mathbb{Z}} \max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)| < \infty, \quad \sum_{\tau \in \mathbb{Z}} \max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)|^2 < \infty. \quad (\text{D.23})$$

Combining (D.20), (D.22), and (D.23), we have

$$B_u(T) := \sum_{\tau=-2(T-1)}^{2(T-1)} \max_{k,k' \in [m_u]} |\gamma_{k,k'}(\tau)|^2 \lesssim \begin{cases} 1, & d_u < 1/4, \\ \log T, & d_u = 1/4, \\ T^{4d_u-1}, & d_u > 1/4, \end{cases} \quad (\text{D.24})$$

In particular, $T^{-1}B_u(T) = O(R(T, d_u))$.

We next bound the covariance between the products. Recall that $k = J_u(t)$, $r = J_u(t+h)$, $k' = J_u(s)$, $r' = J_u(s+h)$, and $\tau = t - s$. If $i \neq j$, independence across coordinates and the zero-mean property give

$$\text{Cov}(u_{t,i}u_{t+h,j}, u_{s,i}u_{s+h,j}) = \gamma_{k,k'}(\tau)\gamma_{r,r'}(\tau).$$

If $i = j$, the fourth-moment expansion gives

$$\text{Cov}(u_{t,i}u_{t+h,i}, u_{s,i}u_{s+h,i}) = \gamma_{k,k'}(\tau)\gamma_{r,r'}(\tau) + \gamma_{k,r'}(\tau-h)\gamma_{r,k'}(\tau+h) + \mathcal{K}_{k,r,k',r'}(\tau, h), \quad (\text{D.25})$$

where $\mathcal{K}_{k,r,k',r'}(\tau, h)$ denotes the fourth-cumulant contribution, viz. after extending each filter by zero on negative indices,

$$\mathcal{K}_{k,r,k',r'}(\tau, h) := \kappa_{4,u} \sum_{\ell \in \mathbb{Z}} \psi_{u,k}(\ell) \psi_{u,r}(\ell+h) \psi_{u,k'}(\ell-\tau) \psi_{u,r'}(\ell+h-\tau),$$

where $\kappa_{4,u} := \mathbb{E}(x_{u,0,1}^4) - 3\{\mathbb{E}(x_{u,0,1}^2)\}^2$ is finite by the bounded fourth moment assumption. Define

$$a_q := \psi_{u,k}(q) \psi_{u,r}(q+h), \quad b_q := \psi_{u,k'}(q) \psi_{u,r'}(q+h).$$

Then, $|\mathcal{K}_{k,r,k',r'}(\tau, h)| \lesssim \sum_{q \in \mathbb{Z}} |a_q| |b_{q-\tau}|$. Summing over τ and applying the Cauchy–Schwarz inequality

gives

$$\begin{aligned} \sum_{\tau \in \mathbb{Z}} |\mathcal{K}_{k,r,k',r'}(\tau, h)| &\lesssim \left(\sum_{q \in \mathbb{Z}} |a_q| \right) \left(\sum_{q \in \mathbb{Z}} |b_q| \right) \\ &\lesssim \left\{ \sum_{q \in \mathbb{Z}} |\psi_{u,k}(q)|^2 \right\}^{1/2} \left\{ \sum_{q \in \mathbb{Z}} |\psi_{u,r}(q+h)|^2 \right\}^{1/2} \left\{ \sum_{q \in \mathbb{Z}} |\psi_{u,k'}(q)|^2 \right\}^{1/2} \left\{ \sum_{q \in \mathbb{Z}} |\psi_{u,r'}(q+h)|^2 \right\}^{1/2} \lesssim 1. \end{aligned}$$

Therefore, we have

$$\sum_{\tau \in \mathbb{Z}} \max_{k,r,k',r' \in [m_u]} |\mathcal{K}_{k,r,k',r'}(\tau, h)| \lesssim 1. \quad (\text{D.26})$$

For the two covariance-product terms in (D.25), we have

$$\sum_{\tau=-(T-1)}^{T-1} \max_{k,r,k',r' \in [m_u]} |\gamma_{k,k'}(\tau) \gamma_{r,r'}(\tau)| \leq \sum_{\tau=-(T-1)}^{T-1} \left\{ \max_{a,b \in [m_u]} |\gamma_{a,b}(\tau)| \right\}^2 \leq B_u(T), \quad (\text{D.27})$$

and, by the Cauchy–Schwarz inequality,

$$\begin{aligned} &\sum_{\tau=-(T-1)}^{T-1} \max_{k,r,k',r' \in [m_u]} |\gamma_{k,r'}(\tau-h) \gamma_{r,k'}(\tau+h)| \\ &\leq \left\{ \sum_{\tau=-(T-1)}^{T-1} \max_{k,r' \in [m_u]} |\gamma_{k,r'}(\tau-h)|^2 \right\}^{1/2} \left\{ \sum_{\tau=-(T-1)}^{T-1} \max_{r,k' \in [m_u]} |\gamma_{r,k'}(\tau+h)|^2 \right\}^{1/2} \leq B_u(T), \end{aligned} \quad (\text{D.28})$$

where the last inequality follows because both shifted ranges $\{\tau-h : |\tau| \leq T-1\}$ and $\{\tau+h : |\tau| \leq T-1\}$ are contained in $\{-2(T-1), \dots, 2(T-1)\}$. For fixed i, j and h , define

$$D_{u,ij,h}(\tau) := \max_{\substack{t,s \in [T-h] \\ t-s=\tau}} |\text{Cov}(u_{t,i} u_{t+h,j}, u_{s,i} u_{s+h,j})|,$$

where the maximum over an empty set is taken to be zero by convention. Combining (D.26), (D.27), and (D.28), we obtain

$$\sum_{\tau=-(T-1)}^{T-1} D_{u,ij,h}(\tau) \lesssim B_u(T) + 1. \quad (\text{D.29})$$

For each τ , the number of pairs $(t, s) \in [T-h]^2$ satisfying $t-s=\tau$ is $\max\{T-h-|\tau|, 0\}$. Therefore, by

(D.16),

$$\begin{aligned}
\text{Var}\left(\tilde{\Gamma}_{u,ij}(h)\right) &\leq T^{-2} \sum_{\tau=-(T-1)}^{T-1} \max\{T-h-|\tau|, 0\} D_{u,ij,h}(\tau) \\
&\leq T^{-1} \sum_{\tau=-(T-1)}^{T-1} D_{u,ij,h}(\tau) \\
&\lesssim T^{-1}(B_u(T) + 1).
\end{aligned}$$

Since $T^{-1}B_u(T) = O(R(T, d_u))$ and $T^{-1} = O(R(T, d_u))$, it follows that

$$\text{Var}\left(\tilde{\Gamma}_{u,ij}(h)\right) = O(R(T, d_u)).$$

Since each entry of $\tilde{\Gamma}_u(h)$ is centered and p_u is bounded,

$$\mathbb{E}\|\tilde{\Gamma}_u(h)\|_F^2 = \sum_{i=1}^{p_u} \sum_{j=1}^{p_u} \mathbb{E}\left(\tilde{\Gamma}_{u,ij}(h)^2\right) = \sum_{i=1}^{p_u} \sum_{j=1}^{p_u} \text{Var}\left(\tilde{\Gamma}_{u,ij}(h)\right) = O(R(T, d_u)).$$

Markov's inequality then gives

$$\|\tilde{\Gamma}_u(h)\|_F^2 = O_P(R(T, d_u)),$$

which proves (D.14).

To prove (D.15), define

$$\tilde{\Sigma}_{uv} := T^{-1} \sum_{t=1}^T \mathbf{u}_t \mathbf{v}_t'.$$

Fix an entry $(i, j) \in [p_u] \times [p_v]$. Since the two processes $\{\mathbf{u}_t\}$ and $\{\mathbf{v}_t\}$ are independent and have mean zero, we have $\mathbb{E}(u_{t,i}v_{t,j}) = 0$ for every $t \in [T]$. Hence,

$$\text{Var}\left(\tilde{\Sigma}_{uv,ij}\right) = T^{-2} \sum_{t=1}^T \sum_{s=1}^T \text{Cov}(u_{t,i}v_{t,j}, u_{s,i}v_{s,j}) = T^{-2} \sum_{t=1}^T \sum_{s=1}^T \text{Cov}(u_{t,i}, u_{s,i}) \text{Cov}(v_{t,j}, v_{s,j}), \quad (\text{D.30})$$

where the second equality follows from the independence of $\{\mathbf{u}_t\}$ and $\{\mathbf{v}_t\}$.

Let $J_u(t)$ and $J_v(t)$ denote the segment indices of $\mathbf{u}_t$ and $\mathbf{v}_t$, respectively. For $a, a' \in [m_u]$ and $\tau \geq 0$, define

$$\gamma_{u,a,a'}(\tau) := \sigma_u^2 \sum_{w \geq 0} \psi_{u,a}(w + \tau) \psi_{u,a'}(w),$$

where $\sigma_u^2 := \mathbb{E}(x_{u,0,1}^2) < \infty$, and set $\gamma_{u,a,a'}(\tau) := \gamma_{u,a',a}(-\tau)$ when $\tau < 0$. Define $\gamma_{v,b,b'}(\tau)$ analogously for $b, b' \in [m_v]$. By the temporal independence of the innovations, whenever $J_u(t) = a$, $J_u(s) = a'$, and $t - s = \tau$, we have $\text{Cov}(u_{t,i}, u_{s,i}) = \gamma_{u,a,a'}(\tau)$, and the analogous equality holds for $\{\mathbf{v}_t\}$. In particular,

$$\max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| = \left\{ \max_{a,a' \in [m_u]} |\gamma_{u,a,a'}(\tau)| \right\} \left\{ \max_{b,b' \in [m_v]} |\gamma_{v,b,b'}(\tau)| \right\}.$$

We first consider the case $d_u > 0$ and $d_v > 0$. Taking square roots in (D.19), applied separately to $\{\mathbf{u}_t\}$ and $\{\mathbf{v}_t\}$, gives

$$\max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| \lesssim (1 + |\tau|)^{2d_u-1} (1 + |\tau|)^{2d_v-1} = (1 + |\tau|)^{4d_{uv}-2}, \quad (\text{D.31})$$

where $d_{uv} := (d_u + d_v)/2$.

Next, consider the case where $d_u \leq 0$ or $d_v \leq 0$. We divide it into two disjoint subcases. First, suppose that $\min\{d_u, d_v\} < 0$. Without loss of generality, let $d_u < 0$. The argument in (D.23) gives

$$\sum_{\tau \in \mathbb{Z}} \max_{a,a' \in [m_u]} |\gamma_{u,a,a'}(\tau)| < \infty.$$

Moreover, by the Cauchy–Schwarz inequality and (D.13),

$$\sup_{\tau \in \mathbb{Z}} \max_{b,b' \in [m_v]} |\gamma_{v,b,b'}(\tau)| \leq \sigma_v^2 \max_{b,b' \in [m_v]} \left\{ \sum_{w \geq 0} \psi_{v,b}(w)^2 \right\}^{1/2} \left\{ \sum_{w \geq 0} \psi_{v,b'}(w)^2 \right\}^{1/2} \lesssim 1,$$

where $\sigma_v^2 := \mathbb{E}(x_{v,0,1}^2)$. It follows that

$$\sum_{\tau \in \mathbb{Z}} \max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| < \infty.$$

It remains to consider the case $\min\{d_u, d_v\} = 0$. Without loss of generality, let $d_u = 0$ and $d_v \geq 0$. The bound in (D.21) gives

$$\max_{a,a' \in [m_u]} |\gamma_{u,a,a'}(\tau)| \lesssim (1 + |\tau|)^{-1} \log(2 + |\tau|).$$

If $d_v > 0$, then

$$\max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| \lesssim (1 + |\tau|)^{2d_v-2} \log(2 + |\tau|),$$

which is summable because $2d_v - 2 < -1$. If $d_v = 0$, the same bound in (D.21) gives

$$\max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| \lesssim (1 + |\tau|)^{-2} \log^2(2 + |\tau|),$$

which is also summable. Hence, whenever $d_u \leq 0$ or $d_v \leq 0$, we have

$$\sum_{\tau \in \mathbb{Z}} \max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| < \infty. \quad (\text{D.32})$$

In this case, since $d_u \leq 0$ or $d_v \leq 0$, while both $d_u, d_v < 1/2$, we have $d_{uv} < 1/4$.

Define

$$B_{uv}(T) := \sum_{\tau=-(T-1)}^{T-1} \max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)|.$$

Combining the two cases (D.31) and (D.32), we obtain

$$B_{uv}(T) \lesssim \begin{cases} 1, & d_{uv} < 1/4, \\ \log T, & d_{uv} = 1/4, \\ T^{4d_{uv}-1}, & d_{uv} > 1/4. \end{cases}$$

Consequently, $T^{-1}B_{uv}(T) = O(R(T, d_{uv}))$. For each $\tau \in \{-(T-1), \dots, T-1\}$, the number of pairs $(t, s) \in [T]^2$ satisfying $t - s = \tau$ is $T - |\tau|$. Therefore, grouping the pairs according to their lag gives

$$\begin{aligned} \text{Var}(\tilde{\Sigma}_{uv,ij}) &\leq T^{-2} \sum_{\tau=-(T-1)}^{T-1} (T - |\tau|) \max_{\substack{a,a' \in [m_u] \\ b,b' \in [m_v]}} |\gamma_{u,a,a'}(\tau) \gamma_{v,b,b'}(\tau)| \\ &\leq T^{-1} B_{uv}(T) = O(R(T, d_{uv})). \end{aligned}$$

Since p_u and p_v are bounded and $\mathbb{E}(\tilde{\Sigma}_{uv,ij}) = 0$,

$$\mathbb{E} \|\tilde{\Sigma}_{uv}\|_F^2 = \sum_{i=1}^{p_u} \sum_{j=1}^{p_v} \text{Var}(\tilde{\Sigma}_{uv,ij}) = O(R(T, d_{uv})).$$

Markov's inequality then gives

$$\|\tilde{\Sigma}_{uv}\|_F^2 = O_P(R(T, d_{uv})),$$

which proves (D.15) and completes the proof. $\square$

Lemma 3 (*Auxiliary bounds for error interactions*) Let Assumptions (E1) and (E2) hold. Let $\{\mathbf{u}_t\}_{t \in [T]} \in \mathcal{G}_T(\{d_{u,1}, \dots, d_{u,m_u}\} | \{\mathbf{x}_{u,i}, i \in \mathbb{Z}\})$ be a p_u -dimensional piecewise-stationary time series, where p_u and m_u are bounded, and assume that $\{\mathbf{u}_t\}_{t \in [T]}$ is independent of $\{\mathbf{E}_t^{(m)} : t \in [T], m \in [M]\}$. Let $d_u := \max_{1 \leq j \leq m_u} d_{u,j}$. For each $m, n \in [M]$ with $m \neq n$, the following bounds hold:

$$\left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{u}_t' \right\|_F^2 = O_P(T^{-1} p_m q_m), \quad (\text{D.33})$$

$$\left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \right\|_F^2 = O_P(T^{-1} p_m p_n q_m q_n), \quad (\text{D.34})$$

$$\left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \right\|_F^2 = O_P(p_m q_m^2 + T^{-1} p_m^2 q_m^2). \quad (\text{D.35})$$

Moreover, for any deterministic vector $\mathbf{w} = (w_1, \dots, w_T)'$ and for each $m \in [M]$,

$$\left\| T^{-1} \sum_{t=1}^T w_t \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \right\|^2 = O_P(T^{-2} p_m q_m \|\mathbf{w}\|^2). \quad (\text{D.36})$$

Proof of Lemma 3. We first prove (D.33).

$$\begin{aligned} \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{u}_t' \right\|_F^2 &= T^{-2} \sum_{t=1}^T \sum_{s=1}^T \mathbb{E} \left(\text{tr}(\mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{u}_t' \mathbf{u}_s' \mathbf{1}_{q_m}' \mathbf{E}_s^{(m)'}) \right) \\ &= T^{-2} \sum_{t=1}^T \sum_{s=1}^T \mathbb{E}(\mathbf{1}_{q_m}' \mathbf{E}_s^{(m)'} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}) \mathbb{E}(\mathbf{u}_t' \mathbf{u}_s). \end{aligned}$$

The second equality follows because the whole process $\{\mathbf{u}_t\}_{t \in [T]}$ is independent of $\{\mathbf{E}_t^{(m)}\}_{t \in [T]}$. Define the admissible index set $I_T(h) := \{t \in [T] : t+h \in [T]\}$. Re-indexing by lag $h := s - t$ gives

$$\mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{u}_t' \right\|_F^2 = T^{-2} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \mathbb{E}(\mathbf{1}_{q_m}' \mathbf{E}_{t+h}^{(m)'} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}) \mathbb{E}(\mathbf{u}_t' \mathbf{u}_{t+h}).$$

We bound the two expectations on the r.h.s. separately. First, expand

$$|\mathbb{E}(\mathbf{1}_{q_m}' \mathbf{E}_{t+h}^{(m)'} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m})| = \left| \mathbb{E} \left(\sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} E_{t+h,ij}^{(m)} E_{t,il}^{(m)} \right) \right| \leq \sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} |\mathbb{E}(E_{t+h,ij}^{(m)} E_{t,il}^{(m)})|.$$

By Lemma 1, for each fixed (i, j) ,

$$\sum_{l=1}^{q_m} |\mathbb{E}(E_{t+h,ij}^{(m)} E_{t,il}^{(m)})| \leq \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)})| \leq C \rho(h),$$

uniformly over t, h, m, i, j . Summing over i and j , we obtain

$$|\mathbb{E}(\mathbf{1}'_{q_m} \mathbf{E}_{t+h}^{(m)'} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m})| \leq C p_m q_m \rho(h). \quad (\text{D.37})$$

Moreover, by (D.13) in Lemma 2 and the Cauchy–Schwarz inequality,

$$\sup_{t \in I_T(h)} |\mathbb{E}(\mathbf{u}'_t \mathbf{u}_{t+h})| \leq \sup_{t, s \in [T]} (\mathbb{E} \|\mathbf{u}_t\|^2)^{1/2} (\mathbb{E} \|\mathbf{u}_s\|^2)^{1/2} \leq C. \quad (\text{D.38})$$

Combining (D.37) and (D.38),

$$\begin{aligned} \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{u}'_t \right\|_F^2 &\leq C T^{-2} p_m q_m \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \rho(h) \\ &\leq C T^{-1} p_m q_m \sum_{h \in \mathbb{Z}} \rho(h) \\ &= O(T^{-1} p_m q_m), \end{aligned}$$

and Markov's inequality proves (D.33).

We next prove (D.34). First, for each group $m \in [M]$, define $\zeta_t^{(m)} := q_m^{-1} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$. Then

$$\mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}'_{q_n} \mathbf{E}_t^{(n)'} \right\|_F^2 = q_m^2 q_n^2 \sum_{i=1}^{p_m} \sum_{k=1}^{p_n} \mathbb{E} \left(T^{-1} \sum_{t=1}^T \zeta_{t,i}^{(m)} \zeta_{t,k}^{(n)} \right)^2.$$

We first show that, for $m \neq n$ and $t, s \in [T]$,

$$\mathbb{E} \left(\zeta_{t,i}^{(m)} \zeta_{s,k}^{(n)} \right) = 0. \quad (\text{D.39})$$

Write $\zeta_{t,i}^{(m)} = \zeta_{G,t,i}^{(m)} + \zeta_{F,t,i}^{(m)} + \zeta_{\epsilon,t,i}^{(m)}$, where the three terms are the corresponding row-averaged components induced respectively by $\mathbf{Q}_{e,m} \mathbf{G}_{e,t} \mathbf{J}'_{e,m}$, $\mathbf{R}_{e,m} \mathbf{F}_{e,t}^{(m)} \mathbf{C}'_{e,m}$, and $\mathbf{\Sigma}_{\epsilon,m} \circ \boldsymbol{\epsilon}_t^{(m)}$ in (E1). By the mean-zero assumptions and the mutual independence of the global error, local error, and idiosyncratic error innovation processes in (E1) and (E2), all cross terms vanish, and hence

$$\mathbb{E} \left(\zeta_{t,i}^{(m)} \zeta_{s,k}^{(n)} \right) = \mathbb{E} \left(\zeta_{G,t,i}^{(m)} \zeta_{G,s,k}^{(n)} \right) + \mathbb{E} \left(\zeta_{F,t,i}^{(m)} \zeta_{F,s,k}^{(n)} \right) + \mathbb{E} \left(\zeta_{\epsilon,t,i}^{(m)} \zeta_{\epsilon,s,k}^{(n)} \right).$$

The last two terms are zero by the cross-group independence of the group-specific local error processes and the idiosyncratic error processes in (E2). It remains to consider the global-global term. By (E2),

$$\mathbb{E} ((\mathbf{G}_{e,t})_{ab} (\mathbf{G}_{e,s})_{cd}) = \mathbb{1}\{a = c, b = d\} \sum_{u,v \geq 0} a_{e,u} a_{e,v} \mathbb{1}\{t - u = s - v\}.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\{\zeta_{G,t,i}^{(m)} \zeta_{G,s,k}^{(n)}\} &= (q_m q_n)^{-1} \sum_{j=1}^{q_m} \sum_{l=1}^{q_n} \mathbb{E}((\mathbf{Q}_{e,m} \mathbf{G}_{e,t} \mathbf{J}'_{e,m})_{ij} (\mathbf{Q}_{e,n} \mathbf{G}_{e,s} \mathbf{J}'_{e,n})_{kl}) \\
&= (q_m q_n)^{-1} \left(\sum_{u,v \geq 0} a_{e,u} a_{e,v} \mathbb{1}\{t-u = s-v\} \right) (\mathbf{Q}_{e,m} \mathbf{Q}'_{e,n})_{ik} (\mathbf{1}'_{q_m} \mathbf{J}_{e,m} \mathbf{J}'_{e,n} \mathbf{1}_{q_n}) \\
&= 0,
\end{aligned}$$

where the last equality follows from the cross-group orthogonality $\mathbf{Q}_{e,m} \mathbf{Q}'_{e,n} = \mathbf{0}$ or $\mathbf{J}_{e,m} \mathbf{J}'_{e,n} = \mathbf{0}$ in (E1).

This proves (D.39).

By (D.39), for each fixed $i \in [p_m]$ and $k \in [p_n]$,

$$\begin{aligned}
\mathbb{E} \left(T^{-1} \sum_{t=1}^T \zeta_{t,i}^{(m)} \zeta_{t,k}^{(n)} \right)^2 &= T^{-2} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \text{Cov} \left(\zeta_{t+h,i}^{(m)} \zeta_{t+h,k}^{(n)}, \zeta_{t,i}^{(m)} \zeta_{t,k}^{(n)} \right) \\
&\leq T^{-2} q_m^{-2} q_n^{-2} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \sum_{j=1}^{q_m} \sum_{l=1}^{q_n} \sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(n)}, E_{t,ia}^{(m)} E_{t,kb}^{(n)} \right) \right|.
\end{aligned}$$

By (D.6) in Lemma 1,

$$\sum_{a=1}^{q_m} \sum_{b=1}^{q_n} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(n)}, E_{t,ia}^{(m)} E_{t,kb}^{(n)} \right) \right| \leq C \rho(h),$$

uniformly over all admissible indices. Therefore,

$$\begin{aligned}
\mathbb{E} \left(T^{-1} \sum_{t=1}^T \zeta_{t,i}^{(m)} \zeta_{t,k}^{(n)} \right)^2 &\leq C T^{-2} q_m^{-2} q_n^{-2} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \sum_{j=1}^{q_m} \sum_{l=1}^{q_n} \rho(h) \\
&\leq C T^{-1} q_m^{-1} q_n^{-1} \sum_{h \in \mathbb{Z}} \rho(h) \\
&= O(T^{-1} q_m^{-1} q_n^{-1}).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}'_{q_n} \mathbf{E}_t^{(n)'} \right\|_F^2 &\leq C q_m^2 q_n^2 \sum_{i=1}^{p_m} \sum_{k=1}^{p_n} T^{-1} q_m^{-1} q_n^{-1} \\
&= O(T^{-1} p_m p_n q_m q_n).
\end{aligned}$$

Markov's inequality proves (D.34).

We next prove (D.35). Using the same notation $\zeta_t^{(m)} = q_m^{-1} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$, we have

$$T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} = T^{-1} q_m^2 \sum_{t=1}^T \zeta_t^{(m)} \zeta_t^{(m)'}$$

Hence,

$$\begin{aligned} & \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \right\|_F^2 \\ &= q_m^4 \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \zeta_t^{(m)} \zeta_t^{(m)'} \right\|_F^2 \\ &\leq 2q_m^4 \left\| T^{-1} \sum_{t=1}^T \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right\|_F^2 + 2q_m^4 \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \left[\zeta_t^{(m)} \zeta_t^{(m)'} - \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right] \right\|_F^2. \end{aligned}$$

We first bound the deterministic part. Let $\Sigma_{\zeta,t}^{(m)} := \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right)$. For each fixed $t \in [T]$ and $i \in [p_m]$, Lemma 1 with $h = 0$ gives

$$\begin{aligned} \sum_{k=1}^{p_m} |(\Sigma_{\zeta,t}^{(m)})_{ik}| &= q_m^{-2} \sum_{k=1}^{p_m} \left| \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} \mathbb{E}(E_{t,ij}^{(m)} E_{t,kl}^{(m)}) \right| \\ &\leq q_m^{-2} \sum_{j=1}^{q_m} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} |\mathbb{E}(E_{t,ij}^{(m)} E_{t,kl}^{(m)})| \\ &\leq C q_m^{-1}. \end{aligned}$$

Therefore,

$$\|\Sigma_{\zeta,t}^{(m)}\|_F^2 = \sum_{i=1}^{p_m} \sum_{k=1}^{p_m} |(\Sigma_{\zeta,t}^{(m)})_{ik}|^2 \leq \sum_{i=1}^{p_m} \left\{ \sum_{k=1}^{p_m} |(\Sigma_{\zeta,t}^{(m)})_{ik}| \right\}^2 = O(p_m q_m^{-2}),$$

uniformly over $t \in [T]$. By Jensen's inequality,

$$\left\| T^{-1} \sum_{t=1}^T \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right\|_F^2 \leq T^{-1} \sum_{t=1}^T \|\Sigma_{\zeta,t}^{(m)}\|_F^2 = O(p_m q_m^{-2}).$$

We next bound the centered stochastic part. Write

$$\begin{aligned} & \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \left[\zeta_t^{(m)} \zeta_t^{(m)'} - \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right] \right\|_F^2 \\ &= T^{-2} \sum_{i=1}^{p_m} \sum_{k=1}^{p_m} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \text{Cov} \left(\zeta_{t+h,i}^{(m)} \zeta_{t+h,k}^{(m)}, \zeta_{t,i}^{(m)} \zeta_{t,k}^{(m)} \right). \end{aligned}$$

Taking absolute values and using $\zeta_t^{(m)} = q_m^{-1} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$, we obtain

$$\begin{aligned} & \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \left[\zeta_t^{(m)} \zeta_t^{(m)'} - \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right] \right\|_F^2 \\ & \leq T^{-2} q_m^{-4} \sum_{i=1}^{p_m} \sum_{k=1}^{p_m} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} \sum_{a=1}^{q_m} \sum_{b=1}^{q_m} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(m)}, E_{t,ia}^{(m)} E_{t,kb}^{(m)} \right) \right|. \end{aligned}$$

By (D.6) in Lemma 1, with the two group indices both equal to m ,

$$\sum_{a=1}^{q_m} \sum_{b=1}^{q_m} \left| \text{Cov} \left(E_{t+h,ij}^{(m)} E_{t+h,kl}^{(m)}, E_{t,ia}^{(m)} E_{t,kb}^{(m)} \right) \right| \leq C \rho(h),$$

uniformly over all admissible indices. Hence,

$$\begin{aligned} & \mathbb{E} \left\| T^{-1} \sum_{t=1}^T \left[\zeta_t^{(m)} \zeta_t^{(m)'} - \mathbb{E} \left(\zeta_t^{(m)} \zeta_t^{(m)'} \right) \right] \right\|_F^2 \\ & \leq C T^{-2} q_m^{-4} \sum_{i=1}^{p_m} \sum_{k=1}^{p_m} \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} \rho(h) \\ & \leq C T^{-1} p_m^2 q_m^{-2} \sum_{h \in \mathbb{Z}} \rho(h) \\ & = O(T^{-1} p_m^2 q_m^{-2}). \end{aligned}$$

Combining the deterministic and centered stochastic parts, we obtain

$$\mathbb{E} \left\| T^{-1} \sum_{t=1}^T \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \right\|_F^2 = O(p_m q_m^2 + T^{-1} p_m^2 q_m^2).$$

Markov's inequality proves (D.35).

To prove (D.36), retain the notation $\zeta_t^{(m)} = q_m^{-1} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$. Then,

$$\begin{aligned} \mathbb{E} \left\| T^{-1} \sum_{t=1}^T w_t \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \right\|^2 &= T^{-2} q_m^2 \sum_{t=1}^T \sum_{s=1}^T w_t w_s \mathbb{E} \left(\zeta_t^{(m)'} \mathbf{M}_{p_m} \zeta_s^{(m)} \right) \\ &\leq T^{-2} q_m^2 \sum_{t=1}^T \sum_{s=1}^T |w_t w_s| \left| \mathbb{E} \left(\zeta_t^{(m)'} \mathbf{M}_{p_m} \zeta_s^{(m)} \right) \right|. \end{aligned}$$

For any $h \in \mathbb{Z}$ and $t \in I_T(h)$, write

$$\begin{aligned} \left| \mathbb{E} \left(\zeta_{t+h}^{(m)'} \mathbf{M}_{p_m} \zeta_t^{(m)} \right) \right| &= q_m^{-2} \left| \sum_{i=1}^{p_m} \sum_{k=1}^{p_m} (\mathbf{M}_{p_m})_{ik} \sum_{j=1}^{q_m} \sum_{l=1}^{q_m} \mathbb{E} \left(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)} \right) \right| \\ &\leq q_m^{-2} \sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \sum_{k=1}^{p_m} \sum_{l=1}^{q_m} \left| \mathbb{E} \left(E_{t+h,ij}^{(m)} E_{t,kl}^{(m)} \right) \right| \end{aligned}$$

$$\begin{aligned}
&\leq Cq_m^{-2} \sum_{i=1}^{p_m} \sum_{j=1}^{q_m} \rho(h) \\
&= Cp_m q_m^{-1} \rho(h),
\end{aligned}$$

where the first inequality uses $|(\mathbf{M}_{p_m})_{ik}| \leq 1$, and the last inequality follows from (D.4) in Lemma 1. Therefore,

$$\mathbb{E} \left\| T^{-1} \sum_{t=1}^T w_t \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \right\|^2 \leq CT^{-2} p_m q_m \sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} |w_t w_{t+h}| \rho(h).$$

Recall the $p = 1, q = 2, r = 2$ case of the discrete Young inequality: if $\sum_{h \in \mathbb{Z}} |a_h| < \infty$ and $\sum_{t \in \mathbb{Z}} b_t^2 < \infty$, then

$$\sum_{s \in \mathbb{Z}} \left(\sum_{t \in \mathbb{Z}} |a_{s-t}| |b_t| \right)^2 \leq \left(\sum_{h \in \mathbb{Z}} |a_h| \right)^2 \sum_{t \in \mathbb{Z}} b_t^2.$$

Extending w_t by zero outside $[T]$, and taking $a_h = \rho(h)$ and $b_t = w_t$ in the preceding inequality, we have

$$\begin{aligned}
&\sum_{h=-(T-1)}^{T-1} \sum_{t \in I_T(h)} |w_t w_{t+h}| \rho(h) = \sum_{s \in \mathbb{Z}} |w_s| \sum_{t \in \mathbb{Z}} \rho(s-t) |w_t| \\
&\leq \left(\sum_{s \in \mathbb{Z}} w_s^2 \right)^{1/2} \left\{ \sum_{s \in \mathbb{Z}} \left(\sum_{t \in \mathbb{Z}} \rho(s-t) |w_t| \right)^2 \right\}^{1/2} \leq \left(\sum_{h \in \mathbb{Z}} \rho(h) \right) \sum_{t \in \mathbb{Z}} w_t^2 \leq C \|\mathbf{w}\|^2,
\end{aligned}$$

where the second inequality follows from the discrete Young inequality, and the last inequality follows from $\sum_{h \in \mathbb{Z}} \rho(h) < \infty$ in Lemma 1. Consequently,

$$\mathbb{E} \left\| T^{-1} \sum_{t=1}^T w_t \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \right\|^2 \leq CT^{-2} p_m q_m \|\mathbf{w}\|^2.$$

Markov's inequality proves (D.36). This completes the proof. $\square$

Lemma 4 (*Asymptotic orthogonality of projected rotation*) Let $\mathbf{U}, \hat{\mathbf{U}} \in \mathbb{R}^{d \times r}$ satisfy $\mathbf{U}'\mathbf{U} = \mathbf{I}_r$ and $\hat{\mathbf{U}}'\hat{\mathbf{U}} = \mathbf{I}_r$. Define $\mathbf{P}_U := \mathbf{U}\mathbf{U}'$, $\hat{\mathbf{P}}_U := \hat{\mathbf{U}}\hat{\mathbf{U}}'$, and $\mathbf{H}_U := \mathbf{U}'\hat{\mathbf{U}}$. Then,

$$\hat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U = (\mathbf{I}_d - \mathbf{P}_U)\hat{\mathbf{U}}. \quad (\text{D.40})$$

Moreover, if $\|\hat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U\|_F = o_P(1)$, then

$$\mathbf{H}_U' \mathbf{H}_U = \mathbf{I}_r + o_P(1), \quad \mathbf{H}_U \mathbf{H}_U' = \mathbf{I}_r + o_P(1), \quad (\text{D.41})$$

and $\mathbf{H}_U$ is invertible with probability tending to one. In addition,

$$\|\widehat{\mathbf{P}}_U - \mathbf{P}_U\|_F^2 = 2\|\widehat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U\|_F^2. \quad (\text{D.42})$$

Proof of Lemma 4. By the definition of $\mathbf{H}_U$, $\mathbf{U}\mathbf{H}_U = \mathbf{U}\mathbf{U}'\widehat{\mathbf{U}} = \mathbf{P}_U\widehat{\mathbf{U}}$. Therefore,

$$\widehat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U = \widehat{\mathbf{U}} - \mathbf{P}_U\widehat{\mathbf{U}} = (\mathbf{I}_d - \mathbf{P}_U)\widehat{\mathbf{U}}.$$

Let $\mathbf{E}_U := \widehat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U$. Since $\mathbf{U}'\mathbf{E}_U = 0$, we have

$$\mathbf{I}_r = \widehat{\mathbf{U}}'\widehat{\mathbf{U}} = (\mathbf{U}\mathbf{H}_U + \mathbf{E}_U)'(\mathbf{U}\mathbf{H}_U + \mathbf{E}_U) = \mathbf{H}_U'\mathbf{H}_U + \mathbf{E}_U'\mathbf{E}_U.$$

Hence

$$\|\mathbf{H}_U'\mathbf{H}_U - \mathbf{I}_r\| = \|\mathbf{E}_U\|^2 \leq \|\mathbf{E}_U\|_F^2 = o_P(1).$$

Since $\mathbf{H}_U$ is square, $\mathbf{H}_U'\mathbf{H}_U$ and $\mathbf{H}_U\mathbf{H}_U'$ have the same eigenvalues, given by the squared singular values of $\mathbf{H}_U$. Hence

$$\|\mathbf{H}_U\mathbf{H}_U' - \mathbf{I}_r\| = \max_{1 \leq j \leq r} |\sigma_j^2(\mathbf{H}_U) - 1| = \|\mathbf{H}_U'\mathbf{H}_U - \mathbf{I}_r\| = o_P(1).$$

Also, since $\lambda_{\min}(\mathbf{H}_U'\mathbf{H}_U) = 1 + o_P(1)$, $\mathbf{H}_U$ is invertible with probability tending to one.

Finally, write

$$\begin{aligned} \|\widehat{\mathbf{P}}_U - \mathbf{P}_U\|_F^2 &= \text{tr}(\widehat{\mathbf{P}}_U) + \text{tr}(\mathbf{P}_U) - 2\text{tr}(\widehat{\mathbf{P}}_U\mathbf{P}_U) \\ &= 2r - 2\text{tr}(\widehat{\mathbf{U}}'\mathbf{U}\mathbf{U}'\widehat{\mathbf{U}}) \\ &= 2r - 2\text{tr}(\mathbf{H}_U'\mathbf{H}_U). \end{aligned}$$

On the other hand, from $\mathbf{I}_r = \mathbf{H}_U'\mathbf{H}_U + \mathbf{E}_U'\mathbf{E}_U$, we have

$$\|\mathbf{E}_U\|_F^2 = \text{tr}(\mathbf{E}_U'\mathbf{E}_U) = r - \text{tr}(\mathbf{H}_U'\mathbf{H}_U).$$

Therefore,

$$\|\widehat{\mathbf{P}}_U - \mathbf{P}_U\|_F^2 = 2\|\mathbf{E}_U\|_F^2 = 2\|\widehat{\mathbf{U}} - \mathbf{U}\mathbf{H}_U\|_F^2.$$

This completes the proof.

Lemma 5 (*Bounds for $\mathbf{S}_{A_1,mn}$ and $\Delta_{A_1,mn}$*) For any $m, n \in [M]$ with $m \neq n$, let $\mathbf{S}_{A_1,mn}$ and $\widehat{\mathbf{S}}_{A_1,mn}$ be defined in (C.9) and (C.10), respectively, and define $\Delta_{A_1,mn} := \widehat{\mathbf{S}}_{A_1,mn} - \mathbf{S}_{A_1,mn}$. Under the assumptions of Theorem 3,

$$\|\mathbf{S}_{A_1,mn}\|_F^2 = O(p_m^{\gamma_r} p_n^{\gamma_r}), \quad (\text{D.43})$$

$$\|\Delta_{A_1,mn}\|_F^2 = o_P(p_m^{\gamma_r} p_n^{\gamma_r}). \quad (\text{D.44})$$

Furthermore, defining $\mathbf{P}_{A_1,m} := \check{\mathbf{A}}_{1m} \check{\mathbf{A}}'_{1m}$, we have

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \Delta_{A_1,mn}\|_F^2 = O_P\left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} + T^{-1} p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1}\right). \quad (\text{D.45})$$

Proof of Lemma 5. Recall from (C.9) and (C.10) that

$$\begin{aligned} \mathbf{S}_{A_1,mn} &:= T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \mathbb{E} \left(\mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}'_{q_n} \mathbf{Y}_t^{(n)'} \right) \mathbf{M}_{p_n}, \\ \widehat{\mathbf{S}}_{A_1,mn} &:= T^{-1} q_m^{-1} q_n^{-1} \mathbf{M}_{p_m} \sum_{t=1}^T \left(\mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}'_{q_n} \mathbf{Y}_t^{(n)'} \right) \mathbf{M}_{p_n}. \end{aligned}$$

We first prove (D.43). By (D.13) in Lemma 2,

$$\left\| T^{-1} \sum_{t=1}^T \mathbb{E}(\alpha_t \alpha_t') \right\|_F \leq T^{-1} \sum_{t=1}^T \mathbb{E} \|\alpha_t \alpha_t'\|_F = T^{-1} \sum_{t=1}^T \mathbb{E} \|\alpha_t\|^2 = O(1).$$

By (C.2), we have $\mathbf{S}_{A_1,mn} = \mathbf{A}_{1m} \left(T^{-1} \sum_{t=1}^T \mathbb{E}(\alpha_t \alpha_t') \right) \mathbf{A}'_{1n}$. Hence, by (ME1),

$$\|\mathbf{S}_{A_1,mn}\|_F^2 \leq \|\mathbf{A}_{1m}\|_F^2 \left\| T^{-1} \sum_{t=1}^T \mathbb{E}(\alpha_t \alpha_t') \right\|_F^2 \|\mathbf{A}_{1n}\|_F^2 = O(p_m^{\gamma_r} p_n^{\gamma_r}),$$

which proves (D.43).

We next prove (D.44) and (D.45). By (IC1), for each $m \in [M]$,

$$q_m^{-1} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} = \mathbf{a}_t^{(m)} + q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}.$$

Substituting these identities into the definitions of $\mathbf{S}_{A_1,mn}$ and $\widehat{\mathbf{S}}_{A_1,mn}$ gives

$$\Delta_{A_1,mn} = \widehat{\mathbf{S}}_{A_1,mn} - \mathbf{S}_{A_1,mn} = \mathcal{I}_{A_1,1} + \mathcal{I}_{A_1,2} + \mathcal{I}_{A_1,3} + \mathcal{I}_{A_1,4},$$

where

$$\begin{aligned}\mathcal{I}_{A_1,1} &:= T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(n)'} - \mathbf{S}_{A_1,mn}, & \mathcal{I}_{A_1,2} &:= (Tq_n)^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n}, \\ \mathcal{I}_{A_1,3} &:= (Tq_m)^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{a}_t^{(n)'}, & \mathcal{I}_{A_1,4} &:= (Tq_m q_n)^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n}.\end{aligned}$$

We bound the four components separately. Throughout the remainder of the proof, we repeatedly use the contraction properties (D.1) in Proposition 1 to remove orthogonal projectors from upper bounds without further mention. By (C.2), since $\mathbf{S}_{A_1,mn} = T^{-1} \sum_{t=1}^T \mathbb{E}(\mathbf{a}_t^{(m)} \mathbf{a}_t^{(n)'})$, we have

$$\mathcal{I}_{A_1,1} = T^{-1} \sum_{t=1}^T \left\{ \mathbf{a}_t^{(m)} \mathbf{a}_t^{(n)'} - \mathbb{E}(\mathbf{a}_t^{(m)} \mathbf{a}_t^{(n)'}) \right\}.$$

Using $\mathbf{a}_t^{(m)} = \mathbf{A}_{1m} \boldsymbol{\alpha}_t + \mathbf{A}_{2m} \boldsymbol{\alpha}_t^{(m)}$ and $\mathbf{a}_t^{(n)} = \mathbf{A}_{1n} \boldsymbol{\alpha}_t + \mathbf{A}_{2n} \boldsymbol{\alpha}_t^{(n)}$, together with the sample orthogonality and cross-group uncorrelatedness conditions in (IC2), we obtain

$$\mathcal{I}_{A_1,1} = \mathbf{A}_{1m} \left\{ T^{-1} \sum_{t=1}^T (\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t' - \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t')) \right\} \mathbf{A}_{1n}' + \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(n)'} \right) \mathbf{A}_{2n}'.$$

It follows from (D.14) and (D.15) in Lemma 2, together with (ME1) and (ME2), that

$$\|\mathcal{I}_{A_1,1}\|_F^2 = O_P \left(R(T, d_{r,\star}) \left(p_m^{\gamma_r} p_n^{\gamma_r} + p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} \right) \right). \quad (\text{D.46})$$

Moreover, since $(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{A}_{1m} = \mathbf{0}$,

$$(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathcal{I}_{A_1,1} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(n)'} \right) \mathbf{A}_{2n}'.$$

Thus,

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}) \mathcal{I}_{A_1,1}\|_F^2 = O_P \left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} \right). \quad (\text{D.47})$$

For $\mathcal{I}_{A_1,2}$, write

$$\mathcal{I}_{A_1,2} = \mathbf{A}_{1m} \left\{ (Tq_n)^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right\} + \mathbf{A}_{2m} \left\{ (Tq_n)^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right\}.$$

Applying (D.33) in Lemma 3 to $\{\alpha_t\}$ and $\{\alpha_t^{(m)}\}$, together with (ME1),

$$\|\mathcal{I}_{A_1,2}\|_F^2 = O_P\left(T^{-1}\left(p_m^{\gamma_r} + p_m^{\gamma_r^{(m)}}\right)p_n q_n^{-1}\right). \quad (\text{D.48})$$

Since left-multiplication by $(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})$ eliminates the term involving $\mathbf{A}_{1m}$,

$$(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})\mathcal{I}_{A_1,2} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})\mathbf{A}_{2m} \left\{ (Tq_n)^{-1} \sum_{t=1}^T \alpha_t^{(m)} \mathbf{1}_{q_n}' \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right\}.$$

Consequently,

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})\mathcal{I}_{A_1,2}\|_F^2 = O_P\left(T^{-1}p_m^{\gamma_r^{(m)}}p_n q_n^{-1}\right). \quad (\text{D.49})$$

For $\mathcal{I}_{A_1,3}$, we have

$$\mathcal{I}_{A_1,3} = q_m^{-1} \left(T^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \alpha_t' \right) \mathbf{A}_{1n}' + q_m^{-1} \left(T^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \alpha_t^{(n)'} \right) \mathbf{A}_{2n}'.$$

Applying (D.33) in Lemma 3 and (ME1) gives

$$\|\mathcal{I}_{A_1,3}\|_F^2 = O_P\left(T^{-1}p_m q_m^{-1} \left(p_n^{\gamma_r} + p_n^{\gamma_r^{(n)}}\right)\right) = O_P\left(T^{-1}p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1}\right). \quad (\text{D.50})$$

Consequently,

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})\mathcal{I}_{A_1,3}\|_F^2 \leq \|\mathcal{I}_{A_1,3}\|_F^2 = O_P\left(T^{-1}p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1}\right). \quad (\text{D.51})$$

For $\mathcal{I}_{A_1,4}$, by (D.34) in Lemma 3,

$$\|\mathcal{I}_{A_1,4}\|_F^2 = O_P\left(T^{-1}p_m p_n q_m^{-1} q_n^{-1}\right), \quad (\text{D.52})$$

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1,m})\mathcal{I}_{A_1,4}\|_F^2 = O_P\left(T^{-1}p_m p_n q_m^{-1} q_n^{-1}\right). \quad (\text{D.53})$$

Using $\|\sum_{j=1}^4 \mathcal{I}_{A_1,j}\|_F^2 \lesssim \sum_{j=1}^4 \|\mathcal{I}_{A_1,j}\|_F^2$, together with (D.46), (D.48), (D.50), and (D.52), we obtain

$$\begin{aligned} \|\Delta_{A_1,mn}\|_F^2 &= O_P\left(R(T, d_{r,\star}) \left(p_m^{\gamma_r} p_n^{\gamma_r} + p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}}\right) + T^{-1} p_m^{\gamma_{r,\max}^{(m)}} p_n q_n^{-1} \right. \\ &\quad \left. + T^{-1} p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1} + T^{-1} p_m p_n q_m^{-1} q_n^{-1}\right). \end{aligned} \quad (\text{D.54})$$

To deduce (D.44), first note that $T^{-1} = O(R(T, d_{r,\star}))$. Moreover, by (R1), $p_s^{1-\gamma_r} q_s^{-1}$ and $p_s^{1-\gamma_r^{(s)}} q_s^{-1}$ are both $o(1)$ uniformly in $s \in [M]$. Since $p_m^{\gamma_{r,\max}^{(m)}} \leq p_m^{\gamma_r} + p_m^{\gamma_r^{(m)}}$, we have

$$\begin{aligned} T^{-1} p_m^{\gamma_{r,\max}^{(m)}} p_n q_n^{-1} &\lesssim R(T, d_{r,\star}) \left\{ p_m^{\gamma_r} p_n^{\gamma_r} p_n^{1-\gamma_r} q_n^{-1} + p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} p_n^{1-\gamma_r^{(n)}} q_n^{-1} \right\} \\ &= o \left(R(T, d_{r,\star}) \left\{ p_m^{\gamma_r} p_n^{\gamma_r} + p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} \right\} \right). \end{aligned}$$

The term $T^{-1} p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1}$ follows symmetrically, while the term $T^{-1} p_m p_n q_m^{-1} q_n^{-1}$ is of still smaller order since $p_m^{1-\gamma_{r,\max}^{(m)}} q_m^{-1} = o(1)$. Hence, (D.54) reduces to

$$\|\Delta_{A_1, mn}\|_F^2 = O_P \left(R(T, d_{r,\star}) \left\{ p_m^{\gamma_r} p_n^{\gamma_r} + p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} \right\} \right).$$

After division by $p_m^{\gamma_r} p_n^{\gamma_r}$, the r.h.s. is $O_P(R(T, d_{r,\star})\{1 + \nu_{r,m} \nu_{r,n}\})$. For $m \neq n$,

$$1 + \nu_{r,m} \nu_{r,n} \leq 2(\nu_{r,m} \vee \nu_{r,m}^{-1})(1 \vee \nu_{r,n}),$$

so (R1) gives $R(T, d_{r,\star})\{1 + \nu_{r,m} \nu_{r,n}\} = o(1)$ uniformly over $m \neq n$. This proves (D.44).

Finally, by $T^{-1} = O(R(T, d_{r,\star}))$ and (R1), we have $T^{-1} p_m^{\gamma_r^{(m)}} p_n q_n^{-1} = o \left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} \right)$, while $T^{-1} p_m p_n q_m^{-1} q_n^{-1} = o \left(T^{-1} p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1} \right)$. Combining (D.47), (D.49), (D.51), and (D.53), we obtain

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_1, m}) \Delta_{A_1, mn}\|_F^2 = O_P \left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)}} p_n^{\gamma_r^{(n)}} + T^{-1} p_m p_n^{\gamma_{r,\max}^{(n)}} q_m^{-1} \right),$$

which proves (D.45). This completes the proof. $\square$

Lemma 6 (Bounding $\widehat{\mathbf{D}}_{A_1, m}^{-1}$) Under the assumptions of Theorem 3, define $\omega_{A_1, m} := p_m^{\gamma_r} \sum_{n \neq m} p_n^{\gamma_r}$, and let $\widehat{\mathbf{D}}_{A_1, m}$ be defined in (C.13). Then, for each $m \in [M]$,

$$\lambda_{r_1}(\mathbf{\Omega}_{A_1, m}) \asymp \omega_{A_1, m}, \quad \lambda_{r_1}(\widehat{\mathbf{\Omega}}_{A_1, m}) \asymp_P \omega_{A_1, m}, \quad \|\widehat{\mathbf{D}}_{A_1, m}^{-1}\|_F^2 = O_P(\omega_{A_1, m}^{-2}). \quad (\text{D.55})$$

Proof of Lemma 6. By (ME1), we have $\lambda_{r_1}(\mathbf{A}'_{1m} \mathbf{A}_{1m}) \asymp \lambda_1(\mathbf{A}'_{1m} \mathbf{A}_{1m}) \asymp p_m^{\gamma_r}$. Moreover, since $\mathbf{A}'_{1n} \mathbf{A}_{1n} \succeq 0$ for every $n \neq m$,

$$\begin{aligned} \lambda_{r_1} \left(\sum_{n \neq m} \mathbf{A}'_{1n} \mathbf{A}_{1n} \right) &\geq \sum_{n \neq m} \lambda_{r_1}(\mathbf{A}'_{1n} \mathbf{A}_{1n}) \asymp \sum_{n \neq m} p_n^{\gamma_r}, \\ \lambda_1 \left(\sum_{n \neq m} \mathbf{A}'_{1n} \mathbf{A}_{1n} \right) &\leq \sum_{n \neq m} \lambda_1(\mathbf{A}'_{1n} \mathbf{A}_{1n}) \asymp \sum_{n \neq m} p_n^{\gamma_r}. \end{aligned}$$

By (S1), $\lambda_{r_1} \left(T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t') \right) \asymp \lambda_1 \left(T^{-1} \sum_{t=1}^T \mathbb{E}(\boldsymbol{\alpha}_t \boldsymbol{\alpha}_t') \right) \asymp 1$. Together with the preceding bounds and (ME1), we have

$$\lambda_{r_1}(\boldsymbol{\Omega}_{A_1, m}) \asymp \lambda_1(\boldsymbol{\Omega}_{A_1, m}) \asymp p_m^{\gamma_r} \sum_{n \neq m} p_n^{\gamma_r} = \omega_{A_1, m}.$$

By (C.14), Lemma 5,

$$\begin{aligned} \|\widehat{\boldsymbol{\Omega}}_{A_1, m} - \boldsymbol{\Omega}_{A_1, m}\|_F &\lesssim \sum_{n \neq m} \left\{ \|\boldsymbol{\Delta}_{A_1, mn}\|_F \|\mathbf{S}_{A_1, mn}\|_F + \|\boldsymbol{\Delta}_{A_1, mn}\|_F^2 \right\} \\ &= o_P \left(p_m^{\gamma_r} \sum_{n \neq m} p_n^{\gamma_r} \right) = o_P(\omega_{A_1, m}). \end{aligned}$$

Therefore, Weyl's inequality gives

$$|\lambda_{r_1}(\widehat{\boldsymbol{\Omega}}_{A_1, m}) - \lambda_{r_1}(\boldsymbol{\Omega}_{A_1, m})| \leq \|\widehat{\boldsymbol{\Omega}}_{A_1, m} - \boldsymbol{\Omega}_{A_1, m}\| \leq \|\widehat{\boldsymbol{\Omega}}_{A_1, m} - \boldsymbol{\Omega}_{A_1, m}\|_F = o_P(\omega_{A_1, m}),$$

and hence $\lambda_{r_1}(\widehat{\boldsymbol{\Omega}}_{A_1, m}) \asymp_P \omega_{A_1, m}$. Thus, $\widehat{\mathbf{D}}_{A_1, m}$ is invertible with probability tending to one and, by (C.13),

$$\|\widehat{\mathbf{D}}_{A_1, m}^{-1}\|_F^2 = \sum_{i=1}^{r_1} \lambda_i^{-2}(\widehat{\boldsymbol{\Omega}}_{A_1, m}) \lesssim \lambda_{r_1}^{-2}(\widehat{\boldsymbol{\Omega}}_{A_1, m}) = O_P(\omega_{A_1, m}^{-2}),$$

which completes the proof. $\square$

Lemma 7 (Bounds for $\mathbf{S}_{A_2, m}$ and $\boldsymbol{\Delta}_{A_2, m}$) For any $m \in [M]$, let $\mathbf{S}_{A_2, m}$ and $\widehat{\mathbf{S}}_{A_2, m}$ be defined in (C.19), and define $\boldsymbol{\Delta}_{A_2, m} := \widehat{\mathbf{S}}_{A_2, m} - \mathbf{S}_{A_2, m}$. Under the assumptions of Theorem 4,

$$\|\mathbf{S}_{A_2, m}\|_F^2 = O_P(p_m^{2\gamma_r^{(m)}}), \quad (\text{D.56})$$

$$\|\boldsymbol{\Delta}_{A_2, m}\|_F^2 = o_P(p_m^{2\gamma_r^{(m)}}). \quad (\text{D.57})$$

Furthermore, defining $\mathbf{P}_{A_2, m} := \check{\mathbf{A}}_{2m} \check{\mathbf{A}}_{2m}'$, we have

$$\begin{aligned} \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2, m})\boldsymbol{\Delta}_{A_2, m}\|_F^2 &= O_P \left(R(T, d_r, \star) p_m^{\gamma_r + \gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1} p_m^{\gamma_r + 1} q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r, \max}^{(n)} - \gamma_r} + p_m q_m^{-2} \right). \end{aligned} \quad (\text{D.58})$$

Proof of Lemma 7. Recall from (C.19) that

$$\mathbf{S}_{A_2, m} := T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} \mathbf{P}_{A_1, m}^\perp, \quad \widehat{\mathbf{S}}_{A_2, m} := T^{-1} \sum_{t=1}^T \widehat{\mathbf{a}}_t^{(m)} \widehat{\mathbf{a}}_t^{(m)'} \widehat{\mathbf{P}}_{A_1, m}^\perp,$$

where $\mathbf{P}_{A_1,m}^\perp = \mathbf{I}_{p_m} - \mathbf{P}_{A_1,m}$ and $\widehat{\mathbf{P}}_{A_1,m}^\perp = \mathbf{I}_{p_m} - \widehat{\mathbf{P}}_{A_1,m}$. Consequently,

$$\Delta_{A_2,m} = T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} (\widehat{\mathbf{P}}_{A_1,m}^\perp - \mathbf{P}_{A_1,m}^\perp) + T^{-1} \sum_{t=1}^T (\widehat{\mathbf{a}}_t^{(m)} \widehat{\mathbf{a}}_t^{(m)'} - \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'}) \widehat{\mathbf{P}}_{A_1,m}^\perp. \quad (\text{D.59})$$

We first prove (D.56). By (D.13) and (D.14) in Lemma 2, and (IC2), it follows that

$$T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} = \mathbf{A}_{1m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t' \right) \mathbf{A}_{1m}' + \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \mathbf{A}_{2m}'. \quad (\text{D.60})$$

Since $\mathbf{A}_{1m}' \mathbf{P}_{A_1,m}^\perp = \mathbf{0}$,

$$\mathbf{S}_{A_2,m} = \mathbf{A}_{2m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \mathbf{A}_{2m}' \mathbf{P}_{A_1,m}^\perp. \quad (\text{D.61})$$

Therefore, by (ME1), Proposition 1, and Lemma 2,

$$\|\mathbf{S}_{A_2,m}\|_F^2 \leq \|\mathbf{A}_{2m}\|_F^4 \left\| T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right\|_F^2 = O_P(p_m^{2\gamma_r^{(m)}}),$$

which proves (D.56).

We next prove (D.57) and (D.58). By (IC1), $\widehat{\mathbf{a}}_t^{(m)} - \mathbf{a}_t^{(m)} = q_m^{-1} \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m}$. Using (D.59), write

$$\Delta_{A_2,m} = \mathcal{I}_{A_2,1} + \mathcal{I}_{A_2,2} + \mathcal{I}_{A_2,3} + \mathcal{I}_{A_2,4},$$

where

$$\begin{aligned} \mathcal{I}_{A_2,1} &:= T^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{a}_t^{(m)'} (\widehat{\mathbf{P}}_{A_1,m}^\perp - \mathbf{P}_{A_1,m}^\perp), \quad \mathcal{I}_{A_2,2} := (Tq_m)^{-1} \sum_{t=1}^T \mathbf{a}_t^{(m)} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} \widehat{\mathbf{P}}_{A_1,m}^\perp, \\ \mathcal{I}_{A_2,3} &:= (Tq_m)^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{a}_t^{(m)'} \widehat{\mathbf{P}}_{A_1,m}^\perp, \quad \mathcal{I}_{A_2,4} := (Tq_m^2)^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} \widehat{\mathbf{P}}_{A_1,m}^\perp. \end{aligned}$$

We bound these four terms separately.

By Theorem 3 and (C.18),

$$\begin{aligned} \|\widehat{\mathbf{P}}_{A_1,m} - \mathbf{P}_{A_1,m}\|_F^2 &= O_P \left(R(T, d_{r,\star}) p_m^{\gamma_r^{(m)} - \gamma_r} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1} p_m^{1 - \gamma_r} q_m^{-1} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_{r,\max}^{(n)} - \gamma_r} \right). \end{aligned} \quad (\text{D.62})$$

For $\mathcal{I}_{A_2,1}$, since $\widehat{\mathbf{P}}_{A_1,m}^\perp - \mathbf{P}_{A_1,m}^\perp = \mathbf{P}_{A_1,m} - \widehat{\mathbf{P}}_{A_1,m}$, together with (D.60) and (ME1), we have

$$\|\mathcal{I}_{A_2,1}\|_F^2 = O_P \left((p_m^{2\gamma_r} + p_m^{2\gamma_r^{(m)}}) \|\widehat{\mathbf{P}}_{A_1,m} - \mathbf{P}_{A_1,m}\|_F^2 \right). \quad (\text{D.63})$$

Furthermore, since $(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})\mathbf{A}_{2m} = \mathbf{0}$,

$$(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})\mathcal{I}_{A_2,1} = (\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})\mathbf{A}_{1m} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t' \right) \mathbf{A}_{1m}' (\mathbf{P}_{A_1,m} - \widehat{\mathbf{P}}_{A_1,m}).$$

Consequently,

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})\mathcal{I}_{A_2,1}\|_F^2 = O_P \left(p_m^{2\gamma_r} \|\widehat{\mathbf{P}}_{A_1,m} - \mathbf{P}_{A_1,m}\|_F^2 \right). \quad (\text{D.64})$$

For $\mathcal{I}_{A_2,2}$, write

$$\begin{aligned} \mathcal{I}_{A_2,2} &= \mathbf{A}_{1m} \left\{ (Tq_m)^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} \right\} \widehat{\mathbf{P}}_{A_1,m}^\perp \\ &\quad + \mathbf{A}_{2m} \left\{ (Tq_m)^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \mathbf{1}_{q_m}' \mathbf{E}_t^{(m)'} \mathbf{M}_{p_m} \right\} \widehat{\mathbf{P}}_{A_1,m}^\perp. \end{aligned}$$

Applying (D.33) in Lemma 3, together with (ME1) and Proposition 1, gives

$$\|\mathcal{I}_{A_2,2}\|_F^2 = O_P \left(T^{-1} p_m q_m^{-1} (p_m^{\gamma_r} + p_m^{\gamma_r^{(m)}}) \right). \quad (\text{D.65})$$

Since left-multiplication by $\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m}$ eliminates the term involving $\mathbf{A}_{2m}$,

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2,m})\mathcal{I}_{A_2,2}\|_F^2 = O_P \left(T^{-1} p_m^{\gamma_r+1} q_m^{-1} \right). \quad (\text{D.66})$$

For $\mathcal{I}_{A_2,3}$, we have

$$\begin{aligned} \mathcal{I}_{A_2,3} &= q_m^{-1} \left\{ T^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \boldsymbol{\alpha}_t' \right\} \mathbf{A}_{1m}' \widehat{\mathbf{P}}_{A_1,m}^\perp \\ &\quad + q_m^{-1} \left\{ T^{-1} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{1}_{q_m} \boldsymbol{\alpha}_t^{(m)'} \right\} \mathbf{A}_{2m}' \widehat{\mathbf{P}}_{A_1,m}^\perp. \end{aligned}$$

By (D.33) in Lemma 3, (ME1), and Proposition 1,

$$\|\mathcal{I}_{A_2,3}\|_F^2 = O_P \left(T^{-1} p_m q_m^{-1} (p_m^{\gamma_r} + p_m^{\gamma_r^{(m)}}) \right) = O_P \left(T^{-1} p_m^{\gamma_{r,\max}^{(m)}+1} q_m^{-1} \right). \quad (\text{D.67})$$

The same bound remains valid after left-multiplication by $\mathbf{I}_{p_m} - \mathbf{P}_{A_2, m}$, and hence

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2, m})\mathcal{I}_{A_2, 3}\|_F^2 = O_P\left(T^{-1}p_m^{\gamma_{r, \max}^{(m)}+1}q_m^{-1}\right). \quad (\text{D.68})$$

For $\mathcal{I}_{A_2, 4}$, by (D.35) in Lemma 3 and Proposition 1,

$$\|\mathcal{I}_{A_2, 4}\|_F^2 = O_P(p_m q_m^{-2} + T^{-1}p_m^2 q_m^{-2}), \quad (\text{D.69})$$

$$\|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2, m})\mathcal{I}_{A_2, 4}\|_F^2 = O_P(p_m q_m^{-2} + T^{-1}p_m^2 q_m^{-2}). \quad (\text{D.70})$$

Using $\|\sum_{j=1}^4 \mathcal{I}_{A_2, j}\|_F^2 \lesssim \sum_{j=1}^4 \|\mathcal{I}_{A_2, j}\|_F^2$, together with (D.63), (D.65), (D.67), and (D.69), we obtain

$$\begin{aligned} \|\boldsymbol{\Delta}_{A_2, m}\|_F^2 &= O_P\left((p_m^{2\gamma_r} + p_m^{2\gamma_r^{(m)}})\|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|_F^2 \right. \\ &\quad \left. + T^{-1}p_m^{\gamma_{r, \max}^{(m)}+1}q_m^{-1} + p_m q_m^{-2} + T^{-1}p_m^2 q_m^{-2}\right). \end{aligned} \quad (\text{D.71})$$

To deduce (D.57), Theorem 3 and (C.18) give

$$\begin{aligned} (1 + \nu_{r, m}^{-2})\|\widehat{\mathbf{P}}_{A_1, m} - \mathbf{P}_{A_1, m}\|_F^2 &= O_P\left(R(T, d_{r, \star})(\nu_{r, m} + \nu_{r, m}^{-1}) \max_{n \in [M] \setminus \{m\}} \nu_{r, n} \right. \\ &\quad \left. + \frac{1}{T} \left\{ \frac{p_m^{1-\gamma_r}}{q_m} + \frac{p_m^{1-\gamma_r^{(m)}}}{q_m} \nu_{r, m}^{-1} \right\} \left(1 \vee \max_{n \in [M] \setminus \{m\}} \nu_{r, n}\right) \right) = o_P(1), \end{aligned}$$

where the last equality follows from (R1) and $T^{-1} = O(R(T, d_{r, \star}))$. Moreover,

$$T^{-1}p_m^{\gamma_{r, \max}^{(m)}+1-2\gamma_r^{(m)}}q_m^{-1} = T^{-1}p_m^{1-\gamma_r^{(m)}}q_m^{-1}(1 \vee \nu_{r, m}^{-1}) = o(1),$$

while the remaining two terms in (D.71), after division by $p_m^{2\gamma_r^{(m)}}$, are also $o(1)$ by (R1). Hence (D.57) follows.

Finally, combining (D.64), (D.66), (D.68), and (D.70) with (D.62), we obtain

$$\begin{aligned} \|(\mathbf{I}_{p_m} - \mathbf{P}_{A_2, m})\boldsymbol{\Delta}_{A_2, m}\|_F^2 &= O_P\left(R(T, d_{r, \star})p_m^{\gamma_r + \gamma_r^{(m)}} \max_{n \in [M] \setminus \{m\}} p_n^{\gamma_r^{(n)} - \gamma_r} \right. \\ &\quad \left. + T^{-1}p_m^{\gamma_r + 1}q_m^{-1} \max_{n \in [M]} p_n^{\gamma_{r, \max}^{(n)} - \gamma_r} + p_m q_m^{-2}\right), \end{aligned}$$

where $T^{-1}p_m^2 q_m^{-2}$ is absorbed into the second term under (R1). This proves (D.58), and completes the proof. $\square$

Lemma 8 (Bounding $\widehat{\mathbf{D}}_{A_2, m}^{-1}$) Under the assumptions of Theorem 4, define $\omega_{A_2, m} := p_m^{2\gamma_r^{(m)}}$, and let $\widehat{\mathbf{D}}_{A_2, m}$

be defined in (C.23). Then, for each $m \in [M]$,

$$\lambda_{r_2^{(m)}}(\boldsymbol{\Omega}_{A_2,m}) \asymp_P \omega_{A_2,m}, \quad \lambda_{r_2^{(m)}}(\widehat{\boldsymbol{\Omega}}_{A_2,m}) \asymp_P \omega_{A_2,m}, \quad \|\widehat{\mathbf{D}}_{A_2,m}^{-1}\|_F^2 = O_P(\omega_{A_2,m}^{-2}). \quad (\text{D.72})$$

Proof of Lemma 8. By Lemma 2, (S1), and Weyl's inequality,

$$\lambda_{r_2^{(m)}} \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \asymp_P \lambda_1 \left(T^{-1} \sum_{t=1}^T \boldsymbol{\alpha}_t^{(m)} \boldsymbol{\alpha}_t^{(m)'} \right) \asymp_P 1.$$

Since $\mathbf{P}_{A_1,m}^\perp = \mathbf{A}_{1m}^\perp \mathbf{A}_{1m}^{\perp'}$,

$$\begin{aligned} & (\mathbf{A}_{2m}' \mathbf{A}_{2m})^{-1/2} \mathbf{A}_{2m}' \mathbf{P}_{A_1,m}^\perp \mathbf{A}_{2m} (\mathbf{A}_{2m}' \mathbf{A}_{2m})^{-1/2} \\ &= \left\{ \mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m} (\mathbf{A}_{2m}' \mathbf{A}_{2m})^{-1/2} \right\}' \left\{ \mathbf{A}_{1m}^{\perp'} \mathbf{A}_{2m} (\mathbf{A}_{2m}' \mathbf{A}_{2m})^{-1/2} \right\}. \end{aligned}$$

By (S1), (ME1), and Proposition 1, this gives

$$\lambda_{r_2^{(m)}}(\mathbf{A}_{2m}' \mathbf{P}_{A_1,m}^\perp \mathbf{A}_{2m}) \asymp \lambda_1(\mathbf{A}_{2m}' \mathbf{P}_{A_1,m}^\perp \mathbf{A}_{2m}) \asymp p_m^{\gamma_r^{(m)}}.$$

Therefore, (ME1) gives

$$\lambda_{r_2^{(m)}}(\boldsymbol{\Omega}_{A_2,m}) \asymp_P \lambda_1(\boldsymbol{\Omega}_{A_2,m}) \asymp_P p_m^{2\gamma_r^{(m)}} = \omega_{A_2,m}.$$

Let $\boldsymbol{\Delta}_{A_2,m} := \widehat{\mathbf{S}}_{A_2,m} - \mathbf{S}_{A_2,m}$. Since $\widehat{\boldsymbol{\Omega}}_{A_2,m} = \widehat{\mathbf{S}}_{A_2,m} \widehat{\mathbf{S}}_{A_2,m}'$ and $\boldsymbol{\Omega}_{A_2,m} = \mathbf{S}_{A_2,m} \mathbf{S}_{A_2,m}'$,

$$\widehat{\boldsymbol{\Omega}}_{A_2,m} - \boldsymbol{\Omega}_{A_2,m} = \boldsymbol{\Delta}_{A_2,m} \mathbf{S}_{A_2,m}' + \mathbf{S}_{A_2,m} \boldsymbol{\Delta}_{A_2,m}' + \boldsymbol{\Delta}_{A_2,m} \boldsymbol{\Delta}_{A_2,m}'.$$

Lemma 7 gives $\|\mathbf{S}_{A_2,m}\|_F = O_P(p_m^{\gamma_r^{(m)}})$ and $\|\boldsymbol{\Delta}_{A_2,m}\|_F = o_P(p_m^{\gamma_r^{(m)}})$. Therefore,

$$\|\widehat{\boldsymbol{\Omega}}_{A_2,m} - \boldsymbol{\Omega}_{A_2,m}\|_F \lesssim \|\boldsymbol{\Delta}_{A_2,m}\|_F \|\mathbf{S}_{A_2,m}\|_F + \|\boldsymbol{\Delta}_{A_2,m}\|_F^2 = o_P(p_m^{2\gamma_r^{(m)}}) = o_P(\omega_{A_2,m}).$$

By Weyl's inequality,

$$|\lambda_{r_2^{(m)}}(\widehat{\boldsymbol{\Omega}}_{A_2,m}) - \lambda_{r_2^{(m)}}(\boldsymbol{\Omega}_{A_2,m})| \leq \|\widehat{\boldsymbol{\Omega}}_{A_2,m} - \boldsymbol{\Omega}_{A_2,m}\|_F = o_P(\omega_{A_2,m}),$$

so that $\lambda_{r_2^{(m)}}(\hat{\boldsymbol{\Omega}}_{A_2,m}) \asymp_P \omega_{A_2,m}$. Thus, $\hat{\mathbf{D}}_{A_2,m}$ is invertible with probability tending to one and, by (C.23),

$$\|\hat{\mathbf{D}}_{A_2,m}^{-1}\|_F^2 = \sum_{i=1}^{r_2^{(m)}} \lambda_i^{-2}(\hat{\boldsymbol{\Omega}}_{A_2,m}) \leq r_2^{(m)} \lambda_{r_2^{(m)}}^{-2}(\hat{\boldsymbol{\Omega}}_{A_2,m}) = O_P(\omega_{A_2,m}^{-2}),$$

which completes the proof. $\square$

Lemma 9 (Properties of $\Omega_{Q,m}$) For any $m, n \in [M]$ with $n \neq m$, define

$$\begin{aligned}\Omega_{Q,m} &:= \frac{1}{M-1} \sum_{n=1; n \neq m}^M \mathbf{S}_{Q,mn}^{\otimes 2}, \quad \mathbf{S}_{Q,mn} := \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \mathbf{Q}_n' \right) \\ \hat{\mathbf{S}}_{Q,mn} &:= \frac{1}{T} \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} , \quad \Delta_{Q,mn} := \hat{\mathbf{S}}_{Q,mn} - \mathbf{S}_{Q,mn}.\end{aligned}$$

Then under Assumptions (IC1), (E1), (E2), (F1), and (L1), we have

$$\begin{aligned}\|\Delta_{Q,mn}\|_F^2 &\lesssim_P p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} s_{J,m}^2 s_{J,n}^2 R(T, d_{fac,*}) + p_m^{\delta_{r,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 q_m^{\delta_{c,1}^{(m)}} R(T, d_{fac,*}) \\ &\quad + p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{r,1}^{(n)}} s_{J,m}^2 q_n^{\delta_{c,1}^{(n)}} R(T, d_{fac,*}) + p_m^{\delta_{r,1}^{(m)}} p_n^{\delta_{r,1}^{(n)}} q_m^{\delta_{c,1}^{(m)}} q_n^{\delta_{c,1}^{(n)}} R(T, d_{fac,*}) + T^{-1} p_m p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 \\ &\quad + T^{-1} p_m p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} + T^{-1} p_m^{\delta_{q,1}^{(m)}} s_{J,m}^2 p_n + T^{-1} p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} p_n + T^{-1} p_m p_n.\end{aligned}$$

Further define $\check{\mathbf{Q}}_m := \mathbf{Q}_m (\mathbf{Q}_m' \mathbf{Q}_m)^{-1/2}$. Then under Assumptions (L2) and (R2), there exists an asymptotically orthogonal matrix $\mathbf{H}_{Q,m}$ with $\mathbf{H}_{Q,m} \mathbf{H}_{Q,m}' = \mathbf{I} + o_P(1)$, such that

$$\|\hat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m}\|^2 = O_P \left\{ R(T, d_{fac,*}) p_m^{2(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right)^2 \right\}.$$

Analogous results hold for $\hat{\mathbf{J}}_m$, with asymptotically orthogonal matrix $\mathbf{H}_{J,m}$.

Proof of Lemma 9. By Assumption (F1), we note that $T^{-1} \sum_{t=1}^T \mathbb{E}[\mathbf{z}_{G,t}^{\otimes 2}]$ and $T^{-1} \sum_{t=1}^T \mathbb{E}[(\mathbf{z}_{F,t}^{(m)})^{\otimes 2}]$ are diagonal due to the independence between cross-sectional entries. Without loss of generality, we further normalize these long-run covariance matrices to be the identity matrix. Namely, we will assume throughout the proof that

$$\frac{1}{T} \sum_{t=1}^T \mathbb{E}[\mathbf{z}_{G,t}^{\otimes 2}] = \mathbf{I}_{k_r k_c}, \quad \frac{1}{T} \sum_{t=1}^T \mathbb{E}[(\mathbf{z}_{F,t}^{(m)})^{\otimes 2}] = \mathbf{I}_{k_r^{(m)} k_c^{(m)}}. \quad (\text{D.73})$$

This is equivalent to absorbing the entry-wise variance into $\Sigma_{Q,m}$, $\Sigma_{J,m}$, $\Sigma_{R,m}$ and $\Sigma_{C,m}$ (from Assumption (L1)) which are arbitrary positive-definite matrices with eigenvalues uniformly bounded away from 0 and infinity. Then for any group $m \in [M]$, by definitions we have

$$\begin{aligned}\check{\mathbf{Y}}_t^{(m)} &= \mathbf{Y}_t^{(m)} - \hat{\mathbf{c}}_t^{(m)} \mathbf{1}_{p_m} \mathbf{1}_{q_m}' - \hat{\mathbf{a}}_t^{(m)} \mathbf{1}_{q_m}' - \mathbf{1}_{p_m} \hat{\mathbf{b}}_t^{(m)'} \\ &= \mathbf{Y}_t^{(m)} - (p_m q_m)^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' - q_m^{-1} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' - p_m^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} \\ &= \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} + p_m^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} + q_m^{-1} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' - (p_m q_m)^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' \\ &\quad - q_m^{-1} \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{1}_{q_m} \mathbf{1}_{q_m}' - p_m^{-1} \mathbf{1}_{p_m} \mathbf{1}_{p_m}' \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} = \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m}.\end{aligned}$$

Together with our model (3.1), Assumption (IC1), and the fact that $\mathbf{M}_a \mathbf{1}_a = \mathbf{0}$ for any positive integer a , we further arrive at

$$\begin{aligned}\check{\mathbf{Y}}_t^{(m)} &= \mathbf{M}_{p_m} \mathbf{Y}_t^{(m)} \mathbf{M}_{q_m} = \mathbf{M}_{p_m} \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{M}_{q_m} + \mathbf{M}_{p_m} \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{M}_{q_m} + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \\ &= \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m}.\end{aligned}\quad (\text{D.74})$$

Then it holds that

$$\begin{aligned}\hat{\mathbf{S}}_{Q,mn} &= \frac{1}{T} \sum_{t=1}^T \left(\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{v}_b^{(m)} + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \right) \\ &\quad \cdot \left(\mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \mathbf{Q}_n' + \mathbf{v}_b^{(n)'} \mathbf{C}_n \mathbf{F}_t^{(n)'} \mathbf{R}_n' + \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \right) \\ &= \frac{1}{T} \sum_{t=1}^T \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \mathbf{Q}_n' + \frac{1}{T} \sum_{t=1}^T \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \mathbf{Q}_n' \\ &\quad + \frac{1}{T} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \mathbf{Q}_n' + \frac{1}{T} \sum_{t=1}^T \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{C}_n \mathbf{F}_t^{(n)'} \mathbf{R}_n' \\ &\quad + \frac{1}{T} \sum_{t=1}^T \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{C}_n \mathbf{F}_t^{(n)'} \mathbf{R}_n' + \frac{1}{T} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{C}_n \mathbf{F}_t^{(n)'} \mathbf{R}_n' \\ &\quad + \frac{1}{T} \sum_{t=1}^T \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} + \frac{1}{T} \sum_{t=1}^T \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \\ &\quad + \frac{1}{T} \sum_{t=1}^T \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{M}_{q_n} \mathbf{E}_t^{(n)'} \mathbf{M}_{p_n} \\ &=: \mathcal{I}_{Q,mn} + \mathcal{II}_{Q,mn} + \mathcal{III}_{Q,mn} + \mathcal{IV}_{Q,mn} + \mathcal{V}_{Q,mn} + \mathcal{VI}_{Q,mn} + \mathcal{VII}_{Q,mn} + \mathcal{VIII}_{Q,mn} + \mathcal{IX}_{Q,mn}.\end{aligned}\quad (\text{D.75})$$

Consider first $\mathcal{I}_{Q,mn} - \mathbf{S}_{Q,mn}$. Define $\mathbf{z}_{G,t} := \text{vec}(\mathbf{G}_t)$ for any $t \in [T]$. As the dimensions of $\mathbf{G}_t$ are bounded, Lemma 2 implies that

$$\left\| \frac{1}{T} \sum_{t=1}^T \{ \mathbf{z}_{G,t} \mathbf{z}_{G,t}' - \mathbb{E}[\mathbf{z}_{G,t} \mathbf{z}_{G,t}'] \} \right\|_F^2 = O_P(R(T, d_{\text{fac},*})), \quad (\text{D.76})$$

where $R(\cdot, \cdot)$ is defined in (4.1) and $d_{\text{fac},*}$ is the largest long-memory exponent among all processes in $\{\mathbf{G}_t\}$ and $\{\mathbf{F}_t^{(m)}\}$ for $m \in [M]$. Then from Assumption (L1), by the triangle inequality we have

$$\left\| \mathcal{I}_{Q,mn} - \mathbf{S}_{Q,mn} \right\|_F^2 = \left\| \mathbf{Q}_m \frac{1}{T} \sum_{t=1}^T \left(\mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' - \mathbb{E}[\mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t'] \right) \mathbf{Q}_n' \right\|_F^2$$

$$\begin{aligned}
&\lesssim \frac{1}{T^2} \cdot p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot \left\| \sum_{t=1}^T \left(\mathbf{G}_t \mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}'_t - \mathbb{E} \left[\mathbf{G}_t \mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}'_t \right] \right) \right\|_F^2 \\
&= \frac{1}{T^2} \cdot p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot \left\| \sum_{t=1}^T \left\{ (\mathbf{G}_t \otimes \mathbf{G}_t) \text{vec}(\mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n) - \mathbb{E} \left[(\mathbf{G}_t \otimes \mathbf{G}_t) \text{vec}(\mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n) \right] \right\} \right\|_F^2 \\
&\leq \frac{1}{T^2} \cdot p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot \left\| \sum_{t=1}^T \{ \mathbf{G}_t \otimes \mathbf{G}_t - \mathbb{E}[\mathbf{G}_t \otimes \mathbf{G}_t] \} \right\|_F^2 \cdot \left\| \mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \right\|_F^2 \\
&\asymp p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot \left\| \frac{1}{T} \sum_{t=1}^T \{ \mathbf{z}_{G,t} \mathbf{z}'_{G,t} - \mathbb{E}[\mathbf{z}_{G,t} \mathbf{z}'_{G,t}] \} \right\|_F^2 \cdot \left\| \mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \right\|_F^2 \\
&\lesssim_P p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot R(T, d_{\text{fac}, \star}) \cdot \left\| \mathbf{J}'_m \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \right\|_F^2 \\
&\lesssim_P p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot s_{J,m}^2 \cdot s_{J,n}^2 \cdot R(T, d_{\text{fac}, \star}), \tag{D.77}
\end{aligned}$$

where the first equivalence follows from the fact that the Frobenius norm of $\mathbf{G}_t \otimes \mathbf{G}_t$ equals that of $\mathbf{z}_{G,t} \mathbf{z}'_{G,t}$ under a fixed permutation of their entries, the second last inequality used (D.76), and the last used the definitions of $s_{J,m}$ and $s_{J,n}$.

By Assumption (L1), analogous arguments in (D.77), and the fact $\mathbf{v}_b^{(m)}$ and $\mathbf{v}_b^{(n)}$ are unit vectors by construction, we may also obtain

$$\|\mathcal{II}_{Q,mn}\|_F^2 \lesssim_P p_m^{\delta_r^{(m)}} \cdot p_n^{\delta_q^{(n)}} \cdot s_{J,n}^2 \cdot q_m^{\delta_c^{(m)}} \cdot R(T, d_{\text{fac}, \star}), \tag{D.78}$$

$$\|\mathcal{IV}_{Q,mn}\|_F^2 \lesssim_P p_m^{\delta_q^{(m)}} \cdot p_n^{\delta_r^{(n)}} \cdot s_{J,m}^2 \cdot q_n^{\delta_c^{(n)}} \cdot R(T, d_{\text{fac}, \star}), \tag{D.79}$$

$$\|\mathcal{V}_{Q,mn}\|_F^2 \lesssim_P p_m^{\delta_r^{(m)}} \cdot p_n^{\delta_r^{(n)}} \cdot q_m^{\delta_c^{(m)}} \cdot q_n^{\delta_c^{(n)}} \cdot R(T, d_{\text{fac}, \star}). \tag{D.80}$$

We next bound the noise-involved terms in (D.75), i.e., $\mathcal{III}_{Q,mn}$, $\mathcal{VI}_{Q,mn}$, $\mathcal{VII}_{Q,mn}$, $\mathcal{VIII}_{Q,mn}$, and $\mathcal{IX}_{Q,mn}$. For each $m \in [M]$, define $a_m := \mathbf{M}_{q_m} \mathbf{v}_b^{(m)}$ and $b_m := \mathbf{J}'_m \mathbf{v}_b^{(m)}$. By construction, $\|a_m\| \leq 1$ and $\|b_m\| = s_{J,m}$. We first present a simple but useful bound. For any deterministic vector $a \in \mathbb{R}^{q_m}$ with $\|a\| \leq 1$, Lemma 1 implies that, uniformly in t ,

$$\mathbb{E} \left\| \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} a \right\|^2 = O(p_m), \tag{D.81}$$

which used that for each row $i \in [p_m]$, the covariance matrix of the row vector $(\mathbf{E}_t^{(m)})_i$ has row-wise ℓ_1 -norms bounded by a constant under Lemma 1, and as the covariance matrix is symmetric and positive semidefinite, its spectral norm is bounded by the same constant, and thus the above bound follows by summing over $i \in [p_m]$. Also, by Assumptions (F1) and (L1), we have

$$\mathbb{E} \left\| \mathbf{Q}_n \mathbf{G}_t b_n \right\|^2 \lesssim p_n^{\delta_q^{(n)}} s_{J,n}^2, \quad \mathbb{E} \left\| \mathbf{R}_n \mathbf{F}_t^{(n)} \mathbf{C}'_n \mathbf{v}_b^{(n)} \right\|^2 \lesssim p_n^{\delta_r^{(n)}} q_n^{\delta_c^{(n)}}. \tag{D.82}$$

Consider $\mathcal{III}_{Q,mn}$. For neatness, define $u_{n,t} := \mathbf{Q}_n \mathbf{G}_t b_n$ and $e_{m,t} := \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} a_m$. Then we can write $\mathcal{III}_{Q,mn} = T^{-1} \sum_{t=1}^T e_{m,t} u'_{n,t}$. By Assumption (E2), the processes $\{e_{m,t}\}_{t \in [T]}$ and $\{u_{n,t}\}_{t \in [T]}$ are independent. Hence we have

$$\mathbb{E} \|\mathcal{III}_{Q,mn}\|_F^2 = \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \mathbb{E}(u'_{n,t} u_{n,s}) \mathbb{E}(e'_{m,t} e_{m,s}). \quad (\text{D.83})$$

We bound the two expectations on the right-hand side separately. First, by the Cauchy–Schwarz inequality and (D.82), we have

$$\left| \mathbb{E}(u'_{n,t} u_{n,s}) \right| \leq \left(\mathbb{E} \|u_{n,t}\|^2 \right)^{1/2} \left(\mathbb{E} \|u_{n,s}\|^2 \right)^{1/2} \lesssim p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2. \quad (\text{D.84})$$

Second, by the lagged version of (D.81), for any $t, s \in [T]$,

$$\left| \mathbb{E}(e'_{m,t} e_{m,s}) \right| \lesssim p_m \rho(|t - s|), \quad (\text{D.85})$$

where $\rho(\cdot)$ is the summable sequence from Lemma 1. Substituting (D.84) and (D.85) into (D.83) gives

$$\begin{aligned} \mathbb{E} \|\mathcal{III}_{Q,mn}\|_F^2 &\lesssim T^{-2} p_m p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 \sum_{t=1}^T \sum_{s=1}^T \rho(|t - s|) \\ &\lesssim T^{-2} \cdot p_m p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 \cdot T = T^{-1} p_m p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2, \end{aligned} \quad (\text{D.86})$$

where the second line follows from evaluating the double sum, viz. setting $h = t - s$ gives

$$\sum_{t=1}^T \sum_{s=1}^T \rho(|t - s|) = \sum_{h=-(T-1)}^{T-1} (T - |h|) \rho(|h|) \leq T \sum_{h=-\infty}^{\infty} \rho(|h|) = O(T),$$

where the final equality follows from the summability of $\rho(\cdot)$ as in Lemma 1. We mark that (D.86) carries T^{-1} rather than $R(T, d_{\text{fac}, \star})$ as in e.g. (D.77), since here we consider the cross-moment between the factor and noise processes, whereas previously, say (D.77), the sample covariance of factor processes is considered.

The same arguments can be applied to $\mathcal{VI}_{Q,mn}$, $\mathcal{VII}_{Q,mn}$, $\mathcal{VIII}_{Q,mn}$ and $\mathcal{IX}_{Q,mn}$, yielding

$$\|\mathcal{VI}_{Q,mn}\|_F^2 = O_P \left(T^{-1} p_m p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} \right), \quad (\text{D.87})$$

$$\|\mathcal{VII}_{Q,mn}\|_F^2 = O_P \left(T^{-1} p_m^{\delta_{q,1}^{(m)}} s_{J,m}^2 p_n \right), \quad (\text{D.88})$$

$$\|\mathcal{VIII}_{Q,mn}\|_F^2 = O_P \left(T^{-1} p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} p_n \right), \quad (\text{D.89})$$

$$\|\mathcal{IX}_{Q,mn}\|_F^2 = O_P \left(T^{-1} p_m p_n \right), \quad (\text{D.90})$$

where in each case, the bound follows by separating the independent vector processes, applying (D.81) to the error part, (D.82) to the factor part, and using the short-memory property of the error process to sum over t, s . Lastly, combining (D.75), (D.77), (D.78), (D.86), (D.79), (D.80), (D.87), (D.88), (D.89), and (D.90), together with the definitions of $\Delta_{Q,mn}$, we may conclude that

$$\begin{aligned}
\|\Delta_{Q,mn}\|_F^2 &= \|\widehat{\mathbf{S}}_{Q,mn} - \mathbf{S}_{Q,mn}\|_F^2 \\
&\lesssim \left\| \mathcal{I}_{Q,mn} - \mathbf{S}_{Q,mn} \right\|_F^2 + \|\mathcal{II}_{Q,mn}\|_F^2 + \|\mathcal{III}_{Q,mn}\|_F^2 + \|\mathcal{IV}_{Q,mn}\|_F^2 + \|\mathcal{V}_{Q,mn}\|_F^2 \\
&\quad + \|\mathcal{VI}_{Q,mn}\|_F^2 + \|\mathcal{VII}_{Q,mn}\|_F^2 + \|\mathcal{VIII}_{Q,mn}\|_F^2 + \|\mathcal{IX}_{Q,mn}\|_F^2 \\
&\lesssim_P p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} s_{J,m}^2 s_{J,n}^2 R(T, d_{\text{fac},*}) + p_m^{\delta_{r,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 q_m^{\delta_{c,1}^{(m)}} R(T, d_{\text{fac},*}) + p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{r,1}^{(n)}} s_{J,m}^2 q_n^{\delta_{c,1}^{(n)}} R(T, d_{\text{fac},*}) \\
&\quad + p_m^{\delta_{r,1}^{(m)}} p_n^{\delta_{r,1}^{(n)}} q_m^{\delta_{c,1}^{(m)}} q_n^{\delta_{c,1}^{(n)}} R(T, d_{\text{fac},*}) + T^{-1} p_m p_n^{\delta_{q,1}^{(n)}} s_{J,n}^2 + T^{-1} p_m p_n^{\delta_{r,1}^{(n)}} q_n^{\delta_{c,1}^{(n)}} \\
&\quad + T^{-1} p_m^{\delta_{q,1}^{(m)}} s_{J,m}^2 p_n + T^{-1} p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} p_n + T^{-1} p_m p_n.
\end{aligned}$$

This completes the proof on $\|\Delta_{Q,mn}\|_F^2$.

Next, we focus on the k_r -th eigenvalue of $\Omega_{Q,m}$. First, by Assumption (F1), for any $m, n \in [M]$, we may simplify and write

$$\begin{aligned}
\mathbf{S}_{Q,mn} &= \mathbf{Q}_m \mathbb{E} \left(\mathbf{G}_t \mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{G}_t' \right) \mathbf{Q}_n' \\
&= \text{tr} \left(\mathbf{J}_m' \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \mathbf{J}_n \right) \mathbf{Q}_m \mathbf{Q}_n' = \left(\mathbf{v}_b^{(n)'} \mathbf{J}_n \mathbf{J}_m' \mathbf{v}_b^{(m)} \right) \mathbf{Q}_m \mathbf{Q}_n' =: w_{mn} \mathbf{Q}_m \mathbf{Q}_n',
\end{aligned} \tag{D.91}$$

where by Assumption (L2), the selected projection directions satisfy

$$|w_{mn}| \asymp s_{J,m} s_{J,n} \tag{D.92}$$

for all $m, n \in [M]$ with $n \neq m$. Hence we have $\mathbf{S}_{Q,mn}$ of rank k_r , and

$$\|\mathbf{S}_{Q,mn}\| \asymp_P \left(s_{J,m} s_{J,n} p_m^{\delta_{q,1}^{(m)}/2} p_n^{\delta_{q,1}^{(n)}/2} \right). \tag{D.93}$$

Now plugging (D.91) in $\Omega_{Q,m}$, we have

$$\begin{aligned}
\Omega_{Q,m} &= \frac{1}{M-1} \sum_{n=1; n \neq m}^M w_{mn}^2 \mathbf{Q}_m \mathbf{Q}_n' \mathbf{Q}_n \mathbf{Q}_m' \\
&= \mathbf{Q}_m \left\{ \frac{1}{M-1} \sum_{n=1; n \neq m}^M w_{mn}^2 \mathbf{Q}_n' \mathbf{Q}_n \right\} \mathbf{Q}_m' =: \mathbf{Q}_m \mathbf{G}_{Q,m} \mathbf{Q}_m'.
\end{aligned} \tag{D.94}$$

We next derive the order of the k_r -th eigenvalue of $\mathbf{\Omega}_{Q,m}$. For any nonzero vector $\mathbf{x} \in \mathbb{R}^{k_r}$,

$$\begin{aligned} \mathbf{x}' \mathbf{G}_{Q,m} \mathbf{x} &= \frac{1}{M-1} \sum_{n=1; n \neq m}^M w_{mn}^2 \mathbf{x}' \mathbf{Q}'_n \mathbf{Q}_n \mathbf{x} = \frac{1}{M-1} \sum_{n=1; n \neq m}^M w_{mn}^2 \|\mathbf{Q}_n \mathbf{x}\|^2 \\ &\geq \frac{1}{M-1} \sum_{n=1; n \neq m}^M w_{mn}^2 \lambda_{\min}(\mathbf{Q}'_n \mathbf{Q}_n) \|\mathbf{x}\|^2. \end{aligned} \quad (\text{D.95})$$

By Assumption (L1), $\lambda_{\min}(\mathbf{Q}'_n \mathbf{Q}_n) \asymp p_n^{\delta_{q,k_r}^{(n)}}$ for each $n \in [M]$. Combining this with (D.92) and (D.95) gives

$$\lambda_{\min}(\mathbf{G}_{Q,m}) \gtrsim \frac{s_{J,m}^2}{M-1} \sum_{n=1; n \neq m}^M s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}} \asymp s_{J,m}^2 \max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}, \quad (\text{D.96})$$

where also used that M is fixed. As the right-hand side of (D.96) is bounded away from zero, the matrix $\mathbf{G}_{Q,m}$ is positive definite with all eigenvalues bounded away from zero. Then from (D.94), the nonzero eigenvalues of $\mathbf{\Omega}_{Q,m}$ are the same as those of $\mathbf{G}_{Q,m}^{1/2} \mathbf{Q}'_m \mathbf{Q}_m \mathbf{G}_{Q,m}^{1/2}$. Therefore, we may conclude that

$$\lambda_{k_r}(\mathbf{\Omega}_{Q,m}) \geq \lambda_{\min}(\mathbf{G}_{Q,m}) \lambda_{\min}(\mathbf{Q}'_m \mathbf{Q}_m) \asymp \lambda_{\min}(\mathbf{G}_{Q,m}) p_m^{\delta_{q,k_r}^{(m)}} \gtrsim p_m^{\delta_{q,k_r}^{(m)}} s_{J,m}^2 \max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}, \quad (\text{D.97})$$

by Assumption (L1) and (D.96), which particularly implies $\mathbf{\Omega}_{Q,m}$ has rank k_r .

Next, note that by Assumption (R2), we may further simplify $\|\mathbf{\Delta}_{Q,mn}\|_F^2$ as

$$\|\mathbf{\Delta}_{Q,mn}\|_F^2 = O_P\left(R(T, d_{\text{fac},*}) p_m^{\delta_{q,1}^{(m)}} p_n^{\delta_{q,1}^{(n)}} s_{J,m}^2 s_{J,n}^2\right) = O_P\left(R(T, d_{\text{fac},*}) \|\mathbf{S}_{Q,mn}\|_F^2\right),$$

where the last equality used (D.93). Hence by (3.8) and the triangle inequality,

$$\begin{aligned} \|\mathbf{X}_{Q,m} - \mathbf{\Omega}_{Q,m}\| &= \left\| \frac{1}{T^2(M-1)} \sum_{n=1; n \neq m}^M \left\{ \sum_{t=1}^T \check{\mathbf{Y}}_t^{(m)} \mathbf{v}_b^{(m)} \mathbf{v}_b^{(n)'} \check{\mathbf{Y}}_t^{(n)'} \right\}^{\otimes 2} - \mathbf{\Omega}_{Q,m} \right\| \\ &= \left\| \left(\frac{1}{M-1} \sum_{n=1; n \neq m}^M \hat{\mathbf{S}}_{Q,mn}^{\otimes 2} \right) - \mathbf{\Omega}_{Q,m} \right\| \leq \frac{1}{M-1} \sum_{n=1; n \neq m}^M \left\| \hat{\mathbf{S}}_{Q,mn} \hat{\mathbf{S}}'_{Q,mn} - \mathbf{S}_{Q,mn} \mathbf{S}'_{Q,mn} \right\| \\ &\leq \frac{1}{M-1} \sum_{n=1; n \neq m}^M \left\| \mathbf{\Delta}_{Q,mn} \hat{\mathbf{S}}'_{Q,mn} + \mathbf{S}_{Q,mn} \mathbf{\Delta}'_{Q,mn} \right\| \\ &= \frac{1}{M-1} \sum_{n=1; n \neq m}^M \left\| \mathbf{\Delta}_{Q,mn} \mathbf{\Delta}'_{Q,mn} + \mathbf{\Delta}_{Q,mn} \mathbf{S}'_{Q,mn} + \mathbf{S}_{Q,mn} \mathbf{\Delta}'_{Q,mn} \right\| \\ &\lesssim \max_{n \in [M] \setminus \{m\}} (\|\mathbf{\Delta}_{Q,mn}\|^2 + \|\mathbf{S}_{Q,mn}\| \|\mathbf{\Delta}_{Q,mn}\|) \\ &\lesssim R(T, d_{\text{fac},*})^{1/2} \max_{n \in [M] \setminus \{m\}} \|\mathbf{S}_{Q,mn}\|_F^2 \asymp R(T, d_{\text{fac},*})^{1/2} p_m^{\delta_{q,1}^{(m)}} s_{J,m}^2 \max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}, \end{aligned}$$

which, combined with (D.97), applying the Davis–Kahan theorem (Yu et al., 2015, Theorem 2) gives the

desired result on $\widehat{\mathbf{Q}}_m$, and thus concludes this proof. $\square$

Lemma 10 *Let all assumptions in Theorem 7 hold. Fix any $m \in [M]$. Recall $\check{\mathbf{Q}}_m = \mathbf{Q}_m(\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2}$ from Lemma 9, and similarly let $\check{\mathbf{J}}_m := \mathbf{J}_m(\mathbf{J}'_m \mathbf{J}_m)^{-1/2}$. Also, recall $\mathbf{H}_{Q,m}$ and $\mathbf{H}_{J,m}$ from Lemma 9. For each $m \in [M]$, with $\check{\mathbf{Y}}_{G,t}^{(m)} = \check{\mathbf{Y}}_t^{(m)} - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}'_m$ from Step 3 of Section 3.3, define*

$$\check{\mathbf{G}}_{t,m} := \check{\mathbf{Q}}'_m \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m \check{\mathbf{J}}_m, \quad \widehat{\mathbf{G}}_{t,m} := \widehat{\mathbf{Q}}'_m \check{\mathbf{Y}}_{G,t}^{(m)} \widehat{\mathbf{J}}_m.$$

Then for every $t \in [T]$, $\widehat{\mathbf{G}}_{t,m}$ satisfies that

$$\begin{aligned} & \left\| \widehat{\mathbf{G}}_{t,m} - \mathbf{H}'_{Q,m} \check{\mathbf{G}}_{t,m} \mathbf{H}_{J,m} \right\|_F^2 \\ &= O_P \left(\min(p_m, q_m) + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + (\eta_{Q,m} + \eta_{J,m}) p_m q_m \right), \end{aligned}$$

where $\Delta_{R,m}$, $\Delta_{C,m}$, $\Delta_{F,m}$, $\eta_{Q,m}$, and $\eta_{J,m}$ are defined in the statement of Theorem 7.

Proof of Lemma 10. By construction, we may decompose

$$\widehat{\mathbf{G}}_{t,m} = \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m \mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m \check{\mathbf{J}}_m \mathbf{H}_{J,m} + \mathcal{R}_{Q,m} + \mathcal{R}_{J,m} + \mathcal{R}_{E,m} + \mathcal{R}_{\text{local},m}, \quad (\text{D.98})$$

where

$$\begin{aligned} \mathcal{R}_{Q,m} &:= (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m})' (\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}'_m) \widehat{\mathbf{J}}_m, \\ \mathcal{R}_{J,m} &:= \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m (\mathbf{Q}_m \mathbf{G}_t \mathbf{J}'_m + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}'_m) (\widehat{\mathbf{J}}_m - \check{\mathbf{J}}_m \mathbf{H}_{J,m}), \\ \mathcal{R}_{E,m} &:= \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \check{\mathbf{J}}_m \mathbf{H}_{J,m}, \\ \mathcal{R}_{\text{local},m} &:= \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m (\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}'_m - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}'_m) \check{\mathbf{J}}_m \mathbf{H}_{J,m}. \end{aligned}$$

Note that $\check{\mathbf{Q}}_m$ and $\check{\mathbf{J}}_m$ have orthonormal columns, $\mathbf{H}_{Q,m}$ and $\mathbf{H}_{J,m}$ are asymptotically orthogonal, and $\widehat{\mathbf{J}}_m$ has orthonormal columns. Hence we have

$$\|\check{\mathbf{Q}}_m\| = \|\check{\mathbf{J}}_m\| = \|\widehat{\mathbf{J}}_m\| = 1, \quad \|\mathbf{H}_{Q,m}\| = \|\mathbf{H}_{J,m}\| = O_P(1). \quad (\text{D.99})$$

Then consider $\mathcal{R}_{E,m}$. As the columns of $\check{\mathbf{J}}_m \mathbf{H}_{J,m}$ are asymptotically orthonormal, by (D.81), it holds that

$$\|\mathcal{R}_{E,m}\|_F^2 \leq \|\mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m\|^2 \cdot \|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \check{\mathbf{J}}_m \mathbf{H}_{J,m}\|_F^2 \lesssim_P \min(p_m, q_m), \quad (\text{D.100})$$

where the equality follows from (D.99) and the fact that $\|\mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} \check{\mathbf{J}}_m \mathbf{H}_{J,m}\|_F^2 = O_P(p_m)$ as in the

proof of Theorem 7, and a symmetrical argument which also gives the rate in $O_P(q_m)$.

Next, consider $\mathcal{R}_{\text{local},m}$. By Theorem 7, we have

$$\begin{aligned}
& \|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}_m'\|_F^2 \\
& \lesssim \|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \widehat{\mathbf{R}}_m \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m} \widehat{\mathbf{C}}_m'\|_F^2 + \|\widehat{\mathbf{R}}_m (\widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}) \widehat{\mathbf{C}}_m'\|_F^2 \\
& = \|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \widehat{\mathbf{R}}_m \mathbf{H}_{R,m}^{-1} (\mathbf{R}_m' \mathbf{R}_m)^{1/2} \mathbf{F}_t^{(m)} (\mathbf{C}_m' \mathbf{C}_m)^{1/2} \mathbf{H}_{C,m} \widehat{\mathbf{C}}_m'\|_F^2 \\
& \quad + \|\widehat{\mathbf{R}}_m (\widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}) \widehat{\mathbf{C}}_m'\|_F^2 \\
& \lesssim_P \|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \mathbf{R}_m (\mathbf{R}_m' \mathbf{R}_m)^{-1/2} (\mathbf{R}_m' \mathbf{R}_m)^{1/2} \mathbf{F}_t^{(m)} (\mathbf{C}_m' \mathbf{C}_m)^{1/2} (\mathbf{C}_m' \mathbf{C}_m)^{-1/2} \mathbf{C}_m'\|_F^2 \\
& \quad + \|\widehat{\mathbf{R}}_m (\widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}) \widehat{\mathbf{C}}_m'\|_F^2 \\
& \lesssim_P \|\widehat{\mathbf{R}}_m (\widehat{\mathbf{F}}_t^{(m)} - \mathbf{H}_{R,m}^{-1} \check{\mathbf{F}}_t^{(m)} \mathbf{H}_{C,m}) \widehat{\mathbf{C}}_m'\|_F^2 \lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}), \tag{D.101}
\end{aligned}$$

where $\Delta_{R,m}$, $\Delta_{C,m}$ and $\Delta_{F,m}$ are defined in the statement of Theorem 7. Combining (D.99) and (D.101), we thus have

$$\begin{aligned}
\|\mathcal{R}_{\text{local},m}\|_F^2 & \leq \|\mathbf{H}_{Q,m}' \check{\mathbf{Q}}_m'\|_F^2 \cdot \|\mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}_m'\|_F^2 \cdot \|\check{\mathbf{J}}_m \mathbf{H}_{J,m}\|_F^2 \\
& \lesssim_P p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}). \tag{D.102}
\end{aligned}$$

It remains to consider $\mathcal{R}_{Q,m}$ and $\mathcal{R}_{J,m}$. For $\mathcal{R}_{Q,m}$, by Assumptions (F1), (L1), (D.35), and (D.101), we have

$$\begin{aligned}
& \|\mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' + \mathbf{M}_{p_m} \mathbf{E}_t^{(m)} \mathbf{M}_{q_m} + \mathbf{R}_m \mathbf{F}_t^{(m)} \mathbf{C}_m' - \widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}_m'\|_F^2 \\
& \lesssim_P p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} + p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}),
\end{aligned}$$

which, combined with Lemma 9 and (D.99), give

$$\|\mathcal{R}_{Q,m}\|_F^2 \lesssim_P \eta_{Q,m} \left\{ p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} + p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) \right\}. \tag{D.103}$$

Similarly, $\mathcal{R}_{J,m}$ satisfies that

$$\|\mathcal{R}_{J,m}\|_F^2 \lesssim_P \eta_{J,m} \left\{ p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} + p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) \right\}. \tag{D.104}$$

Finally, combining (D.98), (D.100), (D.102), (D.103), and (D.104), we conclude that

$$\|\widehat{\mathbf{G}}_{t,m} - \mathbf{H}_{Q,m}' \check{\mathbf{Q}}_m' \mathbf{Q}_m \mathbf{G}_t \mathbf{J}_m' \check{\mathbf{J}}_m \mathbf{H}_{J,m}\|_F^2$$

$$\begin{aligned}
&\lesssim_P \min(p_m, q_m) + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) \\
&\quad + (\eta_{Q,m} + \eta_{J,m}) \left\{ p_m^{\delta_{q,1}^{(m)}} q_m^{\delta_{j,1}^{(m)}} + p_m q_m + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) \right\} \\
&\lesssim_P \min(p_m, q_m) + p_m^{\delta_{r,1}^{(m)}} q_m^{\delta_{c,1}^{(m)}} (\Delta_{R,m} + \Delta_{C,m} + \Delta_{F,m}) + (\eta_{Q,m} + \eta_{J,m}) p_m q_m,
\end{aligned}$$

which completes the proof of Lemma 10. $\square$

Lemma 11 *By construction in Step 3 of Section 3.3, it holds that*

$$\mathcal{R}(\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}) = \widehat{\sigma}_{\Xi,m} \widehat{\mathbf{u}}_{\Xi,m} \widehat{\mathbf{v}}'_{\Xi,m}.$$

Proof of Lemma 11. Recall from Step 3 of Section 3.3 the definition of the notation $\mathcal{R}(\mathbf{W})$ for a block matrix $\mathbf{W} = (\mathbf{W}_{ij})_{i \in [k_c], j \in [k_c]}$, with each sub-matrix $\mathbf{W}_{ij}$ of size $k_r \times k_r$, such that

$$\mathcal{R}(\mathbf{W}) = (\mathbf{vec}(\mathbf{W}_{11}), \dots, \mathbf{vec}(\mathbf{W}_{k_c 1}), \dots, \mathbf{vec}(\mathbf{W}_{1 k_c}), \dots, \mathbf{vec}(\mathbf{W}_{k_c k_c})).$$

Then note that each sub-matrix in $\widehat{\Xi}_{J,m} \otimes \widehat{\Xi}_{Q,m}$ is exactly the corresponding element in $\widehat{\Xi}_{J,m}$ multiplied by the matrix $\widehat{\Xi}_{Q,m}$; hence the vectorization of such sub-matrix would be the corresponding element in $\widehat{\Xi}_{J,m}$ multiplied by $\mathbf{vec}(\widehat{\Xi}_{Q,m}) = \widehat{\sigma}_{\Xi,m}^{1/2} \widehat{\mathbf{u}}_{\Xi,m}$ by construction. Together with the construction of $\widehat{\Xi}_{J,m}$, we complete the proof of this lemma. $\square$

Lemma 12 *Let all assumptions in Lemma 9 hold, and recall the matrices $\mathbf{H}_{Q,m}$ and $\mathbf{H}_{J,m}$ within. Then*

$$\mathbf{H}'_{Q,m} \mathbf{H}_{Q,m} = \mathbf{I}_{k_r} + O_P(\eta_{Q,m}^{1/2}), \quad \mathbf{H}'_{J,m} \mathbf{H}_{J,m} = \mathbf{I}_{k_c} + O_P(\eta_{J,m}^{1/2}),$$

where $\eta_{Q,m}$ and $\eta_{J,m}$ are defined in Theorem 7.

Proof of Lemma 12. For $\mathbf{H}_{Q,m}$, note that

$$\begin{aligned}
\mathbf{I}_{k_r} &= \widehat{\mathbf{Q}}'_m \widehat{\mathbf{Q}}_m = (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m})' \widehat{\mathbf{Q}}_m + \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m \widehat{\mathbf{Q}}_m \\
&= (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m})' \widehat{\mathbf{Q}}_m + \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m}) + \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m \check{\mathbf{Q}}_m \mathbf{H}_{Q,m} \\
&= (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m})' \widehat{\mathbf{Q}}_m + \mathbf{H}'_{Q,m} \check{\mathbf{Q}}'_m (\widehat{\mathbf{Q}}_m - \check{\mathbf{Q}}_m \mathbf{H}_{Q,m}) \\
&\quad + \mathbf{H}'_{Q,m} (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{Q}'_m \mathbf{Q}_m (\mathbf{Q}'_m \mathbf{Q}_m)^{-1/2} \mathbf{H}_{Q,m} \\
&= \mathbf{H}'_{Q,m} \mathbf{H}_{Q,m} + O_P \left\{ R(T, d_{\text{fac},*})^{1/2} p_m^{(\delta_{q,1}^{(m)} - \delta_{q,k_r}^{(m)})} \left(\frac{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,1}^{(n)}}}{\max_{n \in [M] \setminus \{m\}} s_{J,n}^2 p_n^{\delta_{q,k_r}^{(n)}}} \right) \right\} \\
&= \mathbf{H}'_{Q,m} \mathbf{H}_{Q,m} + O_P(\eta_{Q,m}^{1/2}),
\end{aligned}$$

where the last equality used Lemma 9, and the $O_P(\cdot)$ bound is in terms of the Frobenius norm. Analogous arguments hold for $\mathbf{H}_{J,m}$. This thus completes the proof of this lemma. $\square$

E Additional Numerical Results

E.1 Simulation

E.1.1 Heavy-tailed robustness

We repeat the main effect settings (Ia)–(Id) and common component settings (IIa)–(IIc) under heavy-tailed innovations. The DGP is identical to that in Section 6.1, except that the Gaussian innovations in the latent time-series processes are replaced by Student $t(6)$ innovations. Each setting is crossed with $T \in \{100, 400\}$, with 500 Monte Carlo replications.

The main effect results in Table E.1 remain close to their Gaussian counterparts. The direct moment estimators are essentially unchanged, while the loading-space and reconstructed-component errors retain the same improvement with T ; heavy tails mainly increase finite-sample dispersion in some entries.

Table E.1: Main effect estimation accuracy. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses. The latent-process innovations follow a Student $t(6)$ distribution.

Settings	Ia		Ib		Ic		Id	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$
$\text{MSE}(\hat{c}_t^{(m)})$	0.05(0.01)	0.05(0.00)	0.04(0.01)	0.04(0.00)	0.05(0.01)	0.05(0.00)	0.02(0.00)	0.02(0.00)
$\text{MSE}(\hat{\mathbf{a}}_t^{(m)})$	0.60(0.08)	0.58(0.06)	0.49(0.07)	0.47(0.05)	1.19(0.17)	1.14(0.12)	0.30(0.04)	0.29(0.03)
$\text{MSE}(\hat{\mathbf{b}}_t^{(m)})$	0.60(0.08)	0.58(0.06)	0.49(0.07)	0.47(0.05)	1.19(0.17)	1.16(0.12)	0.60(0.07)	0.58(0.04)
$\text{MSE}(\hat{\mu}_t)$	0.03(0.01)	0.03(0.00)	0.02(0.01)	0.02(0.00)	0.03(0.01)	0.03(0.00)	0.01(0.00)	0.01(0.00)
$\text{MSE}(\hat{\mu}_t^{(m)})$	0.08(0.02)	0.08(0.01)	0.06(0.01)	0.06(0.01)	0.08(0.02)	0.08(0.01)	0.04(0.01)	0.04(0.00)
$\mathcal{D}(\mathbf{A}_{1m}, \hat{\mathbf{A}}_{1m})$	7.04(5.98)	2.80(1.13)	9.81(10.06)	3.22(1.73)	7.37(5.11)	2.97(1.01)	7.38(6.06)	2.70(1.15)
$\mathcal{D}(\mathbf{A}_{1m}, \tilde{\mathbf{A}}_{1m})$	7.03(6.01)	2.79(1.13)	9.80(10.07)	3.21(1.73)	7.34(5.10)	2.96(1.00)	7.37(6.06)	2.69(1.15)
$\mathcal{D}(\mathbf{A}_{2m}, \hat{\mathbf{A}}_{2m})$	6.33(4.15)	2.74(0.92)	8.23(6.34)	3.15(1.16)	6.68(2.95)	2.94(0.82)	6.41(3.80)	2.58(0.87)
$\mathcal{D}(\mathbf{B}_{1m}, \hat{\mathbf{B}}_{1m})$	6.95(5.47)	2.74(1.10)	9.76(10.80)	3.26(1.57)	7.73(5.95)	2.89(1.04)	5.12(3.82)	2.03(0.65)
$\mathcal{D}(\mathbf{B}_{1m}, \tilde{\mathbf{B}}_{1m})$	6.93(5.47)	2.73(1.10)	9.74(10.80)	3.26(1.57)	7.70(5.94)	2.87(1.04)	5.11(3.82)	2.02(0.64)
$\mathcal{D}(\mathbf{B}_{2m}, \hat{\mathbf{B}}_{2m})$	6.19(3.70)	2.70(0.96)	8.25(7.19)	3.19(1.23)	6.90(3.66)	2.89(0.88)	4.57(2.53)	1.99(0.58)
$\text{MSE}(\tilde{\mathbf{A}}_{1m}\hat{\alpha}_t)$	0.79(5.71)	0.10(0.07)	1.58(6.01)	0.12(0.15)	0.62(1.22)	0.14(0.05)	0.73(3.14)	0.08(0.06)
$\text{MSE}(\hat{\mathbf{A}}_{2m}\hat{\alpha}_t^{(m)})$	0.83(2.89)	0.19(0.08)	1.56(4.20)	0.20(0.13)	0.84(1.12)	0.28(0.07)	0.78(2.17)	0.13(0.07)
$\text{MSE}(\tilde{\mathbf{B}}_{1m}\hat{\beta}_t)$	0.65(2.98)	0.10(0.06)	1.60(5.63)	0.12(0.10)	0.83(3.69)	0.13(0.06)	0.29(0.94)	0.05(0.02)
$\text{MSE}(\hat{\mathbf{B}}_{2m}\hat{\beta}_t^{(m)})$	0.77(2.69)	0.18(0.08)	1.75(5.34)	0.21(0.13)	0.98(2.72)	0.28(0.08)	0.38(1.30)	0.09(0.03)

The common component results in Table E.2 show the same robustness. Refitting remains effective and the estimation errors continue to decrease with T , while some small-sample entries exhibit greater Monte Carlo variation.

We also repeat the $S_{G,m}$ behavior experiment in Section 6.1.4 under the baseline design (IIIa), using 400 Monte Carlo replications and replacing the Gaussian innovations by Student $t(6)$ innovations. Figure E.1 shows that the active and inactive group averages remain clearly separated at $T = 200$, with a wider gap as more groups become active. Figure E.2 shows the same pattern over T : the inactive averages decrease, while the active averages remain comparatively stable. These results closely match the Gaussian baseline.

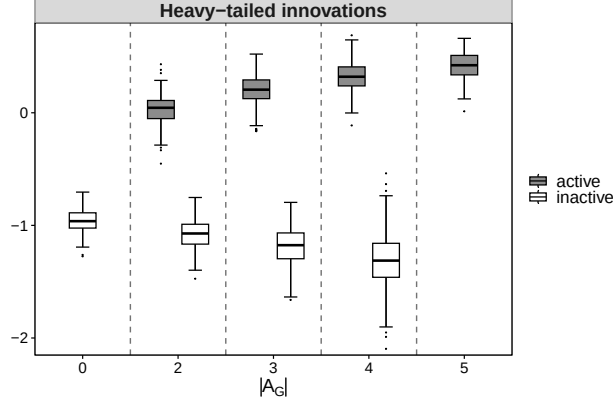


Figure E.1: Within-replication averages of $S_{G,m}$ over the active and inactive groups at $T = 200$ under the baseline design with Student $t(6)$ innovations. Each boxplot is formed across Monte Carlo replications. The y-axis reports the base-10 logarithm of the group average.

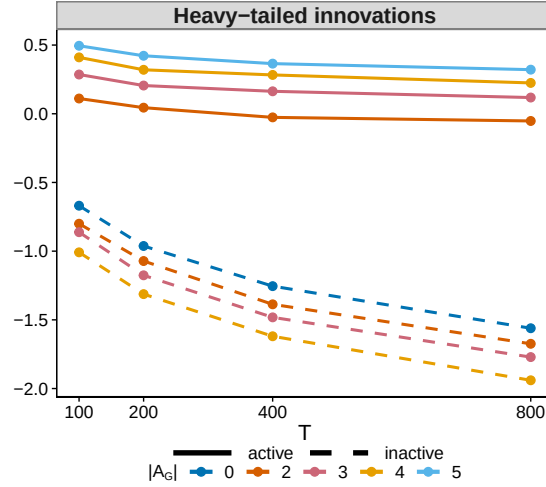


Figure E.2: Monte Carlo medians of the within-replication averages of $S_{G,m}$ over the active and inactive groups, respectively, over $T \in \{100, 200, 400, 800\}$ under Student $t(6)$ innovations. The y-axis reports the statistics on the $\log_{10}$ scale. Solid and dashed lines correspond to active and inactive groups, respectively.

Table E.2: Common component estimation accuracy. All entries are Monte Carlo means multiplied by 10^2 , with standard deviations in parentheses. The latent-process innovations follow a Student $t(6)$ distribution.

Settings	IIa		IIb		IIc		IIId	
	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$	$T = 100$	$T = 400$
$\mathcal{D}(\mathbf{Q}_m, \widehat{\mathbf{Q}}_m)$	8.14(3.45)	3.72(1.32)	9.66(3.98)	4.69(1.89)	9.28(3.41)	4.45(1.44)	5.85(2.47)	2.75(1.01)
$\mathcal{D}(\mathbf{Q}_m, \widetilde{\mathbf{Q}}_m)$	1.84(0.46)	0.79(0.14)	1.97(1.54)	0.81(0.21)	3.22(0.64)	1.45(0.19)	1.16(0.20)	0.50(0.07)
$\mathcal{D}(\mathbf{J}_m, \widehat{\mathbf{J}}_m)$	8.41(3.41)	3.85(1.45)	9.92(4.38)	4.61(1.76)	9.43(3.51)	4.33(1.41)	7.92(3.03)	3.87(1.43)
$\mathcal{D}(\mathbf{J}_m, \widetilde{\mathbf{J}}_m)$	1.60(0.23)	0.71(0.08)	1.67(1.42)	0.67(0.08)	3.08(0.71)	1.40(0.17)	1.48(0.16)	0.67(0.06)
$\mathcal{D}(\mathbf{R}_m, \widehat{\mathbf{R}}_m)$	2.60(0.75)	1.01(0.17)	3.23(1.89)	1.04(0.22)	4.44(1.87)	2.43(0.57)	2.20(0.70)	0.78(0.15)
$\mathcal{D}(\mathbf{C}_m, \widehat{\mathbf{C}}_m)$	2.57(0.71)	1.00(0.17)	3.14(1.78)	1.05(0.21)	4.33(0.96)	2.41(0.57)	2.13(0.51)	0.85(0.11)
$\text{MSE}(\widetilde{\mathbf{Q}}_m \widehat{\mathbf{G}}_t \widetilde{\mathbf{J}}_m')$	0.07(0.03)	0.04(0.01)	0.19(2.72)	0.03(0.01)	0.26(0.31)	0.16(0.02)	0.04(0.01)	0.02(0.00)
$\text{MSE}(\widehat{\mathbf{R}}_m \widehat{\mathbf{F}}_t^{(m)} \widehat{\mathbf{C}}_m')$	0.27(0.15)	0.13(0.03)	0.42(2.05)	0.12(0.03)	0.76(0.60)	0.48(0.12)	0.14(0.06)	0.06(0.01)

E.1.2 Order-statistic calibration for detection and screening

We further examine the choice of order statistic in the bootstrap procedures of Section 5. For existence testing, we compare tests based on $S_{G,(1)}$ and $S_{G,(2)}$, each calibrated by its corresponding bootstrap distribution. For screening, we compare two procedures indexed by $j \in \{1, 2\}$. The $S_{G,(j)}$ -based procedure sets both $\tau_1^{(j)}$ and $\tau_2^{(j)}$ equal to the empirical 95% quantile of $S_{G,(j)}^*$ and applies the corresponding existence gate and groupwise screening rule. Since $\tau_1^{(j)} = \tau_2^{(j)}$, the existence gate is automatically satisfied whenever at least two groups exceed the screening cutoff. All results below use the baseline setting (IIIa), 400 Monte Carlo replications, and $B = 199$ bootstrap replications.

Table E.3 shows that the two existence tests have very similar finite-sample performance. Both maintain size close to the nominal 5% level and have nearly identical power across the alternatives.

Table E.3: Bootstrap existence tests based on $S_{G,(1)}$ and $S_{G,(2)}$ in the baseline setting (IIIa). Each statistic is calibrated by its corresponding bootstrap distribution. Entries are empirical rejection rates, with $|\mathcal{A}_G| = 0$ representing size and $|\mathcal{A}_G| \geq 2$ representing power.

T	Statistic	$ \mathcal{A}_G = 0$	$ \mathcal{A}_G = 2$	$ \mathcal{A}_G = 3$	$ \mathcal{A}_G = 4$	$ \mathcal{A}_G = 5$
100	$S_{G,(1)}$	0.048	0.800	0.993	1.000	1.000
	$S_{G,(2)}$	0.043	0.828	0.993	1.000	1.000
200	$S_{G,(1)}$	0.058	0.998	1.000	1.000	1.000
	$S_{G,(2)}$	0.058	0.995	1.000	1.000	1.000
400	$S_{G,(1)}$	0.058	0.998	1.000	1.000	1.000
	$S_{G,(2)}$	0.050	0.998	1.000	1.000	1.000

The difference is more pronounced for screening. As shown in Table E.4, the $S_{G,(1)}$ -based procedure is substantially more conservative: it further reduces false positives under the null but yields noticeably lower true positive and exact-selection rates, especially when T is small or few groups are active. The gap narrows as T and the number of active groups increase. For $|\mathcal{A}_G| = 3$ and 4, false positives are zero under both procedures in the baseline design, so the differences in exact recovery are driven by missed active

groups.

Table E.4: Bootstrap screening procedures based on $S_{G,(1)}$ and $S_{G,(2)}$ in the baseline setting (IIIa). For each procedure, the same bootstrap order statistic is used to calibrate both the existence and screening cutoffs.

$ \mathcal{A}_G $	Metric	$T = 100$		$T = 200$		$T = 400$	
		$S_{G,(1)}$	$S_{G,(2)}$	$S_{G,(1)}$	$S_{G,(2)}$	$S_{G,(1)}$	$S_{G,(2)}$
0	TPR	–	–	–	–	–	–
	FPR	0.001	0.019	0.000	0.026	0.002	0.021
	Exact	0.998	0.958	1.000	0.943	0.995	0.950
2	TPR	0.255	0.826	0.755	0.995	0.973	0.998
	FPR	0.000	0.001	0.000	0.002	0.000	0.000
	Exact	0.255	0.825	0.755	0.990	0.973	0.998
3	TPR	0.729	0.951	0.947	0.998	0.994	1.000
	FPR	0.000	0.000	0.000	0.000	0.000	0.000
	Exact	0.497	0.867	0.870	0.993	0.983	1.000
4	TPR	0.860	0.971	0.959	0.996	0.998	0.999
	FPR	0.000	0.000	0.000	0.000	0.000	0.000
	Exact	0.570	0.887	0.855	0.983	0.990	0.998
5	TPR	0.896	0.985	0.971	0.995	0.999	1.000
	FPR	–	–	–	–	–	–
	Exact	0.595	0.927	0.858	0.975	0.993	1.000

Taken together, the results support using the second-largest statistic throughout the proposed procedure. It performs similarly to the largest statistic for existence testing, while providing substantially higher screening recovery with low false positive rates.

E.2 Real data

E.2.1 Data construction and preprocessing

Table E.5 summarizes the construction of the nine indicators used in the empirical analysis. The last indicator is a rolling version of the illiquidity measure of Amihud (2002), based on the ratio of absolute return to dollar trading volume. Rolling indicators are constructed stock by stock using November–December 2022 as warm-up history. ADF tests do not reveal systematic unit-root behavior in the constructed series, so no differencing is applied; we also do not winsorize the data.

For each indicator, the mean and standard deviation are first computed from all finite observed cells in the 2023–2024 primary sample, before any imputation, and used for standardization. The panel is then standardized using these values. Ordinary missing cells are imputed using the same-day sector median, with the same-day market median as a fallback; observations after structural exits are left missing. Finally, each stock-indicator series is demeaned using its 2023–2024 time mean.

Table E.5: Construction of the equity indicators.

Indicator	Construction	Window
Daily return	CRSP daily total return, $r_t = \text{DlyRet}_t$	Daily
20-day cumulative log return	$\sum_{\ell=0}^{19} \log(1 + r_{t-\ell})$	20 days
Log dollar volume	$\log(1 + \text{DlyPrVol}_t)$	Daily
Log turnover	$\log\{\text{DlyPrVol}_t / (1000 \text{DlyCap}_t)\}$	Daily
20-day relative turnover	Current log turnover minus its mean over the previous 20 trading days	20 lagged days
20-day log realized volatility	Log sample standard deviation of daily returns over the latest 20 trading days	20 days
High-low range	$\log(\text{DlyHigh}_t / \text{DlyLow}_t)$	Daily
Log bid-ask spread	$\log[1 + 10^4 \{2(\text{DlyAsk}_t - \text{DlyBid}_t) / (\text{DlyAsk}_t + \text{DlyBid}_t)\}]$	Daily
20-day log Amihud illiquidity	$\log\left[1 + \frac{1}{20} \sum_{\ell=0}^{19} \frac{10^6 r_{t-\ell} }{\text{DlyPrVol}_{t-\ell}}\right]$	20 days

E.2.2 Additional real-data results

More evidence of using MMEFM structure. Under (IC1), the four components, grand mean, row main effect, column main effect, and the rest interaction component are orthogonal, so for each sector m and date t , $\|\mathbf{Y}_t^{(m)}\|_F^2 = p_m q_m (\hat{c}_t^{(m)})^2 + q_m \|\hat{\mathbf{a}}_t^{(m)}\|^2 + p_m \|\hat{\mathbf{b}}_t^{(m)}\|^2 + \|\check{\mathbf{Y}}_t^{(m)}\|_F^2$. After aggregation over sectors and dates, the additive components account for 43.23% of the total squared magnitude. Figure E.3 shows that the interaction component is the largest part in every sector, but the main effects remain substantial and vary across sectors. The grand-mean contribution is small throughout.

Since the fitted global and local components are not exactly orthogonal, we compare their squared

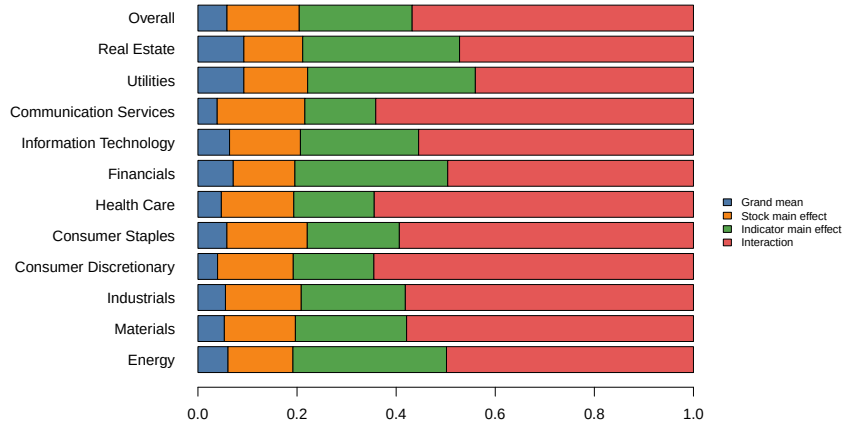


Figure E.3: Orthogonal decomposition of time-varying squared Frobenius norms across sectors.

magnitudes ratios. After aggregation over sectors and dates, the global magnitude ratios are

$$\text{Indicator main effect: } \frac{\sum_{m=1}^M \sum_{t=1}^T \|\mathbf{1}_{p_m} (\tilde{\mathbf{B}}_{1m} \hat{\boldsymbol{\beta}}_t)'\|_F^2}{\sum_{m=1}^M \sum_{t=1}^T \left\{ \|\mathbf{1}_{p_m} (\tilde{\mathbf{B}}_{1m} \hat{\boldsymbol{\beta}}_t)'\|_F^2 + \|\mathbf{1}_{p_m} (\tilde{\mathbf{B}}_{2m} \hat{\boldsymbol{\beta}}_t^{(m)})'\|_F^2 \right\}} = 42.81\%,$$

$$\text{Common component: } \frac{\sum_{m=1}^M \sum_{t=1}^T \|\tilde{\mathbf{Q}}_m \hat{\mathbf{G}}_t \tilde{\mathbf{J}}_m'\|_F^2}{\sum_{m=1}^M \sum_{t=1}^T \left\{ \|\tilde{\mathbf{Q}}_m \hat{\mathbf{G}}_t \tilde{\mathbf{J}}_m'\|_F^2 + \|\hat{\mathbf{R}}_m \hat{\mathbf{F}}_t^{(m)} \hat{\mathbf{C}}_m'\|_F^2 \right\}} = 26.63\%.$$

The sector-level ratios are computed in the same way by restricting the sums to each sector. Figure E.4 reports the corresponding sector-level ratios together with the screening statistics.

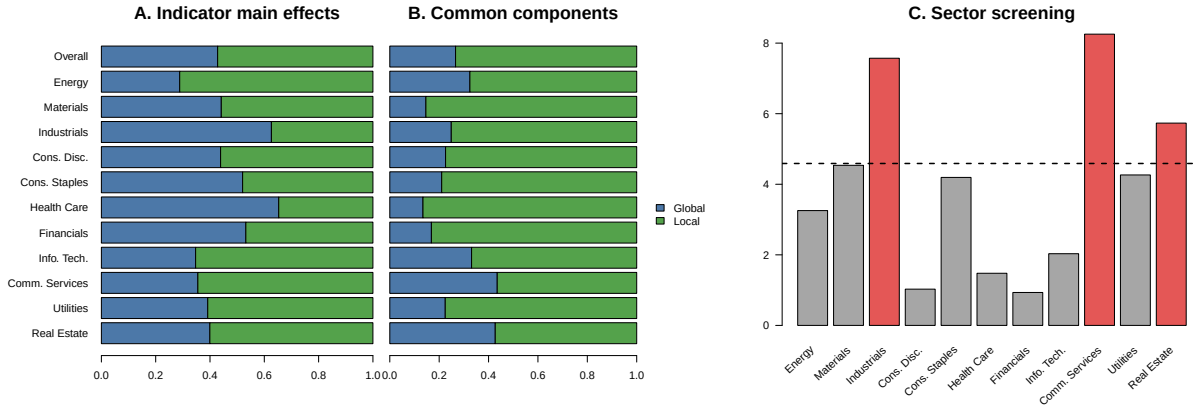


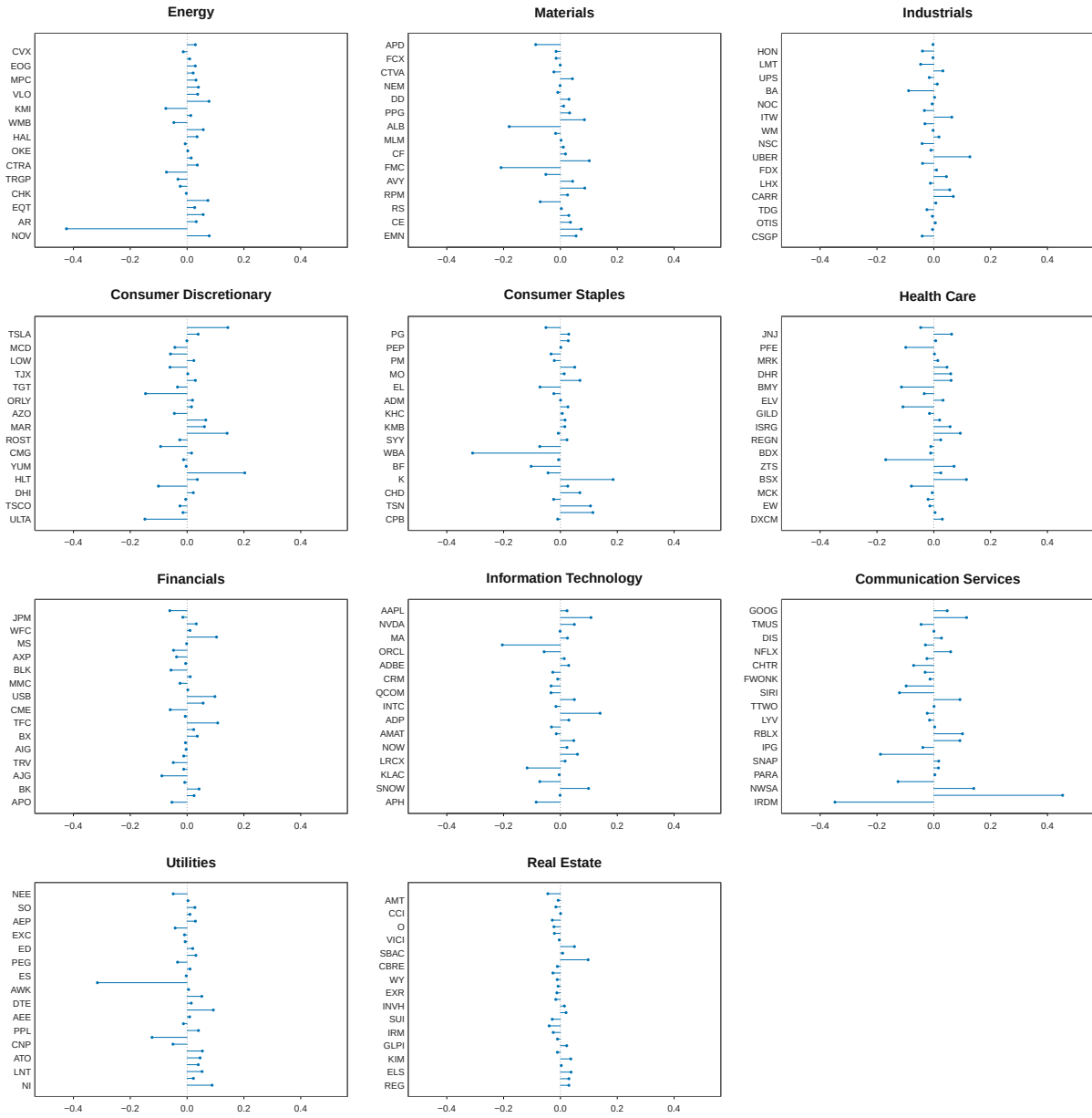
Figure E.4: Evidence for a multilevel structure. Panels A and B report global–local squared Frobenius magnitude ratios for the indicator main effects and common components, with Overall formed after aggregation across sectors. Panel C reports $S_{G,m}$; the dashed line is the 95% cutoff calibrated from $B = 199$ circular-shift bootstrap replications, and red bars denote screened sectors.

More results. Table E.6 reports the selected factor dimensions. The main text reports the global indicator main effect and illustrative local estimates for Financials and Information Technology. For completeness, Figure E.5 reports the global stock main effect, and Figures E.6–E.14 report local MMEFM estimates for the remaining nine sectors. In each sector-specific figure, Panels A–B show local stock-effect loadings and factor series, Panels C–D show local indicator-effect loadings and factor series, and Panels E–G show local common-component stock loadings, factor series, and indicator loadings.

Table E.6: Selected factor dimensions for the MMEFM: (r_1, ℓ_1) and $(r_2^{(m)}, \ell_2^{(m)})$ for the main effects, and (k_r, k_c) and $(k_r^{(m)}, k_c^{(m)})$ for the common component.

	Main effects	Common component
Global (all sectors)	(1, 1)	(1, 1)
<i>Local ranks by sector</i>		
Energy	(1, 2)	(2, 1)
Materials	(1, 2)	(3, 1)
Industrials	(1, 1)	(1, 2)
Consumer Discretionary	(1, 2)	(1, 1)
Consumer Staples	(2, 2)	(2, 1)
Health Care	(1, 1)	(2, 1)
Financials	(1, 1)	(1, 2)
Information Technology	(1, 2)	(1, 1)
Communication Services	(1, 2)	(2, 1)
Utilities	(3, 2)	(2, 1)
Real Estate	(1, 2)	(1, 1)

A. Global stock effect loadings ($\tilde{A}_{1m}$)



B. Global stock effect factor series ($\hat{\delta}_t$)

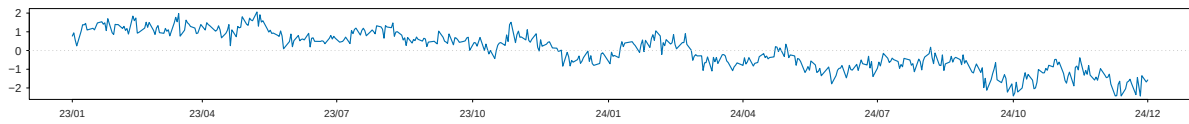


Figure E.5: Estimated global stock main effects.

Energy

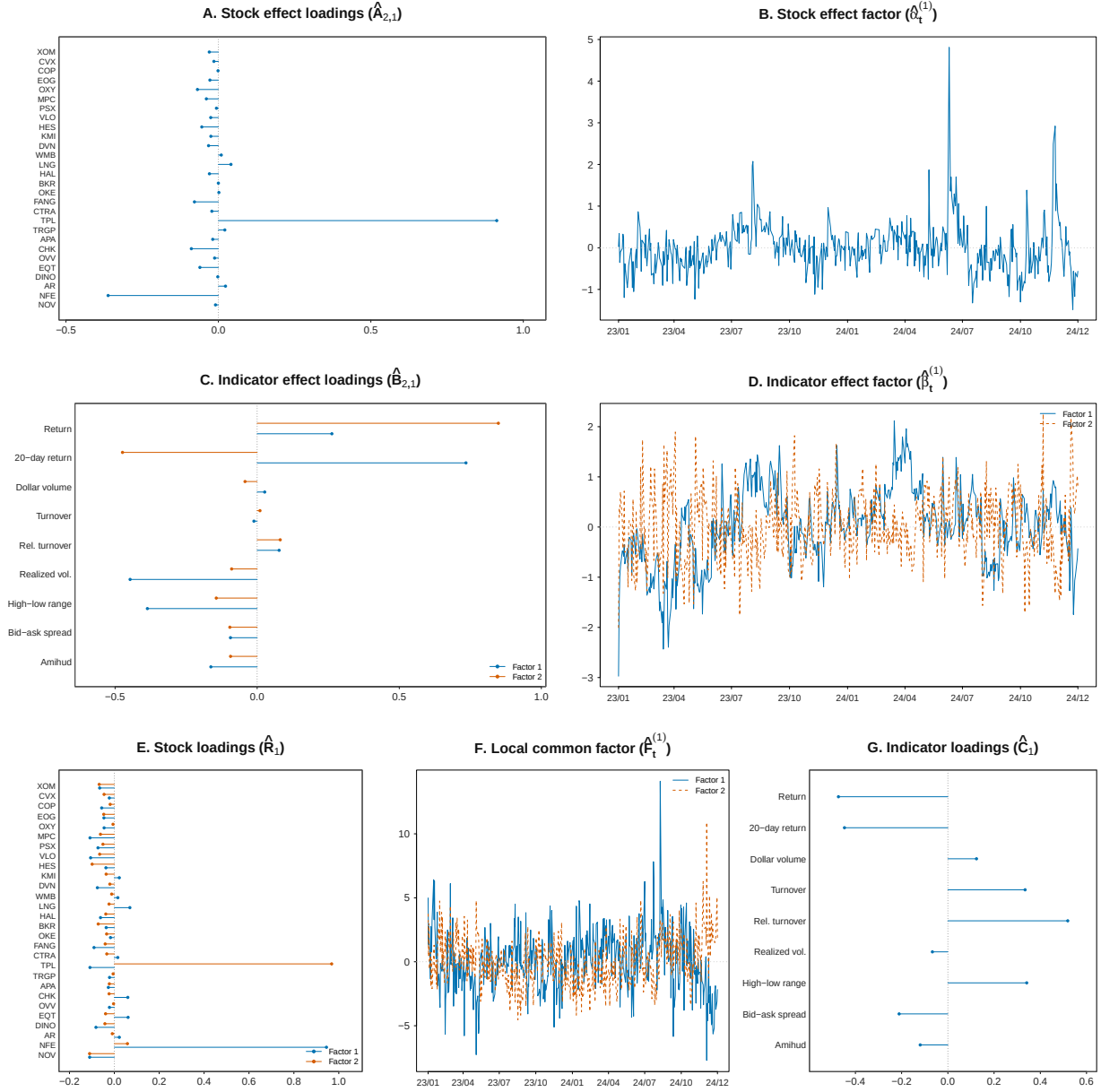


Figure E.6: Estimated local MMEFM components for Energy.

Materials

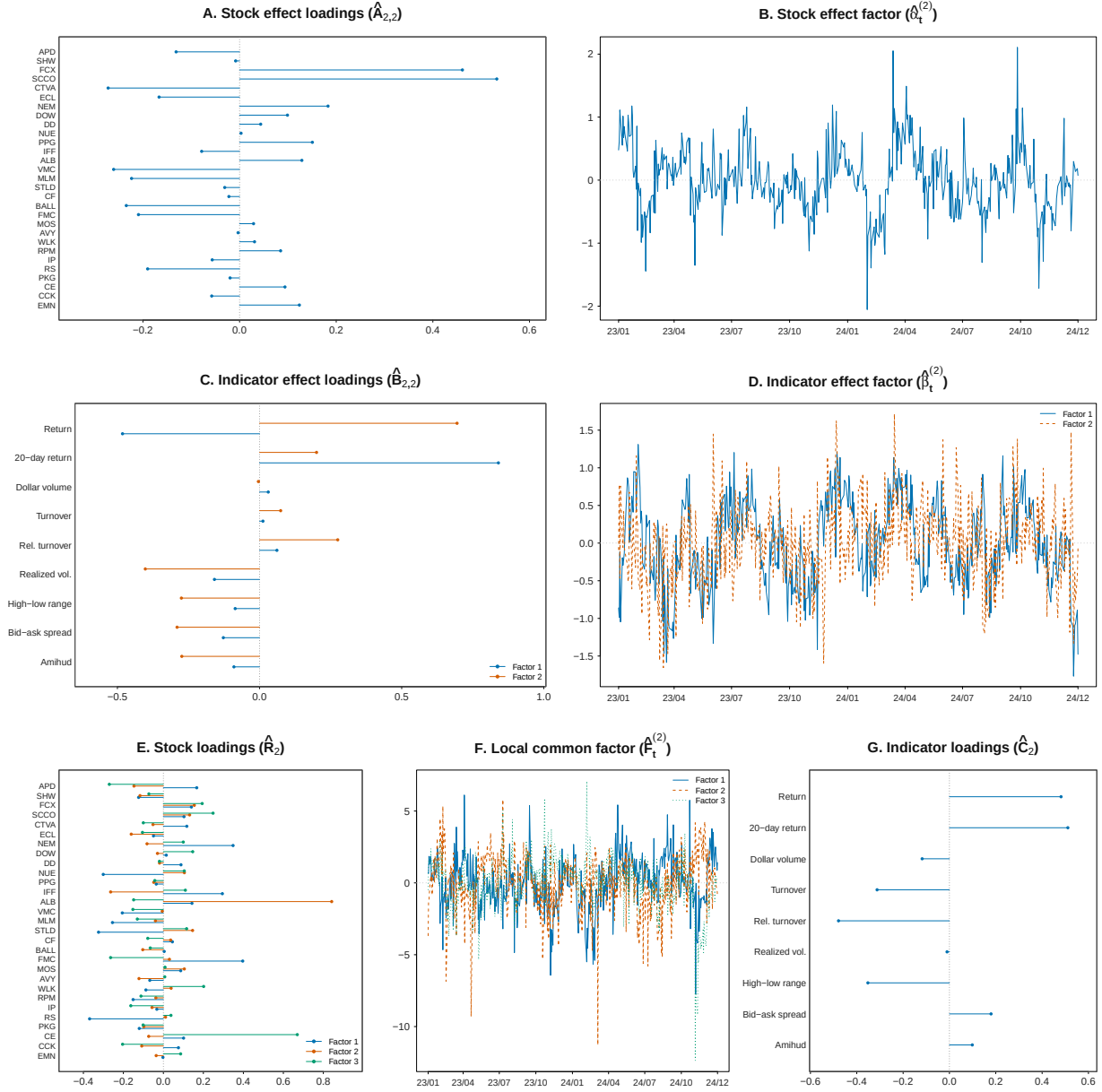


Figure E.7: Estimated local MMEFM components for Materials.

Industrials

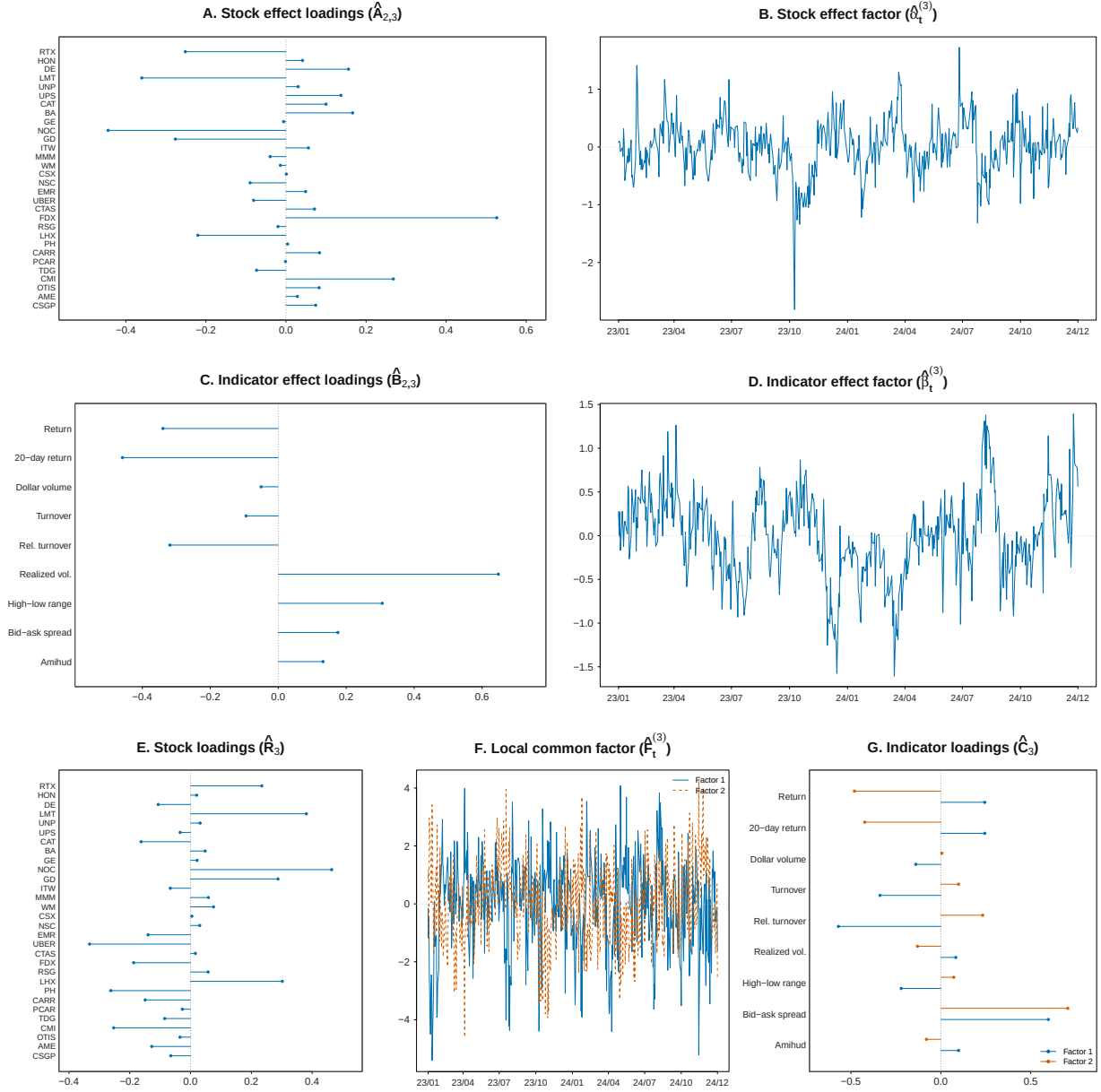


Figure E.8: Estimated local MMEFM components for Industrials.

Consumer Discretionary

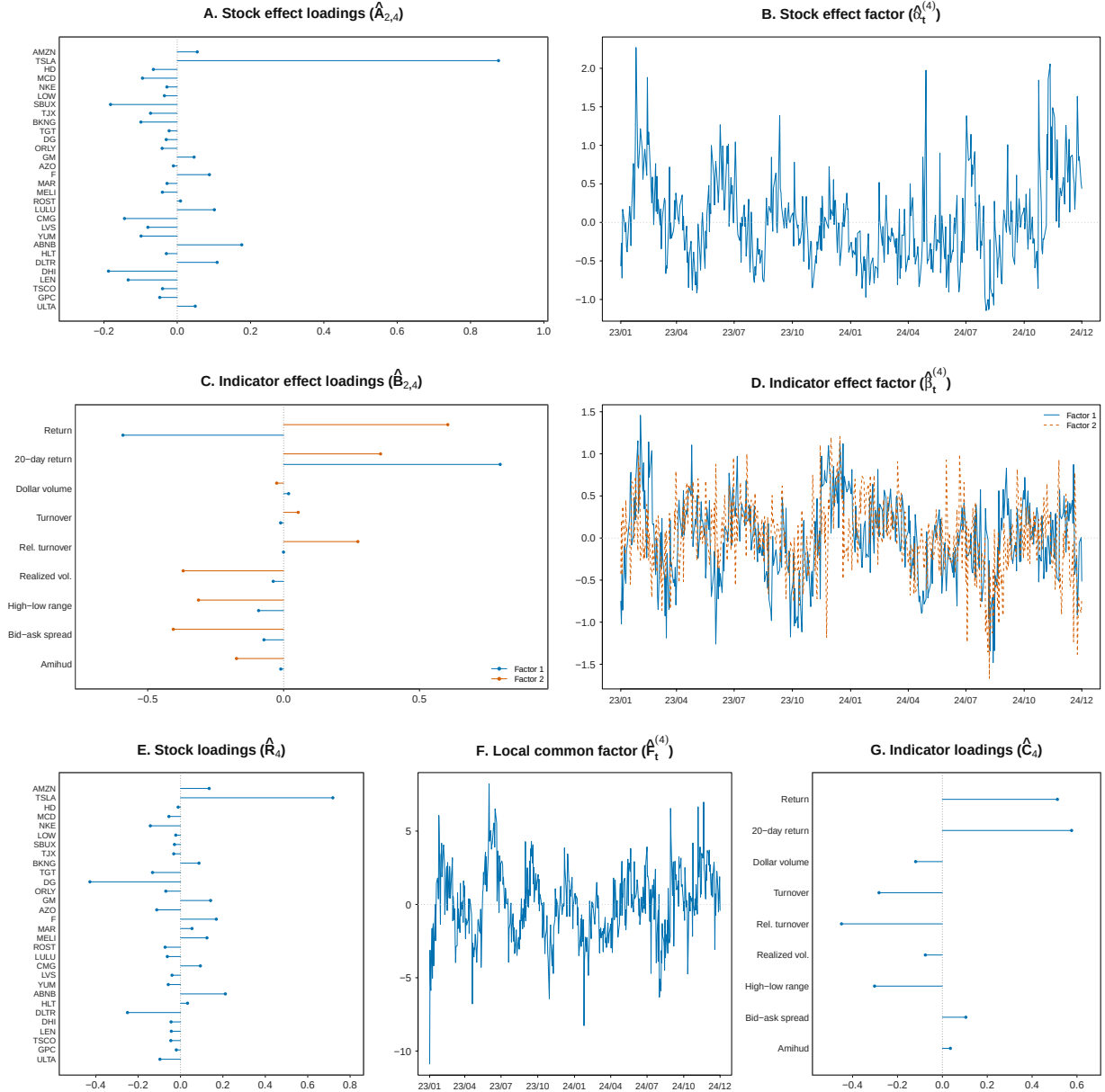


Figure E.9: Estimated local MMEFM components for Consumer Discretionary.

Consumer Staples

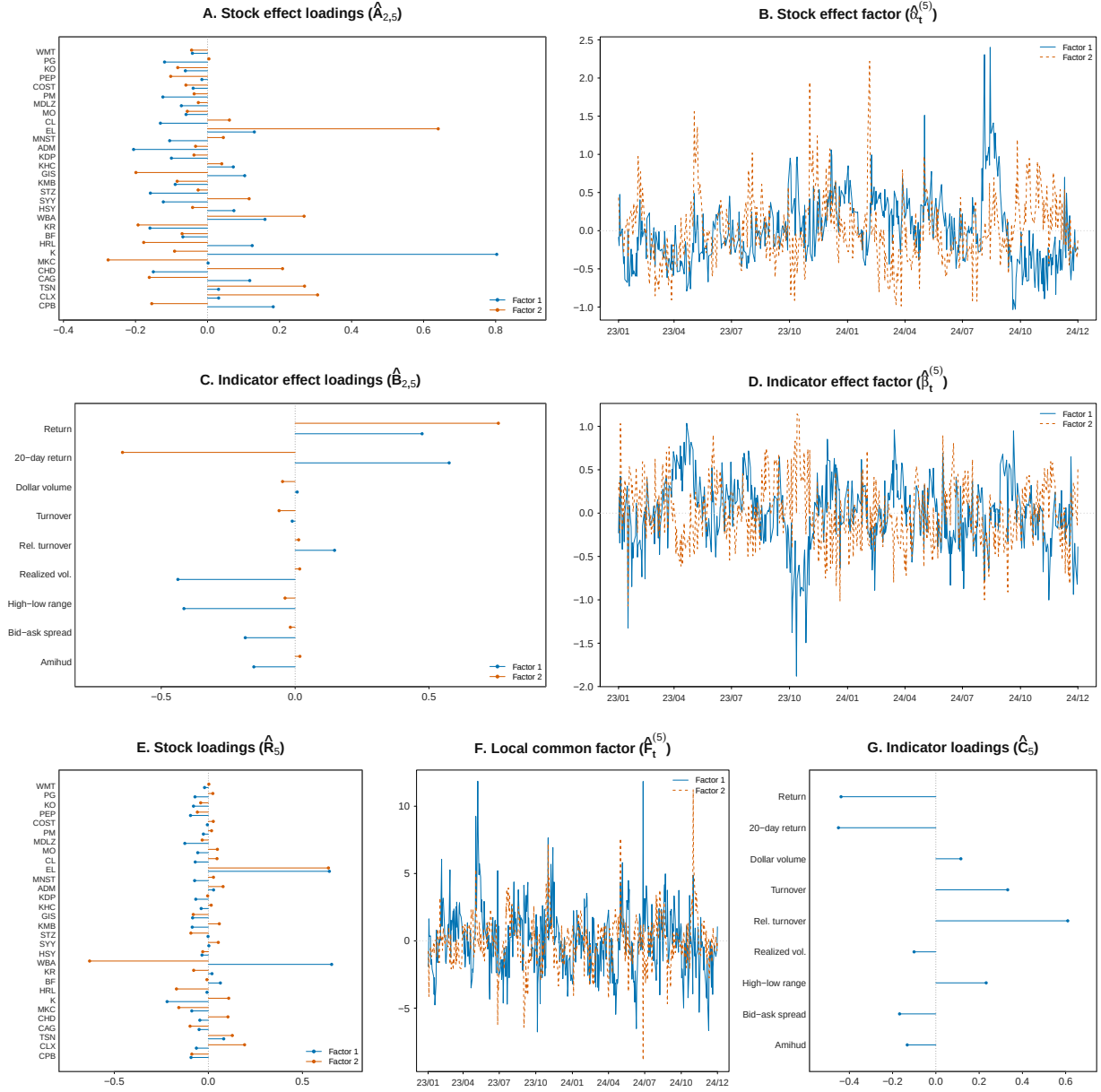


Figure E.10: Estimated local MMEFM components for Consumer Staples.

Health Care

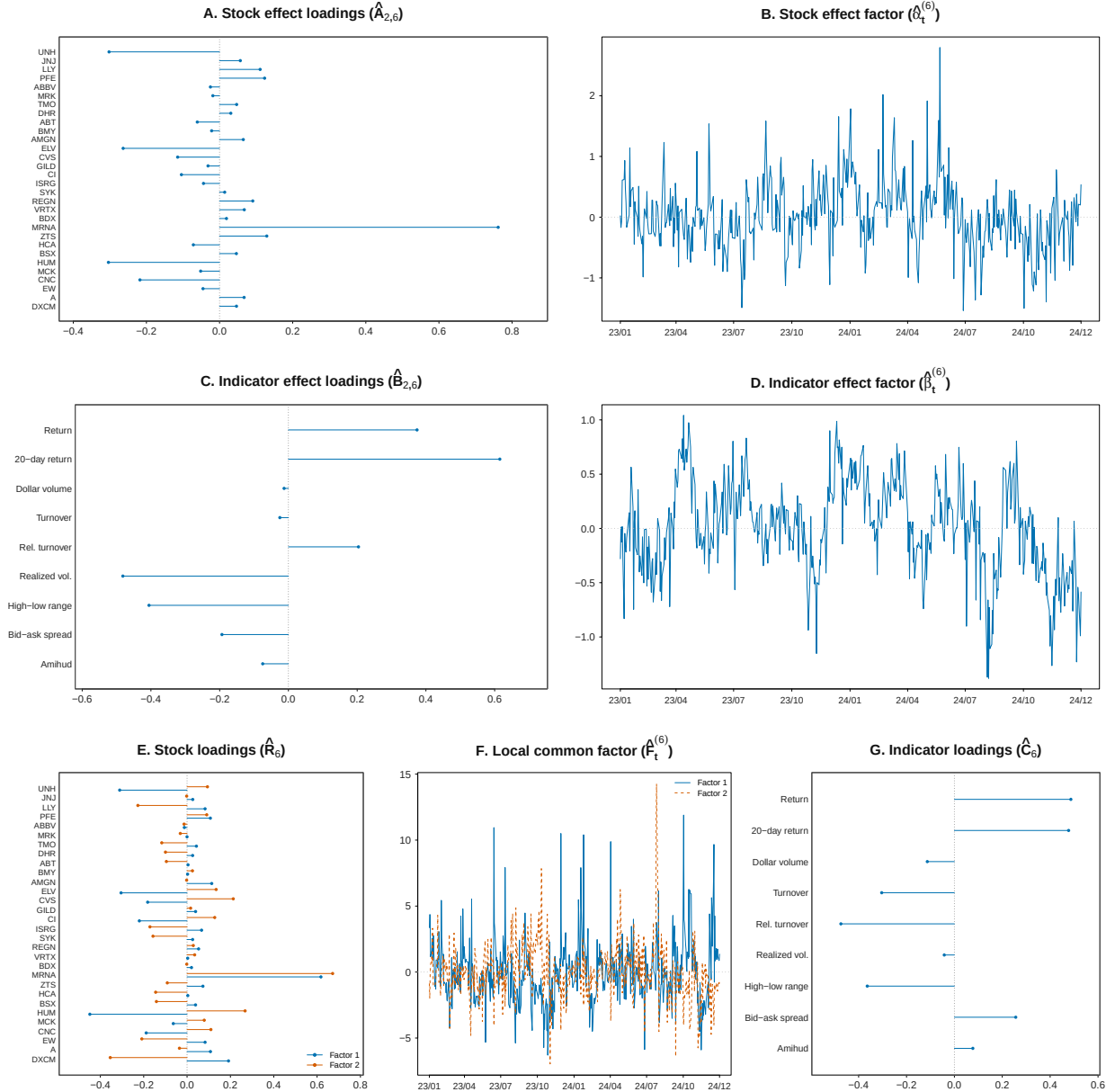


Figure E.11: Estimated local MMEFM components for Health Care.

Communication Services

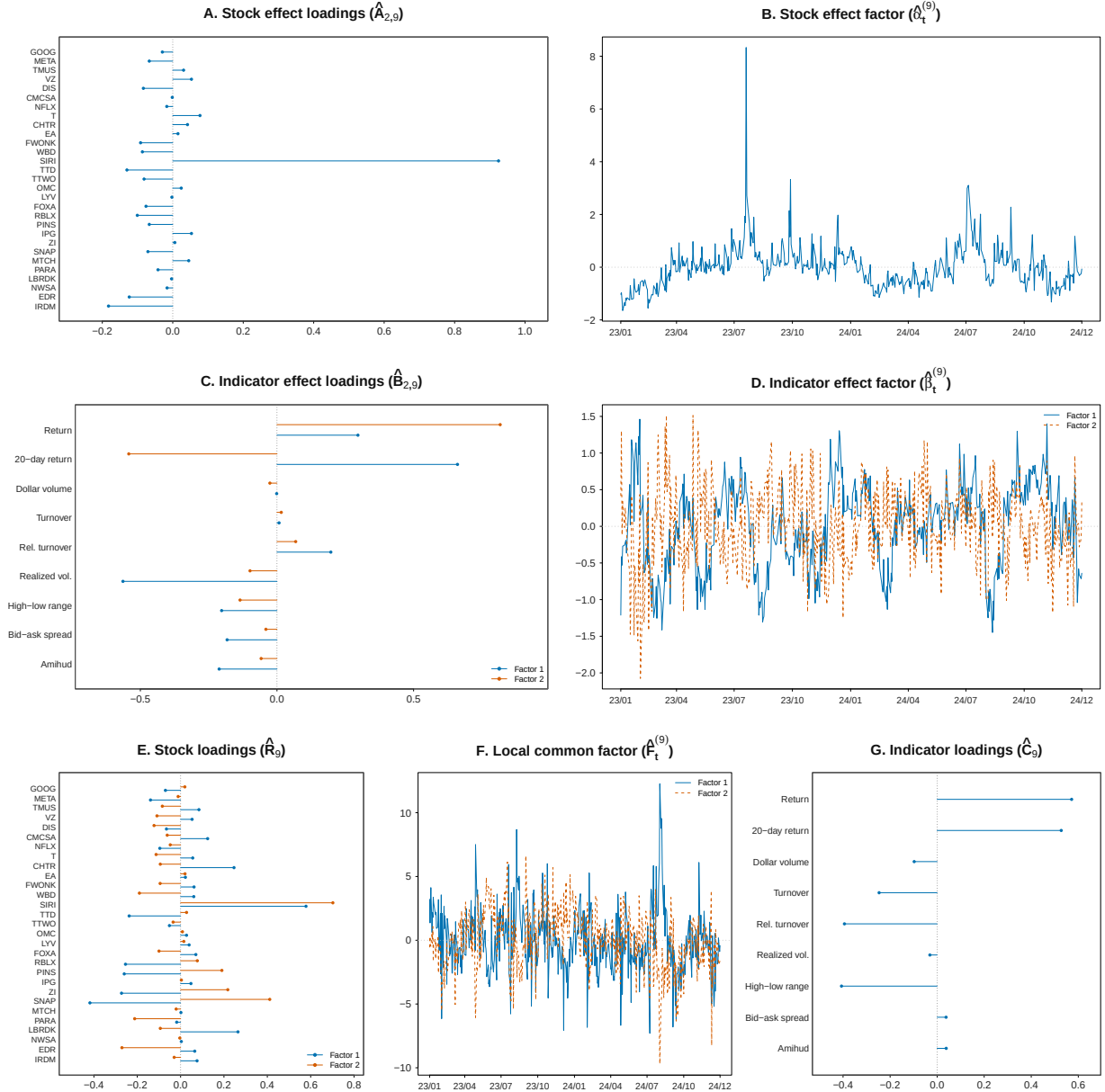


Figure E.12: Estimated local MMEFM components for Communication Services.

Utilities

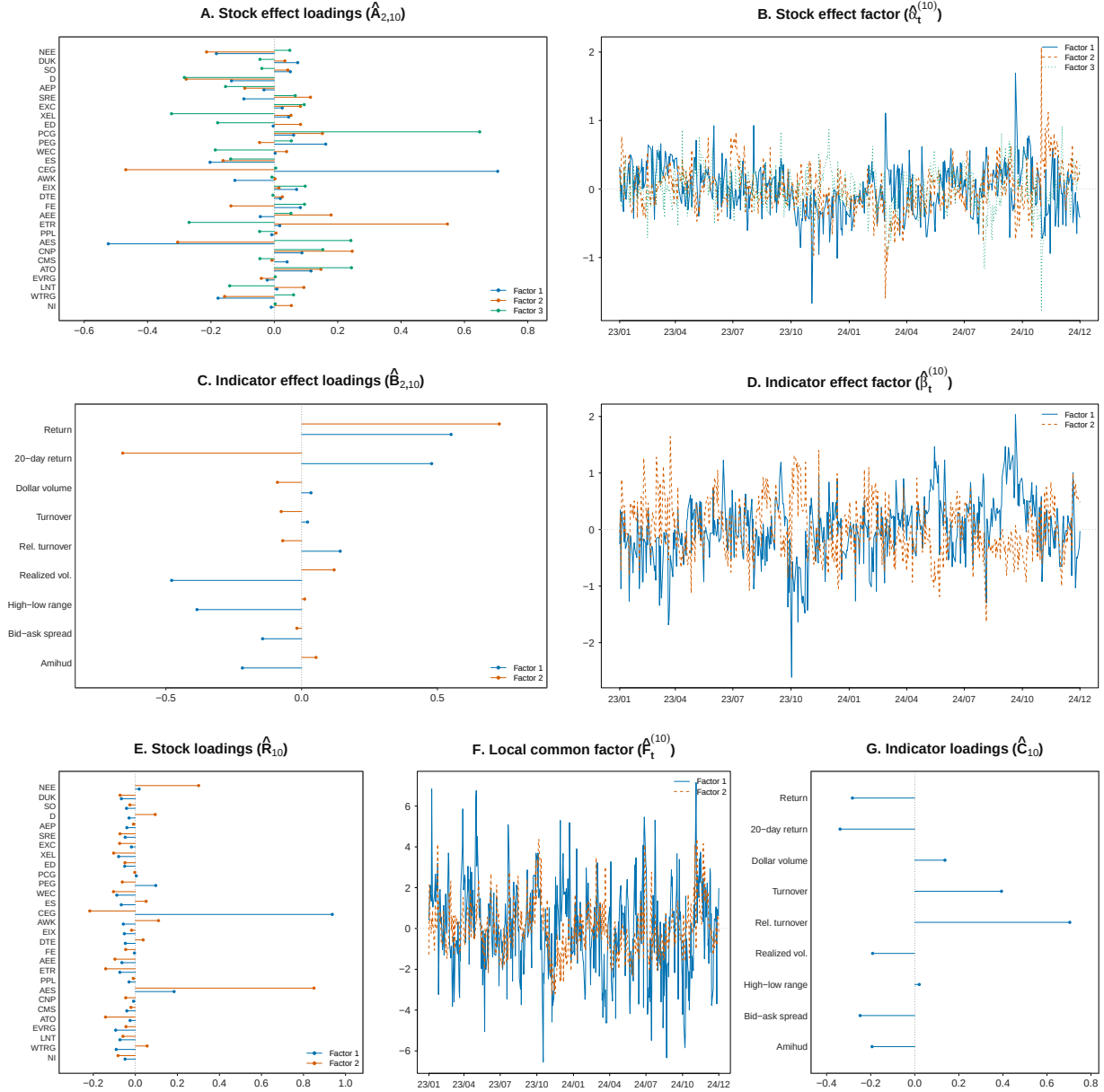


Figure E.13: Estimated local MMEFM components for Utilities.

Real Estate

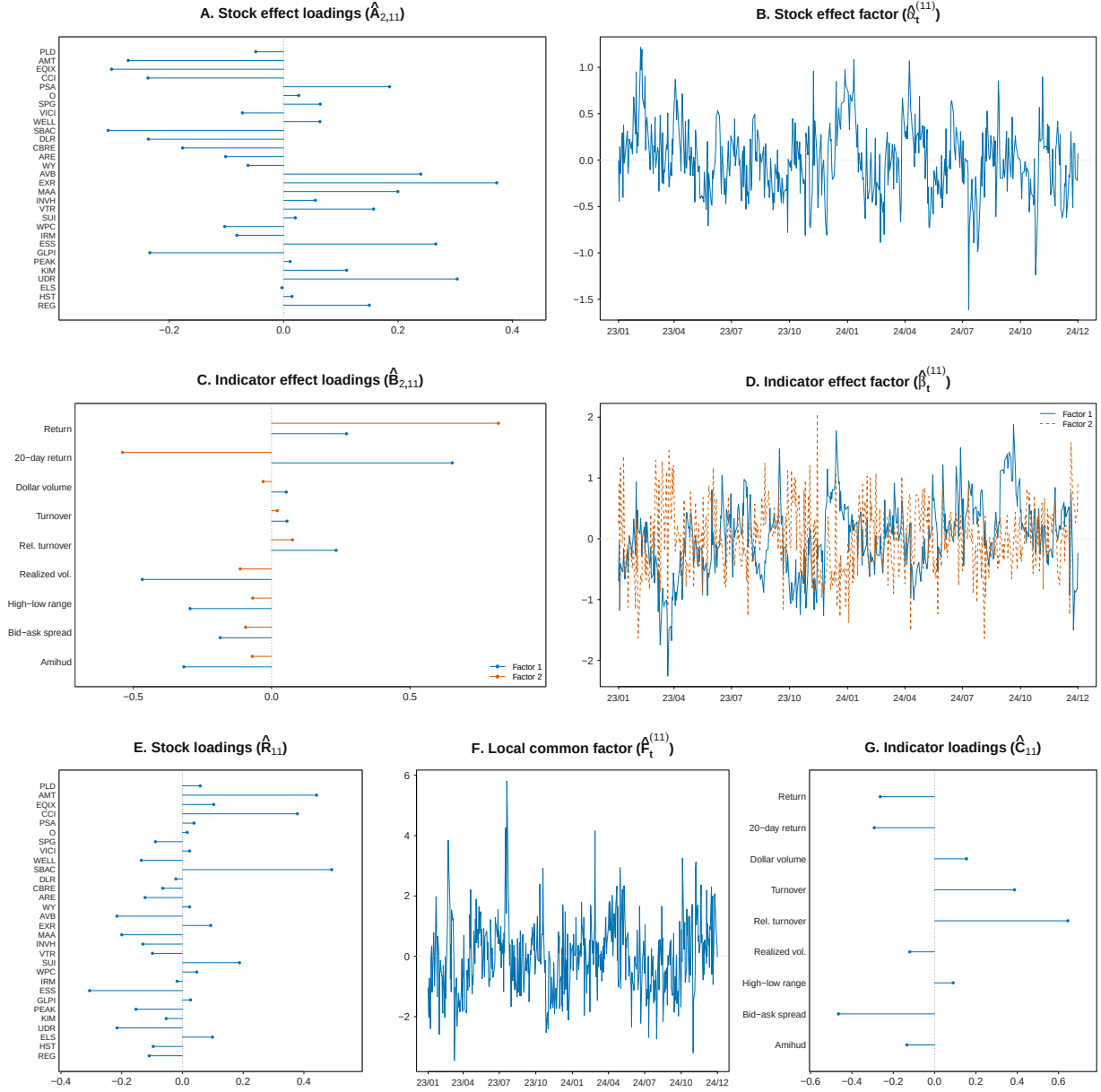


Figure E.14: Estimated local MMEFM components for Real Estate.