Bootstrap inference for group factor models*

Sílvia Gonçalves[†], Julia Koh and Benoit Perron

McGill University, Tilburg University and Université de Montréal

August 23, 2024

Abstract

Andreou et al. (2019) have proposed a test for common factors based on canonical correlations between factors estimated separately from each group. We propose a simple bootstrap test that avoids the need to estimate the bias and variance of the canonical correlations explicitly and provide high-level conditions for its validity. We verify these conditions for a wild bootstrap scheme similar to the one proposed in Gonçalves and Perron (2014). Simulation experiments show that this bootstrap approach leads to null rejection rates closer to the nominal level in all of our designs compared to the asymptotic framework.

1 Introduction

Factor models have been extensively used in the past decades to reduce dimensions of large data sets. They are now widely used in forecasting, as controls in regressions, and as a tool to model cross-sectional dependence.

Andreou et al. (2019) have proposed a test of whether two groups of data contain common factors. The test consists in estimating a set of factors for each subgroup using principal components and testing whether some canonical correlations between these two groups of estimated factors are 1 as they would be if there are factors common to both groups of data. Inference in this situation is complicated by the need to account for the preliminary estimation of the factors. The asymptotic

*This article is based on the Halbert White Memorial Lecture given by Sílvia Gonçalves at the Society for Financial Econometrics Conference on June 15th, 2024 in Rio de Janeiro, Brazil. We thank the JFEC Editors Allan Timmermann and Fabio Trojani for the invitation, and the discussants at the conference, Michael Wolf and Eric Ghysels, for their insightful comments. We also thank participants at the following conferences: Society for Financial Econometrics (2023), North American Summer Meetings of the Econometric Society (2023), Advances in Econometrics conference in honor of Joon Park (2023), International Association for Applied Econometrics (2023), Canadian Economics Association (2023), Société canadienne de science économique (2024), CIREQ Econometrics conference in honor of Eric Ghysels (2024), and International Symposium on Non-Parametric Statistics (2024). Gonçalves and Perron thank the Social Sciences and Humanities Research Council (SSHRC, grants 435-2023-0352 and 435-2020-1349 respectively) and the Fonds de recherche du Québec – société et culture (FRQSC, grant 2020-SE2-269954).

†Corresponding author: silvia.goncalves@mcgill.ca.

theory in Andreou et al. (2019) is highly nonstandard with non-standard rates of convergence and the presence of an asymptotic bias. Under restrictive assumptions, they propose an estimator for this bias and construct a feasible statistic. However, their simulation results suggest that, even under these restrictive assumptions, their statistic can exhibit level distortions.

This approach was applied in Andreou et al. (2022) to sets of returns on individual stocks and on portfolios. In principle, these two sets of returns should share a common set of factors that represent the stochastic discount factor. The authors find a set of 3 common factors that price both individual stocks and sorted portfolios. They also find that 10 principal components from the large number of factors proposed in the literature to price stocks (the factor zoo) are needed to span the space of these three common factors.

This paper proposes the bootstrap as an alternative inference method. Our main contribution is to propose a simple bootstrap test that avoids the need to estimate the bias and variance of the canonical correlations explicitly. We show its validity under a set of high-level conditions that allow for weak dependence on the data generating process. The specific bootstrap scheme that is used depends on the assumptions a researcher is willing to make about this dependence.

For example, if a researcher is willing to assume that the idiosyncratic terms do not exhibit cross-sectional or serial correlation, we show that a wild bootstrap is valid in this context. This is analogous to the results in Gonçalves and Perron (2014) who showed the validity of a wild bootstrap in the context of factor-augmented regression models. If the presence of cross-sectional dependence is important, a researcher could instead use the cross-sectional dependent bootstrap of Gonçalves and Perron (2020). If serial correlation in the idiosyncratic errors is relevant, Koh (2022) proposed an autoregressive sieve bootstrap for factor models. Finally, we also discuss an extension of this method that allows for cross-sectional and serial dependence in the idiosyncratic errors.

The bootstrap has recently been applied in Andreou et al. (2024) to test for the number of common factors. Contrary to our framework which follows Andreou et al. (2019), one set of the factors is assumed to be observed, implying that their bootstrap method is different from ours.

The remainder of the paper is organized as follows. Section 2 describes the model and the testing problem in Andreou et al. (2019). Section 3 introduces a general bootstrap scheme in this context and provides a set of high level conditions under which the bootstrap test is asymptotically valid under the null hypothesis. We also provide a set of sufficient conditions that ensure the bootstrap test is consistent under the alternative hypothesis. Section 4 verifies these conditions for the wild bootstrap method of Gonçalves and Perron (2014) under a set of assumptions similar to those in Andreou et al. (2019). Section 5 provides simulation results and Section 6 concludes. We provide three appendices. Appendix A contains a set of assumptions under which we derive the limiting distribution of the original test statistic as well as auxiliary lemmas used to derive this asymptotic distribution. Appendix B contains a set of bootstrap high level conditions that mirror the primitive assumptions in Appendix A. It also provides the bootstrap analog of the auxiliary lemmas introduced in Appendix A, which are used to prove the bootstrap results in Section 3. Finally, Appendix C contains the proofs of the bootstrap results for the wild bootstrap method proposed in Section 4.

A final word on notation. For a bootstrap sequence, say $X^*_{N,T}$, we use $X^*_{N,T} \overset{p^*}{\to}_p 0$, or, equivalently, or $X^*_{N,T} \overset{p^*}{\to} 0$, as $N, T \to \infty$, in probability, to mean that, for any $\epsilon > 0$, $P^*(|X^*_{N,T}| > \epsilon) \to_p 0$, where P^* denotes the probability measure conditionally on the original data. An equivalent notation is $X^*_{N,T} = o_{p^*}(1)$ (where we omit the qualification "in probability" for brevity). We also write $X^*_{N,T} = O_{p^*}(1)$ to mean that $P^*(|X^*_{N,T}| > M) \to_p 0$ for some large enough M. Finally, we write $X^*_{N,T} \overset{d^*}{\to}_p X$ or, equivalently, $X^*_{N,T} \overset{d^*}{\to} X$, in probability, to mean that, for all continuity points $x \in \mathbb{R}$ of the cdf of X, say $F(x) \equiv P(X \leq x)$, we have that $P^*(X^*_{N,T} \leq x) - F(x) \to_p 0$.

2 Framework

2.1 The group panel factor model

Following Andreou et al. (2019) (henceforth AGGR(2019)), we consider a group factor panel model with two groups, both driven by a set of common factors and a set of specific factors:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \Lambda_1^c & \Lambda_1^s & 0 \\ \Lambda_2^c & 0 & \Lambda_2^s \end{bmatrix} \begin{bmatrix} f_t^c \\ f_{1t}^s \\ f_{2t}^s \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Here, y_{1t} and y_{2t} are $N_1 \times 1$ and $N_2 \times 1$ vectors, respectively. In particular, $y_{jt} = \left[y_{j,1t}, \ldots, y_{j,N_j t}\right]'$ collects the N_j observations in group j at time t. We use a similar notation to denote ε_{jt}. The $k^c \times 1$ vector f_t^c denotes the common factors whereas f_{jt}^s is a $k_j^s \times 1$ vector which contains the factors specific to group j. The matrices Λ_j^c and Λ_j^s contain the factor loadings associated with the common factors and the group specific factors for group j, respectively. Thus, Λ_j^c is $N_j \times k^c$ and Λ_j^s is $N_j \times k_j^s$. We let $k_j \equiv k^c + k_j^s$ denote the total number of factors in each sub-panel and define $\underline{k} \equiv \min(k_1, k_2)$. Finally, we let $f_{jt} = \left((f_t^c)', \left(f_{jt}^s\right)'\right)'$ define the $k_j \times 1$ vector that collects all factors in each group. Their variance-covariance matrices are $V_{jl} \equiv E\left(f_{jt} f_{lt}'\right)$, $j, l = 1, 2$. As in AGGR(2019), we assume that the common and group-specific factors have mean zero, variance-covariance matrix equal to the identity matrix, and that they are orthogonal within each group:

$$E\left(f_{jt}\right) = 0 \text{ and } V_{jj} \equiv Var\left(f_{jt}\right) = Var\begin{pmatrix} f_t^c \\ f_{j,t}^s \end{pmatrix} = \begin{bmatrix} I_{k^c} & 0 \\ 0 & I_{k_j^s} \end{bmatrix} = I_{k_j}.$$

However, we allow for the possibility that f_{1t}^s and f_{2t}^s are correlated with covariance matrix Φ.

2.2 The testing problem

AGGR(2019) propose an inference procedure for determining the number of common factors k^c. Their procedure is based on the fact that the canonical correlations between the two sets of factors f_{1t} and

f_{2t} identify the common factor space. Specifically, let $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_{\underline{k}}$ denote the ordered canonical correlations between f_{1t} and f_{2t}. The squared canonical correlations ρ_l^2 for $l = 1, \ldots, \underline{k}$ are defined as the $\underline{k}$ largest eigenvalues of the matrix $R = V_{11}^{-1} V_{12} V_{22}^{-1} V_{21}$ (or equivalently of $\acute{R} = V_{22}^{-1} V_{21} V_{11}^{-1} V_{12}$). Proposition 1 of AGGR(2019) shows that if $k^c > 0$, the largest k^c canonical correlations are equal to 1 and the remaining $\underline{k} - k^c$ are strictly less than 1. This corresponds to the null hypothesis that there are k^c common factors, i.e.,

$$\mathsf{H}_0 : \rho_1 = \ldots = \rho_{k^c} = 1 > \rho_{k^c+1} \geq \ldots \geq \rho_{\underline{k}}.$$

To test H_0, AGGR(2019) propose a test statistic based on

$$\hat{\xi}\left(k^c\right) = \sum_{l=1}^{k^c} \hat{\rho}_l,$$

where $\hat{\rho}_l$ is the sample analogue of ρ_l (we define these estimators below). Under the null, $\hat{\xi}\left(k^c\right)$ should be close to k^c, whereas it should be less than k^c under the alternative hypothesis H_1. Here, H_1 is defined as

$$\mathsf{H}_1 : \exists\ 0 \leq r < k^c \text{ s.t. } \rho_1 = \ldots = \rho_r = 1 > \rho_{r+1} \geq \ldots \geq \rho_{\underline{k}},$$

with the understanding that if $r = 0$, $\rho_l < 1$ for all $l = 1, \ldots, \underline{k}$, corresponding to the absence of common factors. Thus, we reject the null when $\hat{\xi}\left(k^c\right) - k^c$ is negative and large.

The critical value used in AGGR(2019) is obtained from the asymptotic distribution of the test statistic when N_1, N_2 and $T \to \infty$. Our goal in this paper is to propose an alternative method of inference based on the bootstrap.

2.3 Canonical correlations, common and group-specific factors and their loadings

Here, we define the estimators $\hat{\rho}_l$, $l = 1, \ldots, \underline{k}$. In the process of doing so, we also define the estimators of the common and group-specific factors and factor loadings. These will be used to form our bootstrap data generating process.

We start by extracting the first k_j principal components for each group j, with $j = 1, 2$. In particular, let Y_j denote the observed data matrix of size $T \times N_j$ for group j. The factor model for this group can be written as

$$Y_j = F_j \Lambda_j' + \varepsilon_j, \tag{1}$$

where ε_j is $T \times N_j$, $F_j = (f_{j1}, \ldots, f_{jT})'$ is $T \times k_j$, and Λ_j is $N_j \times k_j$.

Given Y_j, we estimate F_j and Λ_j with the standard method of principal components. In particular, F_j is estimated with the $T \times k_j$ matrix $\hat{F}_j = \begin{pmatrix} \hat{f}_{j1}, & \ldots & , \hat{f}_{jT} \end{pmatrix}'$ composed of $\sqrt{T}$ times the eigenvectors corresponding to the k_j largest eigenvalues of of $Y_j Y_j' / T N_j$ (arranged in decreasing order), where the normalization $\frac{\hat{F}_j' \hat{F}_j}{T} = I_{k_j}$ is used. The factor loading matrix is $\hat{\Lambda}_j = Y_j' \hat{F}_j / T$.

The estimators $\hat{\rho}_l$, $l = 1, \ldots, \underline{k}$ are obtained from the eigenvalues of the sample analogue of R.

Specifically, letting

$$\hat{V}_{jl} = \frac{1}{T}\hat{F}_j'\hat{F}_l = \frac{1}{T}\sum_{t=1}^{T}\hat{f}_{jt}\hat{f}_{lt}', \;\; j,l = 1,2,$$

we can define[1]

$$\hat{R} = \hat{V}_{11}^{-1}\hat{V}_{12}\hat{V}_{22}^{-1}\hat{V}_{21}.$$

The k^c largest eigenvalues of $\hat{R}$ are denoted by $\hat{\rho}_l^2$, $l = 1, \ldots, k^c$. They correspond to the largest k^c sample squared canonical correlations between $\hat{f}_{1t}$ and $\hat{f}_{2t}$.

For our bootstrap data generating process (to be defined in the next section), we also need estimates of the common and group-specific factors and loadings. These estimates are also used to obtain the test statistic proposed by AGGR(2019). Hence, we describe them next.

First, using Definition 1 of AGGR(2019), we can estimate the common factors as follows. Let the k^c associated eigenvectors of $\hat{R}$ (the canonical directions) be collected in the $k_1 \times k^c$ matrix $\hat{W}$, normalized to have length one, e.g. $\hat{W}'\hat{V}_{11}\hat{W} = \hat{W}'\hat{W} = I_{k_1}$ since $\hat{V}_{11} = I_{k_1}$. Given $\hat{W}$, an estimator of the common factors f_t^c is $\hat{f}_t^c = \hat{W}'\hat{f}_{1t}$.

The group-specific factors f_{jt}^s $(j = 1,2)$ can then be estimated using Definition 2 of AGGR(2019). In particular, $\hat{f}_{jt}^s$ are obtained by applying the method of principal components to the $T \times N_j$ matrix of residuals Ξ_j obtained from regressing y_{jt} on the estimated common components $\hat{f}_t^c$. More specifically, given model (1), we can further decompose F_j and Λ_j in terms of common and group-specific factors and factor loadings to write

$$Y_j = F^c\Lambda_j^{c\prime} + F_j^s\Lambda_j^{s\prime} + \varepsilon_j.$$

Let $\hat{F}^c = \begin{pmatrix} \hat{f}_1^c, & \ldots, & \hat{f}_T^c \end{pmatrix}'$ denote the $T \times k^c$ matrix of the k^c largest estimated common factors. The regression of Y_j on $\hat{F}^c$ yields the common factor loadings

$$\hat{\Lambda}_j^c = Y_j'\hat{F}^c\left(\hat{F}^{c\prime}\hat{F}^c\right)^{-1} = \frac{1}{T}Y_j'\hat{F}^c.$$

The matrix Ξ_j is defined as $\Xi_j = Y_j - \hat{F}^c\hat{\Lambda}_j^{c\prime}$. Given Ξ_j, we can now apply the method of principal components to obtain $\hat{F}_j^s = \begin{pmatrix} \hat{f}_{j1}^s & \ldots & \hat{f}_{jT}^s \end{pmatrix}'$, composed of $\sqrt{T}$ times the eigenvectors corresponding to the k_j^s largest eigenvalues of of $\Xi_j\Xi_j'/TN_j$ (arranged in decreasing order), where the normalization $\frac{\hat{F}_j^{s\prime}\hat{F}_j^s}{T} = I_{k_j^s}$ is used.

2.4 The test statistic and its asymptotic distribution

To test H_0, we rely on the statistic

$$\hat{\xi}\left(k^c\right) - k^c = \sum_{l=1}^{k^c}\hat{\rho}_l - k^c.$$

[1]For simplicity, we focus on $\hat{R}$ here. Our results also apply to a test statistic based on the alternative estimator $\hat{R}^*$ defined in AGGR(2019) (this is the sample analogue of $\acute{R}$ in our notation).

Our goal is to propose a bootstrap test based on the bootstrap analogue of $\hat{\xi}(k^c) - k^c$, say $\hat{\xi}^*(k^c) - k^c$. In particular, we focus on obtaining a valid bootstrap p-value $p^* \equiv P^*\left(\hat{\xi}^*(k^c) \leq \hat{\xi}(k^c)\right)$.[2]

To understand the properties that a bootstrap test should have in order to be asymptotically valid, we first review the large sample properties of this test statistic, as studied by AGGR(2019). In the following, we let $N \equiv \min(N_1, N_2) = N_2$ (without loss of generality) and define $\mu_N = \sqrt{N_2/N_1}$. Since $N = N_2 \leq N_1$, $\mu_N \leq 1$. We assume that $\mu_N \to \mu \in [0,1]$. When $N_1 = N_2 = N$, $\mu_N = \mu = 1$.

Appendix A contains a set of assumptions under which we derive the asymptotic distribution of $\hat{\xi}(k^c)$. Compared to AGGR(2019), we impose a stricter rate condition on N relatively to T. In particular, while our Assumption 1 maintains AGGR(2019)'s assumption that $\sqrt{T}/N \to 0$, we require that $N/T^{3/2} \to 0$ as opposed to $N/T^{5/2} \to 0$. The main reason why we adopt a stricter rate condition is that it greatly simplifies both the asymptotic and the bootstrap theory. [3] In addition, we generalize standard assumptions in the literature on factor models (see e.g., Bai (2003), Bai and Ng (2006) and Gonçalves and Perron (2014), henceforth GP(2014)) to the group factor context of interest here. Our assumptions suggest natural bootstrap high level conditions (which we provide in Appendix B) under which the bootstrap asymptotic distribution can be derived. Since some of these bootstrap conditions have already been verified in the previous literature, we can rely on existing results for proving our bootstrap theory. Instead, AGGR(2019)'s assumptions are not easily adapted to proving our bootstrap theory.

Next, we characterize the asymptotic distribution of $\hat{\xi}(k^c)$ under Assumptions 1-6 in Appendix A. We introduce the following notation. First, we let $u_{jt} \equiv \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}$ and note that u_{jt} captures the factors estimation uncertainty for panel j. In particular, as is well known (cf. Bai (2003)), estimation of f_{jt} by principal components implies that each estimator $\hat{f}_{jt}$ is consistent for $H_j f_{jt}$, a rotated version of f_{jt}. The rotation matrix is defined as $H_j = \mathcal{V}_j^{-1}\frac{\hat{F}_j'F_j}{T}\frac{\Lambda_j'\Lambda_j}{N_j}$, where $\mathcal{V}_j$ is a $k_j \times k_j$ diagonal matrix containing the k_j largest eigenvalues of Y_jY_j'/N_jT on the main diagonal, in decreasing order. As shown by Bai (2003), u_{jt} is the leading term in the asymptotic expansion of $\sqrt{N_j}\left(\hat{f}_{jt} - H_j f_{jt}\right)$. We let $u_{jt}^{(c)}$ denote the $k^c \times 1$ vector containing the first k^c rows of $u_{jt} \equiv \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}$ and define $\mathcal{U}_t \equiv \mu_N u_{1t}^{(c)} - u_{2t}^{(c)}$. Finally, we let $\tilde{\Sigma}_{\mathcal{U}} \equiv T^{-1}\sum_{t=1}^T E\left(\mathcal{U}_t\mathcal{U}_t'\right)$ and $\tilde{\Sigma}_{cc} \equiv T^{-1}\sum_{t=1}^T E\left(f_t^c f_t^{c\prime}\right)$.

Theorem 2.1 *Suppose Assumptions 1-6 hold and the null hypothesis is true so that* $f_{jt} = \left(f_t^{c\prime}, f_{jt}^{s\prime}\right)'$

[2] Although we denote the bootstrap p-value by p^*, we should note it is not random with respect to the bootstrap measure P^*. A similar notation is used below to denote the bootstrap bias $\mathcal{B}^*$ and bootstrap variance $\Omega_{\mathcal{U}}^*$ of the bootstrap test statistic $\hat{\xi}^*(k^c)$. This choice of notation allows us to differentiate bootstrap population quantities from other potential estimators that do not rely on the bootstrap.

[3] Under our Assumption 1, the asymptotic expansions of the test statistic (and of its bootstrap analogue) used to derive the limiting distributions need to have remainders of order $O_p\left(\delta_{NT}^{-4}\right)$, with $\delta_{NT} \equiv \min\left(\sqrt{N}, \sqrt{T}\right)$, whereas AGGR(2019) need to obtain expansions up to order $O_p\left(\delta_{NT}^{-6}\right)$.

for $j = 1, 2$. It follows that

$$\hat{\xi}(k^c) - k^c + \frac{1}{2N}\underbrace{tr\left(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}\right)}_{\equiv\mathcal{B}} = -\frac{1}{2N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\underbrace{\left(\mathcal{U}_t'\mathcal{U}_t - E\left(\mathcal{U}_t'\mathcal{U}_t\right)\right)}_{\equiv\mathcal{Z}_{N,t}} + O_p\left(\delta_{NT}^{-4}\right), \tag{2}$$

implying that

$$N\sqrt{T}\Omega_{\mathcal{U}}^{-1/2}\left(\hat{\xi}(k^c) - k^c + \frac{1}{2N}\mathcal{B}\right) \to^d N(0,1). \tag{3}$$

Theorem 2.1 corresponds to Theorem 1 in AGGR(2019) under our Assumptions 1-6. For completeness, we provide a proof of this result in Appendix A. As in AGGR(2019), we obtain an asymptotic expansion of $\hat{R}$ around $\tilde{R} \equiv \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21}$, where $\tilde{V}_{jk} \equiv T^{-1}\sum_{t=1}^{T} f_{jt}f_{kt}'$. We then use the fact that under the null hypothesis, f_{jt} and f_{kt} share a set of common factors f_t^c (i.e. $f_{jt} = \left(f_t^{c\prime}, f_{jt}^{s\prime}\right)'$ for $j = 1, 2$), implying that the k^c largest eigenvalues of $\tilde{R}$ are all equal to 1. This explains why $\hat{\xi}(k^c)$ is centered around k^c under the null. However, the asymptotic distribution of $\hat{\xi}(k^c)$ depends on the contribution of the factors estimation uncertainty to $\hat{V}_{jk} \equiv T^{-1}\sum_{t=1}^{T}\hat{f}_{jt}\hat{f}_{kt}'$, which involves products of $\hat{f}_{jt}$ and $\hat{f}_{kt}$. Using Bai (2003)'s identity for the factor estimation error $\hat{f}_{jt} - H_j f_{jt}$, we rely on Lemma A.2 in Appendix A (which gives an asymptotic expansion of $T^{-1}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_j f_{jt}\right)\left(\hat{f}_{kt} - H_k f_{kt}\right)'$ up to order $O_p\left(\delta_{NT}^{-4}\right)$) to obtain the asymptotic distribution in Theorem 2.1.[4]

Under our assumptions, the leading term of the asymptotic expansion of $\hat{\xi}(k^c) - k^c$ in (2) is given by $\frac{1}{2N}\mathcal{B}$, where $\mathcal{B} \equiv tr\left(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}\right)$. Since $\mathcal{B} = O_p(1)$ under our assumptions, $\frac{1}{2N}\mathcal{B}$ is of order $O_p\left(N^{-1}\right)$. The asymptotic Gaussianity of the test statistic is driven by the first term on the right hand side of (2), which we can rewrite as $-\frac{1}{N\sqrt{T}}\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}$, where $\mathcal{Z}_{N,t} \equiv \mathcal{U}_t'\mathcal{U}_t - E\left(\mathcal{U}_t'\mathcal{U}_t\right)$. Under Assumption 6, $\mathcal{Z}_{N,t}$ satisfies a central limit theorem, i.e. we assume[5] that $\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t} \to^d N(0, \Omega_{\mathcal{U}})$. Hence, $N\sqrt{T}\Omega_{\mathcal{U}}^{-1}\left(\hat{\xi}(k^c) - k^c + \frac{1}{2N}tr\left(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}\right)\right)$ is asymptotically distributed as $N(0,1)$, as stated in (3). Note that in deriving this result we have used the fact that $\sqrt{T}/N \to 0$ and $N/T^{3/2} \to 0$ to show that the remainder is $N\sqrt{T}O_p\left(\delta_{NT}^{-4}\right) = o_p(1)$.

Theorem 2.1 illustrates two crucial features of the asymptotic properties of the test statistic $\hat{\xi}(k^c)$ under the null. First, the test converges at a non-standard rate given by $N\sqrt{T}$. Second, the statistic $\hat{\xi}(k^c)$ is not centered at k^c even under the null. There is an asymptotic bias term of order $O_p\left(N^{-1}\right)$ given by $\mathcal{B}/2N$. When multiplied by $N\sqrt{T}$, this term is of order $O_p\left(\sqrt{T}\right)$. Thus, the bias is diverging but at a slower rate than the convergence rate $N\sqrt{T}$.

The distributional result (3) is infeasible since we do not observe the asymptotic bias $\mathcal{B}$ nor the asymptotic covariance matrix $\Omega_{\mathcal{U}}$. To obtain a feasible test statistic, we need consistent estimators of $\mathcal{B}$ and $\Omega_{\mathcal{U}}$. In particular, suppose that $\hat{\mathcal{B}}$ and $\hat{\Omega}_{\mathcal{U}}$ denote such estimators. Then, a feasible test statistic

[4] In contrast, AGGR(2019) rely on an asymptotic expansion up to order $O_p\left(\delta_{NT}^{-6}\right)$ because they require $N/T^{5/2} \to 0$ rather than $N/T^{3/2} \to 0$ (see their Proposition 3).

[5] AGGR(2019) provide conditions under which this high level condition holds. See in particular their Assumptions A.5 and A.6, which are used to show that $\mathcal{Z}_{N,t}$ is a Near Epoch Dependent (NED) process. Since our contribution is proving the bootstrap validity in this context, we do not provide these more primitive conditions. They are not required to prove our bootstrap theory.

is

$$\tilde{\xi}\left(k^{c}\right) \equiv N\sqrt{T}\hat{\Omega}_{\mathcal{U}}^{-1/2}\left(\hat{\xi}\left(k^{c}\right) - k^{c} + \frac{1}{2N}\hat{\mathcal{B}}\right).$$

Two crucial conditions for showing that $\tilde{\xi}\left(k^{c}\right) \to_{d} N\left(0,1\right)$ are (i) $\sqrt{T}\left(\hat{\mathcal{B}} - \mathcal{B}\right) = o_{p}\left(1\right)$ and (ii) $\hat{\Omega}_{\mathcal{U}} - \Omega_{\mathcal{U}} = o_{p}\left(1\right)$. Under these conditions, we can use a standard normal critical value to test H_0 against H_1. Since under H_1, $\hat{\xi}\left(k^{c}\right) - k^{c}$ is large and negative, the decision rule is to reject H_0 whenever $\tilde{\xi}\left(k^{c}\right) < z_{\alpha}$, where z_{α} is the α-quantile of a $N\left(0,1\right)$ distribution. This is the approach followed by AGGR(2019).

As it turns out, estimating $\mathcal{B}$ and $\Omega_{\mathcal{U}}$ is a difficult task when we allow for general time series and cross-sectional dependence in the idiosyncratic errors ε_{jt}. In particular, we can show that $\mathcal{B}$ depends on the cross-sectional dependence of ε_{1t} and ε_{2t} (but not on their serial dependence) whereas $\Omega_{\mathcal{U}}$ depends on both forms of dependence.

To see this, assume that $k^{c} = 1$ (and $k_{j}^{s} = 0$ for $j = 1, 2$), in which case $\mathcal{B} = \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}$. Assume also that $N = N_2 = N_1$, which implies that $\mu_N = 1$. When the idiosyncratic errors are independent across the two groups, we can write

$$\tilde{\Sigma}_{\mathcal{U}} \equiv T^{-1}\sum_{t=1}^{T} E\left(u_{1t} - u_{2t}\right)^{2} = T^{-1}\sum_{t=1}^{T}[E\left(u_{1t}^{2}\right) + E\left(u_{2t}^{2}\right)].$$

For each group j, $E\left(u_{jt}^{2}\right)$ captures the factor estimation uncertainty in $\hat{f}_{jt}$ and is given by $E\left(u_{jt}^{2}\right) = (N^{-1}\Lambda_{j}'\Lambda_{j})^{-2}\Gamma_{j,t}$, where $\Gamma_{j,t} \equiv Var(N^{-1/2}\sum_{i=1}^{N}\lambda_{j,i}\varepsilon_{j,it})$. It follows that

$$\tilde{\Sigma}_{\mathcal{U}} = (N^{-1}\Lambda_{1}'\Lambda_{1})^{-2}\Gamma_{1} + (N^{-1}\Lambda_{2}'\Lambda_{2})^{-2}\Gamma_{2},$$

where $\Gamma_{j} \equiv T^{-1}\sum_{t=1}^{T}\Gamma_{j,t}$. This shows that $\mathcal{B} = \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{U}$ depends on Γ_1 and Γ_2, the time averages of the variances of the cross-sectional averages of $\lambda_{j,i}\varepsilon_{j,it}$ for $j = 1, 2$. Hence, $\mathcal{B}$ depends on the cross-sectional dependence of each group's idiosyncratic errors, but it does not depend on their serial dependence.

To see that $\Omega_{\mathcal{U}}$ depends on both serial and cross-sectional dependence in $\varepsilon_{j,it}$, note that $\Omega_{\mathcal{U}} \equiv Var\left(\frac{1}{2}T^{-1/2}\sum_{t=1}^{T}\mathcal{Z}_{N,t}\right)$ is the long run variance of $\mathcal{Z}_{N,t} \equiv \left(u_{1t} - u_{2t}\right)^{2} - E\left(u_{1t} - u_{2t}\right)^{2}$, whose form depends on the potential serial dependence of $\varepsilon_{j,it}$. It also depends on the cross-sectional dependence because $\mathcal{Z}_{N,t}$ is a (quadratic) function of u_{jt}, which depends on the cross-sectional averages of $\varepsilon_{j,it}$. Thus, we conclude that $\Omega_{\mathcal{U}}$ is a complicated function of the serial and cross-sectional dependence in the idiosyncratic error terms.

For these reasons, in order to obtain a feasible test statistic, AGGR(2019) assume that each sub-panel follows a strict factor model. Under this assumption (including the assumption of conditional homoskedasticity in the idiosyncratic errors), the form of $\mathcal{B}$ and $\Omega_{\mathcal{U}}$ simplifies considerably. Their Theorem 2 provides consistent estimators of these quantities, allowing for the construction of a feasible test statistic. However, even under these restrictive assumptions, our simulations (to be discussed later) show important level distortions.

This provides the main motivation for using the bootstrap as an alternative method of inference. Our main goal is to propose a simple bootstrap test that avoids the need to estimate $\mathcal{B}$ and $\Omega_{\mathcal{U}}$ explicitly and outperforms the asymptotic theory-based test of AGGR(2019).

3 A general bootstrap scheme

3.1 The bootstrap data generating process and the bootstrap statistics

Let $\hat{\xi}^*(k^c)$ denote the bootstrap analogue of $\hat{\xi}(k^c)$. Our goal is to propose a bootstrap test that rejects H_0 whenever $p^* \leq \alpha$, where α is the significance level of the test and p^* is the bootstrap p-value defined as

$$p^* \equiv P^*\left(N\sqrt{T}\left(\hat{\xi}^*(k^c) - k^c\right) \leq N\sqrt{T}\left(\hat{\xi}(k^c) - k^c\right)\right).$$

The goal of this section is to propose asymptotically valid bootstrap methods. A crucial condition for bootstrap validity is that the bootstrap p-value is asymptotically distributed as $U_{[0,1]}$, a uniform random variable on $[0,1]$, when H_0 holds. Under H_1, the bootstrap p-value should converge to zero in probability to ensure that the bootstrap test has power. We propose a general residual-based bootstrap scheme that resamples the residuals from the two sub-panels in order to create the bootstrap data on y_{1t}^* and y_{2t}^*. We highlight the crucial conditions that the resampled idiosyncratic errors ε_{1t}^* and ε_{2t}^* need to verify in order to produce an asymptotically valid bootstrap test.

We adapt the general residual-based bootstrap method of GP(2014) to the group panel factor model. Specifically, for $j = 1, 2$, let $\left\{\varepsilon_{jt}^* : t = 1, \ldots, N_j\right\}$ denote a resampled version of $\left\{\tilde{\varepsilon}_{jt} = y_{jt} - \hat{\Lambda}_j^c \hat{f}_t^c - \hat{\Lambda}_j^s \hat{f}_{jt}^s\right\}$. The bootstrap data generating process (DGP) is

$$\begin{bmatrix} y_{1t}^* \\ y_{2t}^* \end{bmatrix} = \begin{bmatrix} \hat{\Lambda}_1^c & \hat{\Lambda}_1^s & 0 \\ \hat{\Lambda}_2^c & 0 & \hat{\Lambda}_2^s \end{bmatrix} \begin{bmatrix} \hat{f}_t^c \\ \hat{f}_{1t}^s \\ \hat{f}_{2t}^s \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^* \\ \varepsilon_{2t}^* \end{bmatrix}, \tag{4}$$

or, equivalently, for $j = 1, 2$, we let $Y_j^* = \tilde{F}_j \tilde{\Lambda}_j' + \varepsilon_j^*$, where $\tilde{F}_j = [\tilde{f}_{j1}, \ldots, \tilde{f}_{jT}]'$ is $T \times k_j$ and $\tilde{\Lambda}_j = (\tilde{\lambda}_{j,1}, \cdots, \tilde{\lambda}_{j,N_j})'$ is $N_j \times k_j$. An important feature of (4) is that it imposes the null hypothesis of k^c common factors between the two panels since the conditional mean of y_{jt}^* relies on the restricted estimated factors $\tilde{f}_{jt} = \left(\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime}\right)'$ for each $j = 1, 2$. This mimics the fact that y_{jt} depends on $f_{jt} = \left(f_t^{c\prime}, f_{jt}^{s\prime}\right)'$ under the null hypothesis. Similarly, ε_{jt}^* are a resampled version of the restricted residuals $\tilde{\varepsilon}_{jt}$. Although other bootstrap schemes that do not impose the null hypothesis could be considered[6], we focus on the null restricted bootstrap DGP in (4) for two main reasons. First, the fact that we impose the null hypothesis implies that the factors underlying the bootstrap DGP satisfy the normalization conditions imposed on the group factor model (see Assumption 2(a)). In particular,

[6] For example, we could use the principal components estimators $\hat{f}_{jt}$ and $\hat{\Lambda}_j$ when generating y_{jt}^*. To distinguish these estimators from their restricted versions, we denote the latter by $\tilde{f}_{jt}$ and $\tilde{\Lambda}_j$.

by construction $\hat{f}_t^c$ is orthogonal in-sample to $\hat{f}_{jt}^s$ for both $j = 1, 2$ when we use Definition 2 of AGGR(2019), and $T^{-1} \sum_{t=1}^T \hat{f}_t^c \hat{f}_t^{c\prime} = I_{k^c}$ and $T^{-1} \sum_{t=1}^T \hat{f}_{jt}^s \hat{f}_{jt}^{s\prime} = I_{k_j}$ for both $j = 1, 2$. These properties are crucial in showing the asymptotic Gaussianity of the bootstrap test statistic. Second, imposing the null hypothesis in the bootstrap DGP when doing hypothesis testing has been shown to be important to minimize the probability of type I error (see e.g., Davidson and MacKinnon (1999)).

Estimation in the bootstrap world proceeds as in the original sample. First, we extract the largest k_j principal components for each group j, with $j = 1, 2$, by applying the method of principal components to each sub-panel. In particular, the $T \times k_j$ matrix $\hat{F}_j^* = \begin{pmatrix} \hat{f}_{j1}^*, & \ldots, & \hat{f}_{jT}^* \end{pmatrix}'$ contains the estimated factors for each bootstrap sample generated from $Y_j^* = \tilde{F}_j \tilde{\Lambda}_j' + \varepsilon_j^*$. The matrix $\hat{F}_j^*$ collects the eigenvectors corresponding to the k_j largest eigenvalues of of $Y_j^* Y_j^{*\prime}/TN_j$ (arranged in decreasing order and multiplied by $\sqrt{T}$), where we impose that $\frac{\hat{F}_j^{*\prime}\hat{F}_j^*}{T} = I_{k_j}$. We then compute $\hat{R}^* = \hat{V}_{11}^{*-1}\hat{V}_{12}^*\hat{V}_{22}^{*-1}\hat{V}_{21}^*$, where $\hat{V}_{jk}^* = \hat{F}_j^{*\prime}\hat{F}_k^*/T = T^{-1}\sum_{t=1}^T \hat{f}_{jt}^*\hat{f}_{kt}^{*\prime}$. The bootstrap test statistic is $\hat{\xi}^*(k^c) = \sum_{l=1}^{k^c} \hat{\rho}_l^* = tr\left(\hat{\Lambda}^{*1/2}\right)$, where $\hat{\Lambda}^* = diag\left(\hat{\rho}_l^{*2} : l = 1, \ldots, k^c\right)$ is a $k^c \times k^c$ diagonal matrix containing the k^c largest eigenvalues of $\hat{R}^*$ obtained from the eigenvalue-eigenvector problem $\hat{R}^*\hat{W}^* = \hat{W}^*\hat{\Lambda}^*$, where $\hat{W}^*$ is the $k_1 \times k^c$ matrix eigenvector matrix.

As in the original sample, estimation by principal components using the bootstrap data Y_j^* implies that each estimator $\hat{f}_{jt}^*$ is consistent for $H_j^*\tilde{f}_{jt}$, a rotated version of $\tilde{f}_{jt}$. The bootstrap rotation matrix is defined as $H_j^* = \mathcal{V}_j^{*-1}\frac{\hat{F}_j^{*\prime}\tilde{F}_j}{T}\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}$, where $\mathcal{V}_j^*$ is a $k_j \times k_j$ diagonal matrix containing the k_j largest eigenvalues of $Y_j^*Y_j^{*\prime}/N_jT$ on the main diagonal, in decreasing order. Contrary to H_j, H_j^* is observed and could be used for inference on the factors as in Gonçalves and Perron (2014). Here, the bootstrap test statistic $\hat{\xi}^*(k^c)$ is invariant to H_j^*, but it shows up in the bootstrap theory. The bootstrap p-value p^* is based on $N\sqrt{T}\left(\hat{\xi}^*(k^c) - k^c\right)$, where $\hat{\xi}^*(k^c)$ is centered around k^c because we have imposed the null hypothesis in the bootstrap DGP in (4).

Next, we characterize the bootstrap distribution of $\hat{\xi}^*(k^c)$. Following the proof of Theorem 2.1, we expand $\hat{R}^*$ around $\tilde{R}^* \equiv \tilde{V}_{11}^{*-1}\tilde{V}_{12}^*\tilde{V}_{22}^{*-1}\tilde{V}_{21}^*$, where $\tilde{V}_{jk}^* \equiv T^{-1}\sum_{t=1}^T \tilde{f}_{jt}\tilde{f}_{kt}'$ is the bootstrap analogue of $\tilde{V}_{jk} \equiv T^{-1}\sum_{t=1}^T f_{jt}f_{kt}'$.[7] Given (4), $\tilde{f}_{jt}$ and $\tilde{f}_{kt}$ share a set of common factors $\hat{f}_t^c$ (i.e. $\tilde{f}_{jt} = \left(\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime}\right)'$ for $j = 1, 2$), implying that the k^c largest eigenvalues of $\tilde{R}^*$ are all equal to 1 and $\hat{\xi}^*(k^c)$ is centered around k^c. Note that this holds by construction, independently of whether the null hypothesis H_0 is true or not. As argued for the original statistic, the bootstrap distribution of $\hat{\xi}^*(k^c)$ is driven by the contribution of the factors estimation uncertainty (as measured by $\hat{f}_{jt}^* - H_j^*\tilde{f}_{jt}$) to $\hat{V}_{jk}^* \equiv T^{-1}\sum_{t=1}^T \hat{f}_{jt}^*\hat{f}_{kt}^{\prime *}$. In particular, following the proof of Theorem 2.1, the asymptotic distribution of $\hat{\xi}^*(k^c)$ is based on an asymptotic expansion of $T^{-1}\sum_{t=1}^T\left(\hat{f}_{jt}^* - H_j^*\tilde{f}_{jt}\right)\left(\hat{f}_{kt}^* - H_k^*\tilde{f}_{kt}\right)'$ up to order $O_{p^*}\left(\delta_{NT}^{-4}\right)$. This crucial result is given in Lemma B.2 in Appendix B. It relies on Conditions A*, B* and C*, which are the bootstrap analogues of Assumptions 3, 4 and 5. We call these bootstrap high level conditions because they apply to any bootstrap method that is used to generate the bootstrap

[7] Although $\tilde{V}_{jk}^*$ is defined as a function of $\tilde{f}_{kt}$ and does not depend on resampled data, we use this notation to indicate that it is the bootstrap analogue of $\tilde{V}_{jk}$.

draws ε^*_{jt}. We will verify these conditions for the wild bootstrap in the next section.

The following result follows under Conditions A*-C*. We let $\mathcal{U}^*_t \equiv \mu_N u^{(c)*}_{1t} - u^{(c)*}_{2t}$, where $u^{(c)*}_{jt}$ denotes the first k^c rows of $u^*_{jt} \equiv \left(N_j^{-1}\tilde{\Lambda}_j'\tilde{\Lambda}_j\right)^{-1} N_j^{-1/2}\sum_{i=1}^{N_j}\tilde{\lambda}_{j,i}\varepsilon^*_{j,it}$. Similarly, we let $\tilde{\Sigma}^*_{\mathcal{U}} \equiv T^{-1}\sum_{t=1}^T E^*\left(\mathcal{U}^*_t\mathcal{U}^{*\prime}_t\right)$, which is the bootstrap analogue of $\tilde{\Sigma}_{\mathcal{U}} \equiv T^{-1}\sum_{t=1}^T E\left(\mathcal{U}_t\mathcal{U}_t'\right)$.

Lemma 3.1 *Suppose that Conditions A*, B*, and C* hold. It follows that*

$$\hat{\xi}^*\left(k^c\right) - k^c + \frac{1}{2N}\underbrace{tr\left(\tilde{\Sigma}^*_{\mathcal{U}}\right)}_{\equiv \mathcal{B}^*} = -\frac{1}{2N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^T\underbrace{\left(\mathcal{U}^{*\prime}_t\mathcal{U}^*_t - E^*\left(\mathcal{U}^{*\prime}_t\mathcal{U}^*_t\right)\right)}_{\equiv \mathcal{Z}^*_{N,t}} + O_{p^*}\left(\delta_{NT}^{-4}\right). \qquad (5)$$

Lemma 3.1 gives the asymptotic expansion of $\hat{\xi}^*\left(k^c\right)$ and is the bootstrap analogue of (2) in Theorem 2.1. The leading term in the expansion of $\hat{\xi}^*(k^c) - k^c$ in (5) is given by $\frac{1}{2N}\mathcal{B}^*$, where $\mathcal{B}^* \equiv tr\left(\tilde{\Sigma}^*_{\mathcal{U}}\right)$ is the bootstrap analogue of $\mathcal{B} \equiv tr\left(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}\right)$. Note that in the bootstrap world, $\tilde{\Sigma}^*_{cc} \equiv T^{-1}\sum_{t=1}^T \hat{f}^c_t\hat{f}^{c\prime}_t = I_{k^c}$, which explains why $\tilde{\Sigma}^{*-1}_{cc}$ is omitted from the definition of $\mathcal{B}^*$. Under our bootstrap high level conditions, $\frac{1}{2N}\mathcal{B}^*$ is of order $O_{p^*}\left(N^{-1}\right)$.

To show the asymptotic validity of the bootstrap test, we impose the following additional bootstrap high level conditions. We define $\mathcal{Z}^*_{N,t} \equiv \mathcal{U}^{*\prime}_t\mathcal{U}^*_t - E^*\left(\mathcal{U}^{*\prime}_t\mathcal{U}^*_t\right)$, and let $\Omega^*_{\mathcal{U}} \equiv Var^*\left(\frac{1}{2\sqrt{T}}\sum_{t=1}^T \mathcal{Z}^*_{N,t}\right)$.

Condition D* $\sqrt{T}\left(\mathcal{B}^* - \mathcal{B}\right) \to_p 0$.

Condition E* $\Omega^{*-1/2}_{\mathcal{U}}\frac{1}{2\sqrt{T}}\sum_{t=1}^T \mathcal{Z}^*_{N,t} \overset{d^*}{\to}_p N\left(0,1\right)$, where $\Omega^*_{\mathcal{U}} \equiv Var^*\left(\frac{1}{2\sqrt{T}}\sum_{t=1}^T \mathcal{Z}^*_{N,t}\right)$ is such that $\Omega^*_{\mathcal{U}} - \Omega_{\mathcal{U}} \to_p 0$.

Theorem 3.1 *Assume Assumptions 1-6 hold and* H_0 *is true. Then, any bootstrap scheme that verifies Conditions A*-E* is such that*

$$N\sqrt{T}\Omega_{\mathcal{U}}^{-1/2}\left(\hat{\xi}^*\left(k^c\right) - k^c + \frac{1}{2N}\mathcal{B}\right) \overset{d^*}{\to}_p N\left(0,1\right),$$

which implies that $p^* \overset{d}{\to} U_{[0,1]}$.

Condition D* requires the bootstrap bias $\mathcal{B}^*$ to mimic the bias term $\mathcal{B}$. In particular, $\mathcal{B}^*$ needs to be a $\sqrt{T}$-convergent estimator of $\mathcal{B}$. Having $\mathcal{B}^* - \mathcal{B} = o_p\left(1\right)$ does not suffice. The main reason for the faster rate of convergence requirement is that the asymptotic bias term $(2N)^{-1}\mathcal{B}$ is of order $O_p\left(N^{-1}\right)$ and since the convergence rate is $N\sqrt{T}$, this induces a shift of the center of the distribution of order $O_p\left(\sqrt{T}\right)$. So, contrary to more standard settings where the asymptotic bias is of order $O\left(1\right)$, here the asymptotic bias diverges. However, any $\sqrt{T}$-consistent estimator of $\mathcal{B}$ can be used to recenter $\hat{\xi}\left(k^c\right) - k^c$ and yield a random variable whose limiting distribution is $N\left(0, \Omega_{\mathcal{U}}\right)$. Condition D* requires that the bootstrap bias $\mathcal{B}^*$ has this property. Condition E* requires that the bootstrap array $\mathcal{Z}^*_{N,t}$ satisfies a central limit theorem in the bootstrap world with an asymptotic variance-covariance matrix $\Omega^*_{\mathcal{U}}$ that converges in probability to $\Omega_{\mathcal{U}}$. This condition is the bootstrap analogue of Assumption 6-(b) in Appendix A.

We discuss a few implications of our bootstrap high level conditions. The first one is that for the bootstrap to mimic the asymptotic bias term $\mathcal{B}$ (as implied by Condition D*) we need to generate ε_{jt}^* in a way that preserves the cross-sectional dependence of ε_{jt}. Serial dependence in ε_{jt} is asymptotically irrelevant for this term. The reason for this is that $\mathcal{B}$ depends only on the cross-sectional dependence but not on the serial dependence of ε_{jt}, as we explained in the previous section.

The second implication is that in order for the bootstrap to replicate the covariance $\Omega_{\mathcal{U}}$ (as required by Condition E*) we need to design a bootstrap method that generates ε_{jt}^* with serial dependence (in addition to cross-sectional dependence). This can be seen by noting that $\Omega_{\mathcal{U}}$ is the long run variance of $\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}$, which depends on both the serial and the cross-sectional dependence properties of $\{\varepsilon_{jt}\}$.

The overall conclusion is that the implementation of the bootstrap depends on the serial and cross-sectional dependence assumptions we make on the idiosyncratic errors of each sub-panel. Different assumptions will lead to different bootstrap algorithms. Theorem 3.1 is useful because it gives a set of high-level conditions that can be used to prove the asymptotic validity of the bootstrap for any bootstrap scheme used to obtain ε_{jt}^*.

To end this section, we discuss the asymptotic power of our bootstrap test. Although Conditions A*-E* suffice to show that $p^* \xrightarrow{p} 0$ under H_1, a weaker set of assumptions suffices. In particular, the following high level condition is sufficient to ensure that any bootstrap test based on $\hat{\xi}^*(k^c)$ is consistent.

Condition F* $\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}^* = O_p(1)$ and $\mathcal{B}^* = O_p\left(N^{1-\epsilon}\right)$, where ϵ is some positive number.

Proposition 3.1 *Under Assumptions 1-6, any bootstrap method that verifies Conditions A*-C* and F* satisfies* $p^* \xrightarrow{p} 0$ *under* H_1.

Since we reject H_0 if $p^* \leq \alpha$, Proposition 3.1 ensures that $P(p^* \leq \alpha) \to 1$ when H_1 is true.

4 Specific bootstrap schemes

4.1 The wild bootstrap method

Here, we discuss a wild bootstrap method and show that it verifies Conditions A*-F* under a set of assumptions similar to those of Theorem 2 in AGGR(2019). Algorithm 1 below contains a description of this method.

Algorithm 1 : Wild Bootstrap

1. For $t = 1, \ldots, T$, and $j = 1, 2$, let
$$y_{jt}^{*} = \tilde{\Lambda}_j \tilde{f}_{jt} + \varepsilon_{jt}^{*},$$
where $\tilde{f}_{jt} = (\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime})'$ and $\varepsilon_{jt}^{*} = \left(\varepsilon_{j,1t}^{*}, \ldots, \varepsilon_{j,N_j t}^{*}\right)'$ is such that
$$\varepsilon_{j,it}^{*} = \tilde{\varepsilon}_{j,it} \eta_{j,it},$$
and $\eta_{j,it}$ are i.i.d. $N(0,1)$ across (j, i, t).

2. For $j = 1, 2$, estimate the bootstrap factors $\hat{F}_j^{*}$ by extracting the first k_j principal components from y_{jt}^{*}, and set
$$\hat{V}_{jl}^{*} = \frac{1}{T} \hat{F}_j^{*\prime} \hat{F}_l^{*}, \ \ j, l = 1, 2,$$
and
$$\hat{R}^{*} = \hat{V}_{11}^{*-1} \hat{V}_{12}^{*} \hat{V}_{22}^{*-1} \hat{V}_{21}^{*}.$$

3. Compute the k^c largest eigenvalues of $\hat{R}^{*}$ and denote these by $\hat{\rho}_l^{*2}$, $l = 1, \ldots, k^c$.

4. Compute the bootstrap test statistic $\hat{\xi}^{*}(k^c) = \sum_{l=1}^{k^c} \hat{\rho}_l^{*}$.

5. Repeat steps 1-4 M times and then compute the bootstrap p-value as
$$p^{*} = \frac{1}{M} \sum_{b=1}^{M} 1\left(\hat{\xi}^{*(b)}(k^c) \leq \hat{\xi}(k^c)\right),$$
where $\hat{\xi}^{*(b)}(k^c)$ is the value of the bootstrap test for replication $b = 1, \ldots, M$.

6. Reject the null hypothesis of k^c common factors at level α if $p^{*} \leq \alpha$.

To prove the asymptotic validity of the wild bootstrap p-value, we strengthen the primitive assumptions given in Appendix A as follows.

Assumption WB1 For $j = 1, 2$, $\{f_{jt}\}$ and $\{\varepsilon_{j,it}\}$ are mutually independent such that $E\|f_{jt}\|^{32} \leq M < \infty$ and $E|\varepsilon_{j,it}|^{32} \leq M < \infty$ for all (i, t).

Assumption WB2 (a) $Cov(\varepsilon_{j,it}, \varepsilon_{k,ls}) = 0$ if $j \neq k$ or $i \neq l$ or $t \neq s$, and (b) $E(\varepsilon_{j,it}^2) = \gamma_{j,ii} > 0$.

Assumption WB3 For each $j = 1, 2$,

(a) $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_{j,it}^2 \varepsilon_{j,kt}^2 - E(\varepsilon_{j,it}^2 \varepsilon_{j,kt}^2) = O_p(1)$ for any i, k.

(b) $\max_{i \leq N_j} \left\| \frac{1}{T} \sum_{t=1}^{T} f_{jt} \varepsilon_{j,it} \right\| = O_p\left(\sqrt{\frac{\log N_j}{T}}\right)$.

(c) $E\left\| \frac{1}{N_j} \Lambda_j' \varepsilon_{jt} \right\|^2 \leq M$.

Assumption WB1 strengthens the moment conditions in Assumption 2 and Assumption 3-(a). A larger number of moments of f_{jt} and $\varepsilon_{j,it}$ is required here than in GP (2014) (who require the existence of 12 moments rather than 32). As explained above, our bootstrap test statistic $\hat{\xi}^*(k^c)$ involves products and cross products of bootstrap estimated factors from each sub-panel. The derivation of the bootstrap asymptotic distribution of $\hat{\xi}^*(k^c)$ relies on Lemma B.2 which obtains an asymptotic expansion of $T^{-1}\sum_{t=1}^{T}\left(\hat{f}_{jt}^* - H_j^* \tilde{f}_{jt}\right)\left(\hat{f}_{kt}^* - H_k^* \hat{f}_{kt}\right)'$ up to order $O_{p^*}\left(\delta_{NT}^{-4}\right)$. This requires not only the verification of Conditions A* and B* from GP (2014) (who obtain an asymptotic expansion of $T^{-1}\sum_{t=1}^{T}\left(\hat{f}_{jt}^* - H_j^* \tilde{f}_{jt}\right)\left(\hat{f}_{kt}^* - H_k^* \hat{f}_{kt}\right)'$ up to order $O_{p^*}\left(\delta_{NT}^{-2}\right)$), but also of Condition C*, which is new to this paper. The large number of moments is used in verifying this condition. In particular, we rely on repeated applications of Cauchy-Schwarz's inequality, and bound sums such as $\frac{1}{N_j T}\sum_{i=1}^{N_j}\sum_{t=1}^{T}|\tilde{\varepsilon}_{j,it}|^p$ for $p \leq 16$, which requires the existence of $2p$ moments of f_{jt} and $\varepsilon_{j,it}$ (see Lemma C.1).

Assumption WB2 rules out cross-sectional and serial correlation in the idiosyncratic errors of each sub-panel as well as correlation among ε_{jt} and ε_{kt} for $j \neq k$. These assumptions are similar to the assumptions used by AGGR(2019) to justify their feasible test statistic (see their Theorem 2). For simplicity, we assume the external random variable $\eta_{j,it}$ to be Gaussian, but the result generalizes to any i.i.d. draw that is mean zero and variance one with finite eight moments and a symmetric distribution.

Theorem 4.1 *Assume that Assumptions 1-6 strengthened by Assumptions WB1, WB2, and WB3 hold. Then, if Algorithm 1 is used to generate ε_{jt}^* for $j = 1, 2$, the conclusions of Theorem 3.1 and Proposition 3.1 apply.*

Theorem 4.1 justifies theoretically using the wild bootstrap p-value p^* to test the null hypothesis of k^c common factors. Although Assumption WB2 rules out dependence in ε_{jt} in both dimensions, as in Theorem 2 of AGGR(2019), this bootstrap test does not require an explicit bias correction nor a variance estimator. We show in Section 5 that the feasible test statistic AGGR(2019) can be oversized even under these restrictive assumptions. The wild bootstrap essentially eliminates these level distortions.

4.2 An extension: AR-CSD bootstrap method

Here, we discuss an extension of the wild bootstrap that allows for cross-sectional and serial dependence in the idiosyncratic error terms of each sub-panel. In particular, we assume that for each $j = 1, 2$, and $i = 1, \ldots, N_j$, the idiosyncratic errors $\varepsilon_{j,it}$ follow an AR(p) model (autoregressive model of order p):

$$\varepsilon_{j,it} = a_{ji}(L)\,\varepsilon_{j,i,t-1} + v_{j,it}, \tag{6}$$

where $a_{ji}(L) = a_{ji,1} + a_{ji,2}L + \ldots + a_{ji,p-1}L^{p-1}$. If we collect all observations i for panel j, we can write this as $\varepsilon_{jt} = A_j(L)\,\varepsilon_{j,t-1} + v_{jt}$, where $A_j(L) = A_{j,1} + A_{j,2}L + \ldots + A_{j,p-1}L^{p-1}$ and $A_{j,k}$ are $N_j \times N_j$ diagonal matrices with coefficients $a_{ji,k}$ along the main diagonal. Since N_j is large, consistent estimation of $A_{j,k}$ is not feasible unless we impose some form of sparsity. Assuming that each series $\varepsilon_{j,it}$ is an autoregressive process with possibly heterogeneous coefficients is a restrictive form of sparsity which allows the use of OLS[8]. In addition, we assume that

$$v_{jt} \sim \text{i.i.d.}\left(0, \Sigma_{v,j}\right), \quad \Sigma_{v,j} = \left(\gamma_{j,il}\right)_{i,l=1,\ldots,N_j}.$$

The fact that we allow for a possibly non-diagonal covariance matrix $\Sigma_{v,j}$ means that we allow for cross-sectional dependence in the innovations v_{jt}.

Our proposal is to create bootstrap observations ε_{jt}^* using a residual-based bootstrap procedure that resamples the residuals of the AR model (6). Resampling the vector of the AR(p) residuals $\tilde{v}_{jt}$ allowing for unrestricted cross-sectional dependence is complicated due to the fact that the covariance matrix $\Sigma_{v,j}$ is high dimensional. In particular, i.i.d. resampling of $\tilde{v}_{j,it}$ is not valid, as shown by Gonçalves and Perron (2020) in the context of factor augmented regression models. Our bootstrap algorithm (described in Algorithm 2) relies on the cross-sectional dependent (CSD) bootstrap of Gonçalves and Perron (2020). In the following, we let $\tilde{\Sigma}_{v,j}$ denote any consistent estimator of $\Sigma_{v,j}$ under the spectral norm. Examples include the thresholding estimator of Bickel and Levina (2008a) and the banding estimator of Bickel and Levina (2008b).

Algorithm 2 : AR-CSD Bootstrap

1. For $t = 1, \ldots, T$, and $j = 1, 2$, let

$$y_{jt}^* = \tilde{\Lambda}_j \tilde{f}_{jt} + \varepsilon_{jt}^*,$$

where $\tilde{f}_{jt} = (\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime})'$ and $\varepsilon_{jt}^* = \left(\varepsilon_{j,1t}^*, \ldots, \varepsilon_{j,N_jt}^*\right)'$ is such that

$$\varepsilon_{j,it}^* = \tilde{a}_{ji}\left(L\right)\varepsilon_{j,i,t-1}^* + v_{j,it}^*, \text{ for } t = 1, \ldots, T$$

with $\varepsilon_{j,i0}^* = 0$ for $i = 1, \ldots, N_j$ and where $v_{j,it}^*$ is i-th element of v_{jt}^* obtained as

$$v_{jt}^* = \tilde{\Sigma}_{v,j}^{1/2}\eta_{jt}, \quad \text{where } \eta_{jt} \text{ is i.i.d.} N\left(0, I_{N_j}\right) \text{ over } t.$$

2. Repeat steps 2 through 6 of Algorithm 1.

The wild bootstrap algorithm (Algorithm 1) is a special case of Algorithm 2 when we set $\tilde{a}_{ji}\left(L\right) = 0$ for all i and $\tilde{\Sigma}_{v,j} = diag\left(\tilde{\varepsilon}_{j,it}^2\right)$. Another special case is the cross-sectional dependent (CSD) bootstrap of Gonçalves and Perron (2020), which sets $\tilde{a}_{ji}\left(L\right) = 0$ and lets $\tilde{\Sigma}_{v,j}$ denote the thresholding estimator based on the sample covariances of $\tilde{\varepsilon}_{j,it}$. Finally, a generalization of Algorithm 2 is the sieve bootstrap proposed by Koh (2022) in the context of MIDAS factor models. Although it would be interesting to extend the sieve bootstrap to our testing problem, we focus on a class of finite order AR models here

[8] We could allow for richer dynamics by assuming a sparse VAR model for the idiosyncratic error vector ε_{jt}, as in Kock and Callot (2015), Krampe and Paparoditis (2021), and Krampe and Margaritella (2021). Under sparsity, we would estimate $A_j\left(L\right)$ by a regularized OLS estimator such as LASSO rather than OLS. The remaining steps of our bootstrap method would remain the same.

in order to simplify the analysis.

The proof of the asymptotic validity of Algorithm 2 follows from Theorem 3.1 and Proposition 3.1 by verifying Conditions A* - F*. Since $\varepsilon^*_{j,it}$ is both serially and cross-sectionally correlated, the verification of these bootstrap high-level conditions is much more involved than for the wild bootstrap and beyond the scope of this paper. However, we evaluate by simulation the performance of both Algorithms 1 and 2 in the next section.

5 Simulations

In this section, we compare the performance of the bootstrap methods discussed in the previous sections. Our data generating process (DGP) is a simple model with one factor for each group:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{7}$$

where y_{jt} and ε_{jt} are $N_j \times 1$ for $t = 1, \ldots, T$. As opposed to Andreou et al. (2019), we assume that both groups have the same frequency.

For level experiments, we let $f_{1t} = f_{2t} = f^c_t$. As in Gonçalves and Perron (2014), this common factor is generated independently over time from a standard normal distribution, $f^c_t \sim$ i.i.d.$N(0,1)$. For power experiments, each group has a specific factor $f_{1t} = f^s_{1t}$ and $f_{2t} = f^s_{2t}$. These two group-specific factors are also generated independently over time from a bivariate normal with unit variance and correlation $\phi = 0.99$. In all cases, the factor loadings are drawn independently from a standard normal distribution, $\Lambda_j \sim$ i.i.d.$N(0,1)$, $j = 1, 2$.

The idiosyncratic error terms in the model, $\varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})'$, are such that

$$\varepsilon_t = A_\varepsilon \varepsilon_{t-1} + v_t$$

where A_ε is a block-diagonal matrix

$$A_\varepsilon = \begin{bmatrix} a_{\varepsilon,1} I_{N_1} & \mathbf{0}_{(N_1 \times N_2)} \\ \mathbf{0}_{(N_2 \times N_1)} & a_{\varepsilon,2} I_{N_2} \end{bmatrix}$$

and $a_{\varepsilon,j}$ is the AR(1) coefficient in group j (we assume that all individual series in each group share the same autoregressive coefficient). The innovations in the idiosyncratic error terms, $v_t = (v'_{1t}, v'_{2t})'$, are such that:

$$v_{1t} \sim N(0, (1 - a^2_{\varepsilon,1})\Sigma_{v,1}), \ v_{2t} \sim N(0, (1 - a^2_{\varepsilon,2})\Sigma_{v,2}),$$

where $\Sigma_{v,1}$ is the first diagonal block and $\Sigma_{v,2}$ is the second diagonal block of

$$\Sigma_v = \begin{bmatrix} \{\beta^{|i-j|}\}_{i,j=1,\ldots,N_1} & \mathbf{0}_{(N_1 \times N_2)} \\ \mathbf{0}_{(N_2 \times N_1)} & \{\beta^{|i-j|}\}_{i,j=1,\ldots,N_2} \end{bmatrix}.$$

The scalar β induces cross-sectional dependence in each group among the idiosyncratic innovations.

Table 1: Data generating processes

DGP	$a_{\varepsilon,1}$	$a_{\varepsilon,2}$	β
Design 1 (no serial & no cross-sectional dependence)	0	0	0
Design 2 (only serial dependence)	0.5	0.3	0
Design 3 (only cross-sectional dependence)	0	0	0.2
Design 4 (serial & cross-sectional dependence)	0.5	0.3	0.2

Table 2: Sample sizes in simulation experiment

N_1	N_2	T
50	50	50
50	50	100
50	50	200
100	100	50
100	100	100
100	100	200
200	200	50
200	200	100
200	200	200

This is similar to the design in Bai and Ng (2006). Note that we assume that Σ_v is a block diagonal matrix, so we do not consider dependence between the two groups. In Table 1, we report the parameter settings we consider.

In Design 1, we assume that there is no serial correlation and no cross-sectional dependence and that the idiosyncratic errors are homoskedastic. The idiosyncratic error terms in Design 2 are serially correlated in each group where the AR(1) coefficient in group 1 is larger than the one in group 2. In the third design, we consider cross-sectional dependence without serial correlation in the idiosyncratic error term. Finally, in the last design, the idiosyncratic innovation terms are both serially and cross-sectionally correlated.

We consider sample sizes $N_1 = N_2 = N$ between 50 and 200 and T between 50 and 200. We simulate each design 5000 times, and the bootstrap replication number is set at 399. We use the bootstrap algorithms proposed in Sections 3 and 4 with four different bootstrap methods: the wild bootstrap, the AR(1)-CSD bootstrap proposed earlier and two variants: a parametric AR(1) bootstrap with no cross-sectional dependence and a CSD bootstrap with no serial dependence. The CSD and AR(1)-CSD bootstraps involve an estimator of the covariance matrix of the idiosyncratic errors. We rely on the banding estimator of Bickel and Levina (2008b) with the banding parameter k chosen by their cross-validation procedure. We focus our results on $\alpha = 0.05$ and report rejection rates for each

Table 3: Rejection rate of 5% test - level

		$N = 50$			$N = 100$			$N = 200$		
		$T = 50$	100	200	$T = 50$	100	200	$T = 50$	100	200
Design 1 i.i.d.	AGGR	6.5	4.9	3.3	7.4	6.2	5.0	8.3	7.2	6.2
	WB	5.3	4.9	4.3	5.8	5.7	5.1	6.7	6.4	5.6
	AR (1)	5.1	4.9	4.1	5.9	5.6	5.1	6.7	6.1	5.8
	CSD	6.1	5.9	5.6	6.0	5.7	5.6	6.7	6.2	6.0
	AR(1)-CSD	7.3	6.7	5.8	7.5	6.8	6.3	8.4	7.3	6.7
Design 2 AR	AGGR	14.3	10.0	7.7	15.2	12.4	9.8	17.7	13.8	10.8
	WB	9.8	8.7	8.0	10.4	9.8	8.9	12.5	10.7	9.3
	AR (1)	4.9	4.7	4.2	5.9	5.7	4.8	6.9	5.9	5.4
	CSD	11.9	12.5	14.7	11.5	11.5	11.4	13.0	11.0	10.1
	AR(1)-CSD	6.4	6.2	6.1	7.2	7.0	6.3	7.8	7.1	6.2
Design 3 CSD	AGGR	21.6	18.9	20.5	20.4	17.5	16.9	20.8	16.8	14.2
	WB	15.9	16.5	20.3	15.0	14.5	15.8	15.2	13.4	12.6
	AR (1)	9.6	11.7	15.0	8.9	9.5	10.7	8.7	8.0	8.4
	CSD	7.2	8.2	8.2	8.8	8.7	7.4	10.5	9.4	8.3
	AR(1)-CSD	5.8	5.3	5.3	6.3	5.6	4.9	6.9	6.2	5.8
Design 4 AR + CSD	AGGR	21.6	18.9	20.5	20.4	17.5	16.9	20.8	16.8	14.2
	WB	15.7	16.6	20.1	15.1	14.7	15.8	15.2	13.8	12.4
	AR (1)	9.6	11.6	15.0	8.6	9.4	10.6	8.7	8.2	8.2
	CSD	7.5	8.0	8.0	8.9	8.2	7.7	10.5	9.4	8.4
	AR(1)-CSD	5.4	5.3	5.1	6.5	5.8	4.9	7.1	6.1	5.7

design, bootstrap method, and sample size.

The simulation results for the level experiments are shown in Table 3. The row labeled "AGGR" reports results based on the asymptotic standard normal critical values. The other four rows contain the results for the bootstrap methods: WB for wild bootstrap and AR(1) for parametric AR(1) bootstrap method, CSD for the cross-sectional bootstrap, and AR-CSD for the bootstrap that combines the autoregressive and cross-sectional dependent bootstrap.

Under the restrictive Design 1 where the assumptions of Theorem 2 of Andreou et al. (2019) are satisfied, the asymptotic theory performs reasonably well, although some distortions appear for the smaller value of T. For the other three designs, we find severe over-rejections for all sample sizes, as expected given that the statistic is computed assuming away autocorrelation and cross-sectional dependence.

In all sets of samples and designs, bootstrap methods provide more reliable inference than standard normal inference. The bootstrap method that performs best is typically the one tailored to the properties of the DGP. For example, in Design 1, both the wild bootstrap and the AR(1) bootstrap perform similarly, and they reject the null hypothesis at a rate close to 5%. To illustrate, for $N_1 =$

$N_2 = 100$ and $T = 50$, the test rejects in 7.4% of the replications using the standard normal critical values. The rejection rates for the wild bootstrap and AR(1) bootstrap are 5.8% and 5.9% respectively. On the other hand, the cross-sectional bootstrap and combined AR(1) and CSD bootstrap reject in 6.0% and 7.5% of the replications. This higher rejection rate is the cost of using a more robust method than necessary.

As mentioned above, in Designs 2 - 4, the feasible statistic in Andreou et al. (2019) leads to large level distortions since it is not robust to serial correlation or cross-sectional dependence. Because there is serial dependence in the idiosyncratic error terms in Design 2, the wild bootstrap and CSD bootstrap are no longer valid while still improving on the use of the standard normal critical values. In this case, both the AR(1) and AR(1)-CSD bootstraps are valid and provide similar results with a slight preference for the simple AR(1) bootstrap. To illustrate, with the same $N_1 = N_2 = 100$ and $T = 50$ as above, the standard normal critical values lead to a rejection rate of 15.2% for a 5% test. The (invalid) wild and CSD bootstraps have rejection rates of 10.4% and 11.5% respectively. On the other hand, the (valid) AR(1) and AR(1)-CSD bootstraps have rejection rates of 5.9% and 7.2%.

In Designs 3 and 4, where we introduce cross-sectional dependence, neither the wild bootstrap nor the AR(1) bootstrap are valid and they are not performing well, as expected. In the most general design with both serial and cross-sectional dependence, only the AR(1)-CSD bootstrap provides reliable results. While the asymptotic theory in the $N_1 = N_2 = 100$ and $T = 50$ case shows a rejection rate of 20.4%, the AR(1)-CSD bootstrap has a rejection rate of 6.3% compared with 8.8% for the CSD bootstrap, 8.9% for the AR(1) bootstrap, and 15.0% for the simple wild bootstrap.

Our power results are presented in Table 4. These results must be interpreted with caution given the large level distortions documented in some cases. For the simple i.i.d. case (Design 1) where all tests have reasonable rejection rates under the null, we see that the bootstrap entails a small reduction in power relative to the AGGR test. The largest loss occurs for $N_1 = N_2 = T = 50$ where the AGGR test has a power of 65.2% while the wild bootstrap rejects in 61.5% of the cases. The gap between the two methods disappears as sample size increases in both dimensions.

It is interesting to note that power increases faster in the cross-sectional than in the time series dimension. Going from $N = 50$ to $N = 100$ for given T has more impact on power than going from $T = 50$ to $T = 100$ for given N. This is consistent with the different rates of convergence of the statistic in the two dimensions.

Finally, we see that more complex idiosyncratic dependencies lead to a reduction in power for bootstrap methods that control level. Nevertheless, power approaches one rather quickly.

Overall, our results suggest that except for the simple case with no serial or cross-sectional dependence and large sample sizes, the use of standard normal critical values leads to large level distortions. On the other hand, a bootstrap method that adapts to the properties of the idiosyncratic terms provides excellent coverage rates, while a misspecified bootstrap still improves matters noticeably. The use of more robust bootstrap methods has a small cost in terms of power.

Table 4: Rejection rate of 5% test - power

		$N = 50$			$N = 100$			$N = 200$		
		$T = 50$	100	200	$T = 50$	100	200	$T = 50$	100	200
Design 1 i.i.d.	AGGR	65.2	83.5	96.4	96.4	99.7	100.0	100.0	100.0	100.0
	WB	61.5	83.0	97.2	95.5	99.7	100.0	100.0	100.0	100.0
	AR (1)	60.8	83.4	97.4	95.2	99.6	100.0	99.9	100.0	100.0
	CSD	58.9	79.6	92.9	94.9	99.6	100.0	100.0	100.0	100.0
	AR(1)-CSD	62.0	81.0	93.7	95.6	99.7	100.0	100.0	100.0	100.0
Design 2 AR	AGGR	70.1	84.9	96.0	96.7	99.8	100.0	100.0	100.0	100.0
	WB	61.3	82.3	96.0	95.1	99.6	100.0	99.9	100.0	100.0
	AR (1)	48.9	74.2	93.3	90.1	99.3	100.0	99.8	100.0	100.0
	CSD	61.5	79.9	92.7	94.3	99.5	100.0	99.9	100.0	100.0
	AR(1)-CSD	50.0	71.6	87.5	90.3	99.2	100.0	99.9	100.0	100.0
Design 3 CSD	AGGR	68.4	84.3	94.4	96.3	99.6	100.0	100.0	100.0	100.0
	WB	64.7	83.9	95.0	95.2	99.5	100.0	100.0	100.0	100.0
	AR (1)	63.9	83.4	95.1	95.1	99.5	100.0	100.0	100.0	100.0
	CSD	46.1	66.3	83.1	91.7	99.0	99.9	99.9	100.0	100.0
	AR(1)-CSD	51.3	69.1	84.6	92.9	99.1	100.0	100.0	100.0	100.0
Design 4 AR + CSD	AGGR	73.3	85.0	94.3	96.9	99.7	100.0	100.0	100.0	100.0
	WB	65.4	82.6	94.4	94.9	99.5	100.0	100.0	100.0	100.0
	AR (1)	53.5	75.1	91.7	90.5	99.1	100.0	99.9	100.0	100.0
	CSD	48.8	66.5	83.8	91.3	99.0	100.0	99.9	100.0	100.0
	AR(1)-CSD	40.0	57.2	75.4	85.8	98.0	99.0	99.8	100.0	100.0

6 Conclusions

In this paper, we have proposed the bootstrap as an inference method on the number of common factors in two groups of data. We propose and theoretically justify under weak conditions a simple bootstrap test that avoids the need to estimate the bias and variance of the canonical correlations explicitly. We have verified these conditions in the case of the wild bootstrap under conditions similar to those in AGGR(2019). However, other approaches tailored to more general data generating processes are possible. Our simulation experiment shows that the bootstrap leads to rejection rates closer to the nominal level in all of the designs we considered compared to the asymptotic framework of AGGR(2019).

A Asymptotic theory

This Appendix is organized as follows. In Appendix A.1, we provide a set of primitive assumptions under which we derive the asymptotic distribution of $\hat{\xi}(k^c)$. Appendix A.2 contains auxiliary lemmas used to derive this limiting distribution. Appendix A.3 provides a proof of the results in Section 2.4. When describing our assumptions below, it is convenient to collect the vectors f_t^c, f_{1t}^s and f_{2t}^s into a vector $G_t = (f_t^{c\prime}, f_{1t}^{s\prime}, f_{2t}^{s\prime})'$, whose dimension is $k^c + k_1^s + k_2^s$.

A.1 Primitive assumptions

Assumption 1 We let $N, T \to \infty$ such that $\frac{\sqrt{T}}{N} \to 0$, and $\frac{N}{T^{3/2}} \to 0$, where $N = \min(N_1, N_2) = N_2$ and $\mu_N \equiv \sqrt{N_2/N_1} \to \mu \in [0, 1]$.

Assumption 2

(a) $E(G_t) = 0$ and $E\|G_t\|^4 \le M$ such that $\frac{1}{T}\sum_{t=1}^T G_t G_t' \to_p \Sigma_G > 0$, where Σ_G is a non-random positive definite matrix defined as

$$\Sigma_G \equiv \begin{pmatrix} I_{k^c} & 0 & 0 \\ 0 & I_{k_1^s} & \Phi \\ 0 & \Phi' & I_{k_2^s} \end{pmatrix}.$$

(b) For each $j = 1, 2$, the factor loadings matrix $\Lambda_j \equiv (\lambda_{j,1}, \cdots, \lambda_{j,N_j})'$ is deterministic such that $\|\lambda_{j,i}\| \le M$ and $\Sigma_{\Lambda,j} \equiv \lim_{N_j \to \infty} \Lambda_j' \Lambda_j / N_j > 0$ has distinct eigenvalues.

Assumption 3 For each $j = 1, 2$,

(a) $E(\varepsilon_{j,it}) = 0$, $E(|\varepsilon_{j,it}|^8) \le M$ for any i, t.

(b) $E(\varepsilon_{j,it}\varepsilon_{j,ls}) = \sigma_{j,il,ts}$, $|\sigma_{j,il,ts}| \le \bar{\sigma}_{j,il}$ for all (t, s) and $|\sigma_{j,il,ts}| \le \tau_{j,ts}$ for all (i, l) such that $\frac{1}{N_j}\sum_{i,l=1}^{N_j} \bar{\sigma}_{j,il} \le M$, $T^{-1}\sum_{t,s=1}^T \tau_{j,ts} \le M$, and $\frac{1}{N_j T}\sum_{t,s,i,l} |\sigma_{j,il,ts}| \le M$.

(c) $E\left|\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j} (\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it}))\right|^4 \le M$ for every (t, s).

Assumption 4 For each $j = 1, 2$,

(a) $E\left(\frac{1}{N_j}\sum_{i=1}^{N_j} \|\frac{1}{\sqrt{T}}\sum_{t=1}^T G_t \varepsilon_{j,it}\|^2\right) \le M$, where $E(G_t \varepsilon_{j,it}) = 0$ for all (i, t).

(b) For each s, $E\left\|\frac{1}{\sqrt{TN_j}}\sum_{t=1}^T \sum_{i=1}^{N_j} G_t(\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it}))\right\|^2 \le M$.

(c) $E\left\|\frac{1}{\sqrt{TN_j}}\sum_{t=1}^T G_t \varepsilon_{jt}' \Lambda_j\right\|^2 \le M$.

(d) $E\left(\frac{1}{T}\sum_{t=1}^T \left\|\frac{1}{\sqrt{N_j}}\Lambda_j' \varepsilon_{jt}\right\|^2\right) \le M$.

Assumptions 2-4 are standard in the factor literature. In particular, Assumptions 2-(a) and 2-(b) impose standard conditions on the factors and factor loadings, respectively. They are identical to Assumptions A.2 and A.3 of AGGR(2019). Assumption 3 imposes standard time and cross-section dependence and heteroskedasticity in the idiosyncratic errors of each panel and corresponds to Assumption 2 of GP(2014). Finally, Assumption 4 imposes conditions on moments and weak dependence among $\{G_t\}$ and $\{\varepsilon_{j,it}\}$. This assumption corresponds to Assumption 3(a)-(d) in GP(2014). Note that given Assumption 2-(b), which assumes the factor loadings to be deterministic, we can show that Assumption 4-(d) is implied by Assumptions 2-(b) and 3-(b). To see this, note that we can write

$$E\left\|\frac{1}{\sqrt{N_j}}\Lambda_j'\varepsilon_{jt}\right\|^2 = \frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{l=1}^{N_j}\lambda_{j,i}'\lambda_{j,l}\underbrace{E\left(\varepsilon_{j,it}\varepsilon_{j,lt}\right)}_{\leq\bar{\sigma}_{j,il}\text{ by Ass-3(b)}} \leq \frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{l=1}^{N_j}\left|\lambda_{j,i}'\lambda_{j,l}\right|\bar{\sigma}_{j,il} \leq M\frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{l=1}^{N_j}\bar{\sigma}_{j,il},$$

given that Assumption 2-(b) and Cauchy-Schwartz's inequality imply that we can bound $\left|\lambda_{j,i}'\lambda_{j,l}\right| = \left|\sum_{k=1}^{k_j}\lambda_{j,ik}\lambda_{j,lk}\right| \leq \left(\sum_{k=1}^{k_j}\lambda_{j,ik}^2\right)^{1/2}\left(\sum_{k=1}^{k_j}\lambda_{j,lk}^2\right)^{1/2} = \|\lambda_{j,i}\|\,\|\lambda_{j,l}\| \leq M$. Assumption 4-(d) then follows from Assumption 3-(b) which bounds $\frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{l=1}^{N_j}\bar{\sigma}_{j,il}$ by M for all t. The reason why we keep Assumption 4-(d) is that we will give its bootstrap analogue in Appendix B.1. Note also that, as stated in GP(2014), we can show that Assumptions 4-(a) and (c) are implied by Assumptions 2 and 3 if we assume that the factors and the idiosyncratic errors are mutually independent. Assumption 4-(b) in turn holds if we assume in addition that $T^{-2}N_j^{-1}\sum_{s,q=1}^{T}\sum_{i=1}^{N_j}\left|Cov\left(\varepsilon_{j,it}\varepsilon_{j,is},\varepsilon_{j,it}\varepsilon_{j,iq}\right)\right| \leq M$, which follows if $\varepsilon_{j,it}$ is i.i.d. and $E\left(\varepsilon_{j,it}^4\right) \leq M$.

A key step in deriving the asymptotic distribution of the AGGR(2019) test statistic (and of its bootstrap analogue) under our Assumption 1 is to obtain an asymptotic expansion of the factors estimation uncertainty (as characterized by $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_jf_{jt}\right)\left(\hat{f}_{kt}-H_kf_{kt}\right)'$ for $j,k\in\{1,2\}$ up to order [9] $O_p\left(\delta_{NT}^{-4}\right)$). See Lemma A.5 in Appendix A.2. As it turns out, Assumptions 1-4 are not sufficient to ensure this fast rate of convergence. For this reason, we strengthen Assumptions 1-4 as follows.

Assumption 5

(a) For each t and $j=1,2$, $\sum_{s=1}^{T}|\gamma_{j,st}| \leq M$, where $\gamma_{j,st} \equiv E(\frac{1}{N_j}\sum_{i=1}^{N_j}\varepsilon_{j,is}\varepsilon_{j,it})$ and $\sum_{l=1}^{N_j}\bar{\sigma}_{j,il} \leq M$.

(b) For any j,k, $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}f_{js}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}} = O_p(1)$.

(c) For any j,k, $\frac{1}{T}\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\right\|^2 = O_p(1)$.

(d) For any j,k, $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}f_{js}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{k,i}\varepsilon_{k,it}(\varepsilon_{j,is}\varepsilon_{j,it}-E(\varepsilon_{j,is}\varepsilon_{j,it}))\right) = O_p(1)$, where $N = \min(N_1,N_2)$.

[9] This means that it contains terms of order $O_p\left(\delta_{NT}^{-2}\right)$ and a remainder of order $O_p\left(\delta_{NT}^{-4}\right)$.

(e) For any j,k, $\frac{1}{\sqrt{T}}\sum_{s=1}^{T} f_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\lambda_{k,i_2}\varepsilon_{k,i_2t}(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t}-E(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t})) = O_p(1)$.

(f) For any j,k, $\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{k,i}\varepsilon_{k,it}(\varepsilon_{j,is}\varepsilon_{j,it}-E(\varepsilon_{j,is}\varepsilon_{j,it}))\right)\right\|^2 = O_p(1)$, where $N = \min(N_1, N_2)$.

(g) For any j,k, $\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\lambda_{k,i_2}\varepsilon_{k,i_2t}(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t}-E(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t}))\right)\right\|^2 = O_p(1)$.

Assumption 5-(a) is a strengthening of Assumption 3-(b) and corresponds to Assumption E.1 of Bai (2003). A similar assumption has been used by AGGR(2019). See in particular their Assumption A.7(c) on $\beta_{j,t}$. As explained by Bai (2003), this assumption is satisfied when we rule out serial dependence, implying that $\gamma_{j,st}=0$ for $s\neq t$. In this case, Assumption 5-(a) is equivalent to requiring that $\frac{1}{N_j}\sum_{i=1}^{N_j}E\left(\varepsilon_{j,it}^2\right)\leq M$. More generally, this condition holds whenever for each panel j and each series i, the autocovariance function of $\{\varepsilon_{j,it}\}$ is absolutely summable (thus covering all finite order stationary ARMA models).

To interpret Assumptions 5-(b) and (c), let $v_{kt}\equiv\frac{\Lambda_k'\varepsilon_{kt}}{\sqrt{N_k}}$ and $m_{jk,s}\equiv\sum_{t=1}^{T}\gamma_{j,st}v_{kt}$. With this notation, we can rewrite part (b) as $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}f_{js}m_{jk,s}'=O_p(1)$ and part (c) as $\frac{1}{T}\sum_{s=1}^{T}\|m_{jk,s}\|^2=O_p(1)$. The latter condition holds if $E\|m_{jk,s}\|^2\leq M$ for all j,k,s, which follows if part (a) holds and if we assume that $E\|v_{kt}\|^2\leq M$ for all k,t. To see this, note that $E\|m_{jk,s}\|^2 = E\left[\left(\sum_{t=1}^{T}\gamma_{j,st}v_{kt}'\right)\left(\sum_{l=1}^{T}\gamma_{j,sl}v_{kl}\right)\right] = \sum_{t=1}^{T}\sum_{l=1}^{T}\gamma_{j,st}\gamma_{j,sl}E\left(v_{kt}'v_{kl}\right)$, which is bounded by $\sum_{t=1}^{T}\sum_{l=1}^{T}|\gamma_{j,st}|\,|\gamma_{j,sl}|\left(E\|v_{kt}\|^2\right)^{1/2}\left(E\|v_{kl}\|^2\right)^{1/2}$ by Cauchy-Schwarz's inequality. If $E\|v_{kt}\|^2\leq M$ for all k,t, we can use Assumption 5-(a) to verify Assumption 5-(c). The assumption that $E\|v_{kt}\|^2\leq M$ for all k,t is a strengthening of Assumption 4-(d) and both are equivalent if we assume stationarity of $\{\varepsilon_{kt}\}$. Hence, Assumption 5-(c) holds under general serial and cross-sectional dependence in the idiosyncratic error terms.

A sufficient condition for Assumption 5-(b) is that $E\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^{T}f_{js}m_{jk,s}'\right\|^2\leq M$. We can show that this condition is implied by Assumptions 3-(b) and 5-(a) if we assume that $\{f_{js}\}$ and $\{\varepsilon_{kt}\}$ are mutually independent. We can verify Assumptions 5-(d) and (e) by showing that $\frac{1}{T}\sum_{s=1}^{T}\sum_{l=1}^{T}E(A_{jk,l}'A_{jk,s})\leq M$ and $\frac{1}{T}\sum_{s=1}^{T}\sum_{l=1}^{T}E(B_{jk,l}'B_{jk,s})\leq M$, where $A_{jk,s}\equiv\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{k,i}\varepsilon_{k,it}(\varepsilon_{j,is}\varepsilon_{j,it}-E(\varepsilon_{j,is}\varepsilon_{j,it}))\right)$ and $B_{jk,s}\equiv\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\lambda_{k,i_2}\varepsilon_{k,i_2t}(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t}-E(\varepsilon_{j,i_1s}\varepsilon_{j,i_1t}))$, which holds for instance if $\varepsilon_{j,it}$ is i.i.d. with $E(\varepsilon_{j,it}^3)=0$ and $E(\varepsilon_{j,it}^6)\leq M$ for $j=1,2$. Similarly, we can show that Assumptions 5-(d) and (e) are verified under similar conditions on $\varepsilon_{j,it}$.

Our next assumption is a high level condition that allows us to obtain the asymptotic normal distribution for the AGGR test statistic.

Assumption 6

(a) $\tilde{\Sigma}_{cc}\equiv\frac{1}{T}\sum_{t=1}^{T}f_t^c f_t^{c\prime}$ is such that $\tilde{\Sigma}_{cc}-I_{k^c}=O_p\left(T^{-1/2}\right)$.

(b) $\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\left(\mathcal{U}_t'\mathcal{U}_t - E\left(\mathcal{U}_t'\mathcal{U}_t\right)\right) \to^d N\left(0, \Omega_{\mathcal{U}}\right)$, where $\mathcal{U}_t \equiv \mu_N u_{1t}^{(c)} - u_{2t}^{(c)}$ and $u_{jt}^{(c)}$ is a $k^c \times 1$ vector containing the first k^c rows of $u_{jt} \equiv \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}$.

Assumption 6-(a) strengthens Assumption 2-(a) by requiring that $\frac{1}{T}\sum_{t=1}^{T} f_t^c f_t^{c\prime}$ converges to I_{k^c} at rate $O_p\left(T^{-1/2}\right)$. This assumption is implied by standard mixing conditions on f_t^c by a maximal inequality for mixing processes and has been used in this literature. See e.g., Gonçalves, McCracken, and Perron (2017). AGGR(2019) assume factors to be mixing, explaining why they do not explicitly write this assumption. It is used to omit $\tilde{\Sigma}_{cc}$ from the term $\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\left(\mathcal{U}_t'\mathcal{U}_t - E\left(\mathcal{U}_t'\mathcal{U}_t\right)\right)$ that appears in the asymptotic expansion of the test statistic. Assumption 6-(b) is a high level condition that requires the time series process $\mathcal{Z}_{N,t} \equiv \left(\mathcal{U}_t'\mathcal{U}_t - E\left(\mathcal{U}_t'\mathcal{U}_t\right)\right)$ to satisfy a CLT. AGGR(2019) provide conditions under which this high level condition holds. See in particular their Assumptions A.5 and A.6, which are used to show that $\mathcal{Z}_{N,t}$ is a NED process. Since our contribution is proving the bootstrap validity in this context, we do not provide these more primitive conditions. They are not required to prove our bootstrap theory.

Note that our assumptions (in particular, Assumptions 2-(b) and 4-(d)) imply that

$$\tilde{\Sigma}_{\mathcal{U}} \equiv T^{-1}\sum_{t=1}^{T} E\left(\mathcal{U}_t\mathcal{U}_t'\right) = \mu_N^2\tilde{\Sigma}_{\mathcal{U},11} + \tilde{\Sigma}_{\mathcal{U},22} - \mu_N\tilde{\Sigma}_{\mathcal{U},12} - \mu_N\tilde{\Sigma}_{\mathcal{U},21}, \text{ with } \tilde{\Sigma}_{\mathcal{U},jk} \equiv T^{-1}\sum_{t=1}^{T} E\left(u_{jt}^{(c)}u_{kt}^{(c)\prime}\right),$$

is $O(1)$. This term enters the bias $\mathcal{B} \equiv tr\left(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}}\right)$ that appears in the asymptotic distribution of the test statistic.

A.2 Asymptotic expansion of the sample covariance of the factors estimation error

The main goal of this section is to provide an asymptotic expansion of $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_j f_{jt}\right)\left(\hat{f}_{kt} - H_k f_{kt}\right)'$ for $j, k \in \{1, 2\}$ up to order $O_p\left(\delta_{NT}^{-4}\right)$, which is then used to characterize the bias term. See Lemma A.5 in Appendix A.3.

To derive this result, we use the following identity for each group j, which follows from Bai (2003):

$$\hat{f}_{jt} - H_j f_{jt} = \mathcal{V}_j^{-1}\left(A_{j,1t} + A_{j,2t} + A_{j,3t} + A_{j,4t}\right). \tag{8}$$

Each of the terms $A_{j,1t}$ through $A_{j,4t}$ is defined as follows:

$$\begin{aligned}
A_{j,1t} &= \frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\gamma_{j,st}, \text{ with } \gamma_{j,st} = E(\frac{1}{N_j}\sum_{i=1}^{N_j}\varepsilon_{j,is}\varepsilon_{j,it});\\
A_{j,2t} &= \frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\zeta_{j,st}, \text{ with } \zeta_{j,st} = \frac{1}{N_j}\sum_{i=1}^{N_j}(\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it}));\\
A_{j,3t} &= \frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\eta_{j,st}, \text{ with } \eta_{j,st} = \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}' f_{js}\varepsilon_{j,it} = f_{js}'\frac{\Lambda_j'\varepsilon_{jt}}{N_j}; \text{ and}\\
A_{j,4t} &= \frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\xi_{j,st}, \text{ with } \xi_{j,st} = f_{jt}'\frac{\Lambda_j'\varepsilon_{js}}{N_j} = \eta_{j,ts}.
\end{aligned}$$

The following auxiliary lemma is used to prove Lemma A.2.

Lemma A.1 *Suppose Assumptions 1-4 strengthened by Assumption 5 hold. Then, for any* $j,k \in \{1,2\}$*: (a)* $\frac{1}{T}\sum_{t=1}^{T}A_{j,1t}A_{k,1t}' = O_p\left(\delta_{NT}^{-4}\right)$*; (b)* $\frac{1}{T}\sum_{t=1}^{T}A_{j,2t}A_{k,2t}' = O_p\left(\delta_{NT}^{-4}\right)$*; (c)* $\frac{1}{T}\sum_{t=1}^{T}A_{j,4t}A_{k,4t}' = O_p\left(\delta_{NT}^{-4}\right)$*; (d)* $\frac{1}{T}\sum_{t=1}^{T}A_{j,mt}A_{k,nt}' = O_p\left(\delta_{NT}^{-4}\right)$ *for* $m \neq n$*, where* $m,n \in \{1,2,3,4\}$*; and (e)*

$$\frac{1}{T}\sum_{t=1}^{T}A_{j,3t}A_{k,3t}' = \frac{1}{\sqrt{N_jN_k}}\mathcal{V}_jH_j\frac{1}{T}\sum_{t=1}^{T}u_{jt}u_{kt}'H_k'\mathcal{V}_k' = O_p\left(N^{-1}\right), \quad \textit{where} \quad u_{jt} \equiv \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}.$$

Lemma A.2 *Suppose Assumptions 1-4 strengthened by Assumption 5 hold. Then, for* $j,k \in \{1,2\}$,

$$\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_jf_{jt}\right)\left(\hat{f}_{kt} - H_kf_{kt}\right)' = \frac{1}{\sqrt{N_jN_k}}H_j\left(\frac{1}{T}\sum_{t=1}^{T}u_{jt}u_{kt}'\right)H_k' + O_p\left(\delta_{NT}^{-4}\right),$$

where u_{jt} *is as defined in Lemma A.1.*

Proof of Lemma A.1. Part (a): We can bound the norm of $\frac{1}{T}\sum_{t=1}^{T}A_{j,1t}A_{k,1t}'$ by

$$\frac{1}{T}\sum_{t=1}^{T}\left\|A_{j,1t}A_{k,1t}'\right\| = \frac{1}{T}\sum_{t=1}^{T}\left\|A_{j,1t}\right\|\left\|A_{k,1t}\right\| \leq \left(\frac{1}{T}\sum_{t=1}^{T}\left\|A_{j,1t}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|A_{k,1t}\right\|^2\right)^{1/2},$$

where the first equality follows by the fact that for any vectors A and B we have that $\|AB'\|^2 = tr\left(AB'BA'\right) = tr\left(A'AB'B\right) = \|A\|^2\|B\|^2$ given the definitions of the Frobenius norm of a matrix and of the Euclidean norm of a vector. The inequality then follows by the Cauchy-Schwartz inequality. Next, we show that $\frac{1}{T}\sum_{t=1}^{T}\|A_{j,1t}\|^2 = O_p\left(\delta_{NT}^{-4}\right)$ for any j, which implies the result. To show this, write $A_{j,1t} = A_{j,1t}^{(1)} + A_{j,1t}^{(2)}$, where

$$A_{j,1t}^{(1)} \equiv \frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}_{js} - H_jf_{js}\right)\gamma_{j,st} \quad\text{and}\quad A_{j,1t}^{(2)} \equiv H_j\frac{1}{T}\sum_{s=1}^{T}f_{js}\gamma_{j,st}.$$

Since $\left\|A_{j,1t}^{(1)} + A_{j,1t}^{(2)}\right\|^2 \leq 2\left(\left\|A_{j,1t}^{(1)}\right\|^2 + \left\|A_{j,1t}^{(2)}\right\|^2\right)$, we have that $\frac{1}{T}\sum_{t=1}^{T}\|A_{j,1t}\|^2 \leq 2\frac{1}{T}\sum_{t=1}^{T}\left\|A_{j,1t}^{(1)}\right\|^2 + 2\frac{1}{T}\sum_{t=1}^{T}\left\|A_{j,1t}^{(2)}\right\|^2 \equiv 2I_1 + 2II_1$. We analyse each term separately. First, by an application of the triangle inequality and the Cauchy-Schwartz inequality,

$$\|I_1\| \leq \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}-H_jf_{js})\gamma_{j,st}\right\|^2 \leq \frac{1}{T}\sum_{s=1}^{T}\|\hat{f}_{js}-H_jf_{js}\|^2\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma_{j,st}|^2 = O_p(\delta_{NT}^{-2}T^{-1}) = O_p(\delta_{NT}^{-4}),$$

since $T^{-1}\sum_{s=1}^{T}\|\hat{f}_{js}-H_jf_{js}\|^2 = O_p\left(\delta_{NT}^{-2}\right)$ and $\sum_{s=1}^{T}|\gamma_{j,st}|^2 = O\left(1\right)$ given Assumptions 1-5. Similarly, we can show that $\|II_1\| = O_p\left(T^{-2}\right) = O_p\left(\delta_{NT}^{-4}\right)$ by using Markov's inequality and noting that

$$\frac{1}{T}\sum_{t=1}^{T}E\left\|\frac{1}{T}\sum_{s=1}^{T}f_{js}\gamma_{j,st}\right\|^2 \leq \frac{1}{T}\sum_{t=1}^{T}E\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}f_{jl}'f_{js}\gamma_{j,st}\gamma_{j,lt}\right),$$

where

$$E\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}f_{jl}'f_{js}\gamma_{j,st}\gamma_{j,lt}\right) = \frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}\underbrace{E\left(f_{jl}'f_{js}\right)}_{\leq\Delta\text{ by Ass-2(a)}}\gamma_{j,st}\gamma_{j,lt} \leq \Delta\frac{1}{T^2}\underbrace{[\sum_{s=1}^{T}|\gamma_{j,st}|]^2}_{\leq M^2\text{ by Ass-5(a)}} \leq C\frac{1}{T^2}.$$

Note that to obtain this last bound, we impose Assumption 5-(c), which is a strengthening of Assumption 3.

Part (b): we proceed as in part (a) and show that $T^{-1}\sum_{t=1}^{T}\|A_{j,2t}\|^2 = O_p(\delta_{NT}^{-4})$ for any $j\in\{1,2\}$. Adding and subtracting appropriately, $T^{-1}\sum_{t=1}^{T}\|A_{j,2t}\|^2 \leq 2T^{-1}\sum_{t=1}^{T}\|A_{j,2t}^{(1)}\|^2+2T^{-1}\sum_{t=1}^{T}\|A_{j,2t}^{(2)}\|^2 \equiv 2I_2+2II_2$, where $A_{j,2t}^{(1)} \equiv T^{-1}\sum_{s=1}^{T}(\hat{f}_{js}-H_jf_{js})\zeta_{j,st}$ and $A_{j,2t}^{(2)} \equiv T^{-1}\sum_{s=1}^{T}H_jf_{js}\zeta_{j,st}$, with $\zeta_{j,st} \equiv N_j^{-1}\sum_{i=1}^{N_j}(\varepsilon_{j,is}\varepsilon_{j,it}-E(\varepsilon_{j,is}\varepsilon_{j,it}))$. First, note that

$$I_2 \leq \left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{f}_{js}-H_jf_{js}\|^2\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\zeta_{j,st}|^2\right) = O_p(\delta_{NT}^{-2}N_j^{-1}) = O_p(\delta_{NT}^{-4}),$$

since $\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\zeta_{j,st}|^2 = O_p(N_j^{-1})$ by Assumption 3-(c). Second, by Assumption 4-(b),

$$II_2 \leq \|H_j\|^2\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}f_{js}\zeta_{j,st}\right\|^2 = O_p((TN_j)^{-1}) = O_p(\delta_{NT}^{-4}).$$

Part (c): Following the same arguments as above, the result follows by showing that $T^{-1}\sum_{t=1}^{T}\|A_{j,4t}\|^2 = O_p(\delta_{NT}^{-4})$ for any $j\in\{1,2\}$. Adding and subtracting appropriately, we can write $A_{j,4t} = A_{j,4t}^{(1)}+A_{j,4t}^{(2)}$, where $A_{j,4t}^{(1)} \equiv \frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}-H_jf_{js})\xi_{j,st}$ and $A_{j,4t}^{(2)} \equiv \frac{1}{T}\sum_{s=1}^{T}H_jf_{js}\xi_{j,st}$, with $\xi_{j,st} \equiv f_{jt}'\frac{\Lambda_j'\varepsilon_{js}}{N_j}$. We show that $I_4 \equiv T^{-1}\sum_{t=1}^{T}\left\|A_{j,4t}^{(1)}\right\|^2$ and $II_4 \equiv T^{-1}\sum_{t=1}^{T}\left\|A_{j,4t}^{(2)}\right\|^2$ are both $O_p\left(\delta_{NT}^{-4}\right)$ under our assumptions. For the first term, using the definition of $\xi_{j,st}$, we have that

$$I_4 = \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}-H_jf_{js})\frac{\varepsilon_{js}'\Lambda_j}{N_j}f_{jt}\right\|^2 \leq \underbrace{\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}-H_jf_{js})\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2}_{=O_p\left(\delta_{NT}^{-2}\right)O_p\left(N_j^{-1}\right)}\underbrace{\frac{1}{T}\sum_{t=1}^{T}\|f_{jt}\|^2}_{=O_p(1)\text{ by Ass-2(a)}} = O_p(\delta_{NT}^{-4}),$$

since $E\|f_{jt}\|^2 \leq \Delta$, and by Cauchy-Schwartz's inequality,

$$\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}-H_j f_{js})\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2 \leq \underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js}-H_j f_{js}\right\|^2}_{=O_p\left(\delta_{NT}^{-2}\right)}\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2}_{=O_p\left(N_j^{-1}\right)} = O_p\left(\delta_{NT}^{-4}\right),$$

given that $\frac{1}{T}\sum_{s=1}^{T}\|\frac{\Lambda_j'\varepsilon_{js}}{\sqrt{N_j}}\|^2 = O_p(1)$ under Assumption 4-(d). For II_4, using the definition of $\xi_{j,st} \equiv \frac{\varepsilon_{js}'\Lambda_j}{N_j}f_{jt}$, we have that

$$\begin{aligned} II_4 &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}H_j f_{js}\frac{\varepsilon_{js}'\Lambda_j}{N_j}f_{jt}'\right\|^2 \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}H_j f_{js}\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2\|f_{jt}\|^2 \\ &\leq \|H_j\|^2\underbrace{\left\|\frac{1}{T}\sum_{s=1}^{T}f_{js}\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2}_{\substack{=O_p\left((N_jT)^{-1}\right)\\ \text{by Ass-4(c)}}}\underbrace{\frac{1}{T}\sum_{t=1}^{T}\|f_{jt}\|^2}_{\substack{=O_p(1)\\ \text{by Ass-2(a)}}} = O_p\left(\delta_{NT}^{-4}\right). \end{aligned}$$

Part (d): Given parts (a), (b), and (c), all the cross terms that involve $A_{j,1t}$, $A_{j,2t}$ and $A_{j,4t}$ are $O_p\left(\delta_{NT}^{-4}\right)$ by an application of Cauchy-Schwartz's inequality. Hence, we only need to show that $T^{-1}\sum_{t=1}^{T}A_{j,mt}A_{k,3t}'$ is $O_p\left(\delta_{NT}^{-4}\right)$ for $m = 1,2,4$. Using the definition of $A_{k,3t}$, we have that

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{j,mt}A_{k,3t}' &= T^{-1}\sum_{t=1}^{T}A_{j,mt}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}\eta_{k,st}\right)', \text{ where } \eta_{k,st} \equiv f_{ks}'\frac{\Lambda_k'\varepsilon_{kt}}{N_k} \\ &= T^{-1}\sum_{t=1}^{T}A_{j,mt}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}f_{ks}'\frac{\Lambda_k'\varepsilon_{kt}}{N_k}\right)' \\ &= \left[T^{-1}\sum_{t=1}^{T}A_{j,mt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right]\underbrace{\frac{F_k'\hat{F}_k}{T}}_{=O_p(1)}, \end{aligned}$$

implying that it suffices to show that $T^{-1}\sum_{t=1}^{T}A_{j,mt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = O_p\left(\delta_{NT}^{-4}\right)$. To show this, an application of Cauchy Schwartz's inequality is not enough because $T^{-1}\sum_{t=1}^{T}\left\|\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2 = O_p\left(N_k^{-1}\right) = O_p\left(\delta_{NT}^{-2}\right)$. Hence, using the fact that $T^{-1}\sum_{t=1}^{T}\|A_{j,mt}\|^2 = O_p\left(\delta_{NT}^{-4}\right)$ for $m \neq 3$ implies by Cauchy-Schwartz's inequality that the term in square bracket is $O_p\left(\delta_{NT}^{-3}\right)$, which is larger than $O_p\left(\delta_{NT}^{-4}\right)$. We need a more refined analysis, which in turn requires a strengthening of Assumptions 1-4 as given by Assumption 5. Starting with $m = 1$, by the definition of $A_{j,1t}$, we have that

$$\frac{1}{T}\sum_{t=1}^{T}A_{j,1t}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = \frac{1}{T}\sum_{t=1}^{T}A_{j,1t}^{(1)}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} + \frac{1}{T}\sum_{t=1}^{T}A_{j,1t}^{(2)}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} \equiv (a_1) + (b_1),$$

where

$$(a_1) = T^{-1}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}_{js} - H_j f_{js}\right)\gamma_{j,st}\right)\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = \frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}_{js} - H_j f_{js}\right)\left(\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right),$$

and

$$(b_1) = H_j\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T} f_{js}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = H_j\frac{1}{T}\sum_{s=1}^{T} f_{js}\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}.$$

Note that we can rewrite (b_1) as

$$(b_1) = H_j\frac{1}{T}\frac{1}{\sqrt{TN_k}}\underbrace{\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T} f_{js}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\right]}_{=O_p(1)\text{ by Ass-5(b)}} = O_p\left(\delta_{NT}^{-4}\right),$$

if we assume that the term in the square bracket is $O_p(1)$. We impose this as a new assumption, cf. Assumption 5-(b). In addition,

$$\|(a_1)\| \leq \left(\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js} - H_j f_{js}\right\|^2}_{=O_p\left(\delta_{NT}^{-2}\right)}\right)^{1/2}\left(\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2}_{=O_p\left(\delta_{NT}^{-6}\right)}\right)^{1/2},$$

where

$$\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2 = \frac{1}{N_k}\frac{1}{T^2}\underbrace{\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\gamma_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\right\|^2\right]}_{=O_p(1)\text{ by Ass-5(c)}} = O_p\left(\delta_{NT}^{-6}\right),$$

provided we assume that the term in square bracket is $O_p(1)$. We impose this as a new assumption, cf. Assumption 5-(c). Consider next $m = 2$. Using using the decomposition of $A_{j,2t} = A_{j,2t}^{(1)} + A_{j,2t}^{(2)}$, we can write

$$\frac{1}{T}\sum_{t=1}^{T} A_{j,2t}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = \frac{1}{T}\sum_{t=1}^{T} A_{j,2t}^{(1)}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} + \frac{1}{T}\sum_{t=1}^{T} A_{j,2t}^{(2)}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} \equiv (a_2) + (b_2),$$

where

$$(a_2) = T^{-1}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}_{js} - H_j f_{js}\right)\zeta_{j,st}\right)\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = \frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}_{js} - H_j f_{js}\right)\left(\frac{1}{T}\sum_{t=1}^{T}\zeta_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right),$$

and

$$(b_2) = H_j\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T} f_{js}\zeta_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} = H_j\frac{1}{T}\sum_{s=1}^{T} f_{js}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}(\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it}))\right)\frac{\varepsilon_{kt}'\Lambda_k}{N_k}.$$

Note that

$$(b_2) = \frac{1}{\sqrt{T}} \frac{\sqrt{N}}{N_j N_k} \underbrace{\left[\frac{1}{\sqrt{T}} \sum_{s=1}^{T} f_{js} \frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \lambda_{k,i} \varepsilon_{k,it} (\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it})) \right) \right]}_{=O_p(1) \text{ by Ass-5(d)}}$$

$$+ \frac{1}{T} \frac{1}{\sqrt{N_j N_k}} \underbrace{\left[\frac{1}{\sqrt{T}} \sum_{s=1}^{T} f_{js} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{1}{\sqrt{N_j N_k}} \sum_{i_1=1}^{N_j} \sum_{i_2 \neq i_1}^{N_k} \lambda_{k,i_2} \varepsilon_{k,i_2 t} (\varepsilon_{j,i_1 s}\varepsilon_{j,i_1 t} - E(\varepsilon_{j,i_1 s}\varepsilon_{j,i_1 t})) \right]}_{=O_p(1) \text{ by Ass-5(e)}} = O_p\left(\delta_{NT}^{-4}\right).$$

By Cauchy-Schwartz's inequality, we can bound (a_2) by

$$\underbrace{\left(\frac{1}{T} \sum_{s=1}^{T} \left\| \hat{f}_{js} - H_j f_{js} \right\|^2 \right)^{1/2}}_{=O_p\left(\delta_{NT}^{-1}\right)} \left(\underbrace{\frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{T} \sum_{t=1}^{T} \zeta_{j,st} \frac{\varepsilon_{kt}' \Lambda_k}{N_k} \right\|^2}_{=(a2-ii)} \right)^{1/2} = O_p\left(\delta_{NT}^{-1}\right) O_p\left(\delta_{NT}^{-3}\right),$$

where we show that $(a2-ii) = O_p(\delta_{NT}^{-6})$ by noting that

$$(a2-ii) = \underbrace{\frac{N}{N_j^2 N_k^2}}_{=O\left(\delta_{NT}^{-6}\right)} \underbrace{\left[\frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \lambda_{k,i} \varepsilon_{k,it} (\varepsilon_{j,is}\varepsilon_{j,it} - E(\varepsilon_{j,is}\varepsilon_{j,it})) \right) \right\|^2 \right]}_{=O_p(1) \text{ by Ass-5(f)}}$$

$$+ \underbrace{\frac{1}{N_j N_k} \frac{1}{T}}_{=O\left(\delta_{NT}^{-6}\right)} \underbrace{\left[\frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{N_j N_k}} \sum_{i_1=1}^{N_j} \sum_{i_2 \neq i_1}^{N_k} \lambda_{k,i_2} \varepsilon_{k,i_2 t} (\varepsilon_{j,i_1 s}\varepsilon_{j,i_1 t} - E(\varepsilon_{j,i_1 s}\varepsilon_{j,i_1 t})) \right) \right\|^2 \right]}_{=O_p(1) \text{ by Ass-5(g)}}.$$

Finally, consider $m = 4$. Using the decomposition of $A_{j,4t} = A_{j,4t}^{(1)} + A_{j,4t}^{(2)}$, we can write

$$\frac{1}{T} \sum_{t=1}^{T} A_{j,4t} \frac{\varepsilon_{kt}' \Lambda_k}{N_k} = \frac{1}{T} \sum_{t=1}^{T} A_{j,4t}^{(1)} \frac{\varepsilon_{kt}' \Lambda_k}{N_k} + \frac{1}{T} \sum_{t=1}^{T} A_{j,4t}^{(2)} \frac{\varepsilon_{kt}' \Lambda_k}{N_k} \equiv (a_4) + (b_4),$$

where

$$\begin{aligned} (a_4) &= T^{-1} \sum_{t=1}^{T} \left(\frac{1}{T} \sum_{s=1}^{T} \left(\hat{f}_{js} - H_j f_{js} \right) \xi_{j,st} \right) \frac{\varepsilon_{kt}' \Lambda_k}{N_k} = \frac{1}{T} \sum_{s=1}^{T} \left(\hat{f}_{js} - H_j f_{js} \right) \left(\frac{1}{T} \sum_{t=1}^{T} \xi_{j,st} \frac{\varepsilon_{kt}' \Lambda_k}{N_k} \right), \text{ and} \\ (b_4) &= H_j \frac{1}{T} \sum_{t=1}^{T} \frac{1}{T} \sum_{s=1}^{T} f_{js} \xi_{j,st} \frac{\varepsilon_{kt}' \Lambda_k}{N_k}. \end{aligned}$$

Note that

$$\begin{aligned}
(b_4) &= H_j \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T} f_{js}\xi_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k} \\
&= H_j \frac{1}{T}\sum_{s=1}^{T} f_{js}\frac{1}{T}\sum_{t=1}^{T}\left(f_{jt}'\frac{\Lambda_j'\varepsilon_{js}}{N_j}\right)\frac{\varepsilon_{kt}'\Lambda_k}{N_k} \\
&= H_j \underbrace{\left[\frac{1}{T}\sum_{s=1}^{T} f_{js}\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right]}_{=O_p\left(\frac{1}{\sqrt{TN_j}}\right)} \underbrace{\left[\frac{1}{T}\sum_{t=1}^{T} f_{jt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right]}_{=O_p\left(\frac{1}{\sqrt{TN_k}}\right)\text{by Ass-4(c)}} = O_p\left(\frac{1}{T\sqrt{N_jN_k}}\right) = O_p\left(\delta_{NT}^{-4}\right).
\end{aligned}$$

In addition,

$$\|(a_4)\| \le \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js} - H_j f_{js}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\xi_{j,st}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2\right)^{1/2},$$

where

$$\begin{aligned}
\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\varepsilon_{js}'\Lambda_j}{N_j}f_{jt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2 &= \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon_{js}'\Lambda_j}{N_j}\frac{1}{T}\sum_{t=1}^{T}f_{jt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2 \\
&\le \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon_{js}'\Lambda_j}{N_j}\right\|^2\left\|\frac{1}{T}\sum_{t=1}^{T}f_{jt}\frac{\varepsilon_{kt}'\Lambda_k}{N_k}\right\|^2 \\
&= \frac{1}{N_k}\frac{1}{TN_j}\underbrace{\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon_{js}'\Lambda_j}{\sqrt{N_j}}\right\|^2\right)}_{=O_p(1)\text{ by Ass-4(d)}}\underbrace{\left\|\frac{1}{\sqrt{TN_k}}\sum_{t=1}^{T}f_{jt}\varepsilon_{kt}'\Lambda_k\right\|^2}_{=O_p(1)\text{ by Ass-4(c)}} \\
&= O_p\left(\frac{1}{TN_jN_k}\right) = O_p\left(\delta_{NT}^{-6}\right),
\end{aligned}$$

implying that $\|(a_4)\| = O_p\left(\delta_{NT}^{-1}\right)O_p\left(\delta_{NT}^{-3}\right) = O_p\left(\delta_{NT}^{-4}\right)$.

Part (e): By definition, $A_{j,3t} \equiv \frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\eta_{j,st}$, where $\eta_{j,st} \equiv f_{js}'\frac{\Lambda_j'\varepsilon_{jt}}{N_j}$. Using the definition of the rotation matrix, $H_j \equiv \mathcal{V}_j^{-1}\frac{\hat{F}_j'F_j}{T}\frac{\Lambda_j'\Lambda_j}{N_j}$, we can rewrite this term as

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}A_{j,3t}A_{k,3t}' &= \frac{1}{\sqrt{N_jN_k}}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}f_{js}'\right)\frac{1}{T}\sum_{t=1}^{T}\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\left(\frac{1}{T}\sum_{l=1}^{T}f_{kl}\hat{f}_{kl}'\right) \\
&= \frac{1}{\sqrt{N_jN_k}}\left(\frac{\hat{F}_j'F_j}{T}\right)\frac{1}{T}\sum_{t=1}^{T}\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\left(\frac{\hat{F}_k'F_k}{T}\right)' \\
&= \frac{1}{\sqrt{N_jN_k}}\mathcal{V}_jH_j\left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}\frac{\varepsilon_{kt}'\Lambda_k}{\sqrt{N_k}}\left(\frac{\Lambda_k'\Lambda_k}{N_k}\right)^{-1}H_k'\mathcal{V}_k' = O_p(N^{-1}),
\end{aligned}$$

by Assumption 2-(b) and Assumption 4-(d). ■

Proof of Lemma A.2. Using Bai (2003)'s identity to express $\hat{f}_{jt} - H_jf_{jt} = \mathcal{V}_j^{-1}\left(A_{j,1t} + A_{j,2t} + A_{j,3t} + A_{j,4t}\right)$, we can write

$$
\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_j f_{jt}\right)\left(\hat{f}_{kt}-H_k f_{kt}\right)' \\
= & \mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(A_{j,1t}+A_{j,2t}+A_{j,3t}+A_{j,4t}\right)\left(A_{k,1t}+A_{k,2t}+A_{k,3t}+A_{k,4t}\right)'\mathcal{V}_k^{-1} \\
= & \mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}A_{j,1t}A_{k,1t}'\mathcal{V}_k^{-1}+\mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}A_{j,2t}A_{k,2t}'\mathcal{V}_k^{-1}+\mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}A_{j,3t}A_{k,3t}'\mathcal{V}_k^{-1} \\
& +\mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}A_{j,4t}A_{k,4t}'\mathcal{V}_k^{-1}+\text{cross terms.}
\end{aligned}
$$

Given Lemma A.1, the dominant term is the third term. All other terms are $O_p\left(\delta_{NT}^{-4}\right)$ under our assumptions, given Lemma A.1. This implies that

$$
\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_j f_{jt}\right)\left(\hat{f}_{kt}-H_k f_{kt}\right)' \\
= & \mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}A_{j,3t}A_{k,3t}'\mathcal{V}_k^{-1}+O_p\left(\delta_{NT}^{-4}\right) \\
= & \mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}\eta_{j,st}\right)\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}\eta_{k,st}\right)'\mathcal{V}_k^{-1}+O_p\left(\delta_{NT}^{-4}\right) \\
= & \mathcal{V}_j^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{js}f_{js}'\frac{\Lambda_j'\varepsilon_{jt}}{N_j}\right)\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}f_{ks}'\frac{\Lambda_k'\varepsilon_{kt}}{N_k}\right)'\mathcal{V}_k^{-1}+O_p\left(\delta_{NT}^{-4}\right) \\
= & \mathcal{V}_j^{-1}\frac{\hat{F}_j'F_j}{T}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\Lambda_j'\varepsilon_{jt}}{N_j}\right)\left(\frac{\Lambda_k'\varepsilon_{kt}}{N_k}\right)'\frac{F_k'\hat{F}_k}{T}\mathcal{V}_k^{-1}+O_p\left(\delta_{NT}^{-4}\right) \\
= & \frac{1}{\sqrt{N_jN_k}}H_j\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\left(\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}\right)\left[\left(\frac{\Lambda_k'\Lambda_j}{N_k}\right)^{-1}\left(\frac{\Lambda_k'\varepsilon_{kt}}{\sqrt{N_k}}\right)\right]'H_k'+O_p\left(\delta_{NT}^{-4}\right) \\
= & \frac{1}{\sqrt{N_jN_k}}H_j\left(\frac{1}{T}\sum_{t=1}^{T}u_{jt}u_{kt}'\right)H_k'+O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

completing the proof . ■

A.3 Proof of Theorem 2.1

Following AGGR(2019), we define $\hat{R}=\hat{V}_{11}^{-1}\hat{V}_{12}\hat{V}_{22}^{-1}\hat{V}_{21}$, where $\hat{V}_{jk}=\frac{1}{T}\sum_{t=1}^{T}\hat{f}_{jt}\hat{f}_{kt}'$. The test statistic is given by $\hat{\xi}\left(k^c\right)\equiv\sum_{l=1}^{k^c}\hat{\rho}_l=tr\left(\hat{\Lambda}^{1/2}\right)$, where $\hat{\Lambda}=diag\left(\hat{\rho}_l^2:l=1,\ldots,k^c\right)$ is a $k^c\times k^c$ diagonal matrix containing the k^c largest eigenvalues of $\hat{R}$ obtained from the eigenvalue-eigenvector problem $\hat{R}\hat{W}=\hat{W}\hat{\Lambda}$, where $\hat{W}$ is the $k_1\times k^c$ eigenvector matrix. The main idea of the proof is to obtain an expansion of $\hat{R}$ through order[10] $O_p\left(\delta_{NT}^{-2}\right)$, where $\delta_{NT}=\min\left(\sqrt{N},\sqrt{T}\right)$, from which we obtain an asymptotic expansion of $\hat{\Lambda}$ and of $tr\left(\hat{\Lambda}^{1/2}\right)$.

The asymptotic expansion of $\hat{R}$ is based on expanding $\hat{V}_{jk}$ around $\tilde{V}_{jk}\equiv\frac{1}{T}\sum_{t=1}^{T}f_{jt}f_{kt}'$ and using the fact that under the null hypothesis f_{jt} and f_{kt} share a set of common factors f_t^c (i.e. $f_{jt}=\left(f_t^{c\prime},f_{jt}^{s\prime}\right)'$ for $j=1,2$). Adding and subtracting appropriately yields

$$\begin{aligned}\hat{V}_{jk} &= \frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_jf_{jt}+H_jf_{jt}\right)\left(\hat{f}_{kt}-H_kf_{kt}+H_kf_{kt}\right)'\\ &= H_j\frac{1}{T}\sum_{t=1}^{T}f_{jt}f_{kt}'H_k'+\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_jf_{jt}\right)\left(\hat{f}_{kt}-H_kf_{kt}\right)'+\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt}-H_jf_{jt}\right)f_{kt}'H_k'\\ &\quad +H_j\frac{1}{T}\sum_{t=1}^{T}f_{jt}\left(\hat{f}_{kt}-H_kf_{kt}\right)'\\ &\equiv \ddot{V}_{jk}+\ddot{X}_{jk},\end{aligned}$$

with $\ddot{V}_{jk}\equiv H_j\tilde{V}_{jk}H_k'$, $\tilde{V}_{jk}\equiv\frac{1}{T}\sum_{t=1}^{T}f_{jt}f_{kt}'$, and $\ddot{X}_{jk}=H_j\hat{X}_{jk}H_k'$, where letting $\psi_{jt}\equiv H_j^{-1}\left(\hat{f}_{jt}-H_jf_{jt}\right)$,

$$\hat{X}_{jk}\equiv\frac{1}{T}\sum_{t=1}^{T}\psi_{jt}\psi_{kt}'+\frac{1}{T}\sum_{t=1}^{T}\psi_{jt}f_{kt}'+\frac{1}{T}\sum_{t=1}^{T}f_{jt}\psi_{kt}'.$$

We can show that $\hat{X}_{jk}=O_p\left(\delta_{NT}^{-2}\right)$ under Assumptions 1-4 (see Lemma A.3(a) below). Using this result, we can show that $\hat{R}=\ddot{R}+O_p\left(\delta_{NT}^{-2}\right)$, where $\ddot{R}=\ddot{V}_{11}^{-1}\ddot{V}_{12}\ddot{V}_{22}^{-1}\ddot{V}_{21}=\left(H_1'\right)^{-1}\tilde{R}H_1'$, where $\tilde{R}\equiv\tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21}$. The following auxiliary lemma states this result and characterizes the term of order $O_p\left(\delta_{NT}^{-2}\right)$ under Assumptions 1-4. Note that for this result we do not need Assumptions 5 and 6. Nor do we need to impose the null hypothesis of k^c common factors between the two panels.

Lemma A.3 *Let Assumptions 1-4 hold. Then, (a)* $\ddot{X}_{jk}=O_p\left(\delta_{NT}^{-2}\right)$ *and* $\hat{X}_{jk}=O_p\left(\delta_{NT}^{-2}\right)$*; and (b)* $\hat{R}=\left(H_1'\right)^{-1}\left[\tilde{R}+\tilde{V}_{11}^{-1}\hat{\Psi}\right]H_1'+O_p\left(\delta_{NT}^{-4}\right)$, *where*

$$\hat{\Psi}\equiv-\hat{X}_{11}\tilde{R}+\hat{X}_{12}\tilde{B}+\tilde{B}'\hat{X}_{21}-\tilde{B}'\hat{X}_{22}\tilde{B},\qquad \tilde{B}\equiv\tilde{V}_{22}^{-1}\tilde{V}_{21},\ \ \textit{and}$$

$$\hat{X}_{jk}\equiv\frac{1}{T}\sum_{t=1}^{T}\psi_{jt}\psi_{kt}'+\frac{1}{T}\sum_{t=1}^{T}\psi_{jt}f_{kt}'+\frac{1}{T}\sum_{t=1}^{T}f_{jt}\psi_{kt}',\quad \textit{where}\quad \psi_{jt}\equiv H_j^{-1}\left(\hat{f}_{jt}-H_jf_{jt}\right).$$

Remark 1 *Lemma A.3(a) is the analogue of Lemma B.1 of AGGR(2019). Contrary to AGGR(2019), we rely on Bai (2003)'s asymptotic expansion for* $\hat{f}_{jt}-H_jf_{jt}$*, which explains why our set of assumptions is different from those of AGGR(2019). Lemma A.3(b) is the analogue of Lemma B.2 of AGGR(2019) under our Assumptions 1-4. Note that the order of magnitude of the remainder term follows from expressing* $\hat{R}$ *as a function of the inverse matrices of* $\hat{V}_{jj}=\ddot{V}_{jj}(I_{k_j}+\ddot{V}_{jj}^{-1}\ddot{X}_{jj})$ *and then using the expansion* $(I-X)^{-1}=I+X+O\left(X^2\right)$ *to obtain* $(I_{k_j}+\ddot{V}_{jj}^{-1}\ddot{X}_{jj})^{-1}=I_{k_j}-\ddot{V}_{jj}^{-1}\ddot{X}_{jj}+O_p\left(\delta_{NT}^{-4}\right)$ *given that*

[10] This means that it contains terms of order $O_p\left(\delta_{NT}^{-2}\right)$ and a remainder of order $O_p\left(\delta_{NT}^{-4}\right)$. Instead, AGGR(2019) need to obtain higher order expansions with remainders of order $O_p\left(\delta_{NT}^{-6}\right)$ because they replace our assumption $\frac{N}{T^{3/2}}\to0$ with $\frac{N}{T^{5/2}}\to0$.

$\ddot{X}_{jk} = O_p\left(\delta_{NT}^{-2}\right)$. Instead AGGR(2019) use a second-order expansion $(I-X)^{-1} = I+X+X^2+O\left(X^3\right)$ to obtain their equation (B.5). They require a higher order asymptotic expansion than ours because their rate conditions on N and T are weaker than those we assume under Assumption 1.

The next step is to obtain an asymptotic expansion of the k^c largest eigenvalues of $\hat{R}$ when the two panels share k^c common factors, i.e. when $f_{jt} = [f_t^{c\prime}, f_{jt}^{s\prime}]'$ for $j = 1,2$ (hence, when the null hypothesis of k^c common factors is true). We summarize these results in the following lemma.

Lemma A.4 *Suppose that Assumptions 1-4 hold and assume that $f_{jt} = [f_t^{c\prime}, f_{jt}^{s\prime}]'$ for $j = 1,2$. Letting $\hat{\Psi}_{cc}$ denote the first $k^c \times k^c$ block obtained from $\hat{\Psi}$ defined in Lemma A.3, it follows that*

$$\sum_{l=1}^{k^c} \hat{\rho}_l = k^c + \frac{1}{2} tr\left(\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}\right) + O_p\left(\delta_{NT}^{-4}\right).$$

Remark 2 *Lemma A.4 gives the asymptotic expansion of $\hat{\xi}\left(k^c\right) = \sum_{l=1}^{k^c} \hat{\rho}_l$ through order $O_p\left(\delta_{NT}^{-2}\right)$ under the null hypothesis that there are k^c factors that are common between the two groups. This result is a simplified version of equation (B.13) of AGGR since it only contains terms of order $O_p\left(\delta_{NT}^{-2}\right)$ (their expansion contains terms of order $O_p\left(\delta_{NT}^{-4}\right)$).*

Next, we can use Lemma A.2 to expand $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_j f_{jt}\right)\left(\hat{f}_{kt} - H_k f_{kt}\right)'$ up to a remainder of order $O_p\left(\delta_{NT}^{-4}\right)$. We can then obtain the following result using the definition of $\hat{\Psi}_{cc}$ given above.

Lemma A.5 *Suppose Assumptions 1-4 strengthened by Assumption 5 hold. Then, letting $u_{jt}^{(c)}$ denote the $k^c \times 1$ vector containing the first k^c rows of $u_{jt} \equiv \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)\frac{\Lambda_j'\varepsilon_{jt}}{\sqrt{N_j}}$ and defining $\mathcal{U}_t \equiv \mu_N u_{1t}^{(c)} - u_{2t}^{(c)}$, we have that under the null hypothesis of k^c common factors,*

$$\hat{\Psi}_{cc} = -\frac{1}{TN}\sum_{t=1}^{T}\mathcal{U}_t\mathcal{U}_t' + O_p\left(\delta_{NT}^{-4}\right).$$

The asymptotic distribution of the test statistic given Theorem 2.1 follows from the previous lemmas by adding Assumption 6 (in addition to Assumptions 1-5).

Proof of Lemma A.3. Part (a): This follows from Lemma A.2 of Goncalves and Perron (2014) and the fact that the rotation matrices are $O_p(1)$. Assumptions 1-4 are sufficient to apply this result.

Part (b): We follow AGGR(2019) but only consider a first-order asymptotic expansion of $\hat{R}$. In particular, we write

$$\hat{R} = \hat{V}_{11}^{-1}\hat{V}_{12}\hat{V}_{22}^{-1}\hat{V}_{21} = (I_{k_1} + \ddot{V}_{11}^{-1}\ddot{X}_{11})^{-1}\ddot{V}_{11}^{-1}(\ddot{V}_{12} + \ddot{X}_{12})(I_{k_2} + \ddot{V}_{22}^{-1}\ddot{X}_{22})^{-1}\ddot{V}_{22}^{-1}(\ddot{V}_{21} + \ddot{X}_{21}),$$

where we used $\hat{V}_{jj} = \ddot{V}_{jj}(I_{k_j} + \ddot{V}_{jj}^{-1}\ddot{X}_{jj})$. We then use the expansion $(I-X)^{-1} = I + X + O\left(X^2\right)$ to obtain $(I_{k_j} + \ddot{V}_{jj}^{-1}\ddot{X}_{jj})^{-1} = I_{k_j} - \ddot{V}_{jj}^{-1}\ddot{X}_{jj} + O_p\left(\delta_{NT}^{-4}\right)$. Contrary to AGGR(2019), we only keep terms up to order $O_p\left(\delta_{NT}^{-4}\right)$. Thus, the asymptotic expansion of $\hat{R}$ in part (b) only considers terms that are linear in $\ddot{X}_{jk}$. Terms involving products or squares of $\ddot{X}_{jk}$ are of order $O_p\left(\delta_{NT}^{-4}\right) = O_p\left(1/\min\left(N^2, T^2\right)\right)$, which is either $O_p\left(N/T^{3/2}\right)$ if $\delta_{NT} = \sqrt{N}$ or $O_p\left(\sqrt{T}/N\right)$ if $\delta_{NT} = \sqrt{T}$. Since we assume that $\sqrt{T}/N \to 0$ and $N/T^{3/2} \to 0$, the remainder is $O_p\left(\delta_{NT}^{-4}\right) = o_p(1)$. ■

Proof of Lemma A.4. We follow closely the derivations of AGGR(2019) leading to their equation (B.13) in Section B.1.4. Specifically, consider the eigenvector-eigenvalue problem associated with $\hat{R}$, $\hat{R}\hat{W} = \hat{W}\hat{\Lambda}$, where we let $\hat{W}$ denote the $k_1 \times k^c$ matrix containing the k^c eigenvectors of $\hat{R}$ associated to its largest k^c eigenvalues $\hat{\rho}_1^2, \ldots, \hat{\rho}_{k^c}^2$, which we collect into the diagonal matrix $\hat{\Lambda} = diag\left(\hat{\rho}_l^2 : l = 1, \ldots, k^c\right)$. We can replace $\hat{R}$ from its asymptotic expansion in Lemma A.3(b):

$$\left[\left(H_1'\right)^{-1}\left(\tilde{R} + \tilde{V}_{11}^{-1}\hat{\Psi}\right)H_1' + O_p\left(\delta_{NT}^{-4}\right)\right]\hat{W} = \hat{W}\hat{\Lambda}.$$

Pre-multiplying this equation by H_1' gives

$$\left(\tilde{R} + \tilde{V}_{11}^{-1}\hat{\Psi}\right)\underbrace{H_1'\hat{W}}_{=\tilde{W}_1} = \underbrace{H_1'\hat{W}}_{=\tilde{W}_1}\hat{\Lambda} + O_p\left(\delta_{NT}^{-4}\right).$$

Since $\hat{\Psi} = O_p\left(\delta_{NT}^{-2}\right)$, $\hat{R}$ converges to $\tilde{R}$, implying that they share the same eigenvectors and eigenvalues asymptotically. The next step is to use the fact that under the null when $f_{jt} = [f_t^{c\prime}, f_{jt}^{s\prime}]'$ for $j = 1, 2$, $\tilde{R}$ can be expressed as a block triangular matrix of the form

$$\tilde{R} = \begin{bmatrix} I_{k^c} & \tilde{R}_{cs} \\ 0 & \tilde{R}_{ss} \end{bmatrix},$$

where $\tilde{R}_{cs} = \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}(I_{k_1-k^c} - \tilde{R}_{ss})$, with $\tilde{\Sigma}_{cc} = T^{-1}\sum_{t=1}^{T} f_t^c f_t^{c\prime}$, $\tilde{\Sigma}_{c1} = T^{-1}\sum_{t=1}^{T} f_t^c f_{1t}^{s\prime}$ and $\tilde{R}_{ss}$ is as defined in Lemma B.3 of AGGR(2019). This result is an algebraic result that only relies on the assumption that $f_{jt} = [f_t^{c\prime}, f_{jt}^{s\prime}]'$ for $j = 1, 2$. Hence, it holds under our Assumptions 1-4. The fact that $\tilde{R}$ has this special form is key for deriving the asymptotic distribution of the test statistic under the null hypothesis. In particular, because $\tilde{R}$ is block triangular, its eigenvalues are equal to the eigenvalues of I_{k^c} and $\tilde{R}_{ss}$, and we can show that the largest k^c eigenvalues are all equal to 1. Similarly, the first k^c eigenvectors of $\tilde{R}$ can be shown to be of the form $(x^{c\prime}, 0')'$, where x^c is a $k^c \times 1$ vector of constants and 0 is a $(k_1 - k^c) \times 1$ vector of zeros. Hence, letting

$$\underset{k_1 \times k^c}{E_c} = \begin{pmatrix} I_{k^c} \\ 0 \end{pmatrix} \text{ and } \underset{k_1 \times (k_1 - k^c)}{E_s} = \begin{pmatrix} 0 \\ I_{k_1 - k_s} \end{pmatrix},$$

we can follow AGGR(2019) and decompose the eigenvector and eigenvalue matrices of $\hat{R}$ as

$$\tilde{W}_1 = E_c\hat{U} + E_s\hat{\alpha} \quad \text{and} \quad \hat{\Lambda} = I_{k^c} + \hat{M},$$

where $\hat{U}$ is a $k^c \times k^c$ nonsingular matrix, and $\hat{M}$ and $\hat{\alpha}$ are also stochastic matrices. Because E_c and E_s span $\mathbb{R}^{k_1}$, the decomposition of $\tilde{W}_1$ is true by definition. The same applies to the decompositon of $\hat{\Lambda}$. However, under the null hypothesis, and because $\tilde{W}_1$ and $\hat{\Lambda}$ are also the eigenvector and eigenvalue matrices of $\tilde{R}$, $\hat{\alpha}$ and $\hat{M}$ converge to zero at rate $O_p\left(\delta_{NT}^{-2}\right)$. In particular, replacing $\tilde{W}_1$ and $\hat{\Lambda}$ into the eigenvector-eigenvalue equation for $\hat{R}$ and letting $\hat{\Phi} \equiv V_{11}^{-1}\hat{\Psi}$ gives:

$$\left(\tilde{R} + \hat{\Phi}\right)\left(E_c\hat{U} + E_s\hat{\alpha}\right) = \left(E_c\hat{U} + E_s\hat{\alpha}\right)\left(I_{k^c} + \hat{M}\right) + O_p\left(\delta_{NT}^{-4}\right), \text{ and}$$
$$\tilde{R}E_c\hat{U} + \hat{\Phi}E_c\hat{U} + \tilde{R}E_s\hat{\alpha} + \hat{\Phi}E_s\hat{\alpha} = E_c\hat{U} + E_c\hat{U}\hat{M} + E_s\hat{\alpha} + E_s\hat{\alpha}\hat{M} + O_p\left(\delta_{NT}^{-4}\right).$$

Using the fact that $\tilde{R}E_c = E_c$ under the null hypothesis and the fact that $\hat{\Phi}E_s\hat{\alpha}$ and $E_s\hat{\alpha}\hat{M}$ are of order $O_p\left(\delta_{NT}^{-4}\right)$ implies that

$$\hat{\Phi}E_c\hat{U} + \tilde{R}E_s\hat{\alpha} = E_c\hat{U}\hat{M} + E_s\hat{\alpha} + O_p\left(\delta_{NT}^{-4}\right). \tag{9}$$

Pre-multiplying equation (9) by E_c' gives

$$\underbrace{E_c'\tilde{R}E_s}_{\equiv\tilde{R}_{cs}}\hat{\alpha} + \underbrace{E_c'\hat{\Phi}E_c}_{\equiv\hat{\Phi}_{cc}}\hat{U} = \underbrace{E_c'E_c}_{=I_{k^c}}\hat{U}\hat{M} + \underbrace{E_c'E_s}_{=0}\hat{\alpha} + O_p\left(\delta_{NT}^{-4}\right),$$

from which we obtain

$$\hat{M} = \hat{U}^{-1}\left(\tilde{R}_{cs}\hat{\alpha} + \hat{\Phi}_{cc}\hat{U}\right) + O_p\left(\delta_{NT}^{-4}\right).$$

Pre-multiplying equation (9) by E_s' gives $\tilde{R}_{ss}\hat{\alpha} + \hat{\Phi}_{sc}\hat{U} = \hat{\alpha} + O_p\left(\delta_{NT}^{-4}\right)$, from which we obtain

$$\hat{\alpha} = \left(I_{k_1-k^c} - \tilde{R}_{ss}\right)^{-1}\hat{\Phi}_{sc}\hat{U} + O_p\left(\delta_{NT}^{-4}\right). \tag{10}$$

Plugging $\hat{\alpha}$ into the expansion for $\hat{M}$ gives

$$\hat{M} = \hat{U}^{-1}\left(\hat{\Phi}_{cc} + \tilde{R}_{cs}\left(I_{k_1-k^c} - \tilde{R}_{ss}\right)^{-1}\hat{\Phi}_{sc}\right)\hat{U} + O_p\left(\delta_{NT}^{-4}\right). \tag{11}$$

The expansions (10) and (11) correspond to equations (C.61) and (C.62) in AGGR(2019)'s Online Appendix (proof of their Lemma B.4). Given the definition of $\tilde{R}_{cs}$, we can write $\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,1} = \tilde{R}_{cs}\left(I_{k_1-k^c} - \tilde{R}_{ss}\right)^{-1}$, from which it follows that $\hat{\Phi}_{cc}+\tilde{R}_{cs}\left(I_{k_1-k^c} - \tilde{R}_{ss}\right)^{-1}\hat{\Phi}_{sc} = \tilde{\Sigma}_{cc}^{-1}\left[\tilde{\Sigma}_{cc}\hat{\Phi}_{cc} + \tilde{\Sigma}_{c,1}\hat{\Phi}_{sc}\right]$. Letting $f_{1t} = (f_t^{c\prime}, f_{1t}^{s\prime})'$, we can write

$$\tilde{V}_{11} \equiv T^{-1}\sum_{t=1}^{T} f_{1t}f_{1t}' = \begin{pmatrix} \tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c,1} \\ \tilde{\Sigma}_{1,c} & \tilde{\Sigma}_{11} \end{pmatrix}.$$

Partioning $\hat{\Phi}$ accordingly, i.e. letting $\hat{\Phi} = \begin{pmatrix} \hat{\Phi}_{cc} & \hat{\Phi}_{cs} \\ \hat{\Phi}_{sc} & \hat{\Phi}_{ss} \end{pmatrix}$, implies that $\tilde{\Sigma}_{cc}\hat{\Phi}_{cc} + \tilde{\Sigma}_{c,1}\hat{\Phi}_{sc} = \left(\tilde{V}_{11}\hat{\Phi}\right)_{(cc)}$, where we use the notation $(A)_{(cc)}$ to denote the upper-left $k^c \times k^c$ block of any matrix A. Since $\hat{\Phi} = V_{11}^{-1}\hat{\Psi}$, we obtain that $\left(\tilde{V}_{11}\hat{\Phi}\right)_{(cc)} = \left(\hat{\Psi}\right)_{(cc)} \equiv \hat{\Psi}_{cc}$, the upper-left $k^c \times k^c$ block of $\hat{\Psi}$ as defined in Lemma A.3(b). Hence,

$$\hat{M} = \hat{U}^{-1}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}\hat{U} + O_p\left(\delta_{NT}^{-4}\right).$$

This implies that

$$\hat{\Lambda} = I_{k^c} + \hat{M} = I_{k^c} + \hat{U}^{-1}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}\hat{U} + O_p\left(\delta_{NT}^{-4}\right),$$

from which it follows that

$$\hat{\Lambda}^{1/2} = I_{k^c} + \frac{1}{2}\hat{U}^{-1}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}\hat{U} + O_p\left(\delta_{NT}^{-4}\right),$$

by using the expansion $(I+X)^{1/2} = I + \frac{1}{2}X + O_p\left(X^2\right)$ with $X = \hat{M}$. Taking the trace of $\hat{\Lambda}^{1/2}$ yields the asymptotic expansion of $\hat{\xi}\left(k^c\right) = \sum_{l=1}^{k^c}\hat{\rho}_l$. ■

Proof of Lemma A.5. This result follows by replacing $\hat{\Psi}_{cc}$ with the expression from Lemma A.3(b). In particular, recall that $\hat{\Psi}$ is defined as

$$\hat{\Psi} \equiv -\hat{X}_{11}\tilde{R} + \hat{X}_{12}\tilde{B} + \tilde{B}'\hat{X}_{21} - \tilde{B}'\hat{X}_{22}\tilde{B},$$

where $\tilde{B} \equiv \tilde{V}_{22}^{-1}\tilde{V}_{21}$, and $\hat{X}_{jk}$ is as defined in Lemma A.3(b). Under the null hypothesis, both $\tilde{R}$ and $\tilde{B}$ have the same structure $[E_c \vdots *]$, which implies that the upper-left $k^c \times k^c$ block $\hat{\Psi}_{cc}$ is equal to

$$\hat{\Psi}_{cc} = -\hat{X}_{11,cc} + \hat{X}_{12,cc} + \hat{X}_{21,cc} - \hat{X}_{22,cc},$$

as argued by AGGR(2019) (see their equation (C.69) in the Online Appendix). As explained by AGGR(2019), we can rewrite the expression of $\hat{\Psi}_{cc}$ as

$$\begin{aligned}\hat{\Psi}_{cc} &= -\frac{1}{T}\sum_{t=1}^{T}\left(\psi_{1t}^{(c)} - \psi_{2t}^{(c)}\right)\left(\psi_{1t}^{(c)} - \psi_{2t}^{(c)}\right)' \\ &= -\left\{\frac{1}{T}\sum_{t=1}^{T}\psi_{1t}^{(c)}\psi_{1t}^{(c)\prime} - \frac{1}{T}\sum_{t=1}^{T}\psi_{1t}^{(c)}\psi_{2t}^{(c)\prime} + \frac{1}{T}\sum_{t=1}^{T}\psi_{2t}^{(c)}\psi_{1t}^{(c)\prime} - \frac{1}{T}\sum_{t=1}^{T}\psi_{2t}^{(c)}\psi_{2t}^{(c)\prime}\right\},\end{aligned}$$

where $\psi_{jt}^{(c)}\psi_{kt}^{(c)\prime}$ denotes the upper-left $k^c \times k^c$ block of the matrix $\psi_{jt}\psi_{kt}'$, where $\psi_{jt} \equiv H_j^{-1}\left(\hat{f}_{jt} - H_j f_{jt}\right)$. For any $j, k \in \{1, 2\}$, we can write

$$\frac{1}{T}\sum_{t=1}^{T}\psi_{jt}\psi_{kt}' = H_j^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_j f_{jt}\right)\left(\hat{f}_{kt} - H_k f_{kt}\right)'\left(H_k'\right)^{-1}.$$

The result follows by replacing $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}_{jt} - H_j f_{jt}\right)\left(\hat{f}_{kt} - H_k f_{kt}\right)'$ with the asymptotic expansion given in Lemma A.2. ■

Proof of Theorem 2.1. The proof of this result follows from Lemmas A.3, A.4, and A.5 under Assumptions 1-6, when the null hypothesis is true. ■

B Bootstrap results

We organize this appendix as follows. In Appendix B.1, we provide a set of bootstrap high level conditions which are the bootstrap analogues of Assumptions 3, 4, and 5. These conditions are used to prove two auxiliary lemmas in Appendix B.2. Appendix B.3 provides the proofs of the results in Section 3.

B.1 Bootstrap high level conditions

Here, we propose a set of high level conditions on $\varepsilon^*_{j,it}$ under which we can characterize the asymptotic distribution of the bootstrap test statistic $\hat{\xi}^*(k^c)$. These conditions can be verified for any resampling scheme.

Condition A*

(i) $E^*\left(\varepsilon^*_{j,it}\right) = 0$, for all i, t.

(ii) $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma^*_{j,st}\right|^2 = O_p(1)$, where $\gamma^*_{j,st} \equiv E^*(\frac{1}{N_j}\sum_{i=1}^{N_j}\varepsilon^*_{j,is}\varepsilon^*_{j,it})$.

(iii) $\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}E^*\left|\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\left(\varepsilon^*_{j,it}\varepsilon^*_{j,is} - E^*\left(\varepsilon^*_{j,it}\varepsilon^*_{j,is}\right)\right)\right|^2 = O_p(1)$.

Condition B*

(i) $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{f}_{js}\tilde{f}'_{jt}\gamma^*_{j,st} = O_p(1)$.

(ii) $\frac{1}{T}\sum_{t=1}^{T}E^*\left\|\frac{1}{\sqrt{TN_j}}\sum_{s=1}^{T}\sum_{i=1}^{N_j}\tilde{f}_{js}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))\right\|^2 = O_p(1)$.

(iii) $E^*\left\|\frac{1}{\sqrt{TN_j}}\sum_{t=1}^{T}\tilde{f}_{jt}\varepsilon^{*\prime}_{jt}\tilde{\Lambda}_j\right\|^2 = O_p(1)$.

(iv) $\frac{1}{T}\sum_{t=1}^{T}E^*\left\|\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{\sqrt{N_j}}\right\|^2 = O_p(1)$.

Condition C*

(i) $\frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}\tilde{f}_{js}\gamma^*_{j,st}\right\|^2 = O_p(1)$.

(ii) $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\sum_{t=1}^{T}\gamma^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}} = O_{p^*}(1)$.

(iii) $\frac{1}{T}\sum_{s=1}^{T}E^*\left\|\sum_{t=1}^{T}\gamma^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\right\|^2 = O_p(1)$.

(iv) $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\lambda}_{k,i}\varepsilon^*_{k,it}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))\right) = O_{p^*}(1)$, where $N = \min(N_1, N_2)$.

(v) $\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}_{k,i_2}\varepsilon^*_{k,i_2t}(\varepsilon^*_{j,i_1s}\varepsilon^*_{j,i_1t} - E^*(\varepsilon^*_{j,i_1s}\varepsilon^*_{j,i_1t}))\right) = O_{p^*}(1)$.

(vi) $\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\lambda}_{k,i}\varepsilon^*_{k,it}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))\right)\right\|^2 = O_{p^*}(1)$, where $N = \min(N_1, N_2)$.

(vii) $\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}_{k,i_2}\varepsilon^*_{k,i_2t}(\varepsilon^*_{j,i_1s}\varepsilon^*_{j,i_1t} - E^*(\varepsilon^*_{j,i_1s}\varepsilon^*_{j,i_1t}))\right)\right\|^2 = O_{p^*}(1)$.

Remark 3 *Conditions A^* and B^* are used in GP(2014) and Gonçalves and Perron (2020) and have been verified for the wild bootstrap and the cross-sectional dependent bootstrap, respectively, when $\tilde{f}_{jt}$ and $\tilde{\lambda}_{j,i}$ are the PCA estimators. Here, they are obtained as in AGGR(2019) under the null. Condition C^* is new to the group factor model and needs to verified.*

B.2 Asymptotic expansion of the sample covariance of the bootstrap factors estimation error

For each group j, we have that

$$\hat{f}_{jt}^* - H_j^* \tilde{f}_{jt} = \mathcal{V}_j^{*-1} \left(A_{j,1t}^* + A_{j,2t}^* + A_{j,3t}^* + A_{j,4t}^* \right), \tag{12}$$

where

$$\begin{aligned}
A_{j,1t}^* &= \frac{1}{T} \sum_{s=1}^{T} \hat{f}_{js}^* \gamma_{j,st}^*, \text{ with } \gamma_{j,st}^* = E^* \left(\frac{1}{N_j} \sum_{i=1}^{N_j} \varepsilon_{j,is}^* \varepsilon_{j,it}^* \right); \\
A_{j,2t}^* &= \frac{1}{T} \sum_{s=1}^{T} \hat{f}_{js}^* \zeta_{j,st}^*, \text{ with } \zeta_{j,st} = \frac{1}{N_j} \sum_{i=1}^{N_j} (\varepsilon_{j,is}^* \varepsilon_{j,it}^* - E^*(\varepsilon_{j,is}^* \varepsilon_{j,it}^*)); \\
A_{j,3t}^* &= \frac{1}{T} \sum_{s=1}^{T} \hat{f}_{js}^* \eta_{j,st}^*, \text{ with } \eta_{j,st}^* = \frac{1}{N_j} \sum_{i=1}^{N_j} \tilde{\lambda}_{j,i}' \tilde{f}_{js} \varepsilon_{j,it}^* = \tilde{f}_{js}' \frac{\tilde{\Lambda}_j' \varepsilon_{jt}^*}{N_j}; \text{ and} \\
A_{j,4t}^* &= \frac{1}{T} \sum_{s=1}^{T} \hat{f}_{js}^* \xi_{j,st}^*, \text{ with } \xi_{j,st}^* = \tilde{f}_{jt}' \frac{\tilde{\Lambda}_j' \varepsilon_{js}^*}{N_j} = \eta_{j,ts}^*.
\end{aligned}$$

First, note that $\frac{1}{T}\sum_{s=1}^{T} \|\hat{f}_{js}^* - H_j^* \tilde{f}_{js}\|^2 = O_{p^*}\left(\delta_{NT}^{-2}\right)$ under Conditions A* and B* below, which are all from GP(2014).

The following auxiliary lemmas are the bootstrap analogues of Lemmas A.1 and A.2.

Lemma B.1 *Suppose Conditions A*, B* and C* hold. Then, for any* $j,k \in \{1,2\}$*: (a)* $\frac{1}{T}\sum_{t=1}^{T} A_{j,1t}^* A_{k,1t}^{*\prime} = O_{p^*}\left(\delta_{NT}^{-4}\right)$*; (b)* $\frac{1}{T}\sum_{t=1}^{T} A_{j,2t}^* A_{k,2t}^{*\prime} = O_{p^*}\left(\delta_{NT}^{-4}\right)$*; (c)* $\frac{1}{T}\sum_{t=1}^{T} A_{j,4t}^* A_{k,4t}^{*\prime} = O_{p^*}\left(\delta_{NT}^{-4}\right)$*; (d)* $\frac{1}{T}\sum_{t=1}^{T} A_{j,mt}^* A_{k,nt}^{*\prime} = O_{p^*}\left(\delta_{NT}^{-4}\right)$ *for* $m \neq n$*, where* $m,n \in \{1,2,3,4\}$*; and (e)*

$$\frac{1}{T}\sum_{t=1}^{T} A_{j,3t}^* A_{k,3t}^{*\prime} = \frac{1}{\sqrt{N_j N_k}} \mathcal{V}_j^* H_j^* \frac{1}{T}\sum_{t=1}^{T} u_{jt}^* u_{kt}^{*\prime} H_k^{*\prime} \mathcal{V}_k^{*\prime} = O_{p^*}\left(N^{-1}\right), \quad \text{where} \quad u_{jt}^* \equiv \left(\frac{\tilde{\Lambda}_j' \tilde{\Lambda}_j}{N_j} \right)^{-1} \frac{\tilde{\Lambda}_j' \varepsilon_{jt}^*}{\sqrt{N_j}}.$$

Lemma B.2 *Suppose Conditions A*, B* and C* hold. Then, for* $j,k \in \{1,2\}$,

$$\frac{1}{T}\sum_{t=1}^{T} \left(\hat{f}_{jt}^* - H_j^* \tilde{f}_{jt} \right) \left(\hat{f}_{kt}^* - H_k^* \tilde{f}_{kt} \right)' = \frac{1}{\sqrt{N_j N_k}} H_j^* \left(\frac{1}{T}\sum_{t=1}^{T} u_{jt}^* u_{kt}^{*\prime} \right) H_k^{*\prime} + O_{p^*}\left(\delta_{NT}^{-4}\right),$$

where u_{jt}^* *is as defined in Lemma B.1.*

Proof of Lemma B.1. This proof follows closely the proof of Lemma A.1. **Part (a):** We can bound the norm of $\frac{1}{T}\sum_{t=1}^{T} A_{j,1t}^* A_{k,1t}^{*\prime}$ by

$$\frac{1}{T}\sum_{t=1}^{T} \left\| A_{j,1t}^* A_{k,1t}^{*\prime} \right\| \le \left(\frac{1}{T}\sum_{t=1}^{T} \left\| A_{j,1t}^* \right\|^2 \right)^{1/2} \left(\frac{1}{T}\sum_{t=1}^{T} \left\| A_{k,1t}^* \right\|^2 \right)^{1/2},$$

thus we show that $\frac{1}{T}\sum_{t=1}^{T} \left\| A_{j,1t}^* \right\|^2 = O_p\left(\delta_{NT}^{-4}\right)$ for any j. To show this, we write $A_{j,1t}^* = A_{j,1t}^{*(1)} + A_{j,1t}^{*(2)}$, where

$$A^{*(1)}_{j,1t} \equiv \frac{1}{T}\sum_{s=1}^{T}\left(\hat{f}^*_{js} - H^*_j \tilde{f}_{js}\right)\gamma^*_{j,st} \quad \text{and} \quad A^{*(1)}_{j,1t} \equiv H^*_j \frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\gamma^*_{j,st},$$

and we show that $I^*_1 \equiv \frac{1}{T}\sum_{t=1}^{T}\left\|A^{*(1)}_{j,1t}\right\|^2$ and $II^*_1 \equiv \frac{1}{T}\sum_{t=1}^{T}\left\|A^{*(2)}_{j,1t}\right\|^2$ are both of order $O_{p^*}\left(\delta_{NT}^{-4}\right)$ under our bootstrap high level conditions. First, note that

$$\|I^*_1\| \leq \underbrace{\frac{1}{T}\sum_{s=1}^{T}\|\hat{f}^*_{js} - H^*_j\tilde{f}_{js}\|^2}_{=O_{p^*}\left(\delta_{NT}^{-2}\right)}\underbrace{\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma^*_{j,st}|^2}_{=O_{p^*}\left(T^{-1}\right)} = O_{p^*}(\delta_{NT}^{-2}T^{-1}) = O_{p^*}(\delta_{NT}^{-4}),$$

since $T^{-1}\sum_{s=1}^{T}\|\hat{f}^*_{js} - H^*_j\tilde{f}_{js}\|^2 = O_{p^*}\left(\delta_{NT}^{-2}\right)$ under Condition A*, and $T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma^*_{j,st}|^2 = O_p\left(T^{-1}\right)$ under Condition A*-(i). Similarly, ignoring $H^*_j = O_{p^*}(1)$, Condition C*-(i) (which is new) implies that $II^*_1 = O_{p^*}\left(T^{-2}\right) = O_{p^*}\left(\delta_{NT}^{-4}\right)$ since

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\gamma^*_{j,st}\right\|^2 = \frac{1}{T^2}\underbrace{\frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}\tilde{f}_{js}\gamma^*_{j,st}\right\|^2}_{=O_p(1)\text{ by Cond-C*(i)}}.$$

Part (b): We let $I^*_2 \equiv T^{-1}\sum_{t=1}^{T}\|A^{*(1)}_{j,2t}\|^2$ and $II^*_2 \equiv T^{-1}\sum_{t=1}^{T}\|A^{*(2)}_{j,2t}\|^2$, where $A^{*(1)}_{j,2t} \equiv T^{-1}\sum_{s=1}^{T}(\hat{f}^*_{js} - H^*_j\tilde{f}_{js})\zeta^*_{j,st}$ and $A^{*(2)}_{j,2t} \equiv H^*_j T^{-1}\sum_{s=1}^{T}\tilde{f}_{js}\zeta^*_{j,st}$, with $\zeta^*_{j,st} \equiv N_j^{-1}\sum_{i=1}^{N_j}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))$. First, note that

$$I^*_2 \leq \left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{f}^*_{js} - H^*_j\tilde{f}_{js}\|^2\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\zeta^*_{j,st}|^2\right) = O_{p^*}(\delta_{NT}^{-2}N_j^{-1}) = O_p(\delta_{NT}^{-4}),$$

since $\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}|\zeta^*_{j,st}|^2 = O_{p^*}(N_j^{-1})$ as implied by Condition A*-(iii). Second, by Condition B*-(ii),

$$II^*_2 \leq \|H^*_j\|^2\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\zeta^*_{j,st}\right\|^2 = O_p((TN_j)^{-1}) = O_p(\delta_{NT}^{-4}).$$

Part (c): We let $A^{*(1)}_{j,4t} \equiv \frac{1}{T}\sum_{s=1}^{T}(\hat{f}^*_{js} - H^*_j\tilde{f}_{js})\xi^*_{j,st}$ and $A^{*(2)}_{j,4t} \equiv H^*_j\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\xi^*_{j,st}$, with $\xi^*_{j,st} \equiv \tilde{f}'_{jt}\frac{\tilde{\Lambda}'_j\varepsilon^*_{js}}{N_j}$. We show that $I^*_4 \equiv T^{-1}\sum_{t=1}^{T}\left\|A^{*(1)}_{j,4t}\right\|^2$ and $II_4 \equiv T^{-1}\sum_{t=1}^{T}\left\|A^{*(2)}_{j,4t}\right\|^2$ are both $O_{p^*}\left(\delta_{NT}^{-4}\right)$ under our assumptions. For the first term, using the definition of $\xi^*_{j,st}$, we have that

$$I^*_4 = \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}^*_{js} - H^*_j\tilde{f}_{js})\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\tilde{f}_{jt}\right\|^2 \leq \underbrace{\left\|\frac{1}{T}\sum_{s=1}^{T}(\hat{f}^*_{js} - H^*_j\tilde{f}_{js})\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\right\|^2}_{=O_{p^*}\left(\delta_{NT}^{-2}\right)O_{p^*}\left(N_j^{-1}\right)}\underbrace{\frac{1}{T}\sum_{t=1}^{T}\|\tilde{f}_{jt}\|^2}_{=k_j} = O_p(\delta_{NT}^{-4}),$$

since by Cauchy-Schwartz's inequality,

$$\left\| \frac{1}{T}\sum_{s=1}^{T}(\hat{f}_{js}^{*}-H_j^{*}\tilde{f}_{js})\frac{\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j}{N_j}\right\|^2 \leq \underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js}^{*}-H_j^{*}\tilde{f}_{js}\right\|^2}_{=O_{p^*}\left(\delta_{NT}^{-2}\right)}\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j}{N_j}\right\|^2}_{=O_{p^*}\left(N_j^{-1}\right)} = O_{p^*}\left(\delta_{NT}^{-4}\right),$$

given that $\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\tilde{\Lambda}_j'\varepsilon_{js}^{*}}{\sqrt{N_j}}\right\|^2 = O_{p^*}(1)$ under Condition B*-(iv). For II_4^*, using the definition of $\xi_{j,st}^{*} \equiv \frac{\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j}{N_j}\tilde{f}_{jt}$ (and ignoring $H_j^{*}=O_{p^*}(1)$), we have that

$$\begin{aligned} II_4^* &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\frac{\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j}{N_j}\tilde{f}_{jt}'\right\|^2 \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\frac{\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j}{N_j}\right\|^2\left\|\tilde{f}_{jt}\right\|^2 \\ &\leq \|H_j\|^2\frac{1}{TN_j}\underbrace{\left\|\frac{1}{\sqrt{TN_j}}\sum_{s=1}^{T}\tilde{f}_{js}\varepsilon_{js}^{*\prime}\tilde{\Lambda}_j\right\|^2}_{=O_{p^*}(1)\text{ by Cond-B*(iii)}}\underbrace{\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{f}_{jt}\right\|^2}_{=k_j} = O_{p^*}\left(\delta_{NT}^{-4}\right). \end{aligned}$$

Part (d): Given parts (a), (b), and (c), all the cross terms that involve $A_{j,1t}^{*}$, $A_{j,2t}^{*}$ and $A_{j,4t}^{*}$ are $O_{p^*}\left(\delta_{NT}^{-4}\right)$ by an application of Cauchy-Schwartz's inequality. Hence, we only need to show that $T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}A_{k,3t}^{*\prime}$ is $O_{p^*}\left(\delta_{NT}^{-4}\right)$ for $m=1,2,4$. Using the definition of $A_{k,3t}^{*}$, we have that

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}A_{k,3t}^{*\prime} &= T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}^{*}\eta_{k,st}^{*}\right)', \text{ where } \eta_{k,st}^{*}\equiv\tilde{f}_{ks}'\frac{\tilde{\Lambda}_k'\varepsilon_{kt}^{*}}{N_k} \\ &= T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}_{ks}^{*}\tilde{f}_{ks}'\frac{\tilde{\Lambda}_k'\varepsilon_{kt}^{*}}{N_k}\right)' \\ &= \left[T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k}\right]\underbrace{\frac{\tilde{F}_k'\hat{F}_k^{*}}{T}}_{=O_{p^*}(1)}. \end{aligned}$$

Thus, it suffices to show that $T^{-1}\sum_{t=1}^{T}A_{j,mt}^{*}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k}=O_{p^*}\left(\delta_{NT}^{-4}\right)$. Starting with $m=1$, by the definition of $A_{j,1t}^{*}$, we have that

$$\frac{1}{T}\sum_{t=1}^{T}A_{j,1t}^{*}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} = \frac{1}{T}\sum_{t=1}^{T}A_{j,1t}^{*(1)}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} + \frac{1}{T}\sum_{t=1}^{T}A_{j,1t}^{*(2)}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} \equiv (a_1^*)+(b_1^*).$$

Note that we can rewrite (b_1^*) as

$$(b_1^*) = H_j^{*}\frac{1}{T}\frac{1}{\sqrt{TN_k}}\underbrace{\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\sum_{t=1}^{T}\gamma_{j,st}^{*}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{\sqrt{N_k}}\right]}_{=O_{p^*}(1)\text{ by Cond-C*(ii)}} = O_{p^*}\left(\delta_{NT}^{-4}\right).$$

In addition,

$$\|(a_1^*)\| \leq \left(\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js}^* - H_j^*\tilde{f}_{js}\right\|^2}_{=O_{p^*}\left(\delta_{NT}^{-2}\right)}\right)^{1/2} \left(\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}^*\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k}\right\|^2}_{=O_{p^*}\left(\delta_{NT}^{-6}\right)}\right)^{1/2},$$

where

$$\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma_{j,st}^*\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k}\right\|^2 = \frac{1}{N_k}\frac{1}{T^2}\underbrace{\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\gamma_{j,st}^*\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{\sqrt{N_k}}\right\|^2\right]}_{=O_{p^*}(1)\text{ by Cond-C*(iii)}} = O_{p^*}\left(\delta_{NT}^{-6}\right)$$

provided the term in square bracket is $O_{p^*}(1)$, which follows under Condition C*-(iii). Consider next $m=2$. Using the decomposition of $A_{j,2t}^* = A_{j,2t}^{*(1)} + A_{j,2t}^{*(2)}$, we can write

$$\frac{1}{T}\sum_{t=1}^{T}A_{j,2t}^*\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} = \frac{1}{T}\sum_{t=1}^{T}A_{j,2t}^{*(1)}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} + \frac{1}{T}\sum_{t=1}^{T}A_{j,2t}^{*(2)}\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k} \equiv (a_2^*) + (b_2^*).$$

Note that

$$(b_2^*) = \frac{1}{\sqrt{T}}\frac{\sqrt{N}}{N_jN_k}\underbrace{\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\lambda}_{k,i}\varepsilon_{k,it}^*(\varepsilon_{j,is}^*\varepsilon_{j,it}^* - E^*(\varepsilon_{j,is}^*\varepsilon_{j,it}^*))\right)\right]}_{=O_{p^*}(1)\text{ by Cond-C*(iv)}}$$

$$+\frac{1}{T}\frac{1}{\sqrt{N_kN_j}}\underbrace{\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}_{k,i_2}\varepsilon_{k,i_2t}^*(\varepsilon_{j,i_1s}^*\varepsilon_{j,i_1t}^* - E^*(\varepsilon_{j,i_1s}^*\varepsilon_{j,i_1t}^*))\right)\right]}_{=O_{p^*}(1)\text{ by Cond-C*(v)}} = O_{p^*}\left(\delta_{NT}^{-4}\right).$$

By Cauchy-Schwartz's inequality, we can bound (a_2^*) by

$$\underbrace{\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}_{js}^* - H_j^*\tilde{f}_{js}\right\|^2\right)^{1/2}}_{=O_{p^*}\left(\delta_{NT}^{-1}\right)}\left(\underbrace{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\zeta_{j,st}^*\frac{\varepsilon_{kt}^{*\prime}\tilde{\Lambda}_k}{N_k}\right\|^2}_{=(a_2^*-ii)}\right)^{1/2} = O_{p^*}\left(\delta_{NT}^{-1}\right)O_{p^*}\left(\delta_{NT}^{-3}\right),$$

since

$$(a_2^* - ii) = \underbrace{\frac{N}{N_k^2N_j^2}}_{=O\left(\delta_{NT}^{-6}\right)}\underbrace{\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\lambda}_{k,i}\varepsilon_{k,it}^*(\varepsilon_{j,is}^*\varepsilon_{j,it}^* - E^*(\varepsilon_{j,is}^*\varepsilon_{j,it}^*))\right)\right\|^2\right]}_{=O_{p^*}(1)\text{ by Cond-C*(vi)}}$$

$$+\underbrace{\frac{1}{N_kN_j}\frac{1}{T}}_{=O\left(\delta_{NT}^{-6}\right)}\underbrace{\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_jN_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}_{k,i_2}\varepsilon_{k,i_2t}^*(\varepsilon_{j,i_1s}^*\varepsilon_{j,i_1t}^* - E^*(\varepsilon_{j,i_1s}^*\varepsilon_{j,i_1t}^*))\right)\right\|^2\right]}_{=O_{p^*}(1)\text{ by Cond-C*(vii)}} = O_{p^*}(\delta_{NT}^{-6}).$$

Finally, consider $m=4$. Using using the decomposition of $A^*_{j,4t} = A^{*(1)}_{j,4t} + A^{*(2)}_{j,4t}$, we can write

$$\frac{1}{T}\sum_{t=1}^{T} A^*_{j,4t}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k} = \frac{1}{T}\sum_{t=1}^{T} A^{*(1)}_{j,4t}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k} + \frac{1}{T}\sum_{t=1}^{T} A^{*(2)}_{j,4t}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k} \equiv (a^*_4) + (b^*_4).$$

Note that

$$\begin{aligned}
(b^*_4) &= H^*_j \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\xi^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k} \\
&= H^*_j \underbrace{\left[\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_{js}\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\right]}_{=O_{p^*}\left(\frac{1}{\sqrt{TN_j}}\right)} \quad \underbrace{\left[\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{jt}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k}\right]}_{=O_{p^*}\left(\frac{1}{\sqrt{TN_k}}\right) \text{ by Cond-B*(iii)}} \quad = O_{p^*}\left(\frac{1}{T\sqrt{N_jN_k}}\right) = O_{p^*}\left(\delta^{-4}_{NT}\right).
\end{aligned}$$

In addition,

$$\|(a^*_4)\| \le \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{f}^*_{js} - H^*_j\tilde{f}_{js}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\xi^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k}\right\|^2\right)^{1/2},$$

where

$$\begin{aligned}
\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\tilde{f}_{jt}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k}\right\|^2 &= \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{jt}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k}\right\|^2 \\
&\le \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{N_j}\right\|^2\left\|\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{jt}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{N_k}\right\|^2 \\
&= \frac{1}{N_k}\frac{1}{TN_j}\underbrace{\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{\varepsilon^{*\prime}_{js}\tilde{\Lambda}_j}{\sqrt{N_j}}\right\|^2\right)}_{=O_{p^*}(1) \text{ by Cond-B*(iv)}}\underbrace{\left\|\frac{1}{\sqrt{TN_k}}\sum_{t=1}^{T}\tilde{f}_{jt}\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k\right\|^2}_{=O_{p^*}(1) \text{ by Cond-B*(iii)}} \\
&= O_{p^*}\left(\frac{1}{TN_jN_k}\right) = O_{p^*}\left(\delta^{-6}_{NT}\right),
\end{aligned}$$

implying that $\|(a^*_4)\| = O_{p^*}\left(\delta^{-4}_{NT}\right)$.

Part (e): By definition, $A^*_{j,3t} \equiv \frac{1}{T}\sum_{s=1}^{T}\hat{f}^*_{js}\eta^*_{j,st}$, where $\eta^*_{j,st} \equiv \tilde{f}'_{js}\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{N_j}$. Using the definition of the bootstrap rotation matrix, $H^*_j \equiv \mathcal{V}^{*-1}_j\frac{\hat{F}^{*\prime}_j\tilde{F}_j}{T}\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}$, we can rewrite this term as

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}A^*_{j,3t}A^{*\prime}_{k,3t} &= \frac{1}{\sqrt{N_jN_k}}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{f}^*_{js}\tilde{f}'_{js}\right)\frac{1}{T}\sum_{t=1}^{T}\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{\sqrt{N_j}}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\left(\frac{1}{T}\sum_{l=1}^{T}\tilde{f}_{kl}\hat{f}^{*\prime}_{kl}\right) \\
&= \frac{1}{\sqrt{N_jN_k}}\left(\frac{\hat{F}^{*\prime}_j\tilde{F}_j}{T}\right)\frac{1}{T}\sum_{t=1}^{T}\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{\sqrt{N_j}}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\left(\frac{\hat{F}^{*\prime}_k\tilde{F}_k}{T}\right)' \\
&= \frac{1}{\sqrt{N_jN_k}}\mathcal{V}^*_jH^*_j\left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{\sqrt{N_j}}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\left(\frac{\tilde{\Lambda}'_k\tilde{\Lambda}_k}{N_k}\right)^{-1}H^{*\prime}_k\mathcal{V}^{*\prime}_k = O_{p^*}(N^{-1}),
\end{aligned}$$

given in particular Condition B*-(iv). ■

Proof of Lemma B.2. This follows immediately from Lemma B.1. ■

B.3 Proof of bootstrap results in Section 3

The section is organized as follows. First, we state several auxiliary lemmas used to prove Lemma 3.1 and Theorem 3.1, followed by their proofs. Then, we prove Lemma 3.1, Theorem 3.1 and Proposition 3.1.

Following AGGR(2019), we define $\hat{R}^* = \hat{V}_{11}^{*-1}\hat{V}_{12}^*\hat{V}_{22}^{*-1}\hat{V}_{21}^*$, where $\hat{V}_{jk}^* = \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^*\hat{f}_{kt}^{*\prime}$. The test statistic is given by $\hat{\xi}^*(k^c) \equiv \sum_{l=1}^{k^c}\hat{\rho}_l^* = tr\left(\hat{\Lambda}^{*1/2}\right)$, where $\hat{\Lambda}^* = diag\left(\hat{\rho}_l^{*2} : l = 1, \ldots, k^c\right)$ is a $k^c \times k^c$ diagonal matrix containing the k^c largest eigenvalues of $\hat{R}^*$ obtained from the eigenvalue-eigenvector problem $\hat{R}^*\hat{W}^* = \hat{W}^*\hat{\Lambda}^*$, where $\hat{W}^*$ is a $k_1 \times k^c$ matrix of eigenvectors associated to k^c largest eigenvalues. The main idea of the proof is to obtain an expansion of $\hat{R}^*$ through order $O_{p^*}\left(\delta_{NT}^{-2}\right)$, where $\delta_{NT} = \min\left(\sqrt{N}, \sqrt{T}\right)$, from which we obtain an asymptotic expansion of $\hat{\Lambda}^*$ and of $tr\left(\hat{\Lambda}^{*1/2}\right)$.

The asymptotic expansion of $\hat{R}^*$ is based on expanding $\hat{V}_{jk}^*$ around[11] $\tilde{V}_{jk}^* \equiv \frac{1}{T}\sum_{t=1}^T \tilde{f}_{jt}\tilde{f}_{kt}'$, where $\tilde{f}_{jt} = \left(\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime}\right)'$ for $j = 1, 2$. Note that $\tilde{f}_{jt}$ imposes the null hypothesis that there are k^c common factors among the two panels and it is different from the vector $\hat{f}_{jt}$, which contains the k_j largest principal components of Y_j. Hence, the need to use different notation. The properties of the bootstrap test rely heavily from imposing the null hypothesis in the bootstrap DGP. Adding and subtracting appropriately yields

$$\hat{V}_{jk}^* = \ddot{V}_{jk}^* + \ddot{X}_{jk}^*, \quad \text{with} \quad \ddot{V}_{jk}^* \equiv H_j^*\tilde{V}_{jk}^*H_k^{*\prime} \quad \text{and} \quad \ddot{X}_{jk}^* = H_j^*\hat{X}_{jk}^*H_k^{*\prime},$$

where letting $\psi_{jt}^* \equiv H_j^{*-1}\left(\hat{f}_{jt}^* - H_j^*\tilde{f}_{jt}\right)$,

$$\hat{X}_{jk}^* \equiv \frac{1}{T}\sum_{t=1}^T \psi_{jt}^*\psi_{kt}^{*\prime} + \frac{1}{T}\sum_{t=1}^T \psi_{jt}^*\tilde{f}_{kt}' + \frac{1}{T}\sum_{t=1}^T \tilde{f}_{jt}\psi_{kt}^{*\prime}.$$

Under Conditions A* and B*, we can show that $\hat{X}_{jk}^* = O_{p^*}\left(\delta_{NT}^{-2}\right)$ (this follows from Lemma B.3 of GP (2014)). Using this result, we can show that $\hat{R}^* = \ddot{R}^* + O_{p^*}\left(\delta_{NT}^{-2}\right)$, where $\ddot{R}^* = \ddot{V}_{11}^{*-1}\ddot{V}_{12}^*\ddot{V}_{22}^{*-1}\ddot{V}_{21}^* = \left(H_1^{*\prime}\right)^{-1}\tilde{R}^*H_1^{*\prime}$, where $\tilde{R}^* \equiv \tilde{V}_{11}^{*-1}\tilde{V}_{12}^*\tilde{V}_{22}^{*-1}\tilde{V}_{21}^*$. Note that $\tilde{R}^*$ is the bootstrap analogue of $\tilde{R} \equiv \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21}$ defined in Lemma B.2 of AGGR(2019).

The following auxiliary lemma provides the asymptotic expansion of $\hat{R}^*$ through order $O_{p^*}\left(\delta_{NT}^{-2}\right)$.

Lemma B.3 *Suppose Conditions A* and B* hold. Under Assumption 1,*

$$\hat{R}^* = \left(H_1^{*\prime}\right)^{-1}\left[\tilde{R}^* + \hat{\Psi}^*\right]H_1^{*\prime} + O_{p^*}\left(\delta_{NT}^{-4}\right),$$

[11] Note that $\tilde{V}_{jk}^*$ is the bootstrap analogue of $\tilde{V}_{jk} \equiv T^{-1}\sum_{t=1}^T f_{jt}f_{kt}'$ defined in eq. (B.3) of AGGR(2019). Although we keep the star notation when defining $\tilde{V}_{jk}^*$, we note that $\tilde{V}_{jk}^*$ is not random when we condition on the original sample. We adopt this notation to be consistent with notation in AGGR(2019).

where $\hat{\Psi}^* \equiv -\hat{X}^*_{11}\tilde{R}^* + \hat{X}^*_{12}\tilde{B}^* + \tilde{B}^{*\prime}\hat{X}^*_{21} - \tilde{B}^{*\prime}\hat{X}^*_{22}\tilde{B}^*, \qquad \tilde{B}^* \equiv \tilde{V}^*_{21},$ *and*

$$\hat{X}^*_{jk} \equiv \frac{1}{T}\sum_{t=1}^{T}\psi^*_{jt}\psi^{*\prime}_{kt} + \frac{1}{T}\sum_{t=1}^{T}\psi^*_{jt}\tilde{f}'_{kt} + \frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{jt}\psi^{*\prime}_{kt}, \quad \textit{where} \quad \psi^*_{jt} \equiv H^{*-1}_j\left(\hat{f}^*_{jt} - H^*_j\tilde{f}_{jt}\right).$$

Remark 4 *Lemma B.3 is the bootstrap analogue of Lemma B.2 of AGGR(2019) when the rate conditions on N and T are as assumed in Assumption 1. Note that under this assumption, we only require an asymptotic expansion through order $O_{p^*}\left(\delta_{NT}^{-2}\right)$, which means its remainder is of order $O_{p^*}\left(\delta_{NT}^{-4}\right)$.*

Remark 5 *Lemma B.3 only requires Conditions A* and B*. Condition C* is not used here. Note that $\tilde{V}^*_{11} = I_{k_1}$ and $\tilde{V}^*_{22} = I_{k_2}$, which explains the differences between the asymptotic expansions of $\hat{R}$ and $\hat{R}^*$ (in particular, we do not need to pre-multiply $\hat{\Psi}^*$ by $\tilde{V}^{*-1}_{11}$).*

Since the bootstrap test statistic is defined as $\hat{\xi}^*(k^c) \equiv tr\left(\hat{\Lambda}^{*1/2}\right)$, where $\hat{\Lambda}^* = diag\left(\hat{\rho}^{*2}_l : l = 1, \ldots, k^c\right)$ contains the first k^c eigenvalues of $\hat{R}^*$, our next result provides an asymptotic of $\hat{\Lambda}^{*1/2}$, from which we obtain an asymptotic expansion of $\hat{\xi}^*(k^c) \equiv \sum_{l=1}^{k^c}\hat{\rho}^*_l$.

Lemma B.4 *Suppose Conditions A* and B* hold. Under Assumption 1,*

(a) $\hat{\Lambda}^{*1/2} = I_{k^c} + \frac{1}{2}\hat{U}^{*-1}\hat{\Psi}^*_{cc}\hat{U}^* + O_{p^*}\left(\delta_{NT}^{-4}\right)$, *where $\hat{\Psi}^*_{cc}$ is upper-left $k^c \times k^c$ block of the matrix $\hat{\Psi}^*$ defined in Lemma B.3 and $\hat{U}^*$ is a $k^c \times k^c$ matrix.*

(b) $tr\left(\hat{\Lambda}^{*1/2}\right) = \sum_{l=1}^{k^c}\hat{\rho}^*_l = k^c + \frac{1}{2}tr\left(\hat{\Psi}^*_{cc}\right) + O_{p^*}\left(\delta_{NT}^{-4}\right)$.

Lemma B.4 is the bootstrap analogue of Lemma B.4 of AGGR(2019) when N and T satisfy the rate conditions of Assumption 1. In contrast to Lemma B.4 in AGGR(2019), which only holds under the null hypothesis, Lemma B.4 holds under both the null and the alternative hypothesis.

Next, we provide an asymptotic expansion of $\hat{\Psi}^*_{cc}$ through order $O_{p^*}\left(\delta_{NT}^{-2}\right)$ (i.e. with remainder of order $O_{p^*}\left(\delta_{NT}^{-4}\right)$). This expansion is based on the asymptotic expansion of $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}^*_{jt} - H^*_j\tilde{f}_{jt}\right)\left(\hat{f}^*_{kt} - H^*_k\hat{f}_{kt}\right)'$ given in Lemma B.2. This result is in Appendix B.2 and it requires the strengthening of Conditions A* and B* with Condition C*. We can then obtain the following result using the definition of $\hat{\Psi}^*_{cc}$ given in Lemma B.3.

Recall that $\mathcal{U}^*_t \equiv \mu_N u^{*(c)}_{1t} - u^{*(c)}_{2t}$, where $u^{*(c)}_{jt}$ denotes the $k^c \times 1$ vector containing the first k^c rows of $u^*_{jt} \equiv \left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)\frac{\tilde{\Lambda}'_j\varepsilon^*_{jt}}{\sqrt{N_j}}$.

Lemma B.5 *Suppose Conditions A*, B* and C* hold and assume that Assumption 1 is verified with $N = N_2 < N_1$. Defining $\mathcal{U}^*_t \equiv \mu_N u^{*(c)}_{1t} - u^{*(c)}_{2t}$, we have that $\hat{\Psi}^*_{cc} = -\frac{1}{TN}\sum_{t=1}^{T}\mathcal{U}^*_t\mathcal{U}^{*\prime}_t + O_{p^*}\left(\delta_{NT}^{-4}\right)$.*

Proof of Lemma B.3. We follow the proof of Lemma B.2 of AGGR(2019), but only consider a first order asymptotic expansion of $\hat{R}^*$. In particular, we write

$$\hat{R}^* = \hat{V}^{*-1}_{11}\hat{V}^*_{12}\hat{V}^{*-1}_{22}\hat{V}^*_{21} = (I_{k_1} + \ddot{V}^{*-1}_{11}\ddot{X}^*_{11})^{-1}\ddot{V}^{*-1}_{11}(\ddot{V}^*_{12} + \ddot{X}^*_{12})(I_{k_2} + \ddot{V}^{*-1}_{22}\ddot{X}^*_{22})^{-1}\ddot{V}^{*-1}_{22}(\ddot{V}^*_{21} + \ddot{X}^*_{21}),$$

where we used $\hat{V}^*_{jj} = \ddot{V}^*_{jj}(I_{k_j} + \ddot{V}^{*-1}_{jj}\ddot{X}^*_{jj})$. We then use the expansion $(I - X)^{-1} = I + X + O\left(X^2\right)$ to obtain $(I_{k_j} + \ddot{V}^{*-1}_{jj}\ddot{X}^*_{jj})^{-1} = I_{k_j} - \ddot{V}^{*-1}_{jj}\ddot{X}^*_{jj} + O_{p^*}\left(\delta^{-4}_{NT}\right)$, where only terms that are linear in $\ddot{X}^*_{jk}$ are larger than $O_{p^*}\left(\delta^{-4}_{NT}\right)$. Terms involving products or squares of $\ddot{X}^*_{jk}$ are of order $O_{p^*}\left(\delta^{-4}_{NT}\right)$ because we can show that $\ddot{X}^*_{jk}$ is of order $O_{p^*}\left(\delta^{-2}_{NT}\right)$ using Lemma B.3 of GP(2014). ■

Proof of Lemma B.4. Part (a): We follow closely the proof of Lemma B.4 of AGGR(2019), but rely on Assumption 1 and the following key features of the bootstrap DGP to simplify their proof. First, note that the eigenvector-eigenvalue problem associated with $\hat{R}^*$ is $\hat{R}^*\hat{W}^* = \hat{W}^*\hat{\Lambda}^*$, where $\hat{\Lambda}^* = diag\left(\hat{\rho}^{*2}_l : l = 1, \ldots, k^c\right)$. We can replace $\hat{R}^*$ from its asymptotic expansion in Lemma B.3:

$$\left[\left(H^{*\prime}_1\right)^{-1}\left(\tilde{R}^* + \hat{\Psi}^*\right)H^{*\prime}_1 + O_{p^*}\left(\delta^{-4}_{NT}\right)\right]\hat{W}^* = \hat{W}^*\hat{\Lambda}^*,$$

where we note that $\tilde{V}^{*-1}_{11} = I_{k_1}$ by construction. Pre-multiplying this equation by $H^{*\prime}_1$ gives

$$\left(\tilde{R}^* + \hat{\Psi}^*\right)\underbrace{H^{*\prime}_1\hat{W}^*}_{=\tilde{W}^*_1} = \underbrace{H^{*\prime}_1\hat{W}}_{=\tilde{W}^*_1}\hat{\Lambda}^* + O_p\left(\delta^{-4}_{NT}\right).$$

Note that

$$\tilde{R}^* = \begin{pmatrix} I_{k^c} & 0 \\ 0 & \tilde{R}^*_{ss} \end{pmatrix},$$

where $\tilde{R}^*_{ss} \equiv \tilde{\Sigma}^*_{12}\tilde{\Sigma}^*_{21}$, with $\tilde{\Sigma}^*_{12} \equiv T^{-1}\sum_{t=1}^T \hat{f}^s_{1t}\hat{f}^{s\prime}_{2t} = \tilde{\Sigma}^{*\prime}_{21}$. This follows by the definition of $\tilde{R}^* \equiv \tilde{V}^{*-1}_{11}\tilde{V}^*_{12}\tilde{V}^{*-1}_{22}\tilde{V}^*_{21}$ and the fact that $\tilde{V}^*_{jk} \equiv T^{-1}\sum_{t=1}^T \tilde{f}_{jt}\tilde{f}'_{kt}$, where $\tilde{f}_{jt} \equiv \left(\hat{f}^{c\prime}_t, \hat{f}^{s\prime}_{jt}\right)'$ for $j = 1, 2$, with $\hat{f}^c_t = \hat{W}'\hat{f}_{1t}$ and $\hat{f}^s_{jt}$ as defined in Definition 2 of AGGR(2019). As argued by AGGR(2019) (specifically their p. 1271),

$$T^{-1}\sum_{t=1}^T \hat{f}^c_t\hat{f}^{c\prime}_t = I_{k^c},\ \ T^{-1}\sum_{t=1}^T \hat{f}^c_t\hat{f}^{s\prime}_{jt} = 0,\ \text{ and } T^{-1}\sum_{t=1}^T \hat{f}^s_{jt}\hat{f}^{s\prime}_{jt} = I_{k^s_j},$$

which implies that for $j = 1, 2$,

$$\tilde{V}^*_{jj} \equiv T^{-1}\sum_{t=1}^T \tilde{f}_{jt}\tilde{f}'_{jt} = \begin{pmatrix} I_{k^c} & 0 \\ 0 & I_{k^s_j} \end{pmatrix} = I_{k_j}.$$

Compared to the matrix $\tilde{R}$ defined in Lemma B.3 of AGGR(2019), here $\tilde{R}^*_{cs}$, the upper-right block of $\tilde{R}^*$, is 0 due to the orthogonality between $\hat{f}^c_t$ and $\hat{f}^s_{jt}$ for both $j = 1, 2$. This in turn simplifies the form of $\tilde{R}^*_{ss}$ as compared to $\tilde{R}_{ss}$ in AGGR(2019). Importantly, the fact that $\tilde{R}^*$ is block diagonal implies that its first k^c eigenvalues are all equal to 1 (since they correspond to the eigenvalues of I_{k^c}), whereas its remaining k^s_1 eigenvalues are those of $\tilde{R}^*_{ss}$, which can be shown to be all smaller than one. This can be seen when f^s_{1t} and f^s_{2t} are both scalars, since then $\tilde{R}^*_{ss} = \hat{\Phi}^2$, where $\hat{\Phi} = T^{-1}\sum_{t=1}^T \hat{f}^s_{1t}\hat{f}^s_{2t}$ is the correlation between the two group specific factors. Moreover, the eigenvectors associated with the first k^c eigenvalues of $\tilde{R}^*$ are spanned by the columns of the matrix $E_c \equiv [I_{k^c}, 0']'$. Thus, letting $E_s \equiv \left(0', I_{k_1 - k^s_1}\right)$, and following AGGR(2019), we can decompose the eigenvector and eigenvalue matrices of $\hat{R}^*$ as

$$\tilde{W}_1^* = E_c\hat{U}^* + E_s\hat{\alpha}^* \qquad \text{and} \qquad \hat{\Lambda}^* = I_{k^c} + \hat{M}^*.$$

Following AGGR(2019), by Lemma B.3, $\hat{\alpha}^*$ and $\hat{M}^*$ converge to zero at rate $O_{p^*}\left(\delta_{NT}^{-2}\right)$. Thus, replacing $\tilde{W}_1^*$ and $\hat{\Lambda}^*$ into the eigenvector-eigenvalue equation for $\hat{R}^*$ gives:

$$\left(\tilde{R}^* + \hat{\Psi}^*\right)\left(E_c\hat{U}^* + E_s\hat{\alpha}^*\right) = \left(E_c\hat{U}^* + E_s\hat{\alpha}^*\right)\left(I_{k^c} + \hat{M}^*\right) + O_{p^*}\left(\delta_{NT}^{-4}\right), \text{ and}$$
$$\tilde{R}^*E_c\hat{U}^* + \hat{\Psi}^*E_c\hat{U}^* + \tilde{R}^*E_s\hat{\alpha}^* + \hat{\Psi}^*E_s\hat{\alpha}^* = E_c\hat{U}^* + E_c\hat{U}^*\hat{M}^* + E_s\hat{\alpha}^* + E_s\hat{\alpha}^*\hat{M}^* + O_p\left(\delta_{NT}^{-4}\right).$$

Using the fact that $\tilde{R}^*E_c = E_c$ and that $\hat{\Psi}^*E_s\hat{\alpha}^*$ and $E_s\hat{\alpha}^*\hat{M}^*$ are of order $O_{p^*}\left(\delta_{NT}^{-4}\right)$ implies that

$$\hat{\Psi}^*E_c\hat{U}^* + \tilde{R}^*E_s\hat{\alpha}^* = E_c\hat{U}^*\hat{M}^* + E_s\hat{\alpha}^* + O_{p^*}\left(\delta_{NT}^{-4}\right). \tag{13}$$

Pre-multiplying this equation (13) by E_c' gives

$$\underbrace{E_c'\tilde{R}^*E_s}_{\equiv\tilde{R}_{cs}^*=0}\hat{\alpha}^* + \underbrace{E_c'\hat{\Psi}^*E_c}_{\equiv\hat{\Psi}_{cc}^*}\hat{U}^* = \underbrace{E_c'E_c}_{\equiv I_{k^c}}\hat{U}^*\hat{M}^* + \underbrace{E_c'E_s}_{=0}\hat{\alpha}^* + O_{p^*}\left(\delta_{NT}^{-4}\right),$$

from which we obtain

$$\hat{M}^* = \hat{U}^{*-1}\hat{\Psi}_{cc}^*\hat{U}^* + O_{p^*}\left(\delta_{NT}^{-4}\right). \tag{14}$$

Expansion (14) is the bootstrap analogue of equation (C.62) in AGGR(2019)'s Online Appendix (proof of their Lemma B.4), where we have used the facts that $\tilde{R}_{cs}^* = 0$ and $\tilde{\Sigma}_{cc}^* \equiv T^{-1}\sum_{t=1}^T \hat{f}_t^c\hat{f}_t^{c\prime} = I_{k^c}$ to simplify the expansion in the bootstrap world. Equation (14) implies that

$$\hat{\Lambda}^* = I_{k^c} + \hat{M}^* = I_{k^c} + \hat{U}^{*-1}\hat{\Psi}_{cc}^*\hat{U}^* + O_{p^*}\left(\delta_{NT}^{-4}\right),$$

from which it follows that

$$\hat{\Lambda}^{*1/2} = I_{k^c} + \frac{1}{2}\hat{U}^{*-1}\hat{\Psi}_{cc}^*\hat{U}^* + O_{p^*}\left(\delta_{NT}^{-4}\right),$$

by using the expansion $(I+X)^{1/2} = I + \frac{1}{2}X + O_p\left(X^2\right)$ with $X = \hat{M}^*$. Part (b): This follows by taking the trace of $\hat{\Lambda}^{*1/2}$ and using the properties of the trace operator. ■

Proof of Lemma B.5. We replace $\hat{\Psi}_{cc}^*$ with the expression from Lemma B.3 and use Lemma B.5. In particular, recall that $\hat{\Psi}^*$ is defined as

$$\hat{\Psi}^* \equiv -\hat{X}_{11}^*\tilde{R}^* + \hat{X}_{12}^*\tilde{B}^* + \tilde{B}^{*\prime}\hat{X}_{21}^* - \tilde{B}^{*\prime}\hat{X}_{22}^*\tilde{B}^*,$$

where $\tilde{B}^* \equiv \tilde{V}_{22}^{*-1}\tilde{V}_{21}^* = \tilde{V}_{21}^*$ since $\tilde{V}_{22}^* = I_{k_2}$, and $\hat{X}_{jk}^*$ is as defined in Lemma B.3. Since the bootstrap DGP for each panel generates bootstrap observations on Y_j^* using $\tilde{f}_{jt} = \left(\hat{f}_t^{c\prime}, \hat{f}_{jt}^{s\prime}\right)'$, we can show that

$$\tilde{R}^* = \begin{pmatrix} I_{k^c} & 0 \\ 0 & \tilde{R}_{ss}^* \end{pmatrix} \text{ and } \tilde{B}^* = \begin{pmatrix} I_{k^c} & 0 \\ 0 & \tilde{\Sigma}_{21}^* \end{pmatrix},$$

where $\tilde{R}_{ss}^* = \tilde{\Sigma}_{12}^*\tilde{\Sigma}_{21}^*$, where $\tilde{\Sigma}_{12}^* \equiv T^{-1}\sum_{t=1}^T \hat{f}_{1t}^s\hat{f}_{2t}^{s\prime} = \tilde{\Sigma}_{21}^{*\prime}$. Thus, the upper-left $k^c \times k^c$ block $\hat{\Psi}_{cc}^*$ is equal to

$$\hat{\Psi}^*_{cc} = -\hat{X}^*_{11,cc} + \hat{X}^*_{12,cc} + \hat{X}^*_{21,cc} - \hat{X}^*_{22,cc},$$

as argued by AGGR(2019) (see their equation (C.69) in the Online Appendix). Given the expressions of $\hat{X}^*_{jk}$ in Lemma B.3 and the fact that $\tilde{f}_{jt} = \left(\hat{f}^{c\prime}_t, \hat{f}^{s\prime}_{jt}\right)'$, we can then use the same arguments of AGGR(2019) to rewrite the expression of $\hat{\Psi}^*_{cc}$ as

$$\begin{aligned}\hat{\Psi}^*_{cc} &= -\frac{1}{T}\sum_{t=1}^{T}\left(\psi^{*(c)}_{1t} - \psi^{*(c)}_{2t}\right)\left(\psi^{*(c)}_{1t} - \psi^{*(c)}_{2t}\right)' \\ &= -\left\{\frac{1}{T}\sum_{t=1}^{T}\psi^{*(c)}_{1t}\psi^{*(c)\prime}_{1t} - \frac{1}{T}\sum_{t=1}^{T}\psi^{*(c)}_{1t}\psi^{*(c)\prime}_{2t} - \frac{1}{T}\sum_{t=1}^{T}\psi^{*(c)}_{2t}\psi^{*(c)\prime}_{1t} + \frac{1}{T}\sum_{t=1}^{T}\psi^{*(c)}_{2t}\psi^{*(c)\prime}_{2t}\right\}\end{aligned}$$

where $\psi^{*(c)}_{jt}\psi^{*(c)\prime}_{kt}$ denotes the upper-left $k^c \times k^c$ block of the matrix $\psi^*_{jt}\psi^{*\prime}_{kt}$, where $\psi^*_{jt} \equiv H^{*-1}_j\left(\hat{f}^*_{jt} - H^*_j\tilde{f}_{jt}\right)$. For any $j,k \in \{1,2\}$, we can write

$$\frac{1}{T}\sum_{t=1}^{T}\psi^*_{jt}\psi^{*\prime}_{kt} = H^{*-1}_j\frac{1}{T}\sum_{t=1}^{T}\left(\hat{f}^*_{jt} - H^*_j\tilde{f}_{jt}\right)\left(\hat{f}^*_{kt} - H^*_k\tilde{f}_{kt}\right)'\left(H^{*\prime}_k\right)^{-1}.$$

The desired result follows by Lemma B.2, noting that $\mu_N = \sqrt{N_2/N_1}$, where $N \equiv \min(N_1, N_2) = N_2$ (without loss of generality), which implies the definition of $\mathcal{U}^*_t \equiv \mu_N u^{*(c)}_{1t} - u^{*(c)}_{2t}$. ■

Proof of Lemma 3.1. This follows from Lemmas B.3, B.4 and B.5 under Conditions A*-C*. ■

Proof of Theorem 3.1. The asymptotic Gaussianity of the bootstrap test statistic follows from Lemma 3.1 when we add Conditions D* and E*. To see that this implies that the bootstrap p-value converges in distribution to a uniform distribution under the null hypothesis, note that

$$\begin{aligned}p^* &\equiv P^*\left(N\sqrt{T}\left(\hat{\xi}^*(k^c) - k^c\right) \le N\sqrt{T}\left(\hat{\xi}(k^c) - k^c\right)\right) \\ &= P^*\left(\Omega_{\mathcal{U}}^{-1/2}N\sqrt{T}\left(\hat{\xi}^*(k^c) - k^c + \frac{\mathcal{B}}{2N}\right) \le \Omega_{\mathcal{U}}^{-1/2}N\sqrt{T}\left(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N}\right)\right) \\ &= \Phi\left(\Omega_{\mathcal{U}}^{-1/2}N\sqrt{T}\left(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N}\right)\right) + o_p(1).\end{aligned}$$

Since $\Omega_{\mathcal{U}}^{-1/2}N\sqrt{T}\left(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N}\right) \xrightarrow{d} N(0,1)$ under the null hypothesis, the random variable inside $\Phi(\cdot)$ in can be written as $\Phi^{-1}\left(U_{[0,1]}\right)$, implying that $p^* \xrightarrow{d} \Phi\left(\Phi^{-1}\left(U_{[0,1]}\right)\right) = U_{[0,1]}$. ■

Proof of Proposition 3.1. We can rewrite p^* as follows

$$\begin{aligned}p^* &= P^*\left(N\sqrt{T}(\hat{\xi}^*(k^c) - k^c + \frac{\mathcal{B}^*}{2N}) \le N\sqrt{T}(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}^*}{2N}) + \sqrt{T}(\mathcal{B}^* - \mathcal{B})\right) \\ &\overset{(1)}{=} P^*\left(-\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}^*_{N,t} \le N\sqrt{T}(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N}) + \sqrt{T}(\mathcal{B}^* - \mathcal{B})\right) + o_p(1) \\ &\overset{(2)}{=} P^*\left(-\frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}^*_{N,t} \le -N\sqrt{T}c_1 + N^{1-\epsilon}\sqrt{T}c_2\right) + o_p(1),\end{aligned}$$

where c_1 and c_2 are positive constants and ϵ is also positive. Note that (1) follows by Lemma 3.1

under Conditions A*- C*, whereas (2) follows by using the fact that under H_1, $\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N} = \sum_{l=1}^{k^c} \rho_l - k^c + o_p(1)$ (since $\mathcal{B} = O_p(1)$ and $\hat{\rho}_l \to_p \rho_l$), where ρ_l denotes the true canonical correlations. Since $\sum_{l=1}^{k^c} \rho_l - k^c < 0$ when there are less than k^c common factors, $N\sqrt{T}(\hat{\xi}(k^c) - k^c + \frac{\mathcal{B}}{2N}) \leq -N\sqrt{T}c_1$ for $c_1 > 0$ under H_1, as argued by AGGR(2019). Finally, we can bound $\sqrt{T}(\mathcal{B}^* - \mathcal{B})$ by $\sqrt{T}N^{1-\epsilon}c_2$ for some positive constant c_2 by using Condition F* and the fact that $\mathcal{B}^*$ and $\mathcal{B}$ are positive. Thus, $\sqrt{T}(\mathcal{B}^* - \mathcal{B})$ is asymptotically negligible with respect to $-N\sqrt{T}c_1$. This together with the fact that $\frac{1}{2\sqrt{T}} \sum_{t=1}^{T} \mathcal{Z}_{N,t}^*$ is $O_p(1)$ as assumed in Condition F* implies that $p^* \to_p 0$. ■

C Proof of wild bootstrap results in Section 4.1

In this appendix, we first provide three auxiliary lemmas, followed by their proofs. Then, we prove Theorem 4.1.

Lemma C.1 *Suppose Assumptions 1-4 hold. If either (1) $\{f_t^c\}$, $\{f_{jt}^s\}$ and $\{\varepsilon_{j,it}\}$ are mutually independent and for some $p \geq 2$, $E|\varepsilon_{j,it}|^{2p} \leq M < \infty$ and $E\|f_{jt}\|^{2p} \leq M < \infty$, or (2) for some $p \geq 2$, $E|\varepsilon_{j,it}|^{4p} \leq M < \infty$ and $E\|f_{jt}\|^{4p} \leq M < \infty$, it follows that*

(i) $\frac{1}{T}\sum_{t=1}^{T} \|\hat{f}_t^c - f_t^c\|^p = O_p(1)$, *and* $\frac{1}{N_j}\sum_{i=1}^{N_j} \|\hat{\lambda}_{j,i}^c - \lambda_{j,i}^c\|^p = O_p(1)$;

(ii) $\frac{1}{T}\sum_{t=1}^{T} \|\hat{f}_{jt}^s - H_j^s f_{jt}^s\|^p = O_p(1)$ *and* $\frac{1}{N_j}\sum_{i=1}^{N_j} \|\hat{\lambda}_{j,i}^s - (H_j^s)^{-1\prime} \lambda_{j,i}^s\|^p = O_p(1)$;

(iii) $\frac{1}{N_j T}\sum_{i=1}^{N_j}\sum_{t=1}^{T} |\tilde{\varepsilon}_{j,it}|^p = O_p(1)$,

where $H_j^s = (\mathcal{V}_j^s)^{-1} \frac{\hat{F}_j^{s\prime} F_j^s}{T} \frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}$ and $\mathcal{V}_j^s$ is the $k_j^s \times k_j^s$ diagonal matrix containing the k_j^s largest eigenvalues of $\Xi_j \Xi_j'/N_j T$ on the main diagonal in descending order.

Lemma C.2 *Assume that Assumptions 1-6 strengthened by Assumption WB1 and WB2 hold. Then Lemma 3.1 follows for Algorithm 1.*

Remark 6 *In Lemma C.2, we verify that the bootstrap method generated by Algorithm 1 satisfies Conditions A* through C*. To verify these conditions, we use Lemma C.1 which is valid under H_0 and H_1. Therefore, Lemma C.2 is satisfied regardless of the fact that either H_0 or H_1 is true.*

In the following Lemma C.3, we obtain the uniform expansions of the group common factors, factor loadings, group specific factors, and group specific factor loadings up to order $o_p(T^{-1/2})$ under H_0 to verify Condition D*. Note that Lemma C.3 is only valid under H_0.

Lemma C.3 *Assume that Assumptions 1-5 hold and H_0 is true. Then, for $j = 1, 2$, we have the following:*

(i) $\hat{f}_t^c = H^c(f_t^c + \frac{1}{\sqrt{N_1}} u_{1t}^{(c)}) + o_p(T^{-1/2})$;

(ii) $\hat{\lambda}^c_{j,i} = (H^c)^{-1\prime}\lambda^c_{j,i} + H^c\frac{1}{T}\sum_{t=1}^T f^c_t\varepsilon_{j,it} + H^c\frac{1}{T}\sum_{t=1}^T f^c_t f^{s\prime}_{jt}\lambda^s_{j,i} + o_p(T^{-1/2})$;

(iii) $\hat{f}^s_{jt} = \tilde{H}^s_j(\tilde{f}^s_{jt} + \frac{1}{\sqrt{N_j}}u^{(s)}_{jt}) + o_p(T^{-1/2})$;

(iv) $\hat{\lambda}^s_{j,i} = (\tilde{H}^s_j)^{-1\prime}\lambda^s_{j,i} + \tilde{H}^s_j\frac{1}{T}\sum_{t=1}^T \tilde{f}^s_{jt}\varepsilon_{j,it} + o_p(T^{-1/2})$,

where $\tilde{f}^s_{jt} \equiv f^s_{jt} - \tilde{\Sigma}_{j,c}\tilde{\Sigma}^{-1}_{cc}f^c_t$ *and* $\tilde{H}^s_j = (\mathcal{V}^s_j)^{-1}\frac{\hat{F}^{s\prime}_j\tilde{F}^s_j}{T}\frac{\Lambda^{s\prime}_j\Lambda^s_j}{N_j}$ *and* $\mathcal{V}^s_j$ *is defined in Lemma C.1.*

Proof of Lemma C.1. **Part (i)**: Recall that $\hat{f}^c_t = \hat{W}'\hat{f}_{1t}$, where $\hat{W}$ is s $k_1 \times k^c$ matrix collecting the eigenvectors of $\hat{R}$ associated to the k^c largest eigenvalues and $\hat{W}'\hat{W} = I_{k^c}$. By following Proposition 1 in AGGR(2019), $f^c_t = W'f_{1t}$, where W is a $k_1 \times k^c$ matrix of eigenvectors of R associated to the k^c largest eigenvalues. Then, by adding and subtracting appropriately, we can write $\hat{f}^c_t - f^c_t = \hat{W}'(\hat{f}_{1t} - H_1 f_{1t}) + (H_1'\hat{W} - W)'f_{1t}$. By the c_r-inequality, we can bound part (i) as follows,

$$\frac{1}{T}\sum_{t=1}^T \left\|\hat{f}^c_t - f^c_t\right\|^p \leq 2^{p-1}\left(\underbrace{\left\|\hat{W}\right\|^p}_{=O_p(1)}\frac{1}{T}\sum_{t=1}^T\left\|\hat{f}_{1t} - H_1 f_{1t}\right\|^p + \underbrace{\left\|\tilde{W}_1 - W\right\|^p}_{=O_p(1)}\frac{1}{T}\sum_{t=1}^T\|f_{1t}\|^p\right),$$

where we let $\tilde{W}_1 = H_1'\hat{W}$. It is sufficient to show that $\frac{1}{T}\sum_{t=1}^T\left\|\hat{f}_{1t} - H_1 f_{1t}\right\|^p = O_p(1)$. By following the arguments in GP(2014) (i.e., their Lemma C.1-(i)), given that $E|\varepsilon_{j,it}|^{2p} \leq M < \infty$ and $E\|f_{jt}\|^{2p} \leq M < \infty$, we have $\frac{1}{T}\sum_{t=1}^T\left\|\hat{f}_{1t} - H_1 f_{1t}\right\|^p = O_p(1)$. If we assume that $\{f_{jt}\}$ and $\{\varepsilon_{j,it}\}$ are independent, then $E\|f_{jt}\|^p \leq M < \infty$ and $E|\varepsilon_{j,it}|^{2p} \leq M < \infty$ are sufficient. Next, we show that $\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}^c_{j,i} - \lambda^c_{j,i}\|^p = O_p(1)$. Since $\hat{\Lambda}^c_j = \frac{1}{T}Y_j'\hat{F}^c$ and $Y_j = F^c\Lambda^{c\prime}_j + F^s_j\Lambda^{s\prime}_j + \varepsilon_j$, we can write $\hat{\Lambda}^c_j$ as follows,

$$\begin{aligned}
\hat{\Lambda}^c_j &= \frac{1}{T}Y_j'\hat{F}^c = \frac{1}{T}(F^c\Lambda^{c\prime}_j + F^s_j\Lambda^{s\prime}_j + \varepsilon_j)'\hat{F}^c \\
&= \frac{1}{T}\Lambda^c_j F^{c\prime}\hat{F}^c + \frac{1}{T}\Lambda^s_j F^{s\prime}_j\hat{F}^c + \frac{1}{T}\varepsilon_j'\hat{F}^c \\
&= \Lambda^c_j\underbrace{\frac{\hat{F}^{c\prime}\hat{F}^c}{T}}_{=I_{k^c}} - \frac{1}{T}\Lambda^c_j(\hat{F}^c - F^c)'\hat{F}^c + \frac{1}{T}\Lambda^s_j F^{s\prime}_j(\hat{F}^c - F^c) + \frac{1}{T}\Lambda^s_j F^{s\prime}_j F^c + \frac{1}{T}\varepsilon_j'(\hat{F}^c - F^c) + \frac{1}{T}\varepsilon_j' F^c.
\end{aligned}$$

Then, $\hat{\lambda}^c_{j,i} - \lambda^c_{j,i} = -\frac{1}{T}\hat{F}^{c\prime}(\hat{F}^c - F^c)\lambda^c_{j,i} + \frac{1}{T}(\hat{F}^c - F^c)'F^s_j\lambda^s_{j,i} + \frac{1}{T}F^{c\prime}F^s_j\lambda^s_{j,i} + \frac{1}{T}(\hat{F}^c - F^c)'\varepsilon_{j,i} + \frac{1}{T}F^{c\prime}\varepsilon_{j,i}$. We apply the c_r-inequality and show that each term is $O_p(1)$. In particular,

$$\begin{aligned}
\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\hat{\lambda}^c_{j,i} - \lambda^c_{j,i}\right\|^p \leq 5^{p-1}&\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}\hat{F}^{c\prime}(\hat{F}^c - F^c)\lambda^c_{j,i}\right\|^p + \frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}(\hat{F}^c - F^c)'F^s_j\lambda^s_{j,i}\right\|^p\right. \\
&\left.+\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}F^{c\prime}F^s_j\lambda^s_{j,i}\right\|^p + \frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}(\hat{F}^c - F^c)'\varepsilon_{j,i}\right\|^p + \frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}F^{c\prime}\varepsilon_{j,i}\right\|^p\right).
\end{aligned}$$

To see that the first term is bounded, note that

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|T^{-1}\hat{F}^{c\prime}(\hat{F}^c - F^c H^{c\prime})\lambda^c_{j,i}\right\|^p \leq \underbrace{\left(\left\|T^{-1/2}\hat{F}^c\right\|^p\right)}_{=(k^c)^{p/2}}\underbrace{\left(\left\|T^{-1/2}(\hat{F}^c - F^c H^{c\prime})\right\|^p\right)}_{=O_p(1)}\underbrace{\frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda^c_{j,i}\|^p}_{=O(1)} = O_p(1).$$

Similarly, for the second term,

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}(\hat{F}^c - F^c H^{c\prime})' F_j^s \lambda_{j,i}^s\right\|^p \le \left\|T^{-1/2}F_j^s\right\|^p \underbrace{\left\|T^{-1/2}(\hat{F}^c - F^c H^{c\prime})\right\|^p}_{=O_p(1)} \underbrace{\frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^s\|^p}_{=O(1)} = O_p(1),$$

where we can show $\left\|T^{-1/2}F_j^s\right\|^p = \left(\frac{1}{T}\|F_j^s\|^2\right)^{p/2} \le \frac{1}{T}\sum_{t=1}^T \|f_{jt}^s\|^p \le M$, provided $E\|f_{jt}^s\|^p \le M < \infty$. We can bound the third term as $\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}F^{c\prime}F_j^s\lambda_{j,i}^s\right\|^p \le \frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^s\|^p\left\|\frac{1}{T}F^{c\prime}F_j^s\right\|^p$, ignoring $\|H^c\|^p = \|\hat{U}\|^p = O_p(1)$. Given $\|\lambda_{j,i}^s\|^p \le M$, it suffices to show that $\left\|\frac{1}{T}F^{c\prime}F_j^s\right\|^p = O_p(1)$. By Markov's inequality, this follows from $E\left\|\frac{1}{T}\sum_{t=1}^T f_t^c f_{jt}^{s\prime}\right\|^p \le \frac{1}{T}\sum_{t=1}^T E\|f_t^c f_{jt}^{s\prime}\|^p$, which is bounded given $E\|f_{jt}\|^{2p} \le M < \infty$, if f_t^c and f_{jt}^s are not independent (otherwise, $E\|f_{jt}\|^p \le M < \infty$ is sufficient). The fourth term can be bounded as $\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}(\hat{F}^c - F^c H^{c\prime})'\varepsilon_{j,i}\right\|^p \le \|T^{-1/2}(\hat{F}^c - F^c H^{c\prime})\|^p \frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1/2}\varepsilon_{j,i}\|^p = O_p(1)$, where

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1/2}\varepsilon_{j,i}\|^p = \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\left\|T^{-1/2}\varepsilon_{j,i}\right\|^2\right)^{p/2} = \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\frac{1}{T}\sum_{t=1}^T \varepsilon_{j,it}^2\right)^{p/2} \le \frac{1}{N_j T}\sum_{i=1}^{N_j}\sum_{t=1}^T |\varepsilon_{j,it}|^p = O_p(1),$$

given $E|\varepsilon_{j,it}|^p \le M < \infty$. Similarly, we can bound the last term as $\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}F^{c\prime}\varepsilon_{j,i}\|^p = \|H^c\|^p\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1/2}\varepsilon_{j,i}\|^p\right)\left(\|T^{-1/2}F^c\|^p\right) = O_p(1)$, given $E|\varepsilon_{j,it}|^p \le M < \infty$ and $E\|f_t^c\|^p \le M < \infty$.

Part (ii): Note that $\hat{f}_{jt}^s$ is the principal component estimator from $\Xi_{jt} = y_{jt} - \hat{\Lambda}_j^c \hat{f}_t^c$. By using the fact that $y_{jt} = \Lambda_j^c f_t^c + \Lambda_j^s f_{jt}^s + \varepsilon_{jt}$, we can write Ξ_{jt} as follows.

$$\Xi_{jt} = y_{jt} - \Lambda_j^c f_t^c + \Lambda_j^c f_t^c - \hat{\Lambda}_j^c \hat{f}_t^c = \Lambda_j^s f_{jt}^s + \underbrace{\varepsilon_{jt} + (\Lambda_j^c f_t^c - \hat{\Lambda}_j^c \hat{f}_t^c)}_{\equiv e_{jt}} = \Lambda_j^s f_{jt}^s + e_{jt}.$$

Then, using the identity from the proof of Theorem 1 in Bai (2003), we have

$$\hat{f}_{jt}^s - H_j^s f_{jt}^s = (\mathcal{V}_j^s)^{-1}\left(\frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \psi_{j,lt}^s + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \eta_{j,lt}^s + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \xi_{j,lt}^s\right),$$

where $\mathcal{V}_j^s$ is the $k_j^s \times k_j^s$ matrix of k_j^s eigenvalues of $\Xi_j \Xi_j'/(TN_j)$ in its diagonal elements and $\psi_{j,lt}^s = \frac{1}{N_j}\sum_{i=1}^{N_j} e_{j,il}e_{j,it}$, $\eta_{j,lt}^s = \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^{s\prime} f_{jl}^s e_{j,it}$, and $\xi_{j,lt}^s = \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^{s\prime} f_{jt}^s e_{j,il}$. Using this identity and the c_r-inequality, we have

$$\frac{1}{T}\sum_{t=1}^T \|\hat{f}_{jt}^s - H_j^s f_{jt}^s\|^p \le 3^{p-1}\|(\mathcal{V}_j^s)^{-1}\|^p\left(\frac{1}{T}\sum_{t=1}^T a_t^s + \frac{1}{T}\sum_{t=1}^T b_t^s + \frac{1}{T}\sum_{t=1}^T c_t^s\right),$$

where $a_t^s = \frac{1}{T^p}\left\|\sum_{l=1}^T \hat{f}_{jl}^s \psi_{j,lt}^s\right\|^p$, $b_t^s = \frac{1}{T^p}\left\|\sum_{l=1}^T \hat{f}_{jl}^s \eta_{j,lt}^s\right\|^p$, and $c_t^s = \frac{1}{T^p}\left\|\sum_{l=1}^T \hat{f}_{jl}^s \xi_{j,lt}^s\right\|^p$. Let $\chi_{j,lt}$ denote either $\psi_{j,lt}^s$, $\eta_{j,lt}^s$, or $\xi_{j,lt}^s$. Then, we can show $\left\|\sum_{l=1}^T \hat{f}_{jl}^s \chi_{j,lt}\right\|^p = \left(\left\|\sum_{l=1}^T \hat{f}_{jl}^s \chi_{j,lt}\right\|^2\right)^{1/2} \le$

$\left(\sum_{l=1}^{T}\|\hat{f}_{jl}^{s}\|^2\sum_{l=1}^{T}|\chi_{j,lt}|^2\right)^{p/2}$. Under this inequality, we can show that

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T^p}\left\|\sum_{l=1}^{T}\hat{f}_{jl}^{s}\chi_{j,lt}\right\|^p \leq (k_j^s)^{p/2}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{l=1}^{T}|\chi_{j,lt}|^2\right)^{p/2} \leq (k_j^s)^{p/2}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{l=1}^{T}|\chi_{j,lt}|^p,$$

where we use the fact that $\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_{jt}^{s}\|^2 = k_j^s$. It suffices to show that $\frac{1}{T^2}\sum_{t=1}^{T}\sum_{l=1}^{T}|\chi_{j,lt}|^p = O_p(1)$. Starting with $\chi_{j,lt} = \psi_{j,lt}^{s}$, we can write as

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{l=1}^{T}\left|\frac{1}{N_j}\sum_{i=1}^{N_j}e_{j,il}e_{j,it}\right|^p &\leq \frac{1}{T^2N_j}\sum_{t=1}^{T}\sum_{l=1}^{T}\sum_{i=1}^{N_j}|e_{j,il}e_{j,it}|^p \\ &\leq 4^{p-1}\frac{1}{T^2N_j}\sum_{t=1}^{T}\sum_{l=1}^{T}\sum_{i=1}^{N_j}\left(|\varepsilon_{j,it}\varepsilon_{j,il}|^p + |\varepsilon_{j,it}\hat{c}_{j,il}|^p + |\hat{c}_{j,it}\varepsilon_{j,il}|^p + |\hat{c}_{j,it}\hat{c}_{j,il}|^p\right),\end{aligned}$$

where we let $e_{j,it} = \varepsilon_{j,it} + \hat{c}_{j,it}$, with $\hat{c}_{j,it} \equiv \lambda_{j,i}^{c\prime}f_t^c - \hat{\lambda}_{j,i}^{c\prime}\hat{f}_t^c$. Using the c_r-inequality, we can show that $\frac{1}{T^2N_j}\sum_{t=1}^{T}\sum_{l=1}^{T}\sum_{i=1}^{N_j}|\varepsilon_{j,it}\varepsilon_{j,il}|^p \leq \frac{1}{T^2N_j}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|\varepsilon_{j,it}|^{2p} = O_p(1)$ given that $E|\varepsilon_{j,it}|^{2p} \leq M < \infty$. For the second term, it suffices to show that $\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^{T}|\hat{c}_{j,it}|^{2p} = O_p(1)$, because $\frac{1}{T^2N_j}\sum_{t=1}^{T}\sum_{l=1}^{T}\sum_{i=1}^{N_j}|\varepsilon_{j,it}\hat{c}_{j,il}|^p$ $\left(\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}|\varepsilon_{j,il}|^{2p}\right)^{1/2}\left(\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}|\hat{c}_{j,it}|^{2p}\right)^{1/2}$ by using the c_r- and Cauchy-Schwarz inequalities. Using the definition of $\hat{c}_{j,it}$, we have

$$\begin{aligned}\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}|\hat{c}_{j,it}|^{2p} &= \frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}\left|-(\hat{\lambda}_{j,i}^{c} - \lambda_{j,i}^{c})'\hat{f}_t^c - \lambda_{j,i}^{c\prime}(\hat{f}_t^c - f_t^c)\right|^{2p} \\ &\leq 2^{2p-1}\left(\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}\left|(\hat{\lambda}_{j,i}^{c} - \lambda_{j,i}^{c})'\hat{f}_t^c\right|^{2p} + \frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{l=1}^{T}\left|\lambda_{j,i}^{c\prime}(\hat{f}_t^c - f_t^c)\right|^{2p}\right).\end{aligned}$$

To show that this term is $O_p(1)$, it suffices that $\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^{c} - \lambda_{j,i}^{c}\|^{2p} = O_p(1)$, and $\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t^c - f_t^c\|^{2p} = O_p(1)$, given that $\|\lambda_{j,i}^{c}\|^{2p} \leq M$. Assuming f_t^c and f_{jt}^s are independent, provided that $E\|f_{jt}\|^{2p} \leq M$ and $\|\lambda_{j,i}^{s}\|^{2p} \leq M$, we have $\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^{c} - \lambda_{j,i}^{c}\|^{2p} = O_p(1)$ (otherwise, we need $E\|f_{jt}\|^{4p} \leq M$). Assuming that f_{jt} and $\varepsilon_{j,it}$ are independent, given that $\|\lambda_{j,i}^{c}\|^{2p} \leq M$, $E|\varepsilon_{j,it}|^{2p} \leq M$ and $E\|f_{jt}\|^{2p} \leq M$, we have $\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t^c - f_t^c\|^{2p} = O_p(1)$ (otherwise, we need $E|\varepsilon_{j,it}|^{4p} \leq M$ and $E\|f_{jt}\|^{4p} \leq M$). The remaining terms can be handled similarly by using $\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^{T}|\varepsilon_{j,it}|^{2p} = O_p(1)$ and $\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^{T}|\hat{c}_{j,it}|^{2p} = O_p(1)$. For instance, letting $\chi_{j,lt} = \eta_{j,lt}^{s}$ and assuming that f_{jt}

and $\varepsilon_{j,it}$ are independent, we have that

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T}\sum_{l=1}^{T}\left|\eta_{j,lt}^{s}\right|^p &\le \frac{1}{N_jT^2}\sum_{t=1}^{T}\sum_{l=1}^{T}\sum_{i=1}^{N_j}\left|\lambda_{j,i}^{s\prime}f_{jl}^{s}e_{j,it}\right|^p \\
&\le \frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^{s}\|^p\left(\frac{1}{T}\sum_{t=1}^{T}|e_{j,it}|^p\right)\left(\frac{1}{T}\sum_{l=1}^{T}\|f_{jl}^{s}\|^p\right) \\
&\le \left(\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\lambda_{j,i}^{s}\right\|^{2p}\right)^{1/2}\left(\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|e_{j,it}|^{2p}\right)^{1/2}\left(\frac{1}{T}\sum_{l=1}^{T}\|f_{jl}^{s}\|^p\right),
\end{aligned}$$

which is $O_p(1)$ by showing that $\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|e_{j,it}|^{2p}=O_p(1)$ given that $\|\lambda_{j,i}^{s}\|^{2p}\le M$ and $E\|f_{jl}^{s}\|\le M$. To show that $\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|e_{j,it}|^{2p}=O_p(1)$, it is sufficient to have $\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|\varepsilon_{j,it}|^{2p}=O_p(1)$ and $\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}|\hat{c}_{j,it}|^{2p}=O_p(1)$. We can use a similar argument when $\chi_{j,lt}=\xi_{j,lt}^{s}$.

Next, we show that $\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^{s}-(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}\|^p=O_p(1)$. Note that $\hat{\Lambda}_j^{s}=\frac{1}{T}\Xi_j^{\prime}\hat{F}_j^{s}$ and $\Xi_j=F_j^{s}\Lambda_j^{s\prime}+e_j$, where $e_j=\varepsilon_j+(F^c\Lambda_j^{c\prime}-\hat{F}^c\hat{\Lambda}_j^{c\prime})$. Then, we can write $\hat{\lambda}_{j,i}^{s}$ as follows,

$$\begin{aligned}
\hat{\lambda}_{j,i}^{s} &= \frac{\hat{F}_j^{s\prime}F_j^{s}}{T}\lambda_{j,i}^{s}+\frac{\hat{F}_j^{s\prime}e_{j,i}}{T} \\
&= \frac{\hat{F}_j^{s\prime}F_j^{s}H_j^{s\prime}}{T}(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}+\frac{\hat{F}_j^{s\prime}e_{j,i}}{T} \\
&= (H_j^{s\prime})^{-1}\lambda_{j,i}^{s}-\frac{\hat{F}_j^{s\prime}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})}{T}(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}+T^{-1}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})^{\prime}e_{j,i}+T^{-1}H_j^{s}F_j^{s\prime}e_{j,i}.
\end{aligned}$$

Under this identity and the c_r-inequality, we can bound $\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^{s}-(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}\|^p$ by

$$\begin{aligned}
3^{p-1}\Bigg(&\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}\hat{F}_j^{s\prime}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}\|^p \\
&+\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})^{\prime}e_{j,i}\|^p+\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}(F_j^{s}H_j^{s\prime})^{\prime}e_{j,i}\|^p\Bigg).
\end{aligned}$$

The first term is $O_p(1)$ since $\|\lambda_{j,i}^{s}\|^p\le M<\infty$ and

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}\hat{F}_j^{s\prime}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})(H_j^{s\prime})^{-1}\lambda_{j,i}^{s}\|^p\le\|T^{-1/2}\hat{F}_j^{s}\|^p\|T^{-1/2}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})\|^p\|(H_j^{s\prime})^{-1}\|^p\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^{s}\|^p\right).$$

For the second term, we have

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})^{\prime}e_{j,i}\|^p\le\underbrace{\|T^{-1/2}(\hat{F}_j^{s}-F_j^{s}H_j^{s\prime})\|^p}_{=O_p(1)}\underbrace{\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1/2}e_{j,i}\|^p\right)}_{=O_p(1)}=O_p(1).$$

Similarly, the third term can be bounded as

$$\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1}(F_j^s H_j^{s\prime})' e_{j,i}\|^p \leq \underbrace{\|T^{-1/2}F_j^s\|^p}_{\substack{\leq \frac{1}{T}\sum_{t=1}^T \|f_{jt}^s\|^p \\ =O_p(1)}} \|H_j^s\|^p \underbrace{\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|T^{-1/2}e_{j,i}\|^p\right)}_{=O_p(1)} = O_p(1),$$

given that $E\|f_{jt}^s\|^p \leq M < \infty$.

Part (iii): To show $\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T |\tilde{\varepsilon}_{j,it}|^p = O_p(1)$, we first rewrite $\tilde{\varepsilon}_{j,it}$ as follows.

$$\begin{aligned}
\tilde{\varepsilon}_{j,it} &= y_{j,it} - \hat{\lambda}_{j,i}^{c\prime}\hat{f}_t^c - \hat{\lambda}_{j,i}^{s\prime}\hat{f}_{jt}^s \\
&= \varepsilon_{j,it} + (\lambda_{j,i}^{c\prime} f_t^c - \hat{\lambda}_{j,i}^{c\prime}\hat{f}_t^c) + (\lambda_{j,i}^{s\prime} f_{jt}^s - \hat{\lambda}_{j,i}^{s\prime}\hat{f}_{jt}^s) \\
&= \varepsilon_{j,it} - \lambda_{j,i}^{c\prime}(\hat{f}_t^c - H^c f_t^c) - (\hat{\lambda}_{j,i}^c - (H^c)^{-1\prime}\lambda_{j,i}^c)'\hat{f}_t^c - \lambda_{j,i}^{s\prime}(H_j^s)^{-1}(\hat{f}_{jt}^s - H_j^s f_{jt}^s) - (\hat{\lambda}_{j,i}^s - (H_j^s)^{-1\prime}\lambda_{j,i}^s)'\hat{f}_{jt}^s.
\end{aligned}$$

Using the identity above and the c_r-inequality, we have

$$\begin{aligned}
\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T |\tilde{\varepsilon}_{j,it}|^p \leq 5^{p-1}\Bigg(&\underbrace{\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T |\varepsilon_{j,it}|^p}_{\equiv(a)} + \underbrace{\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T \left|\lambda_{j,i}^{c\prime}(\hat{f}_t^c - H^c f_t^c)\right|^p}_{\equiv(b)} \\
&+ \underbrace{\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T \left|(\hat{\lambda}_{j,i}^c - (H^c)^{-1\prime}\lambda_{j,i}^c)'\hat{f}_t^c\right|^p}_{\equiv(c)} + \underbrace{\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T \left|\lambda_{j,i}^{s\prime}(H_j^s)^{-1}(\hat{f}_{jt}^s - H_j^s f_{jt}^s)\right|^p}_{\equiv(d)} \\
&+ \underbrace{\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T \left|(\hat{\lambda}_{j,i}^s - (H_j^s)^{-1\prime}\lambda_{j,i}^s)'\hat{f}_{jt}^s\right|^p}_{\equiv(e)}\Bigg).
\end{aligned}$$

To end the proof, we show that (a) through (e) are $O_p(1)$. The fact that (a) is $O_p(1)$ follows from $E|\varepsilon_{j,it}|^p \leq M < \infty$. The term (b) can be bounded by $\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^c\|^p\right)\left(\frac{1}{T}\sum_{t=1}^T\|\hat{f}_t^c - H^c f_t^c\|^p\right) = O_p(1)$ using part (i) and $\|\lambda_{j,i}^c\|^p \leq M$. We can also bound the term (c) as

$$\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^T \left|(\hat{\lambda}_{j,i}^c - (H^c)^{-1\prime}\lambda_{j,i}^c)'\hat{f}_t^c\right|^p \leq \underbrace{\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^c - (H^c)^{-1\prime}\lambda_{j,i}^c\|^p\right)}_{=O_p(1)\text{ by part (ii)}}\left(\frac{1}{T}\sum_{t=1}^T\|\hat{f}_t^c\|^p\right),$$

where

$$\frac{1}{T}\sum_{t=1}^T\|\hat{f}_t^c\|^p \leq 2^{p-1}\left(\|H^c\|^p \underbrace{\frac{1}{T}\sum_{t=1}^T\|f_t^c\|^p}_{\substack{=O_p(1) \\ \text{by } E\|f_t^c\|\leq M<\infty}} + \underbrace{\frac{1}{T}\sum_{t=1}^T\|\hat{f}_t^c - H^c f_t^c\|^p}_{=O_p(1)\text{ by part (i)}}\right) = O_p(1).$$

The term (d) can be bounded by $\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|\lambda_{j,i}^s\|^p\right)\|(H_j^s)^{-1}\|^p\left(\frac{1}{T}\sum_{t=1}^T\|\hat{f}_{jt}^s - H_j^s f_{jt}^s\|^p\right) = O_p(1)$, by part (i) and $\|\lambda_{j,i}^s\|^p \leq M < \infty$. Finally, the last term can be bounded as follows,

$$\frac{1}{N_jT}\sum_{i=1}^{N_j}\sum_{t=1}^{T}\left|(\hat{\lambda}_{j,i}^s - (H_j^s)^{-1\prime}\lambda_{j,i}^s)'\hat{f}_{jt}^s\right|^p \leq \underbrace{\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\|\hat{\lambda}_{j,i}^s - (H_j^s)^{-1\prime}\lambda_{j,i}^s\|^p\right)}_{=O_p(1)\text{ by part (ii)}}\left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{f}_{jt}^s\|^p\right),$$

where

$$\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_{jt}^s\|^p \leq 2^{p-1}\left(\|H_j^s\|^p\underbrace{\frac{1}{T}\sum_{t=1}^{T}\|f_{jt}^s\|^p}_{=O_p(1)} + \underbrace{\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_{jt}^s - H_j^s f_{jt}^s\|^p}_{=O_p(1)\text{ by part (ii)}}\right) = O_p(1).$$

■

Proof of Lemma C.2.

As argued in Remark 3, Condition A* - B* are verified for the wild bootstrap in GP(2014) (for details, see their proof of Theorem 4.1). Therefore, we focus on Condition C*. Part (i): By Cauchy-Schwarz inequality, we can show that

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}\tilde{f}_{js}\gamma_{j,st}^*\right\|^2 \leq \left(\frac{1}{T}\sum_{s=1}^{T}\|\tilde{f}_{js}\|^4\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma_{j,st}^*|^4\right)^{1/2} = O_p(1).$$

Since we have $\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{f}_{js}\right\|^4 = O_p(1)$, by applying Lemma C.1 with $p = 4$, it is sufficient to show that $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma_{j,st}^*|^4 = O_p(1)$. Noting that $\gamma_{j,st}^* = 0$ for $s \neq t$ and using c_r-inequality, we have that

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}|\gamma_{j,st}^*|^4 = \frac{1}{T}\sum_{t=1}^{T}|\gamma_{j,tt}^*|^4 = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N_j}\sum_{i=1}^{N_j}\tilde{\varepsilon}_{j,it}^2\right|^4 \leq \frac{1}{TN_j}\sum_{t=1}^{T}\sum_{i=1}^{N_j}\tilde{\varepsilon}_{j,it}^8 = O_p(1).$$

Lemma C.1 with $p = 8$ implies the last equality. Part (ii): By letting $m_{jk,s}^* \equiv \sum_{t=1}^{T}\gamma_{j,st}^*\frac{\tilde{\Lambda}_k\varepsilon_{kt}^{*\prime}}{\sqrt{N_k}}$, we can write the sufficient condition for part (ii) to be $O_p(1)$ as follows.

$$\begin{aligned}
E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}m_{jk,s}^{*\prime}\right\|^2 &= E^*\left\{tr\left[\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\tilde{f}_{js}m_{jk,s}^{*\prime}\right)'\left(\frac{1}{\sqrt{T}}\sum_{l=1}^{T}\tilde{f}_{jl}m_{jk,l}^{*\prime}\right)\right]\right\}\\
&= E^*\left\{\frac{1}{T}\sum_{s=1}^{T}\sum_{l=1}^{T}\tilde{f}_{js}'\tilde{f}_{jl}m_{jk,l}^{*\prime}m_{jk,s}^*\right\}\\
&= \frac{1}{T}\sum_{s=1}^{T}\sum_{l=1}^{T}\tilde{f}_{js}'\tilde{f}_{jl}E^*(m_{jk,l}^{*\prime}m_{jk,s}^*),
\end{aligned}$$

where

$$E^*(m_{jk,l}^{*\prime}m_{jk,s}^*) = E^*\left[\left(\sum_{t_1=1}^{T}\gamma_{j,lt_1}^*\frac{\varepsilon_{kt_1}^{*\prime}\tilde{\Lambda}_k}{\sqrt{N_k}}\right)\left(\sum_{t_2=1}^{T}\gamma_{j,st_2}^*\frac{\varepsilon_{kt_2}^{*\prime}\tilde{\Lambda}_k}{\sqrt{N_k}}\right)\right] = \frac{1}{N_k}\sum_{t=1}^{T}\sum_{i=1}^{N_k}\tilde{\lambda}_{k,i}'\tilde{\lambda}_{k,i}\tilde{\varepsilon}_{k,it}^2\gamma_{j,lt}^*\gamma_{j,st}^*.$$

By noting that $\gamma^*_{j,st} = 0$ if $s \neq t$, it follows that

$$\frac{1}{TN_k}\sum_{s=1}^{T}\sum_{l=1}^{T}\tilde{f}'_{js}\tilde{f}_{jl}E^*(m^{*\prime}_{jk,l}m^*_{jk,s}) = \frac{1}{TN_k}\sum_{t=1}^{T}\sum_{i=1}^{N_k}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}\tilde{\varepsilon}^2_{k,it}\tilde{f}'_{jt}\tilde{f}_{jt}(\gamma^*_{j,tt})^2$$

$$\leq \left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{N_k}\sum_{i=1}^{N}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}\tilde{\varepsilon}^2_{k,it}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}|\tilde{f}'_{jt}\tilde{f}_{jt}(\gamma^*_{j,tt})^2|^2\right)^{1/2} = O_p(1),$$

where the first parenthesis can be bounded as

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{N_k}\sum_{i=1}^{N}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}\tilde{\varepsilon}^2_{k,it}\right\|^2 \leq \left(\frac{1}{TN_k}\sum_{t=1}^{T}\sum_{i=1}^{N_k}\tilde{\varepsilon}^4_{k,it}\right)\left(\frac{1}{N_k}\sum_{i=1}^{N_k}|\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}|^2\right)$$

$$\leq \left(\frac{1}{TN_k}\sum_{t=1}^{T}\sum_{i=1}^{N_k}\tilde{\varepsilon}^4_{k,it}\right)\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\|\tilde{\lambda}_{k,i}\|^4\right) = O_p(1),$$

given Lemma C.1 with $p = 4$. By Cauchy-Schwarz inequality, we can also bound the second parenthesis as follows.

$$\frac{1}{T}\sum_{t=1}^{T}|\tilde{f}'_{jt}\tilde{f}_{jt}(\gamma^*_{j,tt})^2|^2 \leq \left(\frac{1}{T}\sum_{t=1}^{T}|\tilde{f}'_{jt}\tilde{f}_{jt}|^4\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}(\gamma^*_{j,tt})^8\right)^{1/2}$$

$$\leq \left(\frac{1}{T}\sum_{t=1}^{T}\|\tilde{f}_{jt}\|^8\right)^{1/2}\left(\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}\tilde{\varepsilon}^{16}_{j,it}\right)^{1/2} = O_p(1),$$

where we apply Lemma C.1 with $p = 16$ to obtain $\frac{1}{N_jT}\sum_{t=1}^{T}\sum_{i=1}^{N_j}\tilde{\varepsilon}^{16}_{j,it} = O_p(1)$. Part (iii): We rewrite the term as follows,

$$\frac{1}{T}\sum_{s=1}^{T}E^*\left\|\sum_{t=1}^{T}\gamma^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\right\|^2 = \frac{1}{T}\sum_{s=1}^{T}E^*\left[\left(\sum_{t=1}^{T}\gamma^*_{j,st}\frac{\varepsilon^{*\prime}_{kt}\tilde{\Lambda}_k}{\sqrt{N_k}}\right)\left(\sum_{l=1}^{T}\gamma^*_{j,sl}\frac{\varepsilon^{*\prime}_{kl}\tilde{\Lambda}_k}{\sqrt{N_k}}\right)'\right]$$

$$= \frac{1}{T}\sum_{s=1}^{T}\left[\sum_{t=1}^{T}\sum_{l=1}^{T}\gamma^*_{j,st}\gamma^*_{j,sl}E^*\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\sum_{m=1}^{N_k}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,m}\varepsilon^*_{k,it}\varepsilon^*_{k,ml}\right)\right]$$

$$= \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma^*_{j,st}\gamma^*_{j,st}\frac{1}{N_k}\sum_{i=1}^{N_k}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}\tilde{\varepsilon}^2_{k,it}$$

$$= \frac{1}{T}\sum_{t=1}^{T}(\gamma^*_{j,tt})^2\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\tilde{\lambda}'_{k,i}\tilde{\lambda}_{k,i}\tilde{\varepsilon}^2_{k,it}\right)$$

$$\leq \left(\underbrace{\frac{1}{T}\sum_{t=1}^{T}|\gamma^*_{j,tt}|^4}_{=O_p(1)\text{ by part(i)}}\right)^{1/2}\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\|\tilde{\lambda}_{k,i}\|^4\right)^{1/2}\left(\frac{1}{TN_k}\sum_{t=1}^{T}\sum_{i=1}^{N_k}\tilde{\varepsilon}^4_{k,it}\right)^{1/2} = O_p(1),$$

where the third equality follows since $E^*(\varepsilon^*_{k,it}\varepsilon^*_{k,ml}) = \tilde{\varepsilon}_{k,it}\tilde{\varepsilon}_{k,ml}E^*(\eta_{k,it}\eta_{k,ml}) = 0$ if either $i \neq m$ or $t \neq l$ and the fourth equality follows since $\gamma^*_{j,st} = 0$ for $s \neq t$. To verify part (iv), a sufficient condition is that $E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T \tilde{f}_{js}\frac{1}{T}\sum_{t=1}^T\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \tilde{\lambda}_{k,i}\varepsilon^*_{k,it}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))\right)\right\|^2 = O_p(1)$. To simplify the notation, we define $\psi^*_{1,jk,st} = \frac{1}{\sqrt{N}}\sum_{i=1}^N \tilde{\lambda}_{k,i}\varepsilon^*_{k,it}(\varepsilon^*_{j,is}\varepsilon^*_{j,it} - E^*(\varepsilon^*_{j,is}\varepsilon^*_{j,it}))$.

$$
\begin{aligned}
E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T \tilde{f}_{js}\frac{1}{T}\sum_{t=1}^T \psi^*_{1,jk,st}\right\|^2 &= \frac{1}{T}\sum_{s=1}^T\sum_{l=1}^T \tilde{f}'_{js}\tilde{f}_{jl}\frac{1}{T^2}\sum_{t=1}^T\sum_{q=1}^T E^*(\psi^{*\prime}_{1,jk,st}\psi^*_{1,jk,lq}) \\
&= \frac{1}{T}\sum_{s=1}^T\sum_{l=1}^T \tilde{f}'_{js}\tilde{f}_{jl}\frac{1}{T^2}\sum_{t=1}^T\sum_{q=1}^T \frac{1}{N}\sum_{i_1=1}^N\sum_{i_2=1}^N \tilde{\lambda}'_{k,i_1}\tilde{\lambda}_{k,i_2} \\
&\quad \underbrace{E^*[\varepsilon^*_{k,i_1 s}\varepsilon^*_{k,i_2 q}(\varepsilon^*_{j,i_1 s}\varepsilon^*_{j,i_1 t} - E^*(\varepsilon^*_{j,i_1 s}\varepsilon^*_{j,i_1 t}))(\varepsilon^*_{j,i_2 l}\varepsilon^*_{j,i_2 q} - E^*(\varepsilon^*_{j,i_2 l}\varepsilon^*_{j,i_2 q}))]}_{\equiv X_1}.
\end{aligned}
$$

We simplify the expression of X_1 depending on the choices of j, k, i_1, i_2, and s, t, q, and l. To simplify the notation, we let $j = k$ and ignore the group notation (if $j \neq k$, under the group independence, the proof is simpler). If $i_1 \neq i_2$, we have $X_1 = \tilde{\varepsilon}^3_{i_1 t}\tilde{\varepsilon}^3_{i_2 l}E^*(\eta^3_{i_1 t})E^*(\eta^3_{i_1 l}) = 0$, when $s = t \neq l = q$ or $s = t = l = q$, since $\eta_{it} \sim$ i.i.d.$N(0,1)$. Therefore, we only need to consider the case of $i_1 = i_2 (= i)$. For this case, X_1 takes a non-zero value for three different cases: $s = l \neq t \neq q$ ($X_1 = \tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^2_{iq}\tilde{\varepsilon}^2_{is}$), $s = l \neq t = q$ ($X_1 = 3\tilde{\varepsilon}^4_{it}\tilde{\varepsilon}^2_{is}$), and $s = l = t = q$ ($X_1 = 10\tilde{\varepsilon}^6_{it}$). Considering these cases and using Cauchy-Schwarz inequality and c_r-inequality, we can bound the above condition as follows.

$$
\begin{aligned}
E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T \tilde{f}_s\frac{1}{T}\sum_{t=1}^T \psi^*_{1,st}\right\|^2 &\leq M\frac{1}{T}\sum_{s=1}^T \tilde{f}'_s\tilde{f}_s\frac{1}{T^2}\sum_{t=1}^T\sum_{q=1}^T\frac{1}{N}\sum_{i=1}^N \tilde{\lambda}'_i\tilde{\lambda}_i\tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^2_{is}\tilde{\varepsilon}^2_{iq} \\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T|\tilde{f}'_s\tilde{f}_s|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^T\left|\frac{1}{T^2}\sum_{t=1}^T\sum_{q=1}^T\frac{1}{N}\sum_{i=1}^N\tilde{\lambda}'_i\tilde{\lambda}_i\tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^2_{is}\tilde{\varepsilon}^2_{iq}\right|^2\right)^{1/2} \\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T\|\tilde{f}_s\|^4\right)^{1/2}\left(\frac{1}{NT}\sum_{s=1}^T\sum_{i=1}^N|\tilde{\lambda}'_i\tilde{\lambda}_i\tilde{\varepsilon}^2_{is}|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^N\left|\frac{1}{T}\sum_{t=1}^T\tilde{\varepsilon}^2_{it}\right|^4\right)^{1/2} \\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T\|\tilde{f}_s\|^4\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^N\|\tilde{\lambda}_i\|^4\right)^{1/4}\left(\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T|\tilde{\varepsilon}_{is}|^8\right)^{3/4} \\
&= O_p(1).
\end{aligned}
$$

By applying Lemma C.1 with $p = 8$, we can show that $\frac{1}{T}\sum_{s=1}^T\|\tilde{f}_s\|^4 = O_p(1)$, $\frac{1}{N}\sum_{i=1}^N\|\tilde{\lambda}_i\|^4 = O_p(1)$, and $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T|\tilde{\varepsilon}_{it}|^8 = O_p(1)$. To prove the part (v), $E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T \tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi^*_{2,jk,st}\right\|^2 = O_p(1)$, where $\psi^*_{2,jk,st} = \frac{1}{\sqrt{N_j N_k}}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}_{k,i_2}\varepsilon^*_{k,i_2 t}(\varepsilon^*_{j,i_1 s}\varepsilon^*_{j,i_1 t} - E^*(\varepsilon^*_{j,i_1 s}\varepsilon^*_{j,i_1 t}))$. We have that

$$
E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T \tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi^*_{2,jk,st}\right\|^2 = \frac{1}{T}\sum_{s=1}^T\sum_{l=1}^T\tilde{f}'_{js}\tilde{f}_{jl}\left(\frac{1}{T}\sum_{t=1}^T E^*(\psi^{*\prime}_{2,jk,st}\psi^*_{2,jk,lq})\right).
$$

To show that this is $O_p(1)$, we expand the expression for $E^*(\psi^{*\prime}_{2,jk,st}\psi^*_{2,jk,lq})$. Ignoring the group notation and considering the case where $j=k$, we can rewrite $E^*(\psi^*_{2,jk,st}\psi^*_{2,jk,lq})$ as

$$E^*(\psi^{*\prime}_{2,jk,st}\psi^*_{2,jk,lq}) = \frac{1}{N^2}\sum_{i_1=1}^{N}\sum_{i_2\neq i_1}^{N}\sum_{i_3=1}^{N}\sum_{i_4\neq i_3}^{N}\tilde{\lambda}'_{i_2}\tilde{\lambda}_{i_4}\underbrace{E^*[\varepsilon^*_{i_2t}\varepsilon^*_{i_4q}(\varepsilon^*_{i_1s}\varepsilon^*_{i_1t}-E^*(\varepsilon^*_{i_1s}\varepsilon^*_{i_1t}))(\varepsilon^*_{i_3l}\varepsilon^*_{i_3q}-E^*(\varepsilon^*_{i_3l}\varepsilon^*_{i_3q}))]}_{\equiv X_2}.$$

Since X_2 is non-zero only if $i_1=i_3\neq i_2=i_4$, we consider X_2 depending on s,t,q, and l and $i_1=i_3\neq i_2=i_4$. When $t=q\neq s=l$, $X_2=\tilde{\varepsilon}^2_{i_2t}\tilde{\varepsilon}^2_{i_1s}\tilde{\varepsilon}^2_{i_1t}$ and when $t=q=s=l$, $X_2=2\tilde{\varepsilon}^4_{i_1t}$ using the fact that $\eta_{it}\sim$ i.i.d.$N(0,1)$. For the other combinations of s,t,q, and l, we have $X_2=0$. Considering this and putting the group notation back, we can bound $E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T\tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi^*_{2,jk,st}\right\|^2$ as follows.

$$\begin{aligned}
&E^*\left\|\frac{1}{\sqrt{T}}\sum_{s=1}^T\tilde{f}_{js}\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi^*_{2,jk,st}\right\|^2\\
&\leq M\frac{1}{T}\sum_{s=1}^T\tilde{f}'_{js}\tilde{f}_{js}\frac{1}{T}\sum_{t=1}^T\frac{1}{N_jN_k}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}'_{k,i_2}\tilde{\lambda}_{k,i_2}\tilde{\varepsilon}^2_{k,i_2t}\tilde{\varepsilon}^2_{j,i_1s}\tilde{\varepsilon}^2_{j,i_1t}\\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T|\tilde{f}'_{js}\tilde{f}_{js}|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^T\left|\frac{1}{T}\sum_{t=1}^T\frac{1}{N_jN_k}\sum_{i_1=1}^{N_j}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}'_{k,i_2}\tilde{\lambda}_{k,i_2}\tilde{\varepsilon}^2_{k,i_2t}\tilde{\varepsilon}^2_{j,i_1s}\tilde{\varepsilon}^2_{j,i_1t}\right|^2\right)^{1/2}\\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T\|\tilde{f}_{js}\|^4\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^T\sum_{t=1}^T\left|\frac{1}{N_j}\sum_{i_1=1}^{N_j}\tilde{\varepsilon}^2_{j,i_1s}\tilde{\varepsilon}^2_{j,i_1t}\right|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^T\left|\frac{1}{N_k}\sum_{i_2\neq i_1}^{N_k}\tilde{\lambda}'_{k,i_2}\tilde{\lambda}_{k,i_2}\tilde{\varepsilon}^2_{k,i_2t}\right|^2\right)^{1/2}\\
&\leq M\left(\frac{1}{T}\sum_{s=1}^T\|\tilde{f}_{js}\|^4\right)^{1/2}\left(\frac{1}{N_jT}\sum_{i_1=1}^{N_j}\sum_{t=1}^T\tilde{\varepsilon}^4_{j,i_1t}\right)^{1/2}\left(\frac{1}{N_k}\sum_{i_2\neq i_1}^{N_k}\|\tilde{\lambda}_{k,i_2}\|^4\right)^{1/2}\left(\frac{1}{N_kT}\sum_{i_2\neq i_1}^{N_k}\sum_{t=1}^T\tilde{\varepsilon}^4_{k,i_2t}\right)^{1/2}\\
&=O_p(1).
\end{aligned}$$

Part (vi): As a sufficient condition, we can show that $\frac{1}{T}\sum_{s=1}^T E^*\left\|\frac{1}{T}\sum_{t=1}^T\psi^*_{1,jk,st}\right\|^2=O_p(1)$. Since $E^*\left\|\frac{1}{T}\sum_{t=1}^T\psi^*_{1,jk,st}\right\|^2=\frac{1}{T^2}\sum_{t=1}^T\sum_{q=1}^T E^*(\psi^{*\prime}_{1,jk,st}\psi^*_{1,jk,sq})$, we focus on expanding $E^*(\psi^{*\prime}_{1,jk,st}\psi^*_{1,jk,sq})$ (ignoring the group notation) as follows.

$$\begin{aligned}
E^*(\psi^{*\prime}_{1,jk,st}\psi^*_{1,jk,sq}) &= \frac{1}{N}\sum_{i_1=1}^N\sum_{i_2=1}^N\tilde{\lambda}'_{i_2}\tilde{\lambda}_{i_2}\underbrace{E^*[\varepsilon^*_{i_1t}\varepsilon^*_{i_2t}(\varepsilon^*_{i_1s}\varepsilon^*_{i_1t}-E^*(\varepsilon^*_{i_1s}\varepsilon^*_{i_1t}))(\varepsilon^*_{i_2s}\varepsilon^*_{i_2q}-E^*(\varepsilon^*_{i_2s}\varepsilon^*_{i_2q}))]}_{\equiv X_3}\\
&\leq M\frac{1}{N}\sum_{i=1}^N\tilde{\lambda}'_i\tilde{\lambda}_i\tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^2_{iq}\tilde{\varepsilon}^2_{is}.
\end{aligned}$$

If $i_1\neq i_2$, we have $X_3=0$, since $\eta_{it}\sim$ i.i.d.$N(0,1)$. When $i_1=i_2$, we have five cases to consider: if $s\neq t\neq q$, $X_3=\tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^2_{iq}\tilde{\varepsilon}^2_{is}$; if $s=t\neq q$, $X_3=2\tilde{\varepsilon}^4_{it}\tilde{\varepsilon}^2_{iq}$; if $s=q\neq t$, $X_3=2\tilde{\varepsilon}^2_{it}\tilde{\varepsilon}^4_{is}$; if $q=t\neq s$,

$X_3 = 3\tilde{\varepsilon}_{it}^4\tilde{\varepsilon}_{is}^2$; and if $s = t = q$, $X_3 = 10\tilde{\varepsilon}_{it}^6$. Therefore,

$$\frac{1}{T^3}\sum_{s=1}^T\sum_{t=1}^T\sum_{q=1}^T E^*(\psi_{1,jk,st}^{*\prime}\psi_{1,jk,sq}^*) \leq M\frac{1}{T^3}\sum_{s=1}^T\sum_{t=1}^T\sum_{q=1}^T\left(\frac{1}{N}\sum_{i=1}^N\tilde{\lambda}_i'\tilde{\lambda}_i\tilde{\varepsilon}_{it}^2\tilde{\varepsilon}_{iq}^2\tilde{\varepsilon}_{is}^2\right) = M\frac{1}{N}\sum_{i=1}^N\tilde{\lambda}_i'\tilde{\lambda}_i\left(\frac{1}{T}\sum_{t=1}^T\tilde{\varepsilon}_{it}^2\right)^3$$

$$\leq M\left(\frac{1}{N}\sum_{i=1}^N|\tilde{\lambda}_i'\tilde{\lambda}_i|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^N\left|\frac{1}{T}\sum_{t=1}^T\tilde{\varepsilon}_{it}^2\right|^6\right)^{1/2}$$

$$\leq M\left(\frac{1}{N}\sum_{i=1}^N\|\tilde{\lambda}_i\|^4\right)^{1/2}\left(\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T|\tilde{\varepsilon}_{it}|^{12}\right)^{1/2} = O_p(1),$$

by applying Lemma C.1 with $p = 12$. For part (vii), we use similar arguments and show that $\frac{1}{T}\sum_{s=1}^T E^*\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi_{2,jk,st}^*\right\|^2 = O_p(1)$, where $\psi_{2,jk,st}^*$ is defined in part (v). In particular, ignoring the group notation and considering $j = k$ ($N_j = N_k = N$), we can write $E^*(\psi_{2,st}^{*\prime}\psi_{2,sq}^*)$ as follows.

$$E^*(\psi_{2,st}^{*\prime}\psi_{2,sq}^*) = \frac{1}{N^2}\sum_{i_1=1}^N\sum_{i_2\neq i_1}^N\sum_{i_3=1}^N\sum_{i_4\neq i_3}^N\tilde{\lambda}_{i_2}'\tilde{\lambda}_{i_4}\underbrace{E^*[\varepsilon_{i_2t}^*\varepsilon_{i_4q}^*(\varepsilon_{i_1s}^*\varepsilon_{i_1t}^* - E^*(\varepsilon_{i_1s}^*\varepsilon_{i_1t}^*))(\varepsilon_{i_3s}^*\varepsilon_{i_3q}^* - E^*(\varepsilon_{i_3s}^*\varepsilon_{i_3q}^*))]}_{\equiv X_4},$$

where $X_4 = 0$ when $i_1 = i_4 \neq i_2 = i_3$, and $X_4 \neq 0$ when $i_1 = i_3 \neq i_2 = i_4$ with $s \neq t = q$ or $s = t = q$. It follows that

$$\frac{1}{T}\sum_{s=1}^T E^*\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^T\psi_{2,jk,st}^*\right\|^2 \leq M\frac{1}{T^2}\sum_{s=1}^T\sum_{t=1}^T\left(\frac{1}{N^2}\sum_{i_1=1}^N\sum_{i_2\neq i_1}^N\tilde{\lambda}_{i_2}'\tilde{\lambda}_{i_2}\tilde{\varepsilon}_{i_1s}^2\tilde{\varepsilon}_{i_1t}^2\tilde{\varepsilon}_{i_2t}^2\right)$$

$$= M\frac{1}{T}\sum_{t=1}^T\left(\frac{1}{NT}\sum_{i_1=1}^N\sum_{s=1}^T\tilde{\varepsilon}_{i_1s}^2\tilde{\varepsilon}_{i_1t}^2\right)\left(\frac{1}{N}\sum_{i_2\neq i_1}^N\tilde{\lambda}_{i_2}'\tilde{\lambda}_{i_2}\tilde{\varepsilon}_{i_2t}^2\right)$$

$$\leq M\left(\frac{1}{T}\sum_{t=1}^T\left|\frac{1}{NT}\sum_{i_1=1}^N\sum_{s=1}^T\tilde{\varepsilon}_{i_1s}^2\tilde{\varepsilon}_{i_1t}^2\right|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^T\left|\frac{1}{N}\sum_{i_2\neq i_1}^N\tilde{\lambda}_{i_2}'\tilde{\lambda}_{i_2}\tilde{\varepsilon}_{i_2t}^2\right|^2\right)^{1/2}$$

$$\leq M\left(\frac{1}{NT}\sum_{i_1=1}^N\sum_{s=1}^T|\tilde{\varepsilon}_{i_1s}|^4\right)\left(\frac{1}{N}\sum_{i_2\neq i_1}^N\|\tilde{\lambda}_{i_2}\|^4\right)^{1/2}\left(\frac{1}{NT}\sum_{i_2\neq i_1}^N\sum_{t=1}^T|\tilde{\varepsilon}_{i_2t}|^4\right)^{1/2}$$

$$= O_p(1),$$

given Lemma C.1 with $p = 4$. ■

Proof of Lemma C.3. Part (i): Since $\hat{f}_t^c = \hat{W}'\hat{f}_{1t}$ and we can write the factors estimation error as $\hat{f}_{jt} - H_jf_{jt} = \mathcal{V}_j^{-1}(A_{j,1t} + A_{j,2t} + A_{j,3t} + A_{j,4t})$ as in A.2 in Appendix A, we can write $\hat{f}_t^c$ as follows.

$$\begin{aligned}\hat{f}_t^c = \hat{W}'\hat{f}_{1t} &= \hat{W}'(H_1f_{1t} + \mathcal{V}_1^{-1}(A_{1,1t} + A_{1,2t} + A_{1,3t} + A_{1,4t}))\\ &= \tilde{W}_1'f_{1t} + \hat{W}'\mathcal{V}_1^{-1}(A_{1,1t} + A_{1,2t} + A_{1,3t} + A_{1,4t})\\ &= \hat{U}'[\underbrace{E_c'f_{1t}}_{=f_t^c} + \hat{\Phi}_{sc}'(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\underbrace{E_s'f_{1t}}_{=f_{1t}^s}] + \hat{W}'\mathcal{V}_1^{-1}(A_{1,1t} + A_{1,2t} + A_{1,3t} + A_{1,4t}),\end{aligned}$$

where we use $\tilde{W}_1 = H_1'\hat{W}$ and $\tilde{W}_1 = [E_c + E_s(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Phi}_{sc}]\hat{U}$ from the proof of Lemma A.4 in

Appendix A. Note that $E_c' f_{1t} = f_t^c$ and $E_s' f_{1t} = f_{1t}^s$ under H_0. By letting $H^c = \hat{U}'$, we can rewrite $\hat{f}_t^c$ as follows.

$$\begin{aligned}\hat{f}_t^c &= H^c f_t^c + \hat{W}' \mathcal{V}_1^{-1}(A_{1,1t} + A_{1,2t} + A_{1,3t} + A_{1,4t}) + \hat{\Phi}_{sc}'(I_{k_1-k^c} - \tilde{R}_{ss})^{-1} f_{1t}^s \\ &= H^c f_t^c + \hat{W}' \mathcal{V}_1^{-1} \left(\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\gamma_{1,st} + \frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\zeta_{1,st} + \frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\eta_{1,st} + \frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\xi_{1,st} \right) + o_p(T^{-1/2}),\end{aligned}$$

where we use $\hat{\Phi}_{sc}'(I_{k_1-k^c} - \tilde{R}_{ss})^{-1} = O_p(\delta_{NT}^{-2})$. For the rest of the terms, we use Lemma A.2 in Bai (2003): $\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\gamma_{1,st} = O_p(\delta_{NT}^{-1}T^{-1/2})$; $\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\zeta_{1,st} = O_p(\delta_{NT}^{-1}N^{-1/2})$; $\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\eta_{1,st} = O_p(N^{-1/2})$; and $\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\xi_{1,st} = O_p(\delta_{NT}^{-1}N^{-1/2})$. Since $O_p(\delta_{NT}^{-1}N^{-1/2}) = o_p(T^{-1/2})$ and $O_p(\delta_{NT}^{-1}T^{-1/2}) = o_p(T^{-1/2})$, we can simplify the asymptotic expansion of $\hat{f}_t^c$ up to order $o_p(T^{-1/2})$ as follows.

$$\begin{aligned}\hat{f}_t^c &= H^c f_t^c + \hat{W}'\mathcal{V}_1^{-1}\frac{1}{T}\sum_{s=1}^T \hat{f}_{1s}\eta_{1,st} + o_p(T^{-1/2}) \\ &= H^c f_t^c + \underbrace{\hat{W}'\mathcal{V}_1^{-1}\mathcal{V}_1 H_1}_{=\tilde{W}_1'} \underbrace{\left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-1}\left(\frac{1}{\sqrt{N_1}}\sum_{i=1}^{N_1}\lambda_{1,i}\varepsilon_{1,it}\right)}_{=u_{1t}} \frac{1}{\sqrt{N_1}} + o_p(T^{-1/2}) \\ &= H^c f_t^c + \frac{1}{\sqrt{N_1}}\hat{U}'[\underbrace{E_c' u_{1t}}_{=u_{1t}^{(c)}} + \underbrace{\hat{\Phi}_{sc}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}}_{=O_p(\delta_{NT}^{-2})} \underbrace{E_s' u_{1t}}_{=u_{1t}^{(s)}}] + o_p(T^{-1/2}) \\ &= H^c f_t^c + \frac{1}{\sqrt{N_1}} H^c u_{1t}^{(c)} + o_p(T^{-1/2}),\end{aligned}$$

where we use the fact that $\frac{\hat{F}_1' F_1}{T} = \mathcal{V}_1 H_1 \left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-1}$ by the definition of H_1 in the second equality and use the expression for $\tilde{W}_1$ in the third equality.

Part (ii): Next, we show the asymptotic expansion of $\hat{\lambda}_{j,i}^c$ up to order $o_p(T^{-1/2})$. In particular, by using the fact that $\hat{\Lambda}_j^c = \frac{1}{T}Y_j'\hat{F}^c$ and $Y_j = F^c\Lambda_j^{c\prime} + F_j^s\Lambda_j^{s\prime} + \varepsilon_j$ and substituting appropriately, we can write $\hat{\lambda}_{j,i}^c$ as follows.

$$\begin{aligned}\hat{\lambda}_{j,i}^c &= (H^c)^{-1\prime}\lambda_{j,i}^c + H^c\frac{1}{T}\sum_{t=1}^T f_t^c\varepsilon_{j,it} + H^c\frac{1}{T}\sum_{t=1}^T f_t^c f_{jt}^{s\prime}\lambda_{j,i}^s \\ &\quad + \underbrace{\frac{1}{T}\sum_{t=1}^T(\hat{f}_t^c - H^c f_t^c)\varepsilon_{j,it}}_{=(a)} + \underbrace{\frac{1}{T}\sum_{t=1}^T(\hat{f}_t^c - H^c f_t^c)f_{jt}^{s\prime}}_{=(b)}\lambda_{j,i}^s - \underbrace{\frac{1}{T}\sum_{t=1}^T(\hat{f}_t^c - H^c f_t^c)(\hat{f}_t^c - H^c f_t^c)'(H^c)^{-1\prime}}_{=(c)}\lambda_{j,i}^c \\ &\quad - \underbrace{\frac{1}{T}\sum_{t=1}^T H^c f_t^c(\hat{f}_t^c - H^c f_t^c)'}_{=(d)}(H^c)^{-1\prime}\lambda_{j,i}^c.\end{aligned}$$

Then, to prove part (ii), we show that the terms (a) through (d) are $o_p(T^{-1/2})$. Using the expansion of $\hat{f}_t^c$ from part (i), we can rewrite the term (a) as follows.

$$
\begin{aligned}
(a) &= H^c \frac{1}{T\sqrt{N_1}} \sum_{t=1}^{T} u_{1t}\varepsilon_{j,it} + o_p(T^{-1/2}) \\
&= H^c \left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-1} \frac{1}{TN_1} \sum_{t=1}^{T}\sum_{k=1}^{N_1} \lambda_{1,k}\varepsilon_{1,kt}\varepsilon_{j,it} + o_p(T^{-1/2}) \\
&= H^c \left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-1} \left[\underbrace{\frac{1}{TN_1} \sum_{t=1}^{T}\sum_{k=1}^{N_1} (\varepsilon_{1,kt}\varepsilon_{j,it} - E(\varepsilon_{1,kt}\varepsilon_{j,it}))}_{=O_p((TN_1)^{-1/2})} + \underbrace{\frac{1}{TN_1} \sum_{t=1}^{T}\sum_{k=1}^{N_1} E(\varepsilon_{1,kt}\varepsilon_{j,it})}_{\substack{=O(N_1^{-1}) \\ \text{by Ass 5-(a)}}} \right] + o_p(T^{-1/2}) \\
&= o_p(T^{-1/2}).
\end{aligned}
$$

Similarly, we can show that (b) and (d) are $o_p(T^{-1/2})$ by replacing $\hat{f}_t^c - H^c f_t^c$ with its expansion up to order $o_p(T^{-1/2})$. For example, ignoring $\|H^c\| = O_p(1)$, the term (d) is $\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t^c - H^c f_t^c) f_t^{c\prime} = \left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-1} \frac{1}{\sqrt{N_1 T}} \left(\frac{1}{\sqrt{N_1 T}} \sum_{t=1}^{T} f_t^c \varepsilon_{1t}' \Lambda_1\right) = O_p((TN_1)^{-1/2}) = o_p(T^{-1/2})$ by Assumption 4-(c). Using the proof of Lemma A.1-(e), we can show that $\frac{1}{NT}\sum_{t=1}^{T} u_{1t}u_{1t}' = O_p(N^{-1})$ and show that $(e) = o_p(T^{-1/2})$. Our asymptotic expansions for $\hat{f}_t^c$ and $\hat{\lambda}_{j,i}^c$ are equivalent to those in AGGR(2019) (specifically, (C.92) and (C.94) in their Online Appendix).

Part (iii): To obtain the asymptotic expansion of $\hat{f}_{jt}^s$, we follow the arguments in AGGR(2019) closely. Recall that $\hat{f}_{jt}^s$ are principal components of the residuals such that $\xi_{j,it} = y_{j,it} - \hat{f}_t^{c\prime}\hat{\lambda}_{j,i}^c$. Following the arguments in AGGR (2019), by replacing $\hat{f}_t^c$ and $\hat{\lambda}_{j,i}^c$ with their asymptotic expansions of order up to $o_p(T^{-1/2})$ and using the fact that $H^{c\prime}H^c = \tilde{\Sigma}_{cc}^{-1} + o_p(T^{-1/2})$, we can rewrite $\xi_{j,it}$ as follows.[12]

$$
\xi_{j,it} = \underbrace{\left(f_{jt}^s - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1} f_t^c\right)'}_{=\tilde{f}_{jt}^{s\prime}} \lambda_{j,i}^s + \underbrace{\left(\varepsilon_{j,it} - \frac{1}{\sqrt{N_j}} u_{jt}^{(c)\prime}\lambda_{j,i}^c - f_t^{c\prime}\left(\tilde{\Sigma}_{cc}^{-1}\frac{1}{T}\sum_{t=1}^{T} f_t^c \varepsilon_{j,it}\right)\right)}_{\equiv \tilde{e}_{j,it}} + o_p(T^{-1/2}).
$$

Using the identity from Bai (2003) as in Appendix A.2, we can write $\hat{f}_{jt}^s - \tilde{H}_j^s \tilde{f}_{jt}^s$ as follows.

$$
\hat{f}_{jt}^s - \tilde{H}_j^s \tilde{f}_{jt}^s = (\mathcal{V}_j^s)^{-1}(B_{j,1t} + B_{j,2t} + B_{j,3t}),
$$

where

$$
B_{j,1t} = \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^s \left(\frac{1}{N_j}\sum_{i=1}^{N_j} \tilde{e}_{j,il}\tilde{e}_{j,it}\right); \quad B_{j,2t} = \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^s \tilde{f}_{jl}^{s\prime} \frac{\Lambda_j^{s\prime}\tilde{e}_{jt}}{N_j}; \quad B_{j,3t} = \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^s \tilde{f}_{jt}^{s\prime} \frac{\Lambda_j^{s\prime}\tilde{e}_{jl}}{N_j}.
$$

[12] We can replace $\hat{f}_t^c$ and $\hat{\lambda}_{j,i}^c$ with their expansions based on alternative group common factors such that $\dot{f}_t^c = \dot{W}'\hat{f}_{2t}$, where $\dot{W}$ is $k_1 \times k^c$ matrix collecting eigenvectors of $\hat{R}^*$ associated to k^c eigenvalues. It yields the similar expansion such that $\dot{f}_t^c = \dot{H}^c\left(f_t^c + \frac{1}{N_2}u_{2t}^{(c)}\right) + o_p(T^{-1/2})$.

We show that $B_{j,1t}$ and $B_{j,3t}$ are $o_p(T^{-1/2})$ and $B_{j,2t}$ have a bias term of order $O_p(N_j^{-1/2})$ up to order $o_p(T^{-1/2})$ (ignoring $\|(\mathcal{V}_j^s)^{-1}\| = O_p(1)$). Letting $w_{j,i}^c \equiv \tilde{\Sigma}_{cc}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t^c \varepsilon_{j,it}$, we first rewrite $B_{j,1t}$ as follows.

$$
\begin{aligned}
B_{j,1t} &= \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\varepsilon_{j,il} - \frac{1}{\sqrt{N_j}}u_{jl}^{(c)\prime}\lambda_{j,i}^c - \frac{1}{\sqrt{T}}f_l^{c\prime}w_{j,i}^c\right)\left(\varepsilon_{j,it} - \frac{1}{\sqrt{N_j}}u_{jt}^{(c)\prime}\lambda_{j,i}^c - \frac{1}{\sqrt{T}}f_t^{c\prime}w_{j,i}^c\right) \\
&= \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j}\sum_{i=1}^{N_j}\varepsilon_{j,il}\varepsilon_{j,it} - \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{N_j}}\sum_{i=1}^{N_j}\varepsilon_{j,il}u_{jt}^{(c)\prime}\lambda_{j,i}^c - \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{N_j}}\sum_{i=1}^{N_j}\varepsilon_{j,it}u_{jl}^{(c)\prime}\lambda_{j,i}^c \\
&\quad - \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{T}}\sum_{i=1}^{N_j}\varepsilon_{j,il}f_t^{c\prime}w_{j,i}^c - \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{T}}\sum_{i=1}^{N_j}\varepsilon_{j,it}f_l^{c\prime}w_{j,i}^c + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j^2}\sum_{i=1}^{N_j}u_{jl}^{(c)\prime}\lambda_{j,i}^c u_{jt}^{(c)\prime}\lambda_{j,i}^c \\
&\quad + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{N_jT}}\sum_{i=1}^{N_j}\lambda_{j,i}^{c\prime}u_{jl}^{(c)}f_t^{c\prime}w_{j,i}^c + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j\sqrt{N_jT}}\sum_{i=1}^{N_j}\lambda_{j,i}^{c\prime}u_{jt}^{(c)}f_l^{c\prime}w_{j,i}^c \\
&\quad + \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_jT}\sum_{i=1}^{N_j}f_l^{c\prime}w_{j,i}^c f_t^{c\prime}w_{j,i}^c \\
&= B_{j,1t,(1)} + B_{j,1t,(2)} + B_{j,1t,(3)} + B_{j,1t,(4)} + B_{j,1t,(5)} + B_{j,1t,(6)} + B_{j,1t,(7)} + B_{j,1t,(8)} + B_{j,1t,(9)}.
\end{aligned}
$$

We can show that all nine terms in $B_{j,1t}$ are $o_p(T^{-1/2})$. We can show that $B_{j,1t,(1)} = o_p(T^{-1/2})$ by applying similar arguments in Lemma A.2 in Bai (2003). We can write the next term as follows.

$$
\begin{aligned}
B_{j,1t,(2)} &= \frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \frac{1}{N_j}\left(\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^c\varepsilon_{j,il}\right)u_{jt}^{(c)} \\
&= \frac{1}{N_j}\left[\tilde{H}_j^s\frac{1}{T}\sum_{l=1}^T \tilde{f}_{jl}^s\left(\frac{\Lambda_j^{c\prime}\varepsilon_{jl}}{\sqrt{N_j}}\right)' + \frac{1}{T}\sum_{l=1}^T(\hat{f}_{jl}^s - \tilde{H}_j^s\tilde{f}_{jl}^s)\left(\frac{\Lambda_j^{c\prime}\varepsilon_{jl}}{\sqrt{N_j}}\right)'\right]u_{jt}^{(c)} \\
&\leq \max_{1\leq t\leq T}\|u_{jt}^{(c)}\|\frac{1}{N_j}\left[\underbrace{\left\|\tilde{H}_j^s\frac{1}{T}\sum_{l=1}^T \tilde{f}_{jl}^s\left(\frac{\Lambda_j^{c\prime}\varepsilon_{jl}}{\sqrt{N_j}}\right)'\right\|}_{=(b_1)} + \underbrace{\left\|\frac{1}{T}\sum_{l=1}^T(\hat{f}_{jl}^s - \tilde{H}_j^s\tilde{f}_{jl}^s)\left(\frac{\Lambda_j^{c\prime}\varepsilon_{jl}}{\sqrt{N_j}}\right)'\right\|}_{=(b_2)}\right].
\end{aligned}
$$

Then, $(b_1) = O_p(T^{-1/2})$ since it is equivalent to $\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{TN_j}}\sum_{l=1}^T \tilde{f}_{jl}^s\varepsilon_{jl}'\Lambda_j^c\right)$, which is $O_p(T^{-1/2})$ by Assumption 4-(c). By applying Cauchy-Schwarz inequality,

$$
(b_2) \leq \left(\frac{1}{T}\sum_{l=1}^T\|\hat{f}_{jl}^s - \tilde{H}_j^s\tilde{f}_{jl}^s\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{l=1}^T\left\|\frac{\Lambda_j^{c\prime}\varepsilon_{jl}}{\sqrt{N_j}}\right\|^2\right)^{1/2} = O_p(\delta_{NT}^{-1}).
$$

Since we can show that $\max_{1\leq t\leq T}\|u_{jt}^{(c)}\| = O_p(\sqrt{T})$ by Assumption WB3-(c), we have $B_{j,1t,(2)} = o_p(T^{-1/2})$. We can use similar arguments to show that $B_{j,1t,(3)} = o_p(T^{-1/2})$. Specifically, we can write $B_{j,1t,(3)} = \left(\frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s u_{jl}^{(c)\prime}\right)\left(\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^{(c)}\varepsilon_{j,it}\right)\frac{1}{N_j}$. Then, by using the fact that $\frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s u_{jl}^{(c)} =$

$O_p(\delta_{NT}^{-1})$ and $\max_t \left\| \frac{\Lambda_j^{c\prime}\varepsilon_{jt}}{\sqrt{N_j}} \right\| = O_p(T^{1/2})$, we can show that $B_{j,1t,(3)} = o_p(T^{-1/2})$. We write the next term as

$$
\begin{aligned}
B_{j,1t,(4)} &= \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} \left(\frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c \varepsilon_{j,is}\varepsilon_{j,il} \right)' \tilde{\Sigma}_{cc}^{-1} f_t^c \\
&\leq \max_{1\leq t\leq T} \|f_t^c\| \left\| \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} \left(\frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c \varepsilon_{j,is}\varepsilon_{j,il} \right)' \right\| \|\tilde{\Sigma}_{cc}^{-1}\| \\
&\leq \max_{1\leq t\leq T} \|f_t^c\| \left[\left\| \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} \frac{1}{\sqrt{N_j T}} \left(\frac{1}{\sqrt{N_j T}} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c (\varepsilon_{j,is}\varepsilon_{j,il} - E(\varepsilon_{j,is}\varepsilon_{j,il})) \right) \right\| \right. \\
&\quad \left. + \left\| \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} \frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c E(\varepsilon_{j,is}\varepsilon_{j,il}) \right\| \right]' \left\| \tilde{\Sigma}_{cc}^{-1} \right\|.
\end{aligned}
$$

By Assumption 4-(b), we have $\frac{1}{\sqrt{N_j T}} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c (\varepsilon_{j,is}\varepsilon_{j,il} - E(\varepsilon_{j,is}\varepsilon_{j,il})) = O_p(1)$. Then, we can show that the first term in the square bracket is $O_p(N_j^{-1/2}T^{-1/2})$. We can decompose the second term in the square bracket into two parts as follows.

$$
\begin{aligned}
\left\| \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} \frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c E(\varepsilon_{j,is}\varepsilon_{j,il}) \right\| &\leq \left\| \frac{1}{T}\sum_{l=1}^{T} (\hat{f}_{jl}^{s} - \tilde{H}_j^s \tilde{f}_{jl}^s) \frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c E(\varepsilon_{j,is}\varepsilon_{j,il}) \right\| \\
&\quad + \left\| \frac{1}{T}\sum_{l=1}^{T} \tilde{H}_j^s \tilde{f}_{jl}^{s} \frac{1}{N_j T} \sum_{i=1}^{N_j} \sum_{s=1}^{T} f_s^c E(\varepsilon_{j,is}\varepsilon_{j,il}) \right\| \\
&= O_p(\delta_{NT}^{-1} T^{-1}) + O_p(T^{-1}),
\end{aligned}
$$

where we use the fact that $\frac{1}{T}\sum_{t=1}^{T} \left\| \frac{1}{T}\sum_{s=1}^{T} f_s^c \gamma_{j,sl} \right\| = O_p(T^{-2})$ by following the arguments in the proof of Lemma A.1-(a). Then, by Assumption 2-(a), we can show that $\max_{1\leq t\leq T} \|f_t^c\| = O_p(T^{1/4})$ and hence, $B_{j,1t,(4)} = o_p(T^{-1/2})$. Using similar arguments, we can also show that $B_{j,1t,(5)} = o_p(T^{-1/2})$. Similar to the arguments to show that $B_{j,1t,(2)} = o_p(T^{-1/2})$, we can show that $\frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} u_{jl}^{(c)\prime} = O_p(\delta_{NT}^{-1})$. Then, since $B_{j,1t,(6)}$ is equivalent to write it as $\left(\frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} u_{jl}^{(c)\prime}\right)\left(\frac{\Lambda_j^{c\prime}\Lambda_j^c}{N_j}\right)\frac{1}{N_j} u_{jt}^{(c)}$, we can show this term is $O_p(\delta_{NT}^{-1} N_j^{-1} T^{1/2}) = o_p(T^{-1/2})$. Next, $B_{j,1t,(7)}$ can be bounded as follows.

$$
B_{j,1t,(7)} \leq \underbrace{\left\| \frac{1}{T}\sum_{l=1}^{T} \hat{f}_{jl}^{s} u_{jl}^{(c)\prime} \right\|}_{=O_p(\delta_{NT}^{-1})} \frac{1}{\sqrt{N_j T}} \underbrace{\left\| \frac{1}{N_j}\sum_{i=1}^{N_j} \lambda_{j,i}^c w_{j,i}^{c\prime} \right\|}_{=O_p(N_j^{-1/2})} \underbrace{\max_{1\leq t\leq T} \|f_t^c\|}_{=O_p(T^{1/4})} = O_p(\delta_{NT}^{-1} T^{-1/2} N_j^{-1}) = o_p(T^{-1/2}),
$$

where we use the fact that $\frac{1}{N_j}\sum_{i=1}^{N_j} \lambda_{j,i}^c w_{j,i}^{c\prime} = \frac{1}{\sqrt{N_j}} \left(\frac{1}{\sqrt{N_j T}} \sum_{i=1}^{N_j} \sum_{t=1}^{T} f_t^c \varepsilon_{jt}' \Lambda_j^c \right) \tilde{\Sigma}_{cc}^{-1} = O_p(N_j^{-1/2})$ by Assumption 4-(c). Following similar arguments, we have $B_{j,1t,(8)} = O_p(N_j^{-1}) = o_p(T^{-1/2})$. Next,

$B_{j,1t,(9)}$ can be bounded as follows.

$$B_{j,1t,(9)} \leq \frac{1}{T}\left\|\frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}f_{l}^{c\prime}\right\|\left\|\frac{1}{N_j}\sum_{i=1}^{N_j}w_{j,i}^{c}w_{j,i}^{c\prime}\right\|\max_{1\leq t\leq T}\|f_t^c\| = O_p(T^{-3/4}) = o_p(T^{-1/2}),$$

where we use the fact that $\left\|\frac{1}{N_j}\sum_{i=1}^{N_j}w_{j,i}^{c}w_{j,i}^{c\prime}\right\| \leq \left(\frac{1}{N_j}\sum_{i=1}^{N_j}\left\|\frac{1}{T}\sum_{t=1}^{T}f_t^c\varepsilon_{j,it}\right\|^2\right)\|\tilde{\Sigma}_{cc}^{-1}\|^2 = O_p(1)$ by Assumption 4-(a). Since we show that $B_{j,1t,(i)} = o_p(T^{-1/2})$ for $i = 1,\dots,9$, we have $B_{j,1t} = o_p(T^{-1/2})$.

Next, our goal is to expand $B_{j,2t}$ up to order $o_p(T^{-1/2})$. We first rewrite $B_{j,2t}$ as follows.

$$\begin{aligned}
B_{j,2t} &= \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jl}^{s\prime}\frac{\Lambda_j^{s\prime}\tilde{e}_{jt}}{N_j}\\
&= \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jl}^{s\prime}\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\varepsilon_{j,it} - \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jl}^{s\prime}\left(\frac{1}{N_j\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}u_{jt}^{(c)\prime}\lambda_{j,i}^{c}\right) - \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jl}^{s\prime}\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}f_t^{c\prime}w_{j,i}^{c}\\
&= B_{j,2t,(1)} + B_{j,2t,(2)} + B_{j,2t,(3)}.
\end{aligned}$$

We have $B_{j,2t,(i)} = O_p(N_j^{-1/2})$ for $i = 1, 2$ and $B_{j,2t,(3)} = o_p(T^{-1/2})$. To see this, we can bound $B_{j,2t,(3)}$ as follows.

$$B_{j,2t,(3)} \leq \left(\frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jl}^{s\prime}\right)\underbrace{\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}w_{j,i}^{c\prime}\right)}_{=O_p(N_j^{-1/2})}\frac{1}{\sqrt{T}}\underbrace{\max_{1\leq t\leq T}\|f_t^c\|}_{=O_p(T^{1/4})} = O_p(N_j^{-1/2}T^{-1/2}T^{1/4}) = o_p(T^{-1/2}).$$

Using the definition of $\tilde{e}_{jl}$, we can decompose $B_{j,3t}$ into three parts as follows.

$$\begin{aligned}
B_{j,3t} &= \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jt}^{s\prime}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}\varepsilon_{j,il}\right) - \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jt}^{s\prime}\left(\frac{1}{N_j\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}u_{jl}^{(c)\prime}\lambda_{j,i}^{c}\right)\\
&\quad - \frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\tilde{f}_{jt}^{s\prime}\left(\frac{1}{N_j\sqrt{T}}\sum_{i=1}^{N_j}\lambda_{j,i}^{s}f_l^{c\prime}w_{j,i}^{c}\right)\\
&= B_{j,3t,(1)} + B_{j,3t,(2)} + B_{j,3t,(3)}
\end{aligned}$$

Our next goal is to show that $B_{j,3t,(i)} = o_p(T^{-1/2})$ for $i = 1, 2, 3$. The first term $B_{j,3t,(1)}$ can be bounded by $\left\|\frac{1}{TN_j}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\varepsilon_{jl}'\Lambda_j^s\right\|\max_t\|\tilde{f}_{jt}^{s}\|$ and since we can show that $\left\|\frac{1}{TN_j}\sum_{l=1}^{T}\hat{f}_{jl}^{s}\varepsilon_{jl}'\Lambda_j^s\right\| = O_p(\delta_{NT}^{-2})$, we have $B_{j,3t,(1)} = o_p(T^{-1/2})$. $B_{j,3t,(2)}$ can be shown as $o_p(T^{-1/2})$ by applying that $\frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}u_{jl}^{(c)\prime} = O_p(\delta_{NT}^{-1})$. The last term $B_{j,3t,(3)}$ can be bounded as follows.

$$\begin{aligned}
B_{j,3t,(3)} &\leq \left\|\frac{1}{T}\sum_{l=1}^{T}\hat{f}_{jl}^{s}f_l^{c\prime}\right\|\left\|\frac{1}{N_j}\sum_{i=1}^{N_j}w_{j,i}^{c}\lambda_{j,i}^{s\prime}\right\|\frac{1}{\sqrt{T}}\max_{1\leq t\leq T}\|\tilde{f}_{jt}^{s}\|\\
&= O_p(N_j^{-1/2}T^{-1/2})O_p(T^{1/4}) = o_p(T^{-1/2}),
\end{aligned}$$

by applying that $\frac{1}{N_j}\sum_{i=1}^{N_j}w_{j,i}^{c}\lambda_{j,i}^{s\prime} = O_p(N_j^{-1/2})$. Therefore, we can expand $\hat{f}_{jt}^{s}$ up to order $o_p(T^{-1/2})$

as follows.

$$\hat{f}_{jt}^s = \tilde{H}_j^s \tilde{f}_{jt}^s + (\mathcal{V}_j^s)^{-1}\left(\frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \tilde{f}_{jl}^{s\prime}\right)\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\varepsilon_{j,it} - \frac{1}{N_j\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^s\lambda_{j,i}^{c\prime}u_{jt}^{(c)}\right) + o_p(T^{-1/2}).$$

Noting that $\frac{1}{T}\sum_{l=1}^T \hat{f}_{jl}^s \tilde{f}_{jl}^{s\prime}$ is equivalent to $(\mathcal{V}_j^s)^{-1}\tilde{H}_j^s\left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}$ by $\tilde{H}_j^s = (\mathcal{V}_j^s)^{-1}\left(\frac{\hat{F}_j^{s\prime}\tilde{F}_j^s}{T}\right)\left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)$, we can also write $\hat{f}_{jt}^s$ as follows.

$$\hat{f}_{jt}^s = \tilde{H}_j^s\left[\tilde{f}_{jt}^s + \frac{1}{\sqrt{N_j}}\underbrace{\left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\left(\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}\varepsilon_{j,it} - \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}^s\lambda_{j,i}^{c\prime}u_{jt}^{(c)}\right)}_{\equiv \nu_{jt}^s}\right] + o_p(T^{-1/2}).$$

Then, by following the arguments in AGGR(2019) (specifically, the arguments on page 56 in their Online Appendix), we can show that $\nu_{jt}^s = u_{jt}^{(s)}$ and this completes the proof of part (iii).

Part (iv): Recall that $\hat{\lambda}_{j,i}^s = \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s \xi_{j,it}$. By expanding $\xi_{j,it}$ up to order $o_p(T^{-1/2})$ as in the proof of part (iii), we can rewrite $\hat{\lambda}_{j,i}^s$ as follows.

$$\begin{aligned}
\hat{\lambda}_{j,i}^s &= \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s(\tilde{f}_{jt}^{s\prime}\lambda_{j,i}^s + \tilde{e}_{j,it}) + o_p(T^{-1/2}) \\
&= \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s\tilde{f}_{jt}^{s\prime}\lambda_{j,i}^s + \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s\tilde{e}_{j,it} + o_p(T^{-1/2}) \\
&= \frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s[\hat{f}_{jt}^s - (\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)]'(\tilde{H}_j^{s\prime})^{-1}\lambda_{j,i}^s + \frac{1}{T}\sum_{t=1}^T(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)\tilde{e}_{j,it} + \frac{1}{T}\sum_{t=1}^T \tilde{H}_j^s\tilde{f}_{jt}^s\tilde{e}_{j,it} + o_p(T^{-1/2}) \\
&= (\tilde{H}_j^{s\prime})^{-1}\lambda_{j,i}^s - \underbrace{\frac{1}{T}\sum_{t=1}^T \hat{f}_{jt}^s(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)'(H_j^{s\prime})^{-1}\lambda_{j,i}^s}_{=(a)} + \underbrace{\frac{1}{T}\sum_{t=1}^T(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)\tilde{e}_{j,it}}_{=(b)} + \underbrace{\frac{1}{T}\sum_{t=1}^T \tilde{H}_j^s\tilde{f}_{jt}^s\tilde{e}_{j,it}}_{=(c)} + o_p(T^{-1/2}).
\end{aligned}$$

To expand further, we analyze three terms, (a), (b), and (c). By replacing $\hat{f}_{jt}^s$ with $(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s) + \tilde{H}_j^s\tilde{f}_{jt}^s$, we can decompose (a) further into two parts: $\frac{1}{T}\sum_{t=1}^T(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)'$ and $\tilde{H}_j^s\frac{1}{T}\sum_{t=1}^T \tilde{f}_{jt}^s(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)'$. By using that $\frac{1}{T}\sum_{t=1}^T u_{jt}u_{jt}' = O_p(1)$, we can show that $\frac{1}{T}\sum_{t=1}^T(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)' = \frac{1}{N_j}\left(\frac{1}{T}\sum_{t=1}^T u_{jt}^{(s)}u_{jt}^{(s)\prime}\right) + o_p(T^{-1/2}) = o_p(T^{-1/2})$. By Assumption 4-(c), we can also show that $\tilde{H}_j^s\frac{1}{T}\sum_{t=1}^T \tilde{f}_{jt}^s(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)' = O_p((TN_j)^{-1/2}) = o_p(T^{-1/2})$. Next, we rewrite (b) as follows.

$$\frac{1}{T}\sum_{t=1}^T(\hat{f}_{jt}^s - \tilde{H}_j^s\tilde{f}_{jt}^s)\tilde{e}_{j,it} = \underbrace{\frac{1}{T\sqrt{N_j}}\sum_{t=1}^T u_{jt}^{(s)}\varepsilon_{j,it}}_{=(b_1)} - \underbrace{\frac{1}{TN_j}\sum_{t=1}^T u_{jt}^{(s)}u_{jt}^{(c)\prime}\lambda_{j,i}^c}_{=(b_2)} - \underbrace{\frac{1}{T\sqrt{TN_j}}\sum_{t=1}^T u_{jt}^{(s)}f_t^{c\prime}w_{j,i}^c}_{=(b_3)} + o_p(T^{-1/2}).$$

Since we can write $(b_1) = \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\left(\frac{1}{TN_j}\sum_{t=1}^T\sum_{i=1}^{N_j}\lambda_{j,i}\varepsilon_{j,it}^2 - \left(\frac{\Lambda_j^{s\prime}\Lambda_j^c}{N_j}\right)\frac{1}{T\sqrt{N_j}}\sum_{t=1}^T\varepsilon_{j,it}u_{jt}^{(c)}\right)$, using the similar arguments in the proof of part (ii), we can show that $(b_1) = O_p(\delta_{NT}^{-2}) = o_p(T^{-1/2})$. Also, using the fact that $\frac{1}{T}\sum_{t=1}^{T} u_{jt}u_{jt}' = O_p(1)$, $(b_2) = O_p(N_j^{-1}) = o_p(T^{-1/2})$. We can bound the term (b_3) as follows.

$$(b_3) \leq \frac{1}{\sqrt{TN_j}} \left\| \frac{1}{T}\sum_{t=1}^{T} u_{jt}^{(s)} f_t^{c\prime} \right\| \max_{1\leq i\leq N} w_{j,i}^c = \frac{1}{\sqrt{TN_j}} O_p\left(\frac{1}{\sqrt{T}}\right) O_p(\sqrt{\log N_j}) = o_p(T^{-1/2}),$$

where we use $\max_{1\leq i\leq N} w_{j,i}^c = O_p(\sqrt{\log N_j})$ by Assumption WB3-(b) and $\frac{1}{T}\sum_{t=1}^{T} u_{jt}^{(s)} f_t^{c\prime} = O_p(T^{-1/2})$ by Assumption 4-(c). Next, we expand the term (c) by using the definition of $\tilde{e}_{j,it}$, ignoring $\tilde{H}_j^s = O_p(1)$.

$$\frac{1}{T}\sum_{t=1}^{T} \tilde{f}_{jt}^s \tilde{e}_{j,it} = \frac{1}{T}\sum_{t=1}^{T} \tilde{f}_{jt}^s \varepsilon_{j,it} - \underbrace{\frac{1}{T\sqrt{N_j}}\sum_{t=1}^{T} \tilde{f}_{jt}^s u_{jt}^{(c)\prime}\lambda_{j,i}^c}_{=(c_1)} - \underbrace{\frac{1}{T\sqrt{T}}\sum_{t=1}^{T} \tilde{f}_{jt}^s f_t^{c\prime} w_{j,i}^c}_{=(c_2)} = \frac{1}{T}\sum_{t=1}^{T} \tilde{f}_{jt}^s \varepsilon_{j,it} + o_p(T^{-1/2}).$$

We can show the terms (c_1) and (c_2) are $o_p(T^{-1/2})$. Using the similar arguments above, we can show that $(c_1) = O_p((TN_j)^{-1/2}) = o_p(T^{-1/2})$ by Assumption 4-(c). By using the definition such that $\tilde{f}_{jt}^s = f_{jt}^s - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1} f_t^c$, we can show that $(c_2) = 0$. Finally, by plugging all the terms back into expansion of $\hat{\lambda}_{j,i}^s$ and keeping only the terms non-negligible up to order $o_p(T^{-1/2})$, we have the following expansion for $\hat{\lambda}_{j,i}^s$:

$$\hat{\lambda}_{j,i}^s = (\tilde{H}_j^{s\prime})^{-1}\lambda_{j,i}^s + \tilde{H}_j^s \frac{1}{T}\sum_{t=1}^{T} \tilde{f}_{jt}^s \varepsilon_{j,it} + o_p(T^{-1/2}).$$

Since we can show that $\tilde{H}_j^{s\prime}\tilde{H}_j^s = \left(\frac{1}{T}\sum_{t=1}^{T} \tilde{f}_{jt}^s \tilde{f}_{jt}^{s\prime}\right)^{-1} + o_p(T^{-1/2})$, we can show this expansion is equivalent to the expansion of $\hat{\lambda}_{j,i}^s$ in AGGR(2019) (i.e., equation (C.95) in their Online Appendix). ■

Proof of Theorem 4.1. We first verify Theorem 3.1 and then Proposition 3.1. Given Lemma C.2, it suffices to show that the wild bootstrap in Algorithm 1 satisfies Condition D* and Condition E*.

Condition D*: Recall that $\mathcal{B}^* = tr(\tilde{\Sigma}_{\mathcal{U}}^*)$. By the fact that $\eta_{j,it}$ are i.i.d.$N(0,1)$ across (j,i,t),

$$\tilde{\Sigma}_{\mathcal{U}}^* = \frac{1}{T}\sum_{t=1}^{T} E^*(\mathcal{U}_t^*\mathcal{U}_t^{*\prime}) = \mu_N^2 \underbrace{\frac{1}{T}\sum_{t=1}^{T} E^*(u_{1t}^{(c)*}u_{1t}^{(c)*\prime})}_{\equiv\tilde{\Sigma}_{\mathcal{U},11}^*} + \underbrace{\frac{1}{T}\sum_{t=1}^{T} E^*(u_{2t}^{(c)*}u_{2t}^{(c)*\prime})}_{\equiv\tilde{\Sigma}_{\mathcal{U},22}^*} = \mu_N^2\tilde{\Sigma}_{\mathcal{U},11}^* + \tilde{\Sigma}_{\mathcal{U},22}^*.$$

Since $u_{jt}^* \equiv \left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\frac{\tilde{\Lambda}_j'\varepsilon_{jt}^*}{\sqrt{N_j}}$, we can write $\tilde{\Sigma}_{\mathcal{U},jj}^*$ for $j = 1,2$ as follows.

$$\begin{aligned}\tilde{\Sigma}_{\mathcal{U},jj}^* &= \left\{\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{k=1}^{N_j}\tilde{\lambda}_{j,i}\tilde{\lambda}_{j,k}'\frac{1}{T}\sum_{t=1}^{T}E^*(\varepsilon_{j,it}\varepsilon_{j,kt})\right)\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\right\}_{(cc)} \\ &= \left\{\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\tilde{\lambda}_{j,i}\tilde{\lambda}_{j,i}'\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{j,it}^2\right)\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\right\}_{(cc)} = \left\{\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\tilde{\Gamma}_j\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j}\right)^{-1}\right\}_{(cc)},\end{aligned}$$

where we define $\tilde{\Gamma}_j \equiv \frac{1}{N_j}\sum_{i=1}^{N_j}\tilde{\lambda}_{j,i}\tilde{\lambda}_{j,i}'\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{j,it}^2$. Next, recall that $\mathcal{B} = tr(\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{\mathcal{U}})$. For example, by Assumption WB2, we have $\tilde{\Sigma}_{\mathcal{U}} = \mu_N^2\tilde{\Sigma}_{\mathcal{U},11} + \tilde{\Sigma}_{\mathcal{U},22}$, where

$$\tilde{\Sigma}_{\mathcal{U},jj} = \left\{\left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\lambda_{j,i}'\gamma_{j,ii}\right)\left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)^{-1}\right\}_{(cc)},$$

and $\gamma_{j,ii} \equiv \frac{1}{T}\sum_{t=1}^{T}E(\varepsilon_{j,it}^2)$. Our goal is to show that $\tilde{\Sigma}_{\mathcal{U},jj}^* = \tilde{\Sigma}_{\mathcal{U},jj} + o_p(T^{-1/2})$. In fact, since the asymptotic expansions of $\lambda_{j,i}^c$ and $\lambda_{j,i}^s$ are equivalent to those in AGGR (2019), by applying their Lemma B.8, we can show that $\tilde{\Sigma}_{\mathcal{U},jj}^* = \tilde{\Sigma}_{\mathcal{U},jj} + o_p(T^{-1/2})$. In particular, by using asymptotic expansions in Lemma C.3 and by stacking over i for $\hat{\lambda}_{j,i}^c$ and $\hat{\lambda}_{j,i}^s$, we have the following expansions:

$$\hat{\Lambda}_j^c = \left(\Lambda_j^c + \frac{1}{\sqrt{T}}W_j^c + \Lambda_j^s\tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1}\right)(H^c)^{-1} + o_p(T^{-1/2})$$
$$\hat{\Lambda}_j^s = \left(\Lambda_j^s + \frac{1}{\sqrt{T}}W_j^s\right)(\tilde{H}_j^s)^{-1} + o_p(T^{-1/2}),$$

where $W_j^c = \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{jt}f_t^{c\prime}\right)\tilde{\Sigma}_{cc}^{-1}$ and $W_j^s = \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{jt}\tilde{f}_{jt}^{s\prime}\right)\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{jt}^s\tilde{f}_{jt}^{s\prime}\right)^{-1}$. Then, we have the following expansion, which is equivalent to the equation (C.98) in AGGR (2019):

$$\tilde{\Lambda}_j = \left(\Lambda_j + \frac{1}{\sqrt{T}}G_j + \frac{1}{\sqrt{T}}\Lambda_j Q_j\right)\mathcal{H}_j^{-1} + o_p(T^{-1/2}),$$

where

$$G_j = \begin{bmatrix} W_j^c & \vdots & W_j^s \end{bmatrix},\ Q_j = \begin{bmatrix} 0 & 0 \\ \sqrt{T}\tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1} & 0 \end{bmatrix} \text{ and } \mathcal{H}_j = \begin{bmatrix} H^c & 0 \\ 0 & \tilde{H}_j^s \end{bmatrix}.$$

Similar to the arguments in AGGR (2019) and by our Assumption 4-(a) and (c), we can show that

$$\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N_j} = (\mathcal{H}_j')^{-1}\left(\frac{\Lambda_j'\Lambda_j}{N_j} + \frac{1}{\sqrt{T}}(L_{\Lambda,j} + L_{\Lambda,j}')\right)(\mathcal{H}_j)^{-1} + o_p(T^{-1/2}),$$

where $L_{\Lambda,j} = \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right)Q_j$. The rest of the proof similarly follows the arguments in Lemma B.8 in AGGR (2019).

Condition E*: For simplicity, we assume that $k^c = 1$ and $k_j^s = 0$ for $j = 1, 2$ and $N_1 = N_2 = N$. We first derive $\Omega_{\mathcal{U}}^*$. Using Algorithm 1, we have

$$\Omega_{\mathcal{U}}^* = \frac{1}{4}\Big(\frac{1}{T}\sum_{t=1}^{T}Var^*(\mathcal{Z}_{N,t}^*) + \frac{2}{T}\sum_{t=1}^{T}\sum_{s>t}^{T}Cov^*(\mathcal{Z}_{N,t}^*, \mathcal{Z}_{N,s}^*)\Big) = \frac{1}{4}\frac{1}{T}\sum_{t=1}^{T}Var^*(\mathcal{Z}_{N,t}^*),$$

where we use $Cov^*(\mathcal{Z}_{N,t}^*, \mathcal{Z}_{N,s}^*) = 0$ for $t \neq s$ since u_{jt}^* and u_{ks}^* are independent for either $t \neq s$ or $j \neq k$ under Assumption WB2. We can write $Var^*(\mathcal{Z}_{N,t}^*)$ as

$$Var^*(\mathcal{Z}_{N,t}^*) = E^*[z_{1t}^{*2} + z_{2t}^{*2} + 2z_{1t}^*z_{2t}^* + 4u_{1t}^{*2}u_{2t}^{*2} - 4(z_{1t}^* + z_{2t}^*)u_{1t}^*u_{2t}^*],$$

where $z_{jt}^* = u_{jt}^{*2} - E^*(u_{jt}^{*2})$. By Assumption WB2, $E^*(z_{1t}^*z_{2t}^*) = 0$ and $E^*(z_{jt}^*u_{1t}^*u_{2t}^*) = 0$. In addition, we can show that $E^*(z_{jt}^{*2}) = 2\left(\frac{\tilde{\Lambda}_j'\tilde{\Lambda}_j}{N}\right)^{-4}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{k=1}^{N_j}\tilde{\lambda}_{j,i}^2\tilde{\lambda}_{j,k}^2\tilde{\varepsilon}_{j,it}^2\tilde{\varepsilon}_{j,kt}^2$, and $E^*(u_{1t}^{*2}u_{2t}^{*2}) =$

$\left(\frac{\tilde{\Lambda}_1'\tilde{\Lambda}_1}{N}\right)^{-2}\left(\frac{\tilde{\Lambda}_2'\tilde{\Lambda}_2}{N}\right)^{-2}\frac{1}{N^2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\tilde{\lambda}_{1,i}^2\tilde{\lambda}_{2,j}^2\tilde{\varepsilon}_{1,it}^2\tilde{\varepsilon}_{2,jt}^2$. Using these expressions, we can rewrite $\Omega_{\mathcal{U}}^*$ as follows.

$$\begin{aligned}\Omega_{\mathcal{U}}^* = \frac{1}{2}\Bigg[&\left(\frac{\tilde{\Lambda}_1'\tilde{\Lambda}_1}{N_1}\right)^{-4}\frac{1}{N_1^2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_1}\tilde{\lambda}_{1,i}^2\tilde{\lambda}_{1,j}^2\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{1,it}^2\tilde{\varepsilon}_{1,jt}^2\right)\\ &+\left(\frac{\tilde{\Lambda}_2'\tilde{\Lambda}_2}{N_2}\right)^{-4}\frac{1}{N_2^2}\sum_{i=1}^{N_2}\sum_{j=1}^{N_2}\tilde{\lambda}_{2,i}^2\tilde{\lambda}_{2,j}^2\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{2,it}^2\tilde{\varepsilon}_{2,jt}^2\right)\\ &+2\left(\frac{\tilde{\Lambda}_1'\tilde{\Lambda}_1}{N_1}\right)^{-2}\left(\frac{\tilde{\Lambda}_2'\tilde{\Lambda}_2}{N_2}\right)^{-2}\frac{1}{N_1N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\tilde{\lambda}_{1,i}^2\tilde{\lambda}_{2,j}^2\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{1,it}^2\tilde{\varepsilon}_{2,jt}^2\right)\Bigg]\\ \equiv{}& (I)+(II)+(III).\end{aligned}$$

To show that $\Omega_{\mathcal{U}}^* \xrightarrow{p} \Omega_{\mathcal{U}}$, first note that under Assumption WB2, $\Omega_{\mathcal{U}} = \frac{1}{2}(\Sigma_{\mathcal{U},11}+\Sigma_{\mathcal{U},22})^2$, where $\Sigma_{\mathcal{U},jj} = \lim_{N\to\infty}\tilde{\Sigma}_{\mathcal{U},jj}$. The proof follows by showing that (I) and (II) converge in probability to $\Sigma_{\mathcal{U},11}^2$ and $\Sigma_{\mathcal{U},22}^2$, respectively, and (III) converges in probability to $2\Sigma_{\mathcal{U},11}\Sigma_{\mathcal{U},22}$. For each $j=1,2$, we can show that $\frac{1}{N^2}\sum_{i=1}^{N}\sum_{k=1}^{N}\tilde{\lambda}_{j,i}^2\tilde{\lambda}_{j,k}^2\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{j,it}^2\tilde{\varepsilon}_{j,kt}^2 = \Omega_{jj}^2(0)+o_p(1)$, where $\Omega_{jj} \equiv \frac{1}{N}\sum_{i=1}^{N_j}\lambda_{j,i}\gamma_{j,ii}$. By appropriately adding and subtracting, a detailed proof involves three steps (ignoring the group notation): (i) $\frac{1}{N^2}\sum_{i,j=1}^{N}\tilde{\lambda}_i^2\tilde{\lambda}_j^2\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{it}^2\tilde{\varepsilon}_{jt}^2-\varepsilon_{it}^2\varepsilon_{jt}^2\right)=o_p(1)$, (ii) $\frac{1}{N^2}\sum_{i,j=1}^{N}(\tilde{\lambda}_i^2\tilde{\lambda}_j^2-\lambda_i^2\lambda_j^2)\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2\varepsilon_{jt}^2 = o_p(1)$, and (iii) $\frac{1}{N^2}\sum_{i,j=1}^{N}\lambda_i^2\lambda_j^2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2\varepsilon_{jt}^2-\gamma_{ii}\gamma_{jj}\right)=o_p(1)$. By Assumption WB3-(a), we can show that $\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}^2\varepsilon_{jt}^2-E(\varepsilon_{it}^2\varepsilon_{jt}^2)\right)=O_p(T^{-1/2})$, which gives us (iii) $=o_p(1)$. Next, to show that (ii) $=o_p(1)$, we first rewrite the term as follows:

$$\frac{1}{N^2}\sum_{i,j=1}^{N}(\tilde{\lambda}_i^2-\lambda_i^2)\lambda_j^2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2\varepsilon_{jt}^2\right)+\frac{1}{N^2}\sum_{i,j=1}^{N}(\tilde{\lambda}_j^2-\lambda_j^2)\tilde{\lambda}_i^2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2\varepsilon_{jt}^2\right)\equiv(ii-a)+(ii-b),$$

by using that $\tilde{\lambda}_i^2\tilde{\lambda}_j^2-\lambda_i^2\lambda_j^2=(\tilde{\lambda}_i^2-\lambda_i^2)\lambda_j^2+(\tilde{\lambda}_j^2-\lambda_j^2)\tilde{\lambda}_i^2$. Using Cauchy-Schwarz inequality, we have

$$\begin{aligned}(ii-a) &= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}(\tilde{\lambda}_i^2-\lambda_i^2)\varepsilon_{it}^2\right)\left(\frac{1}{N}\sum_{j=1}^{N}\lambda_j^2\varepsilon_{jt}^2\right)\\ &\le \left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N}\sum_{i=1}^{N}(\tilde{\lambda}_i^2-\lambda_i^2)\varepsilon_{it}^2\right|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N}\sum_{j=1}^{N}\lambda_j^2\varepsilon_{jt}^2\right|^2\right)^{1/2}\equiv(ii-aa)^{1/2}(ii-ab)^{1/2}.\end{aligned}$$

For the first term,

$$(ii-aa)\le\underbrace{\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|\varepsilon_{it}|^4\right)}_{=O_p(1)}\left(\frac{1}{N}\sum_{i=1}^{N}|\tilde{\lambda}_i^2-\lambda_i^2|^2\right)=O_p\left(\frac{1}{N}+\frac{\log N}{T}\right),$$

where we use the following fact

$$\frac{1}{N}\sum_{i=1}^{N}(\tilde{\lambda}_i-\lambda_i)^2(\tilde{\lambda}_i+\lambda_i)^2\le\max_{i\le N}|\tilde{\lambda}_i-\lambda_i|^2\frac{1}{N}\sum_{i=1}^{N}(\tilde{\lambda}_i+\lambda_i)^2=O_p\left(\frac{1}{N}+\frac{\log N}{T}\right)$$

by Assumption WB3-(b) and by following Gonçalves and Perron (2020) (see proof of their Lemma B.2.). Finally, for part (i), using that $\tilde{\varepsilon}_{it}^2\tilde{\varepsilon}_{jt}^2 - \varepsilon_{it}^2\varepsilon_{jt}^2 = \tilde{\varepsilon}_{it}^2(\tilde{\varepsilon}_{jt}^2 - \varepsilon_{jt}^2) + (\tilde{\varepsilon}_{it}^2 - \varepsilon_{it}^2)\varepsilon_{jt}^2$ and decompose (i) into two parts: (i-a) and (i-b). We can rewrite (i-a) as follows:

$$
\begin{aligned}
(i-a) &= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}_i^2\tilde{\varepsilon}_{it}^2\right)\left(\frac{1}{N}\sum_{j=1}^{N}\tilde{\lambda}_j^2(\tilde{\varepsilon}_{jt}^2 - \varepsilon_{jt}^2)\right) \\
&\leq \left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}_i^2\tilde{\varepsilon}_{it}^2\right|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N}\sum_{j=1}^{N}\tilde{\lambda}_j^2(\tilde{\varepsilon}_{jt}^2 - \varepsilon_{jt}^2)\right|^2\right)^{1/2} \equiv (i-aa)^{1/2}(i-ab)^{1/2}.
\end{aligned}
$$

We can show that $(i-aa) = O_p(1)$. Since we can write $(i-ab) \leq \left(\frac{1}{N}\sum_{i=1}^{N}|\tilde{\lambda}_i|^4\right)\left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}|\tilde{\varepsilon}_{it}^2 - \varepsilon_{it}^2|^2\right)$, our goal is to show that $\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}|\tilde{\varepsilon}_{it}^2 - \varepsilon_{it}^2|^2 = o_p(1)$. Since $\tilde{\varepsilon}_{it}^2 - \varepsilon_{it}^2 = (\tilde{\varepsilon}_{it} - \varepsilon_{it})(\tilde{\varepsilon}_{it} + \varepsilon_{it})$, we have

$$
\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}|\tilde{\varepsilon}_{it}^2 - \varepsilon_{it}^2|^2 &\leq \frac{1}{T}\sum_{t=1}^{T}\max_{i\leq N}(\tilde{\varepsilon}_{it} - \varepsilon_{it})^2\frac{1}{N}\sum_{i=1}^{N}(\tilde{\varepsilon}_{it} + \varepsilon_{it})^2 \\
&\left(\frac{1}{T}\sum_{t=1}^{T}\left|\max_{i\leq N}(\tilde{\varepsilon}_{it} - \varepsilon_{it})^2\right|^2\right)^{1/2}\underbrace{\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N}\sum_{i=1}^{N}(\tilde{\varepsilon}_{it} + \varepsilon_{it})^2\right|^2\right)^{1/2}}_{=O_p(1)}.
\end{aligned}
$$

By applying similar arguments in Gonçalves and Perron (2020) (proof of their Lemma B.2), we can show that $\frac{1}{T}\sum_{t=1}^{T}\left|\max_{i\leq N}(\tilde{\varepsilon}_{it} - \varepsilon_{it})^2\right|^2 = O_p\left(\left(\frac{1}{N} + \frac{\log N}{T}\right)^2\right)$ by Lemma C.1 and Assumption WB3-(b). The proof for (i-b) is similar.

Next, we show that $\frac{1}{\sqrt{T}}\Omega_{\mathcal{U}}^{*-1/2}\sum_{t=1}^{T}Z_{N,t}^* \xrightarrow{d^*}_p N(0,1)$. We let $\omega_{N,t}^* \equiv (\Omega_{\mathcal{U}})^{*-1/2}Z_{N,t}^*$ (given that $Z_{N,t}^*$ depends on η_{1t} and η_{2t}, $Z_{N,t}^*$ is an independent array) and apply a CLT for heterogeneous independent random vectors on $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\omega_{N,t}^*$. We have $E^*(\omega_{N,t}^*) = 0$ and $Var^*(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\omega_{N,t}^*) = 1$. Therefore, it suffices to show that $E^*|\omega_{N,t}^*|^{2d} = O_p(1)$ for some $d > 1$ (Lyapunov's condition) and it is sufficient to show that $E^*|Z_{N,t}^*|^{2d} = O_p(1)$ ($E^*|\omega_{N,t}^*|^{2d} \leq |\Omega_{\mathcal{U}}^{*-1/2}|^{2d}E^*|Z_{N,t}^*|^{2d}$). Note that $Z_{N,t}^* = z_{1,Nt}^* + z_{2,Nt}^* - 2u_{1t}^*u_{2t}^*$, where $z_{jt}^* = u_{jt}^{*2} - E^*(u_{jt}^{*2})$. By applying the c_r-inequality, we have

$$
E^*|Z_{N,t}^*|^{2d} \leq 3^{2d-1}\left(E^*|z_{1,Nt}^*|^{2d} + E^*|z_{2,Nt}^*|^{2d} + E^*|2u_{1t}^*u_{2t}^*|^{2d}\right).
$$

We need to show that $E^*|z_{jt}^*|^{2d} = O_p(1)$ and $E^*|u_{1t}^*u_{2t}^*|^{2d} = O_p(1)$. To show that $E^*|u_{1t}^*u_{2t}^*|^{2d} = O_p(1)$, with $d=2$, it suffices to show that $E^*|u^*_{jt}|^4 = O_p(1)$ ($\because E^*|u^*_{1t}u^*_{2t}|^{2d} \le E^*|u^*_{1t}|^{2d}E^*|u^*_{2t}|^{2d}$) as follows.

$$\begin{aligned}
E^*|u^*_{jt}|^4 &\le \left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)^{-4} E^*\left|\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N}\tilde{\lambda}_{j,i}\varepsilon^*_{j,it}\right|^4 \\
&= \left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)^{-4} \frac{1}{N_j^2}\sum_{i_1,i_2,i_3,i_4}^{N_j} \tilde{\lambda}_{j,i_1}\tilde{\lambda}_{j,i_2}\tilde{\lambda}_{j,i_3}\tilde{\lambda}_{j,i_4}E^*(\varepsilon^*_{j,i_1t}\varepsilon^*_{j,i_2t}\varepsilon^*_{j,i_3t}\varepsilon^*_{j,i_4t}) \\
&\le \bar{\eta}\left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)^{-4} \frac{1}{N^2}\sum_{i_1,i_2}^{N}\tilde{\lambda}^2_{j,i_1}\tilde{\lambda}^2_{j,i_2}\tilde{\varepsilon}^2_{j,i_1t}\tilde{\varepsilon}^2_{j,i_2t} = \bar{\eta}\left(\frac{\tilde{\Lambda}'_j\tilde{\Lambda}_j}{N_j}\right)^{-4}\left(\frac{1}{N_j}\sum_{i=1}^{N}\tilde{\lambda}^2_{j,i}\tilde{\varepsilon}^2_{j,it}\right)^2,
\end{aligned}$$

where $E^*(\varepsilon^*_{j,i_1t}\varepsilon^*_{j,i_2t}\varepsilon^*_{j,i_3t}\varepsilon^*_{j,i_4t}) \le \bar{\eta}_4 \equiv \max\{E^*(\eta^4_{j,it}),1\}$ and $E^*(\eta^4_{j,it}) < C$. Next, we show that $E^*|z^*_{jt}|^{2d} = O_p(1)$. Since $z^*_{jt} = u^{*2}_{jt} - E^*(u^{*2}_{jt})$, we have

$$\begin{aligned}
E^*|z^*_{jt}|^{2d} &= E^*|u^{*2}_{jt} - E^*(u^{*2}_{jt})|^{2d} \\
&\le 2^{2d-1}(E^*|u^{*2}_{jt}|^{2d} + E^*|E^*(u^{*2}_{jt})|^{2d}) \le CE^*|u^*_{jt}|^{4d},
\end{aligned}$$

where C is some positive constant. Taking $d=2$, it is sufficient to show that

$$E^*|u^*_{jt}|^8 = \frac{1}{N_j^4}\sum_{i_1,\ldots,i_8}\tilde{\lambda}_{j,i_1}\ldots\tilde{\lambda}_{j,i_8}\tilde{\varepsilon}_{j,i_1t}\ldots\tilde{\varepsilon}_{j,i_8t}E^*(\eta_{j,i_1t}\ldots\eta_{j,i_8t}) = O_p(1).$$

Since $\eta_{j,it} \sim$ i.i.d.$N(0,1)$ we have four cases to consider. If $i_1 = \ldots = i_8$, we have $\frac{1}{N_j^4}\sum_{i=1}^{N}\tilde{\lambda}^8_{j,i}\tilde{\varepsilon}^8_{j,it} = E^*(\eta^8_{j,it})\frac{1}{N^3}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}^8_{j,i}\tilde{\varepsilon}^8_{j,it}\right) = O_p(1)$ since $E^*|\eta_{j,it}|^8 < C$. For the second case, we consider $i_1 = i_2, \ldots, i_7 = i_8$, we have $E^*|u^*_{jt}|^8 = \frac{1}{N^4}\sum_{i\ne m\ne k\ne l}\tilde{\lambda}^2_{j,i}\tilde{\lambda}^2_{j,m}\tilde{\lambda}^2_{j,k}\tilde{\lambda}^2_{j,l}\tilde{\varepsilon}^2_{j,it}\tilde{\varepsilon}^2_{j,mt}\tilde{\varepsilon}^2_{j,kt}\tilde{\varepsilon}^2_{j,lt}$. The third case is $i_1 = i_2 = i_3 = i_4$ and $i_5 = \ldots = i_8$. In this case, we can bound it as $C_1\frac{1}{N^2}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}^4_{j,i}\tilde{\varepsilon}^4_{j,it}\right)^2$ since $E^*|\eta_{j,it}|^4 < C_1$. Finally, we consider when $i_1 = \ldots = i_6$ and $i_7 = i_8$, and in this case, we can bound the term as $C_2\frac{1}{N^2}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}^6_{j,i}\tilde{\varepsilon}^6_{j,it}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}^2_{j,i}\tilde{\varepsilon}^2_{j,it}\right)$, where we use $E^*|\eta_{j,it}|^6 < C_2$.

Finally, we show that under the alternative hypothesis, the wild bootstrap method satisfies Condition F* and therefore Proposition 3.1 follows. Under the alternative hypothesis, we have no group common factor in our simple setting. The one factor that we extract from each group is group-specific factor. Provided that the group common factor, $\hat{f}^c_t$ is estimated from using the first group factor, $\hat{f}_{1t}$, there is an additional bias term appear only in the second group. Particularly, we have

$$\tilde{\lambda}_{2,i} = \Phi\lambda_{2,i} + o_p(1), \tag{15}$$

$$\tilde{\varepsilon}_{2,it} = \varepsilon_{2,it} + \lambda_{2,i}(f_{2t} - \Phi f_{1t}) + o_p(1). \tag{16}$$

Note that $\Phi = corr(f_{1t}, f_{2t})$ and under the alternative hypothesis, the estimated factor loadings in the second group are consistently estimate only a portion of the true factor loadings of the second group. Moreover, the residual term in the second group will be containing the bias term. Using (15) and (16), we can rewrite the term related to the second group in bootstrap bias, $\mathcal{B}^*$ as following:

$$\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\tilde{\lambda}_{2,i}^2\right)^{-2}\frac{1}{N_2}\sum_{i=1}^{N_2}\tilde{\lambda}_{2,i}^2\frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{2,it}^2 = \frac{1}{\Phi^2}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^2\right)^{-2}\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{2,it}^2\right)$$
$$+\frac{2}{\Phi^2}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^2\right)^{-2}\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^3\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{2,it}(f_{2t}-\Phi f_{1t})\right)$$
$$+\frac{1}{\Phi^2}\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^2\right)^{-2}\frac{1}{N_2}\sum_{i=1}^{N_2}\lambda_{2,i}^4\left(\frac{1}{T}\sum_{t=1}^{T}(f_{2t}-\Phi f_{1t})^2\right)+o_p(1).$$

We can show that all the terms are $O_p(1)$ under Assumption 2 and 3.

We also need to show that $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Z_{N,t}^* = O_p(1)$ under the alternative hypothesis. To show this, it is sufficient to show that $Var^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Z_{N,t}^*\right) = O_p(1)$. Under the Assumption WB2, we have $Cov(Z_{N,t}^*, Z_{N,s}^*) = 0$ and $Var^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Z_{N,t}^*\right) = \frac{1}{T}\sum_{t=1}^{T}Var(Z_{N,t}^*)$. Since we showed that $Var(Z_{N,t}^*) = E^*(z_{1t}^{*2})+E^*(z_{2t}^{*2})+4E^*(u_{1t}^{*2}u_{2t}^{*2})$ in the proof of Condition E*, where $z_{jt}^* = u_{jt}^{*2}-E^*(u_{jt}^{*2})$, we focus on three terms: $E^*(z_{1t}^{*2})$, $E^*(z_{2t}^{*2})$ and $E^*(u_{1t}^{*2}u_{2t}^{*2})$. We can show that the first term is $O_p(1)$ as follows.

$$E^*(z_{1t}^{*2}) = 2\left(\frac{\tilde{\Lambda}_1'\tilde{\Lambda}_1}{N_1}\right)^{-2}\frac{1}{N^2}\sum_{i,j=1}^{N_1}\tilde{\lambda}_{1,i}^2\tilde{\lambda}_{1,j}^2\tilde{\varepsilon}_{1,it}^2\tilde{\varepsilon}_{1,jt}^2 = 2\left(\frac{\Lambda_1'\Lambda_1}{N_1}\right)^{-2}\frac{1}{N_1^2}\sum_{i,j=1}^{N}\lambda_{1,i}^2\lambda_{1,j}^2\varepsilon_{1,it}^2\varepsilon_{1,jt}^2+o_p(1),$$

where we can obtain the second equality by using $\tilde{\lambda}_{1,i} = \lambda_{1,i}+o_p(1)$ and $\tilde{\varepsilon}_{1,it} = \varepsilon_{1,it}+o_p(1)$. This is true when we use the factor from the first group as the group common factor. Then, we can show that $\frac{1}{T}\sum_{t=1}^{T}E^*(z_{1t}^{*2}) = O_p(1)$ under our assumptions. For the second term, we have

$$\begin{aligned}
E^*(z_{2t}^{*2}) &= 2\left(\frac{\tilde{\Lambda}_2'\tilde{\Lambda}_2}{N_2}\right)^{-2}\frac{1}{N_2^2}\sum_{i,j=1}^{N_2}\tilde{\lambda}_{2,i}^2\tilde{\lambda}_{2,j}^2\tilde{\varepsilon}_{2,it}^2\tilde{\varepsilon}_{2,jt}^2 \\
&= \frac{2}{\Phi^2}\left(\frac{\Lambda_2'\Lambda_2}{N_2}\right)^{-2}\frac{1}{N_2^2}\sum_{i,j=1}^{N_2}\lambda_{2,i}^2\lambda_{2,j}^2(\varepsilon_{2,it}+\lambda_{2,i}(f_{2t}-\Phi f_{1t}))^2(\varepsilon_{2,jt}+\lambda_{2,j}(f_{2t}-\Phi f_{1t}))^2 \\
&= \frac{2}{\Phi^2}\left(\frac{\Lambda_2'\Lambda_2}{N_2}\right)^{-2}\frac{1}{N_2^2}\sum_{i,j=1}^{N_2}\lambda_{2,i}^2\lambda_{2,j}^2[\varepsilon_{2,it}^2\varepsilon_{2,jt}^2+\lambda_{2,j}^2\varepsilon_{2,it}^2(f_{2t}-\Phi f_{1t})^2+2\lambda_{2,j}\varepsilon_{2,it}^2\varepsilon_{2,jt}(f_{2t}-\Phi f_{1t}) \\
&\quad+\lambda_{2,i}^2\varepsilon_{2,jt}^2(f_{2t}-\Phi f_{1t})^2+\lambda_{2,i}^2\lambda_{2,j}^2(f_{2t}-\Phi f_{1t})^4+2\lambda_{2,i}^2\lambda_{2,j}\varepsilon_{2,jt}(f_{2t}-\Phi f_{1t})^3+2\lambda_{2,i}\varepsilon_{2,it}\varepsilon_{2,jt}^2(f_{2t}-\Phi f_{1t}) \\
&\quad+2\lambda_{2,i}\lambda_{2,j}^2\varepsilon_{2,it}(f_{2t}-\Phi f_{1t})^3+4\lambda_{2,i}\lambda_{2,j}\varepsilon_{2,it}\varepsilon_{2,jt}(f_{2t}-\Phi f_{1t})^2],
\end{aligned}$$

where we use (15) and (16) to obtain the second equality. Using the final equality, we can show that $\frac{1}{T}\sum_{t=1}^{T}E^*(z_{2t}^{*2}) = O_p(1)$. Similarly, we can show that $E^*(u_{1t}^{*2}u_{2t}^{*2}) = O_p(1)$. ■

References

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

Andreou, E., Gagliardini, P., Ghysels, E., & Rubin, M. (2022). *Three Common Factors* (CEPR Discussion Papers No. 17225). C.E.P.R. Discussion Papers. Retrieved from `https://ideas.repec.org/p/cpr/ceprdp/17225.html`

Andreou, E., Gagliardini, P., Ghysels, E., & Rubin, M. (2024). Spanning latent and observable factors. *Journal of Econometrics*, 105743. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0304407624000897` doi: https://doi.org/10.1016/j.jeconom.2024.105743

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

Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, *74*(4), 1133–1150.

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

Bickel, P. J., & Levina, E. (2008a). Covariance regularization by thresholding. *The Annals of Statistics*, *36*(6), 2577–2604.

Bickel, P. J., & Levina, E. (2008b). Regularized estimation of large covariance matrices. *The Annals of Statistics*, *36*(1), 199–227.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory.* SIAM.

Cai, T., & Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association*, *106*(494), 672–684.

Chamberlain, G., & Rothschild, M. (1982). *Arbitrage, factor structure, and mean-variance analysis on large asset markets.* National Bureau of Economic Research Cambridge, Mass., USA.

Davidson, R., & MacKinnon, J. G. (1999). The size distortion of bootstrap tests. *Econometric Theory*, *15*(3), 361–376.

Eichengreen, B., Mody, A., Nedeljkovic, M., & Sarno, L. (2012). How the subprime crisis went global: Evidence from bank credit default swap spreads. *Journal of International Money and Finance*, *31*(5), 1299–1318.

Fan, J., Liao, Y., & Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. *The Annals of Statistics*, *39*(6), 3320.

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

Fan, J., Liao, Y., & Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica*, *83*(4), 1497–1541.

Gagliardini, P., Ossola, E., & Scaillet, O. (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, *84*(3), 985–1046.

Gallant, A., & White, H. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* B. Blackwell. Retrieved from `https://books.google.ca/books?id=VVOqQgAACAAJ`

Gonçalves, S. (2011). The moving blocks bootstrap for panel linear regression models with individual fixed effects. *Econometric Theory*, *27*(5), 1048–1082.

Gonçalves, S., & Perron, B. (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics*, *182*(1), 156–173.

Gonçalves, S., & Perron, B. (2020). Bootstrapping factor models with cross sectional dependence. *Journal of Econometrics*, *218*(2), 476–495.

Gonçalves, S., McCracken, M. W., & Perron, B. (2017). Tests of equal accuracy for nested models with estimated factors. *Journal of Econometrics*, *198*(2), 231-252. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0304407617300180` doi: https://doi.org/10.1016/j.jeconom.2017.01.004

Gospodinov, N., & Ng, S. (2013). Commodity prices, convenience yields, and inflation. *Review of Economics and Statistics*, *95*(1), 206–219.

Kock, A. B., & Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, *186*(2), 325–344.

Koh, J. (2022). *Bootstrapping factor-midas regression models.*

Krampe, J., & Margaritella, L. (2021). Dynamic factor models with sparse var idiosyncratic components. *arXiv preprint arXiv:2112.07149*.

Krampe, J., & Paparoditis, E. (2021). Sparsity concepts and estimation procedures for high-dimensional vector autoregressive models. *Journal of Time Series Analysis*, *42*(5-6), 554–579.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics*, *83*(1), 171–222.

Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies*, *22*(12), 5027–5067.

MacKinnon, J. G. (2009). Bootstrap hypothesis testing. *Handbook of Computational Econometrics*, *183*, 213.

McMurry, T. L., & Politis, D. N. (2010). Banded and tapered estimates for autocovariance matrices and the linear process bootstrap. *Journal of Time Series Analysis*, *31*(6), 471–482.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management science*, *60*(7), 1772–1791.

Rothman, A. J., Levina, E., & Zhu, J. (2009). Generalized thresholding of large covariance matrices. *Journal of the American Statistical Association*, *104*(485), 177–186.

Schott, J. R. (2016). *Matrix Analysis for Statistics.* John Wiley & Sons.

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

Wu, J. C., & Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, *48*(2-3), 253–291.