Supplementary Material for "One-way or Two-way Factor Model for Matrix Sequences?"

Yong He

Shandong University, China

Xinbing Kong

Nanjing Audit University, China

Lorenzo Trapani

University of Leicester, United Kingdom of Great Britain

and

Long Yu

Shanghai University of Finance and Economics, China

In this Supplement, we further discuss some of our main assumptions, also providing examples (Section A); we discuss several aspects of the implementation of our procedures, and report further Monte Carlo evidence (Section B); we report all preliminary lemmas (Section C); finally, all proofs (both of the results in the main paper, and of the ancillary results reported in this Supplement) are relegated to Section D.

Contents

A. Discussion: model and assumptions

A.1. A macroeconomic example for the two-way matrix factor model

In addition to our discussion in Section 1.1, we consider a further toy example, aimed at better understanding the nature of (1.3), and its relationship with a hierarchical factor model. Our example is a re-elaboration of an example in Wang et al. (2019), where X_t is a matrix of time series whose rows contain some macroeconomic indicator (GDP, inflation, interest rates...) and whose columns represent different countries. Each row $X_{i\cdot,t}$ contains the values of macroeconomic indicator i for all the p_2 countries, and we assume that $X_{i\cdot,t}$ follows a vector factor model

$$X_{i\cdot,t} = G_{i\cdot,t} C^{(i)\prime} + H_{i\cdot,t}, \tag{A.1}$$

where the $1 \times k_2$ vector $G_{i\cdot,t}$ contains the common factors that drive $X_{i\cdot,t}$, and $C^{(i)\prime}$ is a $k_2 \times p_2$ matrix of loadings. Each macroeconomic indicator is driven by a k_2-dimensional vector of common factors, $G_{i\cdot,t}$, $1 \leq i \leq p_1$, which can be organised in a $p_1 \times k_2$ matrix, G_t. Such factors may share comovements between the countries, i.e., considering the j-th column of G_t, $G_{\cdot j,t}$ with $1 \leq j \leq p_2$

$$G_{\cdot j,t} = R^{(j)} F_{\cdot j,t} + H^*_{\cdot j,t}.$$

Upon assuming that the loadings $R^{(j)}$ are equal to R for all $1 \leq j \leq p_1$, this entails

$$G_t = R F_t + H^*_t, \tag{A.2}$$

and subsequently, assuming $C^{(i)} = C$ for all $1 \leq i \leq p_1$ and inserting (A.2) into (A.1), it ultimately follows that

$$X_t = R F_t C' + H^*_t C' + H_t = R F_t C' + E_t.$$

This example illustrates the relationship between (1.3) and a multilevel factor model, of which (1.3) is a restricted, more parsimonious, version.

A.2. Assumption B1: serial dependence in factors and errors

We consider some specific examples of serial dependence in F_t and E_t. In particular, we assume that the common factors $\mathrm{Vec}(F_t)$ and the individual errors $e_{ij,t}$ can be represented as causal processes, viz.

$$\begin{aligned}
\mathrm{Vec}\left(F_t\right) &= g^F\left(\eta_t, ..., \eta_{-\infty}\right), && \text{(A.3)}\\
e_{ij,t} &= g^e_{ij}\left(\eta_{ij,t}, ..., \eta_{ij,-\infty}\right), && \text{(A.4)}
\end{aligned}$$

where $g^F : \mathbb{R}^{(-\infty,\infty)\times...\times(-\infty,\infty)} \to \mathbb{R}^{k_1\times k_2}$ and $g^e_{ij} : \mathbb{R}^{(-\infty,\infty)} \to \mathbb{R}$ are measurable functions such that the random variables $\mathrm{Vec}(F_t)$ and $e_{ij,t}$ are well-defined, and $\{\eta_t, -\infty < t < \infty\}$ and $\{\eta_{ij,t}, -\infty < t < \infty\}$ are *i.i.d.*, zero mean sequences. According to (A.3) and (A.4), $\mathrm{Vec}(F_t)$ and $e_{ij,t}$ are stationary and ergodic processes; the causal process/"physical system" representation in (A.3) and (A.4) is now a standard way of modelling dependence, after the seminal contribution by Wu (2005) (see also Aue et al., 2009 and Berkes et al., 2011).

We define the quantities

$$\begin{aligned}
\delta^F_{t,p} &= \left|g^F\left(\eta_t, ..., \eta_0, ..., \eta_{-\infty}\right) - g^F\left(\eta_t, ..., \eta'_0, ..., \eta_{-\infty}\right)\right|_p,\\
\delta^{e,ij}_{t,p} &= \left|g^e_{ij}\left(\eta_{ij,t}, ..., \eta_{ij,0}, ..., \eta_{ij,-\infty}\right) - g^e_{ij}\left(\eta_{ij,t}, ..., \eta'_{ij,0}, ..., \eta_{ij,-\infty}\right)\right|_p,
\end{aligned}$$

where η'_0 and $\eta'_{ij,0}$ are independent copies of η_0 and $\eta_{ij,0}$ respectively, such that $\eta'_0 \overset{D}{=} \eta_0$ and $\eta'_{ij,0} \overset{D}{=} \eta_{ij,0}$ and η'_0 and $\eta'_{ij,0}$ are independent of $\{\eta_t, -\infty < t < \infty\}$ and $\{\eta_{ij,t}, -\infty < t < \infty\}$ respectively.

Lemma A.1. *We assume that (A.3) holds, with $|F_{ij,t}|_p < \infty$ for all $1 \le i \le k_1$ and $1 \le j \le k_2$, and $p = 4 + \epsilon$ for some $\epsilon > 0$ and $\sum_{t=0}^{\infty} \delta^F_{t,p} < \infty$. Then Assumption B1(iii) is satisfied.*

Lemma A.2. *We assume that (A.4) holds, with $|e_{ij,t}|_4 < \infty$ for all $1 \le i \le p_1$ and $1 \le j \le p_2$, and $\sum_{t=1}^{T} \sum_{m=t}^{\infty} \delta^{e,ij}_{m,q} < \infty$ for all $q \le 4$. Then it holds that*

$$\sum_{s=1}^{T} \left|E\left(e_{ij,t}e_{lh,s}\right)\right| \le c_0, \tag{A.5}$$

for all $1 \le i, l \le p_1$ and $1 \le j, h \le p_2$, and

$$\sum_{s=1}^{T} \left|Cov\left(e_{ij,t}e_{i_1j_1,t}, e_{lh,s}e_{l_1h_1,s}\right)\right| \le c_0, \tag{A.6}$$

for all $1 \le i, l, i_1, l_1 \le p_1$ and $1 \le j, h, j_1, h_1 \le p_2$.

Representation (A.3) and (A.4) - together with the assumption that, essentially, the dependence measures $\delta_{t,p}^F$ and $\delta_{t,p}^{e,ij}$ are summable - allows for a wide variety of data generating processes. Indeed, as shown e.g. in Barigozzi and Trapani (2022), many commonly employed models result in having $\delta_{t,p}^F$ and $\delta_{t,p}^{e,ij}$ *exponentially* decay in t. Hereafter, we mention the most important examples of DGPs that satisfy our assumptions.

Liu and Lin (2009) show that (A.3) and (A.4), together with the summability conditions, hold for (multivariate) linear processes: hence, all our theory can be used if either $\mathrm{Vec}(F_t)$ or $e_{ij,t}$ (or both) follow a stationary (V)ARMA model. In addition, Liu and Lin (2009) (see their Corollaries 3.1-3.7) prove that (A.3) and (A.4), and the summability conditions, hold for several nonlinear models, e.g. for stationary Random Coefficient AutoRegressions, nonlinear transformations of VARMA models, threshold autoregressive models, and other non-linear mappings. Aue et al. (2009) study several multivariate GARCH processes, concluding that (A.3) and (A.4), together with the summability conditions, hold for e.g. the CCC-GARCH of Bollerslev (1990) (and for the CCC-GARCH studied by Jeantheau, 1998), and for the multivariate exponential GARCH (Kawakatsu, 2006) - see also Barigozzi and Trapani (2022). In all the cases mentioned above, the summability conditions hold with $\delta_{t,p}^F$ and $\delta_{t,p}^{e,ij}$ decaying exponentially in t.

We also point out that, as far as univariate GARCH models are concerned, Barigozzi and Trapani (2022) (see their Corollary 9) have shown the validity of our assumptions for the broad class of augmented GARCH models (including standard GARCH, the Threshold GARCH model, and the Exponential GARCH model).

A.3. Assumption B1(iv): discussion and alternative asymptotics

We consider an alternative scenario to Assumption B1*(iv)*. In the paper, we have assumed that, when $k_1 = 0$ and $k_2 > 0$, it holds that

$$\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}F_tF_t'\right) = O_{a.s.}\left(\left(1+\sqrt{\frac{p_1}{T}}\right)^2\right), \tag{A.7}$$

and we have made appeal to the literature on Random Matrix Theory (El Karoui, 2005) to justify this assumption. Equation (A.7) is a high-level requirement, which holds under more primitive assumptions; as a leading example, this result holds for *i.i.d.* data when the relative rate of divergence between p_1 and T is restricted to lie in $(0, 1]$ as $\min\{p_1, T\} \to \infty$ (see Theorem 2 in El Karoui, 2005). Extensions of this result have also been derived to consider the case of dependent and even heavy tailed data (see e.g. Davis et al., 2014).

As an alternative, consider the following more primitive version of Assumption B1*(iv)*.

Assumption 1. *Assumption B1 holds, with part (iv) replaced by:*

(a) when $k_2 = 0$ and $k_1 > 0$, it holds that

$$E\left(F_t' F_t\right) = \widetilde{\Sigma}_1^* \quad \textit{and} \quad \frac{1}{Tp_2}\sum_{t=1}^{T} F_t F_t' \xrightarrow{a.s.} \Sigma_1^*,$$

with $\lambda_{\max}\left(\widetilde{\Sigma}_1^*\right) < \infty$, and Σ_1^* a $k_1 \times k_1$ positive definite matrix with distinct eigenvalues and $\lambda_{\max}\left(\Sigma_1^*\right) < \infty$;

(b) when $k_1 = 0$ and $k_2 > 0$, it holds that

$$E\left(F_t' F_t\right) = \widetilde{\Sigma}_2^* \quad \textit{and} \quad \frac{1}{Tp_1}\sum_{t=1}^{T} F_t' F_t \xrightarrow{a.s.} \Sigma_2^*,$$

with $\lambda_{\max}\left(\widetilde{\Sigma}_2^*\right) < \infty$, Σ_2^* a $k_2 \times k_2$ positive definite matrix with distinct eigenvalues and $\lambda_{\max}\left(\Sigma_2^*\right) < \infty$.

In Assumption 1, the RMT-type results on the largest eigenvalue of $T^{-1}\sum_{t=1}^{T} F_t' F_t$ (and $T^{-1}\sum_{t=1}^{T} F_t F_t'$) are replaced by, essentially, assuming that $E\left(F_t' F_t\right)$ (and $E\left(F_t F_t'\right)$) have bounded eigenvalues.

Corollary A.1. *We assume that the assumptions of Theorem 1 are satisfied, with Assumption B1 replaced with Assumption 1. Then the theorem holds, with (2.4) modified as*

$$\widehat{\lambda}_j = c_0 + o_{a.s.}\left(\frac{p_1}{\sqrt{T}}\left(\ln^2 p_1 \ln T\right)^{1/2+\epsilon}\right) + o_{a.s.}\left(\frac{p_1}{\sqrt{Tp_2}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right). \tag{A.8}$$

Further, under the assumptions of Theorem 2, with again Assumption B1 replaced with Assumption 1, the same results in the theorem hold with (2.7) modified as

$$\widetilde{\lambda}_j = o_{a.s.}\left(\frac{p_1}{\sqrt{T}}\left(\ln^2 p_1 \ln T\right)^{1/2+\epsilon}\right) + o_{a.s.}\left(\left(\frac{1}{p_2} + \frac{1}{T} + \frac{p_1}{\sqrt{Tp_2}}\right)\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1+\epsilon}\right). \tag{A.9}$$

The practical consequence of (A.8) and (A.9) is that, when testing for

$$\begin{array}{ll} H_0: & k_1 \geq 1 \\ H_A: & k_1 = 0 \end{array},$$

there is still a (narrower) eigen-gap between the null and the alternative; interestingly, the rates in such an eigen-gap is the same as found for vector time series (see e.g. Trapani, 2018). In this case, the test can be run as described in Section 3.1, using

$$\widehat{\phi}_1 = \exp\left\{\frac{p_1^{-\delta}\widehat{\lambda}_1}{p_1^{-1}\sum_{j=1}^{p_1}\widehat{\lambda}_j}\right\} - 1 \text{ and } \widetilde{\phi}_1 = \exp\left\{p_1^{-\delta}\frac{\widetilde{\lambda}_1}{p_1^{-1}\sum_{j=1}^{p_1}\widetilde{\lambda}_j}\right\} - 1;$$

with

$$\begin{cases} \delta = \varepsilon & \text{if } \beta'' \leq 1/2 \\ \delta = 1 - 1/(2\beta'') + \varepsilon & \text{if } \beta'' > 1/2 \end{cases}, \tag{A.10}$$

where $\varepsilon > 0$ is an arbitrarily small user-chosen quantity, and $\beta'' = \ln p_1 / \ln T$.

A.4. Assumption B4(ii): dependence between factors and idiosyncratic components

We now turn to commenting on Assumption B4*(ii)*. The assumption stipulates that, *inter alia*

$$\Big\| \sum_{l=1}^{p_1}\sum_{h_2=1}^{p_2} Cov(\bar{\zeta}_{ij}\otimes\bar{\zeta}_{ih_1}, \bar{\zeta}_{lj}\otimes\bar{\zeta}_{lh_2}) \Big\|_{\max} \leq c_0,$$

where recall that we have defined $\bar{\zeta}_{ij} = T^{-1/2}\text{Vec}\Big(\sum_{t=1}^{T} F_t e_{ij,t}\Big)$. This high-level requirement is exactly the same as Assumption E2 in Yu et al. (2022), and it serves the same purpose: in particular, it allows for the initial estimator of C, say $\widehat{C}$, to be consistent at the rate stated in Lemma C.2.

Standard (if tedious) calculations reveal that this assumption holds in the simplistic cases of $\{F_t, 1 \leq t \leq T\}$ and $\{e_{ij,t}, 1 \leq t \leq T\}$ being two mutually independent groups, and assuming weak serial and cross-sectional dependence. Hereafter, we provide an example of how this assumption may hold when F_t and $e_{ij,t}$ are not necessarily independent, and we also provide an alternative to Assumption B4*(ii)*. In order not to overshadow our main arguments with algebra, we will henceforth assume $k_1 = k_2 = 1$, and we consider the leading example where F_t and $e_{ij,t}$ are dependent through

$$e_{ij,t} = g\left(F_t\right) u_{ij,t}, \tag{A.11}$$

where $g : \mathbb{R}^{k_1 k_2} \to \mathbb{R}$ is a measurable function, and $\{F_t, 1 \leq t \leq T\}$ and $\{u_{ij,t}, 1 \leq t \leq T\}$ are two mutually independent groups for all i, j. Equation (A.11) is also proposed - in the simpler form $e_{ij,t} = F_t u_{ij,t}$ - as an illustrative example to Assumption D in Bai (2003).

We begin by briefly discussing some examples in which Assumption B4*(ii)* holds.

Example 1 (Weak temporal dependence)**.** *In the first illustrative example, we consider a moving average process, i.e.,*

$$u_{ij,t} = \alpha v_{i-1,jt} + \beta v_{i,j-1,t} + \gamma v_{ij,t-1},$$

assuming that $v_{ij,t}$ is i.i.d. standard normal. Therefore, $u_{ij,t}$ is m-dependent - i.e., it is related to only a finite number of neighbors, and independent of other neighbors. Then,

$$\sum_{l,h_2} Cov(\bar{\zeta}_{ij}\bar{\zeta}_{ih_1}, \bar{\zeta}_{lj}\bar{\zeta}_{lh_2}) = \sum_{l,h_2} \frac{1}{T^2} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_2=1}^{T} Cov(F_{t_1}^2 F_{t_2}^2 u_{ijt_1} u_{ih_1t_2}, F_{s_1}^2 F_{s_2}^2 u_{ljs_1} u_{lh_2s_2})$$
$$\leq \sum_{|s_1-t_1|\leq c_0, |s_2-t_2|\leq c_1, |l-i|\leq c_2, |h_1-h_2|\leq c_3} \frac{1}{T^2} \sum_{t_1} \sum_{t_2} |Cov(F_{t_1}^2 F_{t_2}^2 u_{ijt_1} u_{ih_1t_2}, F_{s_1}^2 F_{s_2}^2 u_{ljs_1} u_{lh_2s_2})| \leq C.$$

Example 2 (Cross-sectional autoregression)**.** *We consider, as a second example, a cross-sectional autoregression, and we assume serial independence for simplicity. Let $\mathbf{u}_t$ denote the p_1p_2-dimensional vector obtained by stacking the columns of $u_{ij,t}$, and let the entries of $\mathbf{u}_t$ satisfy*

$$u_{i+1,j,t} = \alpha_{i,j,t} u_{ij,t} + v_{i+1,j,t}, \quad u_{1,j+1,t} = \alpha u_{p_1,j,t} + v_{1,j+1,t},$$

where $v_{ij,t}$ is i.i.d., and $\alpha \in (0,1)$. We aim to calculate the covariance between $\bar{\zeta}_{ij}\bar{\zeta}_{ih_1}$ and $\bar{\zeta}_{lj}\bar{\zeta}_{lh_2}$. We only consider the cases $i \leq l$ and $j \leq h_1 \leq h_2$; the other cases are similar. Then,

$$u_{lh_2,t} = \alpha^{p_2(h_2-h_1)+l-i} u_{ih_1,t} + R_{lh_2,t},$$

and $R_{lh_2,t}$ is independent of $u_{ij,t}, u_{lj,t}, u_{ih_1,t}$. As a result,

$$\begin{aligned}
&Cov(\bar{\zeta}_{ij}\bar{\zeta}_{ih_1}, \bar{\zeta}_{lj}\bar{\zeta}_{lh_2}) \\
=&Cov\bigg(\frac{1}{T}\sum_{t_1}\sum_{t2} F_{t_1}^2 F_{t_2}^2 u_{ijt_1} u_{ih_1t_2}, \frac{1}{T}\sum_{s_1}\sum_{s_2} F_{s_1}^2 F_{s_2}^2 u_{ljs_1}[\alpha^{p_2(h_2-h_1)+l-i} u_{ih_1s_2} + R_{lh_2s_2}]\bigg) \\
=&\alpha^{p_2(h_2-h_1)+l-i} Cov\bigg(\frac{1}{T}\sum_{t_1}\sum_{t2} F_{t_1}^2 F_{t_2}^2 u_{ijt_1} u_{ih_1t_2}, \frac{1}{T}\sum_{s_1}\sum_{s_2} F_{s_1}^2 F_{s_2}^2 u_{ljs_1} u_{ih_1s_2}\bigg) \\
=&\frac{\alpha^{p_2(h_2-h_1)+l-i}}{T^2} \sum_{t_1,t_2} \sum_{(s_1,s_2)=(t_1,t_2)} \bigg|Cov\bigg(\frac{1}{T}\sum_{t_1}\sum_{t2} F_{t_1}^2 F_{t_2}^2 u_{ijt_1} u_{ih_1t_2}, \frac{1}{T}\sum_{s_1}\sum_{s_2} F_{s_1}^2 F_{s_2}^2 u_{ljs_1} u_{ih_1s_2}\bigg)\bigg| \\
\leq&C\alpha^{p_2(h_2-h_1)+l-i}.
\end{aligned}$$

Therefore, given any i, j, h_1, we have

$$\sum_{l,h_2} Cov(\bar{\zeta}_{ij}\bar{\zeta}_{ih_1}, \bar{\zeta}_{lj}\bar{\zeta}_{lh_2}) \leq C \sum_{l\geq i} \alpha^{l-i} \sum_{h_2 \geq h_1} \alpha^{p_2(h_2-h_1)} \leq C_0.$$

We now turn to discussing an alternative formulation to Assumption B4*(ii)*. The assumption is needed in order to estimate the magnitude of the following term, which appears in the proof of Lemma B3 in Yu et al. (2022) (which, in turn, we invoke in the proof of Theorem 1):

$$\frac{1}{(Tp_2)^2}\sum_{j=1}^{p_2} E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} C_{j_1} F_t F_s e_{ij,s} e_{ij_1,t}\right)^2;$$

specifically, it is required that

$$\frac{1}{(Tp_2)^2}\sum_{j=1}^{p_2} E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} C_{j_1} F_t F_s e_{ij,s} e_{ij_1,t}\right)^2 = O\left(p_1\right) + O\left(\frac{p_1^2}{p_2}\right). \tag{A.12}$$

Lemma A.3. *We assume that (A.11) holds with*

$$E\left|F_t g\left(F_t\right)\right|^4 < \infty, \tag{A.13}$$

$$\sum_{t=1}^{T}\sum_{j_1=1}^{p_2}\left|E\left(u_{ij,s} u_{ij_1,t}\right)\right| \leq c_0, \tag{A.14}$$

$$\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2}\left|E\left(\left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right)\right| \leq c_0, \tag{A.15}$$

for all $1 \leq u \neq t \leq T$ and $1 \leq v \neq s \leq T$. Then (A.12) holds.

Before showing the proof of Lemma A.3, some comments are in order. The lemma states that Assumption B4*(ii)* can be replaced by assuming a more specific form of dependence between F_t and $e_{ij,t}$, at the price of the possibly stronger moment condition (A.13) on the common factors F_t - recall that the theory in the main paper requires that $E\left|F_t\right|^{4+\epsilon}$ is finite, with arbitrarily small ϵ (see e.g. our Assumption B1*(i)*(b)). Equation (A.14) states that the $u_{ij,t}$s are weakly dependent across time and space, and it is very similar to our Assumption B3*(ii)*(b); clearly, the assumption immediately holds if the $u_{ij,t}$s are serial and cross-sectional independent. Finally, (A.15) is again a weak dependence condition on the $u_{ij,t}$s, which mimics and extends, essentially, our Assumption B3*(iii)*. For example, if $u_{ij,t}$ is assumed to be independent across t, then (A.15) follows.

B. Implementation and further numerical evidence

In this section, we report the specifications of the other criteria to estimate k_1 in Section 4; we discuss in greater detail some aspects of the implementation of our procedures; and we report further Monte Carlo evidence.

B.1. Notes on the implementation of the alternative criteria used in Section 4

In Section 4, we have used the following criteria, whose exact specification we describe here.

1. As far as the α-PCA method proposed by Chen and Fan (2021) is concerned, in the simulations in Yu et al. (2022), the α-PCA method with $\alpha = \pm 1$ is found to have comparable finite sample performances with the case $\alpha = 0$; hence, we only report the results for α-PCA with $\alpha = 0$.
2. As far as the iterative versions of the Eigenvalue Ratio and of the Information Criteria algorithms by Han et al. (2022) (denoted as iTIP-ER and iTIP-IC respectively) are concerned, we refer to the original paper for details. We have used the tensorTS package in R (see Chen et al., 2022), and we have run both the iTOP(UP) and the iTIP(UP) iterative procedures, but results are virtually unchanged between the two - hence, in the simulations we have only reported results for the iTIP(UP) case (which are slightly better) to save space. Similarly, different penalty functions for both the ER and the IC versions of the iterative algorithms gave essentially the same results, and therefore we have reported results using only the default penalty functions computed with $\nu = 0$ and $c_0 = 0.1$ and $h_0 = 1$, namely
$$g_1 = \frac{1}{T} \ln \left(\frac{p_1 p_2 T}{p_1 p_2 + T} \right),$$
for IC, and
$$h_1 = 0.1,$$
for ER. We would like to note that, although this is not explicitly mentioned in the package documentation, it seems evident that both approaches, IC an ER, are initialised at $k = 1$, thus making these procedures effectively unable to detect that $k_1 = 0$ when this is the case.
3. As far as the method developed in Lam (2021) (denoted as TCorTh) is concerned, we refer to the original contribution for details. We would like to point out that, in our simulations, we have used the modified version of TCorTh, with threshold $1 + \eta_{T,\lambda}$, with $\lambda = 1$.
4. Finally, as far as the $PC_{p1}(k)$ criterion proposed by Bai and Ng (2002) is concerned, this is based on estimating k as
$$\widehat{k}_{IC} = \arg \min_{0 \leq k \leq k_{\max}} V(k) + g(k, N, T),$$

where $V(k)$ is defined in Section 4 in Bai and Ng (2002). We have used the penalty function

$$g(k,N,T)=k\widehat{\sigma}^2\left(\frac{p_1p_2+T}{p_1p_2T}\right)\ln\left(\frac{p_1p_2T}{p_1p_2+T}\right),$$

where $\widehat{\sigma}^2$ is a consistent estimate of

$$\frac{1}{p_1p_2T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{t=1}^{T}E\left(e_{ij,t}^2\right).$$

B.2. Implementation: the choice of the weight function $F(u)$

We begin with a discussion of the weight function $F(u)$ employed in Step 4 of our randomisation algorithm, and whose properties are summarised in Assumption C1. In essence, the assumption requires that $F(u)$ is a distribution (seeing as it integrates to 1) with finite second moment. Hence, in order to implement Step 4 of the algorithm, any choice of $F(u)$ that satisfies Assumption C1 can, in principle, be used. On account of its simplicity, we use, as a default option, the standard normal distribution; the practical computation of $\widehat{\Psi}_{k_1^0}$ (and $\widetilde{\Psi}_{k_1^0}$) with this choice is discussed in the paper. Here, we consider two related questions: firstly, how to construct $\widehat{\Psi}_{k_1^0}$ (and $\widetilde{\Psi}_{k_1^0}$) when using a different distribution; and, secondly, we evaluate the robustness of our procedures with respect to the choice of $F(u)$.

Let $u\in U=(-\infty,\infty)$ to begin with. Consistently with the main paper, we propose to approximate the integral

$$\widehat{\Psi}_{k_1^0}=\int_{-\infty}^{\infty}\left[\widehat{\nu}_{k_1^0}(u)\right]^2dF(u),$$

by Gauss-Hermite quadrature, for any choice of $F(u)$. Let the first derivative of $F(u)$ be denoted as $f(u)$. Then, by elementary arguments

$$\begin{aligned}\int_{-\infty}^{\infty}\left[\widetilde{\nu}_{k_1^0}(u)\right]^2dF(u) &= \int_{-\infty}^{\infty}\left[\widetilde{\nu}_{k_1^0}(u)\right]^2 f(u)\,du\\ &= \int_{-\infty}^{\infty}\left[\widetilde{\nu}_{k_1^0}(u)\right]^2 f(u)\,e^{u^2}e^{-u^2}du\\ &= \int_{-\infty}^{\infty}h(u)\,e^{-u^2}du, \qquad \text{(B.1)}\end{aligned}$$

having defined for short

$$h(u)=\left[\widetilde{\nu}_{k_1^0}(u)\right]^2 f(u)\,e^{u^2}.$$

The integral in (B.1) can be evaluated numerically via the approximation

$$\int_U h(u)\, e^{-u^2} du \approx \sum_{s=1}^{n_S} w_s h(z_s) = \sum_{s=1}^{n_S} \left[w_s \exp\left(z_s^2\right) f(z_s) \right] \left[\widetilde{\nu}_{k_1^0}(z_s) \right]^2 , \tag{B.2}$$

where z_s, $1 \leq s \leq n_S$, are the roots of the Hermite polynomial $H_{n_S}(z)$ defined in (4.2), and the associated weights w_s are given by

$$w_s = \frac{2^{n_S-1} n_S! \sqrt{\pi}}{n_S^2 \left[H_{s-1}(z_s) \right]}.$$

The values of z_s and w_s are given e.g. in Salzer et al. (1952), and in Table B.1 we report their values for the cases (considered in our paper) of $n_S = 2$ and $n_S = 4$.

Table B.1: Zeros of the Hermite polynomials and weights

n_S	z_s	w_s
2	$z_1 = 0.707$	$w_1 = 0.886$
	$z_2 = -0.707$	$w_2 = 0.886$
4	$z_1 = 0.525$	$w_1 = 0.80$
	$z_2 = 1.651$	$w_2 = 0.08$
	$z_3 = -0.525$	$w_3 = 0.80$
	$z_4 = -1.651$	$w_4 = 0.08$

As an illustrative example, $F(u)$ could be the Student t distribution with β degrees of freedom:

$$f(u) = \frac{1}{\sqrt{\beta} B\left(\frac{1}{2}, \frac{\beta}{2}\right)} \left(1 + \frac{u^2}{\beta}\right)^{-\frac{\beta+1}{2}},$$

where $B(\cdot, \cdot)$ is the Beta function (Olver et al., 2010). Hence we can compute $\int_{-\infty}^{\infty} \left[\widetilde{\nu}_{k_1^0}(u)\right]^2 dF(u)$ as

$$\int_{-\infty}^{\infty} \left[\widetilde{\nu}_{k_1^0}(u)\right]^2 dF(u)$$
$$\approx \sum_{s=1}^{n_S} \left[\frac{1}{\sqrt{\beta} B\left(\frac{1}{2}, \frac{\beta}{2}\right)} w_s \left(1 + \frac{z_s^2}{\beta}\right)^{-\frac{\beta+1}{2}} \exp\left(z_s^2\right) \right] \left[\widetilde{\nu}_{k_1^0}(z_s)\right]^2 = \sum_{s=1}^{n_S} \widetilde{w}_s \left[\widetilde{\nu}_{k_1^0}(z_s)\right]^2$$

As an example, we report z_s and the corresponding weights $\widetilde{w}_s$ for $n_S = 2$ and $n_S = 4$ in Table B.2.

Finally, note that when the domain of $F(u)$ is $[a, b]$, with $-\infty < a < b < \infty$, it is possible to define

$$\left[\widetilde{\nu}_{k_1^0}(u)\right]^2 f(u) = g(u),$$

Table B.2: Values of z_s and $\widetilde{w}_s$ for the Student t distribution

Distribution		t_3		t_5	
	n_S	z_s	$\widetilde{w}_s$	z_s	$\widetilde{w}_s$
	2	0.525	0.394	0.525	0.416
		-0.525	0.394	-0.525	0.416
	4	0.525	0.325	0.525	0.340
		-0.525	0.325	-0.525	0.340
		1.651	0.122	1.651	0.124
		-1.651	0.122	-1.651	0.124

whence

$$\int_a^b \left[\widetilde{\nu}_{k_1^0}(u)\right]^2 f(u)\, du = \int_a^b g(u)\, du.$$

Using the $\tanh(\cdot)$ transformation, it then follows that

$$\int_a^b g(u)\, du = \frac{b-a}{2}\int_{-\infty}^{\infty} g\left(\frac{a+b}{2}+\frac{b-a}{2}\tanh u\right)\left(\operatorname{sech}^2 u\right) du,$$

and therefore, ultimately, we can compute

$$\begin{aligned} &\int_a^b g(u)\, du \\ \approx\ &\sum_{s=1}^{n_S}\left[w_s f\left(\frac{a+b}{2}+\frac{b-a}{2}\tanh z_s\right)\left(\operatorname{sech}^2 z_s\right)\exp\left(z_s^2\right)\right]\left[\widetilde{\nu}_{k_1^0}\left(\frac{a+b}{2}+\frac{b-a}{2}\tanh z_s\right)\right]^2. \end{aligned} \tag{B.3}$$

In the paper, we have used the standard normal distribution as the default choice for $F(u)$. In Table B.3, we report a small Monte Carlo exercise to assess the impact of such choice on our methodology. We compare results obtained with $F(u)$ being: the standard normal distribution, a Student's t distribution with 3 and 5 degrees of freedom, and we also use a Rademacher distribution with $u = \pm\sqrt{2}$ and equal weights (equal to 1/2) - the last choice, in particular, is proposed in Trapani (2018), and we refer to it as a "naive" set of weights. The other specifications of the experiments are the same as in the main paper, and we refer to the table notes for details; we have only reported results for STP_1 and STP_2, since results with STP_3 are essentially the same as those obtained with STP_2. In short, results show that our procedure is virtually unaffected by the choice of $F(u)$. Interestingly, for small samples, the "naive" procedure seems to fare worse than the other choices, indicating that choosing a standard normal specification for $F(u)$ leads to some numerical improvements, as well as having a sounder theoretical justification (since it washed out the dependence on the choice of u).

B.3. Further Monte Carlo experiments

We complement the results in Section 4 of the main paper by reporting further experiments.

Table B.3: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, y and z are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
(1,1)	STP_1	0.648(0.352|0.648)	0.838(0.162|0.838)	0.97(0.03|0.97)	0.85(0.15|0.85)	0.946(0.054|0.946)	0.998(0.002|0.998)
	STP_2	0.968(0.032|0.968)	0.998(0.002|0.998)	1(0|1)	0.998(0.002|0.998)	1(0|1)	1(0|1)
naive	STP_1	0.332(0.668|0.332)	0.578(0.422|0.578)	0.894(0.106|0.894)	0.592(0.408|0.592)	0.838(0.162|0.838)	0.984(0.016|0.984)
	STP_2	0.882(0.118|0.882)	0.996(0.004|0.996)	1(0|1)	0.984(0.016|0.984)	1(0|1)	1(0|1)
t_3	STP_1	0.726(0.274|0.726)	0.876(0.124|0.876)	0.988(0.012|0.988)	0.896(0.104|0.896)	0.964(0.036|0.964)	0.998(0.002|0.998)
	STP_2	0.982(0.018|0.982)	1(0|1)	1(0|1)	0.996(0.004|0.996)	1(0|1)	1(0|1)
t_5	STP_1	0.71(0.29|0.71)	0.874(0.126|0.874)	0.986(0.014|0.986)	0.888(0.112|0.888)	0.962(0.038|0.962)	0.998(0.002|0.998)
	STP_2	0.98(0.02|0.98)	1(0|1)	1(0|1)	0.996(0.004|0.996)	1(0|1)	1(0|1)
(1,3)	STP_1	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
naive	STP_1	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
t_3	STP_1	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
t_5	STP_1	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
(3,1)	STP_1	2.106(0.338|0.662)	2.732(0.12|0.88)	2.994(0.002|0.998)	2.71(0.106|0.894)	2.946(0.018|0.982)	3(0|1)
	STP_2	2.944(0.024|0.976)	3(0|1)	3(0|1)	2.982(0.006|0.994)	3(0|1)	3(0|1)
naive	STP_1	0.948(0.834|0.166)	1.872(0.542|0.458)	2.782(0.134|0.866)	1.764(0.48|0.52)	2.638(0.178|0.822)	2.98(0.008|0.992)
	STP_2	2.602(0.234|0.766)	2.984(0.012|0.988)	3(0|1)	2.894(0.046|0.954)	2.994(0.002|0.998)	3(0|1)
t_3	STP_1	2.388(0.24|0.76)	2.872(0.056|0.944)	3(0|1)	2.806(0.07|0.93)	2.942(0.022|0.978)	3(0|1)
	STP_2	2.97(0.01|0.99)	3(0|1)	3(0|1)	2.994(0.002|0.998)	3(0|1)	3(0|1)
t_5	STP_1	2.348(0.26|0.74)	2.858(0.062|0.938)	3(0|1)	2.77(0.082|0.918)	2.942(0.022|0.978)	3(0|1)
	STP_2	2.97(0.01|0.99)	3(0|1)	3(0|1)	2.988(0.004|0.996)	3(0|1)	3(0|1)
(3,3)	STP_1	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
	STP_2	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
naive	STP_1	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
	STP_2	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
t_3	STP_1	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
	STP_2	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
t_5	STP_1	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
	STP_2	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)

B.3.1. Further simulations with different levels of signal-to-noise

Recall that, in our DGP, the signal-to-noise ratio is given by

$$\sigma = \frac{\sum_{i,j=1}^{p_1 p_2} \left\{ \Sigma_C \right\}_{ij}}{\sum_{i,j=1}^{p_1 p_2} \left\{ \Sigma_E \right\}_{ij}} \simeq \theta \frac{k_1 k_2}{9\left(1+a\right)^2}, \tag{B.4}$$

for large values of p_1 and p_2. Hence, the signal-to-noise is affected by k_2. In order to shed more light on the performance of our methodology, we report a further set of experiments where we make σ constant across experiments; in particular, we have studied the case $k_1 = k_2 = 1$ with $\theta = 3$, and the case $k_1 = 1$ and $k_2 = 3$ with $\theta = 1$; thus, the signal-to-noise ratio is not affected by k_2. Results in Table B.4 show that this has virtually no impact on our methodology.

Table B.4: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2)	Method	$T = 50$			$T = 100$		
		$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$	$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$
(1,1)	STP$_1$	1.396(0|0.604)	1(0|1)	1(0|1)	1.326(0|0.674)	1.372(0|0.628)	1(0|1)
$\theta = 3$	STP$_2$	1.014(0|0.986)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP$_3$	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	IterER	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	α-PCA	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.232(0|0.770)	1.248(0|0.752)	1.416(0|0.772)	1.214(0|0.786)	1.180(0|0.820)	1.184(0|0.816)
	TCorTh	1.996(0|0.004)	1.458(0|0.542)	1(0|1)	2(0|0)	2(0|0)	1.996(0|0.004)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
(1,3)	STP$_1$	1.344(0|0.656)	1(0|1)	1(0|1)	1.294(0|0.706)	1.328(0|0.672)	1(0|1)
$\theta = 1$	STP$_2$	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP$_3$	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	IterER	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	α-PCA	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.192(0|0.808)	1.234(0|0.766)	5.600(0|0.724)	1.084(0|0.916)	1.076(0|0.924)	1.104(0|0.896)
	TCorTh	2(0|0)	1.436(0|0.564)	1(0|1)	2(0|0)	2(0|0)	2(0|0)
	IC	2.780(0.208|0.792)	2.908(0.09|0.91)	2.904(0.094|0.906)	3(0|1)	3(0|1)	3(0|1)

B.3.2. Comparable and small cross sectional dimensions

We consider the case of $p_1 = p_2$ and small. We evaluate the ability of detecting whether there exists a factor structure in Table B.5, and of estimating the number of common factors in Table B.6. We use the same DGP as in Section 4 in the main paper, altering only the values of (p_1, p_2, T). The values of M and S have been tuned to deliver the best performances in this case, based on the sample sizes (p_1, p_2, T).

Table B.5 shows that our approaches, both the one based on the non-projected and the one based on the projected second moment matrix, are very good at determining whether a factor structure does exist or not; we note that even vectorising the series X_t (and applying the methodology by Trapani, 2018) delivers good results, although this approach is always dominated. As far as estimation of k_1 is concerned, Table B.6 shows that, when p_1 and p_2 are small, STP$_3$ delivers the best performance among our proposed statistics, although when $p_1 = p_2 = 10$, even STP$_3$ is often wrong and understates the true number of

Table B.5: Proportions of correctly determining whether there exists factor structure using $\widehat{\Psi}_1^S$ and $\widetilde{\Psi}_1^S$ over 500 replications with $M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$.

α	(k_1, k_2)	Method	$T = 100$		$T = 150$	
			$p_1 = p_2 = 10$	$p_1 = p_2 = 15$	$p_1 = p_2 = 10$	$p_1 = p_2 = 15$
0.01	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1
		Vec	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	1.000	1.000	1.00	1.000
		$\widetilde{\Psi}_1^S$	1.000	1.000	1.00	1.000
		Vec	0.984	0.992	0.99	0.988
	(1,3)	$\widehat{\Psi}_1^S$	1.000	1.0	1.000	1.000
		$\widetilde{\Psi}_1^S$	1.000	1.0	1.000	1.000
		Vec	0.954	0.9	0.988	0.974

common factors. We note that, even with larger sample sizes, when $k_1 = 1$, the STP_3 methodology is dominated by the iTIP-IC approach of Han et al. (2022), which always estimates k_1 correctly (STP_3 does improve when $k_2 = 3$, but iTIP-IC is still better); however, STP_3 is superior when $k_1 = 3$ - in this case, the iTIP-IC approach always understates k_1. Overall, it appears that when p_1 and p_2 are small, the best approach to employ is STP_3, whose results seem comparable with the estimator by Lam (2021). We note, finally, that, at the other end of the spectrum, STP_1 works rather poorly. Combining this finding with the good performance of the IterER approach, we conclude that using the projection-based estimator results in major improvements, in particular in the presence of small samples.

B.3.3. Detection of weak factors

In Section 3.4.2 in the main paper, we showed that our procedure may be able to detect weak factors, as long as these are not "too weak". Here, we report some evidence on the (comparative) ability of our procedures to detect weak common factors. In particular, we focus on the case where $k_1 = 1$ (and this common factor is weak), and we consider $k_2 = 1$ and $k_2 = 3$, to study the impact of k_2 on the ability to detect common factors.

We use exactly the same DGP as in the other experiments, but we generate data using $p_1^{-\kappa} R$ instead of R. Hence, according to Corollary 2, it holds that

$$\lambda_1 = \Omega\left(p_1^{1-2\kappa}\right);$$

similarly, the same result holds for $\widetilde{\lambda}_1$, according to Lemma C.5.

In a similar spirit to the other experiments, we have considered two scenarios: one with "small" $p_2 = \{20, 30, 50\}$ and $p_1 = T = \{100, 150\}$; and one with large $T = \{100, 150\}$ and $p_1 = p_2 = \{30, 40, 50\}$.

Table B.6: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, \phi = \psi = 0.1, f(S) = S^{-1/4}$, y and z are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2)	Method	$T = 100$			$T = 150$		
		$(M, S) = (150, 200)$	$(M, S) = (200, 400)$	$(M, S) = (250, 500)$	$(M, S) = (150, 200)$	$(M, S) = (250, 500)$	$(M, S) = (250, 500)$
		$p_1 = p_2 = 10$	$p_1 = p_2 = 15$	$p_1 = p_2 = 20$	$p_1 = p_2 = 10$	$p_1 = p_2 = 15$	$p_1 = p_2 = 20$
(1,1)	STP_1	0.335(0.875|0.095)	0.11(0.93|0.03)	0.175(0.895|0.035)	0.355(0.89|0.075)	0.01(0.99|0.01)	0.21(0.875|0.04)
	STP_2	0.255(0.81|0.125)	0.705(0.625|0.045)	1.615(0.17|0.045)	0.22(0.845|0.09)	0.415(0.735|0.115)	1.535(0.185|0.095)
	STP_3	0.68(0.545|0.285)	1(0.375|0.32)	0.995(0.19|0.655)	0.59(0.61|0.23)	0.73(0.4|0.47)	0.915(0.235|0.615)
	IterER	0.1(0.925|0.05)	0.21(0.795|0.2)	0.535(0.465|0.535)	0.14(0.905|0.05)	0.245(0.755|0.245)	0.485(0.515|0.485)
	α-PCA	1.36(0.065|0.51)	1.185(0.37|0.075)	1.27(0.355|0.02)	1.425(0.025|0.525)	1.455(0.24|0.065)	1.19(0.39|0.03)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.27(0|0.785)	1.14(0|0.885)	1.155(0|0.85)	1.3(0|0.78)	1.11(0|0.89)	1.125(0|0.89)
	TCorTh	1.685(0|0.315)	1.995(0|0.005)	1.995(0|0.005)	1.745(0|0.255)	1.99(0|0.01)	2(0|0)
	IC	1.82(0|0.18)	1.01(0|0.99)	1(0|1)	1.915(0|0.085)	1.105(0|0.895)	1(0|1)
(1,3)	STP_1	3.105(0.345|0.125)	1.715(0.055|0.175)	1.59(0.005|0.4)	2.285(0.29|0.125)	1.21(0.085|0.62)	1.565(0.02|0.395)
	STP_2	1.415(0.23|0.125)	1.585(0.005|0.405)	1.18(0|0.82)	1.46(0.2|0.14)	1.11(0.015|0.86)	1.135(0|0.865)
	STP_3	1.225(0.31|0.295)	1.21(0.05|0.705)	1.07(0.005|0.92)	1.08(0.33|0.305)	0.96(0.095|0.85)	1.05(0.01|0.93)
	IterER	0.265(0.75|0.235)	0.635(0.365|0.635)	0.87(0.13|0.87)	0.42(0.59|0.4)	0.655 (0.345 — 0.655)	0.845(0.155|0.845)
	α-PCA	1.81(0|0.19)	1.705(0|0.295)	1.415(0|0.585)	1.805(0|0.195)	1.695(0|0.305)	1.535(0|0.465)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.215(0|0.825)	1.11(0|0.9)	1.07(0|0.935)	1.15(0|0.855)	1.095(0|0.93)	1.06(0|0.94)
	TCorTh	1.965(0|0.035)	2(0|0)	2(0|0)	1.955(0|0.045)	1.995(0|0.005)	2(0|0)
	IC	3.1(0.205|0.465)	2.73(0.26|0.735)	2.84(0.16|0.84)	3.435(0.06|0.435)	2.97(0.135|0.755)	2.91(0.09|0.91)
(3,1)	STP_1	0.27(0.97|0)	0.665(0.92|0)	1.665(0.785|0)	0.26(0.97|0)	0.045(0.99|0)	1.595(0.795|0)
	STP_2	0.195(0.95|0.04)	1.77(0.5|0.255)	3.375(0.085|0.285)	0.185(0.955|0.04)	0.81(0.765|0.19)	3.38(0.06|0.38)
	STP_3	1.795(0.47|0.38)	3.135(0.035|0.79)	3.01(0.005|0.98)	2.02(0.415|0.465)	2.905(0.07|0.92)	3.005(0|0.995)
	IterER	0.78(0.795|0.155)	2.025(0.355|0.645)	2.9(0.035|0.965)	0.755(0.81|0.17)	1.89(0.395|0.605)	2.905(0.035|0.965)
	α-PCA	1.6(0.745|0.15)	1.815(0.6|0.085)	2.15(0.465|0.04)	1.655(0.75|0.15)	2.1(0.51|0.1)	2.58(0.355|0.065)
	iTIP-IC	1(1|0)	1(1|0)	1(1|0)	1(1|0)	1(1|0)	1(1|0)
	iTIP-ER	1.86(0.745|0.185)	1.895(0.76|0.175)	2.04(0.665|0.265)	1.82(0.795|0.16)	1.945(0.735|0.205)	2.18(0.615|0.32)
	TCorTh	2.18(0.73|0.27)	2.96(0.165|0.71)	3.535(0.01|0.445)	2.28(0.66|0.33)	3.18(0.06|0.7)	3.695(0|0.305)
	IC	3.15(0.17|0.485)	2.755(0.23|0.77)	2.86(0.135|0.865)	3.375(0.09|0.44)	3(0.115|0.76)	2.945(0.055|0.945)
(3,3)	STP_1	4.64(0.42|0)	5.715(0.04|0)	4.05(0.005|0)	4.645(0.415|0)	3.42(0.155|0)	4.01(0|0)
	STP_2	2.385(0.305|0.47)	3.055(0.01|0.925)	3(0|1)	2.29(0.33|0.45)	2.915(0.075|0.925)	3(0|1)
	STP_3	2.645(0.215|0.64)	3.045(0.01|0.935)	3(0|1)	2.72(0.2|0.605)	2.945(0.055|0.945)	3(0|1)
	IterER	2.775(0.31|0.45)	2.965(0.04|0.95)	3(0|1)	2.91(0.265|0.46)	2.935(0.07|0.915)	3(0|1)
	α-PCA	3.06(0.275|0.265)	3.405(0.145|0.255)	3.595(0.03|0.33)	3.11(0.28|0.2)	3.585(0.09|0.19)	3.67(0.03|0.265)
	iTIP-IC	1(1|0)	1(1|0)	1(1|0)	1(1|0)	1(1|0)	1(1|0)
	iTIP-ER	1.83(0.795|0.16)	2.08(0.615|0.35)	2.28(0.565|0.39)	1.97(0.715|0.225)	2.16(0.63|0.345)	2.415(0.435|0.55)
	TCorTh	2.705(0.315|0.665)	3.16(0.015|0.81)	3.35(0|0.65)	2.82(0.225|0.73)	3.365(0.01|0.615)	3.545(0|0.455)
	IC	6.775(0.885|0.08)	7.465(0.78|0.22)	8.06(0.59|0.41)	7.33(0.76|0.19)	8.095(0.6|0.365)	8.55(0.365|0.635)

Results are in Tables B.7 and B.8.

The tables show that our approach has some, limited, ability to detect weak factors, a least when these are not "too weak".[1] The best performance is achieved by STP_2, i.e. using the projected estimator with a deliberately overfitted k_2. The comparatively worse performance of STP_3 (which anyway outperforms STP_1, the IterER methodology, and also the α-PCA method, in this case at least when $k_2 = 1$) can be explained upon noting that this requires a first-step estimate of k_2, whose properties may be hampered by the weakness of the common factors. As predicted by the theory, detecting weak factors becomes harder and harder as p_1 decreases: the results in Table B.8 show the impact of p_1, indicating that, *ceteris paribus*, the ability to detect weak factors rapidly deteriorates as p_1 decreases.[2] Interestingly, with smaller values of p_1, STP_3 delivers the best performance in terms of percentage of times of correct detection of one common factor among our three criteria, at least when $k_1 = 1$. The STP_2 approach, conversely, has a pronounced tendency to overstate the number of common factors, with STP_1 behaving in the opposite way. As also reported in the other simulations, results improve as k_2 increases, which may be explained, again, as having more signal in the columns which enhances the quality of the projected estimator. Other methodologies offer a comparable performance - the criterion by Lam (2021) seems to overstate the number of common factors, whereas the criteria by Han et al. (2022) almost always (correctly) estimate one common factor. This result should be read with some caution, since, as mentioned in Section B.1, in the tensorTS package, both the IC and the ER approaches are initialised at $k_1 = 1$ and therefore, by construction, they will never understate the number of common factors (although they, correctly, do not overstate it either).

B.3.4. Estimation of k_1 using STP_2 when $k_2 = 0$

We investigate the robustness of our estimator of k_1 when $k_2 = 0$ but the applied user employs (mistakenly) $\widetilde{k}_1$ nonetheless.

We begin by providing a theoretical justification of the properties of $\widetilde{k}_1$ under this mis-specification issue. In the following lemma, we assume that the applied user is using $\widehat{k}_2 > 0$ even though $k_2 = 0$, i.e. that k_2 has not been estimated correctly.

Lemma B.1. *We assume that $k_1 > 0$ and $k_2 = 0$, and let $\widehat{c} \in \mathbb{R}^{p_2 \times 1}$ with $\widehat{c}'\widehat{c} = p_2 I_{\widehat{k}_2}$, $\max_{1\leq i\leq p_2} |\widehat{c}_i| \leq c_0$.*

[1] In an unreported set of simulations, we also considered the case - complementary to the results in Table B.7 - of $p_2 = (20, 30, 50)$ and $p_1 = T = (100, 150)$ with lower values of α_1 - namely, $\alpha_1 = 2/3$ and $1/2$. Results are available upon request, but they are in line with the ones reported here: our procedures require large and larger samples to detect the weak factor, with STP_2 dominating both STP_1 and STP_3.

[2] We have carried out further, unreported simulations for the cases $p1 = p2 = (30, 40, 50)$ and $T = (100, 150)$ and lower values of α_1, i.e. $\alpha_1 = 3/4$ and $2/3$. Whilst results confirm the findings reported in Table B.8 in terms of comparative performance of our criteria, they are very poor, although they do improve as p_1 increases and as k_2 increases, in line with the results above.

Table B.7: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2, α_1)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 20$	$p_2 = 30$	$p_2 = 50$	$p_2 = 20$	$p_2 = 30$	$p_2 = 50$
$(1, 1, 3/4)$	STP_1	0(1|0)	0(1|0)	0.06(0.94|0.06)	0(1|0)	0.004(0.996|0.004)	0.088(0.912|0.088)
	STP_2	0.292(0.708|0.292)	0.968(0.032|0.968)	1.228(0|0.772)	0.47(0.53|0.47)	0.992(0.008|0.992)	1(0|1)
	STP_3	0.098(0.902|0.098)	0.626(0.692|0.234)	3.976(0.244|0.278)	0.062(0.938|0.062)	0.288(0.712|0.288)	2.418(0.332|0.41)
	IterER	0.002(0.998|0.002)	0.002(0.998|0.002)	0.212(0.788|0.212)	0(1|0)	0.004(0.996|0.004)	0.132(0.868|0.132)
	α-PCA	0(1|0)	0(1|0)	0(1|0)	0(1|0)	0(1|0)	0(1|0)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.24(0|0.794)	1.29(0|0.744)	1.316(0|0.684)	1.32(0|0.724)	1.332(0|0.668)	1.336(0|0.664)
	TCorTh	1.998(0|0.002)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
$(1, 3, 3/4)$	STP_1	0.888(0.112|0.888)	1(0|1)	1(0|1)	0.952(0.048|0.952)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_3	0.432(0.568|0.432)	1.45(0.166|0.668)	2.114(0.006|0.814)	0.454(0.546|0.454)	0.746(0.254|0.746)	1.61(0.006|0.896)
	IterER	0.014(0.986|0.014)	0.116(0.884|0.116)	0.728(0.272|0.728)	0.018(0.982|0.018)	0.042(0.958|0.042)	0.636(0.364|0.636)
	α-PCA	0.872(0.128|0.872)	0.956(0.044|0.956)	0.968(0.032|0.968)	0.96(0.04|0.96)	0.992(0.008|0.992)	0.998(0.002|0.998)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.168(0|0.832)	1.178(0|0.822)	1.148(0|0.852)	1.098(0|0.902)	1.144(0|0.856)	1.094(0|0.906)
	TCorTh	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1.238(0.996|0.004)	1.292(0.988|0.012)	1.23(0.988|0.012)	1.65(0.944|0.056)	1.748(0.924|0.076)	1.888(0.864|0.136)
$(1, 1, 4/5)$	STP_1	0.004(0.996|0.004)	0.036(0.964|0.036)	0.302(0.698|0.302)	0.004(0.996|0.004)	0.078(0.922|0.078)	0.492(0.508|0.492)
	STP_2	0.674(0.326|0.674)	0.998(0.002|0.998)	1.11(0|0.89)	0.83(0.17|0.83)	0.998(0.002|0.998)	1(0|1)
	STP_3	0.2(0.8|0.2)	0.918(0.49|0.41)	2.388(0.064|0.704)	0.16(0.84|0.16)	0.486(0.514|0.486)	1.686(0.094|0.786)
	IterER	0.008(0.992|0.008)	0.052(0.948|0.052)	0.626(0.374|0.626)	0.002(0.998|0.002)	0.038(0.962|0.038)	0.586(0.414|0.586)
	α-PCA	0.002(0.998|0.002)	0(1|0)	0.002(0.998|0.002)	0.006(0.994|0.006)	0.01(0.99|0.01)	0.006(0.994|0.006)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.246(0|0.76)	1.33(0|0.696)	1.348(0|0.652)	1.348(0|0.682)	1.354(0|0.648)	1.346(0|0.654)
	TCorTh	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
$(1, 3, 4/5)$	STP_1	0.994(0.006|0.994)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_3	0.65(0.352|0.646)	1.176(0.044|0.884)	1.082(0|0.986)	0.662(0.338|0.662)	0.92(0.08|0.92)	1.026(0|0.996)
	IterER	0.076(0.924|0.076)	0.402(0.598|0.402)	0.972(0.028|0.972)	0.098(0.902|0.098)	0.312(0.688|0.312)	0.958(0.042|0.958)
	α-PCA	0.992(0.008|0.992)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.19(0|0.81)	1.18(0|0.82)	1.148(0|0.852)	1.108(0|0.892)	1.142(0|0.858)	1.104(0|0.896)
	TCorTh	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1.694(0.93|0.07)	1.836(0.878|0.122)	1.912(0.83|0.17)	2.16(0.726|0.274)	2.408(0.568|0.432)	2.614(0.372|0.628)

Table B.8: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = 500, S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2, α_1)	Method	$T = 100$			$T = 150$		
		$p_1 = p_2 = 30$	$p_1 = p_2 = 40$	$p_1 = p_2 = 50$	$p_1 = p_2 = 30$	$p_1 = p_2 = 40$	$p_1 = p_2 = 50$
$(1, 1, 4/5)$	STP_1	0(1|0)	0.004(0.996|0.004)	0.078(0.926|0.07)	0(1|0)	0.008(0.992|0.008)	0.046(0.956|0.042)
	STP_2	1.216(0.2|0.384)	1.704(0|0.296)	1.692(0|0.308)	1.14(0.192|0.476)	1.452(0|0.548)	1.418(0|0.582)
	STP_3	0.666(0.418|0.498)	0.984(0.156|0.718)	1.12(0.044|0.87)	0.626(0.43|0.514)	0.966(0.12|0.794)	0.992(0.04|0.93)
	IterER	0.158(0.842|0.158)	0.444(0.556|0.444)	0.742(0.258|0.742)	0.166(0.834|0.166)	0.516(0.484|0.516)	0.794(0.206|0.794)
	α-PCA	0.012(0.994|0)	0.008(0.996|0)	0.004(0.998|0)	0.04(0.98|0)	0.008(0.996|0)	0.008(0.996|0)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.106(0|0.898)	1.152(0|0.852)	1.202(0|0.8)	1.09(0|0.914)	1.15(0|0.858)	1.18(0|0.822)
	TCorTh	1.99(0|0.01)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
$(1, 3, 4/5)$	STP_1	0.894(0.106|0.894)	1.004(0.002|0.992)	1.002(0|0.998)	0.902(0.098|0.902)	0.998(0.002|0.998)	1(0|1)
	STP_2	1.002(0|0.998)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_3	1.018(0.044|0.894)	1.022(0.004|0.974)	1.004(0|0.996)	0.968(0.05|0.932)	1.01(0.01|0.97)	1.004(0|0.996)
	IterER	0.584(0.416|0.584)	0.876(0.124|0.876)	0.972(0.028|0.972)	0.58(0.42|0.58)	0.844(0.156|0.844)	0.978(0.022|0.978)
	α-PCA	1.778(0.012|0.198)	1.472(0.01|0.508)	1.194(0.002|0.802)	1.844(0.006|0.144)	1.594(0.002|0.402)	1.212(0.002|0.784)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.122(0|0.882)	1.104(0|0.896)	1.104(0|0.896)	1.054(0|0.948)	1.064(0|0.936)	1.042(0|0.958)
	TCorTh	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	2.24(0.616|0.384)	2.306(0.592|0.408)	2.244(0.62|0.38)	2.748(0.238|0.762)	2.822(0.172|0.828)	2.854(0.14|0.86)
$(1, 1, 5/6)$	STP_1	0.002(0.998|0.002)	0.022(0.978|0.022)	0.168(0.842|0.148)	0.002(0.998|0.002)	0.028(0.972|0.028)	0.128(0.88|0.112)
	STP_2	1.33(0.106|0.458)	1.6(0|0.4)	1.57(0|0.43)	1.206(0.098|0.598)	1.322(0|0.678)	1.292(0|0.708)
	STP_3	0.784(0.304|0.608)	1.046(0.062|0.842)	1.054(0.008|0.952)	0.722(0.316|0.646)	0.972(0.074|0.88)	1.004(0.012|0.974)
	IterER	0.308(0.692|0.308)	0.646(0.354|0.646)	0.882(0.118|0.882)	0.306(0.694|0.306)	0.718(0.282|0.718)	0.914(0.086|0.914)
	α-PCA	0.048(0.976|0)	0.024(0.988|0)	0.016(0.992|0)	0.14(0.93|0)	0.054(0.972|0.002)	0.032(0.984|0)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.112(0|0.892)	1.18(0|0.826)	1.212(0|0.792)	1.092(0|0.91)	1.146(0|0.858)	1.188(0|0.814)
	TCorTh	1.998(0|0.002)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
$(1, 3, 5/6)$	STP_1	0.948(0.052|0.948)	0.998(0.002|0.998)	1(0|1)	0.966(0.034|0.966)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_3	1.016(0.018|0.948)	1(0.002|0.996)	1(0|1)	1.01(0.012|0.966)	1.004(0.002|0.992)	1.002(0|0.998)
	IterER	0.758(0.242|0.758)	0.952(0.048|0.952)	0.998(0.002|0.998)	0.746(0.254|0.746)	0.958(0.042|0.958)	0.994(0.006|0.994)
	α-PCA	1.65(0.006|0.338)	1.278(0.002|0.718)	1.054(0|0.946)	1.72(0|0.28)	1.388(0|0.612)	1.072(0.002|0.924)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.126(0|0.876)	1.106(0|0.894)	1.106(0|0.894)	1.048(0|0.952)	1.064(0|0.936)	1.044(0|0.956)
	TCorTh	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)	2(0|0)
	IC	2.474(0.456|0.544)	2.572(0.388|0.612)	2.584(0.384|0.616)	2.842(0.154|0.846)	2.92(0.08|0.92)	2.954(0.044|0.956)

We also assume that

$$E\left(F_t\widehat{c}\widehat{c}'F_t'\right)=p_2\Sigma_a, \tag{B.5}$$

with Σ_a a positive definite $k_1\times k_1$ matrix with $\lambda_{\max}\left(\Sigma_a\right)<\infty$;

$$E\left|\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_hF_{h,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_he_{jh,t}\right)\right|^2\leq c_0Tp_2; \tag{B.6}$$

$$E\left|\left(\sum_{h=1}^{p_2}\widehat{c}_he_{ih,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_he_{jh,t}\right)\right|^2\leq c_0p_2^2, \tag{B.7}$$

for all i,j; and

$$\sum_{i=1}^{p_1}\sum_{t=1}^{T}\sum_{h_1=1}^{p_2}\sum_{h_3=1}^{p_2}\left|E\left(F_{h_1,t}F_{h_2,s}e_{ih_3,t}e_{jh_4,s}\right)\right|\leq c_0, \tag{B.8}$$

for all $i\neq j$, $t\neq s$, $h_1\neq h_2$ and $h_3\neq h_4$. Then it holds that

$$\frac{\widetilde{\lambda}_j}{p_1^{-1}\sum_{h=1}^{p_1}\widetilde{\lambda}_h}\geq c_0p_1, \tag{B.9}$$

for all $j\leq k_1$ and

$$\frac{\widetilde{\lambda}_j}{p_1^{-1}\sum_{h=1}^{p_1}\widetilde{\lambda}_h}\leq c_0+o_{a.s.}\left(\frac{p_1}{\sqrt{T}}\left(\ln^2p_1\ln p_2\ln T\right)^{1/2+\epsilon}\right), \tag{B.10}$$

for all $j>k_1$ and all $\epsilon>0$.

Some comments on Lemma B.1 are in order. The lemma studies the case where the data X_t are projected onto a generic matrix $\widehat{c}$ under $k_2=0$. In such a case, there is no theoretical justification for using any projection, and therefore one may expect that using the projection method will not be helpful. Assumptions (B.5)-(B.8) are similar to the other assumptions in the main paper. In essence, (B.6) and (B.7) state that errors and common factors are weakly dependent over time and across space; (B.8) extends, to the case of a large factor space, Assumption B4*(i)*.

The conclusion of the lemma is that, even when $k_2=0$ and the projection method is used mistakenly, there is an eigen-gap between the first largest eigenvalues (normalised by the trace) of $\widetilde{M}_c$ and the remaining ones. Consequently, it is still possible, in principle, to determine the number of spiked eigenvalues in the matrix $\sum_{t=1}^{T}X_t\widehat{c}\widehat{c}'X_t'$ using our methodology. Note, though, that the eigen-gap is smaller than in the case where a genuine two-way structure is present, which shows that - at least in theory - using the projected data $X_t\widehat{c}$ does not result in any improvements when $k_2=0$.

We complement the theory in Lemma B.1 by running a small Monte Carlo experiment. As in the above, we consider the same DGP as in the main paper, with the only difference that $k_2=0$. The results in Table B.9 illustrate the robustness of our approach: even when $k_2 = 0$, and the applied user mistakenly uses the projection approach (e.g. due to an initial, incorrect estimate of k_2), the projected approach seems to still work very well in terms of estimating k_1. Indeed, when used to revise the estimate of k_2, it also always correctly indicates that $k_2 = 0$.

Table B.9: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, y and z are the proportions of under- and exact- determination of the factor number, respectively.

(k_1, k_2)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
(1,0) k_1	STP_1	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
k_2	STP_1	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)
	STP_2	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)
(3,0)k_1	STP_1	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
	STP_2	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)	3(0|1)
k_2	STP_1	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)
	STP_2	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)	0(0|1)

B.3.5. Further robustness analysis

We investigate the robustness of our results to different specifications of the tuning parameters in our procedures.

We begin by considering altering the values of M and S in STP_1 and STP_2. In particular, in Table B.10 we analyse the ability of our procedure to detect whether there is a factor structure or not; in Tables B.11 and B.12, we study how different choices of $f(S)$ in (3.17) alter the results. Results in Tables B.10-B.12 clearly illustrate that our procedures are robust to the specifications of M, S, and of the threshold function $f(S)$. Finally, in Table B.13, we show that results are essentially unaffected by the individual tests nominal levels, α.

B.3.6. Sensitivity to the Kronecker product assumption

Throughout the paper, we have based our results on the maintained assumption that the factor structure has a Kronecker product structure. In this section, we investigate what happens in the presence of misspecification, i.e. what happens under

$$\mathrm{Vec}\left(X_t\right) = \Pi \mathrm{Vec}\left(F_t\right) + \mathrm{Vec}\left(E_t\right), \tag{B.11}$$

Table B.10: Proportions of correctly determining whether there exists factor structure using (3.15) $\widehat{\Psi}_1^S$ and $\widetilde{\Psi}_1^S$ over 500 replications with $\alpha = 0.01, \phi = \psi = 0.1, f(S) = S^{-1/4}$.

(M,S)	(k_1,k_2)	Method	$(p_1,T)=(100,100)$			$(p_1,T)=(150,150)$		
			$p_2=15$	$p_2=20$	$p_2=30$	$p_2=15$	$p_2=20$	$p_2=30$
(100,300)	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.956	0.99	1	0.984	1	1
		$\widetilde{\Psi}_1^S$	0.998	1.00	1	1.000	1	1
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
(500,300)	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.42	0.666	0.922	0.69	0.882	0.984
		$\widetilde{\Psi}_1^S$	0.92	1.000	1.000	0.98	0.998	1.000
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
(300,100)	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.636	0.866	0.97	0.874	0.972	1
		$\widetilde{\Psi}_1^S$	0.976	1.000	1.00	0.988	1.000	1
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
(300,500)	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.634	0.822	0.966	0.822	0.946	0.992
		$\widetilde{\Psi}_1^S$	0.966	1.000	1.000	0.994	1.000	1.000
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1

where

$$\Pi = (C \otimes R) + \tau H. \tag{B.12}$$

In (B.12), the further term τH represents a deviation from the Kronecker product structure assumed in equation (1.3) in the main paper.

In particular, we have used exactly the same data generating process as in the remainder of the simulations, with the elements of H generated independently as $\mathcal{U}(-1,1)$; hence, we have used τ in order to control the amount of mis-specification, using in particular $\tau = 1/T$. Results in Table B.14 show that, under such "local" mis-specification, our methodology is still robust.

B.3.7. Comparison of computational times

We report the average (across 500 replications) computing times of our procedures, comparing them against those of other procedures developed in the literature. This is important from an applied perspective, since our procedures require several computations (each test requires to be randomised, and

Table B.11: Proportions of correctly determining whether there exists factor structure using (3.15) $\widehat{\Psi}_1^S$ and $\widetilde{\Psi}_1^S$ over 500 replications with $\alpha = 0.01, \phi = \psi = 0.1, M = S = 300$.

$f(S)$	(k_1, k_2)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
			$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
$S^{-1/3}$	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.548	0.778	0.97	0.804	0.936	1
		$\widetilde{\Psi}_1^S$	0.964	1.000	1.00	0.992	1.000	1
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
$S^{-1/5}$	(0,0)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1
	(1,1)	$\widehat{\Psi}_1^S$	0.642	0.836	0.988	0.848	0.95	1
		$\widetilde{\Psi}_1^S$	0.976	1.000	1.000	0.996	1.00	1
	(1,3)	$\widehat{\Psi}_1^S$	1	1	1	1	1	1
		$\widetilde{\Psi}_1^S$	1	1	1	1	1	1

subsequently, due to the "strong rule", performed a few times).[3]

The computational timings reported in Tables B.15-B.17 for different combinations of (p_1, p_2, T) show that our procedures are not computer intensive *per se*, and that their computational times are - comparatively - average, with some procedures, such as the α-PCA estimator by Chen and Fan (2021), being much faster and others, like the iterative procedures proposed by Han et al. (2022), being slower. This is especially true when p_1 and p_2 are large.

[3]We are grateful to an anonymous Referee for suggesting this to us.

Table B.12: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications with $(k_1, k_2) = (3, 3)$, $\phi = \psi = 0.1, M = S = 300, \alpha = 0.01$, y and z are the numbers of underestimation and overestimation, respectively.

$f(S)$	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
$S^{-1/3}$	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
$S^{-1/5}$	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)

Table B.13: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications with $(k_1, k_2) = (3, 3)$, $\phi = \psi = 0.1$, $f(S) = S^{-1/4}$, y and z are the numbers of underestimation and overestimation, respectively.

(α, M, S)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
(0.05,300,300)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
(0.10,300,300)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
(0.01,100,300)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
(0.01,500,300)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
(0.01,300,100)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
(0.01,300,500)	STP_1	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	STP_2	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	IterER	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)
	α-PCA	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)	3(0|0)

Table B.14: Simulation results for estimating k_1 in the form $x(y|z)$, x is the sample mean of the estimated factor numbers based on 500 replications $\alpha = 0.01, M = S = 300, \phi = \psi = 0.1, f(S) = S^{-1/4}$, are the proportions of under- and exact- determination of the factor number, respectively. $\theta = 1$

(k_1, k_2)	Method	$T = 50$			$T = 100$		
		$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$	$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$
(1,1)	STP_1	1.966(0.014|0.006)	0.996(0.004|0.996)	1(0|1)	1.980(0.008|0.004)	2(0|0)	1.058(0|0.942)
	STP_2	1.902(0|0.098)	1.002(0|0.998)	1(0|1)	1.690(0|0.310)	1.084(0|0.916)	1(0|1)
	STP_3	1.044(0|0.956)	1(0|1)	1(0|1)	1.002(0|0.998)	1(0|1)	1(0|1)
	IterER	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	α-PCA	0.602(0.504|0.390)	0.910(0.090|0.910)	0.974(0.026|0.974)	1.108(0.306|0.280)	0.970(0.030|0.970)	0.998(0.002|0.998)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.290(0|0.716)	1.312(0|0.688)	1.808(0|0.756)	1.230(0|0.770)	1.282(0|0.722)	1.250(0|0.750)
	TCorTh	2(0|0)	2(0|0)	1.518(0|0.482)	2(0|0)	2(0|0)	2(0|0)
	IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
(1,3)	STP_1	1.348(0|0.652)	1(0|1)	1(0|1)	1.330(0|0.670)	1.304(0|0.696)	1(0|1)
	STP_2	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	STP_3	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	IterER	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	α-PCA	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-IC	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)	1(0|1)
	iTIP-ER	1.146(0|0.854)	1.194(0|0.806)	1.304(0|0.790)	1.090(0|0.910)	1.108(0|0.892)	1.076(0|0.924)
	TCorTh	2(0|0)	1.404(0|0.596)	1(0|1)	2(0|0)	2(0|0)	1.998(0|0.002)
	IC	2.796(0.198|0.802)	2.878(0.122|0.878)	2.910(0.088|0.912)	3(0|1)	3(0|1)	3(0|1)

Table B.15: Comparison of Average Computing Time (seconds) for different methods over 500 replications, with standard deviation in the parenthesis.

(k_1, k_2)	Method	$(p_1, T) = (100, 100)$			$(p_1, T) = (150, 150)$		
		$p_2 = 15$	$p_2 = 20$	$p_2 = 30$	$p_2 = 15$	$p_2 = 20$	$p_2 = 30$
(1,1)	STP_1	0.262(0.091)	0.351(0.093)	0.462(0.122)	0.661(0.183)	0.841(0.211)	1.302(0.34)
	STP_2	0.251(0.069)	0.336(0.079)	0.456(0.108)	0.635(0.162)	0.821(0.206)	1.297(0.337)
	IterER	0.188(0.08)	0.239(0.085)	0.359(0.116)	0.511(0.174)	0.634(0.195)	1.101(0.321)
	α-PCA	0.06(0.032)	0.081(0.034)	0.113(0.039)	0.169(0.063)	0.219(0.071)	0.343(0.114)
	iTIP-IC	0.597(0.471)	0.757(0.709)	0.923(0.742)	1.153(0.639)	1.503(1.213)	2.257(1.741)
	iTIP-ER	0.723(0.886)	0.783(0.958)	0.955(1.226)	1.185(1.293)	1.471(1.714)	1.998(1.974)
	TCorTh	0.07(0.037)	0.098(0.053)	0.14(0.059)	0.197(0.065)	0.259(0.085)	0.424(0.128)
	IC	0.236(0.069)	0.332(0.079)	0.52(0.162)	0.848(0.282)	1.508(0.64)	3.268(0.882)
(1,3)	STP_1	0.281(0.098)	0.346(0.104)	0.465(0.115)	0.681(0.165)	0.859(0.183)	1.281(0.318)
	STP_2	0.261(0.054)	0.326(0.082)	0.449(0.104)	0.654(0.14)	0.862(0.184)	1.283(0.324)
	IterER	0.234(0.092)	0.301(0.106)	0.371(0.12)	0.646(0.176)	0.792(0.208)	1.114(0.312)
	α-PCA	0.059(0.034)	0.076(0.032)	0.117(0.047)	0.165(0.049)	0.221(0.056)	0.344(0.103)
	iTIP-IC	0.457(0.512)	0.523(0.592)	0.604(0.529)	0.914(0.72)	1.187(1.018)	1.654(1.223)
	iTIP-ER	0.782(1.096)	0.794(1.127)	0.916(1.275)	1.471(2.045)	1.569(1.991)	2.271(2.645)
	TCorTh	0.073(0.048)	0.091(0.037)	0.135(0.053)	0.197(0.066)	0.264(0.075)	0.42(0.132)
	IC	0.231(0.065)	0.313(0.084)	0.509(0.151)	0.828(0.231)	1.484(0.53)	3.246(0.805)
(3,1)	STP_1	0.494(0.137)	0.621(0.155)	0.879(0.196)	1.238(0.282)	1.615(0.35)	2.433(0.571)
	STP_2	0.471(0.12)	0.61(0.148)	0.862(0.19)	1.215(0.282)	1.619(0.361)	2.454(0.608)
	IterER	0.192(0.064)	0.242(0.09)	0.358(0.108)	0.503(0.123)	0.672(0.169)	1.066(0.274)
	α-PCA	0.06(0.034)	0.077(0.039)	0.113(0.04)	0.159(0.047)	0.217(0.061)	0.336(0.093)
	iTIP-IC	0.421(0.519)	0.406(0.305)	0.532(0.57)	0.761(0.457)	0.968(0.711)	1.432(0.746)
	iTIP-ER	0.54(0.684)	0.617(0.863)	0.737(0.992)	1.062(1.313)	1.231(1.437)	1.647(1.151)
	TCorTh	0.07(0.027)	0.09(0.042)	0.133(0.049)	0.192(0.059)	0.261(0.078)	0.409(0.126)
	IC	0.242(0.062)	0.312(0.089)	0.502(0.147)	0.802(0.227)	1.468(0.532)	3.151(0.813)
(3,3)	STP_1	0.521(0.127)	0.657(0.151)	0.874(0.196)	1.285(0.27)	1.629(0.321)	2.508(0.576)
	STP_2	0.506(0.114)	0.626(0.133)	0.866(0.196)	1.265(0.254)	1.629(0.342)	2.524(0.59)
	IterER	0.196(0.079)	0.245(0.078)	0.352(0.106)	0.516(0.126)	0.682(0.171)	1.087(0.285)
	α-PCA	0.058(0.032)	0.076(0.033)	0.113(0.044)	0.16(0.043)	0.216(0.058)	0.349(0.104)
	iTIP-IC	0.546(0.669)	0.665(0.8)	0.935(1.104)	1.148(1.29)	1.361(1.223)	2.111(1.689)
	iTIP-ER	1.046(1.349)	1.236(1.613)	1.605(2.07)	1.777(2.422)	2.037(2.677)	3.048(3.855)
	TCorTh	0.073(0.047)	0.09(0.035)	0.133(0.047)	0.192(0.06)	0.252(0.074)	0.418(0.126)
	IC	0.23(0.066)	0.303(0.079)	0.492(0.142)	0.795(0.22)	1.404(0.492)	3.211(0.798)

Table B.16: Comparison of Average Computing Time (seconds) for different methods over 500 replications, with standard deviation in the parenthesis.

(k_1, k_2)	Method	$T = 50$			$T = 100$		
		$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$	$p_1 = p_2 = 50$	$p_1 = p_2 = 100$	$p_1 = p_2 = 150$
(1,1)	STP_1	3.108(0.428)	2.6(0.44)	3.802(0.802)	3.1(0.473)	3.868(0.843)	6.305(1.802)
	STP_2	3.084(0.402)	2.588(0.442)	3.887(0.883)	3.069(0.46)	3.948(0.907)	6.448(1.939)
	IterER	0.564(0.119)	1.118(0.279)	2.68(0.639)	0.672(0.133)	1.824(0.561)	5.149(1.581)
	α-PCA	0.456(0.068)	0.595(0.12)	1.029(0.24)	0.481(0.081)	0.788(0.199)	1.809(0.581)
	iTIP-IC	0.423(0.349)	1.434(1.095)	3.225(2.072)	0.716(0.596)	2.538(1.688)	5.769(3.434)
	iTIP-ER	0.652(0.903)	1.559(2.028)	34.225(15.013)	1.114(1.476)	2.666(2.594)	5.524(4.487)
	TCorTh	0.457(0.073)	0.649(0.142)	1.201(0.31)	0.485(0.084)	0.953(0.275)	2.276(0.669)
	IC	3.687(0.6)	3.729(0.659)	4.206(0.898)	3.862(0.644)	6.047(1.067)	8.656(1.847)
(1,3)	STP_1	0.781(0.112)	2.283(1.077)	4.572(1.304)	0.806(0.097)	1.312(0.116)	2.481(0.112)
	STP_2	0.78(0.103)	2.279(1.071)	4.556(1.31)	0.803(0.084)	1.322(0.124)	2.504(0.101)
	IterER	0.188(0.028)	1.008(0.466)	3.223(0.959)	0.241(0.03)	0.77(0.055)	2.029(0.102)
	α-PCA	0.147(0.024)	0.522(0.247)	1.252(0.374)	0.159(0.03)	0.315(0.017)	0.717(0.038)
	iTIP-IC	0.115(0.106)	0.78(0.656)	2.627(2)	0.182(0.152)	0.614(0.253)	1.77(1.103)
	iTIP-ER	0.186(0.285)	1.128(1.643)	9.835(14.853)	0.28(0.391)	0.9(0.96)	2.292(2.396)
	TCorTh	0.148(0.018)	0.576(0.269)	1.421(0.424)	0.162(0.013)	0.355(0.022)	0.862(0.051)
	IC	1.192(0.13)	3.355(1.638)	5.113(1.501)	1.29(0.088)	1.77(0.153)	2.995(0.235)
(3,1)	STP_1	4.209(0.595)	4.759(0.69)	7.413(1.246)	4.23(0.569)	5.734(1.219)	10.909(3.014)
	STP_2	4.125(0.608)	4.715(0.714)	7.458(1.316)	4.171(0.56)	5.816(1.336)	11.37(3.308)
	IterER	0.578(0.103)	1.138(0.25)	2.791(0.58)	0.67(0.132)	1.699(0.498)	4.685(1.409)
	α-PCA	0.468(0.08)	0.59(0.113)	1.066(0.231)	0.478(0.075)	0.73(0.182)	1.665(0.534)
	iTIP-IC	0.291(0.356)	0.912(0.885)	2.234(1.608)	0.444(0.468)	1.677(0.891)	4.16(2.042)
	iTIP-ER	0.374(0.558)	1.269(1.587)	8.803(12.514)	0.716(1.024)	2.292(1.96)	4.943(3.537)
	TCorTh	0.463(0.086)	0.645(0.14)	1.22(0.281)	0.48(0.081)	0.916(0.286)	2.145(0.682)
	IC	3.756(0.665)	3.649(0.657)	4.172(0.797)	3.83(0.634)	5.807(1.255)	8.207(2.101)
(3,3)	STP_1	4.778(0.415)	5.847(0.679)	9.31(1.369)	4.866(0.435)	7.402(1.186)	14.654(3.064)
	STP_2	4.731(0.45)	5.809(0.709)	9.338(1.407)	4.838(0.461)	7.344(1.21)	14.69(3.092)
	IterER	0.665(0.092)	1.402(0.264)	3.439(0.558)	0.778(0.117)	2.294(0.514)	6.269(1.383)
	α-PCA	0.538(0.059)	0.728(0.106)	1.329(0.227)	0.554(0.06)	0.962(0.182)	2.209(0.521)
	iTIP-IC	0.754(1.022)	2.741(3.313)	5.588(5.875)	1.061(1.411)	3.621(3.737)	8.697(8.636)
	iTIP-ER	1.182(1.49)	3.728(4.395)	8.07(10.4)	1.543(2.038)	4.616(5.408)	9.154(10.984)
	TCorTh	0.533(0.076)	0.791(0.145)	1.506(0.3)	0.563(0.082)	1.136(0.267)	2.682(0.676)
	IC	4.329(0.606)	4.538(0.687)	5.36(0.913)	4.494(0.647)	7.058(1.199)	9.738(2.037)

Table B.17: Comparison of Average Computing Time (seconds) for different methods over 500 replications, with standard deviation in the parenthesis.

(k_1, k_2)	Method	$T = 100$		$T = 150$	
		$p_1 = p_2 = 10$	$p_1 = p_2 = 15$	$p_1 = p_2 = 10$	$p_1 = p_2 = 15$
(1,1)	STP_1	0.041(0.019)	0.051(0.031)	0.042(0.009)	0.06(0.036)
	STP_2	0.052(0.061)	0.063(0.066)	0.054(0.056)	0.063(0.039)
	IterER	0.014(0.019)	0.021(0.019)	0.018(0.019)	0.029(0.02)
	α-PCA	0.005(0.009)	0.007(0.01)	0.007(0.011)	0.009(0.009)
	iTIP-IC	0.099(0.12)	0.114(0.118)	0.08(0.072)	0.114(0.097)
	iTIP-ER	0.134(0.17)	0.163(0.194)	0.128(0.174)	0.164(0.196)
	TCorTh	0.007(0.01)	0.009(0.009)	0.008(0.01)	0.014(0.018)
	IC	0.031(0.007)	0.043(0.016)	0.072(0.019)	0.101(0.032)
(1,3)	STP_1	0.054(0.05)	0.088(0.069)	0.055(0.031)	0.1(0.07)
	STP_2	0.056(0.045)	0.077(0.019)	0.06(0.044)	0.09(0.034)
	IterER	0.016(0.02)	0.024(0.025)	0.023(0.025)	0.035(0.031)
	α-PCA	0.005(0.01)	0.007(0.012)	0.006(0.009)	0.011(0.016)
	iTIP-IC	0.08(0.09)	0.103(0.121)	0.078(0.075)	0.102(0.095)
	iTIP-ER	0.141(0.176)	0.177(0.229)	0.129(0.174)	0.16(0.215)
	TCorTh	0.006(0.008)	0.009(0.01)	0.008(0.007)	0.013(0.009)
	IC	0.032(0.02)	0.043(0.019)	0.073(0.026)	0.1(0.029)
(3,1)	STP_1	0.041(0.022)	0.057(0.04)	0.043(0.022)	0.062(0.036)
	STP_2	0.053(0.062)	0.065(0.06)	0.054(0.058)	0.071(0.064)
	IterER	0.017(0.024)	0.025(0.022)	0.022(0.023)	0.034(0.022)
	α-PCA	0.006(0.01)	0.007(0.01)	0.008(0.017)	0.01(0.013)
	iTIP-IC	0.086(0.097)	0.098(0.114)	0.083(0.091)	0.112(0.129)
	iTIP-ER	0.137(0.176)	0.156(0.205)	0.129(0.174)	0.172(0.234)
	TCorTh	0.007(0.01)	0.01(0.01)	0.008(0.008)	0.013(0.01)
	IC	0.031(0.009)	0.043(0.017)	0.072(0.021)	0.101(0.026)
(3,3)	STP_1	0.051(0.048)	0.12(0.067)	0.049(0.023)	0.134(0.061)
	STP_2	0.055(0.056)	0.116(0.054)	0.057(0.05)	0.13(0.044)
	IterER	0.021(0.02)	0.032(0.023)	0.026(0.019)	0.045(0.03)
	α-PCA	0.005(0.01)	0.008(0.011)	0.006(0.009)	0.01(0.016)
	iTIP-IC	0.085(0.103)	0.115(0.152)	0.08(0.088)	0.115(0.129)
	iTIP-ER	0.173(0.209)	0.198(0.257)	0.155(0.206)	0.189(0.264)
	TCorTh	0.007(0.01)	0.009(0.009)	0.008(0.008)	0.013(0.01)
	IC	0.032(0.016)	0.042(0.015)	0.071(0.012)	0.099(0.022)

C. Technical lemmas

Given an $m \times n$ matrix A, recall that $\|A\|$ denotes its spectral norm, a_{ij} its element in position (i, j), and we use $\|A\|_{\max}$ to denote the largest $|a_{ij}|$. Further, we denote the Frobenius norm as $\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2\right)^{1/2}$ and the L_1 and L_∞ induced norms as $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^m |a_{ij}|$ and $\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|$ respectively.

Lemma C.1. *Consider a multi-index partial sum process $U_{S_1,..,S_h} = \sum_{i_1=1}^{S_1} \cdots \sum_{i_h=1}^{S_h} \xi_{i_1,...,i_h}$, and assume that, for some $p \geq 1$*

$$E\left|U_{S_1,..,S_h}\right|^p \leq c_0 \prod_{j=1}^h S_j^{d_j},$$

where $d_j \geq 1$ for all $1 \leq j \leq h$. Then it holds that

$$\lim \sup_{\min\{S_1,...,S_h\}\to\infty} \frac{\left|U_{S_1,..,S_h}\right|}{\prod_{j=1}^h S_j^{d_j/p} \left(\ln S_j\right)^{1+\frac{1}{p}+\epsilon}} = 0 \ a.s.,$$

for all $\epsilon > 0$.

Proof. We begin by showing that the function $f(x_1, ..., x_h) = \prod_{j=1}^h x_j^{d_j}$ is superadditive. Consider the vector $(y_1, ..., y_h)$, such that $y_j \geq x_j$ for all $1 \leq j \leq h$, and, for any two non-zero numbers s and t such that

$$x_i + s \leq y_i + t.$$

We show that

$$\frac{1}{s}\left[f\left(x_1, ...x_i + s, ..., x_h\right) - f\left(x_1, ..., x_h\right)\right] \leq \frac{1}{t}\left[f\left(y_1, ...y_i + t, ..., y_h\right) - f\left(y_1, ..., y_h\right)\right]. \tag{C.1}$$

Indeed,

$$f\left(x_1, ...x_i + s, ..., x_h\right) - f\left(x_1, ..., x_h\right) = \left(\prod_{j\neq i} x_j^{d_j}\right)\left(\left(x_i + s\right)^{d_i} - x_i^{d_i}\right),$$

and, by construction, $\prod_{j\neq i} x_j^{d_j} \leq \prod_{j\neq i} y_j^{d_j}$. Also, note that the function $g(x_i) = x_i^{d_i}$ is convex, and therefore it holds that (see Potra, 1985) there exists a mapping $\delta(x_i + s, x_i)$ such that $\delta(u, v) \leq \delta(p, q)$ whenever $u \leq p$ and $v \leq q$, such that

$$\frac{1}{s}\left[\left(x_i + s\right)^{d_i} - x_i^{d_i}\right] = \delta\left(x_i + s, x_i\right).$$

Thus it follows immediately that (C.1) holds, for every $1 \leq i \leq h$. This entails that $f(x_1, ..., x_h)$ is an S-convex function (see Definition 2.1 and Proposition 2.3 in Potra, 1985), with $f(0, ..., 0) = 0$. Hence, by

Proposition 2.9 in Potra (1985), $f(x_1, ..., x_h)$ is superadditive. We can now apply the maximal inequality for rectangular sums in Corollary 4 of Moricz (1983), which stipulates that

$$E \max_{1\leq i_1\leq S_1,\ldots,1\leq i_h\leq S_h} \left|U_{i_1,..,i_h}\right|^p \leq c_0 \prod_{j=1}^{h} S_j^{d_j} \left(\ln S_j\right)^p ,$$

so that

$$\begin{aligned}
& \sum_{S_1=1}^{\infty}\cdots\sum_{S_h=1}^{\infty} \frac{1}{\prod_{j=1}^{h} S_j} P\left(\max_{1\leq i_1\leq S_1,\ldots,1\leq i_h\leq S_h} \left|U_{i_1,..,i_h}\right| \geq \varepsilon \prod_{j=1}^{h} S_j^{d_j/p} \left(\ln S_j\right)^{1+\frac{1}{p}+\epsilon}\right) \\
\leq & \sum_{S_1=1}^{\infty}\cdots\sum_{S_h=1}^{\infty} \frac{1}{\prod_{j=1}^{h} S_j} \left(E\max_{1\leq i_1\leq S_1,\ldots,1\leq i_h\leq S_h} \left|U_{i_1,..,i_h}\right|^p\right) \varepsilon^{-p} \left(\prod_{j=1}^{h} S_j^{d_j/p} \left(\ln S_j\right)^{1+\frac{1}{p}+\epsilon}\right)^{-p} \leq c_0.
\end{aligned}$$

From here onwards, the desired result follows from Lemma A.1 in Barigozzi and Trapani (2022). $\square$

Lemma C.2. *We assume that Assumptions B1-B4 are satisfied, and that $k_2 > 0$. Then it holds that*

$$\left\|\widehat{C} - CH_1\right\|_F = o_{a.s.}\left(p_2^{1/2}\left(\frac{1}{p_2} + \frac{1}{\sqrt{Tp_1}}\right) l'_{p_1,p_2,T}\right),$$

where

$$H_1 = \left(\frac{1}{T}\sum_{t=1}^{T} F_t' F_t\right) \frac{C'\widehat{C}}{p_2} \widehat{\Lambda}_c^{-1},$$

$$l'_{p_1,p_2,T} = \left(\ln T \ln^2 p_1 \ln^2 p_2\right)^{1/2+\epsilon}, \tag{C.2}$$

for all $\epsilon > 0$.

Proof. We consider separately the cases $k_1 > 0$ and $k_1 = 0$; in both cases, the proof is related to the proof of Theorem 3.3 in Yu et al. (2022), and some passages are omitted when possible.

We begin with the case $k_1 > 0$; in order to avoid overburdening the notation, we let $k_1 = k_2 = 1$. Note that, by construction

$$\widehat{C} = \widehat{M}_c \widehat{C} \widehat{\Lambda}_c^{-1},$$

where $\widehat{\Lambda}_c$ is the diagonal matrix containing the leading k_2 eigenvalues of $\widehat{M}_c = (Tp_1p_2)^{-1} \sum_{t=1}^{T} X_t' X_t$.

Therefore

$$
\begin{aligned}
\widehat{C} &= C\left(\frac{1}{T}\sum_{t=1}^{T}F_t'F_t\right)\frac{C'\widehat{C}}{p_2}\widehat{\Lambda}_c^{-1}+\frac{1}{Tp_1p_2}C\sum_{t=1}^{T}F_t'R'E_t\widehat{C}\widehat{\Lambda}_c^{-1} \\
&\quad +\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'RF_tC'\widehat{C}\widehat{\Lambda}_c^{-1}+\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'E_t\widehat{C}\widehat{\Lambda}_c^{-1} \\
&= CH_1+(II+III+IV)\,\widehat{C}\widehat{\Lambda}_c^{-1}.
\end{aligned}
$$

We begin by using exactly the same logic as in the proof of Theorem 2, it can be shown that there exist a positive constant c_0 such that

$$
\lambda_j\left(\frac{1}{Tp_1p_2}\sum_{t=1}^{T}X_t'X_t\right)=\Omega_{a.s.}\left(c_0\right),
$$

for all $j\leq k_2$; in turn, this entails that $\widehat{\Lambda}_c^{-1}$ exists. Also, recall that $\left\|\widehat{C}\right\|_F=p_2^{1/2}$. Consider now

$$
\begin{aligned}
&\frac{1}{p_2}\left\|\frac{1}{Tp_1p_2}C\sum_{t=1}^{T}F_t'R'E_t\widehat{C}\right\|_2^2 \leq \frac{1}{T^2p_1^2p_2^3}\left\|C\right\|_2^2\left\|\widehat{C}\right\|_2^2\left\|\sum_{t=1}^{T}F_t'R'E_t\right\|_2^2 \\
=\ & \frac{1}{T^2p_1^2p_2^3}\left\|C\right\|_2^2\left\|\widehat{C}\right\|_2^2\left\|\sum_{t=1}^{T}F_t'\sum_{j=1}^{p_1}R_jE_{\cdot j,t}\right\|_2^2,
\end{aligned}
$$

where $E_{\cdot j,t}$ denotes the j-th column of E_t. Hence

$$
\frac{1}{p_2}\left\|\frac{1}{Tp_1p_2}C\sum_{t=1}^{T}F_t'R'E_t\widehat{C}\right\|_2^2\leq c_0\frac{1}{T^2p_1^2p_2}\sum_{i=1}^{p_2}\left(\sum_{t=1}^{T}\sum_{j=1}^{p_1}R_je_{ji,t}F_t\right)^2.
$$

Given that (recall $k_1=k_2=1$)

$$
\begin{aligned}
& E\left(\frac{1}{T^2p_1^2p_2}\sum_{i=1}^{p_2}\left(\sum_{t=1}^{T}\sum_{j=1}^{p_1}R_je_{ji,t}F_t\right)^2\right) \\
=\ & E\left(\frac{1}{T^2p_1^2p_2}\sum_{i=1}^{p_2}\sum_{j,h=1}^{p_1}R_jR_h\left(\sum_{t=1}^{T}e_{ji,t}F_t\right)\left(\sum_{t=1}^{T}e_{hi,t}F_t\right)\right) \\
\leq\ & c_0\frac{1}{T^2p_1^2p_2}\sum_{i=1}^{p_2}\sum_{h=1}^{p_1}\left|\sum_{j=1}^{p_1}E\left(\sum_{t=1}^{T}e_{ji,t}F_t\right)\left(\sum_{t=1}^{T}e_{hi,t}F_t\right)\right| \\
\leq\ & c_0\frac{1}{Tp_1},
\end{aligned}
$$

by Assumption B4*(i)*, Lemma C.1 entails that

$$\frac{1}{p_2}\left\|(II)\,\widehat{C}\widehat{\Lambda}_c^{-1}\right\|_2^2 = o_{a.s.}\left(\frac{1}{Tp_1}\left(\ln T\ln p_1\ln p_2\right)^{1+\epsilon}\right),$$

for all $\epsilon > 0$. Similarly

$$\begin{aligned}\frac{1}{p_2}\left\|\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'RF_tC'\widehat{C}\right\|_2^2 &\leq \frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}E_t'RF_t\right\|_2^2\|C\|_2^2\left\|\widehat{C}\right\|_2^2 \\ &\leq c_0\frac{1}{T^2p_1^2p_2}\left\|\sum_{t=1}^{T}E_t'RF_t\right\|_2^2,\end{aligned}$$

and therefore by the same passages as above, it follows that

$$\frac{1}{p_2}\left\|(III)\,\widehat{C}\widehat{\Lambda}_c^{-1}\right\|_2^2 = o_{a.s.}\left(\frac{1}{Tp_1}\left(\ln T\ln p_1\ln p_2\right)^{1+\epsilon}\right).$$

We note that the rates for II and III are not necessarily sharp (see the proof of Theorem 3.3 in Yu et al., 2022), but they suffice for our purposes. Finally, we have

$$\begin{aligned}&\frac{1}{p_2}\left\|\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'E_t\widehat{C}\right\|_2^2 \\ \leq\ &\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2+\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E_t'E_t-E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2.\end{aligned}$$

We have

$$\begin{aligned}\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2 &\leq \frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E\left(E_t'E_t\right)\right)\right\|_2^2\left\|\widehat{C}\right\|_2^2 \\ &\leq c_0\frac{1}{Tp_1^2p_2^2}\sum_{t=1}^{T}\left\|\left(E\left(E_t'E_t\right)\right)\right\|_2^2;\end{aligned}$$

note that $\left\|\left(E\left(E_t'E_t\right)\right)\right\|_2^2 \leq \left\|\left(E\left(E_t'E_t\right)\right)\right\|_1^2$, and

$$\begin{aligned}&\left\|\left(E\left(E_t'E_t\right)\right)\right\|_1^2 \\ \leq\ &\left(\max_{1\leq h\leq p_2}\sum_{k=1}^{p_2}\left|\sum_{i=1}^{p_1}\left(E\left(e_{ih,t}e_{ik,t}\right)\right)\right|\right)^2 \\ \leq\ &\left(\max_{1\leq h\leq p_2}\sum_{i=1}^{p_1}\sum_{k=1}^{p_2}\left|\left(E\left(e_{ih,t}e_{ik,t}\right)\right)\right|\right)^2 \leq c_0p_1^2,\end{aligned}$$

by Assumption B3*(ii)*(a), so that

$$\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2\leq c_0\frac{1}{p_2^2}.$$

Also

$$\begin{aligned}
&\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E_t'E_t-E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2\\
\leq\ & c_0\frac{1}{T^2p_1^2p_2^2}\left\|\sum_{t=1}^{T}\left(E_t'E_t-E\left(E_t'E_t\right)\right)\right\|_2^2\\
=\ & c_0\frac{1}{T^2p_1^2p_2^2}\sum_{i,j=1}^{p_2}\left(\sum_{h=1}^{p_1}\sum_{t=1}^{T}\left(e_{hi,t}e_{hj,t}-E\left(e_{hi,t}e_{hj,t}\right)\right)\right)^2,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{T^2p_1^2p_2^2}E\left(\sum_{i,j=1}^{p_2}\left(\sum_{h=1}^{p_1}\sum_{t=1}^{T}\left(e_{hi,t}e_{hj,t}-E\left(e_{hi,t}e_{hj,t}\right)\right)\right)^2\right)\\
\leq\ & \frac{1}{T^2p_1^2p_2^2}\sum_{i,j=1}^{p_2}\sum_{h,k=1}^{p_1}\sum_{t,s=1}^{T}\left|Cov\left(e_{hi,t}e_{hj,t},e_{ki,s}e_{kj,s}\right)\right|\\
\leq\ & c_0\frac{1}{T^2p_1^2p_2^2}p_2^2Tp_1,
\end{aligned}$$

by Assumption B3*(ii)*. Thus, by Lemma C.1, this term is bounded by

$$\frac{1}{T^2p_1^2p_2^3}\left\|\sum_{t=1}^{T}\left(E_t'E_t-E\left(E_t'E_t\right)\right)\widehat{C}\right\|_2^2=o_{a.s.}\left(\frac{1}{Tp_1}\left(\ln T\ln p_1\ln^2 p_2\right)^{1+\epsilon}\right),$$

for all $\epsilon>0$; again this bound is not necessarily sharp, but it suffices for our results. Putting all together, the desired rate now follows.

We now turn to considering the case $k_1=0$. Recall that, in such a case

$$X_t=\underset{p_1\times k_2}{F_t}\underset{k_2\times p_2}{C'}+E_t,$$

and note that, for $j\leq k_2$

$$\lambda_j\left(\frac{1}{Tp_1p_2}\sum_{t=1}^{T}CF_t'F_tC'\right)\geq\frac{1}{p_2}\lambda_j\left(C'C\right)\lambda_{\min}\left(\frac{1}{Tp_1}\sum_{t=1}^{T}F_t'F_t\right)\geq c_0\frac{1}{Tp_1p_2}p_2p_1T\geq c_0,$$

for some positive c_0, by Assumptions B1*(iv)*(b) and B2. Thus, using the same logic as above, it can be shown that there exist a positive constant c_0 such that

$$\lambda_j\left(\frac{1}{Tp_1p_2}\sum_{t=1}^{T}X_t'X_t\right)=\Omega_{a.s.}\left(c_0\right),$$

for all $j\leq k_2$; in turn, this entails that $\widehat{\Lambda}_c^{-1}$ exists even in this case. Consider now

$$\begin{aligned}\widehat{C} &= C\left(\frac{1}{p_1T}\sum_{t=1}^{T}F_t'F_t\right)\frac{C'\widehat{C}}{p_2}\widehat{\Lambda}_c^{-1}+\frac{1}{Tp_1p_2}C\sum_{t=1}^{T}F_t'E_t\widehat{C}\widehat{\Lambda}_c^{-1}\\ &\quad+\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'F_t'C'\widehat{C}\widehat{\Lambda}_c^{-1}+\frac{1}{Tp_1p_2}\sum_{t=1}^{T}E_t'E_t\widehat{C}\widehat{\Lambda}_c^{-1}\\ &= CH_1^*+\left(II+III+IV\right)\widehat{C}\widehat{\Lambda}_c^{-1},\end{aligned}$$

having set

$$H_1^*=\left(\frac{1}{p_1T}\sum_{t=1}^{T}F_t'F_t\right)\frac{C'\widehat{C}}{p_2}\widehat{\Lambda}_c^{-1}.$$

We already know from the above that

$$IV\widehat{C}\widehat{\Lambda}_c^{-1}=o_{a.s.}\left(p_2^{1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(\ln T\ln p_1\ln^2 p_2\right)^{(1+\epsilon)/2}\right).$$

Further, note that

$$\begin{aligned}\left\|\frac{1}{Tp_1p_2}C\sum_{t=1}^{T}F_t'E_t\widehat{C}\widehat{\Lambda}_c^{-1}\right\|_F &\leq \frac{1}{Tp_1p_2}\left\|C\right\|_F\left\|\widehat{C}\right\|_F\left\|\sum_{t=1}^{T}F_t'E_t\right\|_F\left\|\widehat{\Lambda}_c^{-1}\right\|_F\\ &\leq c_0\frac{1}{Tp_1}\left\|\sum_{t=1}^{T}F_t'E_t\right\|_F,\end{aligned}$$

and (recall that we are assuming $k_1=1$)

$$\left\|\sum_{t=1}^{T}F_t'E_t\right\|_F=\left(\sum_{i=1}^{p_2}\left(\sum_{t=1}^{T}\sum_{j=1}^{p_1}F_{j,t}e_{ji,t}\right)^2\right)^{1/2}.$$

We have

$$\begin{aligned}E\sum_{i=1}^{p_2}\left(\sum_{t=1}^{T}\sum_{j=1}^{p_1}F_{j,t}e_{ji,t}\right)^2 &= E\sum_{i=1}^{p_2}\sum_{t,s=1}^{T}\sum_{j,k=1}^{p_1}F_{j,t}F_{k,s}e_{ji,t}e_{ki,s}\\ &\leq \sum_{i=1}^{p_2}\sum_{t,s=1}^{T}\sum_{j,k=1}^{p_1}\left|E\left(F_{j,t}F_{k,s}e_{ji,t}e_{ki,s}\right)\right|\\ &\leq c_0p_2Tp_1,\end{aligned}$$

by Assumption B4*(iii)*. The lemma now follows from the same passages as above.

$\square$

Lemma C.3. *We assume that Assumptions B1-B4 are satisfied, and let $\overline{\widehat{\lambda}} = p_1^{-1} \sum_{j=1}^{p_1} \widehat{\lambda}_j$. Then it holds that there exist two constants $0 < c_0 \leq c_1 < \infty$ such that $\overline{\widehat{\lambda}} = \Omega_{a.s.}(c_0)$ and $\overline{\widehat{\lambda}} = O_{a.s.}(c_1)$. The same result holds (provided that $k_2 > 0$) for $\overline{\widetilde{\lambda}} = p_1^{-1} \sum_{j=1}^{p_1} \widetilde{\lambda}_j$.*

Proof. We begin by considering the case where both k_1 and k_2 are nonzero. In this case, as above we focus on $k_1 = k_2 = 1$ for simplicity and with no loss of generality. It holds that

$$
\begin{aligned}
\overline{\widehat{\lambda}} &= \frac{1}{p_1} tr \left(\frac{1}{p_2 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} X_{\cdot i,t} X_{\cdot i,t}' \right) \\
&= \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} X_{ji,t}^2 = \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} \left(R_j F_t C_i + e_{ji,t} \right)^2 \\
&= \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} R_j^2 \sum_{i=1}^{p_2} C_i^2 \sum_{t=1}^{T} F_t^2 + \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} e_{ji,t}^2 \\
&\quad + \frac{2}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} R_j F_t C_i e_{ji,t} \\
&= I + II + III.
\end{aligned}
$$

Assumptions B1*(i)* and B2*(ii)* entail that, whenever $k_1 > 0$ and $k_2 > 0$, $I = c_0 + o_{a.s.}(1)$, with $c_0 > 0$. Consider now II

$$
\begin{aligned}
II &= \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} E\left(e_{ji,t}^2\right) + \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} \left(e_{ji,t}^2 - E\left(e_{ji,t}^2\right) \right) \\
&= II_a + II_b.
\end{aligned}
$$

Note that

$$
E \left(\sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} \left(e_{ji,t}^2 - E\left(e_{ji,t}^2\right) \right) \right)^2 = \sum_{i,j=1}^{p_1} \sum_{h,k=1}^{p_2} \sum_{t,s=1}^{T} Cov\left(e_{ih,t}^2, e_{jk,s}^2\right) \leq c_0 p_1 p_2 T,
$$

having used Assumption B3*(iii)*(a). Hence, by Lemma C.1

$$
II_b = o_{a.s.} \left(\frac{1}{\sqrt{p_1 p_2 T}} \left(\ln p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right),
$$

for all $\epsilon > 0$. Also, using Assumption B3*(iv)*

$$
\begin{aligned}
& \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} E\left(e_{ji,t}^2\right) \\
= & \frac{1}{p_1} tr\left(E\left(\frac{1}{p_2 T} \sum_{t=1}^{T} E_t E_t'\right)\right) \\
\geq & \lambda_{\min}\left(E\left(\frac{1}{p_2 T} \sum_{t=1}^{T} E_t E_t'\right)\right) > 0,
\end{aligned}
$$

which entails that $II > 0$ a.s. Finally, recall that, by Assumption B4*(i)*

$$
E\left(\sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} F_t \frac{R}{p_1^{1/2}} E_t \frac{C'}{p_2^{1/2}}\right)^2 \leq c_0 T,
$$

which entails that

$$
III = o_{a.s.}\left(\frac{1}{\sqrt{p_1 p_2 T}} \left(\ln p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),
$$

for all $\epsilon > 0$. Putting all together, the desired result now follows. The result for $\overline{\overline{\lambda}}$ follows from the same passages, and from the same arguments as in the proof of Theorem 1.

We now consider the "boundary" cases, starting from $k_1 = k_2 = 0$. In such a case, it is immediate to see that

$$
\overline{\overline{\lambda}} = \frac{1}{p_1} tr\left(\frac{1}{p_2 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} X_{\cdot i,t} X_{\cdot i,t}'\right) = \frac{1}{p_1 p_2 T} \sum_{j=1}^{p_1} \sum_{i=1}^{p_2} \sum_{t=1}^{T} e_{ji,t}^2 > 0 \quad \text{a.s.},
$$

as already shown above. When $k_1 = 0$ and $k_2 > 0$, we have

$$
X_t = \underset{p_1 \times k_2}{F_t} \underset{k_2 \times p_2}{C'} + E_t,
$$

and it holds that (with $k_2 = 1$)

$$
\begin{aligned}
& \frac{1}{p_2} tr\left(\frac{1}{p_1 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} X_{\cdot i,t} X_{\cdot i,t}'\right) \\
= & \frac{1}{p_2} tr\left(\frac{1}{p_1 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} C_i^2 F_t F_t'\right) + \frac{1}{p_2} tr\left(\frac{1}{p_1 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} C_i F_t E_{\cdot i,t}'\right) \\
& + \frac{1}{p_2} tr\left(\frac{1}{p_1 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} C_i E_{\cdot i,t} F_t'\right) + \frac{1}{p_2} tr\left(\frac{1}{p_1 T} \sum_{i=1}^{p_2} \sum_{t=1}^{T} E_{\cdot i,t} E_{\cdot i,t}'\right) \\
= & I + II + III + IV.
\end{aligned}
$$

We have

$$I = \frac{1}{p_2}\sum_{i=1}^{p_2} C_i^2 tr\left(\frac{1}{p_1 T}\sum_{t=1}^{T} F_t F_t'\right) = c_0 tr\left(\frac{1}{p_1 T}\sum_{t=1}^{T} F_t F_t'\right),$$

by Assumption B2, and

$$tr\left(\frac{1}{p_1 T}\sum_{t=1}^{T} F_t F_t'\right) = tr\left(\frac{1}{p_1 T}\sum_{t=1}^{T} F_t' F_t\right) = c_1 + o_{a.s.}(1),$$

with $c_1 > 0$, in light of Assumption B1*(iv)*. We already know from above that $IV = c_0 + o_{a.s.}(1)$, with $c_0 > 0$. Finally

$$II = \frac{1}{p_1 p_2 T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{t=1}^{T} C_j F_{i,t} e_{ij,t},$$

and

$$\begin{aligned}
& E\left(\frac{1}{p_1 p_2 T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{t=1}^{T} C_j F_{i,t} e_{ij,t}\right)^2 \\
= & \frac{1}{(p_1 p_2 T)^2}\sum_{i,i'=1}^{p_1}\sum_{j,j'=1}^{p_2}\sum_{t,s=1}^{T} C_j C_{j'} E\left(F_{i,t} F_{i',s} e_{ij,t} e_{i'j',s}\right) \\
\leq & c_0 \frac{1}{(p_1 p_2 T)^2}\sum_{i,i'=1}^{p_1}\sum_{j,j'=1}^{p_2}\sum_{t,s=1}^{T} \left|E\left(F_{i,t} F_{i',s} e_{ij,t} e_{i'j',s}\right)\right| \leq c_1 \frac{1}{p_2 T},
\end{aligned}$$

by Assumption B4*(iii)* - this rate is not necessarily sharp, but it sufficies to prove that II (and, by symmetry, III), are both $o_{a.s.}(1)$, whence the desired result. Finally, consider the case $k_1 > 0$ and $k_2 = 0$. In such a case, recall that

$$X_t = \underset{p_1 \times k_1}{R} \underset{k_1 \times p_2}{F_t} + E_t, \tag{C.3}$$

and

$$X_{\cdot i,t} = R F_{\cdot i,t} + E_{\cdot i,t}. \tag{C.4}$$

Setting again $k_1 = 1$ for simplicity, it follows that

$$\begin{aligned}
& \frac{1}{T p_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} X_{\cdot i,t} X_{\cdot i,t}' \\
= & \frac{1}{T p_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} R F_{\cdot i,t}^2 R' + \frac{1}{T p_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} E_{\cdot i,t} F_{\cdot i,t} R' \\
& \frac{1}{T p_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} R F_{\cdot i,t} E_{\cdot i,t}' + \frac{1}{T p_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} E_{\cdot i,t} E_{\cdot i,t}' \\
= & I + II + III + IV.
\end{aligned}$$

We already know that $IV > 0$ a.s. Also note that the trace of I divided by p_1 is given by

$$\frac{1}{p_1} tr\left(I\right) = \frac{1}{Tp_1p_2}\sum_{j=1}^{p_1}\sum_{i=1}^{p_2}\sum_{t=1}^{T} R_j^2 F_{\cdot i,t}^2,$$

having defined the j-th element of the $p_1 \times 1$ vector R as R_j. By Assumptions B1*(iv)* and B2*(ii)*, it is immediate to see that $p_1^{-1} tr\left(I\right) > 0$ a.s. Finally, the same passages as above also yield that II and III are both $o_{a.s.}\left(1\right)$.

$\square$

Lemma C.4. *We assume that the Assumptions of Lemma C.2 hold, with Assumption B2 replaced with Assumption B6, and*

$$E\left\|\sum_{t=1}^{T}\sum_{j=1}^{p_1} R_j e_{ji,t} F_t\right\|^2 \leq c_0 p_1^{\alpha_1} T, \tag{C.5}$$

$$E\left\|\sum_{t=1}^{T}\sum_{j=1}^{p_2} C_j e_{ij,t} F_t\right\|^2 \leq c_0 p_2^{\alpha_2} T, \tag{C.6}$$

for all $1 \leq i \leq p_2$ *and* $1 \leq i \leq p_1$ *respectively. Then, under*

$$p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left(\frac{1}{p_2} + \frac{1}{\sqrt{Tp_1}}\right) \to 0, \tag{C.7}$$

it holds that

$$\left\|\widehat{C} - CH_1\right\|_F = o_{a.s.}\left(p_2^{1/2} p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left(\frac{1}{p_2} + \frac{1}{\sqrt{Tp_1}}\right) l'_{p_1,p_2,T}\right),$$

with H_1 *and* $l'_{p_1,p_2,T}$ *defined in Lemma C.2.*

Proof. The proof is very similar to that of Lemma C.2, and therefore we omit passages when possible to save space. Recall

$$\begin{aligned}
\widehat{C} &= C\left(\frac{1}{Tp_1}\sum_{t=1}^{T} F_t' R' R F_t\right)\frac{C'\widehat{C}}{p_2}\widehat{\Lambda}_c^{-1} + \frac{1}{Tp_1p_2} C \sum_{t=1}^{T} F_t' R' E_t \widehat{C}\widehat{\Lambda}_c^{-1} \\
&\quad + \frac{1}{Tp_1p_2}\sum_{t=1}^{T} E_t' R F_t C' \widehat{C}\widehat{\Lambda}_c^{-1} + \frac{1}{Tp_1p_2}\sum_{t=1}^{T} E_t' E_t \widehat{C}\widehat{\Lambda}_c^{-1} \\
&= CH_1 + \left(II + III + IV\right)\widehat{C}\widehat{\Lambda}_c^{-1},
\end{aligned}$$

having defined

$$H_1 = \left(\frac{1}{Tp_1}\sum_{t=1}^{T} F_t' R' R F_t\right)\frac{C'\widehat{C}}{p_2}\widehat{\Lambda}_c^{-1},$$

and consider the case of $k_1 = k_2 = 1$ for simplicity. We begin by noting that, using the same logic as above, Assumption B6 readily entails that

$$\lambda_j \left(\frac{1}{Tp_1p_2} \sum_{t=1}^{T} RF_tC'CF_t'R' \right) = \Omega_{a.s.} \left(c_0 p_1^{\alpha_1-1} p_2^{\alpha_2-1} \right);$$

for all $j \leq k_1$; we know from the proof of C.2 that

$$\lambda_{\max} \left(\frac{1}{Tp_1p_2} \sum_{t=1}^{T} E_tE_t' \right) = o_{a.s.} \left(\left(\frac{1}{p_2} + \frac{1}{\sqrt{p_1T}} \right) \left(\ln T \ln p_1 \ln^2 p_2 \right)^{1+\epsilon} \right).$$

Finally

$$\left\| \frac{1}{Tp_1p_2} \sum_{t=1}^{T} RF_tC'E_t' \right\|_F \leq \frac{1}{Tp_1p_2} \left\| R \right\|_F \left\| \sum_{t=1}^{T} F_tC'E_t' \right\|_F = o_{a.s.} \left(\frac{1}{\sqrt{p_1T}} p_1^{\alpha_1/2} p_2^{\alpha_2/2-1} \left(\ln T \ln p_1 \ln^2 p_2 \right)^{1+\epsilon} \right),$$

by (C.6). Thus, using (C.7), it follows that $\left\| \widehat{\Lambda}_c \right\| = \Omega_{a.s.} \left(c_0 p_1^{\alpha_1-1} p_2^{\alpha_2-1} \right)$, and also $\left\| \widehat{\Lambda}_c^{-1} \right\| = \Omega_{a.s.} \left(c_0 p_1^{1-\alpha_1} p_2^{1-\alpha_2} \right)$; note that with our set-up ($k_1 = k_2 = 1$), $\widehat{\Lambda}_c$ is a scalar, but we prefer to use the matrix notation for the sake of generality. Hence, recalling that by construction $\left\| \widehat{C} \right\| = p_2^{1/2}$, it follows that

$$H_1 = \left(\frac{1}{Tp_1^{\alpha_1}} \sum_{t=1}^{T} F_t'R'RF_t \right) \frac{C'\widehat{C}}{p_2^{\alpha_2}} \left(p_1^{\alpha_1-1} p_1^{\alpha_2-1} \widehat{\Lambda}_c^{-1} \right) = \Omega_{a.s.} \left(p_2^{1/2-\alpha_2/2} \right). \tag{C.8}$$

As in the proof of Lemma C.2 we have

$$\begin{aligned}
\left\| \frac{1}{Tp_1p_2} C \sum_{t=1}^{T} F_t'R'E_t\widehat{C}\widehat{\Lambda}_c^{-1} \right\|_2^2 &\leq \frac{1}{(Tp_1p_2)^2} \left\| C \right\|_2^2 \left\| \widehat{C} \right\|_2^2 \left\| \widehat{\Lambda}_c^{-1} \right\|_2^2 \left\| \sum_{t=1}^{T} F_t'R'E_t \right\|_2^2 \\
&\leq c_0 \frac{1}{(Tp_1p_2)^2} p_2^{\alpha_2} p_2 p_1^{2-2\alpha_1} p_2^{2-2\alpha_2} \sum_{i=1}^{p_2} \left(\sum_{t=1}^{T} \sum_{j=1}^{p_1} R_j e_{ji,t} F_t \right)^2 \\
&\leq c_0 \frac{1}{(Tp_1p_2)^2} p_2^{\alpha_2} p_2 p_1^{2-2\alpha_1} p_2^{2-2\alpha_2} p_2 p_1^{\alpha_1} T = c_1 \frac{1}{T} p_1^{-\alpha_1} p_2^{2-\alpha_2}
\end{aligned}$$

having used (C.5) and Lemma C.1, whence

$$\left\| \frac{1}{Tp_1p_2} C \sum_{t=1}^{T} F_t'R'E_t\widehat{C}\widehat{\Lambda}_c^{-1} \right\|_2 = o_{a.s.} \left(T^{-1/2} p_1^{-\alpha_1/2} p_2^{1-\alpha_2/2} \left(\ln T \ln p_1 \ln^2 p_2 \right)^{1+\epsilon} \right). \tag{C.9}$$

The same holds for $III\widehat{C}\widehat{\Lambda}_c^{-1}$ by symmetry. Finally, using the results above

$$\begin{aligned}\left\|IV\widehat{C}\widehat{\Lambda}_c^{-1}\right\|_F &\leq \frac{1}{Tp_1p_2}\left\|\sum_{t=1}^{T}E_t'E_t\right\|_F\left\|\widehat{C}\right\|_F\left\|\widehat{\Lambda}_c^{-1}\right\|_F \\ &= o_{a.s.}\left(\left(\frac{1}{p_2}+\frac{1}{\sqrt{p_1T}}\right)\left(\ln T\ln p_1\ln^2 p_2\right)^{1+\epsilon}p_2^{1/2}p_1^{1-\alpha_1}p_2^{1-\alpha_2}\right).\end{aligned} \tag{C.10}$$

The desired result now follows from putting everything together and noting that the rate in (C.9) is always dominated by the one in (C.10). □

Lemma C.5. *We assume that the Assumptions of Lemma C.4 hold, and that*

$$p_1^{\alpha_1}p_2^{\alpha_2-1}\to\infty, \tag{C.11}$$

and

$$\sum_{j=1}^{p_1}\sum_{h,k=1}^{p_2}\left|C_hC_k\right|\left|E\left(e_{ih,t}e_{jk,t}\right)\right|\leq c_0Tp_2^{\alpha_2}, \tag{C.12}$$

for all $1\leq i\leq p_1$ and $1\leq t\leq T$, and

$$\sum_{i,j=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h,k,h_1,k_1=1}^{p_2}\left|C_hC_kC_{h_1}C_{k_1}\right|\left|Cov\left(e_{ih,t}e_{jk,t},e_{ih_1,s}e_{jk_1,s}\right)\right|\leq c_0p_1Tp_2^{3\alpha_2}. \tag{C.13}$$

Then it holds that

$$\lambda_j\left(\widetilde{M}_1\right)=\Omega_{a.s.}\left(c_0p_1^{\alpha_1}p_2^{\alpha_2-1}\right), \tag{C.14}$$

for all $j\leq k_1$, and

$$\begin{aligned}\lambda_j\left(\widetilde{M}_1\right) &= o_{a.s.}(1)+o_{a.s.}\left(\frac{p_1^{1/2+\alpha_1/2}p_2^{\alpha_2/2}}{p_2T^{1/2}}\left(\ln T\ln^2 p_1\ln p_2\right)^{1+\epsilon}\right) \\ &\quad +o_{a.s.}\left(\frac{p_1p_2^{1-\alpha_2/2}}{p_2T^{1/2}}p_1^{1-\alpha_1}p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(l_{p_1,p_2,T}'\right)^2\right),\end{aligned} \tag{C.15}$$

for all $j>k_1$.

Proof. The proof adapts the proof of Theorem 1, and we assume, for the sake of notational simplicity and no loss of generality, that $k_1=k_2=1$, also omitting, when possible, the rotation matrix H_1 defined in Lemma C.2. As in the proof of Theorem 1, we have

$$\begin{aligned}\frac{1}{p_1}\widetilde{M}_1 &= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\widehat{C}\widehat{C}'CF_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\widehat{C}\widehat{C}'E_t' \\ &\quad + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_t\widehat{C}\widehat{C}'CF_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_t\widehat{C}\widehat{C}'E_t' \\ &= \widetilde{M}_{1,1}+\widetilde{M}_{1,2}+\widetilde{M}_{1,3}+\widetilde{M}_{1,4},\end{aligned}$$

and

$$\lambda_j\left(\widetilde{M}_{1,1}\right)+\lambda_{\min}\left(\sum_{k=2}^{4}\widetilde{M}_{1,k}\right)\leq\lambda_j\left(\widetilde{M}_{1,1}\right)\leq\lambda_j\left(\widetilde{M}_{1,1}\right)+\lambda_{\max}\left(\sum_{k=2}^{4}\widetilde{M}_{1,k}\right).$$

By construction, it holds that, for all $j>k_1$

$$\lambda_j\left(\widetilde{M}_{1,1}\right)=0.$$

Conversely, when $j\leq k_1$ we have

$$\begin{aligned}\widetilde{M}_{1,1} &= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'CH_1H_1'C'CF_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-CH_1\right)\widehat{C}'CF_t'R' \\ &\quad + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\widehat{C}\left(\widehat{C}-CH_1\right)'CF_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-CH_1\right)\left(\widehat{C}-CH_1\right)'CF_t'R' \\ &= \widetilde{M}_{1,1,1}+\widetilde{M}_{1,1,2}+\widetilde{M}_{1,1,3}+\widetilde{M}_{1,1,4}.\end{aligned}$$

By Weyl's inequality and equation (C.8) - recall also that $k_1=k_2=1$

$$\lambda_j\left(\widetilde{M}_{1,1,1}\right)\geq p_2^{2\alpha_2-2}\lambda_j\left(p_1^{-1}R\left\|H_1\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T}E\left(F_t^2\right)\right)R'\right)+p_2^{2\alpha_2-2}\lambda_{\min}\left(p_1^{-1}R\left\|H_1\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2-E\left(F_t^2\right)\right)R'\right),$$

and the multiplicative Weyl inequality (Theorem 7 in Merikoski and Kumar, 2004) immediately yields

$$\lambda_j\left(p_1^{-1}R\left\|H_1\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T}E\left(F_t^2\right)\right)R'\right)\geq\left\|H_1\right\|^2\lambda_j\left(p_1^{-1}R'R\right)\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T}E\left(F_t^2\right)\right)\geq c_0p_1^{\alpha_1-1}p_2^{1-\alpha_2}.$$

Also, combining the results above with the proof of Theorem 1, it follows that

$$\lambda_{\max}\left(p_1^{-1}R\left\|H_1\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T}F_tF_t'-E\left(F_tF_t'\right)\right)R'\right)=o_{a.s.}\left(\frac{p_1^{\alpha_1-1}p_2^{1-\alpha_2}}{T^{1/2}}\left(\ln T\right)^{1/2+\epsilon}\right).$$

Thus, it follows that there exists a positive constant c_0 such that $\lambda_j\left(\widetilde{M}_{1,1,1}\right)=\Omega_{a.s.}\left(c_0p_1^{\alpha_1-1}p_2^{\alpha_2-1}\right)$.

Also (we will omit the rotation matrix H_1 when possible)

$$\left\|\widetilde{M}_{1,1,2}\right\|_F = \left\|\frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} RF_t C'\left(\widehat{C}-C\right)C'CF_t'R'\right\|_F \leq \frac{1}{p_1 p_2^{2-\alpha_2}}\left(\frac{1}{T}\sum_{t=1}^{T}\|F_t\|_F^2\right)\|R\|_F^2\left\|\widehat{C}-C\right\|_F\|C\|_F .$$

Using Lemma C.4 and the same arguments as in the proof of Theorem 1, it follows that

$$\begin{aligned}
& \frac{1}{p_1 p_2^{2-\alpha_2}}\left(\frac{1}{T}\sum_{t=1}^{T}\|F_t\|_F^2\right)\|R\|_F^2\left\|\widehat{C}-C\right\|_F\|C\|_F \\
= \quad & o_{a.s.}\left(\frac{1}{p_1 p_2^{2-\alpha_2}}p_1^{\alpha_1}p_2^{\alpha_2/2}\left(p_2^{1/2}p_1^{1-\alpha_1}p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\right)l'_{p_1,p_2,T}\right) \\
= \quad & o_{a.s.}\left(p_2^{\alpha_2/2-1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)l'_{p_1,p_2,T}\right),
\end{aligned}$$

where $l'_{p_1,p_2,T}$ is defined in (C.2); the same holds for $\left\|\widetilde{M}_{1,1,3}\right\|_F$. Note that these two terms are dominated by $\lambda_j\left(\widetilde{M}_{1,1,1}\right)$ on account of (C.7). We now turn to

$$\left\|\widetilde{M}_{1,1,4}\right\|_F \leq \frac{1}{p_1 p_2^2}\left(\frac{1}{T}\sum_{t=1}^{T}\|F_t\|_F^2\right)\|R\|_F^2\left\|\widehat{C}-C\right\|_F^2\|C\|_F^2 ;$$

repeating the same passages as above, it follows that

$$\begin{aligned}
\left\|\widetilde{M}_{1,1,4}\right\|_F &= o_{a.s.}\left(\frac{1}{p_1 p_2^2}p_1^{\alpha_1}p_2^{\alpha_2}\left(p_2 p_1^{2-2\alpha_1}p_2^{2-2\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)^2\right)l'^2_{p_1,p_2,T}\right) \\
&= o_{a.s.}\left(p_1^{1-\alpha_1}p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)^2 l'^2_{p_1,p_2,T}\right),
\end{aligned}$$

which again is dominated by $\lambda_j\left(\widetilde{M}_{1,1,1}\right)$ on account of (C.7). These results entail that there exists a positive constant c_0 such that

$$\lambda_j\left(\widetilde{M}_{1,1}\right) = \Omega_{a.s.}\left(c_0 p_1^{\alpha_1-1}p_2^{\alpha_2-1}\right), \tag{C.16}$$

whenever $j \leq k_1$. By (C.11), we note that the eigenvalue of $\widetilde{M}_1$ diverges, i.e.

$$p_1\lambda_j\left(\widetilde{M}_{1,1}\right) \to \infty \quad \text{a.s.}$$

thus ensuring that it can be detected.

We now turn to $\widetilde{M}_{1,2}$

$$\begin{aligned}
\left\|\widetilde{M}_{1,2}\right\|_F &\leq \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'CH_1H_1'C'E_t'\right\|_F+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-CH_1\right)H_1'C'E_t'\right\|_F \\
&\quad+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'CH_1\left(\widehat{C}-CH_1\right)'E_t'\right\|_F+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-CH_1\right)\left(\widehat{C}-CH_1\right)'E_t'\right\|_F \\
&= \left\|\widetilde{M}_{1,2,1}\right\|_F+\left\|\widetilde{M}_{1,2,2}\right\|_F+\left\|\widetilde{M}_{1,2,3}\right\|_F+\left\|\widetilde{M}_{1,2,4}\right\|_F.
\end{aligned}$$

Building on the proof of Theorem 1 that

$$\begin{aligned}
\left\|\widetilde{M}_{1,2,1}\right\|_F &= \|H_1\|_F^2\|C\|_F^2\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'E_t'\right\|_F \\
&\leq c_0p_2^{1-\alpha_2}p_2^{\alpha_2}\frac{1}{p_1p_2^2T}\left(\sum_{i=1}^{p_1}R_i^2\sum_{j=1}^{p_1}\left|\sum_{t=1}^{T}\sum_{h=1}^{p_2}C_hF_te_{jh,t}\right|^2\right)^{1/2} \\
&\leq c_0\frac{1}{p_1p_2T}p_1^{\alpha_1/2}\left(\sum_{j=1}^{p_1}\left|\sum_{t=1}^{T}\sum_{h=1}^{p_2}C_hF_te_{jh,t}\right|^2\right)^{1/2} \\
&= o_{a.s.}\left(\frac{p_1^{\alpha_1/2}p_2^{\alpha_2/2}}{p_1^{1/2}p_2T^{1/2}}\left(\ln T\ln^2p_1\ln p_2\right)^{1+\epsilon}\right);
\end{aligned}$$

having used Assumption B6 and (C.6). Consider now

$$\begin{aligned}
&\left\|\widetilde{M}_{1,2,2}\right\|_F \\
\leq\ &\frac{1}{p_1p_2^2T}\|C\|_F\|R\|_F\left\|\widehat{C}-C\right\|_F\|H_1\|_F\left\|\sum_{t=1}^{T}F_tC'E_t'\right\|_F \\
=\ &o_{a.s.}\left(\frac{p_1^{\alpha_1/2}p_2^{\alpha_2/2}}{p_1^{1/2}p_2T^{1/2}}p_1^{1-\alpha_1}p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(\ln T\ln^2p_1\ln p_2\right)^{1+\epsilon}\right),
\end{aligned}$$

thus being dominated by $\left\|\widetilde{M}_{1,2,1}\right\|_F$ in light of (C.7). The same holds, by symmetry, for $\left\|\widetilde{M}_{1,2,3}\right\|_F$. Finally, consider

$$\begin{aligned}
\left\|\widetilde{M}_{1,2,4}\right\|_F &= \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-CH_1\right)\left(\widehat{C}-CH_1\right)'E_t'\right\|_F \\
&\leq \frac{1}{p_1p_2^2T}\|R\|_F\|C\|_F\left\|\widehat{C}-CH_1\right\|_F\left\|\sum_{t=1}^{T}F_t\left(\widehat{C}-CH_1\right)'E_t'\right\|_F,
\end{aligned}$$

and

$$
\begin{aligned}
& \left\| \sum_{t=1}^{T} F_t \left(\widehat{C} - CH_1 \right)' E_t' \right\|_F \\
\leq \quad & \left\| \sum_{t=1}^{T} F_t \left(\frac{1}{Tp_1p_2} C \sum_{t=1}^{T} F_t' R' E_t \widehat{C} \widehat{\Lambda}_c^{-1} \right)' E_t' \right\|_F + \left\| \sum_{t=1}^{T} F_t \left(\frac{1}{Tp_1p_2} \sum_{t=1}^{T} E_t' R F_t C' \widehat{C} \widehat{\Lambda}_c^{-1} \right)' E_t' \right\|_F \\
& + \left\| \sum_{t=1}^{T} F_t \left(\frac{1}{Tp_1p_2} \sum_{t=1}^{T} E_t' E_t \widehat{C} \widehat{\Lambda}_c^{-1} \right)' E_t' \right\|_F \\
= \quad & \left\| \widetilde{M}_{1,2,4,1} \right\|_F + \left\| \widetilde{M}_{1,2,4,2} \right\|_F + \left\| \widetilde{M}_{1,2,4,3} \right\|_F .
\end{aligned}
$$

We have

$$
\begin{aligned}
\left\| \widetilde{M}_{1,2,4,1} \right\|_F \quad & \leq \quad \frac{1}{Tp_1p_2} \left\| \widehat{\Lambda}_c^{-1} \right\|_F \left\| \widehat{C} \right\|_F \left\| \sum_{s=1}^{T} E_s' R F_s \right\|_F \left\| \sum_{t=1}^{T} F_t C' E_t' \right\|_F \\
& = \quad o_{a.s.} \left(p_1^{1/2-\alpha_1/2} p_2^{1-\alpha_2/2} \left(\ln p_1 \ln p_2 \ln T \right)^{1+\epsilon} \right),
\end{aligned}
$$

whence

$$
\begin{aligned}
& \frac{1}{p_1 p_2^2 T} \|C\|_F \|R\|_F \left\| \widehat{C} - CH_1 \right\|_F \left\| \widetilde{M}_{1,2,4,1} \right\|_F \\
= \quad & o_{a.s.} \left(\frac{1}{p_1^{1/2} p_2^{1/2} T} p_1^{1-\alpha_1} p_2^{1-\alpha_2} \left(\frac{1}{p_2} + \frac{1}{\sqrt{Tp_1}} \right) \left(\ln p_1 \ln p_2 \ln T \right)^{1+\epsilon} l_{p_1,p_2,T}' \right),
\end{aligned}
$$

having used (C.5) and (C.6). Similarly

$$
\left\| \widetilde{M}_{1,2,4,2} \right\|_F \leq \frac{1}{Tp_1p_2} \left\| \widehat{\Lambda}_c^{-1} \right\|_F \left\| \widehat{C} \right\|_F \|C\|_F \left\| \sum_{t=1}^{T} E_t' F_t \right\|_F \left\| \sum_{s=1}^{T} F_s R' E_s \right\|_F ,
$$

and using

$$
\left\| \sum_{t=1}^{T} E_t' F_t \right\|_F = \left(\sum_{i=1}^{p_1} \sum_{j=1}^{p_2} \left| \sum_{t=1}^{T} F_t e_{ij,t} \right|^2 \right)^{1/2} = o_{a.s.} \left(p_1^{1/2} p_2^{1/2} T^{1/2} \left(\ln p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right),
$$

we have

$$
\left\| \widetilde{M}_{1,2,4,2} \right\|_F = o_{a.s.} \left(p_2^{1/2} p_1^{-1/2} p_1^{1-\alpha_1/2} p_2^{1-\alpha_2/2} \left(\ln p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right),
$$

so that

$$
\begin{aligned}
& \frac{1}{p_1 p_2^2 T} \|R\|_F \|C\|_F \left\|\widehat{C} - CH_1\right\|_F \left\|\widetilde{M}_{1,2,4,2}\right\|_F \\
= \quad & o_{a.s.}\left(\frac{1}{p_1^{1/2} p_2^{1/2} T} p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{T p_1}}\right)(\ln p_1 \ln p_2 \ln T)^{1/2+\epsilon} l_{p_1,p_2,T}^{\prime}\right).
\end{aligned}
$$

Also

$$
\left\|\widetilde{M}_{1,2,4,3}\right\|_F \leq \frac{1}{T p_1 p_2}\left\|\widehat{\Lambda}_c^{-1}\right\|_F\left\|\widehat{C}^{\prime}\right\|_F\left\|\sum_{s=1}^T E_s^{\prime} E_s\right\|_F\left\|\sum_{t=1}^T F_t E_t^{\prime}\right\|_F ;
$$

upon recalling that

$$
\left\|\sum_{s=1}^T E_s^{\prime} E_s\right\|_F = o_{a.s.}\left(\left(\frac{1}{p_2}+\frac{1}{\sqrt{p_1 T}}\right)\left(\ln T \ln p_1 \ln^2 p_2\right)^{1+\epsilon}\right),
$$

it follows that

$$
\left\|\widetilde{M}_{1,2,4,3}\right\|_F = o_{a.s.}\left(\frac{1}{p_1^{1/2} T^{1/2}} p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{p_1 T}}\right)(\ln p_1 \ln p_2 \ln T)^{3/2+\epsilon}\right),
$$

whence

$$
\begin{aligned}
& \frac{1}{p_1 p_2^2 T} \|R\|_F \|C\|_F \left\|\widehat{C} - CH_1\right\|_F \left\|\widetilde{M}_{1,2,4,3}\right\|_F \\
= \quad & o_{a.s.}\left(\frac{p_1^{\alpha_1/2} p_2^{\alpha_2/2}}{p_1^{3/2} p_2^{3/2} T^{3/2}}\left(p_1^{1-\alpha_1} p_2^{1-\alpha_2}\right)^2\left(\frac{1}{p_2}+\frac{1}{\sqrt{p_1 T}}\right)^2(\ln p_1 \ln p_2 \ln T)^{3/2+\epsilon}\right).
\end{aligned}
$$

Putting all together

$$
\left\|\widetilde{M}_{1,2,4}\right\|_F = o_{a.s.}\left(\frac{1}{p_1^{1/2} p_2^{1/2} T} p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{T p_1}}\right)(\ln p_1 \ln p_2 \ln T)^{1/2+\epsilon} l_{p_1,p_2,T}^{\prime}\right),
$$

which finally entails that, as $\min\{p_1, p_2, T\} \rightarrow \infty$

$$
p_1^{1-\alpha_1} p_2^{1-\alpha_2}\left\|\widetilde{M}_{1,2}\right\|_F = o_{a.s.}(1),
$$

with the same being true for $\left\|\widetilde{M}_{1,3}\right\|_F$ due to symmetry. Finally

$$\begin{aligned}
\widetilde{M}_{1,4} &= \frac{1}{p_1p_2^2T}\sum_{t=1}^T E_tCH_1H_1'C'E_t' + \frac{1}{p_1p_2^2T}\sum_{t=1}^T E_tCH_1\left(\widehat{C}-CH_1\right)'E_t' \\
&\quad + \frac{1}{p_1p_2^2T}\sum_{t=1}^T E_t\left(\widehat{C}-CH_1\right)H_1'C'E_t' + \frac{1}{p_1p_2^2T}\sum_{t=1}^T E_t\left(\widehat{C}-CH_1\right)\left(\widehat{C}-CH_1\right)'E_t' \\
&= \widetilde{M}_{1,4,1}+\widetilde{M}_{1,4,2}+\widetilde{M}_{1,4,3}+\widetilde{M}_{1,4,4}.
\end{aligned}$$

We know from the proof of Theorem 1 that

$$\left\|\widetilde{M}_{1,4,1}\right\|_F \leq \frac{1}{p_1p_2^2T}\|H_1\|^2\left\|\sum_{t=1}^T E\left(E_tCC'E_t'\right)\right\|_F + \frac{1}{p_1p_2^2T}\|H_1\|^2\left\|\sum_{t=1}^T E_tCC'E_t' - E\left(E_tCC'E_t'\right)\right\|_F.$$

We now have

$$\left\|\sum_{t=1}^T E\left(E_tCC'E_t'\right)\right\|_F \leq \left\|\sum_{t=1}^T E\left(E_tCC'E_t'\right)\right\|_1 \leq \max_{1\leq i\leq p_1}\sum_{j=1}^{p_1}\left|\sum_{t=1}^T\sum_{h,k=1}^{p_2} C_hC_kE\left(e_{ih,t}e_{jk,t}\right)\right| \leq c_0Tp_2^{\alpha_2},$$

having used Assumption B3*(ii)* and (C.12). Also

$$\begin{aligned}
&\left\|\sum_{t=1}^T E_tCC'E_t' - E\left(E_tCC'E_t'\right)\right\|_F^2 \\
=\;& \sum_{i,j=1}^{p_1}\left(\sum_{t=1}^T\sum_{h,k=1}^{p_2} C_hC_k\left(e_{ih,t}e_{jk,t}-E\left(e_{ih,t}e_{jk,t}\right)\right)\right)^2 \\
\leq\;& \sum_{i,j=1}^{p_1}\sum_{t,s=1}^T\sum_{h,k,h_1,k_1=1}^{p_2}\left|C_hC_kC_{h_1}C_{k_1}\right|\left|Cov\left(e_{ih,t}e_{jk,t},e_{ih_1,s}e_{jk_1,s}\right)\right| \\
\leq\;& c_0p_1Tp_2^{3\alpha_2},
\end{aligned}$$

by (C.13). Hence, putting all together, we have

$$\left\|\widetilde{M}_{1,4,1}\right\|_F = o_{a.s.}\left(\frac{1}{p_1p_2}l_{p_1,p_2,T}'\right) + o_{a.s.}\left(\frac{1}{p_1^{1/2}T^{1/2}}p_2^{\alpha_2/2-1}l_{p_1,p_2,T}'\right).$$

Consider now $\widetilde{M}_{1,4,3}$; as in the proof of Theorem 1, $E_{\cdot,j,t}$ represents the j-th column of E_t, and $\left(\widehat{C}-C\right)_j$ is the j-th element of $\left(\widehat{C}-C\right)$. As in the proof of Theorem 1, it holds that

$$\left\|\widetilde{M}_{1,4,3}\right\|_2 \leq \|H_1\|\left\|\widehat{C}-C\right\|_F\left(\sum_{j=1}^{p_2}\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^T E_{\cdot,j,t}C'E_t'\right\|_2^2\right)^{1/2},$$

and we know from the proof of Theorem 1 that

$$\left(\sum_{j=1}^{p_2}\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_{\cdot,j,t}C'E_t'\right\|_2^2\right)^{1/2}=o_{a.s.}\left(\left(\frac{1}{Tp_2^2}+\frac{1}{p_1^2p_2^3}\right)^{1/2}l_{p_1,p_2,T}'\right);$$

we note that this may not be the sharpest bound, but it suffices for our purposes. Hence, using Lemma C.4

$$\left\|\widetilde{M}_{1,4,3}\right\|_2=o_{a.s.}\left(p_2^{1-\alpha_2/2}p_1^{1-\alpha_1}p_2^{1-\alpha_2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(\frac{1}{Tp_2^2}+\frac{1}{p_1^2p_2^3}\right)^{1/2}\left(l_{p_1,p_2,T}'\right)^2\right),$$

with the same holding for $\left\|\widetilde{M}_{1,4,2}\right\|_2$. Finally, following *verbatim* the proof of Theorem 1 and using Lemma C.4, it follows that

$$\left\|\widetilde{M}_{1,4,4}\right\|_2=o_{a.s.}\left(\frac{1}{p_1}p_1^{2-2\alpha_1}p_2^{2-2\alpha_2}\left(\frac{1}{p_2^2}+\frac{1}{Tp_1}\right)\left(\ln T\ln^2 p_2\right)^{1/2+\epsilon}l_{p_1,p_2,T}'^2\right).$$

Putting all together, it follows that

$$\left\|\widetilde{M}_{1,4}\right\|_2=o_{a.s.}(1).$$

The proof of the final result is now complete. □

D. Proofs

Henceforth, we let E^* and V^* denote the expectation and variance with respect to P^* respectively.

Some arguments in the proof of this theorem are very similar, and in fact easier, than the ones in the next proof, to which we refer for further details.

We begin by considering the case where both k_1 and k_2 are positive. Let C_i denote the i-th column of C', and note

$$\begin{aligned}
& \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} X_{\cdot i,t}X_{\cdot i,t}' \\
= & \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} \left(RF_tC_i + E_{\cdot i,t}\right)\left(RF_tC_i + E_{\cdot i,t}\right)' \\
= & \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} RF_tC_iC_i'F_t'R' + \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} E_{\cdot i,t}E_{\cdot i,t}' \\
& + \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} RF_tC_iE_{\cdot i,t}' + \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} E_{\cdot i,t}C_i'F_t'R' \\
= & I + II + III + IV,
\end{aligned}$$

so that, by Weyl's inequality

$$\lambda_j\left(I\right) + \lambda_{\min}\left(II + III + IV\right) \leq \lambda_j\left(\frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} X_{\cdot i,t}X_{\cdot i,t}'\right) \leq \lambda_j\left(I\right) + \lambda_{\max}\left(II\right) + \lambda_{\max}\left(III + IV\right).$$

Note that

$$\begin{aligned}
I & = \frac{1}{p_2 T}\sum_{t=1}^{T}\sum_{i=1}^{p_2} RF_tC_iC_i'F_t'R' \\
& = R\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)R',
\end{aligned}$$

by Assumption B2*(ii)*. It is easy to see that, for all $j > k_1$, $\lambda_j\left(R\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)R'\right) = 0$; when $j \leq k_1$, the multiplicative Weyl's inequality (see Theorem 7 in Merikoski and Kumar, 2004) yields

$$\lambda_j\left(R\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)R'\right) = \lambda_j\left(\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)R'R\right) \geq \lambda_j\left(R'R\right)\lambda_{\min}\left(\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)\right);$$

by Assumption B1*(i)*, $\lambda_{\min}\left(\left(\frac{1}{T}\sum_{t=1}^{T} F_tF_t'\right)\right) \geq c_0 + o_{a.s.}(1)$ with $c_0 > 0$; using Assumption B2*(ii)*, it follows that

$$\lambda_j \left(R \left(\frac{1}{T} \sum_{t=1}^{T} F_t F_t' \right) R' \right) = \Omega_{a.s.} \left(p_1 \right).$$

Consider now II

$$\begin{aligned}
& \lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} E_{\cdot i,t} E_{\cdot i,t}' \right) \\
= \; & \lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) + \lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \left(E_{\cdot i,t} E_{\cdot i,t}' - E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) \right).
\end{aligned}$$

Using Weyl's inequality and Assumption B3*(ii)*(b)

$$\begin{aligned}
& \lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) \\
\leq \; & \frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \lambda_{\max} \left(E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) \\
\leq \; & \frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \max_{1 \leq h \leq p_1} \sum_{k=1}^{p_1} \left| E \left(e_{hi,t} e_{ki,t} \right) \right| \leq c_0.
\end{aligned} \tag{D.1}$$

Also

$$\begin{aligned}
& \lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \left(E_{\cdot i,t} E_{\cdot i,t}' - E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) \right) \\
\leq \; & \left(\sum_{h,k=1}^{p_1} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \left(e_{hi,t} e_{ki,t} - E \left(e_{hi,t} e_{ki,t} \right) \right) \right)^2 \right)^{1/2}.
\end{aligned}$$

Consider now

$$\begin{aligned}
& E \sum_{h,k=1}^{p_1} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \left(e_{hi,t} e_{ki,t} - E \left(e_{hi,t} e_{ki,t} \right) \right) \right)^2 \\
= \; & \frac{1}{p_2^2 T^2} \sum_{h,k=1}^{p_1} \sum_{i,j=1}^{p_2} \sum_{t,s=1}^{T} Cov \left(e_{hi,t} e_{ki,t}, e_{hj,s} e_{kj,s} \right) \\
\leq \; & c_0 \frac{p_1^2}{p_2 T},
\end{aligned}$$

having used Assumption B3*(iii)* in the last passage. Thus, by Lemma C.1

$$\lambda_{\max} \left(\frac{1}{p_2 T} \sum_{t=1}^{T} \sum_{i=1}^{p_2} \left(E_{\cdot i,t} E_{\cdot i,t}' - E \left(E_{\cdot i,t} E_{\cdot i,t}' \right) \right) \right) = o_{a.s.} \left(\frac{p_1}{\sqrt{p_2 T}} \left(\ln^2 p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right), \tag{D.2}$$

for all $\epsilon > 0$. Thus, there exist a positive, finite constants c_0 such that

$$\lambda_{\max}(II) = c_0 + o_{a.s.}\left(\frac{p_1}{\sqrt{p_2 T}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right).$$

Finally, note that

$$\begin{aligned} & \lambda_{\max}\left(\frac{1}{p_2 T}\sum_{t=1}^{T}\left(RF_t C' E_t' + E_t C F_t' R'\right)\right) \\ = & \; 2\lambda_{\max}\left(\frac{1}{p_2 T}\sum_{t=1}^{T} RF_t C' E_t'\right) \le 2\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t C' E_t' R\right\|_F \end{aligned}$$

and

$$\begin{aligned} & E\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t C' E_t' R\right\|_F^2 \\ = & \; \frac{1}{p_2^2 T^2} E\sum_{h,k=1}^{p_1}\left(\sum_{i=1}^{p_2}\sum_{t=1}^{T} R_h F_t C_i e_{ki,t}\right)^2 \\ = & \; \frac{1}{p_2^2 T^2}\sum_{h,k=1}^{p_1}\sum_{i,j=1}^{p_2}\sum_{t,s=1}^{T} R_h^2 C_i C_j E\left(F_t F_s e_{ki,t} e_{kj,s}\right) \\ \le & \; \frac{p_1^2}{p_2 T} \end{aligned}$$

having used Assumptions B4*(i)*-*(ii)*, so that, using Lemma C.1, it follows that

$$\lambda_{\max}(III + IV) \le o_{a.s.}\left(\frac{p_1}{\sqrt{p_2 T}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),$$

for all $\epsilon > 0$. The desired result follows from putting all together.

We now turn to consider the boundary cases where either $k_1 = 0$, or $k_2 = 0$, or both $k_1 = k_2 = 0$. In all such cases, we set the nonzero number of factors equal to 1 for simplicity and without loss of generality. When $k_1 = k_2 = 0$, it holds that

$$\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} X_{\cdot i,t} X_{\cdot i,t}' = \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} E_{\cdot i,t} E_{\cdot i,t}',$$

and we already know that

$$\lambda_{\max}\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} E_{\cdot i,t} E_{\cdot i,t}'\right) \le c_0 + o_{a.s.}\left(\frac{p_1}{\sqrt{p_2 T}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),$$

from (D.1) and (D.2); hence

$$\lambda_1\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}X_{\cdot i,t}X_{\cdot i,t}'\right)\leq c_0+o_{a.s.}\left(\frac{p_1}{\sqrt{p_2T}}\left(\ln^2 p_1\ln p_2\ln T\right)^{1/2+\epsilon}\right),$$

which completes the proof under this case. When $k_1=1$ and $k_2=0$, we have, as in (C.3)-(C.4)

$$X_t=\underset{p_1\times k_1}{R}\underset{k_1\times p_2}{F_t}+E_t,$$

and

$$X_{\cdot i,t}=RF_{\cdot i,t}+E_{\cdot i,t}.$$

Therefore

$$\begin{aligned}
&\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}X_{\cdot i,t}X_{\cdot i,t}'\\
=&\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}RF_{\cdot i,t}^2R'+\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}RF_{\cdot i,t}E_{\cdot i,t}'\\
&+\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}E_{\cdot i,t}F_{\cdot i,t}R'+\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}E_{\cdot i,t}E_{\cdot i,t}'\\
=&I+II+III+IV.
\end{aligned}$$

From the proof of Lemma C.3, we already know that

$$\lambda_{\max}\left(II+III+IV\right)=c_0+o_{a.s.}\left(\frac{p_1}{\sqrt{p_2T}}\left(\ln^2 p_1\ln p_2\ln T\right)^{1/2+\epsilon}\right),$$

and

$$\lambda_1\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T}RF_{\cdot i,t}^2R'\right)\geq c_0p_1>0\quad\text{a.s.}$$

whence again the desired result. Finally, we consider the more delicate case $k_2>0$ (letting, as customary, $k_2=1$ for simplicity) and $k_1=0$. In such a case, we have

$$X_t=\underset{p_1\times 1}{F_t}\underset{1\times p_2}{C'}+E_t, \tag{D.3}$$

and

$$X_{\cdot i,t}=F_tC_i+E_{\cdot i,t}, \tag{D.4}$$

where C_i represents the i-th column of C' (or, in the case whereby $k_2=1$, the i-th element of C'). Then

we have

$$
\begin{aligned}
& \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} X_{\cdot i,t}X_{\cdot i,t}' \\
= \; & \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i^2 F_t F_t' + \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i F_t E_{\cdot i,t}' \\
& + \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i E_{\cdot i,t} F_t' + \frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} E_{\cdot i,t}E_{\cdot i,t}' \\
= \; & I + II + III + IV.
\end{aligned}
$$

Repeating the same arguments as above, it is immediate to see that, for all $j \leq k_1$

$$
\lambda_{\max}\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i^2 F_t F_t'\right) \leq c_0 \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right),
$$

and

$$
\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right) = O_{a.s.}\left(\left(1+\sqrt{\frac{p_1}{T}}\right)^2\right) \quad \text{a.s.},
$$

by Assumption B1*(iv)*, whence

$$
\lambda_{\max}\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i^2 F_t F_t'\right) = c_0 + O_{a.s.}\left(\frac{p_1}{T}\right). \tag{D.5}
$$

Finally, consider the expectation of the (square) of the Frobenius norm of III

$$
\begin{aligned}
& E\sum_{j,k=1}^{p_1}\left|\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i F_{j,t} E_{ki,t}\right|^2 \\
= \; & E\frac{1}{(Tp_2)^2}\sum_{j,k=1}^{p_1}\sum_{i,l=1}^{p_2}\sum_{t,s=1}^{T} C_i C_l F_{j,t}F_{j,s}E_{ki,t}E_{kl,s} \\
\leq \; & \frac{1}{(Tp_2)^2}\sum_{j,k=1}^{p_1}\sum_{i,l=1}^{p_2}\sum_{t,s=1}^{T} |C_i|\,|C_l|\,|E(F_{j,t}F_{j,s}E_{ki,t}E_{kl,s})| \\
\leq \; & \frac{1}{(Tp_2)^2}\sum_{j,k=1}^{p_1}\sum_{i,l=1}^{p_2}\sum_{t,s=1}^{T} |E(F_{j,t}F_{j,s}E_{ki,t}E_{kl,s})| \leq c_0\frac{p_1^2}{Tp_2},
\end{aligned}
$$

by Assumptions B4*(i)*-*(ii)*. The same holds for II; similarly, it is easy to see that the same rate as for the case $k_1 > 0$ is found for IV when $k_1 = 0$. The desired result now follows from the same arguments as above.

In the proof of the theorem, we assume, for the sake of notational simplicity and no loss of generality, that $k_1 = k_2 = 1$; and we omit the rotation matrix H_1 defined in Lemma C.2. Further, some passages have already been shown by Yu et al. (2022), and when possible we omit them to save space. Recall that

$$\begin{aligned}
\widetilde{Y}_t &= \frac{1}{p_2} X_t \widehat{C}, \\
\frac{1}{p_1} \widetilde{M}_1 &= \frac{1}{p_1 T} \sum_{t=1}^{T} \widetilde{Y}_t \widetilde{Y}_t'.
\end{aligned}$$

Then we can write

$$\begin{aligned}
\widetilde{M}_1 &= \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} \left(R F_t C' \widehat{C} + E_t \widehat{C} \right) \left(R F_t C' \widehat{C} + E_t \widehat{C} \right)' \qquad \text{(D.6)} \\
&= \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' \widehat{C} \widehat{C}' C F_t' R' + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' \widehat{C} \widehat{C}' E_t' \\
&\quad + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} E_t \widehat{C} \widehat{C}' C F_t' R' + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} E_t \widehat{C} \widehat{C}' E_t' \\
&= \widetilde{M}_{1,1} + \widetilde{M}_{1,2} + \widetilde{M}_{1,3} + \widetilde{M}_{1,4}.
\end{aligned}$$

We begin by noting that, by Weyl's inequality

$$\lambda_j \left(\widetilde{M}_{1,1} \right) + \lambda_{\min} \left(\sum_{k=2}^{4} \widetilde{M}_{1,k} \right) \leq \lambda_j \left(\widetilde{M}_{1,1} \right) \leq \lambda_j \left(\widetilde{M}_{1,1} \right) + \lambda_{\max} \left(\sum_{k=2}^{4} \widetilde{M}_{1,k} \right). \qquad \text{(D.7)}$$

By construction, it holds that, for all $j > k_1$

$$\lambda_j \left(\widetilde{M}_{1,1} \right) = 0. \qquad \text{(D.8)}$$

We now study the case $j \leq k_1$; we will use the decomposition

$$\begin{aligned}
\widetilde{M}_{1,1} &= \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' C C' C F_t' R' + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' \left(\widehat{C} - C \right) \widehat{C}' C F_t' R' \\
&\quad + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' \widehat{C} \left(\widehat{C} - C \right)' C F_t' R' + \frac{1}{p_1 p_2^2 T} \sum_{t=1}^{T} R F_t C' \left(\widehat{C} - C \right) \left(\widehat{C} - C \right)' C F_t' R' \\
&= \widetilde{M}_{1,1,1} + \widetilde{M}_{1,1,2} + \widetilde{M}_{1,1,3} + \widetilde{M}_{1,1,4}.
\end{aligned}$$

We have

$$\begin{aligned}\lambda_j\left(\widetilde{M}_{1,1,1}\right) &= \lambda_j\left(\frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} RF_t C' C C' C F_t' R'\right)\\ &= \lambda_j\left(p_1^{-1} R\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right) R'\right)\\ &\geq \lambda_j\left(p_1^{-1} R\left(\frac{1}{T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right) R'\right) + \lambda_{\min}\left(p_1^{-1} R\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t' - E\left(F_t F_t'\right)\right) R'\right).\end{aligned}$$

By the multiplicative Weyl inequality (Theorem 7 in Merikoski and Kumar, 2004), we have

$$\lambda_j\left(p_1^{-1} R\left(\frac{1}{T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right) R'\right) \geq \lambda_j\left(p_1^{-1} R' R\right)\lambda_{\min}\left(\left(\frac{1}{T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right)\right) \geq c_0,$$

having used Assumptions B1 and B2. Also

$$\begin{aligned}&\lambda_{\max}\left(p_1^{-1} R\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t' - E\left(F_t F_t'\right)\right) R'\right)\\ \leq\ & \lambda_{\max}\left(p_1^{-1} R' R\right)\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t' - E\left(F_t F_t'\right)\right)\right)\\ \leq\ & c_0 \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t' - E\left(F_t F_t'\right)\right)\right).\end{aligned}$$

We now have

$$\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t' - E\left(F_t F_t'\right)\right)\right) \leq \left\|\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t' - E\left(F_t F_t'\right)\right)\right\|_F,$$

and

$$\left\|\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t' - E\left(F_t F_t'\right)\right)\right\|_F = \left(\sum_{h,k=1}^{k_1}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_{h,t}F_{k,t} - E\left(F_{h,t}F_{k,t}\right)\right)\right)^2\right)^{1/2}.$$

Assumption B1*(iii)* yields

$$\sum_{h,k=1}^{k_1} E\max_{1\leq\widetilde{t}\leq T}\left|\sum_{t=1}^{T}\left(F_{h,t}F_{k,t} - E\left(F_{h,t}F_{k,t}\right)\right)\right|^2 \leq c_0 T,$$

which in turn, by Lemma C.1, entails

$$\sum_{h,k=1}^{p_1}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_{h,t}F_{k,t} - E\left(F_{h,t}F_{k,t}\right)\right)\right)^2 = o_{a.s.}\left(T^{-1}\left(\ln T\right)^{1+\epsilon}\right),$$

for all $\epsilon > 0$, whence

$$\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\left(F_tF_t' - E\left(F_tF_t'\right)\right)\right) = o_{a.s.}\left(T^{-1/2}\left(\ln T\right)^{1/2+\epsilon}\right).$$

Thus, it follows that there exists a positive constant c_0 such that $\lambda_j\left(\widetilde{M}_{1,1,1}\right) = \Omega_{a.s.}\left(c_0\right)$. Also

$$\begin{aligned}\left\|\widetilde{M}_{1,1,2}\right\|_F &= \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-C\right)C'CF_t'R'\right\|_F \\ &\leq \frac{1}{p_1p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|F_t\right\|_F^2\right)\left\|R\right\|_F^2\left\|\widehat{C}-C\right\|_F\left\|C\right\|_F.\end{aligned}$$

By Assumption B2, $\|R\|_F^2 \leq c_0p_1$ and $\|C\|_F \leq c_0p_2^{1/2}$; further, by similar arguments as above, it is easy to see that $T^{-1}\sum_{t=1}^{T}\|F_t\|_F^2 = O_{a.s.}(1)$. Finally, applying Lemma C.2, it follows that

$$\left\|\widetilde{M}_{1,1,2}\right\|_F = o_{a.s.}\left(\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)l'_{p_1,p_2,T}\right), \tag{D.9}$$

where $l'_{p_1,p_2,T}$ is defined in (C.2). The same result holds for $\left\|\widetilde{M}_{1,1,3}\right\|_F$, by exactly the same passages. We now turn to

$$\left\|\widetilde{M}_{1,1,4}\right\|_F \leq \frac{1}{p_1p_2^2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|F_t\right\|_F^2\right)\left\|R\right\|_F^2\left\|\widehat{C}-C\right\|_F^2\left\|C\right\|_F^2;$$

noting that $\|R\|_F^2 \leq c_0p_1$ and $\|C\|_F^2 \leq c_0p_2$, and using again Lemma C.2, it immediately follows that

$$\left\|\widetilde{M}_{1,1,4}\right\|_F = o_{a.s.}\left(\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)^2\left(l'_{p_1,p_2,T}\right)^2\right).$$

These results entail that there exists a positive constant c_0 such that $\lambda_j\left(\widetilde{M}_{1,1}\right) = \Omega_{a.s.}\left(c_0\right)$, whenever $j \leq k_1$.

We now bound the other terms in (D.6), using the bound

$$\lambda_{\max}\left(\sum_{k=2}^{4}\widetilde{M}_{1,k}\right) \leq \left\|\sum_{k=2}^{4}\widetilde{M}_{1,k}\right\|_2 \leq \sum_{k=2}^{4}\left\|\widetilde{M}_{1,k}\right\|_2.$$

Consider $\widetilde{M}_{1,2}$ first, and note that

$$\begin{aligned}
\left\|\widetilde{M}_{1,2}\right\|_F &\leq \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'CC'E_t'\right\|_F+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-C\right)C'E_t'\right\|_F \\
&\quad+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'C\left(\widehat{C}-C\right)'E_t'\right\|_F+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-C\right)\left(\widehat{C}-C\right)'E_t'\right\|_F \\
&= \left\|\widetilde{M}_{1,2,1}\right\|_F+\left\|\widetilde{M}_{1,2,2}\right\|_F+\left\|\widetilde{M}_{1,2,3}\right\|_F+\left\|\widetilde{M}_{1,2,4}\right\|_F.
\end{aligned}$$

It holds that

$$\left\|\widetilde{M}_{1,2,1}\right\|_F=\frac{1}{p_1p_2^2T}\left(\sum_{i,j=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h=1}^{p_2}C_hR_iF_te_{jh,t}\right)^2\right)^{1/2},$$

and, expanding the square and taking the expectation, we now study

$$\begin{aligned}
&\sum_{i,j=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h_1,h_2=1}^{p_2}C_{h_1}C_{h_2}R_i^2E\left(F_tF_se_{jh_1,t}e_{jh_2,s}\right) \\
\leq\;& c_0\sum_{i,j=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h_1,h_2=1}^{p_2}\left|E\left(F_tF_se_{jh_1,t}e_{jh_2,s}\right)\right|.
\end{aligned}$$

Using Assumptions B4*(i)* and B4*(ii)*(a), it now follows that

$$E\sum_{i,j=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h=1}^{p_2}C_hR_iF_te_{jh,t}\right)^2\leq c_0p_1^2p_2T,$$

so that, by Lemma C.1, it follows that

$$\left\|\widetilde{M}_{1,2,1}\right\|_F=o_{a.s.}\left(\frac{1}{\sqrt{Tp_2}}\left(\ln T\ln^2p_1\ln p_2\right)^{1+\epsilon}\right),$$

for all $\epsilon>0$. Consider now

$$\begin{aligned}
\left\|\widetilde{M}_{1,2,2}\right\|_2 &\leq \frac{1}{p_1p_2^2T}\left\|\sum_{t=1}^{T}RF_tC'\left(\widehat{C}-C\right)C'E_t'\right\|_F \\
&\leq \frac{1}{p_1p_2^2T}\left|C'\left(\widehat{C}-C\right)\right|\left(\sum_{i,k=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h=1}^{p_2}R_iC_hF_te_{kh,t}\right)^2\right)^{1/2}.
\end{aligned}$$

Using the Cauchy-Schwartz inequality and Lemma C.2, it follows that

$$\left|C'\left(\widehat{C}-C\right)\right|=o_{a.s.}\left(p_2\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)l'_{p_1,p_2,T}\right).$$

Also note that

$$
\begin{aligned}
& \sum_{i,k=1}^{p_1} E\left(\sum_{t=1}^{T}\sum_{h=1}^{p_2} R_i C_h F_t e_{kh,t}\right)^2 \\
\leq & \sum_{i,k=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h_1,h_2=1}^{p_2} E\left(R_i^2 C_{h_1} C_{h_2} F_t F_s e_{kh_1,t} e_{kh_1,s}\right) \\
\leq & c_0 p_1^2 p_2 T,
\end{aligned}
$$

having used again Assumptions B4*(i)* and B4*(ii)*(a). Thus, using Lemma C.1 and putting all together, it holds that

$$
\left\|\widetilde{M}_{1,2,2}\right\|_F = o_{a.s.}\left(\frac{1}{p_2^{1/2}T^{1/2}}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right) l'_{p_1,p_2,T}\left(\ln T \ln^2 p_1 \ln p_2\right)^{1/2+\epsilon}\right).
$$

Also

$$
\begin{aligned}
\left\|\widetilde{M}_{1,2,3}\right\|_F &= \frac{1}{p_1 p_2^2 T}\left\|\sum_{t=1}^{T} R F_t C' C\left(\widehat{C}-C\right)' E_t'\right\|_F \\
&= \frac{1}{p_1 p_2 T}\left\|\sum_{t=1}^{T} R F_t\left(\widehat{C}-C\right)' E_t'\right\|_F \\
&\leq \frac{1}{p_1}\left\|R\right\|_F\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t\left(\widehat{C}-C\right)' E_t'\right\|_F;
\end{aligned}
$$

Following the proof of Lemma B.3 in Yu et al. (2022), it can be show that

$$
E\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t\left(\widehat{C}-C\right)' E_t\right\|_F^2 \leq c_0 p_1\left(\frac{1}{Tp_2^2}+\frac{1}{\left(Tp_1\right)^2}\right),
$$

and therefore, by the same arguments as above, it follows that

$$
\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t\left(\widehat{C}-C\right)' E_t'\right\|_F = o_{a.s.}\left(p_1^{1/2}\left(\frac{1}{Tp_2^2}+\frac{1}{\left(Tp_1\right)^2}\right)^{1/2}\left(\ln p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),
$$

for all $\epsilon > 0$, so that ultimately

$$
\left\|\widetilde{M}_{1,2,3}\right\|_F = o_{a.s.}\left(\left(\frac{1}{T^{1/2}p_2}+\frac{1}{Tp_1}\right)\left(\ln p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right).
$$

Finally, by the same logic we have that

$$
\begin{aligned}
\left\|\widetilde{M}_{1,2,4}\right\|_F &= \frac{1}{p_1 p_2^2 T}\left\|\sum_{t=1}^{T} R F_t C'\left(\widehat{C}-C\right)\left(\widehat{C}-C\right)' E_t'\right\|_F \\
&\leq \frac{1}{p_1 p_2}\|R\|_F\|C\|_F\left\|\left(\widehat{C}-C\right)\right\|_F\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} F_t\left(\widehat{C}-C\right)' E_t'\right\|_F \\
&= \Omega_{a.s.}\left[p_1^{-1/2} p_2^{-1/2}\left(p_2^{1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{T p_1}}\right) l'_{p_1,p_2,T}\right)\times\right. \\
&\quad \left.\left(p_1^{1/2}\left(\frac{1}{T p_2^2}+\frac{1}{(T p_1)^2}\right)^{1/2}(\ln p_1 \ln p_2 \ln T)^{1/2+\epsilon}\right)\right] \\
&= \Omega_{a.s.}\left[\left(\frac{1}{p_2}+\frac{1}{\sqrt{T p_1}}\right)\left(\frac{1}{T^{1/2} p_2}+\frac{1}{T p_1}\right) l'_{p_1,p_2,T}(\ln p_1 \ln p_2 \ln T)^{1/2+\epsilon}\right].
\end{aligned}
$$

Putting all together, it follows that

$$
\left\|\widetilde{M}_{1,2}\right\|_F = o_{a.s.}\left(\left(\frac{1}{\sqrt{T p_2}}+\frac{1}{T p_1}\right)\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1+\epsilon}\right);
$$

the same holds for $\left\|\widetilde{M}_{1,3}\right\|_F$. Finally, consider

$$
\begin{aligned}
\widetilde{M}_{1,4} &= \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_t C C' E_t' + \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_t C\left(\widehat{C}-C\right)' E_t' \\
&\quad + \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_t\left(\widehat{C}-C\right) C' E_t' + \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_t\left(\widehat{C}-C\right)\left(\widehat{C}-C\right)' E_t' \\
&= \widetilde{M}_{1,4,1}+\widetilde{M}_{1,4,2}+\widetilde{M}_{1,4,3}+\widetilde{M}_{1,4,4}.
\end{aligned}
$$

We have

$$
\begin{aligned}
\widetilde{M}_{1,4,1} &= \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E\left(E_t C C' E_t'\right) + \frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T}\left(E_t C C' E_t' - E\left(E_t C C' E_t'\right)\right) \\
&= \widetilde{M}_{1,4,1,1}+\widetilde{M}_{1,4,1,2}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\lambda_{\max}\left(\widetilde{M}_{1,4,1,1}\right) &\leq \frac{1}{p_1 p_2^2}\max_{1\leq j\leq p_1}\sum_{i=1}^{p_1}\left|E\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{h_1,h_2=1}^{p_2} C_{h_1} C_{h_2} e_{i h_1,t} e_{j h_2,t}\right)\right| \\
&\leq \frac{1}{T p_1 p_2^2}\max_{1\leq j\leq p_1}\sum_{t=1}^{T}\sum_{h_1,h_2=1}^{p_2}\sum_{i=1}^{p_1}\left|E\left(e_{i h_1,t} e_{j h_2,t}\right)\right| \\
&\leq c_0 p_2\frac{1}{p_1 p_2^2} = c_0\frac{1}{p_1 p_2},
\end{aligned}
$$

having used Assumption B3*(ii)*. Also, letting $\eta_{i,j,h_1,h_2,t} = e_{ih_1,t}e_{jh_2,t} - E\left(e_{ih_1,t}e_{jh_2,t}\right)$

$$\left\|\widetilde{M}_{1,4,1,2}\right\|_F = \frac{1}{p_1 p_2^2 T}\left(\sum_{i,j=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h_1,h_2=1}^{p_2} C_{h_1}C_{h_2}\eta_{t,i,j,h_1,h_2}\right)^2\right)^{1/2},$$

and

$$\begin{aligned}
& E\sum_{i,j=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h_1,h_2=1}^{p_2} C_{h_1}C_{h_2}\eta_{t,i,j,h_1,h_2}\right)^2 \\
\leq\ & c_0\sum_{i,j=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h_1,h_2,h_3,h_4=1}^{p_2}\left|Cov\left(e_{ih_1,t}e_{jh_2,t}, e_{ih_3,s}e_{jh_4,s}\right)\right| \\
\leq\ & c_0 p_1^2 T p_2^3,
\end{aligned}$$

in light of Assumption B3*(iii)*. Using Lemma C.1, we therefore have

$$\sum_{i,j=1}^{p_1}\left(\sum_{t=1}^{T}\sum_{h_1,h_2=1}^{p_2} C_{h_1}C_{h_2}\eta_{t,i,j,h_1,h_2}\right)^2 = o_{a.s.}\left(p_1^2 T p_2^3\left(\ln T\ln^2 p_1\ln^2 p_2\right)^{1+\epsilon}\right),$$

so that

$$\left\|\widetilde{M}_{1,4,1,2}\right\|_F = o_{a.s.}\left(\frac{1}{\sqrt{p_2 T}}\left(\ln T\ln^2 p_1\ln^2 p_2\right)^{1/2+\epsilon}\right),$$

whence

$$\left\|\widetilde{M}_{1,4,1}\right\|_F = O\left(\frac{1}{p_1 p_2}\right) + o_{a.s.}\left(\frac{1}{\sqrt{p_2 T}}\left(\ln T\ln^2 p_1\ln^2 p_2\right)^{1/2+\epsilon}\right).$$

Let now $E_{\cdot,j,t}$ represent the j-th column of E_t, and $\left(\widehat{C}-C\right)_j$ be the j-th element of $\left(\widehat{C}-C\right)$; we have

$$\begin{aligned}
\left\|\widetilde{M}_{1,4,3}\right\|_2 &= \left\|\frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_t\left(\widehat{C}-C\right)C'E_t'\right\|_2 \\
&\leq \left\|\frac{1}{p_1 p_2^2 T}\sum_{j=1}^{p_2}\left(\widehat{C}-C\right)_j\sum_{t=1}^{T} E_{\cdot,j,t}C'E_t'\right\|_2 \\
&\leq \sum_{j=1}^{p_2}\left\|\frac{1}{p_1 p_2^2 T}\left(\widehat{C}-C\right)_j\sum_{t=1}^{T} E_{\cdot,j,t}C'E_t'\right\|_2 \\
&\leq \sum_{j=1}^{p_2}\left|\left(\widehat{C}-C\right)_j\right|\left\|\frac{1}{p_1 p_2^2 T}\sum_{t=1}^{T} E_{\cdot,j,t}C'E_t'\right\|_2.
\end{aligned}$$

Thus, by the Cauchy-Schwartz inequality

$$\left\|\widetilde{M}_{1,4,3}\right\|_2 \leq \left(\sum_{j=1}^{p_2}\left|\left(\widehat{C}-C\right)_j\right|^2\right)^{1/2}\left(\sum_{j=1}^{p_2}\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_{\cdot,j,t}C'E_t'\right\|_2^2\right)^{1/2}.$$

We know from Lemma C.1 that

$$\left(\sum_{j=1}^{p_2}\left|\left(\widehat{C}-C\right)_j\right|^2\right)^{1/2} = \left\|\widehat{C}-C\right\|_F = o_{a.s.}\left(p_2^{1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)l'_{p_1,p_2,T}\right).$$

Further

$$\begin{aligned}
&\sum_{j=1}^{p_2}E\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_{\cdot,j,t}C'E_t'\right\|_2^2\\
=\;&\sum_{j=1}^{p_2}\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_2^2+\sum_{j=1}^{p_2}E\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\left(E_{\cdot,j,t}C'E_t'-E\left(E_{\cdot,j,t}C'E_t'\right)\right)\right\|_2^2;
\end{aligned}$$

we have

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_2^2 \leq \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_1\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_\infty,$$

and

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_1 \leq \frac{1}{p_1p_2^2T}\max_{1\leq i\leq p_2}|C_i|\max_{1\leq k\leq p_2}\sum_{h=1}^{p_1}\sum_{i=1}^{p_2}\sum_{t=1}^{T}\left|E\left(e_{hj,t}e_{ki,t}\right)\right| \leq c_0\frac{1}{p_1p_2^2},$$

by Assumption B3*(ii)*, so that

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_1 \leq c_0\frac{1}{p_1p_2^2};$$

using the same logic, it follows that

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_\infty \leq c_0\frac{1}{p_1p_2^2}.$$

Hence

$$\sum_{j=1}^{p_2}\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E\left(E_{\cdot,j,t}C'E_t'\right)\right\|_2^2 \leq c_0p_2\left(\frac{1}{p_1p_2^2}\right)^2.$$

Also

$$
\begin{aligned}
& E\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\left(E_{\cdot,j,t}C'E_t'-E\left(E_{\cdot,j,t}C'E_t'\right)\right)\right\|_2^2 \\
\leq & \left(\frac{1}{p_1p_2^2T}\right)^2 E\sum_{h,k=1}^{p_1}\left|\sum_{t=1}^{T}\sum_{i=1}^{p_2}\left(e_{hj,t}C_ie_{ki,t}-E\left(e_{hj,t}C_ie_{ki,t}\right)\right)\right|^2 \\
\leq & \left(\frac{1}{p_1p_2^2T}\right)^2 \sum_{h,k=1}^{p_1}\sum_{i,l=1}^{p_2}\sum_{t,s=1}^{T}\left|E\left(\eta_{h,j,k,i,t}\eta_{h,j,k,l,s}\right)\right|,
\end{aligned}
$$

with $\eta_{h,j,k,i,t}=\left(C_ie_{hj,t}e_{ki,t}-E\left(C_ie_{hj,t}e_{ki,t}\right)\right)$. Assumption B3*(iii)* entails that

$$
\sum_{j=1}^{p_2}E\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\left(E_{\cdot,j,t}C'E_t'-E\left(E_{\cdot,j,t}C'E_t'\right)\right)\right\|_2^2\leq c_0\frac{1}{p_2^2T}.
$$

Putting all together, it follows that

$$
E\left\|\widetilde{M}_{1,4,3}\right\|_2\leq c_0p_2^{1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(\frac{1}{Tp_2^2}+\frac{1}{p_1^2p_2^3}\right)^{1/2},
$$

so that Lemma C.1 yields

$$
\begin{aligned}
& \left\|\widetilde{M}_{1,4,3}\right\|_2 \\
= & o_{a.s.}\left(p_2^{1/2}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)\left(\frac{1}{Tp_2^2}+\frac{1}{p_1^2p_2^3}\right)^{1/2}\left(\ln T\ln^2 p_1\ln^2 p_2\right)^{1/2+\epsilon}l_{p_1,p_2,T}'\right),
\end{aligned}
$$

for all $\epsilon>0$. The same rate, by symmetry, is found for $\left\|\widetilde{M}_{1,4,2}\right\|_2$. Finally, note that

$$
\begin{aligned}
\left\{\widetilde{M}_{1,4,4}\right\}_{hk} &= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\left(\sum_{j=1}^{p_2}e_{hj,t}\left(\widehat{C}-C\right)_j\right)\left(\sum_{j=1}^{p_2}\left(\widehat{C}-C\right)_je_{kj,t}\right) \\
&= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\sum_{m,n=1}^{p_2}e_{hm,t}e_{kn,t}\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n \\
&= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\sum_{m,n=1}^{p_2}\left(e_{hm,t}e_{kn,t}-E\left(e_{hm,t}e_{kn,t}\right)\right)\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n \\
&\quad +\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\sum_{m,n=1}^{p_2}E\left(e_{hm,t}e_{kn,t}\right)\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n.
\end{aligned}
$$

Note now that, using Assumption B3*(ii)*

$$
\begin{aligned}
& \max_{1\leq k\leq p_1}\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}\sum_{h=1}^{p_1}\sum_{m,n=1}^{p_2}\left|E\left(e_{hm,t}e_{kn,t}\right)\right|\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n \\
\leq\ & c_0\frac{1}{p_1p_2^2}\sum_{m,n=1}^{p_2}\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n = c_0\frac{1}{p_1p_2^2}\left(\sum_{m=1}^{p_2}\left(\widehat{C}-C\right)_m\right)^2 \\
\leq\ & c_0\frac{1}{p_1p_2}\sum_{m=1}^{p_2}\left(\widehat{C}-C\right)_m^2 = c_0\frac{1}{p_1p_2}\left\|\widehat{C}-C\right\|_F^2,
\end{aligned}
$$

which, by Lemma C.2, yields that this term is bounded by $o_{a.s.}\left(\frac{1}{p_1}\left(\frac{1}{p_2^2}+\frac{1}{Tp_1}\right)l_{p_1,p_2,T}^{\prime 2}\right)$. Finally, letting $\eta_{h,m,k,n,t}=e_{hm,t}e_{kn,t}-E\left(e_{hm,t}e_{kn,t}\right)$, we have

$$
\begin{aligned}
& \frac{1}{p_1p_2^2T}\sum_{m,n=1}^{p_2}\left(\widehat{C}-C\right)_m\left(\widehat{C}-C\right)_n\left(\sum_{t=1}^{T}\eta_{h,m,k,n,t}\right) \\
\leq\ & \frac{1}{p_1p_2^2T}\left(\sum_{m,n=1}^{p_2}\left(\widehat{C}-C\right)_m^2\left(\widehat{C}-C\right)_n^2\right)^{1/2}\left(\sum_{m,n=1}^{p_2}\left(\sum_{t=1}^{T}\eta_{h,m,k,n,t}\right)^2\right)^{1/2}.
\end{aligned}
$$

It is easy to see that $E\sum_{m,n=1}^{p_2}\left(\sum_{t=1}^{T}\eta_{h,m,k,n,t}\right)^2\leq c_0p_2^2T$, so that, by using Lemma C.1

$$
\left(\sum_{m,n=1}^{p_2}\left(\sum_{t=1}^{T}\eta_{h,m,k,n,t}\right)^2\right)^{1/2}=o_{a.s.}\left(p_2T^{1/2}\left(\ln T\ln^2 p_2\right)^{1/2+\epsilon}\right),
$$

for all $\epsilon>0$. Also

$$
\left(\sum_{m,n=1}^{p_2}\left(\widehat{C}-C\right)_m^2\left(\widehat{C}-C\right)_n^2\right)^{1/2}=\sum_{m=1}^{p_2}\left(\widehat{C}-C\right)_m^2=\left\|\widehat{C}-C\right\|_F^2.
$$

Using Lemma C.2 and putting all together, it follows that

$$
\left\|\widetilde{M}_{1,4,4}\right\|_2=o_{a.s.}\left(\left(\frac{1}{p_1}+\frac{1}{\sqrt{T}}\right)\left(\frac{1}{p_2^2}+\frac{1}{Tp_1}\right)\left(\ln T\ln^2 p_2\right)^{1/2+\epsilon}l_{p_1,p_2,T}^{\prime 2}\right),
$$

for all $\epsilon>0$. The proof of the final result is now complete.

We complete the proof by considering the case $k_1=0$; recall that we have $k_2>0$, and we will set $k_2=1$ henceforth. Arguments are rather repetitive, and therefore we only report the main parts of the proof. We begin by noting that

$$
\begin{aligned}
\widetilde{M}_{1,1} &= \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'CF_t' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'\left(\widehat{C}-C\right)\widehat{C}'CF_t' \qquad \text{(D.10)}\\
&+\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'\widehat{C}\left(\widehat{C}-C\right)'CF_t' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'\left(\widehat{C}-C\right)\left(\widehat{C}-C\right)'CF_t'\\
&= \widetilde{M}_{1,1,1}+\widetilde{M}_{1,1,2}+\widetilde{M}_{1,1,3}+\widetilde{M}_{1,1,4}.
\end{aligned}
$$

It holds that

$$
\lambda_{\max}\left(\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'CF_t'\right) = \lambda_{\max}\left(\frac{1}{p_1T}\sum_{t=1}^{T}F_tF_t'\right) = O_{a.s.}\left(\frac{1}{p_1}\left(1+\sqrt{\frac{p_1}{T}}\right)^2\right),
$$

by Assumption B1*(iv)*(a). Using exactly the same logic as above, it can be verified that all the other terms in the expansion of $\widetilde{M}_{1,1}$ are dominated by this one, so that finally

$$
\lambda_{\max}\left(\widetilde{M}_{1,1}\right) = O_{a.s.}\left(\frac{1}{p_1}\left(1+\sqrt{\frac{p_1}{T}}\right)^2\right).
$$

Also

$$
\begin{aligned}
\left\|\widetilde{M}_{1,2}\right\|_F &\leq \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'E_t'\right\|_F + \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'\left(\widehat{C}-C\right)C'E_t'\right\|_F\\
&+\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'C\left(\widehat{C}-C\right)'E_t'\right\|_F + \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'\left(\widehat{C}-C\right)\left(\widehat{C}-C\right)'E_t'\right\|_F\\
&= \left\|\widetilde{M}_{1,2,1}\right\|_F+\left\|\widetilde{M}_{1,2,2}\right\|_F+\left\|\widetilde{M}_{1,2,3}\right\|_F+\left\|\widetilde{M}_{1,2,4}\right\|_F.
\end{aligned}
$$

$$
\left\|\widetilde{M}_{1,2,1}\right\|_F = \left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'E_t'\right\|_F = \frac{1}{p_1p_2T}\left(\sum_{i,j=1}^{p_1}\left|\sum_{t=1}^{T}F_{i,t}\sum_{h=1}^{p_2}C_he_{jh,t}\right|^2\right)^{1/2},
$$

and we already know that

$$
E\sum_{i,j=1}^{p_1}\left|\sum_{t=1}^{T}F_{i,t}\sum_{h=1}^{p_2}C_he_{jh,t}\right|^2 \leq c_0p_1^2p_2T,
$$

so that

$$
\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'E_t'\right\|_F = o_{a.s.}\left(\frac{1}{\sqrt{p_1T}}\left(\ln T\ln^2 p_1\ln p_2\right)^{1/2+\epsilon}\right).
$$

Exactly the same logic as above yields

$$
\left\|\widetilde{M}_{1,2,2}\right\|_F = o_{a.s.}\left(\frac{1}{p_2^{1/2}T^{1/2}}\left(\frac{1}{p_2}+\frac{1}{\sqrt{Tp_1}}\right)l'_{p_1,p_2,T}\left(\ln T\ln^2 p_1\ln p_2\right)^{1/2+\epsilon}\right).
$$

Also, repeating the proof of Lemma B.3 in Yu et al. (2022), it can be show that

$$E\left\|\frac{1}{p_2T}\sum_{t=1}^{T}F_t\left(\widehat{C}-C\right)'E_t\right\|_F^2\leq c_0p_1^2\left(\frac{1}{Tp_2^2}+\frac{1}{\left(Tp_1\right)^2}\right),$$

whence

$$\left\|\widetilde{M}_{1,2,3}\right\|_F=o_{a.s.}\left(\left(\frac{1}{T^{1/2}p_2}+\frac{1}{Tp_1}\right)\left(\ln p_1\ln p_2\ln T\right)^{1/2+\epsilon}\right),$$

as before. The same rate as above can also be verified for $\left\|\widetilde{M}_{1,2,4}\right\|_F$. The same passages can be applied to $\left\|\widetilde{M}_{1,3}\right\|_F$; similarly, $\left\|\widetilde{M}_{1,4}\right\|_F$ has the same magnitude as in the case $k_1>0$. The desired result now follows from putting everything together.

Proof of Proposition 1. We report the proof based on $\widehat{\Psi}_{k_1^0}$ only; the case of $\widetilde{\Psi}_{k_1^0}$ follows exactly from the same arguments. Theorem 2 entails that, under the null and for all $0<\epsilon<1/\overline{\lambda}$

$$P\left(\omega:\lim_{\min(p_1,p_2,T)\to\infty}\exp\left(-\epsilon p_1^{1-\delta}\right)\widehat{\phi}_{k_1^0}=\infty\right)=1, \tag{D.11}$$

and therefore we can assume from now on that

$$\lim_{\min(p_1,p_2,T)\to\infty}\exp\left(-\epsilon p_1^{1-\delta}\right)\widehat{\phi}_{k_1^0}=\infty.$$

Consider, for simplicity, the case $u\geq0$. It holds that

$$\begin{aligned}
&M^{-1/2}\sum_{m=1}^{M}\left[\widehat{\psi}_{k_1^0}^{(m)}(u)-\frac{1}{2}\right]\\
=\;&M^{-1/2}\sum_{m=1}^{M}\left[\widehat{\psi}_{k_1^0}^{(m)}(0)-\frac{1}{2}\right]+M^{-1/2}\sum_{m=1}^{M}\left[\left(\widehat{\psi}_{k_1^0}^{(m)}(u)-\widehat{\psi}_{k_1^0}^{(m)}(0)\right)-\left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right)\right]\\
&+M^{-1/2}\sum_{m=1}^{M}\left[G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right]
\end{aligned}$$

where $G(\cdot)$ denotes the standard normal distribution. By definition, it holds that

$$\begin{aligned}
E^*\left(\widehat{\psi}_{k_1^0}^{(1)}(u)\right)&=G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right),\\
V^*\left(\widehat{\psi}_{k_1^0}^{(1)}(u)\right)&=G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\left(1-G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right).
\end{aligned}$$

It holds that

$$
\begin{aligned}
& E^{*}\int_{-\infty}^{\infty}\left[M^{-1/2}\sum_{m=1}^{M}\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)}\leq u\right]-\left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right)\right)\right]^{2}dF\left(u\right)\\
= & \int_{-\infty}^{\infty}E^{*}\left[\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)}\leq u\right]-\left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right)\right)\right]^{2}dF\left(u\right),
\end{aligned}
$$

where

$$
\begin{aligned}
E^{*}\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)}\leq u\right]\right) & = G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\\
V^{*}\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)}\leq u\right]\right) & = \left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right)\left(\frac{3}{2}-G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right)\leq G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
& \int_{-\infty}^{\infty}E^{*}\left[\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)}\leq u\right]-\left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right)\right)\right]^{2}dF\left(u\right)\\
\leq & \int_{-\infty}^{\infty}\left|E^{*}\left(I\left[0\leq\widehat{\phi}_{k_1^0}^{1/2}\eta^{(1)}\leq u\right]\right)\right|dF\left(u\right)\\
\leq & \frac{1}{\sqrt{2\pi}}\widehat{\phi}_{k_1^0}^{-1/2}\int_{-\infty}^{\infty}\left|u\right|dF\left(u\right)=o_{P^{*}}\left(1\right),
\end{aligned}
$$

having used (D.11). Also

$$
\begin{aligned}
& \int_{-\infty}^{\infty}\left(M^{-1/2}\sum_{m=1}^{M}\left[G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)-\frac{1}{2}\right]\right)^{2}dF\left(u\right)\\
\leq & \frac{M}{2\pi\widehat{\phi}_{k_1^0}}\int_{-\infty}^{\infty}u^{2}dF\left(u\right)\leq c_{0}\frac{M}{\widehat{\phi}_{k_1^0}}=o_{P^{*}}\left(1\right),
\end{aligned}
$$

where the last passage follows from (3.6). Putting all together

$$
\begin{aligned}
& \int_{-\infty}^{\infty}\left(M^{-1/2}\sum_{m=1}^{M}\left[\widehat{\psi}_{k_1^0}^{(m)}\left(u\right)-\frac{1}{2}\right]\right)^{2}dF\left(u\right)\\
= & \int_{-\infty}^{\infty}\left(M^{-1/2}\sum_{m=1}^{M}\left[\widehat{\psi}_{k_1^0}^{(m)}\left(0\right)-\frac{1}{2}\right]\right)^{2}dF\left(u\right)+o_{P^{*}}\left(1\right),
\end{aligned}
$$

and now equation (3.7) follows from the CLT for Bernoulli random variables. We now turn to (3.8). As in the proof of the previous result, we focus on $\widehat{\Psi}_{k_1^0}$ only. By Theorem 2, under the alternative we have

$$P\left(\omega : \lim_{\min(p_1,p_2,T)\to\infty} \widehat{\phi}_{k_1^0} = 0\right) = 1, \tag{D.12}$$

and therefore we can assume from now on that

$$\lim_{\min(p_1,p_2,T)\to\infty} \widehat{\phi}_{k_1^0} = 0. \tag{D.13}$$

We begin by noting that

$$E^* \int_{-\infty}^{\infty} \left[M^{-1/2} \sum_{m=1}^{M} \left(I\left(\widehat{\phi}_{k_1^0}^{1/2}\eta^{(m)} \leq u\right) - G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right)\right]^2 dF(u) \leq c_0. \tag{D.14}$$

Note now that

$$\begin{aligned}
& \int_{-\infty}^{\infty} \left(M^{-1/2} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0}^{(m)}(u) - G(0)\right]\right)^2 dF(u) \\
= & \int_{-\infty}^{\infty} \left(M^{-1/2} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0}^{(m)}(u) - G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right]\right)^2 dF(u) + M \int_{-\infty}^{\infty} \left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right) - \frac{1}{2}\right)^2 dF(u) \\
& +2 \int_{-\infty}^{\infty} M^{1/2} \left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right) - \frac{1}{2}\right) \left(M^{-1/2} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0}^{(m)}(u) - G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right]\right) dF(u).
\end{aligned}$$

Using (D.14) and Markov inequality, it holds that

$$\int_{-\infty}^{\infty} \left(M^{-1/2} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0}^{(m)}(u) - G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right]\right)^2 dF(u) = O_{P^*}(1);$$

also, using (D.13), it follows immediately that

$$\lim_{\min(p_1,p_2,T,M)\to\infty} M \int_{-\infty}^{\infty} \left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right) - \frac{1}{2}\right)^2 dF(u) = \infty.$$

By applying the Cauchy-Schwartz inequality, it also holds that

$$\int_{-\infty}^{\infty} M^{1/2} \left(G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right) - \frac{1}{2}\right) \left(M^{-1/2} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0}^{(m)}(u) - G\left(\frac{u}{\widehat{\phi}_{k_1^0}^{1/2}}\right)\right]\right) dF(u) = O_{P^*}\left(M^{1/2}\right).$$

The desired result now follows. $\square$

Proof of Theorem 3. We begin by noting that $E^* I\left[\widehat{\Psi}_{k_1^0,s} \leq c_\alpha\right] = P^*\left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha\right)$. From the proof of Theorem 2, we know that

$$\widehat{\Psi}_{k_1^0,s} = X_{M,s} + Y_{M,s},$$

where

$$X_{M,s} = \left(\frac{2}{\sqrt{M}} \sum_{m=1}^{M} \left[\widehat{\psi}_{k_1^0,s}^{(m)}(0) - \frac{1}{2}\right]\right)^2,$$

and the remainder $Y_{M,s}$ is such that $E^* Y_{M,s}^2 = c_0\left(\widehat{\phi}_{k_1^0}^{-1/2} + M\widehat{\phi}_{k_1^0}^{-1}\right)$ - henceforth, we omit the dependence on s when possible. Let now $\epsilon_\phi = \left(\widehat{\phi}_{k_1^0}^{-1/2} + M\widehat{\phi}_{k_1^0}^{-1}\right)^{1/3}$, and note that, using elementary arguments

$$P^*\left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha\right) \leq P^*\left(X_{M,s} \leq c_\alpha + \epsilon_\phi\right) + P^*\left(|Y_{M,s}| \geq \epsilon_\phi\right).$$

Thus

$$\begin{aligned} & P^*\left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha\right) - P^*\left(X_{M,s} \leq c_\alpha\right) \\ \leq \quad & P^*\left(X_{M,s} \leq c_\alpha + \epsilon_\phi\right) - P^*\left(X_{M,s} \leq c_\alpha\right) + P^*\left(|Y_{M,s}| \geq \epsilon_\phi\right). \end{aligned}$$

Markov inequality immediately yields $P^*\left(|Y_{M,s}| \geq \epsilon_\phi\right) \leq \epsilon_\phi^{-2} E^* Y_{M,s}^2 = c_0\left(\widehat{\phi}_{k_1^0}^{-1/2} + M\widehat{\phi}_{k_1^0}^{-1}\right)^{1/3}$. Letting Z be a $N(0,1)$ distributed random variable, we can write

$$\begin{aligned} & P^*\left(X_{M,s} \leq c_\alpha + \epsilon_\phi\right) - P^*\left(X_{M,s} \leq c_\alpha\right) \\ = \quad & P^*\left(X_{M,s} \leq c_\alpha + \epsilon_\phi\right) \pm P\left(Z^2 \leq c_\alpha + \epsilon_\phi\right) - \left[P^*\left(X_{M,s} \leq c_\alpha\right) \pm P\left(Z^2 \leq c_\alpha\right)\right]. \end{aligned}$$

Using the mean value theorem, it is easy to see that

$$P\left(Z^2 \leq c_\alpha + \epsilon_\phi\right) - P\left(Z^2 \leq c_\alpha\right) \leq c_0 \epsilon_\phi.$$

Also, by the Berry-Esseen theorem (see e.g. Michel, 1976) it holds that

$$\begin{aligned} \left|P^*\left(X_{M,s} \leq c_\alpha + \epsilon_\phi\right) - P\left(Z^2 \leq c_\alpha + \epsilon_\phi\right)\right| & \leq c_0 \frac{M^{-1/2}}{1 + |c_\alpha + \epsilon_\phi|^{3+\delta}}, \\ \left|P^*\left(X_{M,s} \leq c_\alpha\right) - P\left(Z^2 \leq c_\alpha\right)\right| & \leq c_0 \frac{M^{-1/2}}{1 + |c_\alpha|^{3+\delta}}, \end{aligned}$$

for all $\delta \geq 0$. Putting all together, it follows that

$$\left| P^* \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) - P \left(Z^2 \leq c_\alpha \right) \right| \leq c_0 \left(M^{-1/2} + \epsilon_\phi \right).$$

Hence we have

$$\sqrt{\frac{S}{2 \ln \ln S}} \frac{\widehat{Q}_{k_1^0}(\alpha) - (1-\alpha)}{\sqrt{\alpha(1-\alpha)}} = \sqrt{\frac{S}{2 \ln \ln S}} \frac{S^{-1} \sum_{s=1}^{S} Z_s^{(\alpha)} - (1-\alpha)}{\sqrt{\alpha(1-\alpha)}} + c_0 \sqrt{\frac{S}{2 \ln \ln S}} \left(M^{-1/2} + \epsilon_\phi \right),$$

where $\left\{ Z_s^{(\alpha)}, 1 \leq s \leq S \right\}$ is an *i.i.d.* sequence with common distribution $Z_1^{(\alpha)} \sim N(1-\alpha, \alpha(1-\alpha))$. By the fact that $S = O(M)$, it follows that

$$\sqrt{\frac{S}{2 \ln \ln S}} \left(M^{-1/2} + \epsilon \right) = o_{a.s.}(1);$$

the desired result now follows from the Law of the Iterated Logarithm. Under H_A, the results in the proof of Proposition 1 show that $P^* \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \leq \epsilon$, for all $\epsilon > 0$ and for almost all realisations of $\{X_t, 1 \leq t \leq T\}$. Hence

$$\begin{aligned}
\widehat{Q}_{k_1^0}(\alpha) &= \frac{1}{S} \sum_{s=1}^{S} I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \\
&= \frac{1}{S} \sum_{s=1}^{S} E^* I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) + \frac{1}{S} \sum_{s=1}^{S} \left(I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) - E^* I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \right).
\end{aligned}$$

Note that

$$\begin{aligned}
& V^* \left(\frac{1}{S} \sum_{s=1}^{S} \left(I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) - E^* I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \right) \right) \\
= & \frac{1}{S^2} \sum_{s=1}^{S} V^* I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \leq \frac{1}{S^2} \sum_{s=1}^{S} E^* I \left(\widehat{\Psi}_{k_1^0,s} \leq c_\alpha \right) \\
\leq & S^{-1} P^* \left(\widehat{\Psi}_{k_1^0,1} \leq c_\alpha \right),
\end{aligned}$$

for almost all realisations of $\{X_t, 1 \leq t \leq T\}$. Using Lemma C.1, it follows that

$$\widehat{Q}_{k_1^0}(\alpha) \leq c_0 + c_1 \frac{(\ln S)^{1+\epsilon}}{S^{1/2}},$$

for all $0 < c_0, c_1 < \infty$ and all $\epsilon > 0$, and for almost all realisations of $\{X_t, 1 \leq t \leq T\}$. This shows the desired result.

□

Proof of Theorem 4. The proof is very similar to that of Theorem 3.3 in Trapani (2018), and we only sketch its main passages. We begin by noting that, when $k_1 = 0$, there is nothing to prove as the desired result is already contained in Theorem 3. When $k_1 > 0$, consider the events $\left\{\widehat{k}_1 = j\right\}$, for $0 \leq j \leq k_1 - 1$. By the independence of tests across j (conditional on the sample), it follows that

$$\begin{aligned} P^*\left(\widehat{k}_1 = j\right) &= P^*\left(\widehat{Q}_j(\alpha) < 1 - \alpha - \sqrt{\alpha(1-\alpha)}\sqrt{\tfrac{2\ln\ln S}{S}}\right) \times \\ & \prod_{h=1}^{j-1} P^*\left(\widehat{Q}_h(\alpha) \geq 1 - \alpha - \sqrt{\alpha(1-\alpha)}\sqrt{\tfrac{2\ln\ln S}{S}}\right). \end{aligned}$$

Note that equation (3.13) immediately implies that $P^*\left(\widehat{k}_1 = j\right) = 0$ as $\min\{p_1, p_2, T, M, S\} \to \infty$. Similarly, considering the events $\left\{\widehat{k}_1 = k_1 + j\right\}$, for $1 \leq j \leq k_{\max} - k_1$, it holds that

$$\begin{aligned} & P^*\left(\widehat{k}_1 = k_1 + j\right) \\ = & \left(\prod_{h=1}^{k_1} P^*\left(\widehat{Q}_h(\alpha) \geq 1 - \alpha - \sqrt{\alpha(1-\alpha)}\sqrt{\tfrac{2\ln\ln S}{S}}\right)\right) \times \\ & \left(\prod_{h=k_1+1}^{k_1+j} P^*\left(\widehat{Q}_h(\alpha) \geq 1 - \alpha - \sqrt{\alpha(1-\alpha)}\sqrt{\tfrac{2\ln\ln S}{S}}\right)\right) \times \\ & P^*\left(\widehat{Q}_{k_1+j+1}(\alpha) < 1 - \alpha - \sqrt{\alpha(1-\alpha)}\sqrt{\tfrac{2\ln\ln S}{S}}\right), \end{aligned}$$

for all $j = 1, ..., k_{\max} - k_1$; by equation (3.14), it follows that $P^*\left(\widehat{k}_1 = k_1 + j\right) = 0$ as $\min\{p_1, p_2, T, M, S\} \to \infty$. Putting all together, it is easy to see that $P^*\left(\widehat{k}_1 = k_1\right)$. $\square$

Proof of Corollary 1. The proof is very similar to an argument reported in Footnote 5 in Bai (2003). Recall that we are using an estimator of k_2, $\widehat{k}_2$, such that as $\min\{p_1, p_2, T, M, S\} \to \infty$

$$P^*\left(\widehat{k}_2 \neq k_2\right) = 0, \tag{D.15}$$

for almost all realisations of $\{X_t, 1 \leq t \leq T\}$. Consider now

$$P^*\left(\widehat{k}_1 = k_1\right) = P^*\left(\widehat{k}_1 = k_1, \widehat{k}_2 = k_2\right) + P^*\left(\widehat{k}_1 = k_1, \widehat{k}_2 \neq k_2\right),$$

and note that

$$\begin{aligned} P^*\left(\widehat{k}_1 = k_1, \widehat{k}_2 \neq k_2\right) &= P^*\left(\widehat{k}_1 = k_1 | \widehat{k}_2 \neq k_2\right) P^*\left(\widehat{k}_2 \neq k_2\right) \\ &\leq P^*\left(\widehat{k}_2 \neq k_2\right), \end{aligned}$$

almost surely. Finally, using (D.15) and dominated convergence, it follows that

$$P^*\left(\widehat{k}_1 = k_1\right) = P^*\left(\widehat{k}_1 = k_1, \widehat{k}_2 = k_2\right) + o_{P^*}(1),$$

for almost all realisations of $\{X_t, 1 \leq t \leq T\}$. This passage states that estimating $\widehat{k}_1$ using $\widehat{k}_2$ is the same as if one were to use k_2 directly. The desired result now follows immediately, and the same arguments apply *verbatim* to $\widetilde{k}_1$. □

Proof of Corollary 2. The proof is essentially the same as that of Theorem 2, upon noting that, for $j \leq k_1$

$$\lambda_j\left(\frac{1}{p_2 T}\sum_{t=1}^{T} R F_t C' C F_t' R'\right) \geq \lambda_j\left(R'R\right)\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T} F_t \frac{C'C}{p_2} F_t'\right) = p_2^{\alpha_2 - 1}\lambda_j\left(R'R\right)\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right),$$

and noting that, by Assumption B6, $\lambda_j\left(R'R\right) \geq c_0 p_1^{\alpha_1}$. Thus, it holds that

$$\lambda_j\left(\frac{1}{p_2 T}\sum_{t=1}^{T} R F_t C' C F_t' R'\right) = \Omega\left(p_1^{\alpha_1} p_2^{\alpha_2 - 1}\right), \tag{D.16}$$

whenever $j \leq k_1$. Condition (3.20) ensures that, in this case, the relevant eigenvalue diverges.

When $j > k_1$, the term above is absent and the proof of Theorem 2 can be followed, showing in particular that

$$\lambda_{\max}\left(\frac{1}{p_2 T}\sum_{t=1}^{T} E_t E_t'\right) = o_{a.s.}\left(\frac{p_1}{\sqrt{T p_2}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right). \tag{D.17}$$

Similarly, a (crude) bound for the cross-product terms is given by

$$\left\|\frac{1}{p_2 T}\sum_{t=1}^{T} R F_t C' E_t'\right\|_F \leq \frac{1}{p_2 T}\left\|R\right\|_F \left\|\sum_{t=1}^{T} F_t C' E_t'\right\|_F, \tag{D.18}$$

and

$$\begin{aligned}
E\left\|\sum_{t=1}^{T} F_t C' E_t'\right\|_F^2 &= E\sum_{i=1}^{p_1}\left|\sum_{t=1}^{T}\sum_{j=1}^{p_2} C_j e_{ij,t} F_t\right|^2 \\
&\leq \max_{1\leq j\leq p_2}\left|C_j\right|^2 \sum_{i=1}^{p_1}\sum_{t,s=1}^{T}\sum_{j=1}^{p_2}\left|E\left(e_{ij,t} e_{ih,s} F_t F_s\right)\right| \\
&\leq c_0 p_1 p_2 T,
\end{aligned}$$

using Assumption B4*(ii)* and the (non-optimal) fact that $\max_{1\leq j\leq p_2}\left|C_j\right|^2 < \infty$. Hence, using (D.18) and

Assumption B6

$$\left\| \frac{1}{p_2 T} \sum_{t=1}^{T} R F_t C' E_t' \right\|_F = o_{a.s.} \left(\frac{p_1^{1/2+\alpha_1/2}}{\sqrt{T p_2}} \left(\ln^2 p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right),$$

which is smaller than the rate in (D.17).

Thus, putting all together, it holds that

$$\widehat{\lambda}_j = \Omega_{a.s.} \left(p_1^{\alpha_1} p_2^{\alpha_2 - 1} \right),$$

for all $j \leq k_1$, and

$$\widehat{\lambda}_j = c_0 + o_{a.s.} \left(\frac{p_1}{\sqrt{T p_2}} \left(\ln^2 p_1 \ln p_2 \ln T \right)^{1/2+\epsilon} \right).$$

Henceforth, the proof is the same as above. □

Proof of Lemma A.1. For simplicity, we let $E\left(F_{h_1 h_2, t} F_{l_1 l_2, t}\right) = 0$, and we define the (i, j)-th component of g^F as g_{ij}^F. We have

$$\begin{aligned} F_{h_1 h_2, t} F_{l_1 l_2, t} &= g_{h_1 h_2}^F \left(\eta_t, ..., \eta_0, ..., \eta_{-\infty} \right) g_{l_1 l_2}^F \left(\eta_t, ..., \eta_0, ..., \eta_{-\infty} \right) \\ &= h \left(\eta_t, ..., \eta_0, ..., \eta_{-\infty} \right). \end{aligned}$$

Define now

$$\begin{aligned} F_{h_1 h_2, t}' &= g_{h_1 h_2}^F \left(\eta_t, ..., \eta_0', ..., \eta_{-\infty} \right) \\ F_{l_1 l_2, t}' &= g_{l_1 l_2}^F \left(\eta_t, ..., \eta_0', ..., \eta_{-\infty} \right) \end{aligned}$$

and note that

$$F_{h_1 h_2, t}' F_{l_1 l_2, t}' = h \left(\eta_t, ..., \eta_0', ..., \eta_{-\infty} \right).$$

It is easy to see that

$$
\begin{aligned}
& \left|F_{h_1h_2,t}F_{l_1l_2,t}-F'_{h_1h_2,t}F'_{l_1l_2,t}\right|_{p/2} \\
\leq\ & \left|F'_{h_1h_2,t}\left(F_{l_1l_2,t}-F'_{l_1l_2,t}\right)\right|_{p/2}+\left|F'_{l_1l_2,t}\left(F_{h_1h_2,t}-F'_{h_1h_2,t}\right)\right|_{p/2} \\
& +\left|\left(F_{h_1h_2,t}-F'_{h_1h_2,t}\right)\left(F_{l_1l_2,t}-F'_{l_1l_2,t}\right)\right|_{p/2} \\
\leq\ & \left|F'_{h_1h_2,t}\right|_p\left|F_{l_1l_2,t}-F'_{l_1l_2,t}\right|_p+\left|F'_{l_1l_2,t}\right|_p\left|F_{h_1h_2,t}-F'_{h_1h_2,t}\right|_p \\
& +\left|F_{h_1h_2,t}-F'_{h_1h_2,t}\right|_p\left|F_{l_1l_2,t}-F'_{l_1l_2,t}\right|_p \\
\leq\ & c_0\left|F_{l_1l_2,t}-F'_{l_1l_2,t}\right|_p+c_1\left|F_{h_1h_2,t}-F'_{h_1h_2,t}\right|_p \\
& +\left|F_{h_1h_2,t}-F'_{h_1h_2,t}\right|_p\left|F_{l_1l_2,t}-F'_{l_1l_2,t}\right|_p \\
\leq\ & c_0\delta^F_{t,p}.
\end{aligned}
$$

Lemma A.2 in Liu and Lin (2009) entails that

$$
E\left|\sum_{t=1}^{T}\left(F_{h_1h_2,t}F_{l_1l_2,t}-E\left(F_{h_1h_2,t}F_{l_1l_2,t}\right)\right)\right|^{p/2}\leq c_0T^{p/4}.
$$

Noting that $p>4$, Theorem 1 in Móricz (1976) immediately yields the desired result (see also Corollary 1 in Berkes et al., 2011). □

Proof of Lemma A.2. We only show (A.5), as (A.6) follows from essentially the same arguments. By stationarity

$$
\sum_{s=1}^{T}\left|E\left(e_{ij,t}e_{lh,s}\right)\right|=\left|E\left(e_{ij,0}e_{lh,0}\right)\right|+\sum_{m=1}^{T}\left|E\left(e_{ij,0}e_{lh,m}\right)\right|,
$$

and our assumptions immediately yield $\left|E\left(e_{ij,0}e_{lh,0}\right)\right|\leq\left|e_{ij,0}\right|_2\left|e_{lh,0}\right|_2<\infty$. Define now the sigma-field $\mathcal{F}^0_{ijlh,-k}=\mathcal{F}\left\{\eta_{ij,0},...,\eta_{ij,-k};\eta_{lh,0},...,\eta_{lh,-k}\right\}$, and consider

$$
\begin{aligned}
& \sum_{m=1}^{T}\left|E\left(e_{ij,0}e_{lh,m}\right)\right| \qquad\qquad (D.19)\\
\leq\ & \sum_{m=1}^{T}\left|E\left(e_{lh,m}\left(e_{ij,0}-E\left(e_{ij,0}|\mathcal{F}^0_{ijlh,-\lfloor am\rfloor}\right)\right)\right)\right| \\
& +\sum_{m=1}^{T}\left|E\left(e_{lh,m}E\left(e_{ij,0}|\mathcal{F}^0_{ijlh,-\lfloor am\rfloor}\right)\right)\right| \\
=\ & I+II,
\end{aligned}
$$

for some $0 < a < 1$. We have

$$\begin{aligned}
& \sum_{m=1}^{T} \left| E\left(e_{lh,m}\left(e_{ij,0} - E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right)\right)\right| \\
\leq & \sum_{m=1}^{T} |e_{lh,m}|_2 \left| e_{ij,0} - E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 \\
\leq & \ c_0 \sum_{m=1}^{T} \left| e_{ij,0} - E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 .
\end{aligned}$$

Consider now the projections $Q_j X_0 = E\left(X_0|\mathcal{F}^{0}_{X,-j}\right)$, for a generic process $X_t = g\left(\eta^X_t, ..., \eta^X_{-\infty}\right)$, where $\mathcal{F}^{0}_{X,-k} = \left\{\eta^X_0, ..., \eta^X_{-k}\right\}$, and define $\widetilde{Q}_j X_0 = Q_j X_0 - Q_{j-1} X_0$. Clearly, $X_0 = \lim_{j\to\infty} Q_j X_0$. We have

$$e_{ij,0} - E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right) = \sum_{k=\lfloor am \rfloor}^{\infty} \widetilde{Q}_k e_{ij,0},$$

so that

$$\begin{aligned}
& \sum_{m=1}^{T} \left| e_{ij,0} - E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 \\
= & \sum_{m=1}^{T} \left| \sum_{k=\lfloor am \rfloor}^{\infty} \widetilde{Q}_k e_{ij,0} \right|_2 \\
\leq & \sum_{m=1}^{T} \sum_{k=\lfloor am \rfloor}^{\infty} \left| \widetilde{Q}_k e_{ij,0} \right|_2 .
\end{aligned}$$

Following the proof of Theorem 1 *(iii)* in Wu (2005), it is immediate to verify that $\left|\widetilde{Q}_k e_{ij,0}\right|_2 \leq \delta^{e,ij}_{k,2}$, and therefore term I in (D.19) is bounded. Also

$$\begin{aligned}
& \left| E\left(e_{lh,m} E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right)\right| \\
= & \left| E\left(E\left(e_{lh,m}|\mathcal{F}^{\lfloor am \rfloor}_{ijlh,-\lfloor am \rfloor}\right) E\left(e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right)\right)\right| \\
\leq & \left| e_{ij,0}|\mathcal{F}^{0}_{ijlh,-\lfloor am \rfloor}\right|_2 \left| E\left(e_{lh,m}|\mathcal{F}^{\lfloor am \rfloor}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 \\
\leq & |e_{ij,0}|_2 \left| E\left(e_{lh,m}|\mathcal{F}^{\lfloor am \rfloor}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 \\
\leq & \ c_0 \left| E\left(e_{lh,m}|\mathcal{F}^{\lfloor am \rfloor}_{ijlh,-\lfloor am \rfloor}\right)\right|_2 .
\end{aligned}$$

Consider now the projections $P_0 X_k = E\left(X_k|\mathcal{F}^{0}_{X,-\infty}\right) - E\left(X_k|\mathcal{F}^{-1}_{X,-\infty}\right)$, and note

$$E\left(e_{lh,m}|\mathcal{F}^{\lfloor am \rfloor}_{ijlh,-\lfloor am \rfloor}\right) = \sum_{r=m-\lfloor am \rfloor}^{\infty} P_{m-r} e_{lh,m} = \sum_{r=m-\lfloor am \rfloor}^{\infty} P_0 e_{lh,r},$$

by stationarity. Again, using Theorem 1(i) in Wu (2005), $|P_0 e_{lh,r}|_2 \leq \delta_{r,2}^{e,lh}$, so that

$$\sum_{m=1}^{T} \left| E\left(e_{lh,m} | \mathcal{F}_{ijlh,-\lfloor am \rfloor}^{\lfloor am \rfloor} \right) \right|_2 \leq \sum_{m=1}^{T} \sum_{r=m-\lfloor am \rfloor}^{\infty} |P_0 e_{lh,r}|_2 < \infty,$$

which gives the desired result. $\square$

Proof of Corollary A.1. The proof is exactly the same as that of Theorem 1, with the exception of the term

$$\lambda_{\max}\left(\frac{1}{Tp_2}\sum_{i=1}^{p_2}\sum_{t=1}^{T} C_i^2 F_t F_t'\right) \leq c_0 \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right).$$

To study this, we begin by noting that, on account of Assumption A.1, $F_{i,t}$ is an L_4-decomposable Bernoulli shift an therefore it can be shown that

$$E\left(\sum_{t=1}^{T} F_{i,t}F_{j,t} - E\left(F_{i,t}F_{j,t}\right)\right)^2 \leq c_0 T, \tag{D.20}$$

for $1 \leq i, j \leq p_1$. Then the bound for $\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right)$ comes from

$$\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} F_t F_t'\right) \leq \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right) + \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} \left(F_t F_t' - E\left(F_t F_t'\right)\right)\right).$$

Assumption 1 readily entails

$$\lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right) \leq \frac{1}{T}\sum_{t=1}^{T} \lambda_{\max}\left(E\left(F_t F_t'\right)\right) \leq c_0.$$

Also

$$\left\| \frac{1}{T}\sum_{t=1}^{T} \left(F_t F_t' - E\left(F_t F_t'\right)\right) \right\|_F = \left(\sum_{i,j=1}^{p_1} \left| \frac{1}{T}\sum_{t=1}^{T} \left(F_{i,t}F_{j,t} - E\left(F_{i,t}F_{j,t}\right)\right) \right|^2 \right)^{1/2},$$

and, combining (D.20) with Lemma C.1, it is easy to see that

$$\left\| \frac{1}{T}\sum_{t=1}^{T} \left(F_t F_t' - E\left(F_t F_t'\right)\right) \right\|_F = o_{a.s.}\left(\frac{p_1}{\sqrt{T}} \left(\ln^2 p_1 \ln T\right)^{1/2+\epsilon} \right),$$

for all $\epsilon > 0$. Putting all together, the desired result follows. The proof of (A.9) is similar; indeed, considering $\widetilde{M}_{1,1,1}$ in (D.10)

$$\lambda_{\max}\left(\frac{1}{p_1 T}\sum_{t=1}^{T} F_t F_t'\right) \leq \lambda_{\max}\left(\frac{1}{p_1 T}\sum_{t=1}^{T} E\left(F_t F_t'\right)\right) + \lambda_{\max}\left(\frac{1}{p_1 T}\sum_{t=1}^{T} \left(F_t F_t' - E\left(F_t F_t'\right)\right)\right).$$

We know that

$$\lambda_{\max}\left(\frac{1}{p_1T}\sum_{t=1}^{T}E\left(F_tF_t'\right)\right)\leq c_0p_1^{-1}.$$

Also, consider

$$\left\|\frac{1}{p_1T}\sum_{t=1}^{T}\left(F_tF_t'-E\left(F_tF_t'\right)\right)\right\|_F=p_1^{-1}\left(\sum_{i,j=1}^{p_1}\left|\frac{1}{T}\sum_{t=1}^{T}\left(F_{i,t}F_{j,t}-E\left(F_{i,t}F_{j,t}\right)\right)\right|^2\right)^{1/2},$$

and

$$\sum_{i,j=1}^{p_1}E\left|\frac{1}{T}\sum_{t=1}^{T}\left(F_{i,t}F_{j,t}-E\left(F_{i,t}F_{j,t}\right)\right)\right|^2\leq c_0p_1^2T^{-1},$$

by virtue of Assumption B1*(iii)*. Thus, by Lemma C.1, it follows that

$$\lambda_{\max}\left(\frac{1}{p_1T}\sum_{t=1}^{T}F_tF_t'-E\left(F_tF_t'\right)\right)=o_{a.s.}\left(T^{-1/2}\left(\ln^2p_1\ln T\right)^{1/2+\epsilon}\right),$$

for all $\epsilon>0$. Thus we finally have

$$\lambda_{\max}\left(\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}F_tC'CC'CF_t'\right)=c_0p_1^{-1}+o_{a.s.}\left(T^{-1/2}\left(\ln^2p_1\ln T\right)^{1/2+\epsilon}\right).$$

Repeating now all the passages in the proof of Theorem 2, the desired result follows.

□

Proof of Lemma A.3. We will focus on studying the magnitude of

$$\begin{aligned}&\frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2}F_tF_se_{ij,s}e_{ij_1,t}\right)^2\\ \leq\quad&\frac{2}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2}F_tF_se_{ij,s}e_{ij_1,t}-E\left(F_tF_se_{ij,s}e_{ij_1,t}\right|\left\{F_t\right\}\right)\right)^2\\ &+\frac{2}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2}E\left(F_tF_se_{ij,s}e_{ij_1,t}\right|\left\{F_t\right\}\right)\right)^2.\end{aligned}$$

Under (A.11), it holds that

$$E\left(F_tF_se_{ij,s}e_{ij_1,t}\right|\left\{F_t\right\}\right)=F_tg\left(F_t\right)F_sg\left(F_s\right)E\left(u_{ij,s}u_{ij_1,t}\right),$$

and defining, for short

$$h_t=h\left(F_t\right)=F_tg\left(F_t\right),$$

we will study

$$
\begin{aligned}
& \frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2} E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} h_t h_s u_{ij,s}u_{ij_1,t} - h_t h_s E\left(u_{ij,s}u_{ij_1,t}\right)\right)^2 \\
& + \frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2} E\left(\sum_{t,s=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} h_t h_s E\left(u_{ij,s}u_{ij_1,t}\right)\right)^2 \\
= \quad & I + II.
\end{aligned}
$$

Let us consider I

$$
\begin{aligned}
& I = \frac{1}{\left(Tp_2\right)^2} E\left(\sum_{j=1}^{p_2}\sum_{t,s,u,v=1}^{T}\sum_{i,i'=1}^{p_1}\sum_{j_1,j_1'=1}^{p_2} h_t h_s h_u h_v \left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\right. \\
& \left.\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right) \\
= \quad & \frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}\sum_{t,s,u,v=1}^{T}\sum_{i,i'=1}^{p_1}\sum_{j_1,j_1'=1}^{p_2} E\left(h_t h_s h_u h_v\right) \\
& \times E\left(\left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right),
\end{aligned}
$$

by independence between F_t and $u_{ij,t}$. Now this is bounded by

$$
\begin{aligned}
& \frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}\sum_{t,s,u,v=1}^{T}\sum_{i,i'=1}^{p_1}\sum_{j_1,j_1'=1}^{p_2} \left|E\left(h_t h_s h_u h_v\right)\right| \\
& \times \left|E\left(\left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right)\right|.
\end{aligned}
$$

In light of the fact that $E\left|h\left(F_t\right)\right|^4 < \infty$, using the Cauchy-Schwartz inequality we have the bound

$$
c_0 \frac{1}{\left(Tp_2\right)^2}\sum_{j=1}^{p_2}\sum_{t,s,u,v=1}^{T}\sum_{i,i'=1}^{p_1}\sum_{j_1,j_1'=1}^{p_2} \left|E\left(\left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right)\right|.
$$

Equation (A.15) immediately entails that

$$
\sum_{t,s=1}^{T}\sum_{i,=1}^{p_1}\sum_{j_1=1}^{p_2} \left|E\left(\left(u_{ij,s}u_{ij_1,t} - E\left(u_{ij,s}u_{ij_1,t}\right)\right)\left(u_{i'j,u}u_{i'j_1',v} - E\left(u_{i'j,u}u_{i'j_1',v}\right)\right)\right)\right| \leq c_0,
$$

it should follow that

$$
\begin{aligned}
& c_0 \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} \sum_{t,s,u,v=1}^{T} \sum_{i,i'=1}^{p_1} \sum_{j_1,j_1'=1}^{p_2} \left| E\left(\left(u_{ij,s} u_{ij_1,t} - E\left(u_{ij,s} u_{ij_1,t} \right) \right) \left(u_{i'j,u} u_{i'j_1',v} - E\left(u_{i'j,u} u_{i'j_1',v} \right) \right) \right) \right| \\
\leq \quad & c_0 \frac{1}{(Tp_2)^2} p_2 T^2 p_2 p_1 \\
\leq \quad & c_0 p_1,
\end{aligned}
$$

which is the bound we want for this term, same as in you Lemma B.3.

Similarly, considering II

$$
\begin{aligned}
& \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} E\left(\sum_{t,s=1}^{T} \sum_{i=1}^{p_1} \sum_{j_1=1}^{p_2} h_t h_s E\left(u_{ij,s} u_{ij_1,t} \right) \right)^2 \\
= \quad & \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} E\left(\sum_{t,s,u,v=1}^{T} \sum_{i,i'=1}^{p_1} \sum_{j_1,j_1'=1}^{p_2} h_t h_s h_u h_v E\left(u_{ij,s} u_{ij_1,t} \right) E\left(u_{i'j,u} u_{i'j_1',v} \right) \right) \\
= \quad & \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} \sum_{t,s,u,v=1}^{T} \sum_{i,i'=1}^{p_1} \sum_{j_1,j_1'=1}^{p_2} E\left(h_t h_s h_u h_v \right) E\left(u_{ij,s} u_{ij_1,t} \right) E\left(u_{i'j,u} u_{i'j_1',v} \right) \\
\leq \quad & \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} \sum_{t,s,u,v=1}^{T} \sum_{i,i'=1}^{p_1} \sum_{j_1,j_1'=1}^{p_2} \left| E\left(h_t h_s h_u h_v \right) \right| \left| E\left(u_{ij,s} u_{ij_1,t} \right) \right| \left| E\left(u_{i'j,u} u_{i'j_1',v} \right) \right| \\
\leq \quad & c_0 \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} \sum_{t,s,u,v=1}^{T} \sum_{i,i'=1}^{p_1} \sum_{j_1,j_1'=1}^{p_2} \left| E\left(u_{ij,s} u_{ij_1,t} \right) \right| \left| E\left(u_{i'j,u} u_{i'j_1',v} \right) \right| \\
= \quad & c_0 \frac{1}{(Tp_2)^2} \sum_{j=1}^{p_2} \left(\sum_{t,s=1}^{T} \sum_{i=1}^{p_1} \sum_{j_1=1}^{p_2} \left| E\left(u_{ij,s} u_{ij_1,t} \right) \right| \right)^2 \\
\leq \quad & c_0 \frac{1}{(Tp_2)^2} p_2 T^2 p_1^2 \\
\leq \quad & c_0 \frac{p_1^2}{p_2},
\end{aligned}
$$

having used equation (A.14) in the third line from the bottom. The desired result now follows. □

Proof of Lemma B.1. We consider the case of $\widehat{k}_2 = 1$ for simplicity, and we begin by noting that, by construction, $\|\widehat{c}\|_F = p_2^{1/2}$. Also recall that

$$
X_t = RF_t + E_t.
$$

Consider now

$$\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}X_t\widehat{c}\widehat{c}'X_t' = \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_t\widehat{c}\widehat{c}'F_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_t\widehat{c}\widehat{c}'E_t' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_t\widehat{c}\widehat{c}'F_t'R' + \frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_t\widehat{c}\widehat{c}'E_t'. \tag{D.21}$$

It holds that

$$\lambda_1\left(\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RE\left(F_t\widehat{c}\widehat{c}'F_t'\right)R'\right) \geq \lambda_1\left(\frac{R'R}{p_1}\right)\lambda_{\min}\left(\frac{1}{p_2^2T}E\left(\sum_{t=1}^{T}F_t\widehat{c}\widehat{c}'F_t'\right)\right) \geq c_0p_2^{-1}\lambda_{\min}\left(E\left(\frac{1}{T}\sum_{t=1}^{T}F_t\widehat{c}\widehat{c}'F_t'\right)\right) \geq c_1p_2^{-1},$$

by (B.5). Hence, by similar passages as in the other proofs, it follows that there exists a constant $0 < c_0 < \infty$ and a triplet of random variables $(p_{1,0}, p_{2,0}, T_0)$ such that, for all $p_1 \geq p_{1,0}$, $p_2 \geq p_{2,0}$ and $T \geq T_0$

$$\lambda_1\left(\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_t\widehat{c}\widehat{c}'F_t'R'\right) \geq c_0p_2^{-1}.$$

Also, it is immediate to see that

$$\lambda_j\left(\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_t\widehat{c}\widehat{c}'F_t'R'\right) = 0,$$

for all $j \geq k_1$. Also

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}RF_t\widehat{c}\widehat{c}'E_t'\right\|_F = \frac{1}{p_1p_2^2T}\left(\sum_{i,j=1}^{p_1}\left|\sum_{t=1}^{T}R_i\left(\sum_{h=1}^{p_2}\widehat{c}_hF_{h,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_he_{jh,t}\right)\right|^2\right)^{1/2} \leq c_0\frac{1}{p_2\sqrt{T}},$$

by (B.6); the same holds for the third term in (D.21) by symmetry. Finally

$$\left\|\frac{1}{p_1p_2^2T}\sum_{t=1}^{T}E_t\widehat{c}\widehat{c}'E_t'\right\|_F = \frac{1}{p_1p_2^2T}\left(\sum_{i,j=1}^{p_1}\left|\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_he_{ih,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_he_{jh,t}\right)\right|^2\right)^{1/2} \leq c_0\frac{1}{p_1p_2}.$$

Therefore, it holds that

$$\lambda_j\left(\frac{1}{p_2^2T}\sum_{t=1}^{T}X_t\widehat{c}\widehat{c}'X_t'\right) \geq c_0\frac{p_1}{p_2},$$

for all $j \leq k_1$ and

$$\lambda_j\left(\frac{1}{p_2^2 T}\sum_{t=1}^{T} X_t \widehat{c}\widehat{c}' X_t'\right) = o_{a.s.}\left(\left(\frac{p_1}{p_2\sqrt{T}}+\frac{1}{p_2}\right)\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),$$

for all $j > k_1$.

We now consider the trace divided by p_1, viz.

$$\begin{aligned}
& \frac{1}{p_1 p_2^2 T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_h X_{ih,t}\right)^2 \\
= & \frac{1}{p_1 p_2^2 T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2} R_i \widehat{c}_h F_{h,t}\right)^2 + \frac{1}{p_1 p_2^2 T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_h e_{ih,t}\right)^2 \\
& + \frac{2}{p_1 p_2^2 T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2} R_i \widehat{c}_h F_{h,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_h e_{ih,t}\right) \\
= & I + II + III.
\end{aligned}$$

We have

$$\begin{aligned}
I &= \left(\frac{1}{p_1}\sum_{i=1}^{p_1} R_i^2\right)\frac{1}{p_2^2 T}\sum_{t=1}^{T}\sum_{h,k=1}^{p_2}\widehat{c}_h\widehat{c}_k F_{h,t}F_{k,t} \\
&= c_0 \frac{1}{p_2^2 T}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_h F_{h,t}\right)^2.
\end{aligned}$$

Given that

$$E\left(\sum_{h=1}^{p_2}\widehat{c}_h F_{h,t}\right)^2 = c_0 p_2,$$

by (B.5), the same arguments as above yield

$$\frac{1}{p_2^2 T}\sum_{t=1}^{T}\sum_{h,k=1}^{p_2}\widehat{c}_h\widehat{c}_k F_{h,t}F_{k,t} \geq c_0 p_2^{-1},$$

almost surely, and also

$$I = O_{a.s.}\left(p_2^{-1}\right).$$

Similarly, (B.7) readily entails that the same rates hold for II. Finally, consider

$$
\begin{aligned}
& E\left(\frac{1}{p_1p_2^2T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}R_i\widehat{c}_hF_{h,t}\right)\left(\sum_{h=1}^{p_2}\widehat{c}_he_{ih,t}\right)\right)^2 \\
= \; & E\frac{1}{p_1^2p_2^4T^2}\sum_{i,j=1}^{p_1}R_iR_j\sum_{t,s=1}^{T}\sum_{h_1,h_2,h_3,h_4=1}^{p_2}\widehat{c}_{h_1}\widehat{c}_{h_2}\widehat{c}_{h_3}\widehat{c}_{h_4}F_{h_1,t}F_{h_2,s}e_{ih_3,t}e_{jh_4,s} \\
\leq \; & c_0E\frac{1}{p_1^2p_2^4T^2}\sum_{i,j=1}^{p_1}\sum_{t,s=1}^{T}\sum_{h_1,h_2,h_3,h_4=1}^{p_2}\left|E\left(F_{h_1,t}F_{h_2,s}e_{ih_3,t}e_{jh_4,s}\right)\right| \\
\leq \; & c_0\frac{1}{p_1p_2^2T},
\end{aligned}
$$

by (B.8), which shows that III is dominated. Hence there are two nonzero, finite constants $c_0 \leq c_1$ and a triplet of random variables $(p_{1,0}, p_{2,0}, T_0)$ such that, for all $p_1 \geq p_{1,0}$, $p_2 \geq p_{2,0}$ and $T \geq T_0$

$$
c_0p_2^{-1} \leq \frac{1}{p_1p_2^2T}\sum_{i=1}^{p_1}\sum_{t=1}^{T}\left(\sum_{h=1}^{p_2}\widehat{c}_hX_{ih,t}\right)^2 \leq c_1p_2^{-1}.
$$

This entails that

$$
\frac{\widetilde{\lambda}_j}{p_1^{-1}\sum_{h=1}^{p_1}\widetilde{\lambda}_h} \geq c_0p_1,
$$

for all $j \leq k_1$ and

$$
\frac{\widetilde{\lambda}_j}{p_1^{-1}\sum_{h=1}^{p_1}\widetilde{\lambda}_h} \leq c_0 + o_{a.s.}\left(\frac{p_1}{\sqrt{T}}\left(\ln^2 p_1 \ln p_2 \ln T\right)^{1/2+\epsilon}\right),
$$

for all $j > k_1$. □

References

Aue, A., Hörmann, S., Horváth, L., Reimherr, M., 2009. Break detection in the covariance structure of multivariate time series models. Annals of Statistics 37, 4046–4087.

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

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

Barigozzi, M., Trapani, L., 2022. Testing for common trends in nonstationary large datasets. Journal of Business & Economic Statistics 40, 1107–1122.

Berkes, I., Hörmann, S., Schauer, J., 2011. Split invariance principles for stationary processes. Annals of Probability 39, 2441–2473.

Bollerslev, T., 1990. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. The Review of Economics and Statistics , 498–505.

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

Chen, R., Han, Y., Li, Z., Xiao, H., Yang, D., Yu, R., 2022. Analysis of tensor time series: tensorts. Journal of Statistical Software .

Davis, R.A., Pfaffel, O., Stelzer, R., 2014. Limit theory for the largest eigenvalues of sample covariance matrices with heavy-tails. Stochastic Processes and their Applications 124, 18–50.

El Karoui, N., 2005. Recent results about the largest eigenvalue of random covariance matrices and statistical application. Acta Physica Polonica Series B 36, 2681.

Han, Y., Chen, R., Zhang, C.H., 2022. Rank determination in tensor factor model. Electronic Journal of Statistics 16, 1726–1803.

Jeantheau, T., 1998. Strong consistency of estimators for multivariate ARCH models. Econometric theory 14, 70–86.

Kawakatsu, H., 2006. Matrix exponential GARCH. Journal of Econometrics 134, 95–128.

Lam, C., 2021. Rank determination for time series tensor factor model using correlation thresholding. Technical Report. Working paper LSE.

Liu, W., Lin, Z., 2009. Strong approximation for a class of stationary processes. Stochastic Processes and their Applications 119, 249–280.

Merikoski, J.K., Kumar, R., 2004. Inequalities for spreads of matrix sums and products. Applied Mathematics E-Notes 4, 150–159.

Michel, R., 1976. Nonuniform central limit bounds with applications to probabilities of deviations. The Annals of Probability , 102–106.

Móricz, F., 1976. Moment inequalities and the strong laws of large numbers. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 35, 299–314.

Moricz, F., 1983. A general moment inequality for the maximum of the rectangular partial sums of multiple series. Acta Mathematica Hungarica 41, 337–346.

Olver, F.W., Lozier, D.W., Boisvert, R.F., Clark, C.W., 2010. NIST handbook of mathematical functions hardback and CD-ROM. Cambridge university press.

Potra, F.A., 1985. On superadditive rates of convergence. ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 19, 671–685.

Salzer, H.E., Zucker, R., Capuano, R., 1952. Table of the zeros and weight factors of the first twenty Hermite polynomials. Journal of Research of the National Bureau of Standards 48, 111.

Trapani, L., 2018. A randomized sequential procedure to determine the number of factors. Journal of the American Statistical Association 113, 1341–1349.

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

Wu, W.B., 2005. Nonlinear system theory: Another look at dependence. Proceedings of the National Academy of Sciences of the United States of America 102, 14150–14154.

Yu, L., He, Y., Kong, X., Zhang, X., 2022. Projected estimation for large-dimensional matrix factor models. Journal of Econometrics 229, 201–217.